Universität Trier

Efficient PDE Constrained Shape Optimization in Shape Spaces

DISSERTATION

zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.)

vorgelegt am Fachbereich IV – Mathematik der Universität Trier von

Dipl.-Math. Kathrin Welker

im Juli 2016

Gutachter: **Prof. Dr. Volker H. Schulz**
Prof. Dr. Roland Herzog

Abstract

Shape optimization is of interest in many fields of application. In particular, shape optimization problems arise frequently in technological processes which are modelled by partial differential equations (PDEs). In a lot of practical circumstances, the shape under investigation is parametrized by a finite number of parameters, which, on the one hand, allows the application of standard optimization approaches, but, on the other hand, unnecessarily limits the space of reachable shapes. Shape calculus presents a way to circumvent this dilemma. However, so far shape optimization based on shape calculus is mainly performed using gradient descent methods. One reason for this is the lack of symmetry of second order shape derivatives or shape Hessians.

A major difference between shape optimization and the standard PDE constrained optimization framework is the lack of a linear space structure on shape spaces. If one cannot use a linear space structure, then the next best structure is a Riemannian manifold structure, in which one works with Riemannian shape Hessians. They possess the often sought property of symmetry, characterize well-posedness of optimization problems and define sufficient optimality conditions. In general, shape Hessians are used to accelerate gradient-based shape optimization methods.

This thesis deals with shape optimization problems constrained by PDEs and embeds these problems in the framework of optimization on Riemannian manifolds to provide efficient techniques for PDE constrained shape optimization problems on shape spaces. The Riemannian geometrical point of view on unconstrained shape optimization established in [84] is extended to a Lagrange-Newton and a quasi-Newton technique in shape spaces for PDE constrained shape optimization problems. These techniques are based on the Hadamard-form of shape derivatives, i.e., on the form of integrals over the surface of the shape under investigation. It is often a very tedious, not to say painful, process to derive such surface expressions. Along the way, volume formulations in the form of integrals over the entire domain appear as an intermediate step. This thesis couples volume integral formulations of shape derivatives with optimization strategies on shape spaces in order to establish efficient shape algorithms reducing analytical effort and programming work. In this context, a novel shape space is proposed.

Zusammenfassung

Formoptimierung ist in vielen Anwendungsbereichen von großem Interesse. Insbesondere entstehen Formoptimierungsprobleme in technologischen Prozessen, die mit Hilfe von partiellen Differentialgleichungen (PDEs) modelliert werden. In vielen praktischen Anwendungen ist die zu untersuchende Form durch endlich viele Parameter charakterisiert. Einerseits ermöglicht dies die Anwendung von Standardoptimierungsansätzen, andererseits wird der Raum der erreichbaren Formen unnötig begrenzt. Einen Weg dieses Dilemma zu umgehen liefert das Formenkalkül. Bisher wurden die auf dem Formenkalkül basierenden Formoptimierungprobleme jedoch hauptsächlich mit Gradientenabstiegsverfahren gelöst. Die im Allgemeinen vorhandene Unsymmetrie der zweiten Formableitung bzw. der Form-Hesse-Matrix ist ein Grund hierfür.

Generell sind Formenräume keine linearen Räume, wodurch eine große Lücke zwischen Formoptimierung und der herkömmlichen PDE-beschränkten Optimierung entsteht. Diese Lücke gilt es zu schließen. Falls kein linearer Raum in Frage kommt bzw. vorhanden ist, dann fällt die nächst beste Wahl auf eine Riemannsche Mannigfaltigkeit. Möchte man auf Riemannschen Mannigfaltigkeiten optimieren, so muss man mit dem Riemannschen Form-Gradienten bzw. der Form-Hesse-Matrix arbeiten. Die Riemannsche Form-Hesse-Matrix besitzt im Gegensatz zu der herkömmlichen Form-Hesse-Matrix die häufig gewünschte Eigenschaft der Symmetrie. Solch eine Hesse-Matrix gibt Auskunft über die Wohlgestelltheit eines Optimierungsproblems und definiert die hinreichenden Optimalitätsbedingungen. Im Allgemeinen wird sie verwendet, um gradientenbasierte Formoptimierungsverfahren zu beschleunigen.

Diese Arbeit befasst sich mit PDE-beschränkten Formoptimierungsproblemen, welche als Optimierungsprobleme auf Riemannschen Mannigfaltigkeiten betrachtet werden, um effiziente Methoden auf Formenräumen für diese bereitzustellen. Durch Anknüpfen an die Riemannschen Formoptimierungsansätze, welche in [84] für unbeschränkte Formoptimierungsprobleme entwickelt wurden, erweitert diese Arbeit sie um Lagrange-Newton, sowie quasi-Newton Ansätze für PDE-beschränkte Probleme. In diesen Methoden wird die Hadamard-Form der Formableitung verwendet. Diese Hadamard-Form ist ein Oberflächenintegral und oftmals sehr aufwendig und zeitintensiv herzuleiten. Als Zwischenergebnis dieser Herleitung erhält man ein Volumenintegral, welches aus mehreren Gesichtspunkten viel attraktiver ist als ein Oberflächenintegral. Daher beschäftigt sich diese Arbeit zusätzlich mit der Frage, wie man dieses Volumenintegral in den in dieser Arbeit entwickelten Formoptimierungsmethoden verwenden kann, um Effizienz durch Reduzierung analytischer Arbeit und des Programmieraufwandes zu erreichen. In diesem Zusammenhang wird ein neuer Formenraum definiert.

Acknowledgments

First of all, I would like to express the deepest appreciation to my supervisor Prof. Dr. Volker H. Schulz for many helpful discussions and the ongoing support. Not only the pleasant environment of his group at the Trier University, but also the opportunities to present my research on several national and international conferences and to take part on various seminars and workshops with inspiring discussions, including two seminars at the Oberwolfach Research Institute for Mathematics (MFO) and a three weeks workshop at the Erwin Schrödinger International Institute for Mathematics and Physics in Vienna, were very helpful to finish this thesis. Moreover, he introduced me to many interesting people. In this regard, he introduced me to Prof. Dr. Roland Herzog whom I am very grateful for accepting the position as second referee.

I would also like to thank Dr. Martin Siebenborn with whom I have had the pleasure to work closely over the last years. His detailed knowledge about the implementation of optimization algorithms in C++ has been as valuable to me as his moral support. The three of us, that is, Prof. Dr. Volker H. Schulz, Dr. Martin Siebenborn, and myself, have written three refereed publications [87, 88, 89] and one paper [86] which is accepted for publication. We have been a perfectly complementing team.

Moreover, I would like to thank Ben Anthes whom I have met during an MFO seminar. We have had many discussions about differential geometry until late at night during our stay at the MFO. These were extremely helpful with regard to develop the shape space definition in Chapter 7. Moreover, I am very grateful for proofreading this thesis.

Furthermore, I would like to thank all my colleagues from the mathematics department of the Trier University for creating such a nice working atmosphere. Special thanks go to Laura Somorowski, Christina Müller, Dr. Ulf Friedrich, Bernd Perscheid, Gennadij Heidel, Dennis Kreber and Dr. Matthias Wagner for proofreading this thesis, many discussions and nice coffee breaks. Moreover, I thank Prof. Dr. Leonhard Frerick for helpful discussions.

Finally, I would like to express a deepest thank to my parents Elfriede and Herbert Welker for supporting me in so many ways and encouraging me again and again in all situations of life. Moreover, I thank my closest friends Julia Reinert, Carina Hartmann, née Welsch, and Jessica Eckert for their moral support.

This work has been partly supported by the Deutsche Forschungsgemeinschaft (DFG) within the priority program SPP 1648 “Software for Exascale Computing” under contract number Schu804/12-1, by BMBF (German Federal Ministry of Education and Research) within the collaborative project RŒNOBIO under contract number 05M13UTA and by the DFG research training group 2126 Algorithmic Optimization.

Preface

As part of this thesis, three refereed publications were written, which appear as [87, 88, 89] in the bibliography. Furthermore, one paper [86] is accepted for publication. These works are fundamental for this thesis and therefore integrated. Moreover, therein, the theoretical methods were mainly implemented in C^{++} for test problems by the co-author Dr. Martin Siebenborn. Since these implementations endorse and illustrate the results of this thesis, some figures are taken from [86, 88, 89].

Contents

Chapter 1

Introduction

1.1 Motivation, aim and scope of the thesis

Shape optimization is of high importance in a wide range of applications, in particular in the context of partial differential equations (PDEs). A lot of real world problems can be reformulated as PDE constrained shape optimization problems. Aerodynamic shape optimization [81], acoustic shape optimization [102], optimization of interfaces in transmission problems [34, 68, 73], shape optimization in thermo-elastic processes [95], image restoration and segmentation [43], electrochemical machining [42] and inverse modelling of skin structures [68] can be mentioned as examples. The subject of shape optimization is covered by several fundamental monographs, see for instance [23, 41, 67, 93]. So far, the research on shape optimization was focussed on the theoretical framework, the study of existence of solutions and on the determination of shape derivatives. The optimization methodology is mostly limited to steepest descent methods based on the shape derivative with only very few exceptions (cf. [43, 77, 84]). One reason for this is the lack of symmetry of second order shape derivatives or shape Hessians.

Questions like *"How can shapes be defined?"* or *"How does the set of all shapes look like?"* have been extensively studied in recent decades. Already in 1984, David G. Kendall has introduced the notion of a shape space, which is published in [48]. Often, a shape space is just modelled as a linear (vector) space, which in the simplest case is made up of vectors of landmark positions (cf. [20, 40, 48, 75, 92]). However, there is a large number of different shape concepts, e.g., plane curves [63, 64, 65, 66], surfaces in higher dimensions [11, 12, 50, 54, 62], boundary contours of objects [32, 60, 107], multiphase objects [106], characteristic functions of measurable sets [111] and morphologies of images [25]. In a lot of processes in engineering, medical imaging and science, there is a great interest to equip the space of all shapes with a significant metric to distinguish between different shape geometries. For example, it is very important to compute distances between shapes in computational anatomy. If one knows that a patient's tissue is close to a failing one, e.g., a failing heart or a human cerebral tissue affected by Alzheimer's disease, an early treatment can be

started. In the simplest shape space case (landmark vectors), the distances between shapes can be measured by the Euclidean distance, but in general, the study of shapes and their similarities is a central problem. Now, we can ask how we can tackle natural questions like *"How different are shapes?"*, *"Can we determine the measure of their difference?"* or *"Can we infer some information?"*. In order to answer these questions mathematically, we have to put a metric on the shape space. There are various types of metrics on shape spaces, e.g., inner metrics [11, 12, 64], which can be seen as describing a deformable material that the shape itself is made of. In contrast to these inner metrics, there are also outer metrics [13, 17, 36, 48, 64]. Since the differential operator governing these metrics is defined even outside of the shape, they can be seen as describing some deformable material that the ambient space of the shape is made of. Furthermore, metamorphosis metrics [45, 101], the Wasserstein or Monge-Kantorovic metric on the shape space of probability measures [5, 14, 15], the Weil-Petersson metric [55, 91], current metrics [26, 27, 103] and metrics based on elastic deformations [32, 108] should be mentioned. However, it is a challenging task to model both, the shape space and the associated metric. There does not exist a common shape space or shape metric suitable for all applications. Different approaches lead to diverse models. The suitability of an approach depends on the requirements in a given situation. In this thesis, among all these shape space concepts, we pick the Riemannian shape manifold introduced by Peter W. Michor and David Mumford in which a two-dimensional shape is defined as a smooth embedding from the unit circle into the plane (cf. [63]). Moreover, we consider inner metrics. More precisely, we first work with so-called Sobolev metrics on this shape space. Of course, there are a lot of further metrics on it (cf. [64]), but the Sobolev metric is the most suitable choice for our applications. However, we consider also so-called Steklov-Poincaré metrics in order to enable the usage of volume integral formulations of shape derivatives in optimization strategies. Note that it is not possible to use these formulations by considering Sobolev metrics.

A priori parametrizations of shapes of interest are often used in an industrial context because of the resulting vector space framework. On the one hand, it allows the application of standard optimization software, on the other hand, it severely limits the insight into optimal shapes because only shapes corresponding to the a priori parametrization can be reached. A way to circumvent this dilemma is presented by shape calculus. Therefore, this thesis focuses on shape optimization in the context of shape calculus, which does not suffer from this limitation. *Interpreting PDE constrained shape optimization as optimization on shape manifolds* is one of the main aims of this thesis. Riemannian manifolds are the next best option if one cannot work with vector spaces. Optimization on finite dimensional manifolds is discussed in-depth in [1]. However, in this thesis, we consider infinite dimensional manifolds. As shown by Peter W. Michor and David Mumford in [64] and related papers, the set of shapes can be understood as an infinite dimensional Riemannian manifold with a tangent space defined by normal vector fields. More precisely, a two-dimensional shape is represented by a smooth embedding from the unit circle into the plane

modulo diffeomorphisms from the unit circle into itself. Loosely speaking, we can think of shapes as the images of simple closed smooth curves in the plane. Since a re-parametrization does not affect the image of a shape, one is led to consider equivalence classes of curves by factoring out diffeomorphisms from the unit circle into itself. This moduli space is a manifold and gets its Riemannian structure by introducing appropriate metrics (cf. [10, 11, 12, 62, 63, 64]). However, it consists only of $\mathcal{C}^\infty$-shapes. By a $\mathcal{C}^\infty$-shape we mean a simply connected, non-empty and compact subset of the Euclidean space with infinitely often differentiable boundary.

As discussed in [84], the interpretation of tangent vectors as directional derivatives builds the bridge between differential geometric concepts of Riemannian shape manifolds and shape optimization. In [84], also a Riemannian shape Hessian is introduced, which is symmetric and provides a Taylor series expansion in contrast to the classical notion of a shape Hessian, which is described as a second order shape derivative. Furthermore, connections between both Hessian concepts are discussed. This opens the door to optimization algorithms in the fashion of non-linear programming (NLP) approaches. However, in [84], only unconstrained shape optimization problems are considered. This thesis aims at *extending the Riemannian geometrical point of view on shape optimization established in [84] to PDE constrained shape optimization problems*. Note that the publication [84] utilizes the above-mentioned shape space introduced by Peter W. Michor and David Mumford. Thus, it is only concerned with $\mathcal{C}^\infty$-shapes, which limits the application of methods to a certain extent. Therefore, another aim of this thesis is to *propose a novel shape space*, which consists not only of $\mathcal{C}^\infty$-shapes. In other words, we want to extend the definition of $\mathcal{C}^\infty$-shapes to shapes with kinks in their boundaries.

Recent progresses in PDE constrained optimization on shape manifolds are based on the Hadamard-form of shape derivatives, i.e., on the form of integrals over the boundary of the shape under investigation, as well as intrinsic shape metrics. Major effort in shape calculus has been devoted towards such Hadamard-form expressions (cf. [23, 93]). It is often very tedious to derive boundary formulations of shape derivatives. Along the way, domain formulations in form of integrals over the entire domain appear as an intermediate step. Recently, it has been shown that this intermediate formulation has numerical advantages, see for instance [16, 34, 44, 73]. In [58], also practical advantages of domain shape formulations have been demonstrated. For example, they require less smoothness assumptions. Furthermore, the derivation as well as the implementation of domain integral formulations require less manual and programming work. Thus, there arises the natural aim to *use domain integral forms of shape derivatives in optimization strategies* on shape spaces, which seem tightly coupled with boundary integral formulations of shape derivatives. However, domain integral forms of shape derivatives require an outer metric on the domain surrounding the shape boundary. This thesis aims to *harmonize both points of view by deriving a metric from an outer metric*. Based on this metric, another main aim of this thesis is to *propose efficient shape optimization algorithms*, which also reduce the analytical effort so far involved in the derivation of shape derivatives.

In summary, this thesis deals with PDE constrained shape optimization problems and its main contributions are the following:

- Embedding PDE constrained shape optimization problems in the framework of optimization on shape spaces.
- Extending the Riemannian geometrical point of view on shape optimization established in [84] to PDE constrained shape optimization.
- Coupling domain integral formulations of shape derivatives with optimization techniques.
- Providing efficient shape optimization algorithms on shape manifolds.
- Defining a shape space which consists not only of $\mathcal{C}^\infty$-shapes.

1.2 Structure of the thesis

This thesis is structured as follows:

For the convenience of the reader **Chapter 2** introduces notations and background knowledge required in this thesis. All used symbols and function spaces are listed in Section 2.1. Second-order PDEs are the topic of Section 2.2. Subsection 2.2.1 is concerned with weak solutions of PDEs, more precisely second-order PDEs. After a brief introduction to Sobolev spaces, weak solutions of elliptic and parabolic PDEs are defined. Subsection 2.2.2 provides regularity results of second-order elliptic PDEs, which are relevant in Chapter 6 and 7. Finally, Section 2.3 provides definitions and results from differential topology and geometry. Subsection 2.3.1 presents manifold definitions and introduces related objects like tangent spaces, vector bundles and mappings between manifolds. The Riemannian geometric part of this section, Subsection 2.3.2, defines covariant derivatives, parallel transports, geodesics and the exponential map. It closes with an important remark about quotient manifolds.

Chapter 3 deals with shape spaces. Among all shape space concepts, the space B_e of two-dimensional shapes introduced by Peter W. Michor and David Mumford is essential in this thesis. Thus, Section 3.1 defines this shape space, which is even a shape manifold. Section 3.2 is devoted to Riemannian metrics on this shape space. Very important is the first Sobolev metric, which is mainly considered in Section 3.2. In particular, an expression of the covariant derivative is produced. For the sake of completeness, the short Section 3.3, which generalizes the shape space B_e and its properties to higher dimensions, closes this chapter.

Shape derivatives are the topic of **Chapter 4**. In Section 4.1, notations and definitions from shape calculus are provided. Special attention is paid to the material and shape derivative, which are needed throughout this thesis, and to the theorem of Correa and Seeger, which is applied to deduce shape derivatives. In Section 4.2, we consider problems of finding interfaces between two subdomains. An elliptic

(Subsection 4.2.1) and a parabolic shape interface optimization problem (Subsection 4.2.2) are introduced. Moreover, their shape derivatives – expressed as domain and boundary integrals – are deduced. These expressions are essential in Chapter 5 to 7 because they exemplify and support the theoretical approaches established in these chapters. Section 4.3 focuses on shape optimization in the context of shape calculus. In particular, the connection of Riemannian geometry on the shape space B_e to shape optimization is analyzed by providing expressions of the Riemannian shape gradient and the Riemannian shape Hessian with respect to the Sobolev metric. Finally, Section 4.4 proposes a volume shape derivative formula for a special class PDE constrained shape interface problems.

In **Chapter 5**, the Riemannian geometrical point of view on shape optimization established in [84] is extended to PDE constrained shape optimization problems. Section 5.1 presents a sequential quadratic programming (SQP) approach. The main tool for the development of respective Lagrange-Newton methods is the concept of vector bundles associated with shape manifolds. Subsection 5.1.1 provides such a Riemannian vector bundle framework, which is applied to the elliptic shape interface optimization problem given in Subsection 4.2.1. Furthermore, shape variants of quasi-Newton methods are provided in Section 5.2 and exemplified by the parabolic shape interface optimization problem given in Subsection 4.2.2. Numerical results, which are presented in Section 5.3, endorse the results of the first two sections. In Subsection 5.3.1, quadratic convergence rates of the Lagrange-Newton approach are demonstrated. Finally, Subsection 5.3.2 shows superlinear convergence rates of the quasi-Newton methods.

Chapter 6 is intended to enable the usage of domain integral forms of shape derivatives in optimization strategies. Section 6.1 discusses generalized Steklov-Poincaré operators as basis for scalar products on shape spaces. Section 6.2 rephrases optimization algorithms on shape spaces within the framework of domain integral formulations of shape derivatives and in the context of Steklov-Poincaré metrics introduced in Section 6.1. In this manner, quasi-Newton methods based on volume expressions of shape derivatives and Steklov-Poincaré metrics are provided. Section 6.3 discusses not only algorithmic and implementation details, but also numerical results for the parabolic shape interface optimization problem given in Subsection 4.2.2. Moreover, it compares the approach established in Chapter 5 (based on surface expressions of shape derivatives) with the approach established in this chapter (based on volume formulations of shape derivatives) from a computational point of view.

All chapters mentioned so far are only concerned with $\mathcal{C}^\infty$-shapes. This limits the application of the methods proposed in the previous chapters to a certain extent. As a remedy, **Chapter 7** proposes a novel shape space in order to extend these methods to shapes of it with respect to the shape metric g^S introduced in Section 6.1. Section 7.1 is devoted to a shape space definition. Moreover, its connection to shape calculus is given in order to formulate the shape quasi-Newton methods of Chapter 6 on this shape space. Section 7.2 gives numerical results. Finally, Section 7.3 investigates the connection of geodesics discussed in [32] to geodesics in the novel shape space $\mathcal{B}^{1/2}$

with respect to the Steklov-Poincaré metric g^S in order to obtain a distance measure for shapes of $\mathcal{B}^{1/2}$ with respect to g^S.

This thesis concludes with **Chapter 8**, which gives a summary (Section 8.1) and an outlook towards future research fields (Section 8.2).

Chapter 2

Notations and background knowledge

This chapter provides notations and background knowledge and is organized as follows. Section 2.1 lists all used symbols and function spaces. Second-order partial differential equations and their regularity are the topic of Section 2.2. After a brief introduction to Sobolev spaces of integer and fractional order, special attention is paid to elliptic and parabolic PDEs. Finally, Section 2.3 provides definitions and results from differential topology and differential geometry. This section presents manifold definitions and introduces related objects like tangent spaces, vector bundles and mappings between manifolds. The Riemannian geometric part of this section defines covariant derivatives, parallel transports, geodesics and the exponential map.

2.1 A brief glossary of notations and conventions

For convenience this section gives a brief glossary of all required symbols. The following conventions and notations are needed throughout this thesis.

Sets and geometrical domains

$\mathbb{N}$	set of natural numbers including zero; $\overline{\mathbb{N}} \coloneqq \mathbb{N} \cup \{\infty\}$, $\mathbb{N}^* \coloneqq \mathbb{N} \setminus \{0\}$
$\mathbb{R}^d$	Euclidean space of dimension $d \in \mathbb{N}^*$; $\mathbb{R} = \mathbb{R}^1$
S^d	unit sphere in the Euclidean space $\mathbb{R}^{d+1}$

Norms and scalar products

Let $\mathcal{U}$ and $\mathcal{V}$ be Hilbert spaces.

$\langle \cdot, \cdot \rangle_{\mathcal{U}}$	scalar product in $\mathcal{U}$
$\| \cdot \|_{\mathcal{U}}$	norm in $\mathcal{U}$
$(\cdot, \cdot)_{\mathcal{U} \times \mathcal{V}}$	duality pairing between $\mathcal{U}$ and $\mathcal{V}$

If the space is clear from the context, we denote the respective norm or scalar product by $\|\cdot\|$ or $\langle\cdot,\cdot\rangle$. Analogously for duality pairings. Next, we list some norms and scalar products on particular spaces:

$\|\cdot\|_p$	p-norm in $\mathbb{R}^d$, i.e., $\|x\|_p \coloneqq \sqrt[p]{\sum_{i=1}^d \|x_i\|^p}$, when $p \in [1,\infty)$, and $\|x\|_\infty \coloneqq \max_{i\in\{1,\dots,n\}} \|x_i\|$
$\|\cdot\|_{g_c^n}$	norm induced by the inner product g_c^n defined by the Sobolev metric $g^n = (g_c^n)_c$, i.e., $\|\cdot\|_{g_c^n} \coloneqq \sqrt{g_c^n(\cdot,\cdot)}$

Linear algebra

Let A and B be matrices and let x and y denote two vectors.

A^T	transpose of A
A^{-1}	inverse of A
$\det(A)$	determinant of A
$\operatorname{tr}(A)$	trace of A
$A : B$	sum of component-wise products, i.e., $A : B \coloneqq \sum_{ij} A_{ij}B_{ij}$
$x^T y$	scalar product of x and y

Banach and Hilbert spaces

Unless stated otherwise, Ω denotes a bounded domain in $\mathbb{R}^d$ with Lipschitz boundary $\partial\Omega$. The closure of Ω is denoted by $\overline{\Omega}$.

$\mathcal{C}(\Omega)$	set of continuous functions in Ω
$\mathcal{C}^k(\Omega)$	set of functions with derivatives up to order $k \in \overline{\mathbb{N}}$ in $\mathcal{C}(\Omega)$; $\mathcal{C}^0(\Omega) = \mathcal{C}(\Omega)$
$\mathcal{C}_b^k(\Omega)$	set of $\mathcal{C}^k(\Omega)$-functions with bounded derivatives up to order $k \in \overline{\mathbb{N}}$
$\mathcal{C}_0^k(\Omega)$	set of $\mathcal{C}^k(\Omega)$-functions which vanish on $\partial\Omega$
$\mathcal{C}^{k,r}(\overline{\Omega})$	Hölder spaces, i.e., set of $\mathcal{C}^k(\Omega)$-functions which are Hölder-continuous with exponent $r \in [0,1]$; $\mathcal{C}^{k,0}(\overline{\Omega}) = \mathcal{C}_b^k(\Omega)$
$L^p(\Omega)$	set of p-integrable Lebesgue functions in Ω, when $p \in [0,\infty)$, and $L^\infty(\Omega)$ is the set of essentially bounded Lebesgue functions in Ω
$L^1_{\mathrm{loc}}(\Omega)$	set of all locally integrable functions in Ω
$W^{k,p}(\Omega)$	Sobolev spaces
$H^k(\Omega)$	$H^k(\Omega) \coloneqq W^{k,2}(\Omega)$; $H^0(\Omega) = L^2(\Omega)$
$H_0^k(\Omega)$	set of $H^k(\Omega)$-functions whose derivatives in the sense of traces vanish on $\partial\Omega$ up to order $k-1$

$H^k(\Omega)'$	dual space of $H^k(\Omega)$
$H^k_{\mathrm{loc}}(\Omega)$	set of all locally $H^k(\Omega)$-functions
$H^{1/2}(\partial\Omega)$	set of Dirichlet traces on $\partial\Omega$ for $H^1(\Omega)$
$H^{-1/2}(\partial\Omega)$	dual space of $H^{1/2}(\partial\Omega)$

The vector valued versions of $\mathcal{C}(\Omega)$, $\mathcal{C}^k(\Omega)$, etc. are denoted by $\mathcal{C}(\Omega, \mathbb{R}^d)$, $\mathcal{C}^k(\Omega, \mathbb{R}^d)$, etc. In the following, $\mathcal{C}^\infty$-functions are also called *smooth* functions.

Spaces involving time and related notations

Let $[0, T]$ denote a time interval with final time $T > 0$. Often, it is important to identify functions $\Omega \times [0, T] \to \mathbb{R}$ with maps from $[0, T]$ into a Banach space, e.g., $H^1(\Omega)$. In particular, the space $L^2(0, T; H^1(\Omega))$ plays an especially important role in the study of parabolic equations.

Let H and P denote abstract Banach spaces with dual spaces H' and P'. In order to handle weak solutions of parabolic differential equations, we introduce the following spaces involving time and a weak time derivative notation:

$L^2(0, T; H)$	set of L^2-integrable functions from $[0, T]$ into H
$\mathcal{C}(0, T; H)$	set of all continuous functions from $[0, T]$ into H
$W(0, T; H)$	set of $L^2(0, T; H)$-functions whose first weak time derivatives exist and are elements of $L^2(0, T; H')$
$Y(0, T; H, P)$	set of $L^2(0, T; H)$-functions whose first weak time derivatives exist and are elements of $L^2(0, T; P')$
$\mathring{y}$	weak time derivative of a function $y \in L^2(0, T; H)$

Differential geometry

Let M and N be manifolds and let $f\colon M \to N$ denote a differentiable map between these two manifolds. Moreover, let $\gamma\colon I \to M$ be a curve in M, where $I \subset \mathbb{R}$ denotes an interval. We use the following symbols from differential geometry:

T_pM	tangent space of M at $p \in M$
(E, π, M)	vector bundle of M, where M is called the base space, E is called the total space and $\pi\colon E \to M$ is called the bundle projection
(TM, π, M)	tangent bundle of M, where TM is the disjoint union of all tangent spaces of M
$\Gamma(TM)$	set of all differentiable vector fields on M
$\Gamma_\gamma(TM)$	set of all differentiable vector fields along γ
∇^{cov}	covariant derivative

exp	exponential map
$\mathcal{T}$	vector transport
$\mathcal{R}$	retraction
$\mathcal{P}$	parallel transport
$[\cdot,\cdot]$	Lie bracket
$X_{p,q}^M$	set of all differentiable paths in M connecting two points $p, q \in M$
$I_{p,v}^M$	maximal existence interval of a geodesic through $p \in M$ in direction $v \in T_pM$
grad	Riemannian shape gradient operator
Hess	Riemannian shape Hessian operator

Shape spaces, metrics and related notations

Let M be a compact smooth manifold of dimension $m > 1$ and let N be a connected Riemannian manifold of dimension $n > m$. We use the following spaces:

$\mathrm{Emb}(M,N)$	manifold of all smooth embeddings from M into N
$\mathrm{Emb}^0([0,1],\mathbb{R}^2)$	set of all $\mathcal{C}^\infty([0,1],\mathbb{R}^2)$-functions being injective immersions with fixed endpoints
$\mathrm{Imm}(M,N)$	manifold of all smooth immersions from M into N
$\mathrm{Diff}(M)$	regular Lie group of all diffeomorphisms from M into itself
$\mathrm{Diff}^0([0,1])$	set of all diffeomorphisms from $[0,1]$ into itself with fixed endpoints
$B_e(M,N)$	$B_e(M,N) \coloneqq \mathrm{Emb}(M,N)/\mathrm{Diff}(M)$; manifold of submanifolds of type M in N; $B_e \coloneqq B_e(S^1,\mathbb{R}^2)$
$B_i(M,N)$	$B_i(M,N) \coloneqq \mathrm{Imm}(M,N)/\mathrm{Diff}(M)$; infinite dimensional orbifold whose points are smooth immersed submanifolds of type M in N; $B_i \coloneqq B_i(S^1,\mathbb{R}^2)$
$B_e^0([0,1],\mathbb{R}^2)$	$B_e^0([0,1],\mathbb{R}^2) \coloneqq \mathrm{Emb}^0([0,1],\mathbb{R}^2)/\mathrm{Diff}^0([0,1])$

Let X denote a bounded domain in $\mathbb{R}$ containing Γ_0, where Γ_0 is the boundary of a Lipschitz domain $\mathcal{X}_0 \subset X$. In this case, we use the following sets:

$\mathcal{H}^{1/2}(\Gamma_0,\mathbb{R}^d)$	space of all continuous and injective deformations from Γ_0 in X, where these deformations arise from $H^1(X,X)$-functions
$\mathrm{Homeo}^{1/2}(\Gamma_0)$	set of $\mathcal{H}^{1/2}(\Gamma_0,\mathbb{R}^d)$-functions which are homeomorphisms from Γ_0 into itself
$\mathcal{B}^{1/2}(\Gamma_0,\mathbb{R}^d)$	$\mathcal{B}^{1/2}(\Gamma_0,\mathbb{R}^d) \coloneqq \mathcal{H}^{1/2}(\Gamma_0,\mathbb{R}^d)/\mathrm{Homeo}^{1/2}(\Gamma_0)$; space of all $H^{1/2}$-shapes

In this thesis, by *metrics* we mean families of positive inner products, i.e., Riemannian metrics on manifolds. We consider the following metrics:

g^n	n-th Sobolev metric, where $n \in \mathbb{N}^*$ and g^0 is equal to the standard L^2-scalar product
g^S	Steklov-Poincaré metric

Shape transformations, differential and shape calculus

Let $t \in [0, T]$ with $T > 0$. Moreover, let J be a real-valued shape differentiable functional. We suppose that this functional is well defined for any measurable set Ω in D, where D denotes a non-empty subset of $\mathbb{R}^d$. Moreover, let $p\colon \Omega \to \mathbb{R}$ be a generic function and let V be a sufficiently smooth vector field.

In this thesis, we distinguish between domain and boundary formulations of shape derivatives. Moreover, note the difference between shape derivatives of objective and generic functions. We use the following notation:

$j(\Omega)$	shape functional J without a regularization term
$j_{\mathrm{reg}}(\Omega)$	regularization term of a shape functional J
$DJ(\Omega)[V]$	shape derivative of J at Ω in direction V
$Dj_\Omega(\Omega)[V]$	domain formulation of $Dj(\Omega)[V]$
$Dj_\Gamma(\Omega)[V]$	boundary formulation of $Dj(\Omega)[V]$
$dp(x)[V]$	directional derivative of p at a point $x \in \Omega$ in direction V
$\dot{p}$ or $D_m p$	material derivative of p
p'	shape derivative of p

To handle, for example, shape derivatives, we need (shape) transformations:

F_t	mapping $\overline{D} \to \mathbb{R}^d$, where $F_0 \coloneqq id$
Ω_t	transformed geometrical domain in $\mathbb{R}^d$ of the form $\Omega_t = F_t(\Omega)$, where $\Omega_0 = \Omega$
Γ_t	boundary of Ω_t
p_t	generic real-valued function defined on Ω_t, where $p_0 = p$

Moreover, we use the following derivatives and related objects:

DF_t	Jacobian of F_t
D^α	α-th partial derivative, where α is a multi-index
D_s	arc length derivative

n	normal on $\partial\Omega$
$\frac{\partial y}{\partial x_i}$, y_{x_i}, $\partial_{x_i} y$ or $d_{x_i} y$	partial derivative of a function $y\colon \Omega \to \mathbb{R}$, i.e., for the i-th standard coordinate vector e_i and $x = (x_1, ..., x_d) \in \Omega$, it is given by $\frac{\partial y}{\partial x_i}(x) = \lim_{h\to 0} \frac{y(x+he_i)-y(x)}{h}$, provided this limit exists
$\frac{\partial}{\partial n}$	normal derivative on $\partial\Omega$
$\frac{\partial}{\partial \tau}$	tangential derivative on $\partial\Omega$
$\frac{\partial}{\partial t}$	time derivative
$\frac{d^+}{dt}\big\vert_{t=0}$	total time derivative evaluated at $t = 0^+$

Furthermore, we need the following differential elements and operators:

dx, dx_t	differential element in Ω, Ω_t
ds, ds_t	differential element on $\partial\Omega$, $\partial\Omega_t$
dt	differential element over $[0, T]$
∇	gradient operator
∇_Γ	tangential gradient operator with respect to $\Gamma = \partial\Omega$
div	divergence operator
div_Γ	tangential divergence operator with respect to $\Gamma = \partial\Omega$
κ	mean curvature of $\Gamma = \partial\Omega$; $\kappa := \frac{1}{d-1}\mathrm{div}_\Gamma(n)$
$\triangle$	Laplace operator
E_D	Dirichlet solution operator
E_N	Neumann solution operator
T	Dirichlet-to-Neumann map or Steklov-Poincaré operator
S	Neumann-to-Dirichlet map or Poincaré-Steklov operator
T^{pr}	projected Dirichlet-to-Neumann map
S^{pr}	projected Neumann-to-Dirichlet map

Miscellaneous I

Let the domain Ω be partitioned into two disjoint subdomains $\Omega_1, \Omega_2 \subset \Omega$ separated by an interface $\Gamma_{\text{int}} = \partial\Omega_1 \cap \partial\Omega_2$ as illustrated in Figure 4.1. Moreover, let $f\colon \Omega \to \mathbb{R}$ be a function. In this case, we use the following notation:

Γ_{out}	outer boundary of the split domain
Γ_{int}	interface or interior boundary of the split domain
$f\big\vert_{\Omega_i}$	restriction of f to Ω_i, where $i = 1, 2$

γ_0	trace on Γ_{int}
γ_1	normal derivative, where the normal vector is defined on Γ_{int}
$[\![\cdot]\!]$	jump symbol; $[\![f]\!] \coloneqq f\vert_{\Omega_1} - f\vert_{\Omega_2}$

Miscellaneous II

$\dot\cup$	disjoint union
$\mathscr{L}$	Lagrangian
I or id	identity function
L	length functional
E	energy functional
δ_{ij}	Kronecker delta

2.2 Second order PDEs and their regularity

A partial differential equation is an equation involving an unknown function of two or more variables and certain of its partial derivatives. Often, there are further conditions given, so-called *boundary* or *initial value conditions*. For example, a condition that specifies the value of the dependent variable on the boundary is called a *Dirichlet boundary condition*, a condition that specifies the normal derivative of the dependent variable on the boundary is called a *Neumann boundary condition* and a condition that specifies both, the value and the normal derivative of the dependent variable on the boundary, is called a *Cauchy boundary condition*. These different boundary conditions ensure that the corresponding class of PDEs has a unique and stable solution. More precisely, *elliptic problems* require either Dirichlet or Neumann boundary conditions, *parabolic problems* require Dirichlet or Neumann boundary conditions combined with initial conditions and *hyperbolic problems* require Cauchy boundary conditions.

In this thesis, we consider second order elliptic and parabolic PDEs combined with appropriate boundary conditions. We need a theorem of higher regularity in Chapter 7 and a theorem of infinite differentiability for weak solutions of elliptic PDEs in Chapter 6. Only these theorems are provided in this section. For more details about the regularity theory we refer to the literature, e.g., [31].

This section is organized as follows. Before we address the question whether *a weak solution of an elliptic PDE is in fact smooth*, we introduce second-order elliptic and parabolic PDEs as well as their weak solutions in Subsection 2.2.1. Afterwards, Subsection 2.2.2 presents a *theorem of higher regularity* and a *theorem of infinite differentiability* for weak solutions of elliptic PDEs.

2.2.1 Solutions of PDEs

Let Ω be an open and bounded subset of $\mathbb{R}^d$ with boundary $\Gamma = \partial\Omega$. In this thesis, we mainly study the following elliptic boundary value problem:

$$L_e y = f \qquad \text{in } \Omega \tag{2.1}$$

$$y = 0 \qquad \text{on } \Gamma \tag{2.2}$$

The function $y\colon \overline{\Omega} \to \mathbb{R}$ is the unknown function, $f \in L^2(\Omega)$ is given and L_e denotes an elliptic second-order partial differential operator. Condition (2.2) is called a *Dirichlet boundary condition*. In general, a *second-order partial differential operator* L has one of the following forms:

$$Ly = -\sum_{i,j=1}^{d} \left(a^{ij} y_{x_i}\right)_{x_j} + \sum_{i=1}^{d} b^i y_{x_i} + cy \qquad \text{(divergence form)} \tag{2.3}$$

$$Ly = -\sum_{i,j=1}^{d} d^{ij} y_{x_i x_j} + \sum_{i=1}^{d} e^i y_{x_i} + cy \qquad \text{(non-divergence form)} \tag{2.4}$$

Here the coefficients a^{ij}, d^{ij}, b^i, e^i and c are given, where $i, j \in \{1, ..., d\}$. Note that these two forms are equivalent if the highest order coefficients a^{ij} are $\mathcal{C}^1$-functions. Indeed the divergence form (2.3) becomes

$$Ly = -\sum_{i,j=1}^{d} d^{ij} y_{x_i x_j} + \sum_{i=1}^{d} \left(b^i - \sum_{j=1}^{d} a^{ij}_{x_j} \right) y_{x_i} + cy.$$

We call the partial differential operator L *elliptic* if there exists a constant $\theta > 0$ such that

$$\sum_{i,j=1}^{d} a^{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2$$

for almost every $x \in \Omega$ and all $\xi \in \mathbb{R}^d$. An example for an elliptic equation is the *Poisson equation*, which we achieve by choosing $a^{ij} = \delta_{ij}$, $b^i = 0$, $c = 0$. In this case, we get $L_e = -\triangle$, where $\triangle$ denotes the *Laplace operator*. For $y\colon \Omega \to \mathbb{R}$, $x \in \Omega \subset \mathbb{R}^d$, it is given by

$$\triangle y(x) \coloneqq \sum_{i=1}^{d} y_{x_i x_i}(x).$$

Besides the elliptic boundary value problem (2.1)-(2.2), we also study PDEs involving time. Let $[0, T]$ denote a time interval with final time $T > 0$ and let y be time-dependent. We mainly consider the following parabolic boundary value problem:

$$\frac{\partial y}{\partial t} + L_p y = f \qquad \text{in } \Omega \times (0, T] \tag{2.5}$$

$$y = 0 \qquad \text{on } \Gamma \times (0, T] \tag{2.6}$$

$$y = g \qquad \text{in } \Omega \times \{0\} \tag{2.7}$$

Here the function $y\colon \overline{\Omega} \times [0,T] \to \mathbb{R}$ is the unknown function, $g \in L^2(\Omega)$ and $f \in L^2(0,T;L^2(\Omega))$ are given and $\frac{\partial}{\partial t} + L_p$ denotes a parabolic second-order partial differential operator. The last condition (2.7) is called an *initial condition*. We say $\frac{\partial}{\partial t} + L_p$ is a *parabolic second-order partial differential operator* if L_p is a second-order partial differential operator for each time t and if, furthermore, there exists a constant $\theta > 0$ such that

$$\sum_{i,j=1}^{d} a^{ij}(x,t)\xi_i\xi_j \geq \theta|\xi|^2$$

for all $(x,t) \in \Omega \times (0,T]$ and all $\xi \in \mathbb{R}^d$. An example of a parabolic equation is the *heat equation*, which we get by choosing $a^{ij} = \delta_{ij}$, $b^i = 0$, $c = 0$, $f = 0$.

Now, we address questions like *"How do solutions of PDEs look like?"* and *"What are the properties of these solutions?"*. But *what exactly is a solution of a PDE*? Should we ask, for example, that a solution has to be analytic or infinitely differentiable? This is desirable, but also a too strong requirement. We have to distinguish between *classical* and *weak solutions*. Assuming that a classical solution of a k-th order PDE exists, it is unique, depends continuously on the data given in the problem and is at least k-times continuously differentiable. This means that all derivatives which appear in the PDE exist and are continuous, although higher derivatives maybe do not exist. However, in general, such classical solutions do not exist or it is very hard to find them. As a remedy, we define the notion of weak solutions for a given PDE. For these, not all derivatives may exist, but nonetheless the weak solutions satisfy the equation in some precisely defined sense. In other words, we separate the existence and regularity problems by defining a weak solution with the expectation that if we do not require too much smoothness, we can easier establish its existence, uniqueness and continuous dependence on the given data. In the following, we illustrate the concept of weak solutions for elliptic and parabolic second-order PDEs.

Weak solutions of second-order PDEs

Roughly speaking, the general idea to define weak solutions of a PDE $Ly = f$ is to rewrite it by using a so-called *test function* p such that half of the derivatives of y are shifted to p by applying the integration by parts formula given in the sequel. This new form is called the *weak* or *variational formulation* of the PDE. In our setting of second-order PDEs, this means that we deal only with weak first-order derivatives. Before we focus on these solutions, we give a very brief introduction to Sobolev spaces. In particular, we define weak derivatives and formulate the above-mentioned integration by parts formula.

A brief introduction to Sobolev spaces. We introduce Sobolev spaces very briefly without defining any norms or scalar products which are not required in this thesis. For a detailed introduction to Sobolev spaces and their properties we refer to the literature, e.g., [99].

Weak derivatives are motivated by repeated use of the integration by parts formula which is achieved by applying the Gauß-Green theorem. Before we can formulate this theorem and, thus, the integration by parts formula, we have to introduce domains of class $\mathcal{C}^{k,r}$, where $k \in \overline{\mathbb{N}}$ and $r \in [0,1]$, and in particular Lipschitz domains on which many PDEs are defined. In the following, we denote by $B_d(x,R)$ the ball in $\mathbb{R}^d$ centred at $x \in \mathbb{R}^d$ with radius $R > 0$.

Definition 2.1 (Hölder spaces)**.** *Let $\Omega \subset \mathbb{R}^d$ be open, let $k \in \overline{\mathbb{N}}$ and let $r \in (0,1]$. A function $v \in \mathcal{C}(\Omega)$ is called Hölder-continuous with exponent r if the Hölder coefficient*

$$\|v\|_{\mathcal{C}^{0,r}} \coloneqq \sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|v(x)-v(y)|}{|x-y|^r}$$

is finite. The Banach spaces

$$\mathcal{C}^{k,r}(\overline{\Omega}) \coloneqq \{v \in \mathcal{C}^k(\Omega) \colon \|D^\alpha v\|_{\mathcal{C}^{0,r}} < \infty \ \forall \alpha \in \mathbb{N}^d, |\alpha| \leq k\}$$

are called the Hölder spaces. In particular, the space $\mathcal{C}^{0,1}(\overline{\Omega})$ consists of all Lipschitz-continuous functions. If $r = 0$, the function and its derivatives are bounded. In this case, $\mathcal{C}_b^k(\Omega)$ stands for $\mathcal{C}^{k,0}(\overline{\Omega})$.

Definition 2.2 ($\mathcal{C}^{k,r}$-boundary, Lipschitz domain)**.** *Let $\Omega \subset \mathbb{R}^d$ be open with boundary $\Gamma = \partial\Omega$. Moreover, let $k \in \overline{\mathbb{N}}$ and let $r \in [0,1]$. We say Ω has a $\mathcal{C}^{k,r}$-boundary or Ω is $\mathcal{C}^{k,r}$, where $\mathcal{C}^k$ stands for $\mathcal{C}^{k,0}$, if for any $x \in \Gamma$ there exist local coordinates $y_1, ..., y_d$ centred at x, i.e., such that x is the unique solution of $y_1 = \cdots = y_d = 0$, and constants $a, b > 0$ as well as a mapping $\psi \in \mathcal{C}^{k,r}(B_{d-1}(x,a))$, where $B_{d-1}(x,a)$ is considered in the linear subspace defined by $(y_1, ..., y_{d-1})$, subject to the following conditions:*

(i) $y_d = \psi(\widetilde{y}) \Rightarrow (\widetilde{y}, y_d) \in \Gamma$,

(ii) $\psi(\widetilde{y}) < y_d < \psi(\widetilde{y}) + b \Rightarrow (\widetilde{y}, y_d) \in \Omega$,

(iii) $\psi(\widetilde{y}) - b < y_d < \psi(\widetilde{y}) \Rightarrow (\widetilde{y}, y_d) \notin \overline{\Omega}$.

A domain Ω is called Lipschitz if it is $\mathcal{C}^{0,1}$.

Now, we can formulate the theorem of Gauß-Green.

Theorem 2.3 (Gauß-Green)**.** *Let $i \in \{1, ..., d\}$, let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain and let $y \in \mathcal{C}^1(\overline{\Omega})$. Then the identity*

$$\int_\Omega y_{x_i} dx = \int_{\partial\Omega} y n^i ds$$

holds, where n^i denotes the i-th component of the outward pointing unit normal vector field $n = (n^1, ...n^d)$ to Ω.

Applying this theorem yields the integration by parts formula:

Theorem 2.4 (Integration by parts formula)**.** *Let Ω, i and n be as in Theorem 2.3. Moreover, let $y, p \in \mathcal{C}^1(\overline{\Omega})$. Then the following identity holds:*

$$\int_\Omega y_{x_i} p \, dx = -\int_\Omega y p_{x_i} \, dx + \int_{\partial\Omega} y p n^i ds \tag{2.8}$$

As just seen, in every bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, we can apply the Gauß-Green Theorem 2.3. In particular, if $y, \phi \in \mathcal{C}^k(\overline{\Omega})$ with $\phi = 0$ on $\Gamma = \partial\Omega$, we get the following equation by repeated application of the integration by parts formula (2.8):

$$\int_\Omega y \, D^\alpha \phi \, dx = (-1)^{|\alpha|} \int_\Omega \phi \, D^\alpha y \, dx, \tag{2.9}$$

where $\alpha = (\alpha_1, ..., \alpha_d)$ is a multi-index of order $|\alpha| = k$ with $|\alpha| \coloneqq \sum_{i=1}^d \alpha_i$ and

$$D^\alpha \coloneqq \frac{\partial^{|\alpha|}}{\partial^{\alpha_1} x_1 \cdots \partial^{\alpha_d} x_d}$$

denotes the α-th partial derivative. Note that the right-hand side of (2.9) is well defined whenever y is *locally integrable* in Ω, i.e., Lebesgue integrable over every compact subset of Ω. We denote by $L^1_{\text{loc}}(\Omega)$ the space of all locally integrable functions in Ω. In this setting, functions are not assumed to be differentiable, but Lebesgue integrable only. This motivates a generalization of the standard concepts of derivatives.

Definition 2.5 (Weak derivative)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, let $y \in L^1_{\text{loc}}(\Omega)$ and let α denote a multi-index. A function $w \in L^1_{\text{loc}}(\Omega)$ is called the α-th weak derivative of y if*

$$\int_\Omega y \, D^\alpha \phi \, dx = (-1)^{|\alpha|} \int_\Omega w \, \phi \, dx \quad \forall \phi \in \mathcal{C}_0^\infty(\Omega). \tag{2.10}$$

We denote the α-th weak derivative of y by $w = D^\alpha y$.

Now, we can introduce Sobolev spaces. Note that we have to distinguish between integer and fractional order Sobolev spaces. After introducing them, we recall a special Sobolev embedding theorem.

Definition 2.6 (Integer order Sobolev spaces)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain, let $k \in \mathbb{N}$ and let $1 \leq p \leq \infty$. The Banach spaces*

$$W^{k,p}(\Omega) \coloneqq \{y \in L^1_{\text{loc}}(\Omega) \colon \forall \alpha \in \mathbb{N}^d, |\alpha| \leq k \ \exists D^\alpha y \in L^p(\Omega)\}$$

are called the integer order Sobolev spaces. If $p = 2$, we denote

$$H^k(\Omega) \coloneqq W^{k,2}(\Omega),$$

which are Hilbert spaces with respect to the inner products

$$\langle u, v\rangle_{H^k(\Omega)} = \sum_{|\alpha|=k} \langle D^\alpha u, D^\alpha v\rangle_{L^2(\Omega)}\,.$$

Moreover, we define

$$H_0^1(\Omega) \coloneqq \{y \in H^1(\Omega)\colon y|_\Gamma = 0\}.$$

Definition 2.7 (Fractional order Sobolev spaces)**.** *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain, let $k \in \mathbb{N}$, let $\sigma \in (0,1)$ and let $1 \le p < \infty$. The Banach spaces*

$$W^{\sigma,p}(\Omega) \coloneqq \left\{ v \in L^2(\Omega)\colon \frac{|v(x)-v(y)|}{|x-y|^{\frac{d}{p}+\sigma}} \in L^p(\Omega\times\Omega) \right\},$$
$$W^{\sigma,\infty}(\Omega) \coloneqq \mathcal{C}^{0,\sigma}(\overline{\Omega})$$

are called the fractional order Sobolev spaces or the Sobolev-Slobodeckij spaces. If $s > 1$ is not an integer and $s = k + \sigma$, the Sobolev spaces $W^{s,p}(\Omega)$ consist of those equivalence classes of functions $y \in W^{k,p}(\Omega)$ whose weak derivatives $D^\alpha y$ belong to $W^{\sigma,p}(\Omega)$, where $|\alpha| = k$, i.e.,

$$W^{s,p}(\Omega) \coloneqq \{y \in W^{k,p}(\Omega)\colon \forall \alpha \in \mathbb{N}^d, |\alpha| = k\ \exists D^\alpha y \in W^{\sigma,p}(\Omega)\}.$$

In the case that $p = \infty$, we define

$$W^{s,\infty}(\Omega) \coloneqq \mathcal{C}^{k,\sigma}(\overline{\Omega}).$$

If $p = 2$, we denote

$$H^\sigma(\Omega) \coloneqq W^{\sigma,2}(\Omega),$$
$$H^s(\Omega) \coloneqq W^{s,2}(\Omega),$$

which are Hilbert spaces.

In Chapter 7, we need the second part of the following Sobolev embedding theorem (cf. [24, Theorem 2.72]).

Theorem 2.8. *Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain, let $1 \le p \le \infty$ and let $k \in \mathbb{N}$. If $kp > d$, the following statements hold:*

(i) If $\frac{d}{p} \notin \mathbb{N}$ and j satisfies $(j-1)p < d < jp$, then $W^{k,p}(\Omega)$ is embedded in $\mathcal{C}^{k-j,r}(\overline{\Omega})$ for all $r \le j - \frac{d}{p}$.

(ii) If $\frac{d}{p} \in \mathbb{N}$ and $j = \frac{d}{p} + 1$ satisfies $k \ge j$, then $W^{k,p}(\Omega)$ is embedded in $\mathcal{C}^{k-j,r}(\overline{\Omega})$ for all $r < 1$.

Elliptic PDEs and their weak solutions. In order to define weak solutions of second-order elliptic PDEs (2.1)-(2.2), we consider the operator L_e in divergence form (2.3), assume $a^{ij}, b^i, c \in L^\infty(\Omega)$ and suppose $a^{ij} = a^{ji}$, where $i, j \in \{1, ..., d\}$. Moreover, assuming that $\partial\Omega$ is $\mathcal{C}^1$, an outward pointing unit normal vector field exists. Now, we rewrite equation (2.1) in such a way that only first-order derivatives of the solution y show up. For this purpose, we multiply the PDE by a smooth function p of compact support and integrate over Ω by applying the integration by parts formula (2.8). In this way, we obtain the variational formulation

$$\int_\Omega \sum_{i,j=1}^{d} a^{ij} y_{x_i} p_{x_j} + \sum_{i=1}^{d} b^i y_{x_i} p + cyp \, dx = \int_\Omega fp \, dx \tag{2.11}$$

for $y, p \in H_0^1(\Omega)$. Note that there are no boundary terms in (2.11) because we chose $p \in H_0^1(\Omega)$ due to the Dirichlet boundary condition (2.2). Now, we can define weak solutions of (2.1)-(2.2).

Definition 2.9 (Weak solution)**.** *We call $y \in H_0^1(\Omega)$ a weak solution of the boundary value problem (2.1)-(2.2) if the variational equality*

$$a(y, p) = b(p) \quad \forall p \in H_0^1(\Omega) \tag{2.12}$$

is satisfied, where the bilinear form $a(\cdot, \cdot)$ is given by the left-hand side and the linear form $b(\cdot)$ is given by the right-hand side of the variational formulation (2.11).

Parabolic PDEs and their weak solutions. Analogous to the elliptic case, we consider the operator L_p in divergence form (2.3), assume $a^{ij}, b^i, c \in L^\infty(\Omega)$ and suppose $a^{ij} = a^{ji}$, where $i, j \in \{1, ..., d\}$, in order to define weak solutions of the second-order parabolic PDEs (2.5)-(2.7). Moreover, we assume again that $\partial\Omega$ is $\mathcal{C}^1$. Often, it is easier to identify functions $y \colon \Omega \times [0, T] \to \mathbb{R}$ with mappings from $[0, T]$ into a Banach space. In other words, we switch our point of view by considering y not as a function of $x \in \Omega$ and $t \in [0, T]$ together, but rather as a mapping from $[0, T]$ into a Banach space of functions on Ω. As in the elliptic case, we have to rewrite equation (2.5) in variational form. Note that there is a time derivative in (2.5) such that we do not only have to integrate over Ω, but also in time. Recall that in the elliptic case, we require merely the existence of weak first-order partial derivatives. The boundary value problem is brought into a variational form in which half of the derivatives are shifted to the test function p. With respect to t, only weak derivatives come into question. There are two possibilities which lead to a source of asymmetry in the treatment of y and p:

1. We postulate the existence of the weak time derivative of y, which then is not needed for the test function p.

2. We do not postulate the existence of the weak time derivative of y, which then is needed for the test function p.

Before we can give the variational formulation of (2.5)-(2.7), we have to introduce suitable function spaces.

Definition 2.10. *Let H be a Hilbert space and let H' denote its dual space. Moreover, let $[0,T]$ be a time interval with final time $T > 0$. The space $L^2(0,T;H)$ consists of L^2-integrable functions $y\colon [0,T] \to H$ and the space $\mathcal{C}(0,T;L^2(\Omega))$ consists of continuous functions $y\colon [0,T] \to H$. A weak time derivative of a function $y \in L^2(0,T;H)$ is denoted by $\mathring{y}$ and given by the following condition:*

$$\int_0^T \phi\, \mathring{y}\, dt = -\int_0^T \frac{\partial \phi}{\partial t} y\, dt \quad \forall \phi \in \mathcal{C}_0^\infty([0,T])$$

We define the linear space of all $L^2(0,T;H)$-functions having weak first time derivatives in $L^2(0,T;H')$ by

$$W(0,T;H) \coloneqq \{y \in L^2(0,T;H)\colon \mathring{y} \in L^2\left(0,T;H'\right) \text{ exists}\}. \tag{2.13}$$

In this thesis, we consider $H = H^1(\Omega)$ or $H = H_0^1(\Omega)$. Note that in this setting, the formula of integration by parts holds for all elements of $W(0,T;H)$. This is because every element of $W(0,T;H)$ coincides – possibly after a suitable modification on a set of zero measure – with an element of $\mathcal{C}(0,T;L^2(\Omega))$, i.e., $W(0,T;H)$ is continuously embedded in the space $\mathcal{C}(0,T;L^2(\Omega))$. The proof of this statement can be found, for example, in [109, 110]. For further properties of the space $W(0,T;H)$ we refer to the literature, e.g., [37, 100].

We choose the second possibility of the above-mentioned two in this thesis, i.e., we postulate the existence of the weak time derivative of p, which then is not needed for y. This means, after multiplying the PDE by a smooth function of compact support and after integrating over Ω by applying the integration by parts formula (2.8), we integrate by parts over time in order to transfer the time derivative of y to the test function. In summary, we choose $p \in W(0,T;H_0^1(\Omega))$ and $y \in L^2(0,T;H_0^1(\Omega))$. Note that we choose $H_0^1(\Omega)$ due to the Dirichlet boundary condition (2.6). Moreover, note that the term $y(T)$ is not necessarily defined, because functions $y \in L^2(0,T;H_0^1(\Omega))$ need not be continuous in time. The adjoint variable $p \in W(0,T;H_0^1(\Omega))$ has higher regularity. In particular, $p(0)$ and $p(T)$ are well defined as traces in $L^2(\Omega)$ for all $p \in W(0,T;H_0^1(\Omega))$ (cf. [56]). Therefore, we require $p(T) = 0$ in order to get rid of the term $y(T)$ which would arise by applying integration by parts in time. The variational formulation can be given in analogy to the elliptic case:

$$a(y,p) = b(p) \quad \forall p \in W(0,T;H_0^1(\Omega)), \tag{2.14}$$

where the bilinear form is given by

$$\begin{aligned} a(y,p) = &-\int_\Omega g p(0)\, dx - \int_0^T \int_\Omega y \frac{\partial p}{\partial t}\, dx\, dt \\ &+ \int_0^T \int_\Omega \sum_{i,j=1}^d a^{ij} y_{x_i} p_{x_j} + \sum_{i=1}^d b^i y_{x_i} p + cyp\, dx\, dt \end{aligned} \tag{2.15}$$

and the linear form is given by

$$b(p) = \int_0^T \int_\Omega fp \, dx \, dt. \tag{2.16}$$

Now, we define weak solutions of (2.5)-(2.7).

Definition 2.11 (Weak solution)**.** *We call $y \in L^2(0,T;H_0^1(\Omega))$ a weak solution of the boundary value problem (2.5)-(2.7) if the variational equality*

$$a(y,p) = b(p) \quad \forall p \in W(0,T;H_0^1(\Omega)) \text{ s.t. } p(T) = 0 \tag{2.17}$$

is satisfied, where the bilinear form $a(\cdot,\cdot)$ is given by the left-hand side and the linear form $b(\cdot)$ is given by the right-hand side of the variational formulation (2.14).

2.2.2 Regularity

In this brief subsection, we address the question whether *a weak solution of an elliptic PDE is in fact smooth.* We touch only a few aspects of regularity theory. For the proofs, more details or the regularity results of parabolic PDEs we refer to the literature, e.g., [31].

A theorem of higher regularity is needed in Chapter 7.

Theorem 2.12 (Higher regularity)**.** *Let $m \in \mathbb{N}$. Assume $a^{ij}, b^i, c \in \mathcal{C}^m(\overline{\Omega})$ for all $i,j \in \{1,...,d\}$ and $f \in H^m(\Omega)$. Suppose $u \in H^1(\Omega)$ is a weak solution of (2.1). Then $u \in H^{m+2}_{\text{loc}}(\Omega)$.*

In the above theorem, $H^{m+2}_{\text{loc}}(\Omega)$ denotes the space of all locally H^{m+2}-functions in Ω. A function is called *locally H^{m+2} in Ω* if it is a H^{m+2}-function in every compact subset of Ω. Theorem 2.12 can be extended up to the boundary.

Theorem 2.13 (Higher regularity up to the boundary)**.** *Let $m \in \mathbb{N}$. Assume $a^{ij}, b^i, c \in \mathcal{C}^{m+1}(\overline{\Omega})$ for all $i,j \in \{1,...,d\}$, $f \in H^m(\Omega)$ and $\partial\Omega$ is $\mathcal{C}^{m+2}$. Suppose $u \in H_0^1(\Omega)$ is a weak solution of the boundary value problem (2.1)-(2.2). Then $u \in H^{m+2}(\Omega)$.*

Besides the theorems of higher regularity, we require theorems of infinite differentiability in Chapter 6.

Theorem 2.14 (Infinite differentiability)**.** *Let $a^{ij}, b^i, c \in \mathcal{C}^\infty(\Omega)$ for all $i,j \in \{1,...,d\}$ and $f \in \mathcal{C}^\infty(\Omega)$. Suppose $u \in H^1(\Omega)$ is a weak solution of (2.1). Then $u \in \mathcal{C}^\infty(\Omega)$.*

This theorem can also be extended up to the boundary.

Theorem 2.15 (Infinite differentiability up to the boundary)**.** *Let $a^{ij}, b^i, c \in \mathcal{C}^\infty(\overline{\Omega})$ for all $i,j \in \{1,...,d\}$, $f \in \mathcal{C}^\infty(\overline{\Omega})$ and $\partial\Omega$ is $\mathcal{C}^\infty$. Suppose $u \in H_0^1(\Omega)$ is a weak solution of the boundary value problem (2.1)-(2.2). Then $u \in \mathcal{C}^\infty(\overline{\Omega})$.*

2.3 Fundamentals of differential topology and geometry

In *differential topology*, homotopy classes of maps and the possibility of finding suitable differentiable maps in them, e.g., immersions, embeddings or isomorphisms, are studied. In contrast to this, in *differential geometry*, additional structures, e.g., vector fields or Riemannian metrics, are put on differentiable manifolds and properties connected with these objects are studied. This section provides definitions and theorems from differential topology and geometry which are required in this thesis. For more details we refer to the literature, e.g., [52, 53, 57, 59]. Subsection 2.3.1 presents manifold definitions and introduces related objects like tangent spaces, vector bundles and mappings between manifolds. The Riemannian geometric part of this section, Subsection 2.3.2, defines covariant derivatives, parallel transports, geodesics and the exponential map. It closes with an important remark about quotient manifolds.

2.3.1 A few manifold definitions and related objects

We start with some basic manifold definitions and related objects. Manifolds are based on topological spaces. For the convenience of the reader we repeat their definition in the sequel.

Definition 2.16 (Topological space)**.** *A set X together with a collection $\mathcal{V}$ of subsets of X is called a topological space if the following axioms are fulfilled:*

(i) *$\emptyset$ and X belong to $\mathcal{V}$.*

(ii) *The union of any collection of members of $\mathcal{V}$ belongs to $\mathcal{V}$.*

(iii) *The intersection of any finite number of elements of $\mathcal{V}$ belongs to $\mathcal{V}$.*

It is assumed that the reader is familiar with basic topological concepts as there are, for example, homomorphisms, isomorphisms, Hausdorff spaces, Fréchet spaces, etc. Roughly speaking, we get a manifold by glueing together open subsets of topological vector spaces with isomorphisms. Let $\mathcal{E}$ be a class of topological vector spaces. More precisely, let it be either of the following classes:

- Let $\mathcal{E}_{fd}$ denote the class which consists of finite dimensional vector spaces with the Euclidean topology.
- Let $\mathcal{E}_H$ denote the class which consists of Hilbert spaces.
- Let $\mathcal{E}_B$ denote the class which consists of Banach spaces.
- Let $\mathcal{E}_F$ denote the class which consists of Fréchet spaces.

Each of these classes is a subclass of $\mathcal{E}_F$. There is a well-developed theory of differentiability for maps between open subsets of Fréchet spaces, which we use in the following definition.

Definition 2.17 (Manifold)**.** *Let $\mathcal{E}$ be a class of topological vector spaces.*

(i) Let M be a topological space. An $\mathcal{E}$-chart for M is a pair (U_α, ϕ_α) consisting of an open subset $U_\alpha \subset M$ and a map $\phi_\alpha \colon U_\alpha \to \phi_\alpha(U_\alpha) \subset E_{U_\alpha}$ such that E_{U_α} is of class $\mathcal{E}$, $\phi_\alpha(U_\alpha) \subset E_{U_\alpha}$ is an open subset and ϕ_α is an isomorphism. An $\mathcal{E}$-atlas for M is a collection $\{(U_\alpha, \phi_\alpha)\}_\alpha$ of $\mathcal{E}$-charts such that $\{U_\alpha\}_\alpha$ is an open covering of M. A topological space together with an $\mathcal{E}$-atlas is called a (topological) manifold modelled on $\mathcal{E}$ or $\mathcal{E}$-manifold.

(ii) Let M be a manifold modelled on $\mathcal{E}$. Then M is called a finite dimensional manifold if $\mathcal{E} = \mathcal{E}_{fd}$, a Hilbert manifold if $\mathcal{E} = \mathcal{E}_H$, a Banach manifold if $\mathcal{E} = \mathcal{E}_B$ and a Fréchet manifold if $\mathcal{E} = \mathcal{E}_F$.

(iii) Let M be an $\mathcal{E}$-manifold with atlas $\{(U_\alpha, \phi_\alpha)\}_\alpha$ and let $k \in \overline{\mathbb{N}}$. For each pair of indices α, β the map $\phi_\beta \circ \phi_\alpha^{-1} \colon \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$ is an isomorphism between open subsets of the $\mathcal{E}$-vector spaces E_{U_α} and E_{U_β}. These maps are called the transition maps. An $\mathcal{E}$-manifold is called a differentiable manifold when its transition maps are diffeomorphisms. The atlas $\{(U_\alpha, \phi_\alpha)\}_\alpha$ of M is called a $\mathcal{C}^k$-atlas if for each pair of indices α, β the transition maps are diffeomorphisms of class $\mathcal{C}^k$. A chart (U, ϕ) of M is called $\mathcal{C}^k$-compatible with a $\mathcal{C}^k$-atlas $\{(U_\alpha, \phi_\alpha)\}_\alpha$ of M if $\{(U_\alpha, \phi_\alpha)\}_\alpha \cup (U, \phi)$ is a $\mathcal{C}^k$-atlas of M. A $\mathcal{C}^k$-atlas of M is called maximal if any $\mathcal{C}^k$-compatible chart of M is contained in it. A manifold with a maximal $\mathcal{C}^k$-atlas is called a $\mathcal{C}^k$-manifold.

(iv) A subset N of a manifold M is called a submanifold if for each $q \in N$ there is a chart (U_α, ϕ_α) of M such that $\phi_\alpha(U_\alpha \cap N) = \phi_\alpha(U_\alpha) \cap F_{U_\alpha}$, where F_{U_α} is a closed linear subspace of E_{U_α}. Then N is clearly itself a manifold with $\left(U_\alpha \cap N, \phi_\alpha|_{U_\alpha \cap N}\right)$ as charts.

In this thesis, we work with infinite dimensional manifolds. In general, they are modelled on Banach spaces. If we have Riemannian geometry structures, which are defined in Subsection 2.3.2, these infinite dimensional manifolds are modelled on Hilbert spaces. Manifolds of smooth mappings between finite dimensional manifolds are the foremost examples of infinite dimensional manifolds. In the following, we work with the space of embeddings and immersions from the unit circle into the plane, which are infinite dimensional manifolds, more precisely Fréchet manifolds. Before we can define embeddings and immersions, we have to introduce tangent spaces. There are various ways of defining them:

- *Geometric* via velocities of curves.
- *Algebraic* via derivations.
- *Physical* via cotangent spaces.

Among these three possibilities, we pick the geometric definition. For the other two we refer to the literature, e.g., [53].

Definition 2.18 (Tangent space)**.** *Let M be a differentiable manifold with atlas $\{(U_\alpha, \phi_\alpha)\}_\alpha$ and let $p \in M$. A curve $\gamma\colon \mathbb{R} \to M$ with $\gamma(0) = p$ is called differentiable if the composition $\phi_\alpha \circ \gamma$ is differentiable at $t = 0$ for all α with $p \in U_\alpha$. Two differentiable curves $\gamma_0, \gamma_1\colon \mathbb{R} \to M$ with $\gamma_0(0) = \gamma_1(0) = p$ are called p-equivalent if*

$$\frac{d}{dt}\Big|_{t=0} \phi_\alpha(\gamma_0(t)) = \frac{d}{dt}\Big|_{t=0} \phi_\alpha(\gamma_1(t))$$

holds for all α with $p \in U_\alpha$. We write $\gamma_0 \overset{p}{\sim} \gamma_1$ if γ_0 is p-equivalent to γ_1. The (geometric) tangent space of M at p is defined as the set of equivalence classes

$$T_pM \coloneqq \{\gamma\colon \mathbb{R} \to M\colon \gamma \text{ differentiable, } \gamma(0) = p\}/\overset{p}{\sim}. \tag{2.18}$$

An element of T_pM is called a tangent vector of M with foot point p.

The disjoint union of all tangent spaces of a differentiable manifold leads to a so-called tangent bundle. It is an example of a vector bundle.

Definition 2.19 (Vector bundle)**.** *Let M be a connected $\mathcal{C}^k$-manifold with $k \in \overline{\mathbb{N}}$ and let $\{U_\alpha\}_\alpha$ be an open covering of M. A vector bundle over M is a triple (E, π, M) consisting of a topological space E, the manifold M and a continuous surjective map $\pi\colon E \to M$ such that for each α there is a mapping $\tau_\alpha\colon \pi^{-1}(U_\alpha) \to U_\alpha \times E_0$, where E_0 denotes a Fréchet space, satisfying the following conditions:*

(i) The map τ_α is a $\mathcal{C}^k$-isomorphism commuting with the projection on U_α, i.e., the following diagram is commutative:

$$\begin{array}{ccc} \pi^{-1}(U_\alpha) & \longrightarrow & U_\alpha \\ {\scriptstyle \tau_\alpha}\big\downarrow & \nearrow & \\ U_\alpha \times E_0 & & \end{array}$$

In particular, we obtain an isomorphism $\tau_\alpha(x)\colon \pi^{-1}(x) \to \{x\} \times E_0$ on each fiber $E_x \coloneqq \pi^{-1}(x)$.

(ii) On each fiber $\pi^{-1}(x)$, there is given a structure of a Fréchet space. Moreover, the trivializing map $\tau_\alpha(x)\colon \pi^{-1}(x) = E_x \to E_0$ is a linear isomorphism for all $x \in U_\alpha$.

(iii) If U_α and U_β are two members of the open covering $\{U_\alpha\}_\alpha$, then the mapping $U_\alpha \cap U_\beta \to L(E_0, E_0)$, $x \mapsto (\tau_\beta \circ \tau_\alpha^{-1})(x)$ is a continuous linear map, where $L(E_0, E_0)$ denotes the space of continuous linear maps from the Fréchet space E_0 into itself.

We call M the base space and E the total space of the bundle. Moreover, the mapping π is called the bundle projection.

Note that in the finite dimensional case, condition (iii) in the above definition is implied by condition (ii) as shown in [57]. An example of a vector bundle is the *tangent bundle* (TM, π, M) of a manifold M, where $TM \coloneqq \bigcup_{p\in M} T_pM$. Note that there is a natural projection $\pi\colon TM \to M,\ T_pM \ni v \mapsto p$ since each $v \in TM$ is in one and only one tangent space T_pM. A *vector field* on a manifold M is a function from M into TM that assigns to each point $p \in M$ a tangent vector $v \in T_pM$.

Suppose that $f\colon M \to N$ is a map between two differentiable manifolds M and N. Moreover, let $p \in M$ be given. We define the differential of f at p in the next definition. It is a linear map $df_p\colon T_pM \to T_{f(p)}N$. The application of df_p to a tangent vector is called the pushforward. The definition of a pushforward depends on the used tangent space. We define it for our setting, where tangent vectors are defined as equivalence classes of curves.

Definition 2.20 (Differential of a mapping between manifolds, pushforward)**.** *Let M and N denote differentiable manifolds with atlases $\{(U_\alpha, \phi_\alpha)\}_\alpha$ and $\{(V_\beta, \psi_\beta)\}_\beta$. A mapping $f\colon M \to N$ is called $\mathcal{C}^k$, where $k \in \overline{\mathbb{N}}$, if $\psi_\beta \circ f \circ \phi_\alpha^{-1}$ is $\mathcal{C}^k$ for all charts (U_α, ϕ_α) and (V_β, ψ_β) with $f(U_\alpha) \subset V_\beta$. We denote by $\mathcal{C}^k(M, N)$ the space of all k-times continuously differentiable functions $f\colon M \to N$. The derivative of f at p is defined by*

$$df_p\colon T_pM \to T_{f(p)}N,\ \gamma \mapsto df_p(\gamma) \coloneqq \frac{d}{dt}\Big|_{t=0} (f \circ \gamma)(t).$$

The tangent vector $df_p(\gamma)$ is called the pushforward of a vector $\gamma \in T_pM$ by f.

Now, we are able to define the above-mentioned immersions and embeddings. Moreover, we need the definition of a submersion.

Definition 2.21 (Immersion, embedding, submersion)**.** *Let M and N be manifolds. Moreover, let $f\colon M \to N$ be a $\mathcal{C}^\infty$-differentiable mapping.*

(i) The map f is called an immersion if its differential $df_p\colon T_pM \to T_{f(p)}N$ is injective at every point $p \in M$.

(ii) The map f is called an embedding if it is a proper injective immersion, where f is proper if the preimage $f^{-1}(K) = \{p \in M\colon f(p) \in K\}$ is compact for every compact subset $K \subset N$.

(iii) The map f is called a submersion if its differential $df_p\colon T_pM \to T_{f(p)}N$ is surjective at every point $p \in M$.

The set of all immersions $f\colon M \to N$ is denoted by $\mathrm{Imm}(M, N)$ *and the set of all embeddings $f\colon M \to N$ is denoted by* $\mathrm{Emb}(M, N)$.

An example of a proper immersion is the map $f\colon S^1 \to \mathbb{R}^2,\ (x, y) \mapsto (x, xy)$, which is not an embedding because the points $(0, 1)$, $(0, -1)$ have the same image under f. In Figure 2.1, an example of an immersion and an embedding are visualized.

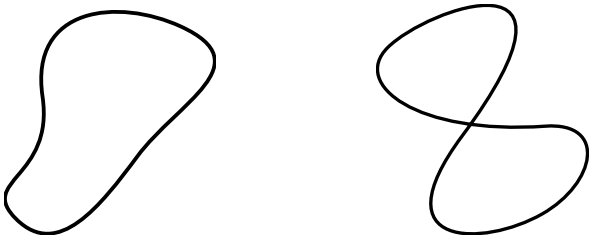


Figure 2.1: Example of an embedding (left) and an immersion which is not an embedding (right).

As already mentioned, we work with the space of embeddings and immersions from the unit circle into the plane. These spaces are Fréchet manifolds due to the following theorem.

Theorem 2.22. *Let M and N be differentiable finite dimensional manifolds.*

(i) The inclusions $\mathrm{Emb}(M,N) \subset \mathrm{Imm}(M,N) \subset \mathcal{C}^\infty(M,N)$ *hold.*

(ii) If $\dim(M) \leq \dim(N)$ *and M is compact, then the space* $\mathcal{C}^\infty(M,N)$ *is a Fréchet manifold.*

(iii) If $\dim(M) \leq \dim(N)$ *and M is compact, then the spaces* $\mathrm{Emb}(M,N)$ *and* $\mathrm{Imm}(M,N)$ *are Fréchet manifolds.*

Proof. The inclusions in (i) are obvious. For (ii) see Proposition 42.3 in [52]. If $\dim(M) \leq \dim(N)$ and M is compact, the spaces $\mathrm{Emb}(M,N)$, $\mathrm{Imm}(M,N)$ are open in the Fréchet manifold $\mathcal{C}^\infty(M,N)$ as shown in [52, Theorem 44.1]. Thus, they are also Fréchet manifolds. □

2.3.2 Riemannian geometry

In the sequel, we introduce the notion of a Riemannian manifold and give related notations as well as basic theorems without proofs. A Riemannian metric provides us with an inner product on each tangent space and can be used to measure the length of curves in a manifold. In this context, a distance function can be defined. This turns the manifold into a metric space in a natural way.

Definition 2.23 (Riemannian manifold)**.** *Let M be a differentiable manifold. A Riemannian metric on M is a collection $g = (g_p)_{p\in M}$ of inner products*

$$g_p \colon T_pM \times T_pM \to \mathbb{R},\ (v,w) \mapsto g_p(v,w),$$

one for every $p \in M$, such that the map $M \to \mathbb{R},\ p \mapsto g_p(X(p),Y(p))$ is smooth for every pair of vector fields X, Y on M. A differentiable manifold equipped with a Riemannian metric is called a Riemannian manifold and denoted by (M,g).

It is not completely obvious that a manifold admits a Riemannian metric. The next theorem gives an answer to this question.

Theorem 2.24. *All paracompact connected Hausdorff manifolds admit Riemannian metrics.*

In the following, let M be connected and equipped with a Riemannian metric $g = (g_p)_{p\in M}$. We define the length of a $\mathcal{C}^1$-curve $\gamma\colon [0,1] \to M$ by

$$L(\gamma) \coloneqq \int_0^1 \sqrt{g_p(\dot\gamma(t), \dot\gamma(t))}\, dt. \tag{2.19}$$

Note that it is invariant under re-parametrizations. Moreover, we define a distance function by

$$d\colon M \times M \to [0,\infty),\ (p,q) \mapsto \inf_{\gamma\in X_{p,q}^M} L(\gamma), \tag{2.20}$$

where

$$X_{p,q}^M \coloneqq \{\gamma\colon [0,1] \to M\colon \gamma \text{ differentiable with } \gamma(0) = p,\ \gamma(1) = q\} \tag{2.21}$$

denotes the space of differentiable paths in M connecting two points $p, q \in M$. Note that this space is non-empty if M is connected. The distance function d defines a metric in M in the traditional sense, the *Riemannian distance*, and induces the same topology as the given atlas. Now, we can ask whether *there is a smooth curve γ with fixed endpoints p, q such that γ minimizes the length functional $L\colon X_{p,q}^M \to \mathbb{R}$.* Geodesics give an answer to this question. In the following, we see that shortest curves between two fixed points in a manifold are geodesics (cf. Theorem 2.31). Thus, the Riemannian distance is also called the *geodesic distance*. To define geodesics, we have to introduce the covariant derivative. Note that there always exists a unique Riemannian connection on a Riemannian manifold (see, for example, [53]).

Definition 2.25 (Covariant derivative)**.** *Let (TM, π, M) be the tangent bundle of a Riemannian manifold (M, g) and let $\Gamma(TM)$ denote the set of all differentiable vector fields on M. A covariant derivative on (TM, π, M), which is also called connection, is a map*

$$\nabla^{\text{cov}}\colon \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM),\ (X,Y) \mapsto \nabla_X^{\text{cov}} Y \tag{2.22}$$

such that

(i) $\nabla_X^{\text{cov}}(\lambda Y + \mu Z) = \lambda \nabla_X^{\text{cov}} Y + \mu \nabla_X^{\text{cov}} Z$,

(ii) $\nabla_X^{\text{cov}}(fY) = X(f) \cdot Y + f \cdot \nabla_X^{\text{cov}} Y$,

(iii) $\nabla_{f\cdot X + g\cdot Y}^{\text{cov}} Z = f \cdot \nabla_X^{\text{cov}} Z + g \cdot \nabla_Y^{\text{cov}} Z$

hold for all $\lambda, \mu \in \mathbb{R}$, $X, Y, Z \in \Gamma(TM)$ and $f, g \in \mathcal{C}^\infty(M)$. In (ii), $X(f) = \nabla_X^{\text{cov}} f$ denotes the directional derivative of f in direction X. A connection ∇^{cov} is called torsion-free if

(iv) $[X,Y] = \nabla^{\mathrm{cov}}_X Y - \nabla^{\mathrm{cov}}_Y X$

holds for all $X, Y \in \Gamma(TM)$*, where* $[\cdot,\cdot]$ *denotes the Lie bracket on* $\Gamma(TM)$ *defined by* $[X,Y] \coloneqq X(Y(f)) - Y(X(f))$ *for all* $f \in \mathcal{C}^\infty(M)$*. A connection is called compatible with the Riemannian metric* g *if*

(v) $X(g(Y,Z)) = g(\nabla^{\mathrm{cov}}_X Y, Z) + g(Y, \nabla^{\mathrm{cov}}_X Z)$

holds for all $X, Y, Z \in \Gamma(TM)$*. A torsion-free connection which is also compatible with a Riemannian metric is called the Riemannian connection or the Levi-Civita connection.*

In order to define geodesics on Riemannian manifolds, we need the following proposition which gives not just a rule for differentiating a vector field along a curve, but also a relation to the Levi-Civita connection.

Theorem 2.26. *Let* (M,g) *denote a differentiable Riemannian manifold and let* $\gamma\colon I \to M$ *be a curve in* M*, where* $I \subset \mathbb{R}$ *denotes an interval. Moreover, let* (TM,π,M) *be the tangential bundle and let* $\Gamma_\gamma(TM)$ *be the set of all differentiable vector fields along* γ*, where a vector field* $X \in \Gamma_\gamma(TM)$ *on* γ *is defined as a curve* $X\colon I \to TM$ *such that* $\pi \circ X = \gamma$*. Furthermore,* ∇^{cov} *denotes the Levi-Civita connection on* (TM,π,M)*. Then there exists a unique operator*

$$\frac{D}{dt}\colon \Gamma_\gamma(TM) \to \Gamma_\gamma(TM)$$

such that for all $\lambda, \mu \in \mathbb{R}$ *and* $f \in \mathcal{C}^\infty(I)$ *the following conditions are fulfilled:*

(i) $\frac{D}{dt}(\lambda X + \mu Y) = \lambda \frac{DX}{dt} + \mu \frac{DY}{dt}$,

(ii) $\frac{D}{dt}(fY) = \frac{Df}{dt} \cdot Y + f \cdot \frac{DY}{dt}$,

(iii) $\frac{D}{dt}(X \circ \gamma) = \nabla^{\mathrm{cov}}_{\dot\gamma} X$.

The Levi-Civita connection can be used to define parallel vector fields and geodesics.

Definition 2.27 (Parallel vector field)**.** *A vector field* X *along a differentiable curve* $\gamma\colon \mathbb{R} \to M$ *in a Riemannian manifold* (M,g) *is called parallel if* $\nabla^{\mathrm{cov}}_{\dot\gamma} X = 0$*, where* ∇^{cov} *denotes the Levi-Civita connection on* (TM,π,M)*.*

The following theorem states that for given initial values at a point in a Riemannian manifold we get a parallel vector field defined globally along any curve through this point.

Theorem 2.28 (Parallel transport)**.** *Let* (M,g) *be a Riemannian manifold and let* $I = (a,b) \subset \mathbb{R}$ *be an open interval. Moreover, let* $\gamma\colon [a,b] \to M$ *be a smooth curve on* I*, let* $t_0 \in I$ *and let* $X_0 \in T_{\gamma(t_0)}M$*. Then there exists a unique parallel vector field* Y *along* γ *such that* $X_0 = Y(t_0)$*. For* $t_0, t \in I$ *we define the map*

$$\mathcal{P}_\gamma(t_0,t)\colon T_{\gamma(t_0)}M \to T_{\gamma(t)}M,\ X_0 \mapsto Y(t),$$

where $Y \in \Gamma_\gamma(TM)$ *is the unique parallel vector field along* γ *satisfying* $X_0 = Y(t_0)$*. The collection of maps* $\mathcal{P}_\gamma(t_0,t)$ *for* $t_0, t \in I$ *is called the parallel transport along* γ*.*

Now, we define geodesics, which are curves $\gamma\colon \mathbb{R} \to M$ in a manifold M whose tangent field $\dot{\gamma}$ is parallel along γ.

Definition 2.29 (Geodesic)**.** *A geodesic on a Riemannian manifold (M, g) is a smooth curve $\gamma\colon \mathbb{R} \to M$ which satisfies the equation $\nabla^{\mathrm{cov}}_{\dot{\gamma}}\dot{\gamma} = 0$, where ∇^{cov} denotes the Levi-Civita connection on (TM, π, M).*

The following fundamental uniqueness and existence result holds for geodesics.

Theorem 2.30. *Let (M, g) be a Riemannian manifold. If $p \in M$ and $v \in T_pM$, then there exists a unique geodesic $\gamma\colon I \to M$ such that $\gamma(0) = p$ and $\dot{\gamma}(0) = v$ in an open interval $I = (-\epsilon, \epsilon) \subset \mathbb{R}$.*

Due to the following theorem, geodesics in a Riemannian manifold (M, g), which connect two different points $p, q \in M$, are critical points of the energy functional

$$E(\gamma) \coloneqq \frac{1}{2}\int_0^1 g(\dot{\gamma}(t), \dot{\gamma}(t))\, dt \tag{2.23}$$

on the space $X^M_{p,q}$ defined in (2.21).

Theorem 2.31. *Let (M, g) be a Riemannian manifold and let $p, q \in M$. Moreover, let $\gamma \in X^M_{p,q}$ be a curve in M, where $X^M_{p,q}$ is defined in (2.21). Furthermore, L denotes the length functional (2.19) and E denotes the energy functional (2.23). Then the following conditions are equivalent:*

(i) $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$*, i.e., γ is a geodesic.*

(ii) Either $p = q$ or $g(\dot{\gamma}, \dot{\gamma}) = c \neq 0$ and γ is a critical point of L, where c is an arbitrary constant.

(iii) γ is a critical point of E.

Due to Theorem 2.30, through each point $p \in M$ and in each direction $v \in T_pM$ there exists a unique geodesic in a Riemannian manifold (M, g). We define the maximal existence interval of such a geodesic by

$$I^M_{p,v} \coloneqq \bigcup_{I \in \tilde{I}^M_{p,v}} I, \tag{2.24}$$

where

$$\begin{aligned}\tilde{I}^M_{p,v} \coloneqq \{I \subset \mathbb{R}\colon I \text{ open}, 0 \in I, \text{ there exists a geodesic } \gamma\colon I \to M \text{ satisfying} \\ \gamma(0) = p \in M, \dot{\gamma}(0) = v \in T_pM\}.\end{aligned} \tag{2.25}$$

This motivates the definition of the exponential map.

Definition 2.32 (Exponential map)**.** *Let (M, g) denote a Riemannian manifold, $p \in M$ and $v \in T_pM$. Moreover, let $I_{p,v}^M$ denote the maximal existence interval (2.24) and $V_p \coloneqq \{v \in T_pM \colon 1 \in I_{p,v}^M\}$. The exponential map is defined by*

$$\exp \colon \bigcup_{p \in M} \{p\} \times V_p \to M,\ (p, v) \mapsto \exp_p(v) \coloneqq \gamma(1), \tag{2.26}$$

where $\exp_p(v)$ denotes the exponential map of M at p which assigns to every tangent vector $v \in V_p$ the point $\gamma(1)$ and $\gamma \colon I_{p,v}^M \to M$ is the unique geodesic satisfying $\gamma(0) = p$ and $\dot{\gamma}(0) = v$.

The following theorem states that a geodesic is given by the exponential map.

Theorem 2.33. *Let (M, g) denote a Riemannian manifold, $p \in M$ and $v \in V_p$ with V_p given as in Definition 2.32. Then $I_{p,v}^M = \{t \in \mathbb{R} \colon tv \in V_p\}$ and the geodesic $\gamma \colon I_{p,v}^M \to M$ with $\gamma(0) = p$, $\dot{\gamma}(0) = v$ is given by $\gamma \colon I_{p,v}^M \to M,\ t \mapsto \gamma(t) = \exp_p(tv)$.*

Besides exponential maps, this thesis deals with retractions. They can be used to locally reduce an optimization problem on a manifold to an optimization problem on its tangent space.

Definition 2.34 (Retraction)**.** *A retraction on a manifold M is a smooth mapping $\mathcal{R} \colon TM \to M$ with the following properties:*

(i) $\mathcal{R}_p(0_p) = p$, where $\mathcal{R}_p$ denotes the restriction of $\mathcal{R}$ to T_pM and 0_p denotes the zero element of T_pM.

(ii) $d\mathcal{R}_p(0_p) = id_{T_pM}$, where id_{T_pM} denotes the identity mapping on T_pM and $d\mathcal{R}_p(0_p)$ denotes the pushforward of $0_p \in T_pM$ by $\mathcal{R}$.

Condition (ii) is called the *local rigidity condition*. Equivalently, for every $v \in T_pM$, the curve $\gamma \colon \mathbb{R} \to M$, $t \mapsto \mathcal{R}_p(tv)$ satisfies $\dot{\gamma}(0) = v$. Moving along this curve γ is thought of as moving in direction v while being constrained to stay on M. Note that every manifold which admits a Riemannian metric also admits a retraction defined by the exponential mapping (cf. [1, Section 5.4]).

In Theorem 2.28, the parallel transport along a curve in a manifold M is defined. However, we need a more general notion, the so-called *vector transport*, which specifies how to transport a tangent vector η form a point $p \in M$ to a point $\mathcal{R}_p(\eta)$, where $\mathcal{R}$ denotes a retraction on M.

Definition 2.35 (Vector transport)**.** *Let (M, g) be a Riemannian manifold and let $\oplus$ denote the Whitney sum, i.e.,*

$$TM \oplus TM \coloneqq \{(\xi_p, \eta_p) \colon \xi_p, \eta_p \in T_pM,\ p \in M\}. \tag{2.27}$$

A vector transport on M is a differentiable mapping

$$\mathcal{T} \colon TM \oplus TM \to TM,\ (\xi_p, \eta_p) \mapsto \mathcal{T}_{\xi_p}(\eta_p)$$

satisfying the following properties for all $p \in M$:

(i) There exists a retraction $\mathcal{R}\colon TM \to M$, called the retraction associated with $\mathcal{T}$, such that the following diagram commutes:

$$\begin{array}{ccc} TM \oplus TM & \xrightarrow{\mathcal{T}} & TM \\ {\scriptstyle pr_1}\downarrow & & \downarrow{\scriptstyle \pi} \\ TM & \xrightarrow{\mathcal{R}} & M \end{array}$$

where $pr_1\colon TM \oplus TM \to TM$ is the projection onto the first component and $\pi\colon TM \to M$ denotes the foot point map.

(ii) $\mathcal{T}_{0_p}(\eta_p) = \eta_p$ for all $\eta_p \in T_pM$, where 0_p denotes the zero element of T_pM.

(iii) $\mathcal{T}_{\xi_p}(\lambda\eta_p + \mu\zeta_p) = \lambda\mathcal{T}_{\xi_p}(\eta_p) + \mu\mathcal{T}_{\xi_p}(\zeta_p)$ for all $\eta_p, \zeta_p \in T_pM$ and all $\lambda, \mu \in \mathbb{R}$.

If a manifold is endowed with a Riemannian metric, we expect that manifolds induced by this manifold, e.g., submanifolds or quotient manifolds, inherit a Riemannian metric in a natural way. We are mainly interested in quotient manifolds. In the following remark, we clarify some quotient manifold concepts.

Remark 2.36 (Quotient manifold)**.** Let $\overline{M}$ be a manifold equipped with an equivalence relation $\sim$ and let M be the corresponding quotient, i.e.,

$$M = \overline{M}/\!\sim \coloneqq \{[x]\colon x \in \overline{M}\},$$

where

$$[x] \coloneqq \{y \in \overline{M}\colon y \sim x\}$$

denotes the equivalence class of an element $x \in \overline{M}$. We call an equivalence relation $\sim$ *admissible* if M is a manifold such that

$$\pi\colon \overline{M} \to \overline{M}/\!\sim,\ x \mapsto [x]$$

is a submersion. The mapping π is called the *canonical projection*. In the following, let $\sim$ be admissible. In this case, M is called the *quotient manifold* of $\overline{M}$ by $\sim$. Let $p \in M$ and $v \in T_pM$. Moreover, let $\overline{p}$ be an element of the equivalence class $\pi^{-1}(p)$. Note that $\pi^{-1}(p)$ admits a tangent space $\mathcal{V}_{\overline{p}} = T_{\overline{p}}(\pi^{-1}(p))$ called the *vertical space* at $\overline{p}$ (cf. [1]). In the following, let the space $\overline{M}$ be endowed with a Riemannian metric $\overline{g}$. A mapping $\mathcal{H}$ is called a *horizontal distribution* on $\overline{M}$ if it assigns to each element $\overline{p} \in \overline{M}$ a subspace $\mathcal{H}_{\overline{p}}$ of $T_{\overline{p}}\overline{M}$ complementary to $\mathcal{V}_{\overline{p}}$ such that $T_{\overline{p}}\overline{M} = \mathcal{H}_{\overline{p}} \oplus \mathcal{V}_{\overline{p}}$. The subspace $\mathcal{H}_{\overline{p}}$ is defined by

$$\mathcal{H}_{\overline{p}} \coloneqq \{\eta \in T_{\overline{p}}\overline{M}\colon \overline{g}(\eta, \theta) = 0\ \forall \theta \in \mathcal{V}_{\overline{p}}\}$$

and called the *horizontal space* at $\overline{p}$. If $\overline{M}$ is endowed with a horizontal distribution, then there exists one and only one element $\overline{v} \in T_{\overline{p}}\overline{M}$ that belongs to $\mathcal{H}_{\overline{p}}$ and satisfies

$d\pi_{\overline{p}}(\overline{v}) = v$ (cf. [1]). This unique vector $\overline{v}$ is called the *horizontal lift* of v at $\overline{p}$. If for every $p \in M$ and every $u, v \in T_pM$ the expression $\overline{g}_{\overline{p}}(\overline{u}, \overline{v})$ does not depend on $\overline{p} \in \pi^{-1}(p)$, then $g_p(u, v) \coloneqq \overline{g}_{\overline{p}}(\overline{u}, \overline{v})$ defines a Riemannian metric on M. Endowed with this Riemannian metric, M is called a *Riemannian quotient manifold* of $\overline{M}$ and the natural projection $\pi\colon \overline{M} \to M$ is a *Riemannian submersion*. In the next chapter, Riemannian metrics are defined on the structure space $\overline{M}$. In this way, the map to the quotient space becomes a Riemannian submersion. In general, assuming that a surjective mapping $\widetilde{\pi}\colon A \to B$ between two manifolds A, B with a surjective tangent map and a Riemannian metric g on A are given, $\widetilde{\pi}$ is a *submersion* if it has the following properties (cf. [64]):

(i) The tangent bundle to A splits into the subbundle $TA^{\top}$ tangent to the fibers of $\widetilde{\pi}$ and its orthogonal complement $TA^{\perp}$ with respect to g. Here $TA^{\top}$ is called the *vertical bundle* and $TA^{\perp}$ is called the *horizontal bundle*.

(ii) Under the isomorphism $d\widetilde{\pi}_x\colon T_xA^{\perp} \to T_{\pi(x)}B$, the restriction of the Riemannian metric g to the horizontal bundle is required to define a Riemannian metric on T_yB independent of the choice of the point $x \in \widetilde{\pi}^{-1}(y)$ in the fiber. In this way, the submersion $\widetilde{\pi}$ becomes a Riemannian submersion.

Chapter 3

Riemannian metrics on a shape manifold

Shapes and their similarities has been extensively studied in recent decades. David G. Kendall [48] has already introduced the notion of a shape space in 1984. In [48], shapes are characterized by labeled points in the Euclidean space, so-called landmarks, and the author investigates Riemannian structures on this space. However, there is a large number of different shape concepts, e.g., landmark vectors [20, 40, 48, 75, 92], plane curves [63, 64, 65, 66], surfaces [11, 12, 50, 54, 62], boundary contours of objects [32, 60, 107], multiphase objects [106], characteristic functions of measurable sets [111] and morphologies of images [25]. In order to answer natural questions like *"How different are shapes?", "Can we determine the measure of their difference?"* or *"Can we infer any information?"* mathematically, we put a metric on the space of shapes. There are various different types of metrics on shape spaces, e.g., inner metrics [11, 12, 64], outer metrics [13, 17, 36, 48, 64], metamorphosis metrics [45, 101], the Wasserstein or Monge-Kantorovic metric on the shape space of probability measures [5, 14, 15], the Weil-Peterson metric [55, 91], current metrics [26, 27, 103] and metrics based on elastic deformations [32, 108]. However, in general, the modelling of both, the shape space and the associated metric, is a challenging task and different approaches lead to diverse models. The suitability of an approach depends on the demands in a given situation. There exists no common shape space or shape metric suitable for all applications.

In this thesis, among all above-mentioned shape space concepts, we pick the shape space introduced by Peter W. Michor and David Mumford. In [63], the authors consider smooth embeddings from the unit circle into the plane modulo diffeomorphisms from the unit circle into itself as the space of all two-dimensional shapes and propose an L^2-metric regularized by the curvature of shape boundaries. This curvature weighted L^2-metric is motivated by the observation that the standard L^2-metric induces vanishing geodesic distance in their shape space (cf. [62]). This chapter is devoted to the definition of this essential Riemannian shape manifold and to one special Riemannian metric on it, the first Sobolev metric. This Riemannian metric is crucial in the following chapters.

This chapter is organized as follows. Section 3.1 introduces the Riemannian shape space of Peter W. Michor and David Mumford, which is essential in Chapter 5. Section 3.2 is devoted to Riemannian metrics on this shape space. Special attention is paid to the first Sobolev metric for which an expression of the covariant derivative is provided. These two first sections are concerned with two-dimensional shapes, which are mainly considered in this thesis. However, for the sake of completeness, this chapter closes with the very short Section 3.3, which generalizes the shape space and its properties to higher dimensions.

3.1 Definition of a shape space

First, we concentrate on two-dimensional shapes. In this subsection, a shape of dimension two is defined as a simply connected and compact subset of $\mathbb{R}^2$ with $\mathcal{C}^\infty$-boundary. Since the boundary of an object or a shape is all that matters, we can think of two-dimensional shapes as the images of simple closed smooth curves in the plane of the unit circle. Such simple closed smooth curves can be represented by embeddings from the circle S^1 into the plane $\mathbb{R}^2$, see for instance [53]. Therefore, the set of all embeddings from S^1 into $\mathbb{R}^2$, denoted by $\mathrm{Emb}(S^1, \mathbb{R}^2)$, represents all simple closed smooth curves in $\mathbb{R}^2$. However, note that we are only interested in the shape itself and that images are not changed by re-parametrizations. Thus, all simple closed smooth curves which differ only by re-parametrizations can be considered equal to each other because they lead to the same image. Let $\mathrm{Diff}(S^1)$ denote the set of all diffeomorphisms from S^1 into itself. This set contains all the smooth re-parametrizations mentioned above. In [63], the *set of all two-dimensional shapes* is characterized by

$$B_e = B_e(S^1, \mathbb{R}^2) \coloneqq \mathrm{Emb}(S^1, \mathbb{R}^2)/\mathrm{Diff}(S^1). \tag{3.1}$$

A particular point on the shape space B_e is represented by a curve

$$c\colon S^1 \to \mathbb{R}^2,\ \theta \mapsto c(\theta)$$

and illustrated in the left picture of Figure 2.1. Because of the equivalence relation $\mathrm{Diff}(S^1)$, the tangent space is isomorphic to the set of all smooth normal vector fields along c, i.e.,

$$T_c B_e \cong \left\{h\colon h = \alpha n,\ \alpha \in \mathcal{C}^\infty(S^1)\right\}, \tag{3.2}$$

where n denotes the exterior unit normal field to the shape boundary c such that $n(\theta) \perp c_\theta(\theta)$ for all $\theta \in S^1$, where $c_\theta = \frac{\partial c}{\partial \theta}$. Since we are dealing with parametrized curves, we have to work with the arc length and its derivative. Therefore, we use the following terminology:

$$ds = |c_\theta| d\theta \qquad \text{(arc length)}$$

$$D_s = \frac{\partial_\theta}{|c_\theta|} \qquad \text{(arc length derivative)}$$

3.2 Riemannian metrics on the shape space

In [52], it is proven that the shape space B_e is a smooth manifold. *Is it even perhaps a Riemannian shape manifold?* This question was investigated by Peter W. Michor and David Mumford. They show in [63] that the standard L^2-metric on the tangent space is too weak because it induces geodesic distance equals zero. This phenomenon is called the *vanishing geodesic distance phenomenon*. The authors employ a curvature weighted L^2-metric as a remedy and prove that the vanishing phenomenon does not occur for this metric. Several Riemannian metrics on this shape space are examined in further publications, e.g., [11, 62, 64]. All these metrics arise from the L^2-metric by putting weights, derivatives or both in it. In this manner, we get three groups of metrics:

- *Almost local metrics* arise by putting weights in the L^2-metric (cf. [10, 12, 64]).
- *Sobolev metrics* arise by putting derivatives in the L^2-metric (cf. [11, 64]).
- *Weighted Sobolev metrics* arise by putting both, weights and derivatives, in the L^2-metric (cf. [12]).

It can be shown that all these metrics do not induce the phenomenon of vanishing geodesic distance under special assumptions. To list all these goes beyond the scope of this thesis, but they can be found in the above-mentioned publications. All Riemannian metrics mentioned above are *inner metrics*. This means that the metric is directly defined on the deformation vector field such that the deformation is prescribed on the shape itself and the ambient space stays fixed. In [32, 54, 66, 106], further inner metrics can be found. In contrast, *outer metrics* define a deformation vector field on the ambient space such that the deformation of the ambient space induces a deformation of the shape. Outer metrics are used, for example, in the method of large deformation diffeomorphic metric matching (LDDMM), e.g., applied to problems in computational anatomy. For any information about LDDMM we refer to the literature, e.g., [13, 36, 38, 78].

In the following, we clarify how the above-mentioned inner Riemannian metrics can be defined on the shape space B_e. For this purpose, we need a larger space of shapes, which has a nice property. We introduce the space

$$B_i = B_i(S^1, \mathbb{R}^2) \coloneqq \mathrm{Imm}(S^1, \mathbb{R}^2)/\mathrm{Diff}(S^1) \supset B_e, \tag{3.3}$$

where $\mathrm{Imm}(S^1, \mathbb{R}^2)$ denotes the set of all immersions from S^1 into $\mathbb{R}^2$. Note that we can identify a closed smooth curve with an immersion from the circle into the plane (cf. [53]). Thus, the space B_i consists of all closed smooth curves which differ only by re-parametrizations. Figure 2.1 illustrates shapes of both spaces, B_e and B_i. The embedding shows a shape which is an element of $B_e \subset B_i$ and the immersion illustrates a shape which is an element of B_i, but not an element of B_e. In [19],

it is shown that B_i is an orbifold, which is a generalization of a manifold (for an introduction to orbifolds see, for example, [2]). Furthermore, the mapping

$$\pi \colon \mathrm{Imm}(S^1, \mathbb{R}^2) \to B_i \tag{3.4}$$

is a submersion. Now, we use the orbifold structure to define Riemannian metrics on the shape space. The important point to note here is that we want to define an inner metric. This means that we have to define a Riemannian metric on $\mathrm{Imm}(S^1, \mathbb{R}^2)$ such that the submersion π induces a Riemannian metric on the shape space. A Riemannian metric on $\mathrm{Imm}(S^1, \mathbb{R}^2)$ is a family $g = (g_c(h,k))_{c \in \mathrm{Imm}(S^1,\mathbb{R}^2)}$ of inner products $g_c(h,k)$, where h and k denote vector fields along $c \in \mathrm{Imm}(S^1, \mathbb{R}^2)$. The most simple inner product on the tangent bundle to $\mathrm{Imm}(S^1, \mathbb{R}^2)$ is the standard L^2-inner product $g_c(h,k) \coloneqq \int_{S^1} \langle h, k \rangle \, ds$. Note that $T_c\mathrm{Imm}(S^1, \mathbb{R}^2) \cong \mathcal{C}^\infty(S^1, \mathbb{R}^2)$ for all $c \in \mathrm{Imm}(S^1, \mathbb{R}^2)$. Moreover, note that the differential ds is invariant under the action of $\mathrm{Diff}(S^1)$. Because of the properties of submersions (cf. Remark 2.36), the tangent bundle to $\mathrm{Imm}(S^1, \mathbb{R}^2)$ splits into subbundles, tangent and horizontal to the fibers of π with respect to the metric. The restriction of the metric to this horizontal subspace gives the quotient metric such that the submersion π is Riemannian. More precisely, the quotient metric induced by the L^2-metric is given by

$$\begin{aligned} g^0 \colon T_cB_e \times T_cB_e &\to \mathbb{R}, \\ (h,k) &\mapsto \int_{S^1} \langle \alpha, \beta \rangle \, ds = \langle \alpha, \beta \rangle_{L^2(S^1)}, \end{aligned} \tag{3.5}$$

where $h = \alpha n$ and $k = \beta n$ denote two elements of the tangent space T_cB_e given in (3.2). Unfortunately, in [63], Peter W. Michor and David Mumford show that this L^2-metric induces vanishing geodesic distance, as already mentioned above.

For the discussion in the following chapters, among all the above-mentioned Riemannian metrics, we pick the Sobolev metric g^1 defined in the sequel. In the following, some results on it are provided. For more details about the other Riemannian metrics we refer to the literature, e.g., [10, 11, 12, 62, 64].

Definition 3.1 (Sobolev metric)**.** *Let $n \in \mathbb{N}^*$. The n-th Sobolev metric is defined by*

$$\begin{aligned} g^n \colon T_c\mathrm{Imm}(S^1, \mathbb{R}^2) \times T_c\mathrm{Imm}(S^1, \mathbb{R}^2) &\to \mathbb{R}, \\ (h,k) &\mapsto \int_{S^1} \left\langle \left(I + (-1)^n A D_s^{2n}\right) h, k \right\rangle ds, \end{aligned}$$

where $A > 0$ denotes the metric parameter.

In particular, due to Definition 3.1 and the isomorphism of the tangent space T_cB_e given in (3.2), the first Sobolev metric on B_e is defined by

$$\begin{aligned} g^1 \colon T_cB_e \times T_cB_e &\to \mathbb{R}, \\ (h,k) &\mapsto \int_{S^1} \langle (I - A\triangle_c)\, \alpha, \beta \rangle \, ds = \langle (I - A\triangle_c)\, \alpha, \beta \rangle_{L^2(S^1)}, \end{aligned} \tag{3.6}$$

where $h = \alpha n$ and $k = \beta n$ denote two elements of the tangent space $T_c B_e$ and $\triangle_c$ denotes the Laplace-Beltrami operator on the surface c. The following theorem, which is proven in [11], states under which assumptions the phenomenon of the vanishing geodesic distance does not occur for the Sobolev metric.

Theorem 3.2. *Let $n \in \mathbb{N}^*$. The Sobolev metric g^n induces non-vanishing geodesic distance on B_e if it is stronger or as strong as the g^1-metric, i.e., if there is a constant $k > 0$ such that*

$$\|h\|_{g_c^n} \geq k \, \|h\|_{g_c^1} \tag{3.7}$$

holds for all $h \in T_c\mathrm{Imm}(S^1, \mathbb{R}^2)$ and all $c \in \mathrm{Imm}(S^1, \mathbb{R}^2)$.

An essential operation in Riemannian geometry is the covariant derivative. In differential geometry, it is often written in terms of the Christoffel symbols. In [11], Christoffel symbols associated with the Sobolev metrics g^n are provided. However, in order to provide a relation with shape calculus in the next chapters, another representation of the covariant derivative in terms of the Sobolev metric g^1 is needed. The Riemannian connection provided in the following theorem makes it possible to specify the Riemannian shape Hessian in Section 4.3.

Theorem 3.3. *Let $A > 0$ and let $h, m \in T_c\mathrm{Imm}(S^1, \mathbb{R}^2)$ denote vector fields along $c \in \mathrm{Imm}(S^1, \mathbb{R}^2)$. Moreover, $L_1 \coloneqq I - AD_s^2$ is a differential operator on $\mathcal{C}^\infty(S^1, \mathbb{R}^2)$ and L_1^{-1} denotes its inverse operator. The covariant derivative associated with the Sobolev metric g^1 can be expressed as*

$$\nabla_m h = L_1^{-1}(K_1(h)) \text{ with } K_1 \coloneqq \frac{1}{2} \langle D_s m, v \rangle \left(I + AD_s^2 \right), \tag{3.8}$$

where $v = \frac{c_\theta}{|c_\theta|}$ denotes the unit tangent vector.

Proof. Let h, k, m be vector fields on $\mathbb{R}^2$ along $c \in \mathrm{Imm}(S^1, \mathbb{R}^2)$. Moreover, $d(\cdot)[m]$ denotes the directional derivative in direction m. By [63],

$$d(|c_\theta|)[m] = \frac{\langle m_\theta, c_\theta \rangle}{|c_\theta|} \tag{3.9}$$

holds. From

$$d(|c_\theta|)[m] \overset{(3.9)}{=} \left\langle m_\theta, \underbrace{\frac{c_\theta}{|c_\theta|}}_{=v} \right\rangle = \left\langle \underbrace{\frac{\partial_\theta}{|c_\theta|}}_{=D_s} m|c_\theta|, v \right\rangle = \langle D_s m, v \rangle \, |c_\theta| \tag{3.10}$$

we get

$$d(|c_\theta|^k)[m] = k|c_\theta|^{k-1} d(|c_\theta|)[m] \overset{(3.10)}{=} k \, \langle D_s m, v \rangle \, |c_\theta|^k. \tag{3.11}$$

Applying (3.10), we obtain

$$\begin{aligned} d(L_1)[m] &= d\left(I - AD_s^2\right)[m] = d\left(I - A\frac{\partial_\theta^2}{|c_\theta|^2}\right)[m] = -A \, d\left(|c_\theta|^{-2}\right)[m] \, \partial_\theta^2 \\ &\overset{(3.10)}{=} 2A \, \langle D_s m, v \rangle \, \partial_\theta^2 |c_\theta|^{-2} = 2A \, \langle D_s m, v \rangle \, D_s^2 \, . \end{aligned} \tag{3.12}$$

Combining (3.10) with (3.12) we get

$$\begin{aligned}
&d\left(g_c^1(h,k)\right)[m] = d\left(\int_{S^1}\langle L_1(h),k\rangle\, ds\right)[m] \\
&= \int_{S^1}\langle d\left(L_1(h)\right)[m],k\rangle\, ds + \int_{S^1}\langle L_1(h),k\rangle\, d\left(|c_\theta|\right)[m]d\theta \\
&\overset{(3.10)}{\underset{(3.12)}{=}} \int_{S^1}\left\langle 2A\left\langle D_s m,v\right\rangle D_s^2 h,k\right\rangle ds + \int_{S^1}\langle L_1(h),k\rangle\left\langle D_s m,v\right\rangle \underbrace{|c_\theta|d\theta}_{=ds} \\
&= \int_{S^1} 2A\left\langle D_s m,v\right\rangle\left\langle D_s^2 h,k\right\rangle ds + \int_{S^1}\langle h,k\rangle\left\langle D_s m,v\right\rangle ds \\
&\quad - \int_{S^1} A\left\langle D_s^2 h,k\right\rangle\left\langle D_s m,v\right\rangle ds \\
&= \int_{S^1}\left\langle D_s m,v\right\rangle\left(\langle h,k\rangle + A\left\langle D_s^2 h,k\right\rangle\right) ds.
\end{aligned} \tag{3.13}$$

Since the differential operator D_s is anti self-adjoint for the L^2-metric g^0, i.e., $\int_{S^1}\langle D_s h,k\rangle\, ds = \int_{S^1}\langle h,-D_s k\rangle\, ds$,

$$\int_{S^1}\left\langle D_s^2 h,k\right\rangle ds = \int_{S^1}\left\langle h,D_s^2 k\right\rangle ds \tag{3.14}$$

holds. We proceed analogously to the proof of Theorem 2.1 in [84], which exploits the product rule for Riemannian connections. Thus, we conclude from

$$\begin{aligned}
&d\left(g_c^1(h,k)\right)[m] \\
&\overset{(3.13)}{=} \int_{S^1}\left\langle D_s m,v\right\rangle\left[\frac{1}{2}\left(\langle h,k\rangle + A\left\langle D_s^2 h,k\right\rangle\right) + \frac{1}{2}\left(\langle h,k\rangle + A\left\langle D_s^2 h,k\right\rangle\right)\right] ds \\
&\overset{(3.14)}{=} \int_{S^1}\left\langle \frac{1}{2}\left\langle D_s m,v\right\rangle\left(I + AD_s^2\right)h,k\right\rangle + \left\langle h,\frac{1}{2}\left\langle D_s m,v\right\rangle\left(I + AD_s^2\right)k\right\rangle ds \\
&= \int_{S^1}\left\langle L_1\left[L_1^{-1}\left(\frac{1}{2}\left\langle D_s m,v\right\rangle\left(I + AD_s^2\right)h\right)\right],k\right\rangle ds \\
&\quad + \int_{S^1}\left\langle h,L_1\left[L_1^{-1}\left(\frac{1}{2}\left\langle D_s m,v\right\rangle\left(I + AD_s^2\right)k\right)\right]\right\rangle ds \\
&= g_c^1\left(L_1^{-1}\left(\frac{1}{2}\left\langle D_s m,v\right\rangle\left(I + AD_s^2\right)h\right),k\right) \\
&\quad + g_c^1\left(h,L_1^{-1}\left(\frac{1}{2}\left\langle D_s m,v\right\rangle\left(I + AD_s^2\right)k\right)\right)
\end{aligned}$$

that the covariant derivative associated with g^1 is given by (3.8). □

Remark 3.4. As stated in [64], the inverse operator L_1^{-1} is an integral operator whose kernel has an expression in terms of the arc length distance between two points on a curve and their unit normal vectors. For the existence and more details about L_1^{-1} we refer to [64].

3.3 Shape space in higher dimensions

This thesis is principally concerned with two-dimensional shapes, but for the sake of completeness it should be mentioned that the shape space B_e and its theoretical results given in the previous section can be generalized to higher dimensions.

Let M be a compact manifold and let N denote a Riemannian manifold with $\dim(M) < \dim(N)$. In [62], the *space of all submanifolds* of type M in N is defined by

$$B_e(M, N) := \mathrm{Emb}(M, N)/\mathrm{Diff}(M), \tag{3.15}$$

i.e., by the set of all equivalence classes of embeddings from M into N, where the equivalence relation is given by the set of all diffeomorphisms from M into itself. In Figure 3.1, a three-dimensional shape which is an element of the shape space $B_e(S^2, \mathbb{R}^3)$ and a three-dimensional shape which is not an element of this shape space are illustrated. Moreover, the vanishing geodesic distance phenomenon occurs also for the L^2-metric in higher dimensions as verified in [62]. For the definition of the Sobolev metric g^1 in higher dimensions we refer to [11].

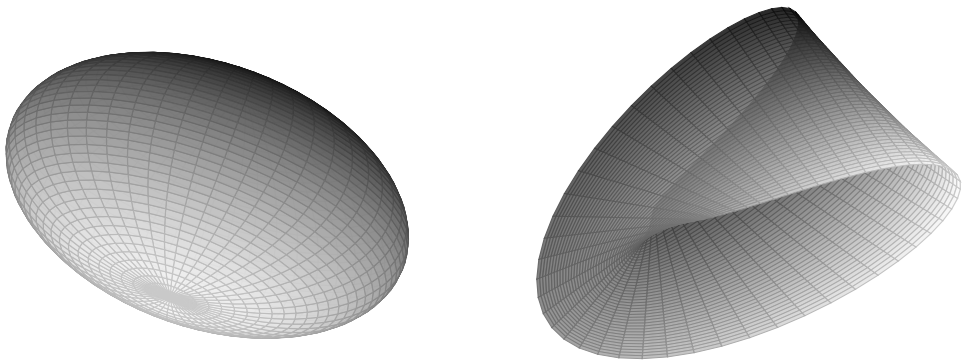


Figure 3.1: The left picture illustrates a shape in $B_e(S^2, \mathbb{R}^3)$. In contrast, the right picture shows no element of $B_e(S^2, \mathbb{R}^3)$.

Chapter 4

Shape derivative

Shape optimization is of interest in many fields of application, in particular, in the context of partial differential equations. Aerodynamic shape optimization [81], acoustic shape optimization [102] or optimization of interfaces in transmission problems [34, 68, 73] can be mentioned as examples. In industry, shapes are often represented within a finite dimensional design space. However, this reduction is often seen as being too restrictive (cf. [94]), which motivates shape optimization based on shape calculus. Major effort in shape calculus has been devoted towards expressions for shape derivatives in a so-called Hadamard-form, i.e., in a surface integral form (cf. [23, 93]). It is often a very tedious process to derive surface formulations of shape derivatives. Along the way, there appear volume formulations in the form of integrals over the entire domain as an intermediate step. Recently, it has been shown that this intermediate formulation has numerical advantages, see, for example, [34, 73]. Furthermore, the derivation as well as the implementation of volume integral formulations require less manual and programming work than the derivation and implementation of surface integral formulations. Hence, there arises the natural aim to use volume expressions of shape derivatives. This chapter aims at providing both, volume and surface formulations of shape derivatives, which are required in the next chapters. The usage of volume shape derivatives expressions in shape optimization is the topic of Chapter 6.

This chapter is organized as follows. We first set up notations and terminology in Section 4.1. Afterwards, we focus on problems of finding the interface of two subdomains (Section 4.2). We consider an elliptic shape interface optimization problem in Subsection 4.2.1 and a parabolic one in Subsection 4.2.2. Moreover, we provide their shape derivatives expressed as surface and volume integrals in each subsection. In Section 4.3, we embed shape optimization problems in the framework of optimization on Riemannian manifolds. Special attention is paid to the Riemannian shape gradient and the Riemannian shape Hessian. Finally, Section 4.4 provides a volume shape derivative formula for a special class of PDE constrained shape interface problems.

4.1 Basic concepts

This section provides basic definitions and terminology to prepare this chapter. Special attention is paid to material and shape derivatives which are needed throughout this thesis.

The main focus of shape optimization is to investigate shape functionals. First, we give the definition of a shape functional.

Definition 4.1 (Shape functional)**.** *Let D denote a non-empty subset of $\mathbb{R}^d$, where $d \in \mathbb{N}^*$. Moreover, $\mathcal{A} \subset \{\Omega \colon \Omega \subset D\}$ denotes a set of subsets. A function*

$$J \colon \mathcal{A} \to \mathbb{R},\ \Omega \mapsto J(\Omega)$$

is called a shape functional.

Let D be as in the above definition. Throughout this thesis any shape functional is denoted by J. We suppose that this functional is well defined for any measurable set Ω in D. Moreover, let $\{F_t\}_{t\in[0,T]}$ be a family of mappings $F_t \colon \overline{D} \to \mathbb{R}^d$ such that $F_0 = id$, where $T > 0$. This family transforms the domain Ω into new *perturbed domains*

$$\Omega_t \coloneqq F_t(\Omega) = \{F_t(x) \colon x \in \Omega\} \text{ with } \Omega_0 = \Omega$$

and the boundary Γ into new *perturbed boundaries*

$$\Gamma_t \coloneqq F_t(\Gamma) = \{F_t(x) \colon x \in \Gamma\} \text{ with } \Gamma_0 = \Gamma.$$

Considering the domain Ω as a collection of material particles which are changing their position in the time-interval $[0, T]$, the family $\{F_t\}_{t\in[0,T]}$ describes the motion of each particle. This means that at time $t \in [0, T]$ a material particle $x \in \Omega$ has the new position $x_t \coloneqq F_t(x) \in \Omega_t$ with $x_0 = x$. The motion of each such particle x can be described by the *velocity method* or by the *perturbation of identity*.

Definition 4.2 (Velocity method)**.** *For a sufficiently smooth vector field V the velocity method defines the family of the above-mentioned mappings as the flow $F_t(x) \coloneqq \xi(t, x)$ determined by the following initial value problem:*

$$\begin{aligned} \frac{d\xi(t,x)}{dt} &= V(\xi(t,x)) \\ \xi(0,x) &= x \end{aligned}$$

Definition 4.3 (Perturbation of identity)**.** *The perturbation of identity is defined by $F_t(x) \coloneqq x + tV(x)$, where V denotes a sufficiently smooth vector field.*

Remark 4.4. We use the perturbation of identity throughout this thesis.

This thesis deals with PDE constrained shape optimization problems, i.e., shape optimization problems constrained by equations involving an unknown function of two or more variables and at least one partial derivative of this function. Often, we have further conditions as outlined in Section 2.2. A shape optimization problem is defined as follows:

Definition 4.5 (Shape optimization problem)**.** *A shape optimization problem is given by*

$$\min_{\Omega} J(\Omega),$$

where J is a shape functional. When J depends on a solution of a PDE, we call the shape optimization problem PDE constrained.

To solve PDE constrained shape optimization problems, we need their shape derivatives.

Definition 4.6 (Shape derivative)**.** *Let $D \subset \mathbb{R}^d$ be open, where $d \geq 2$ is a natural number. Moreover, let $k \in \overline{\mathbb{N}}$ and let $\Omega \subset D$ be measurable. The Eulerian derivative of a shape functional J at Ω in direction $V \in \mathcal{C}_0^k(D, \mathbb{R}^d)$ is defined by*

$$DJ(\Omega)[V] \coloneqq \lim_{t \to 0^+} \frac{J(\Omega_t) - J(\Omega)}{t}. \tag{4.1}$$

If for all directions $V \in \mathcal{C}_0^k(D, \mathbb{R}^d)$ the Eulerian derivative (4.1) exists and the mapping

$$G(\Omega) \colon \mathcal{C}_0^k(D, \mathbb{R}^d) \to \mathbb{R},\ V \mapsto DJ(\Omega)[V]$$

is linear and continuous, the expression $DJ(\Omega)[V]$ is called the shape derivative of J at Ω in direction $V \in \mathcal{C}_0^k(D, \mathbb{R}^d)$. In this case, J is called shape differentiable of class $\mathcal{C}^k$ at Ω.

The following so-called Hadamard Structure Theorem is very important. Among other things, it states that only the normal part of a vector field on the boundary has an impact on the value of the shape derivative.

Theorem 4.7 (Hadamard Structure Theorem)**.** *Let D and Ω be as in Definition 4.6. Moreover, let the shape functional J be shape differentiable of class $\mathcal{C}^k$ at every domain $\Omega \subset D$ with $\mathcal{C}^{k-1}$-boundary $\Gamma = \partial\Omega$. Then there exists a scalar distribution $r \in \mathcal{C}_0^k(\Gamma)'$ such that $G(\Omega) \in \mathcal{C}_0^k(\Omega, \mathbb{R}^d)'$ of J at Ω is given by*

$$G(\Omega) = \gamma_\Gamma'(r \cdot n). \tag{4.2}$$

Here $\mathcal{C}_0^k(\Gamma)'$ and $\mathcal{C}_0^k(\Omega, \mathbb{R}^d)'$ denote the dual spaces of $\mathcal{C}_0^k(\Gamma)$ and $\mathcal{C}_0^k(\Omega, \mathbb{R}^d)$. Moreover,

$$\gamma_\Gamma \colon \mathcal{C}_0^k(\overline{D}, \mathbb{R}^d) \to \mathcal{C}_0^k(\Gamma, \mathbb{R}^d),\ v \to v\big|_\Gamma$$

denotes the trace operator and γ_Γ' its adjoint operator.

Proof. See Theorem 2.27 in [93]. □

Note that the Hadamard Structure Theorem 4.7 actually states the existence of a scalar distribution $r = r(\Omega)$ on the boundary Γ of a domain Ω. However, in this thesis, we always assume that r is an integrable function. In general, if $r \in L^1(\Gamma)$, then r is obtained in the form of the trace on Γ of an element $G \in W^{1,1}(\Omega)$. This

means that it follows from (4.2) that the shape derivative can be expressed more conveniently as

$$DJ(\Omega)[V] = \int_\Gamma r \, \langle V, n \rangle \, ds.$$

Thus, in many cases and particularly in this thesis, the shape derivative arises in two equivalent notational forms:

$$DJ_\Omega[V] \coloneqq \int_\Omega RV(x) \, dx \qquad \text{(volume formulation)} \tag{4.3}$$

$$DJ_\Gamma[V] \coloneqq \int_\Gamma r(s) \, \langle V(s), n(s) \rangle \, ds \qquad \text{(surface formulation)} \tag{4.4}$$

Here R is a (differential) operator acting linearly on the vector field V and $r \in L^1(\Gamma)$ with $DJ_\Omega[V] = DJ(\Omega)[V] = DJ_\Gamma[V]$. In the following, the volume formulation is also called the *domain formulation* and the surface formulation is also called the *boundary formulation*.

There are a lot of options to prove shape differentiability of shape functionals, which depend on a solution of a PDE, and to derive the shape derivative of a PDE constrained shape optimization problem. The min-max approach [23], the chain rule approach [93], the Lagrange method of Céa [18] and the rearrangement method [47] have to be mentioned in this context. A nice overview about these approaches is given in [98]. Note that the approach of Céa is frequently used to derive shape derivatives, but itself gives no proof of shape differentiability. Indeed, there are cases where the method of Céa fails (cf. [74, 96]). However, we use the min-max formulation of the Lagrangian corresponding to the PDE constrained shape optimization problem under consideration. We deduce its shape derivative by an application of the theorem of Correa and Seeger, which moreover guarantees its existence. Before we formulate this theorem, we focus on a rule for differentiating perturbed volume integrals, which is needed to derive shape derivatives in the sequel. We have to deal with so-called material derivatives whose definition is given in the following. For a material derivative free approach we refer to [96, 97].

Definition 4.8 (Material derivative)**.** *Let Ω, Ω_t, F_t and T be as above. Moreover, let $\{p_t \colon \Omega_t \to \mathbb{R} \colon t \leq T\}$ denote a family of mappings. The material derivative of a generic function $p(= p_0) \colon \Omega \to \mathbb{R}$ at $x \in \Omega$ is denoted by $D_m p$ or $\dot{p}$ and given by the derivative of the composed function $p_t \circ F_t \colon \Omega \to \Omega_t \to \mathbb{R}$ defined in the fixed domain Ω, i.e.,*

$$\dot{p}(x) \coloneqq \lim_{t \to 0^+} \frac{(p_t \circ F_t)(x) - p(x)}{t} = \frac{d^+}{dt} (p_t \circ F_t)(x) \Big|_{t=0}.$$

The classical chain rule for differentiation applied to $\dot{p}$ gives the relation between material and shape derivatives.

Definition 4.9 (Shape derivative)**.** *Let p be as in the previous definition. The shape derivative of p in the direction of a vector field V is denoted by p' and given by*

$$p' = \dot{p} - V^T \nabla p \quad \text{in } \Omega. \tag{4.5}$$

In the next sections, the following rules for the material derivative are needed to derive shape derivatives of objective functions depending on solutions of PDEs.

Theorem 4.10. *Let $p, q \colon \Omega \to \mathbb{R}$ be two functions and let D_m denote the material derivative with respect to the perturbation of identity $F_t = id + tV$.*

(i) For the material derivative the product rule holds, i.e.,

$$D_m(p\,q) = D_m p\, q + p\, D_m q. \tag{4.6}$$

(ii) While the shape derivative commutes with the gradient, the material derivative does not, but the following equality holds:

$$D_m \nabla p = \nabla D_m p - \nabla V^T \nabla p \tag{4.7}$$

(iii) The following identity holds:

$$D_m \left(\nabla q^T \nabla p \right) = \nabla D_m p^T \nabla q - \nabla q^T \left(\nabla V + \nabla V^T \right) \nabla p + \nabla p^T \nabla D_m q \tag{4.8}$$

Proof. See [16, Section 5]. □

We aim to deduce shape derivative formulas. For this purpose, we have to consider a perturbed objective function due to Definition 4.6. In the following theorem, the above-mentioned rule for differentiating perturbed domain integrals is given.

Theorem 4.11. *Let Ω, Ω_t, p and p_t be as in Definition 4.8. Then*

$$\frac{d^+}{dt} \left(\int_{\Omega_t} p_t \; dx_t \right) \Big|_{t=0} = \int_\Omega \dot{p} + \operatorname{div}(V) p \; dx \tag{4.9}$$

holds.

Proof. Using the theorem of substitution for integrals and applying the classical rule of differentiation of integrals with respect to parameters, we get

$$\begin{aligned} \frac{d^+}{dt} \left(\int_{\Omega_t} p_t \; dx_t \right) \Big|_{t=0} &= \frac{d^+}{dt} \left(\int_\Omega (p_t \circ F_t) \cdot \det(DF_t) \; dx \right) \Big|_{t=0} \\ &= \int_\Omega \frac{d^+}{dt} \left((p_t \circ F_t) \cdot \det(DF_t) \right) \Big|_{t=0} dx, \end{aligned}$$

where DF_t denotes the Jacobian of F_t. The derivative of the deformation determinant deduced in [80] is given by

$$\frac{d^+}{dt}\left(\det(DF_t)\right)\Big|_{t=0} = \operatorname{div}(V).$$

Applying the product rule for differentiation combined with this deformation determinant derivative gives

$$\frac{d^+}{dt}\left((p_t \circ F_t)\cdot \det(DF_t)\right)\Big|_{t=0} = \dot{p} + \operatorname{div}(V)p,$$

which completes the poof. □

Remark 4.12. The concept of material and shape derivatives of a scalar-valued function $p\colon \Omega \to \mathbb{R}$ can be extended to its boundary $\Gamma = \partial\Omega$. Since this goes beyond the scope of this thesis, we mention only a few aspects required in this thesis. For more details we refer to the literature, e.g., [70]. Let $z\colon \Gamma \to \mathbb{R}$ be the trace on the boundary Γ of p. In this setting, the boundary shape derivative z' is defined by

$$z' = \dot{p} - V^T \nabla_\Gamma p, \tag{4.10}$$

where ∇_Γ denotes the tangential gradient given by

$$\nabla_\Gamma p = \nabla p - \frac{\partial p}{\partial n} n.$$

Here $\frac{\partial}{\partial n}$ denotes the derivative normal to Γ. Combining (4.5) with (4.10) gives the correlation of boundary and domain shape derivatives:

$$z' = p' + V^T \frac{\partial p}{\partial n} n$$

In the next sections, we do not only have to deal with domain integrals, but also with boundary integrals. Therefore, we also need a rule for differentiating perturbed boundary integrals, which is similar to (4.9) and given by the following theorem.

Theorem 4.13. *Let Γ_t and T be as assumed above and let $\{z_t\colon \Gamma_t \to \mathbb{R}\colon t \le T\}$ be a family of mappings. Then*

$$\frac{d^+}{dt}\left(\int_{\Gamma_t} z_t \; ds_t\right)\Big|_{t=0} = \int_\Gamma \dot{z} + \operatorname{div}_\Gamma(V) z \; ds \tag{4.11}$$

holds, where $\operatorname{div}_\Gamma(V)$ denotes the tangential divergence of V defined by

$$\operatorname{div}_\Gamma(V) = \operatorname{div}(V) - n^T \frac{\partial V}{\partial n}.$$

Proof. Lemma 11.2 in [46] states

$$\int_{\Gamma_t} z_t \, ds_t = \int_{\Gamma} (z_t \circ F_t) \cdot \det(DF_t) \left| (DF_t)^{-T} n \right| ds$$

if $z_t \in L^1(\Gamma_t)$. From this, we get

$$\frac{d^+}{dt} \left(\int_{\Gamma_t} z_t(s_t) \, ds_t \right) \Big|_{t=0} = \int_{\Gamma} \frac{d^+}{dt} \left((z_t \circ F_t) \cdot \det(DF_t) \left| (DF_t)^{-T} n \right| \right) \Big|_{t=0} ds$$

by applying the classical rule of differentiation of integrals with respect to parameters. Applying the product rule combined with the derivative

$$\det(DF_t) \left| (DF_t)^{-T} n \right| \Big|_{t=0} = \operatorname{div}_\Gamma(V)$$

given in [46] yields

$$\frac{d^+}{dt} \left((z_t \circ F_t) \cdot \det(DF_t) \left| (DF_t)^{-T} n \right| \right) \Big|_{t=0} = \dot{z} + \operatorname{div}_\Gamma(V) z,$$

which completes the proof. □

Besides (4.11), which includes the material derivative, the next theorem gives a rule for differentiating perturbed boundary integrals, which contains the shape derivative instead. For the proof, which is quite technical, we refer to [51].

Theorem 4.14. *If Γ_t and z_t be as above. Then*

$$\frac{d^+}{dt} \left(\int_{\Gamma_t} z_t \, ds_t \right) \Big|_{t=0} = \int_{\Gamma} z' + \left(\frac{\partial z}{\partial n} + z\kappa \right) \langle V, n \rangle \, ds$$

holds, where κ denotes the mean curvature of Γ.

In this thesis, the objective functions include so-called perimeter regularization terms denoted by j_{reg}. More precisely,

$$J(\Omega) = j(\Omega) + j_{\text{reg}}(\Omega), \text{ where } j_{\text{reg}}(\Omega) \coloneqq \mu \int_{\Gamma} 1 \, ds \text{ with } \mu > 0. \tag{4.12}$$

The following theorem provides the first and second order shape derivative of the regularization term $j_{\text{reg}}(\Omega)$ in the two-dimensional case.

Theorem 4.15. *Let Γ be the boundary of a bounded domain $\Omega \subset \mathbb{R}^2$. Moreover, V and W are two sufficiently smooth vector fields. For*

$$j_{\text{reg}}(\Omega) \coloneqq \int_{\Gamma} 1 \, ds$$

we have

$$Dj_{\mathrm{reg}}(\Omega)[V] = \int_\Gamma \kappa \langle V, n\rangle \, ds, \tag{4.13}$$

$$D^2 j_{\mathrm{reg}}(\Omega)[V,W] = \int_\Gamma \frac{\partial V}{\partial \tau}\frac{\partial W}{\partial \tau} \langle V, n\rangle \langle W, n\rangle \, ds, \tag{4.14}$$

where κ denotes the mean curvature of Γ and $\frac{\partial}{\partial \tau}$ is the derivative tangential to Γ.

Proof. See Proposition 5.1 and Remark 5.3 in [71]. □

As already mentioned, we deduce shape derivatives of PDE constrained shape optimization problems by an application of the theorem of Correa and Seeger. This theorem (in its original version) is given in the sequel.

Remark 4.16. In its original version, the theorem of Correa and Seeger is restricted to locally convex vector spaces. In this thesis, shape spaces are not locally convex spaces. More precisely, they can be identified as infinite dimensional Riemannian manifolds (cf. Chapter 3). Thus, a reformulation of Theorem 4.19 is required to make it applicable in order to deduce shape derivative formulas. In [23, Chapter 10], such a reformulation is provided.

To formulate the theorem of Correa and Seeger, a few notations and definitions have to be introduced:
Let $U_0 \neq \emptyset$ and $V_0 \neq \emptyset$ be subsets of Hausdorff spaces U and V and let X be a Hausdorff locally convex real topological vector space. Moreover, let L be an extended real valued function on $X \times U \times V$. For $x \in X$ we define

$$V(x) \coloneqq \left\{ v \in V_0 \colon \sup_{v\in V_0} \inf_{u\in U_0} L(x,u,v) = \inf_{u\in U_0} L(x,u,v) \right\},$$

$$U(x) \coloneqq \left\{ u \in U_0 \colon \inf_{u\in U_0} \sup_{v\in V_0} L(x,u,v) = \sup_{v\in V_0} L(x,u,v) \right\}.$$

Definition 4.17 (Dini directional derivatives)**.** *With the above notation and a direction $W \in X$, the Dini directional derivatives of $L(\cdot, u, v)$ are defined by*

$$\overline{D}L(x,u,v)[W] = \varlimsup_{t\to 0^+} \frac{L(x_t,u,v) - L(x,u,v)}{t},$$

$$\underline{D}L(x,u,v)[W] = \varliminf_{t\to 0^+} \frac{L(x_t,u,v) - L(x,u,v)}{t},$$

where $x_t = x_t(x) = x + tW(x)$ is the perturbation of $x \in X$ determined by the perturbation of identity.

Definition 4.18 (u.s.c., l.s.c., s.s.c.)**.** *Let T and Z be Hausdorff spaces and let $A\colon T \to Z$ be a multifunction.*

(i) The multifunction A is upper semi-continuous (u.s.c.) at $t_0 \in T$, if whenever $\mathcal{O} \subset Z$ is an open subset which contains $A(t_0)$, then the set $\{t\colon A(t) \subset \mathcal{O}\}$ contains a neighbourhood of t_0.

(ii) The multifunction A is lower semi-continuous (l.s.c.) at $t_0 \in T$, if whenever an open subset $\mathcal{O} \subset Z$ satisfies $\mathcal{O} \cap A(t_0) \neq \emptyset$, then $\{t\colon A(t) \cap \mathcal{O} \neq \emptyset\}$ contains a neighbourhood of t_0.

(iii) The multifunction A is sequentially semi-continuous (s.s.c.) at $t_0 \in T$, if for every sequence $\{t_k\}$ converging to t_0, there exist $z_0 \in A(t_0)$ and a sequence $\{z_k\}$ accumulating at z_0 such that $z_k \in A(t_k)$ for all k sufficiently large.

With these notations and definitions in mind, we can state the original version of the theorem of Correa and Seeger:

Theorem 4.19 (Correa and Seeger)**.** *Assume that the multifunctions $t \mapsto U(x_t)$, $t \mapsto V(x_t)$ are s.s.c. at 0, where $x_t = x_t(x) = x + tW(x)$ is the perturbation of $x \in X$ determined by the perturbation of identity and $t \in \mathbb{R}^+$. Moreover, let the following properties hold:*

(P1) There exists a $\delta > 0$ such that the function $t \mapsto L(x_t, u, v)$ is finite and continuous on $[0, \delta[$ for every $(u, v) \in U_0 \times V_0$.

(P2) The function $\mathbb{R}^+ \times V_0 \ni (t, v) \mapsto \underline{D}L(x_t, u_0, v)[W]$ is u.s.c. at $\{0\} \times V(x)$ and finite for all $u_0 \in U(x)$. The function $\mathbb{R}^+ \times U_0 \ni (t, u) \mapsto \overline{D}L(x_t, u, v_0)[W]$ is l.s.c. at $\{0\} \times U(x)$ and finite for all $v_0 \in V(x)$.

(P3) There exists a $\delta > 0$ such that

$$\sup_{v \in V_0} \inf_{u \in U_0} L(x_t, u, v) = \inf_{u \in U_0} \sup_{v \in V_0} L(x_t, u, v) \quad \forall t \in [0, \delta[.$$

Set $h(x_t) \coloneqq \sup_{v \in V_0} \inf_{u \in U_0} L(x_t, u, v) = \inf_{u \in U_0} \sup_{v \in V_0} L(x_t, u, v)$.

Then the directional derivative $Dh(x)[W]$ exists and is characterized by

$$\sup_{v \in V_0} \inf_{u \in U_0} DL(x, u, v)[W] = Dh(x)[W] = \inf_{u \in U_0} \sup_{v \in V_0} DL(x, u, v)[W].$$

Proof. See Theorem 2.1 in [21]. □

Remark 4.20. We use a min-max formulation of the Lagrangian corresponding to a PDE constrained shape optimization problem. This means that we express the objective function of a PDE constrained shape optimization problem as a min-max of the Lagrangian corresponding to this optimization problem. In this way, the problem of shape differentiability of an objective function is reduced to the differentiability of the min-max function. Theorem 4.19 ensures the differentiability of a min-max function.

4.2 PDE constrained shape interface problems

In this section, we consider problems of finding interfaces between two subdomains. These problems are used to illustrate the connection of shape optimization to Riemannian geometry in Chapter 5. We consider an elliptic shape interface optimization problem in Subsection 4.2.1 and a parabolic one in Subsection 4.2.2. Their shape derivatives, expressed as domain and boundary integrals, are deduced in preparation for the theoretical discussions in the next chapters. More precisely, we need boundary expressions of their shape derivatives in Chapter 5 and we require shape derivatives expressed as domain integrals in Chapter 6.

4.2.1 Elliptic problem

In this subsection, we consider a PDE constrained shape optimization problem which is inspired by the standard tracking-type elliptic optimal control problem and motivated by electrical impedance tomography. It is very close to the model problem in [18, Example 2] and to the inverse interface problem in [46]. Preparing the theoretical discussions in Section 5.1, we deduce its shape derivative expressed as domain and boundary integrals.

Problem formulation

Let the domain $\Omega \coloneqq (0,1)^2 \subset \mathbb{R}^2$ be partitioned into two disjoint subdomains $\Omega_1, \Omega_2 \subset \Omega$ such that $\Omega_1 \cup \Gamma_{\text{int}} \cup \Omega_2 = \Omega$ and $\partial\Omega_1 \cap \partial\Omega_2 = \Gamma_{\text{int}}$, where $\cup$ denotes the disjoint union. In contrast to the outer boundary $\Gamma_{\text{out}} \coloneqq \partial\Omega$, which is assumed to be fixed, the inner boundary Γ_{int}, which is also called the interface, is variable. Let the interface Γ_{int} be an element of

$$B_e^0 = B_e^0\left([0,1],\mathbb{R}^2\right) \coloneqq \text{Emb}^0([0,1],\mathbb{R}^2)/\text{Diff}^0([0,1]), \tag{4.15}$$

where

$$\begin{aligned}\text{Emb}^0([0,1],\mathbb{R}^2) \coloneqq \{\phi \in C^\infty([0,1],\mathbb{R}^2)\colon\ &\phi(0) = (0.5,0),\ \phi(1) = (0.5,1),\\ &\phi \text{ proper injective immersion}\},\end{aligned}$$

$$\text{Diff}^0([0,1]) \coloneqq \{\phi\colon [0,1] \to [0,1]\colon \phi(0) = 0,\ \phi(1) = 1,\ \phi \text{ diffeomorphism}\}.$$

The shape space $B_e^0\left([0,1],\mathbb{R}^2\right)$ is constructed in analogy to the manifold $B_e(S^1,\mathbb{R}^2)$ introduced in Chapter 3. Consequently, a particular point of $B_e^0\left([0,1],\mathbb{R}^2\right)$ is represented by a curve $c\colon [0,1] \to \mathbb{R}^2$, $\theta \mapsto c(\theta)$. Because of the equivalence relation $\text{Diff}^0([0,1])$, the tangent space is isomorphic to the set of all smooth vector fields along c, i.e.,

$$T_c B_e^0\left([0,1],\mathbb{R}^2\right) \cong \{h\colon h = \alpha n, \alpha \in \mathcal{C}^\infty([0,1])\},$$

where n is the unit outward normal to Ω_1 at the interface Γ_{int}. Now, we replace Γ_{int} by u and consider Ω depending on u. Thus, we denote it by $\Omega(u) = \Omega_1(u) \cup u \cup \Omega_2(u)$. In the left picture of Figure 4.1, the construction of $\Omega(u)$ is illustrated.

We define the following PDE constrained shape optimization problem:

$$\min_u \; J(u) = j(u) + j_{\text{reg}}(u) \coloneqq \frac{1}{2}\int_{\Omega(u)} (y - \overline{y})^2 \; dx + \mu \int_u 1 \; ds \tag{4.16}$$

$$\text{s.t.} \quad -\triangle y = f \quad \text{in } \Omega(u) \tag{4.17}$$

$$y = 0 \quad \text{on } \partial\Omega(u) \tag{4.18}$$

The right-hand side of the Poisson equation (4.17) is given by a jumping coefficient

$$f \coloneqq \begin{cases} f_1 = \text{const.} & \text{in } \Omega_1(u) \\ f_2 = \text{const.} & \text{in } \Omega_2(u) \end{cases} . \tag{4.19}$$

Of course, formulation (4.17) of the PDE has to be understood only formally because of the jumping coefficient f. The second term $j_{\text{reg}}(u)$ in the objective function J defined in (4.16) is a perimeter regularization with $\mu > 0$. It is frequently used to overcome ill-posedness of optimization problems (cf. [6]). As already mentioned, n is the unit outward normal to $\Omega_1(u)$ at u. We observe that the unit outward normal to $\Omega_2(u)$ at u is equal to $-n$, which enables us to use only one normal n for the subsequent discussions. Furthermore, we have conditions on the interface u. We formulate explicitly the continuity of the state and the flux at the interface u as

$$[\![y]\!] = 0, \qquad \left[\!\!\left[\frac{\partial y}{\partial n}\right]\!\!\right] = 0 \quad \text{on } u. \tag{4.20}$$

The jump symbol $[\![\cdot]\!]$ denotes the discontinuity across the interface and is defined for $v \in \Omega$ by

$$[\![v]\!] \coloneqq v|_{\Omega_1} - v|_{\Omega_2}. \tag{4.21}$$

For the observation $\overline{y}$ we assume $\overline{y} \in L^2(\Omega)$.

It is well known that the boundary value problem (4.17)-(4.18) admits a weak solution $y \in H_0^1(\Omega(u))$ (cf. for example [31, Chapter 6]). Let us consider such weak solutions in the following. Note that these solutions have higher regularity. To be more precise, a smooth boundary combined with $f \in L^2(\Omega)$ and $\overline{y} \in L^2(\Omega)$ put the solution $y \in H_0^1(\Omega)$ in $H^2(\Omega)$ due to Theorem 2.13. Thus, it is guaranteed that $\frac{\partial y|_{\Omega_1(u)}}{\partial n}, \frac{\partial y|_{\Omega_2(u)}}{\partial n} \in H^{1/2}(u)$ by the trace theorem for Sobolev spaces.

In the setting above, the boundary value problem (4.17)-(4.18) becomes, in its weak form,

$$a_u(y, p) = b_u(p) \quad \forall p \in H_0^1(\Omega(u)), \tag{4.22}$$

where

$$a_u(y, p) = \int_{\Omega(u)} \nabla y^T \nabla p \; dx - \int_u \left[\!\!\left[\frac{\partial y}{\partial n} p\right]\!\!\right] ds, \tag{4.23}$$

$$b_u(p) = \int_{\Omega(u)} fp \; dx. \tag{4.24}$$

Let

$$F \coloneqq \left\{(y,u,p)\colon y,p \in H_0^1(\Omega(u)), u \in B_e^0([0,1],\mathbb{R}^2)\right\}. \tag{4.25}$$

The Lagrangian corresponding to (4.16)-(4.20) is defined by

$$\mathscr{L}\colon F \to \mathbb{R},\ (y,u,p) \mapsto J(y,u) + a_u(y,p) - b_u(p), \tag{4.26}$$

where a_u, b_u are defined in (4.23)-(4.24) and J is given in (4.16).

Remark 4.21. The integral over Ω is understood as the sum of the integrals over Ω_1 and Ω_2 in this thesis.

Shape derivative formulas

We first consider the objective function J defined in (4.16) without perimeter regularization. The shape derivative can be expressed as an integral over the domain $\Omega(u)$ and an integral over the interface u. We first deduce the domain integral by an application of Theorem 4.19. Afterwards, we convert it into an interface integral by applying integration by parts on u. Due to the theorem of Correa and Seeger, the shape derivative of $\mathscr{L}$ evaluated in its saddle point is equal to the one of J. A saddle point $(y,u,p) \in F$ of the Lagrangian $\mathscr{L}$ is given by

$$\frac{\partial \mathscr{L}(y,u,p)}{\partial y} = \frac{\partial \mathscr{L}(y,u,p)}{\partial p} = 0, \tag{4.27}$$

which leads to the adjoint equation given in strong form by

$$-\triangle p = -(y - \overline{y}) \quad \text{in } \Omega(u) \tag{4.28}$$

$$p = 0 \quad \text{on } \partial\Omega(u) \tag{4.29}$$

and to the state equation given in strong form by

$$-\triangle y = f \quad \text{in } \Omega(u). \tag{4.30}$$

We formulate explicitly the interface conditions of (4.28)-(4.29) by

$$[\![p]\!] = 0, \quad \left[\!\!\left[\frac{\partial p}{\partial n}\right]\!\!\right] = 0 \qquad \text{on } u. \tag{4.31}$$

Remark 4.22. Note that only the normal part of a vector field V on the variable boundary – in our case, on the interface u – has an impact on the value of the shape derivative expressed as boundary integral due the Hadamard Structure Theorem 4.7.

Now, we formulate a theorem which provides a representation of the shape derivative expressed as domain integral. Later on, this domain integral allows us to calculate the boundary expression of the shape derivative.

Theorem 4.23. *Let $y \in H_0^1(\Omega(u))$ be the weak solution of (4.17)-(4.18). Moreover, let $p \in H_0^1(\Omega(u))$ denote the weak solution of the adjoint equation (4.28)-(4.29). Then the shape derivative of the objective function J without perimeter regularization, i.e., the shape derivative of j, in direction V is given by*

$$\begin{aligned} Dj_{\Omega(u)}(u)[V] = \int_{\Omega(u)} & -\nabla y^T \left(\nabla V + \nabla V^T\right) \nabla p - pV^T \nabla f \\ & + \operatorname{div}(V) \left(\frac{1}{2}(y-\overline{y})^2 + \nabla y^T \nabla p - fp\right) dx. \end{aligned} \tag{4.32}$$

Proof. Let $(y,u,p) \in H_0^1(\Omega(u)) \times B_e^0([0,1],\mathbb{R}^2) \times H_0^1(\Omega(u))$. Moreover, let $\widetilde{\mathscr{L}}$ be defined by

$$\widetilde{\mathscr{L}}(y,u,p) = j(u) + a_u(y,p) - b_u(p), \tag{4.33}$$

where a_u, b_u are defined in (4.23)-(4.24) and j is given in (4.16). In analogy to [23, Chapter 10, Subsection 5.2], we can verify that

$$j(u) = \min_{y \in H_0^1(\Omega(u))} \max_{p \in H_0^1(\Omega(u))} \widetilde{\mathscr{L}}(y,u,p) \tag{4.34}$$

holds. We apply the theorem of Correa and Seeger on the right-hand side of (4.34). The verification of the assumptions of this theorem can be checked in the same way as in [23, Chapter 10, Subsection 6.4]. We obtain formula (4.32) by evaluation of the shape derivative of the Lagrangian (4.33) in its saddle point. Applying the rule for differentiating domain and boundary integrals given in (4.9) and (4.11) yields

$$\begin{aligned} D\widetilde{\mathscr{L}}(y,u,p)[V] = \int_{\Omega(u)} & D_m\left(\frac{1}{2}(y-\overline{y})^2\right) + D_m\left(\nabla y^T \nabla p\right) - D_m(fp) \\ & + \operatorname{div}(V)\left(\frac{1}{2}(y-\overline{y})^2 + \nabla y^T \nabla p - fp\right) dx \\ - \int_u & D_m\left(\left[\!\left[\frac{\partial y}{\partial n} p\right]\!\right]\right) + \operatorname{div}_u(V)\left[\!\left[\frac{\partial y}{\partial n} p\right]\!\right] ds. \end{aligned} \tag{4.35}$$

By combining (4.35) with the product rule and (4.8) we obtain

$$\begin{aligned} D\widetilde{\mathscr{L}}(y,u,p)[V] = \int_{\Omega(u)} & (y-\overline{y})\dot{y} + \nabla \dot{y}^T \nabla p + \nabla y^T \nabla \dot{p} \\ & - \nabla y^T \left(\nabla V + \nabla V^T\right) \nabla p - \dot{f} p - f\dot{p} \\ & + \operatorname{div}(V)\left(\frac{1}{2}(y-\overline{y})^2 + \nabla y^T \nabla p - fp\right) dx \\ - \int_u & \left[\!\left[D_m\left(\frac{\partial y}{\partial n}\right) p + \frac{\partial y}{\partial n}\dot{p}\right]\!\right] + \operatorname{div}_u(V)\left[\!\left[\frac{\partial y}{\partial n} p\right]\!\right] ds. \end{aligned}$$

Now, we consider the saddle point condition (4.27) or, equivalently, the weak form of the adjoint equation (4.28)-(4.29) and the design equation (4.30). From this we

get

$$D\widetilde{\mathscr{L}}(y,u,p)[V] = \int_{\Omega(u)} -\nabla y^T \left(\nabla V + \nabla V^T\right) \nabla p - \dot{f} p + \operatorname{div}(V) \left(\frac{1}{2}(y-\overline{y})^2 + \nabla y^T \nabla p - fp\right) dx + \int_u \left[\!\!\left[\frac{\partial p}{\partial n} \dot{y} - D_m\left(\frac{\partial y}{\partial n}\right) p \right]\!\!\right] + \operatorname{div}_u(V) \left[\!\!\left[\frac{\partial y}{\partial n} p \right]\!\!\right] ds, \tag{4.36}$$

where the term $\dot{f}p$ in the domain integral is equal to $pV^T\nabla f$ in the both subdomains $\Omega_1(u)$ and $\Omega_2(u)$ because of (4.5). Due to (4.20) and (4.31), we get

$$\left[\!\!\left[\frac{\partial p}{\partial n} \dot{y} \right]\!\!\right] = \dot{y} \left[\!\!\left[\frac{\partial p}{\partial n} \right]\!\!\right] = 0 \quad \text{on } u, \tag{4.37}$$

$$\left[\!\!\left[D_m\left(\frac{\partial y}{\partial n}\right) p \right]\!\!\right] = D_m\left(\frac{\partial y}{\partial n}\right) [\![p]\!] = 0 \quad \text{on } u. \tag{4.38}$$

The identity

$$\left[\!\!\left[\frac{\partial y}{\partial n} p \right]\!\!\right] = 0 \quad \text{on } u \tag{4.39}$$

follows from (4.20), (4.31) and the identity

$$[\![ab]\!] = [\![a]\!]\, b_1 + a_2\, [\![b]\!] = a_1\, [\![b]\!] + [\![a]\!]\, b_2, \tag{4.40}$$

which implies

$$[\![ab]\!] = 0 \text{ if } [\![a]\!] = 0 \wedge [\![b]\!] = 0. \tag{4.41}$$

By combining (4.36)-(4.39) and the theorem of Correa and Seeger, we obtain (4.32). □

The domain integral (4.32) can be converted into a boundary integral. The following theorem provides a representation of the shape derivative expressed as boundary integral.

Theorem 4.24. *Assume that the elliptic PDE problem (4.17)-(4.18) is H^2-regular, i.e., a solution y exists and is at least in $H^2(\Omega(u))$. Moreover, assume that the adjoint equation (4.28)-(4.29) admits a solution $p \in H^2(\Omega(u))$. Then the shape derivative of the objective function J without perimeter regularization, i.e., the shape derivative of j, in direction V is given by*

$$Dj_u(u)[V] = -\int_u [\![f]\!]\, p\, \langle V, n\rangle\, ds. \tag{4.42}$$

Proof. Integration by parts in (4.32) yields

$$\begin{aligned}
&\int_{\Omega(u)} \operatorname{div}(V) \left(\frac{1}{2}(y-\overline{y})^2 + \nabla y^T \nabla p - fp \right) dx \\
&= -\int_{\Omega(u)} V^T \left((y-\overline{y}) \nabla y + \nabla \left(\nabla y^T \nabla p \right) - \nabla (fp) \right) dx \\
&\quad + \int_u \left[\!\!\left[\left(\frac{1}{2}(y-\overline{y})^2 + \nabla y^T \nabla p - fp \right) \langle V, n \rangle \right]\!\!\right] ds \\
&\quad + \int_{\partial\Omega(u)} \left(\frac{1}{2}(y-\overline{y})^2 + \nabla y^T \nabla p - fp \right) \langle V, n \rangle \, ds.
\end{aligned} \tag{4.43}$$

Since the outer boundary $\partial\Omega(u)$ is fixed, we can choose the deformation vector field V equals zero in small neighbourhoods of $\partial\Omega(u)$. Therefore, in (4.43), the outer integral disappears. Combining (4.32), (4.43) and the vector calculus identity

$$\nabla y^T \left(\nabla V + \nabla V^T \right) \nabla p + V^T \nabla \left(\nabla y^T \nabla p \right) = \nabla p^T \nabla \left(V^T \nabla y \right) + \nabla y^T \nabla \left(V^T \nabla p \right),$$

which is proved in [16], gives

$$\begin{aligned}
Dj_u(u)[V] = \int_{\Omega(u)} & -\nabla p^T \nabla \left(V^T \nabla y \right) - \nabla y^T \nabla \left(V^T \nabla p \right) \\
& - (y-\overline{y}) V^T \nabla y + f V^T \nabla p \; dx \\
& + \int_u \left[\!\!\left[\left(\frac{1}{2}(y-\overline{y})^2 + \nabla y^T \nabla p - fp \right) \langle V, n \rangle \right]\!\!\right] ds.
\end{aligned} \tag{4.44}$$

By applying integration by parts in (4.44) we get

$$\begin{aligned}
&\int_{\Omega(u)} \nabla y^T \nabla \left(V^T \nabla p \right) dx \\
&= -\int_{\Omega(u)} \triangle y V^T \nabla p \; dx + \int_u \left[\!\!\left[\frac{\partial y}{\partial n} V^T \nabla p \right]\!\!\right] ds + \int_{\partial\Omega(u)} \frac{\partial y}{\partial n} V^T \nabla p \; ds
\end{aligned} \tag{4.45}$$

and analogously

$$\begin{aligned}
&\int_{\Omega(u)} \nabla p^T \nabla \left(V^T \nabla y \right) dx \\
&= -\int_{\Omega(u)} \triangle p V^T \nabla y \; dx + \int_u \left[\!\!\left[\frac{\partial p}{\partial n} V^T \nabla y \right]\!\!\right] ds + \int_{\partial\Omega(u)} \frac{\partial p}{\partial n} V^T \nabla y \; ds.
\end{aligned} \tag{4.46}$$

In (4.45)-(4.46), the outer integrals vanish due to the fixed outer boundary $\partial\Omega(u)$ as in (4.43). Thus, it follows that

$$\begin{aligned}
Dj_u(u)[V] = \int_{\Omega(u)} & V^T \nabla p \left(\triangle y + f \right) + V^T \nabla y \left(\triangle p - (y-\overline{y}) \right) dx \\
& + \int_u \left[\!\!\left[\left(\frac{1}{2}(y-\overline{y})^2 + \nabla y^T \nabla p - fp \right) \langle V, n \rangle \right]\!\!\right] \\
& - \left[\!\!\left[\frac{\partial y}{\partial n} V^T \nabla p \right]\!\!\right] - \left[\!\!\left[\frac{\partial p}{\partial n} V^T \nabla y \right]\!\!\right] ds
\end{aligned} \tag{4.47}$$

holds. In (4.47), the domain integral vanishes due to (4.28) and (4.30). Moreover, the term $\left[\!\left[\frac{1}{2}(y-\overline{y})^2 \langle V, n\rangle\right]\!\right]$ disappears because of (4.20) and the term $\left[\!\left[\nabla y^T \nabla p \langle V, n\rangle\right]\!\right]$ vanishes because of the continuity of ∇y and ∇p. The identity

$$\left[\!\left[\frac{\partial y}{\partial n} V^T \nabla p\right]\!\right] = \left[\!\left[\frac{\partial p}{\partial n} V^T \nabla y\right]\!\right] = \langle V, n\rangle \left[\!\left[\frac{\partial y}{\partial n}\frac{\partial p}{\partial n}\right]\!\right] = 0 \tag{4.48}$$

follows from (4.20), (4.31) and (4.41). If we summarize all this, we obtain (4.42). □

Now, we consider the objective function (4.16) with perimeter regularization. The following theorem gives the boundary integral expression of its shape derivative.

Theorem 4.25. *Under the assumptions of Theorem 4.24 the shape derivative of the objective function J in direction V is given by*

$$DJ_u(u)[V] = \int_u (-[\![f]\!]\, p + \mu\kappa)\, \langle V, n\rangle \, ds, \tag{4.49}$$

where κ denotes the curvature corresponding to the normal n.

Proof. Combining (4.42) with formula (4.13) in Theorem 4.15 we get (4.49). □

Remark 4.26. In Theorem 4.23, the volume shape derivative expression (4.32) is obtained under the assumptions that both, the solution y of the elliptic problem (4.17)-(4.18) and the solution p of the adjoint equation (4.28)-(4.29), belong to $H_0^1(\Omega)$. In contrast to this, we need a higher regularity to provide the surface shape derivative expressions (4.42) and (4.49). In Theorem 4.24 and Theorem 4.25, the shape derivative expressions are obtained under the assumption that $y, p \in H^2(\Omega)$. The crucial point is that for the elliptic problem (4.17)-(4.18), a smooth boundary combined with $f \in L^2(\Omega)$ and $\overline{y} \in L^2(\Omega)$ put the solutions $y, p \in H_0^1(\Omega)$ in $H^2(\Omega)$ due to Theorem 2.13.

4.2.2 Parabolic problem

In this subsection, we introduce a model problem which is given by a parabolic shape interface optimization problem. Among other things, it is motivated by the question *"How is the absorption of a medical cream into the human skin?"*. This is a topic in [68]. Preparing the theoretical discussions in Section 5.2 and in Chapter 6, its shape derivative, expressed as domain and boundary integral, is deduced by an application of Theorem 4.19 and a generalization of the approach in [73] for parabolic problems.

Problem formulation

As in the previous subsection, we denote by u the interface between two subdomains. The domain $X := (0,1)^2 \subset \mathbb{R}^2$ with fixed Lipschitz boundary $\Gamma_{\text{out}} = \partial X$ is partitioned into two disjoint subdomains $X_1, X_2 \subset X$ such that $u = \partial X_1 \cap \partial X_2$,

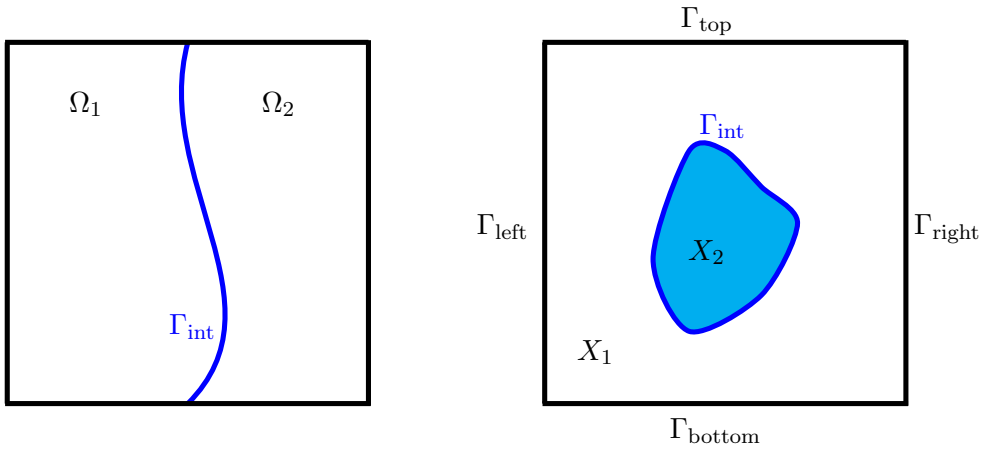


Figure 4.1: Example of a domain $\Omega = \Omega_1 \cup \Gamma_{\text{int}} \cup \Omega_2$ and $X = X_1 \cup \Gamma_{\text{int}} \cup X_2$. Both domains depend on the interface Γ_{int}.

$u = \partial X_2 (= \Gamma_{\text{int}})$, $X_1 \cup u \cup X_2 = X$ and $\Gamma_{\text{bottom}} \cup \Gamma_{\text{left}} \cup \Gamma_{\text{right}} \cup \Gamma_{\text{top}} = \partial X \ (= \Gamma_{\text{out}})$. An example of such a domain is illustrated in the right picture of Figure 4.1. The interface u is assumed to be smooth, variable and an element of the shape space B_e introduced in Chapter 3. We consider X depending on u and denote it by $X(u) = X_1(u) \cup u \cup X_2(u)$.

We introduce the following parabolic PDE constrained shape optimization problem in strong form:

$$\min_u \ J(u) = j(u) + j_{\text{reg}}(u) \coloneqq \int_0^T \int_{X(u)} (y - \overline{y})^2 \, dx \, dt + \mu \int_u 1 \, ds \tag{4.50}$$

$$\text{s.t.} \ \frac{\partial y}{\partial t} - \operatorname{div}(k \nabla y) = f \quad \text{in } X(u) \times (0, T] \tag{4.51}$$

$$y = 1 \quad \text{on } \Gamma_{\text{top}} \times (0, T] \tag{4.52}$$

$$\frac{\partial y}{\partial n} = 0 \quad \text{on } (\Gamma_{\text{bottom}} \cup \Gamma_{\text{left}} \cup \Gamma_{\text{right}}) \times (0, T] \tag{4.53}$$

$$y = y_0 \quad \text{in } X(u) \times \{0\} \tag{4.54}$$

with

$$k \coloneqq \begin{cases} k_1 = \text{const.} & \text{in } X_1(u) \times (0, T] \\ k_2 = \text{const.} & \text{in } X_2(u) \times (0, T] \end{cases} \tag{4.55}$$

denoting a jumping coefficient, n being the unit outward normal vector to $X_2(u)$ and $y_0 \in H^1(X(u))$. Of course, formulation (4.51) of the PDE has to be understood only formally because of the jumping coefficient k. The unit outward normal to $X_1(u)$ is equal to $-n$, which enables us to use only one normal n for the subsequent discussions. Furthermore, we have conditions on the interface u. We formulate

explicitly the continuity of the state and the flux at the interface as

$$[\![y]\!] = 0\,, \quad \left[\!\!\left[k\frac{\partial y}{\partial n}\right]\!\!\right] = 0 \quad \text{on } u \times (0,T], \tag{4.56}$$

where the jump symbol $[\![\cdot]\!]$ is defined as in (4.21). As in the previous section, the second term $j_{\text{reg}}(u)$ in the objective function in (4.50) is a perimeter regularization with $\mu > 0$. It is frequently used in this kind of problems. However, in [96, 97], a weaker, but more complicated regularization is used in order to show existence of solutions. As in [100], we assume $f \in L^2(0,T;L^2(X(u)))$. Moreover, we assume for the observation $\overline{y} \in L^2(0,T;L^2(X(u)))$.

In the following, we consider a weak solution $y \in L^2(0,T;H^1(X(u)))$ of problem (4.51)-(4.54). Note that it is known that such a solution exists (cf. for example [31, Chapter 7]). Recall that there are two possibilities for y which lead to a source of asymmetry in the treatment of the state variable y and the adjoint variable p (cf. Subsection 2.2.1). This means, if we consider $y \in L^2(0,T;H^1(X(u)))$, then we have to choose $p \in W(0,T;H^1(X(u)))$. Note that the values $y(0,x)$ and $y(T,x)$, which arise by applying integration by parts in time (cf. Remark 4.27 and 4.28), are not necessarily defined because functions $y \in L^2(0,T;H^1(X(u)))$ need not be continuous in time. The given value y_0 can be inserted for $y(0,x)$, but the final value $y(T,x)$ cannot be eliminated so easily. The adjoint variable $p \in W(0,T;H^1(X(u)))$ has higher regularity. In particular, $p(0,\cdot)$ and $p(T,\cdot)$ are well defined as traces in $L^2(X(u))$ for all $p \in W(0,T;H^1(X(u)))$ (cf. [56]). Therefore, it makes sense to require $p(T,x) = 0$ in order to get rid of the term $y(T,x)$. In the following, we assume $p = 0$ in $X(u) \times \{T\}$. We have also higher regularity for weak solutions of parabolic problems. More precisely, a solution $y \in L^2(0,T;H^1(X(u)))$ of (4.51)-(4.54) is an element of $L^2(0,T;H^2(X(u)))$ if $f \in L^2(0,T;L^2(X(u)))$ and $y_0 \in H^1(X(u))$. See for instance [31, Theorem 6, Subsection 7.1.3]). Thus, it is guaranteed that $\frac{\partial y|_{X_1(u)}}{\partial n}, \frac{\partial y|_{X_2(u)}}{\partial n} \in L^2(0,T;H^{1/2}(u))$ by the trace theorem for Sobolev spaces.

In the setting above, the boundary value problem (4.51)-(4.54) is written in weak form as

$$a(y,p) = b(y,p,p_1,p_2) \quad \forall p \in W(0,T;H^1(X(u))) \tag{4.57}$$

and for all $p_1 \in L^2(0,T;H^{-1/2}(\Gamma_{\text{top}}))$ and $p_2 \in L^2(0,T;H^{1/2}(\Gamma_{\text{bottom}} \cup \Gamma_{\text{left}} \cup \Gamma_{\text{right}}))$. Here the bilinear form is given by

$$\begin{aligned} a(y,p) = &- \int_{X(u)} y_0\, p(0,x)\, dx \\ &- \int_0^T \int_{X(u)} \frac{\partial p}{\partial t} y \, dx\, dt + \int_0^T \int_{X(u)} k \nabla y^T \nabla p \, dx\, dt \\ &- \int_0^T \int_u \left[\!\!\left[k\frac{\partial y}{\partial n} p\right]\!\!\right] ds\, dt - \int_0^T \int_{\Gamma_{\text{out}}} k_1 \frac{\partial y}{\partial n} p \, ds\, dt \end{aligned} \tag{4.58}$$

and the linear form is given by

$$b(y,p,p_1,p_2) = b_1(p) + b_2(y,p_1,p_2) \tag{4.59}$$

with

$$b_1(p) = \int_0^T \int_{X(u)} fp \; dx\, dt, \tag{4.60}$$

$$b_2(y,p_1,p_2) = \int_0^T \int_{\Gamma_{\text{top}}} p_1(y-1) \; ds\, dt + \int_0^T \int_{\Gamma_{\text{out}}\setminus\Gamma_{\text{top}}} p_2 \frac{\partial y}{\partial n} \; ds\, dt. \tag{4.61}$$

For $(y,u,p) \in L^2(0,T;H^1(X(u))) \times B_e \times W(0,T;H^1(X(u)))$ the Lagrangian corresponding to (4.50)-(4.56) is defined by

$$\mathscr{L}(y,u,p) \coloneqq J(u) + a(y,p) - b(y,p,p_1,p_2) \tag{4.62}$$

with a, b defined in (4.58)-(4.61), J given in (4.50), $p_1 \in L^2(0,T;H^{-1/2}(\Gamma_{\text{top}}))$ and $p_2 \in L^2(0,T;H^{1/2}(\Gamma_{\text{bottom}} \cup \Gamma_{\text{left}} \cup \Gamma_{\text{right}}))$.

Remark 4.27. For all $y, p \in W(0,T;H^1(X(u)))$ the integration by parts formula

$$\begin{aligned} &\int_0^T \int_{X(u)} \frac{\partial y}{\partial t} p \; dx\, dt \\ &= \int_{X(u)} y(T,x)\, p(T,x) \; dx - \int_{X(u)} y(0,x)\, p(0,x) \; dx - \int_0^T \int_{X(u)} \frac{\partial p}{\partial t} y \; dx\, dt \end{aligned} \tag{4.63}$$

holds (cf. [100, Theorem 3.11]). Note that the integrals are to be understood as duality pairings.

Remark 4.28. Let H denote a Banach space and H' its dual. In [100], it is shown that a weak solution $y \in L^2(0,T;H)$ to a special initial boundary value problem exists (cf. [100, Theorem 3.9]). Moreover, possibly after a modification of y on a set of zero measure, the weak time derivative of this solution $y \in L^2(0,T;H)$ exists and is an element of $L^2(0,T;H')$ (cf. [100, Theorem 3.12]). Arguments identical to those in the proof of Theorem 3.12 in [100] ensure that a solution $y \in L^2(0,T;H^1(X(u)))$ of (4.51)-(4.54) is – possibly after a modification on a set of zero measure – an element of $W(0,T;H^1(X(u)))$. In particular, this means that the integration by parts formula (4.63) can be applied for a solution $y \in L^2(0,T;H^1(X(u)))$ of (4.51)-(4.54).

Remark 4.29. The integral over X is understood as the sum of the integrals over X_1 and X_2 in this thesis.

Shape derivative formulas

As in the previous subsection, we first consider the objective function J in (4.50) without perimeter regularization, i.e., the objective function j. The shape derivative

of j can be expressed as domain and boundary integral. Analogous to the previous subsection, we first deduce the domain integral of the shape derivative by an application of Theorem 4.19. Afterwards, we convert it into an interface integral by applying integration by parts on u. Due to Theorem 4.19, the shape derivative of $\mathscr{L}$ evaluated in its saddle point is equal to the one of J. A saddle point

$$(y, u, p) \in L^2(0,T; H^1(X(u))) \times B_e \times W(0,T; H^1(X(u)))$$

of the Lagrangian (4.62) is given by

$$\frac{\partial \mathscr{L}(y,u,p)}{\partial y} = \frac{\partial \mathscr{L}(y,u,p)}{\partial p} = 0, \tag{4.64}$$

which leads to the adjoint equation given in strong form by

$$-\frac{\partial p}{\partial t} - \operatorname{div}(k\nabla p) = -(y - \overline{y}) \quad \text{in } X(u) \times [0,T) \tag{4.65}$$

$$p = 0 \quad \text{in } X(u) \times \{T\} \tag{4.66}$$

$$\frac{\partial p}{\partial n} = 0 \quad \text{on } (\Gamma_{\text{bottom}} \cup \Gamma_{\text{left}} \cup \Gamma_{\text{right}}) \times [0,T) \tag{4.67}$$

$$p = 0 \quad \text{on } \Gamma_{\text{top}} \times [0,T) \tag{4.68}$$

$$p_1 = -k_1 p \quad \text{on } (\Gamma_{\text{bottom}} \cup \Gamma_{\text{left}} \cup \Gamma_{\text{right}}) \times [0,T) \tag{4.69}$$

$$p_2 = k_1 \frac{\partial p}{\partial n} \quad \text{on } \Gamma_{\text{top}} \times [0,T) \tag{4.70}$$

and to the state equation given in strong form by

$$\frac{\partial y}{\partial t} - \operatorname{div}(k\nabla y) = f \quad \text{in } X(u) \times (0,T]. \tag{4.71}$$

We formulate explicitly the interface conditions of (4.65)-(4.70) by

$$\left[\!\left[k\frac{\partial p}{\partial n} \right]\!\right] = 0, \quad [\![p]\!] = 0 \qquad \text{on } u \times [0,T) \tag{4.72}$$

The next theorem gives a domain expression of the shape derivative of j, which we need in Chapter 6. Later on, this domain expression is used in order to express the shape derivative of j and J as a boundary integral.

Theorem 4.30. *Let $y \in L^2(0,T; H^1(X(u)))$ be the weak solution of the parabolic PDE problem (4.51)-(4.54). Moreover, let $p \in W(0,T; H^1(X(u)))$ denote the weak solution of the adjoint equation (4.65)-(4.70). Then the shape derivative of the objective function J without perimeter regularization, i.e., the shape derivative of j, in direction V is given by*

$$\begin{aligned} &Dj_{X(u)}(u)[V] \\ &= \int_0^T \int_{X(u)} -k\nabla y^T \left(\nabla V + \nabla V^T\right) \nabla p - p \nabla f^T V \\ &\qquad + \operatorname{div}(V) \left(\frac{1}{2}(y - \overline{y})^2 + \frac{\partial y}{\partial t} p + k \nabla y^T \nabla p - fp \right) dx\, dt. \end{aligned} \tag{4.73}$$

Proof. The solution $y \in L^2(0,T;H^1(X(u)))$ of the parabolic PDE problem (4.51)-(4.54) is – possibly after a modification on a set of zero measure – an element of $W(0,T;H^1(X(u)))$ (cf. Remark 4.28). Thus, the integration by parts formula (4.63) can be applied.

Let

$$\begin{aligned}(y,u,p) &\in L^2(0,T;H^1(X(u))) \times B_e \times W(0,T;H^1(X(u))),\\ p_1 &\in L^2(0,T;H^{-1/2}(\Gamma_{\text{top}})),\\ p_2 &\in L^2(0,T;H^{1/2}(\Gamma_{\text{bottom}} \cup \Gamma_{\text{left}} \cup \Gamma_{\text{right}})).\end{aligned}$$

Moreover, let $\widetilde{\mathscr{L}}$ be defined by

$$\widetilde{\mathscr{L}}(y,u,p) = j(u) + a(y,p) - b(y,p,p_1,p_2), \tag{4.74}$$

where a, b are given in (4.58)-(4.61) and j is given in (4.50). In analogy to [23, Chapter 10, Subsection 5.2], we can verify that

$$j(u) = \min_{y \in L^2(0,T;H^1(X(u)))} \max_{p \in W(0,T;H^1(X(u)))} \widetilde{\mathscr{L}}(y,u,p) \tag{4.75}$$

holds. We apply Theorem 4.19 on the right-hand side of (4.75). This means that we obtain formula (4.73) by evaluation of the shape derivative of the Lagrangian (4.74) in its saddle point. The verification of the assumptions of this theorem can be checked in the same way as in [23, Chapter 10, Subsection 6.4]. Applying the rule for differentiating domain and boundary integrals given in (4.9) and (4.11) combined with (4.63) yields

$$\begin{aligned}&D\widetilde{\mathscr{L}}(y,u,p)[V]\\ &= -\int_{X(u)} D_m\left(y_0 p(0,x)\right)\, dx\\ &\quad + \int_0^T \int_{X(u)} \frac{1}{2} D_m\left((y-\overline{y})^2\right) - D_m\left(\frac{\partial p}{\partial t} y\right) + D_m\left(k \nabla y^T \nabla p\right) - D_m\left(fp\right)\\ &\qquad\qquad + \operatorname{div}(V)\left(\frac{1}{2}(y-\overline{y})^2 + \frac{\partial y}{\partial t} p + k \nabla y^T \nabla p - fp\right) dx\, dt\\ &\quad - \int_0^T \int_u D_m\left(\left[\!\left[k \frac{\partial y}{\partial n} p\right]\!\right]\right) + \operatorname{div}_u(V)\left[\!\left[k \frac{\partial y}{\partial n} p\right]\!\right] ds\, dt\\ &\quad - \int_0^T \int_{\Gamma_{\text{out}}} D_m\left(k_1 \frac{\partial y}{\partial n} p\right) + \operatorname{div}_{\Gamma_{\text{out}}}(V) k_1 \frac{\partial y}{\partial n} p\, ds\, dt\\ &\quad - \int_0^T \int_{\Gamma_{\text{top}}} D_m(p_1(y-1)) + \operatorname{div}_{\Gamma_{\text{top}}}(V) p_1 (y-1)\, ds\, dt\\ &\quad - \int_0^T \int_{\Gamma_{\text{out}} \setminus \Gamma_{\text{top}}} D_m\left(p_2 \frac{\partial y}{\partial n}\right) + \operatorname{div}_{\Gamma_{\text{out}} \setminus \Gamma_{\text{top}}}(V) p_2 \frac{\partial y}{\partial n}\, ds\, dt\end{aligned}$$

Applying (4.6) and (4.8) gives

$$
\begin{aligned}
&D\widetilde{\mathscr{L}}(y,u,p)[V] \\
&= -\int_{X(u)} y_0 \dot{p}(0,x)\, dx \\
&\quad + \int_0^T \int_{X(u)} (y-\overline{y})\dot{y} - D_m\left(\frac{\partial p}{\partial t}\right) y - \frac{\partial p}{\partial t}\dot{y} - \dot{f}p - f\dot{p} \\
&\qquad\qquad + k\nabla\dot{y}^T\nabla p + k\nabla y^T\nabla\dot{p} - k\nabla y^T\left(\nabla V + \nabla V^T\right)\nabla p \\
&\qquad\qquad + \operatorname{div}(V)\left(\frac{1}{2}(y-\overline{y})^2 + \frac{\partial y}{\partial t}p + k\nabla y^T\nabla p - fp\right) dx\, dt \\
&\quad - \int_0^T \int_u \left[\!\!\left[D_m\left(k\frac{\partial y}{\partial n}\right) p + k\frac{\partial y}{\partial n}\dot{p} \right]\!\!\right] + \operatorname{div}_u(V) \left[\!\!\left[k\frac{\partial y}{\partial n}p \right]\!\!\right] ds\, dt \\
&\quad - \int_0^T \int_{\Gamma_{\text{out}}} D_m\left(k_1\frac{\partial y}{\partial n}\right) p + k_1\frac{\partial y}{\partial n}\dot{p} + \operatorname{div}_{\Gamma_{\text{out}}}(V) k_1\frac{\partial y}{\partial n}p\; ds\, dt \\
&\quad - \int_0^T \int_{\Gamma_{\text{top}}} \dot{p}_1(y-1) + p_1\dot{y} + \operatorname{div}_{\Gamma_{\text{top}}}(V) p_1(y-1)\; ds\, dt \\
&\quad - \int_0^T \int_{\Gamma_{\text{out}}\setminus\Gamma_{\text{top}}} \dot{p}_2\frac{\partial y}{\partial n} + p_2 D_m\left(\frac{\partial y}{\partial n}\right) + \operatorname{div}_{\Gamma_{\text{out}}\setminus\Gamma_{\text{top}}}(V) p_2\frac{\partial y}{\partial n}\; ds\, dt.
\end{aligned}
$$

Note that the outer boundary Γ_{out} is fixed. Thus, we can choose the deformation vector field V equals zero in small neighbourhoods of Γ_{out}. From (4.64) we obtain

$$
\begin{aligned}
&D\widetilde{\mathscr{L}}(y,u,p)[V] \\
&= \int_0^T \int_{X(u)} -k\nabla y^T\left(\nabla V + \nabla V^T\right)\nabla p - \dot{f}p \\
&\qquad\qquad + \operatorname{div}(V)\left(\frac{1}{2}(y-\overline{y})^2 + \frac{\partial y}{\partial t}p + k\nabla y^T\nabla p - fp\right) dx\, dt \\
&\quad + \int_0^T \int_u \left[\!\!\left[k\frac{\partial p}{\partial n}\dot{y} \right]\!\!\right] - \left[\!\!\left[D_m\left(k\frac{\partial y}{\partial n}\right) p \right]\!\!\right] - \operatorname{div}_u(V) \left[\!\!\left[k\frac{\partial y}{\partial n}p \right]\!\!\right] ds\, dt
\end{aligned}
$$

where the term $\dot{f}p$ is equal to $p\nabla f^T V$ in $X(u)$ due to (4.5). We get

$$
\begin{aligned}
\left[\!\!\left[k\frac{\partial p}{\partial n}\dot{y} \right]\!\!\right] &= \dot{y}\left[\!\!\left[k\frac{\partial p}{\partial n} \right]\!\!\right] = 0 \quad \text{on } u, \\
\left[\!\!\left[D_m\left(k\frac{\partial y}{\partial n}\right) p \right]\!\!\right] &= D_m\left(k\frac{\partial y}{\partial n}\right)[\![p]\!] = 0 \quad \text{on } u
\end{aligned}
$$

due to (4.56) and (4.72). Moreover, due to (4.56), (4.72) and (4.41), we get the identity

$$
\left[\!\!\left[k\frac{\partial y}{\partial n}p \right]\!\!\right] = 0 \quad \text{on } u.
$$

Applying Theorem 4.19 completes the proof. □

In Theorem 4.30, the domain shape derivative expression is obtained under the assumptions that the solution y of the parabolic PDE problem (4.51)-(4.54) belongs to $L^2(0,T;H^1(X(u)))$ and the solution p of the adjoint equation (4.65)-(4.70) belongs to $W(0,T;H^1(X(u)))$. Note that the domain integral (4.73) can be converted into a boundary integral. However, as in the previous subsection (cf. Remark 4.26), we need a higher regularity to provide boundary shape derivative expressions. More precisely, p has to be an $L^2(0,T;H^2(X(u)))$-function having weak first time derivatives in $L^2(0,T;H^1(X(u))')$ and y has to be an element of $L^2(0,T;H^2(X(u)))$. We introduce the space

$$Y(0,T;H,P) \coloneqq \left\{ y \in L^2(0,T;H) \colon \mathring{y} \in L^2(0,T;P') \text{ exists} \right\}.$$

This space consists of all $L^2(0,T;H)$-functions whose weak first time derivatives exist and are elements of $L^2(0,T;P')$, where H and P denote Banach spaces and P' denotes the dual space of P. The following theorem is a generalization of Lemma 1 in [73] for parabolic problems and provides two boundary expressions of the shape derivative.

Theorem 4.31. *Assume that a solution y of the parabolic PDE problem (4.51)-(4.54) exists and is at least in $L^2(0,T;H^2(X(u)))$. Moreover, assume that the adjoint equation (4.65)-(4.70) admits a solution $p \in Y(0,T;H^2(X(u)),H^1(X(u)))$. Then the shape derivative of the objective function J without perimeter regularization, i.e., the shape derivative of j, in direction V is given by*

$$Dj_u(u)[V] = \int_0^T \int_u \langle V,n\rangle \left[\!\!\left[-2k \frac{\partial y}{\partial n}\frac{\partial p}{\partial n} + k \nabla y^T \nabla p \right]\!\!\right] ds\,dt. \tag{4.76}$$

Let $y_1 \coloneqq y|_{X_1(u)}$ and $p_2 \coloneqq p|_{X_2(u)}$. Then the shape derivative of the objective function j in direction V can be expressed as

$$Dj_u(u)[V] = \int_0^T \int_u [\![k]\!]\, \nabla y_1^T \nabla p_2 \, \langle V,n\rangle \, ds\,dt. \tag{4.77}$$

Proof. In (4.73), let us consider

$$\begin{aligned}
&\int_{X(u)} \operatorname{div}(V) \left(\frac{1}{2}(y-\overline{y})^2 + \frac{\partial y}{\partial t} p + k \nabla y^T \nabla p - fp \right) dx \\
&= \int_{X_1(u)} \operatorname{div}(V) \left(\frac{1}{2}(y-\overline{y})^2 + \frac{\partial y}{\partial t} p + k_1 \nabla y^T \nabla p - fp \right) dx \\
&\quad + \int_{X_2(u)} \operatorname{div}(V) \left(\frac{1}{2}(y-\overline{y})^2 + \frac{\partial y}{\partial t} p + k_2 \nabla y^T \nabla p - fp \right) dx.
\end{aligned}$$

Applying integration by parts yields

$$\begin{aligned}
&\int_{X(u)} \operatorname{div}(V)\left(\frac{1}{2}(y-\overline{y})^2+\frac{\partial y}{\partial t}p+k\nabla y^T\nabla p-fp\right)dx\\
&=\int_{u\cup\Gamma_{\text{out}}}\left(\frac{1}{2}(y-\overline{y})^2+\frac{\partial y}{\partial t}p+k_1\nabla y^T\nabla p-fp\right)\langle V,n\rangle\,ds\\
&\quad-\int_{X_1(u)} V^T\left((y-\overline{y})\nabla y+\nabla\left(\frac{\partial y}{\partial t}p\right)+k_1\nabla\left(\nabla y^T\nabla p\right)-\nabla\left(fp\right)\right)dx\\
&\quad+\int_u\left(\frac{1}{2}(y-\overline{y})^2+\frac{\partial y}{\partial t}p+k_2\nabla y^T\nabla p-fp\right)\langle V,-n\rangle\,ds\\
&\quad-\int_{X_2(u)} V^T\left((y-\overline{y})\nabla y+\nabla\left(\frac{\partial y}{\partial t}p\right)+k_2\nabla\left(\nabla y^T\nabla p\right)-\nabla\left(fp\right)\right)dx\\
&=-\int_{X(u)} V^T\left((y-\overline{y})\nabla y+\nabla\left(\frac{\partial y}{\partial t}p\right)+k\nabla\left(\nabla y^T\nabla p\right)-\nabla fp-f\nabla p\right)dx\\
&\quad+\int_u\left[\!\!\left[\left(\frac{1}{2}(y-\overline{y})^2+\frac{\partial y}{\partial t}p+k\nabla y^T\nabla p-fp\right)\langle V,n\rangle\right]\!\!\right]ds\\
&\quad+\int_{\Gamma_{\text{out}}}\left(\frac{1}{2}(y-\overline{y})^2+\frac{\partial y}{\partial t}p+k_1\nabla y^T\nabla p-fp\right)\langle V,n\rangle\,ds.
\end{aligned}$$

Combining this with (4.73) and the vector calculus identity

$$\nabla y^T\left(\nabla V+\nabla V^T\right)\nabla p+V^T\nabla\left(\nabla y^T\nabla p\right)=\nabla p^T\nabla\left(V^T\nabla y\right)+\nabla y^T\nabla\left(V^T\nabla p\right),$$

which is proved in [16], gives

$$\begin{aligned}
&Dj(u)[V]\\
&=\int_0^T\Bigg[\int_{X(u)}-k\nabla p^T\nabla\left(V^T\nabla y\right)-k\nabla y^T\nabla\left(V^T\nabla p\right)-(y-\overline{y})V^T\nabla y\\
&\qquad\qquad-V^T\nabla\left(\frac{\partial y}{\partial t}p\right)+fV^T\nabla p\,dx\\
&\qquad+\int_u\left[\!\!\left[\left(\frac{1}{2}(y-\overline{y})^2+\frac{\partial y}{\partial t}p+k\nabla y^T\nabla p-fp\right)\langle V,n\rangle\right]\!\!\right]ds\\
&\qquad+\int_{\Gamma_{\text{out}}}\left(\frac{1}{2}(y-\overline{y})^2+\frac{\partial y}{\partial t}p+k_1\nabla y^T\nabla p-fp\right)\langle V,n\rangle\,ds\Bigg]dt.
\end{aligned}\tag{4.78}$$

Let us consider

$$\begin{aligned}
&\int_{X(u)} k\nabla y^T\nabla\left(V^T\nabla p\right)dx\\
&=\int_{X_1(u)} k_1\nabla y^T\nabla\left(V^T\nabla p\right)dx+\int_{X_2(u)} k_2\nabla y^T\nabla\left(V^T\nabla p\right)dx
\end{aligned}$$

in (4.78). Then by applying integration by parts we get

$$\begin{aligned}
&\int_{X(u)} k\nabla y^T \nabla \left(V^T \nabla p\right) dx \\
&= \int_{X_1(u)} k_1 \nabla y^T \nabla \left(V^T \nabla p\right) dx + \int_{X_2(u)} k_2 \nabla y^T \nabla \left(V^T \nabla p\right) dx \\
&= \int_{u \cup \Gamma_{\text{out}}} k_1 \frac{\partial y}{\partial n} V^T \nabla p \; ds - \int_{X_1(u)} \operatorname{div}(k_1 \nabla y) \nabla p^T V \, dx + \int_u k_2 \frac{\partial y}{\partial n} V^T \nabla p \; ds \\
&\quad - \int_{X_2(u)} \operatorname{div}(k_2 \nabla y) \nabla p^T V \, dx \\
&= - \int_{X(u)} \operatorname{div}(k \nabla y) \nabla p^T V \; dx + \int_u \left[\!\left[k \frac{\partial y}{\partial n} V^T \nabla p \right]\!\right] ds + \int_{\Gamma_{\text{out}}} k_1 \frac{\partial y}{\partial n} V^T \nabla p \; ds.
\end{aligned}$$

Analogously, we get

$$\begin{aligned}
&\int_{X(u)} k\nabla p^T \nabla \left(V^T \nabla y\right) dx \\
&= - \int_{X(u)} \operatorname{div}(k \nabla p) \nabla y^T V \, dx + \int_u \left[\!\left[k \frac{\partial p}{\partial n} V^T \nabla y \right]\!\right] ds + \int_{\Gamma_{\text{out}}} k_1 \frac{\partial p}{\partial n} V^T \nabla y \; ds.
\end{aligned}$$

Applying the integration by parts formula (4.63) in (4.78) yields

$$\int_0^T \int_{X(u)} V^T \nabla \frac{\partial y}{\partial t} p \; dx \, dt = - \int_0^T \int_{X(u)} V^T \nabla y \frac{\partial p}{\partial t} \; dx \, dt.$$

Thus, it follows that

$$\begin{aligned}
&Dj(u)[V] \\
&= \int_0^T \Bigg[\int_{X(u)} \nabla p^T V \left(-\frac{\partial y}{\partial t} + \operatorname{div}(k \nabla y) + f \right) \\
&\qquad + \nabla y^T V \left(\frac{\partial p}{\partial t} + \operatorname{div}(k \nabla p) - (y - \overline{y}) \right) dx \\
&\qquad + \int_u \left[\!\left[\left(\frac{1}{2}(y - \overline{y})^2 + \frac{\partial y}{\partial t} p - k \nabla y^T \nabla p - f p \right) \langle V, n \rangle \right]\!\right] \\
&\qquad\qquad - \left[\!\left[k \frac{\partial y}{\partial n} V^T \nabla p \right]\!\right] - \left[\!\left[k \frac{\partial p}{\partial n} V^T \nabla y \right]\!\right] ds \\
&\qquad + \int_{\Gamma_{\text{out}}} \left(\frac{1}{2}(y - \overline{y})^2 + \frac{\partial y}{\partial t} p - k_1 \nabla y^T \nabla p - f p \right) \langle V, n \rangle \\
&\qquad\qquad - k_1 \frac{\partial y}{\partial n} V^T \nabla p - k_1 \frac{\partial p}{\partial n} V^T \nabla y \; ds \Bigg] dt \\
&\quad + \int_{X(u)} V^T \nabla y(T, x) p(T, x) \; dx
\end{aligned} \tag{4.79}$$

holds. The domain integrals in (4.79) vanish due to (4.65), (4.66) and (4.71). Moreover, the term $\left[\!\left[\left(\frac{\partial y}{\partial t} - f\right) p\right]\!\right]$ vanishes because of (4.72) and the term $\left[\!\left[\frac{1}{2}(y-\overline{y})^2\right]\!\right]$ disappears because of (4.56). Then the identity

$$\left[\!\left[k\frac{\partial y}{\partial n}V^T\nabla p\right]\!\right] = \left[\!\left[k\frac{\partial p}{\partial n}V^T\nabla y\right]\!\right] = \langle V,n\rangle \left[\!\left[k\frac{\partial y}{\partial n}\frac{\partial p}{\partial n}\right]\!\right] \quad \text{on } u \tag{4.80}$$

follows from (4.56) and (4.72). Since the outer boundary Γ_{out} is not variable, we can choose the deformation vector field V equals zero in small neighbourhoods of Γ_{out}. Therefore, the outer integral in (4.79) disappears and we obtain the interface integral (4.76). It is easy to verify that

$$\int_u \langle V,n\rangle \left[\!\left[-2k\frac{\partial y}{\partial n}\frac{\partial p}{\partial n} + k\nabla y^T\nabla p\right]\!\right] ds = \int_u [\![k]\!]\, \nabla y_1^T \nabla p_2 \,\langle V,n\rangle \, ds, \tag{4.81}$$

which completes the proof. For a detailed computation of (4.80) and (4.81) we refer to [46, Subsection 11.3.2]. □

Now, we consider the objective function (4.50) with perimeter regularization. The following theorem provides two boundary integral expressions of its shape derivative.

Theorem 4.32. *Under the assumptions of Theorem 4.31, the shape derivative of the objective function J in direction V is given by*

$$DJ_u(u)[V] = \int_u \left[\int_0^T \langle V,n\rangle \left[\!\left[-2k\frac{\partial y}{\partial n}\frac{\partial p}{\partial n} + k\nabla y^T\nabla p\right]\!\right] dt + \langle V,n\rangle\, \mu\kappa\right] ds, \tag{4.82}$$

where κ denotes the curvature corresponding to the normal n. Let $y_1 \coloneqq y|_{X_1(u)}$ and $p_2 \coloneqq p|_{X_2(u)}$. Then the shape derivative of the objective function J in direction V can be expressed as

$$DJ_u(u)[V] = \int_u \left[\int_0^T \langle V,n\rangle\, [\![k]\!]\, \nabla y_1^T \nabla p_2 \, dt + \langle V,n\rangle\, \mu\kappa\right] ds. \tag{4.83}$$

Proof. Combining Theorem 4.31 with formula (4.13) in Theorem 4.15 we get (4.82) and (4.83). □

Remark 4.33. In Theorem 4.30, the volume shape derivative expression is obtained under the assumption that a solution y of the parabolic equation (4.51)-(4.54) belongs to $L^2(0,T;H^1(X(u)))$ and that a solution p of the adjoint equation (4.65)-(4.70) belongs to $W(0,T;H^1(X(u)))$. In contrast to this, we need a higher regularity to convert this volume expression to a surface integral. In Theorem 4.31 and Theorem 4.32, the shape derivative expressions are obtained under the assumptions that $y \in L^2(0,T;H^2(X(u)))$ and $p \in Y(0,T;H^2(X(u)),H^1(X(u)))$. A solution $y \in L^2(0,T;H^1(X(u)))$ of the parabolic problem (4.51)-(4.54) is an element of $L^2(0,T;H^2(X(u)))$ if $y_0 \in H^1(X(u))$ and $f \in L^2(0,T;L^2(X(u)))$. Furthermore, a solution $p \in W(0,T;H^1(X(u)))$ of the adjoint equation (4.65)-(4.70) is an element of $Y(0,T;H^2(X(u)),H^1(X(u)))$ if $y-\overline{y} \in L^2(0,T;L^2(X(u)))$. See for instance [31, Theorem 6, Subsection 7.1.3]).

4.3 Riemannian shape calculus

As pointed out in [84], shape optimization can be viewed as optimization on Riemannian shape manifolds and the resulting optimization methods can be constructed and analyzed within this framework. This combines algorithmic ideas from [1] with the differential geometrical point of view established in [11]. In this short section, we analyze the connection of Riemannian geometry on the shape space B_e introduced in the previous chapter to shape optimization in order to prepare Chapter 5.

Remark 4.34. In Chapter 6, the shape space B_e containing $\mathcal{C}^\infty$-shapes unnecessarily limits the application of methods established in this chapter. More precisely, numerical investigations have shown that the optimization techniques also work on shapes with kinks in the boundary. This means that they are not limited to elements of B_e and another shape space definition is required. Thus, in Chapter 7, we extend the definition of $\mathcal{C}^\infty$-shapes to shapes of class $H^{1/2}$ and propose a novel shape space. Its connection to shape calculus can be found in Section 7.1.

As already mentioned in Chapter 3, we mainly consider the Sobolev metric g^1 on the shape space B_e. The Riemannian connection with respect to this metric is given in Theorem 3.3. This connection makes it possible to specify the Riemannian shape Hessian of an optimization problem. First, we have to detail the *Riemannian shape gradient*. Due to the Hadamard Structure Theorem 4.7, there exists a scalar distribution r on the boundary Γ of the domain Ω under consideration. If we assume $r \in L^1(\Gamma)$, the shape derivative can be expressed on the boundary Γ of Ω as

$$DJ_\Gamma[V] = \int_\Gamma r \langle V, n\rangle \, ds. \tag{4.84}$$

The distribution r is often called the *shape gradient*. This terminology is also used in this thesis. However, note that gradients depend always on chosen scalar products defined on the space under consideration. Thus, it rather means that r is the usual L^2-shape gradient. Since we want to optimize on shape manifolds, we have to find a representation of the shape gradient with respect to a Riemannian metric defined on the shape manifold under consideration. This representation is called the Riemannian shape gradient. The shape derivative can be expressed (more concisely) as

$$DJ_\Gamma[V] = \int_\Gamma \alpha r \, ds$$

if $V|_{\partial\Omega} = \alpha n$. In order to get an expression of the Riemannian shape gradient with respect to the Sobolev metric g^1, we look at the isomorphism

$$T_c B_e \cong \left\{h \colon h = \alpha n,\, \alpha \in \mathcal{C}^\infty(S^1)\right\}$$

given in (3.2). Due to this isomorphism, a tangent vector $h \in T_\Gamma B_e$ is given by $h = \alpha n$ with $\alpha \in \mathcal{C}^\infty(\Gamma)$. This leads to the following definition.

Definition 4.35 (Riemannian shape gradient with respect to the Sobolev metric)**.** *A Riemannian representation of the shape derivative, i.e., the Riemannian shape gradient of a shape differentiable objective function J in terms of the Sobolev metric g^1, is given by*

$$\operatorname{grad}(J) = qn \quad \text{with} \quad (I - A\triangle_\Gamma)q = r, \tag{4.85}$$

where $\Gamma \in B_e$, $A > 0$, $q \in \mathcal{C}^\infty(\Gamma)$ and r denotes the (standard) shape gradient given in (4.84).

Now, we can specify the *Riemannian shape Hessian*. It is based on the Riemannian connection ∇ related to the Sobolev metric g^1. This Riemannian connection is given in Theorem 3.3. In analogy to [1], we can define the Riemannian shape Hessian as follows:

Definition 4.36 (Riemannian shape Hessian)**.** *In the setting above, the Riemannian shape Hessian of a two times shape differentiable objective function J is defined as the linear mapping*

$$T_\Gamma B_e \to T_\Gamma B_e,\ h \mapsto \operatorname{Hess}(J)[h] \coloneqq \nabla_h \operatorname{grad}(J). \tag{4.86}$$

Now, the Riemannian shape gradient and the Riemannian shape Hessian are defined. These two objects are required in Chapter 5, in which we are concerned with PDE constrained optimization on shape manifolds.

4.4 A view on volume shape derivative formulas

In this section, we provide a volume shape derivative formula for a special class of PDE constrained shape interface problems. For the convenience of the reader we do not consider regularization terms in the objective function. Of course, if necessary, they can be added in the same way as in the problems of the previous sections.

Problem formulation

Let $d, l \in \mathbb{N}^*$ and $T > 0$. We denote by $X \subset \mathbb{R}^d$ a bounded domain with Lipschitz boundary ∂X. Let this domain be partitioned into two subdomains $X_1, X_2 \subset X$ separated by an interface $\Gamma_{\text{int}} = \partial X_1 \cap \partial X_2$ such that $X_1 \cup \Gamma_{\text{int}} \cup X_2 = X$. This interface is assumed to be smooth and variable. In contrast, the outer boundary $\Gamma_{\text{out}} = \partial X$ is assumed to be fixed. Since the domain X depends on Γ_{int}, we write $X(\Gamma_{\text{int}})$. To be in line with Subsection 4.2.2, we replace Γ_{int} by u. Examples of such partitioned domains are illustrated for $d = 2$ in Figure 4.1.

In the following, let $a_1(t, x; \cdot, \cdot)$ denote a bilinear form on $\mathbb{R}^{d\times l} \times \mathbb{R}^{d\times l}$ for all $(t, x) \in [0, T] \times \overline{X(u)}$. Moreover, let $a_2(t, x; \cdot)\colon \mathbb{R}^{d\times l} \to \mathbb{R}^l$ and $i(t, x; \cdot)\colon \mathbb{R}^l \to \mathbb{R}$ be functions for all $(t, x) \in [0, T] \times \overline{X(u)}$. We assume that $a_1(\cdot, \cdot; \xi, \eta)$ and $a_2(\cdot, \cdot; \xi)$ are continuously differentiable for all $\xi, \eta \in \mathbb{R}^{d\times l}$. Furthermore, $i(\cdot, \cdot; \zeta)$ is assumed to be continuously differentiable for all $\zeta \in \mathbb{R}^l$.

Let $y\colon [0,T]\times X(u)\to\mathbb{R}^l$ be an element of $L^2(0,T;H^1(X(u),\mathbb{R}^l))$. Moreover, let a_1, a_2 and i be integrable in $[0,T]\times\overline{X(u)}$ such that the following integrals exist. We consider the following PDE constrained shape interface problem in weak form:

$$\min_u\ J(u) \coloneqq \int_0^T\int_{X(u)} i(t,x;y)\,dx\,dt \tag{4.87}$$

$$\text{s.t. } a(y,p)=b(p)\quad \forall p\in W(0,T;H^1(X(u),\mathbb{R}^l)) \tag{4.88}$$

with bilinear form

$$\begin{aligned} a(y,p) = &-\int_{X(u)} y_0^T p(0,x)\,dx \\ &-\int_0^T\int_{X(u)} y^T\frac{\partial p}{\partial t}\,dx\,dt+\int_0^T\int_{X(u)} a_1(t,x;M\nabla y,\nabla p)\,dx\,dt \\ &-\int_0^T\int_{\Gamma_{\text{out}}} a_2(t,s;M_1\nabla y)p\,ds\,dt-\int_0^T\int_u [\![a_2(t,s;M\nabla y)p]\!]\,ds\,dt \end{aligned} \tag{4.89}$$

and linear form

$$b(p)=\int_0^T\int_{X(u)} f^T p\,dx\,dt, \tag{4.90}$$

where

$$f\coloneqq\begin{cases} f_1=\text{const.} & \text{in } X_1(u)\times(0,T]\\ f_2=\text{const.} & \text{in } X_2(u)\times(0,T]\end{cases}$$

and $[\![\cdot]\!]$ denotes the jump symbol defined in (4.21). Note that the weak equation (4.88)-(4.90) contains jump conditions such that there is a discontinuity across the interface u. Among other things, these conditions are described by a $(d\times d)$-matrix denoted by M. We formulate explicitly the continuity of the state and the flux at the interface as

$$[\![y]\!]=0\,,\quad [\![a_2(t,s;M\nabla y)]\!]=0\quad \text{on } u\times(0,T], \tag{4.91}$$

Note that the flux condition is given by a_2. In general, this flux condition includes the outward unit normal n to $X_2(u)$. Examples of such flux conditions can be found at the end of this section. Note that there is the initial condition

$$y=y_0\quad \text{in } X(u)\times\{0\}$$

included in (4.89). Let $y_0\in H^1(X(u),\mathbb{R}^l)$ be given. Intentionally, the weak formulation of the PDE (4.87)-(4.90) is chosen for ease of presentation. Note that the boundary integrals in (4.89) are to be understood as duality pairings.

Remark 4.37. As in Subsection 4.2.2, we have to assume $p=0$ in $X(u)\times\{T\}$ in order to get rid of the term $y(T,x)$ in (4.89). Note that this value is not necessarily defined because $y\in L^2(0,T;H^1(X(u),\mathbb{R}^l))$ need not to be continuous in time.

Remark 4.38. The integration by parts in time can be applied. More precisely,

$$\begin{aligned}&\int_0^T \int_{X(u)} \frac{\partial y}{\partial t}^T p \, dx \, dt \\ &= \int_{X(u)} y(T,x)^T p(T,x) \, dx - \int_{X(u)} y(0,x)^T p(0,x) \, dx - \int_0^T \int_{X(u)} y^T \frac{\partial p}{\partial t} \, dx \, dt\end{aligned}$$

holds for all $y, p \in W(0,T; H^1(X(u), \mathbb{R}^l))$.

Shape derivative formula

As in the previous sections, we deduce the shape derivative of the problem above by applying Theorem 4.19. Note that this subsection provides a volume shape derivative formula with regard to Chapter 6. This chapter establishes optimization algorithms, which require only boundary shape derivative formulas and, thus, do not need surface shape derivative formulas. Note that we do not convert the volume integral expression of the shape derivative into an interface integral as in the sections before.

First, we build the Lagrangian corresponding to the problem above. We define

$$K \coloneqq L^2(0,T; H^1(X(u), \mathbb{R}^l)) \times B_e(S^{l-1}, \mathbb{R}^l) \times W(0,T; H^1(X(u), \mathbb{R}^l)). \tag{4.92}$$

For $(y,u,p) \in K$ the Lagrangian is given by

$$\mathscr{L}(y,u,p) \coloneqq J(u) + a(y,p) - b(p), \tag{4.93}$$

where J is defined in (4.87) and a, b are given in (4.89)-(4.90). Due to the theorem of Correa and Seeger, the shape derivative of $\mathscr{L}$ evaluated in its saddle point is equal to the one of J. A saddle point $(y,u,p) \in K$ of the Lagrangian (4.93) is given by

$$\frac{\partial \mathscr{L}(y,u,p)}{\partial y} = \frac{\partial \mathscr{L}(y,u,p)}{\partial p} = 0, \tag{4.94}$$

which leads to the adjoint equation given in weak form by

$$\begin{aligned}&\int_0^T \int_{X(u)} \frac{\partial p}{\partial t}^T q \, dx \, dt + \int_0^T \int_{X(u)} a_1(t,x; M\nabla p, \nabla q) \, dx \, dt \\ &- \int_0^T \int_{\Gamma_{\text{out}}} a_2(t,s; M_1 \nabla p) \, q \, ds \, dt - \int_0^T \int_u [\![a_2(t,s; M\nabla p) \, q]\!] \, ds \, dt \\ &= - \int_0^T \int_{X(u)} \frac{\partial i(t,x;y)}{\partial y} q \, dx \, dt \quad \forall q \in L^2(0,T; H^1(X(u), \mathbb{R}^l))\end{aligned} \tag{4.95}$$

and to the design equation given in weak form by (4.88)-(4.90). Note that we have interface conditions on the interface for the adjoint p. We formulate them explicitly as

$$[\![p]\!] = 0\,, \quad [\![a_2(t,s; M\nabla p)]\!] = 0 \quad \text{on } u \times [0,T). \tag{4.96}$$

Now, we are able to formulate the following theorem which provides a volume shape derivative formula.

Theorem 4.39. *Assume that a solution y of the PDE problem (4.88)-(4.90) exists and is at least in $L^2(0,T;H^1(X(u),\mathbb{R}^l))$. Moreover, assume that the adjoint equation (4.95) admits a solution p which is at least in $W(0,T;H^1(X(u),\mathbb{R}^l))$. If the assumptions of the theorem of Correa and Seeger are fulfilled, the shape derivative of the objective function J defined in (4.16) is given by*

$$\begin{aligned} &DJ_{X(u)}(u)[V] \\ &= \int_0^T \int_{X(u)} - a_1(t,x;M\nabla V^T\nabla y,\nabla p) - a_1(t,x;M\nabla y,\nabla V^T\nabla p) - V^T\nabla f p \\ &\qquad + \operatorname{div}(V)\Big(i(t,x;y) + \frac{\partial y}{\partial t}^T p + a_1(t,x;M\nabla y,\nabla p) - f^T p\Big)\,dx\,dt. \end{aligned} \tag{4.97}$$

Proof. Due to arguments identical to those in the proof of Theorem 3.12 in [100], a solution $y \in L^2(0,T;H^1(X(u),\mathbb{R}^l))$ of the PDE problem (4.88)-(4.90) is – possibly after a modification on a set of zero measure – an element of $W(0,T;H^1(X(u),\mathbb{R}^l))$. This means that integration by parts in time can be applied (cf. Remark 4.38).

Let $u \in B_e(S^{l-1},\mathbb{R}^l)$ and let $\mathscr{L}$ be the Lagrangian defined in (4.93). In analogy to [23, Chapter 10, Subsection 5.2], we can verify that

$$J(u) = \min_{y\in L^2(0,T;H^1(X(u),\mathbb{R}^l))} \max_{p\in W(0,T;H^1(X(u),\mathbb{R}^l))} \mathscr{L}(y,u,p) \tag{4.98}$$

holds. We apply Theorem 4.19 on the right-hand side of (4.98). This means that we obtain formula (4.97) by evaluation of the shape derivative of the Lagrangian (4.93) in its saddle point.

Applying the rules for differentiating domain and boundary integrals given in (4.9) and (4.11) combined with integration by parts in time yields

$$\begin{aligned} &D\mathscr{L}(y,u,p)[V] \\ &= -\int_{X(u)} D_m\left(y_0^T p(0,x)\right)\,dx + \int_0^T\int_{X(u)} D_m i(t,x;y)\,dx\,dt \\ &\quad + \int_0^T\int_{X(u)} -D_m\left(y^T\frac{\partial p}{\partial t}\right) + D_m\left(a_1(t,x;M\nabla y,\nabla p)\right) - D_m\left(f^T p\right) \\ &\qquad\qquad + \operatorname{div}(V)\left(i(t,x;y) + \frac{\partial y}{\partial t}^T p + a_1(t,x;M\nabla y,\nabla p) - f^T p\right)dx\,dt \\ &\quad - \int_0^T\int_u D_m\left([\![a_2(t,s;M\nabla y)p]\!]\right) + \operatorname{div}_u(V)\,[\![a_2(t,s;M\nabla y)p]\!]\,ds\,dt \\ &\quad - \int_0^T\int_{\Gamma_{\text{out}}} D_m\left(a_2(t,s;M_1\nabla y)p\right) + \operatorname{div}_{\Gamma_{\text{out}}}(V)\,a_2(t,s;M_1\nabla y)p\,ds\,dt. \end{aligned}$$

From this, by applying the rules for material derivatives given in Section 4.1 we

obtain

$$
\begin{aligned}
&D\mathscr{L}(y,u,p)[V] \\
&= \int_{X(u)} -\dot{y}_0^T\, p(0,x) - y_0^T\, \dot{p}(0,x)\, dx \\
&\quad + \int_0^T \int_{X(u)} \frac{\partial i(t,x;y)}{\partial y}\dot{y} - y^T \frac{\partial \dot{p}}{\partial t} - \dot{y}^T \frac{\partial p}{\partial t} + \underbrace{a_1(t,x;M\nabla\dot{y} - M\nabla V^T \nabla y, \nabla p)}_{=a_1(t,x;M\nabla\dot{y},\nabla p) - a_1(t,x;M\nabla V^T\nabla y,\nabla p)} \\
&\qquad\qquad + \underbrace{a_1(t,x;M\nabla y, \nabla\dot{p} - \nabla V^T \nabla p)}_{=a_1(t,x;M\nabla y,\nabla\dot{p}) - a_1(t,x;M\nabla y,\nabla V^T\nabla p)} - \dot{f}^T p - f^T \dot{p} \\
&\qquad\qquad + \operatorname{div}(V)\left(i(t,x;y) + \frac{\partial y}{\partial t}^T p + a_1(t,x;M\nabla y,\nabla p) - f^T p \right) dx\, dt \\
&\quad - \int_0^T \int_u [\![D_m a_2(t,s;M\nabla y)p + a_2(t,s;M\nabla y)\dot{p}]\!] + \operatorname{div}_u(V)\, [\![a_2(t,s;M\nabla y)p]\!]\, ds\, dt \\
&\quad - \int_0^T \int_{\Gamma_{\text{out}}} D_m a_2(t,s;M_1\nabla y)p + a_2(t,s;M_1\nabla y)\dot{p} + \operatorname{div}_{\Gamma_{\text{out}}}(V)\, a_2(t,s;M_1\nabla y)p\, ds\, dt,
\end{aligned}
$$

where the term $\dot{f}^T p$ is equal to $V^T \nabla f p$ due to (4.5). The outer boundary Γ_{out} is not variable. Thus, we can choose the deformation vector field V equals zero in small neighbourhoods of Γ_{out}. Now, let us consider the saddle point condition (4.94). This gives

$$
\begin{aligned}
&d\mathscr{L}(y,u,p)[V] \\
&= \int_0^T \int_{X(u)} -a_1(t,x;M\nabla V^T\nabla y,\nabla p) - a_1(t,x;M\nabla y,\nabla V^T\nabla p) - V^T\nabla f p \\
&\qquad\qquad + \operatorname{div}(V)\left(i(t,x;y) + \frac{\partial y}{\partial t}^T p + a_1(t,x;M\nabla y,\nabla p) - f^T p \right) dx\, dt \\
&\quad + \int_0^T \int_u [\![a_2(t,s;M\nabla p)\dot{y}]\!] - [\![D_m a_2(t,s;M\nabla y)p]\!] \\
&\qquad\qquad - \operatorname{div}_u(V)\, [\![a_2(t,s;M\nabla y)p]\!]\, ds\, dt
\end{aligned}
$$

Due to (4.91) and (4.96), we get

$$
\begin{aligned}
[\![a_2(t,s;M\nabla p)\dot{y}]\!] = [\![a_2(t,s;M\nabla p)]\!]\, \dot{y} = 0 \quad \text{on } u \times (0,T], \\
[\![D_m a_2(t,x;M\nabla y)p]\!] = D_m a_2(t,x;M\nabla y)\, [\![p]\!] = 0 \quad \text{on } u \times (0,T].
\end{aligned}
$$

The identity

$$
[\![a_2(t,s;M\nabla y)p]\!] = 0 \quad \text{on } u \times (0,T),
$$

which follows from (4.40), (4.91) and (4.96), completes the proof. □

The shape derivative formula (4.97) can be applied, for example, to the elliptic problem (4.16)-(4.20) of Subsection 4.2.1 or to the parabolic problem (4.50)-(4.56) of Subsection 4.2.2, where we ignore the regularization term j_{reg} in the objective functions J given in (4.16) and (4.50):

- **Elliptic problem**
  Without time dependence and by choosing $l = 1$, $d = 2$, $M = I$ and

  $$a_1(x;\cdot,\cdot)\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R},\ (\xi,\eta) \mapsto \xi^T\eta$$

  we get (4.32). Note that the flux condition is given by

  $$a_2(s;\cdot,\cdot)\colon \mathbb{R}^2 \to \mathbb{R},\ \xi \mapsto \xi^T n.$$

- **Parabolic problem**
  By choosing $l = 1$, $d = 2$, $M = kI$, $f = f_1 = f_2 = \text{const.}$ in $[0,T] \times X(u)$ and

  $$a_1(t,x;\cdot,\cdot)\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R},\ (\xi,\eta) \mapsto \xi^T\eta$$

  we get (4.73). Note that the flux condition is given by

  $$a_2(s;\cdot,\cdot)\colon \mathbb{R}^2 \to \mathbb{R},\ \xi \mapsto \xi^T n.$$

Another example for the application of the shape derivative formula (4.97) is the following shape optimization problem (without regularization term) constrained by the linear elasticity equation:

$$\min_{u \in B_e(S^2,\mathbb{R}^3)} J(u) \coloneqq \int_{X(u)} \sigma(y) : \epsilon(y)\, dx \tag{4.99}$$

$$\text{s.t.} - \operatorname{div}(\sigma(y)) = f \quad \text{in } X(u) \tag{4.100}$$

$$y = 0 \quad \text{on } \Gamma_{\text{out}} \tag{4.101}$$

with

$$\sigma(y) \coloneqq \lambda \operatorname{tr}(\epsilon(y)) I + 2\mu\epsilon(y), \tag{4.102}$$

$$\epsilon(y) \coloneqq \frac{1}{2}(\nabla y + \nabla y^T). \tag{4.103}$$

In linear elasticity, σ is called the stress tensor and ϵ is called the strain tensor. λ and μ denote the Lamé parameters, which can be expressed in terms of Young's modulus E and Poisson's ratio ν as

$$\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}, \tag{4.104}$$

$$\mu = \frac{E}{2(1+\nu)}. \tag{4.105}$$

In the objective function (4.99), we use the inner product of two second-order tensors. It is given as the sum of component-wise products, i.e.,

$$\sigma(y) : \epsilon(y) \coloneqq \sum_{ij} \sigma(y)_{ij}\epsilon(y)_{ij}.$$

We formulate explicitly the continuity of the state and the flux at the interface by

$$[\![y]\!] = 0, \quad [\![\sigma(y) \cdot n]\!] = 0 \quad \text{on } u. \tag{4.106}$$

Note that we have no time dependence. We set $l = d = 3$ and $M = I$. Moreover, we choose $f = f_1 = f_2 = \text{const.}$ in $X(u)$. The bilinear form $a_1(x, \cdot, \cdot)$ is given by

$$a_1(x; \cdot, \cdot) \colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}, \ (\nabla y, \nabla p) \mapsto \sigma(y) : \nabla p.$$

Furthermore, the flux condition is given by

$$a_2(s; \cdot) \colon \mathbb{R}^{3\times 3} \to \mathbb{R}^3, \ \nabla y \mapsto \sigma(y) \cdot n.$$

Applying the shape derivative formula (4.97) on the PDE constrained problem (4.99)-(4.101) yields

$$\begin{aligned} DJ_{X(u)}(u)[V] = \int_{X(u)} \Big[&- \sigma(y) : (\nabla V^T \nabla p) - \sigma(p) : (\nabla V^T \nabla y) - V^T \nabla f p \\ &+ \operatorname{div}(V)\big(\sigma(y) : \epsilon(y) + \sigma(p) : \nabla y - f^T p\big) \Big] \, dx. \end{aligned} \tag{4.107}$$

For the sake of completeness, we give the adjoint equation corresponding to the PDE constrained problem (4.99)-(4.101):

$$-\operatorname{div}(\sigma(p)) = -f \quad \text{in } X(u) \tag{4.108}$$

$$p = 0 \quad \text{on } \Gamma_{\text{out}} \tag{4.109}$$

Remark 4.40. Note that

$$\sigma(y) : \nabla p = \sigma(y) : \epsilon(p) \tag{4.110}$$

holds because $\sigma(y)$ is symmetric, see for instance [90]. From this we get

$$\sigma(y) : \nabla p = \sigma(p) : \nabla y \tag{4.111}$$

by transformations. We use (4.110)-(4.111) to specify the adjoint equation (4.108)-(4.109). Moreover, we require (4.111) to specify the shape derivative (4.107). Note that the right-hand side of (4.108) arises from the weak formulation of the linear elasticity equation (4.100)-(4.101), more precisely:

$$\begin{aligned} J(u) &= \int_{X(u)} \sigma(y) : \epsilon(y) \, dx = \int_{X(u)} f y \, dx \\ \Rightarrow \int_{X(u)} \frac{\partial}{\partial y} \left(\sigma(y) : \epsilon(y)\right) \, dx &= \int_{X(u)} f \, dx \end{aligned}$$

Chapter 5

Lagrange-Newton and quasi-Newton approach

Many real world problems can be formulated as shape optimization problems. In particular, shape optimization problems arise frequently in technological processes, which are modelled by PDEs as in [7, 8, 9, 30, 35, 79, 82, 83]. In a lot of practical circumstances, the shape under investigation is parametrized by a finite number of parameters, which, on the one hand, allows the application of standard optimization approaches, but, on the other hand, unnecessarily limits the space of reachable shapes. Shape calculus, which has been the subject of several monographs [23, 67, 93] and which has been a topic in the previous chapter, presents a way to circumvent this dilemma. However, so far optimization based on shape calculus is mainly performed using gradient descent methods, which can be shown to converge. The major difference between shape optimization and the standard PDE constrained optimization framework is the lack of a linear space structure in shape spaces. If one cannot work with linear spaces, then Riemannian manifolds are the next best option. Riemannian manifolds for shape spaces are discussed in [11, 12, 62, 63, 64] and in Chapter 3. The publication [84] links shape calculus with shape manifolds and, thus, enables the usage of optimization techniques on manifolds in the context of shape optimization. This chapter is devoted to the extension of the Riemannian geometrical point of view on shape optimization established in [84] to a Lagrange-Newton and a quasi-Newton approach for PDE constrained shape optimization. Generally, a Lagrange-Newton method is obtained by applying a Newton method to find stationary points of the Lagrangian. In contrast to this method, which requires the Hessian of the Lagrangian in each iteration, quasi-Newton methods only need an approximation in each iteration. Such an approximation is realized, for example, by a limited memory Broyden-Fletcher-Goldfarb-Shanno (BFGS) update.

This chapter is organized as follows. In Section 5.1, the above-mentioned Lagrange-Newton approach is formulated. It is based on optimization on Riemannian vector space bundles and exemplified for the elliptic shape interface optimization problem

introduced in Subsection 4.2.1. Section 5.2 presents a limited memory BFGS quasi-Newton technique in shape spaces and discusses the theoretical background from optimization on Riemannian manifolds. Finally, Section 5.3 gives numerical results for the interface problems introduced in Section 4.2.

5.1 Lagrange-Newton approach

This section is devoted to a Lagrange-Newton approach for PDE constrained shape optimization problems. Newton-type methods have been used in shape optimization for many years, e.g., [28, 72]. Here we specify them for the particular case of shape manifolds. Subsection 5.1.1 presents a vector bundle framework on which the Lagrange-Newton approach is based. This framework is based on the Riemannian framework established in [84], which enables the discussion of Lagrange-Newton methods within the shape calculus framework for PDE constrained shape optimization. Afterwards, in Subsection 5.1.2, this Riemannian vector bundle framework is applied to the elliptic shape interface optimization problem considered in Subsection 4.2.1. The numerical results for this application are presented in Subsection 5.3.1.

5.1.1 Riemannian vector bundle framework

Definition 2.19 of a vector bundle can be generalized to a Riemannian one. Instead of Banach spaces, we have to deal with Hilbert spaces. Let (M, g) be a Riemannian manifold of class $\mathcal{C}^q$ with $q \in \overline{\mathbb{N}}$ and suppose that (H, π, M) is a Riemannian vector bundle. This means that for each $u \in M$ the fiber $H(u) \coloneqq H_u$ is a Hilbert space, there is a bundle-projection $\pi\colon H \to M$ and for an open covering $\{U_i\}_i$ of M there is a local $\mathcal{C}^q$-isomorphism

$$\tau_i\colon \pi^{-1}(U_i) \to U_i \times H_0.$$

Moreover, each τ_i induces a linear isometric isomorphism

$$\tau_i(u)\colon H(u) \to H_0 \tag{5.1}$$

on each fiber $\pi^{-1}(u) = H(u)$ for all $u \in U_i$. The total space H of the vector bundle (H, π, M) is by itself a Riemannian manifold, where the tangent bundle TH satisfies $T_yH \cong H(u) \oplus T_uM$. Let

$$E \coloneqq \{(y, u)\colon y \in H(u), u \in M\}.$$

Moreover, let a_u define a bilinear form and let b_u be a linear form. We consider the following PDE constrained optimization problem in weak form:

$$\min_{(y,u)\in E} \; J(y, u)\,, \;\; J\colon E \to \mathbb{R} \tag{5.2}$$

$$\text{s.t. } a_u(y, p) = b_u(p) \quad \forall p \in H(u) \tag{5.3}$$

The scalar valued function J is assumed to be $\mathcal{C}^q$. Intentionally, the weak formulation of the PDE is chosen for ease of presentation. Now, we define the Lagrangian $\mathscr{L}$ in order to formulate the adjoint and design equation to the PDE constrained optimization problem (5.2)-(5.3).

Definition 5.1. *The Lagrangian corresponding to (5.2)-(5.3) is defined by*

$$\mathscr{L}\colon F \to \mathbb{R},\ (y,u,p) \mapsto J(y,u) + a_u(y,p) - b_u(p),$$

where $F \coloneqq \{(y,u,p)\colon y,p \in H(u), u \in M\}$ *with* $T_{(y,u,p)}F \cong H(u) \times T_uM \times H(u)$.

Let $(\hat{y},\hat{u}) \in E$ solve the optimization problem (5.2)-(5.3). Then the (adjoint) variational problem, which we obtain by differentiating $\mathscr{L}$ with respect to y, is given by

$$a_{\hat{u}}(z,p) = -\frac{\partial}{\partial y}\Big|_{y=\hat{y}} J(y,\hat{u})z \quad \forall z \in H(\hat{u}) \tag{5.4}$$

and the design problem, which we obtain by differentiating $\mathscr{L}$ with respect to u, is given by

$$\frac{\partial}{\partial u}\Big|_{u=\hat{u}} \left[J(\hat{y},u) + a_u(\hat{y},\hat{p}) - b_u(\hat{p})\right] w = 0 \quad \forall w \in T_{\hat{u}}M, \tag{5.5}$$

where $\hat{p} \in H(\hat{u})$ solves (5.4). The derivative on the left-hand side of (5.5) means the shape derivative $D\mathscr{L}(\hat{y},u,\hat{p})[w]\big|_{u=\hat{u}}$. If we differentiate $\mathscr{L}$ with respect to p, we obtain the state equation (5.3). The conditions (5.3)-(5.5) can be collected in the following condition:

$$d\mathscr{L}(\hat{y},\hat{u},\hat{p})[h] = 0 \quad \forall h \in T_{(\hat{y},\hat{u},\hat{p})}F, \tag{5.6}$$

where $d\mathscr{L}(\hat{y},\hat{u},\hat{p})[h]$ denotes the directional derivative of $\mathscr{L}$ at $(\hat{y},\hat{u},\hat{p})$ in direction h. For $h = (h_y,h_u,h_p)$ with $\|h\|_{T_{(y,u,p)}F} = 1$, it is given by

$$d\mathscr{L}(\hat{y},\hat{u},\hat{p})[h] = d_y\mathscr{L}(\hat{y},\hat{u},\hat{p})h_y + d_u\mathscr{L}(\hat{y},\hat{u},\hat{p})h_u + d_p\mathscr{L}(\hat{y},\hat{u},\hat{p})h_p, \tag{5.7}$$

where $d_y = \frac{\partial}{\partial y}$, $d_u = \frac{\partial}{\partial u}$, $d_p = \frac{\partial}{\partial p}$ and $d_u\mathscr{L}(\hat{y},\hat{u},\hat{p})h_u$ is another notation for the shape derivative $D\mathscr{L}(\hat{y},\hat{u},\hat{p})[h_u]$.

Remark 5.2. The condition (5.6) or, equivalently, the conditions (5.3)-(5.5) are the well-known *Karush-Kuhn-Tucker (KKT) conditions.*

By using a Riemannian metric $g = (g_u)_{u\in M}$ on M and a smoothly varying scalar product $\langle\cdot,\cdot\rangle_u$ on the Hilbert space $H(u)$, we can envision $T_{(y,u,p)}F$ as a Hilbert space with the canonical scalar product

$$\left\langle \begin{pmatrix} z_1 \\ w_1 \\ q_1 \end{pmatrix}, \begin{pmatrix} z_2 \\ w_2 \\ q_2 \end{pmatrix} \right\rangle_{T_{(y,u,p)}F} = \langle z_1,z_2\rangle_u + g_u(w_1,w_2) + \langle q_1,q_2\rangle_u . \tag{5.8}$$

This scalar product can be used to find a representation of the Riemannian gradient

$$\mathrm{grad}\mathscr{L}(y,u,p) = \begin{pmatrix} \mathrm{grad}_y\mathscr{L}(y,u,p) \\ \mathrm{grad}_u\mathscr{L}(y,u,p) \\ \mathrm{grad}_p\mathscr{L}(y,u,p) \end{pmatrix} \in T_{(y,u,p)}F$$

by the condition

$$\langle \mathrm{grad}\mathscr{L}(y,u,p), h\rangle_{T_{(y,u,p)}F} = d\mathscr{L}(y,u,p)[h] \quad \forall h \in T_{(y,u,p)}F.$$

Now, similarly to standard non-linear programming, we can solve the problem of finding $(y,u,p) \in F$ with

$$\mathrm{grad}\mathscr{L}(y,u,p) = 0 \tag{5.9}$$

in order to find solutions of the optimization problem (5.2)-(5.3). The non-linear problem (5.9) has exactly the form of the root finding problems discussed in [84]. Exploiting the Riemannian structure on TF, we can formulate a Newton iteration. Let $\mathrm{Hess}\mathscr{L}$ denote the Riemannian Hessian, which is based on the resulting Riemannian connection $\nabla^{\mathrm{cov}}\colon \Gamma(TF) \times \Gamma(TF) \to \Gamma(TF)$. Moreover, let $\mathcal{R}$ denote a retraction map on F. With this notation the Newton iteration is formulated in Algorithm 1.

Algorithm 1. k-th Newton iteration.

(1) Compute the increment $\Delta\xi$ as solution of

$$\mathrm{Hess}\mathscr{L}(\xi^k)\Delta\xi = -\mathrm{grad}\mathscr{L}(\xi^k). \tag{5.10}$$

(2) Compute the increment $\xi^{k+1} = \mathcal{R}_{\xi^k}(\alpha^k \cdot \Delta\xi)$ for some steplength α^k.

In the following, we have to specify both, the Riemannian Hessian and the retraction map. However, we first have to specify the scalar product on the involved Hilbert space. We want to choose a metric that is as simple as possible in the Hilbert space parts. Therefore, we use a metric defined on the Hilbert space $(H_0, \langle\cdot,\cdot\rangle_0)$ and transfer this canonically to the Hilbert spaces $H(u)$. Thus, in the following, we assume that we have to deal with one particular chart (U_i, τ_i) of the covering $\{U_i\}_i$ only, where $\tau_i\colon \pi^{-1}(U_i) \to H_0 \times U_i$, and define

$$\langle z_1, z_2\rangle_u \coloneqq \langle \tau_i(u)z_1, \tau_i(u)z_2\rangle_0 \quad \forall u \in U_i\,.$$

Now, we devote to the retraction $\mathcal{R}$ on F in Algorithm 1. Retractions can be represented by geodesics, which are given by the exponential map due to Theorem 2.33. In the Hilbert space parts of F, they are represented by straight lines in H_0. Thus,

if $\exp^M$ denotes the exponential map on the manifold M, the exponential map on F can be expressed in the form

$$\begin{aligned}&\exp_{(y,u,p)}(z,w,q)\\&= \left(\left(\tau_i\left(\exp_u^M(w)\right)^{-1}\circ\tau_i(u)\right)(y+z), \exp_u^M(w), \left(\tau_i\left(\exp_u^M(w)\right)^{-1}\circ\tau_i(u)\right)(p+q)\right).\end{aligned}$$

Within iteration (5.10) in Algorithm 1, the Riemannian Hessian has to be discussed. It is based on the Riemannian connection $\nabla^{\mathrm{cov}}\colon \Gamma(TF)\times\Gamma(TF)\to\Gamma(TF)$. Let ∇^M denote the Riemannian covariant derivative on $(TM,\widetilde{\pi},M)$, where $\widetilde{\pi}$ is the natural projection (cf. Subsection 2.3.1). Due to the observation that mixed covariant derivatives of vectors from H with respect to tangential vectors in TM are reduced to simple directional derivatives, which is the case for derivatives in linear spaces anyway, for $(h_y,h_u,h_p)\in T_{(y,u,p)}F$ the covariant derivative is given as follows:

$$\begin{aligned}\nabla^{\mathrm{cov}}_{(h_y,h_u,h_p)}\colon T_{(y,u,p)}F &\to T_{(y,u,p)}F,\\ \begin{pmatrix}\overline{h_y}\\ \overline{h_u}\\ \overline{h_p}\end{pmatrix} &\mapsto \begin{pmatrix} d_y\overline{h_y}\,h_y + d_u\overline{h_y}\,h_u + d_p\overline{h_y}\,h_p\\ d_y\overline{h_u}\,h_y + \nabla^M_{h_u}\overline{h_u} + d_p\overline{h_u}\,h_p\\ d_y\overline{h_p}\,h_y + d_u\overline{h_p}\,h_u + d_p\overline{h_p}\,h_p\end{pmatrix}\end{aligned} \tag{5.11}$$

In our vector bundle setting, the Riemannian Hessian of the Lagrangian is defined by the mapping

$$T_{(y,u,p)}F\to T_{(y,u,p)}F,\ l\mapsto \mathrm{Hess}\mathscr{L}[l]\coloneqq\nabla^{\mathrm{cov}}_l\mathrm{grad}\mathscr{L}. \tag{5.12}$$

From (5.11)-(5.12) we conclude the following block structure of the Hessian:

$$\mathrm{Hess}\mathscr{L}=\begin{pmatrix} d_y\mathrm{grad}_y\mathscr{L} & d_u\mathrm{grad}_y\mathscr{L} & d_p\mathrm{grad}_y\mathscr{L}\\ d_y\mathrm{grad}_u\mathscr{L} & \nabla^M\mathrm{grad}_u\mathscr{L} & d_p\mathrm{grad}_u\mathscr{L}\\ d_y\mathrm{grad}_p\mathscr{L} & d_u\mathrm{grad}_p\mathscr{L} & 0\end{pmatrix}. \tag{5.13}$$

Note the difference of the Hessian in the vector bundle setting (cf. (5.12)) and the Riemannian shape Hessian defined on a Riemannian manifold (M,g). In the case $(M,g)=(B_e^0,g^1)$ or $(M,g)=(B_e,g^1)$, the Riemannian shape Hessian is given by (4.86). In the following, we denote the Riemannian shape Hessian defined on M by Hess^M to avoid confusion. The representation of the covariant derivative ∇^{B_e} with respect to the Sobolev metric g^1 is given in Theorem 3.3.

Let $\triangle\xi=(z,w,q)^T$ in (5.10). Then, in order to solve equation (5.10) in Algorithm 1, the following equations have to be satisfied for all $(\overline{z},\overline{w},\overline{q})^T\in T_{(y,u,p)}F$:

$$H_{11}(z,\overline{z})+H_{12}(w,\overline{z})+H_{13}(q,\overline{z})=-a_u(\overline{z},p)-\frac{\partial}{\partial y}J(y,u)\overline{z} \tag{5.14}$$

$$H_{21}(z,\overline{w})+H_{22}(w,\overline{w})+H_{23}(q,\overline{w})=-\frac{\partial}{\partial u}\left[J(y,u)+a_u(y,p)-b_u(p)\right]\overline{w} \tag{5.15}$$

$$H_{31}(z,\overline{q})+H_{32}(w,\overline{q})+H_{33}(q,\overline{q})=-a_u(y,\overline{q})+b_u(\overline{q}) \tag{5.16}$$

In (5.14)-(5.16), the terms H_{ij}, where $i, j \in \{1, 2, 3\}$, are given by

$$
\begin{aligned}
H_{11}(z, \overline{z}) &= \frac{\partial^2}{\partial y^2} J(y, u) z \overline{z}, \\
H_{12}(w, \overline{z}) &= \frac{\partial}{\partial u} \left[a_u(\overline{z}, p) + \frac{\partial}{\partial y} J(y, u) \overline{z} \right] w, \\
H_{13}(q, \overline{z}) &= a_u(\overline{z}, q), \\
H_{21}(z, \overline{w}) &= \frac{\partial}{\partial y} \frac{\partial}{\partial u} \left(\left[J(y, u) + a_u(y, p) \right] \overline{w} \right) z, \\
H_{22}(w, \overline{w}) &= g \left(\mathrm{Hess}^M \mathscr{L}(y, u, p)[w], \overline{w} \right), \\
H_{23}(q, \overline{w}) &= \frac{\partial}{\partial p} \frac{\partial}{\partial u} \left(\left[a_u(y, p) - b_u(p) \right] \overline{w} \right) q, \\
H_{31}(z, \overline{q}) &= a_u(z, \overline{q}), \\
H_{32}(w, \overline{q}) &= \frac{\partial}{\partial u} \left[a_u(y, \overline{q}) - b_u(\overline{q}) \right] w, \\
H_{33}(q, \overline{q}) &= 0.
\end{aligned}
$$

The covariant derivative ∇^{cov} on F reveals natural symmetry properties and, thus, obvious symmetries can be observed in the above components not involving second shape derivatives. A key observation in [84] is that even the expression H_{22} is symmetric in the solution of the shape optimization problem. This motivates a shape sequential quadratic programming (SQP) method as outlined below, where away from the solution only expressions in H_{22} are used which are non-zero at the solution. Its basis is the following observation:

If the term H_{22} is replaced by an approximation $\hat{H}_{22}$ such that

(i) the approximation $\hat{H}_{22}$ omits all terms in H_{22} which are zero at the solution and

(ii) the reduced Hessian of (5.13) built with this approximation is coercive,

then equation (5.10) is equivalent to the following linear quadratic problem (QP):

$$
\min_{(z,w)} \frac{1}{2} \Big(H_{11}(z, z) + 2 H_{12}(w, z) + \hat{H}_{22}(w, w) \Big) + \frac{\partial}{\partial y} J(y, u) z + \frac{\partial}{\partial u} J(y, u) w \tag{5.17}
$$

$$
\text{s.t. } a_u(z, \overline{q}) + \frac{\partial}{\partial u} \left[a_u(y, \overline{q}) - b_u(\overline{q}) \right] w = -a_u(y, \overline{q}) + b_u(\overline{q}) \quad \forall \overline{q} \in H(u) \tag{5.18}
$$

In the next subsections, we also omit terms in H_{21} and H_{12} which are equal to zero when evaluated at the solution of the optimization problem. Nevertheless, quadratic convergence of the resulting SQP method is to be expected and indeed observed in Subsection 5.3.1.

5.1.2 Application of the Riemannian vector bundle framework

In this section, we apply the theoretical discussion of the previous section to the elliptic PDE constrained shape interface optimization problem (4.16)-(4.20) discussed in Subsection 4.2.1. Consequently, we consider shapes $u \in B_e^0$, where the space B_e^0 is defined in (4.15). For our discussions, among the various Riemannian metrics mentioned in Section 3.2, we pick the first Sobolev metric g^1.

Now, F defined in Definition 5.1 is given by (4.25). The metric in the vector space parts is constructed by employing a mesh deformation. Mesh deformations are often used to deform a computational mesh smoothly in accordance with a boundary deformation of the computational domain. In our setting, we deform the computational domain rather than only the mesh. Moreover, we assume that there is a bijective $\mathcal{C}^\infty$-mapping $\Phi_u \colon (0,1)^2 \to \Omega(u)$, e.g., Φ_u is the deformation given by the solution of a linear elasticity equation. Thus, we can construct the linear isomorphism (5.1) by the bijective identification $\tau(u) \colon H_0^1(\Omega(u)) \to H_0^1\big((0,1)^2\big)$, $g \mapsto g \circ \Phi_u$.

We have to detail the expressions in equation (5.10). The Lagrangian $\mathscr{L}$ is defined in (4.26). The shape derivative of $\mathscr{L}$ in the direction of a continuous vector field V is given in Theorem 4.25. We now focus on the weak formulation (5.14)-(5.16). In the case of (4.16)-(4.20), we observe the following expressions for the right-hand sides:

$$-a_u(\overline{z}, p) - \frac{\partial}{\partial y} J(y,u)\overline{z} = -\int_{\Omega(u)} \nabla \overline{z}^T \nabla p + (y - \overline{y})\overline{z}\, dx \tag{5.19}$$

$$-\frac{\partial}{\partial u}\left[J(y,u) + a_u(y,p) - b_u(p)\right]\overline{w} = \int_u \left(\llbracket f \rrbracket\, p - \mu\kappa\right) \langle \overline{w}, n\rangle\, ds \tag{5.20}$$

$$-a_u(y, \overline{q}) + b_u(\overline{q}) = \int_{\Omega(u)} -\nabla y^T \nabla \overline{q} + f\overline{q}\, dx \tag{5.21}$$

These expressions are set to zero in order to define the necessary optimality conditions. In the following, we discuss more details about the Hessian. First, we consider H_{ij}, where $i, j \in \{1,2,3\}$, without the term H_{22}, which requires special care. The other terms can be expressed at the solution $(y,u,p) \in F$ of the optimization problem (4.16)-(4.20) for all $h = (\overline{z}, \overline{w}, \overline{q})^T \in T_{(y,u,p)}F$ as

$$\begin{aligned}
H_{11}(z,\overline{z}) &= \frac{\partial^2}{\partial y^2} J(y,u) z\overline{z} = \int_{\Omega(u)} z\overline{z}\, dx,\\
H_{12}(w,\overline{z}) &= \frac{\partial}{\partial u}\left[a_u(\overline{z}, p) + \frac{\partial}{\partial y} J(y,u)\overline{z}\right] w = 0,\\
H_{13}(q,\overline{z}) &= a_u(\overline{z}, q) = \int_{\Omega(u)} \nabla \overline{z}^T \nabla q\, dx,\\
H_{21}(z,\overline{w}) &= \frac{\partial}{\partial y}\frac{\partial}{\partial u}\left(\left[J(y,u) + a_u(y,p)\right]\overline{w}\right) z = 0,\\
H_{23}(q,\overline{w}) &= \frac{\partial}{\partial p}\frac{\partial}{\partial u}\left(\left[a_u(y,p) - b_u(p)\right]\overline{w}\right) q = -\int_u \llbracket f \rrbracket\, q\, \langle \overline{w}, n\rangle\, ds,
\end{aligned}$$

$$
\begin{aligned}
H_{31}(z,\overline{q}) &= a_u(z,\overline{q}) = \int_{\Omega(u)} \nabla z^T \nabla \overline{q}\, dx,\\
H_{32}(w,\overline{q}) &= \frac{\partial}{\partial u}\left[a_u(y,\overline{q}) - b_u(\overline{q})\right] w = -\int_u [\![f]\!]\, \overline{q}\, \langle w,n\rangle\, ds,\\
H_{33}(q,\overline{q}) &= 0.
\end{aligned}
$$

Now, we consider the term H_{22}. We evaluate it at the solution of the optimization problem. Arguments identical to those in the proof of Proposition 5.5.2 in [1] can be used to state that the Riemannian shape Hessian based on the metric g^1 on B_e^0 satisfies the relation

$$
g^1\left(\mathrm{Hess}^{B_e^0}\mathscr{L}(y,u,p)[w],\overline{w}\right) = D^2\mathscr{L}(y,u,p)[w,\overline{w}] - D\mathscr{L}(y,u,p)\left[\nabla_w^{B_e^0}\overline{w}\right]. \quad (5.22)
$$

Since H_{22} is evaluated at the solution of the optimization problem,

$$
D\mathscr{L}(y,u,p)\left[\nabla_w^{B_e^0}\overline{w}\right] = 0 \quad (5.23)
$$

holds. In Subsection 5.3.1, the solution of (4.16)-(4.20) is a straight line connection of the points $(0.5,0)$, $(0.5,1)$. In this special case, the curvature is equal to zero and, thus, the necessary optimality condition (5.20) reduces to

$$
-\frac{\partial}{\partial u}\left[J(y,u) + a_u(y,p) - b_u(p)\right]\overline{w} = \int_u [\![f]\!]\, p\, \langle \overline{w},n\rangle\, ds. \quad (5.24)
$$

Combining (5.22)-(5.24) with Theorem 4.14 and Theorem 4.15 gives

$$
\begin{aligned}
&H_{22}(w,\overline{w})\\
&= g^1\left(\mathrm{Hess}^{B_e^0}\left(J(y,u) + a_u(y,p) - b_u(p)\right) w,\overline{w}\right)\\
&= \int_u -D\left([\![f]\!]\, p\right)[\overline{w}]\, \langle w,n\rangle - [\![f]\!]\left(\kappa p + \frac{\partial p}{\partial n}\right)\langle \overline{w},n\rangle\, \langle w,n\rangle\\
&\qquad + \mu \frac{\partial w}{\partial \tau}\frac{\partial \overline{w}}{\partial \tau}\, \langle \overline{w},n\rangle\, \langle w,n\rangle\, ds,
\end{aligned} \quad (5.25)
$$

where $\frac{\partial}{\partial \tau}$ denotes the derivative tangential to u. In our setting, we observe

$$
p = 0 \quad \text{in } \Omega(u) \quad (5.26)
$$

in the solution of the tracking-type optimization problem (4.16)-(4.20). Thus, in the solution, all derivatives of p are equal to zero. Consequently,

$$
D\left([\![f]\!]\, p\right)[\overline{w}] = [\![f]\!]\, Dp[\overline{w}] = 0 \quad \text{and} \quad \frac{\partial p}{\partial n} = 0 \quad \text{on } u. \quad (5.27)
$$

Due to (5.25)-(5.27), H_{22} reduces to

$$
\hat{H}_{22}(w,\overline{w}) = \int_u \left(\mu \frac{\partial w}{\partial \tau}\frac{\partial \overline{w}}{\partial \tau} - [\![f]\!]\, \kappa p\right)\langle w,n\rangle\, \langle \overline{w},n\rangle\, ds. \quad (5.28)
$$

By using the expressions above, we can formulate the QP (5.17)-(5.18) at the solution in the following form:

$$\min_{(z,w)} \mathcal{F}(z,w,y,p) \tag{5.29}$$

$$\begin{aligned} \text{s.t.} \quad & \int_{\Omega(u)} \nabla z^T \nabla \overline{q}\, dx - \int_u [\![f]\!]\, \overline{q} w\, ds \\ & = -\int_{\Omega(u)} \nabla y^T \nabla \overline{q}\, dx + \int_{\Omega(u)} f \overline{q}\, dx \quad \forall \overline{q} \in H_0^1(\Omega(u)) \end{aligned} \tag{5.30}$$

with the objective function

$$\begin{aligned} \mathcal{F}(z,w,y,p) = & \int_{\Omega(u)} \frac{z^2}{2} + (y-\overline{y})z\, dx + \int_u \mu\kappa w - [\![f]\!]\, pw\, ds \\ & + \frac{1}{2}\int_u \mu \left(\frac{\partial w}{\partial \tau}\right)^2 - [\![f]\!]\, \kappa p w^2\, ds. \end{aligned} \tag{5.31}$$

This QP is given in weak form, but it can be rewritten in the more intelligible strong form of an optimal control problem as follows:

$$\min_{(z,w)} \mathcal{F}(z,w,y,p) \tag{5.32}$$

$$\text{s.t.} \quad -\triangle z = \triangle y + f \quad \text{in } \Omega(u) \tag{5.33}$$

$$z = 0 \quad \text{on } \partial\Omega(u) \tag{5.34}$$

$$\left[\!\left[\frac{\partial z}{\partial n}\right]\!\right] = [\![f]\!]\, w \quad \text{on } u \tag{5.35}$$

Note that the interface condition (5.35) can also be expressed as

$$\frac{\partial z}{\partial n} = f_1 w \quad \text{and} \quad -\frac{\partial z}{\partial n} = f_2 w \quad \text{on } u.$$

The adjoint problem to the optimal control problem (5.32)-(5.35) is the following boundary value problem:

$$-\triangle q = -z - (y - \overline{y}) \quad \text{in } \Omega(u) \tag{5.36}$$

$$q = 0 \quad \text{on } \partial\Omega(u) \tag{5.37}$$

The resulting design equation for the optimal control problem (5.32)-(5.35) is given by

$$0 = -[\![f]\!]\,(p + \kappa p w + q) + \mu\kappa - \mu \frac{\partial^2 w}{\partial \tau^2} \quad \text{on } u. \tag{5.38}$$

5.2 Quasi-Newton approach

This section is devoted to a quasi-Newton approach for PDE constrained shape optimization problems. Quasi-Newton methods on general manifolds have already been discussed in [1, 33, 77]. Here we specify them for the particular case of shape manifolds. From a different standpoint, the discussion in this section can be viewed as a generalization of the elliptic structured inverse modelling in the publications [46, 73] to the parabolic case. The methodology and algorithm established in this section apply, for example, to the problem of inversely determining cell shapes in the human skin as investigated in [68].

As already mentioned in Section 4.3, shape optimization can be viewed as optimization on Riemannian shape manifolds and the resulting optimization methods can be constructed and analyzed within this framework. We consider connected and compact subsets $X_2 \subset \mathbb{R}^2$ with $X_2 \neq \emptyset$ and variable $\mathcal{C}^\infty$-boundary ∂X_2 (cf. Figure 4.1). The boundary ∂X_2 can be identified with a simple closed curve $c\colon S^1 \to \mathbb{R}^2$ and is an element of the shape space B_e defined in (3.1). In this thesis, among all Riemannian metrics mentioned in Section 3.2, we pick the Sobolev metric g^1 given in (3.6). With the shape space (B_e, g^1) and its tangent space T_cB_e, which is given in (3.2), we can state an expression of the Riemannian shape gradient corresponding to a shape derivative. In this section, we mean always the Riemannian shape gradient with respect to (B_e, g^1) by $\operatorname{grad}(\cdot)$. In our setting, the Riemannian shape gradient is given in Section 4.3, more precisely in (4.85).

In the sequel, we need the concept of the covariant derivative corresponding to the Sobolev metric g^1. This covariant derivative is given in Theorem 3.3. Moreover, we need the exponential map $\exp_c\colon\ T_cB_e \to B_e$, $h \mapsto \exp_c(h)$ which defines a local diffeomorphism between the tangent space and the manifold by following the locally uniquely defined geodesic $\exp_c(h)$ starting in $c \in B_e$ with velocity $h \in T_cB_e$.

Recall that a Lagrange-Newton method is obtained by applying a Newton method on the KKT-conditions of a PDE constrained optimization problem. Its application is based on the tangent condition. In contrast to this method, the application of a quasi-Newton method is based on the secant condition. The secant condition is formulated on the Riemannian manifold B_e analogously to [1] for a step $c_{j+1} := \mathcal{R}_{c_j}(\eta_j)$ resulting from an increment $\eta_j \in T_{c_j}B_e$ in iteration j via a retraction $\mathcal{R}$ as

$$\operatorname{grad} J(c_{j+1}) - \mathcal{T}_{\eta_j} \operatorname{grad} J(c_j) = G_{j+1}[\mathcal{T}_{\eta_j}\eta_j].$$

Here $J\colon B_e \to \mathbb{R}$ denotes the objective function of a PDE constrained shape optimization problem, G_{j+1} is intended to approximate the Riemannian shape Hessian $\operatorname{Hess}^{B_e} J(c_{j+1}) = \nabla^{B_e} \operatorname{grad} J(c_{j+1})$ and

$$\mathcal{T}\colon TB_e \oplus TB_e \to TB_e,\ (h_c, k_c) \mapsto \mathcal{T}_{h_c}k_c$$

denotes a vector transport associated with the chosen retraction $\mathcal{R}$ (cf. Definition 2.35). Note that, in a quasi-Newton method, we need only an approximation of the Hessian. Such an approximation is realized, for example, by a limited memory

Algorithm 2. Inverse limited memory BFGS update in (B_e, g^1).

$\rho_j \leftarrow g^1(y_j, s_j)^{-1}$
$q \leftarrow \mathrm{grad}J(c_j)$
for $i = j-1, \ldots, j-m$ **do**
  $s_i \leftarrow \mathcal{T}_q s_i$
  $y_i \leftarrow \mathcal{T}_q y_i$
  $\alpha_i \leftarrow \rho_i g^1(s_i, q)$
  $q \leftarrow q - \alpha_i y_i$
end for
$q \leftarrow \frac{g^1(y_{j-1}, s_{j-1})}{g^1(y_{j-1}, y_{j-1})} q$
for $i = j-m, \ldots, j-1$ **do**
  $\beta_i \leftarrow \rho_i g^1(y_i, z)$
  $q \leftarrow q + (\alpha_i - \beta_i) s_i$
end for
return q

BFGS update technique. Moreover, note that we want to approximate the Riemannian shape Hessian Hess^{B_e} by such a technique. In order to formulate a limited memory BFGS update in a concise way, we have to introduce the following notation for a typical linear operator associated with the Sobolev metric:

$$h \otimes k \colon T_c B_e \to T_c B_e,\ v \mapsto g^1(k, v)h,$$

where $h, k \in T_c B_e$. With this notation and the abbreviations

$$s_j = \mathcal{T}_{\eta_j} \eta_j \in T_{c_{j+1}} B_e, \tag{5.39}$$
$$y_j = \mathrm{grad}J(c_{j+1}) - \mathcal{T}_{\eta_j} \mathrm{grad}J(c_j) \in T_{c_{j+1}} B_e \tag{5.40}$$

we can rephrase the BFGS update on the shape manifold B_e endowed with the Riemannian metric g^1 as

$$G_{j+1} = \widetilde{G}_j - \frac{(\widetilde{G}_j s_j) \otimes (\widetilde{G}_j s_j)}{g^1(s_j, \widetilde{G}_j s_j)} + \frac{y_j \otimes y_j}{g^1(s_j, y_j)},$$

where $\widetilde{G}_j = \mathcal{T}_{\eta_j} \circ G_j \circ \mathcal{T}_{\eta_j}^{-1}$. In [77], superlinear convergence properties for BFGS quasi-Newton methods on manifolds are analyzed for the case that $\mathcal{T}_{\eta_j}$ is an isometry. This requirement is satisfied if $\mathcal{T}$ and $\mathcal{R}$ are, for example, the parallel transport and the exponential map. The corresponding update of the inverse operator can be written in the form

$$G_{j+1}^{-1} = \left(I - \frac{s_j \otimes y_j}{g^1(y_j, s_j)}\right) \widetilde{G}_j^{-1} \left(I - \frac{y_j \otimes s_j}{g^1(y_j, s_j)}\right) + \frac{s_j \otimes s_j}{g^1(y_j, s_j)},$$

see for instance [69]. This is the most convenient update formulation in an infinite dimensional setting. In standard formulations, update formulas require the storage

of the whole convergence history up to the current iteration. Limited memory update techniques have been developed in order to reduce the amount of storage (cf. [69]). In the current situation, an inverse one is formulated in Algorithm 2, whose output is $q = G_j^{-1}\text{grad}J(c_j)$. Algorithm 2 is conceptually similar to the double loop algorithm in finite dimensional Euclidean spaces. Yet, the inner products are now given by the Sobolev metric and vector transports are considered. Moreover, in the j-th iteration, s_j given in (5.39) denotes the distance between two iterated shapes and y_j given in (5.40) denotes the difference of iterated Riemannian shape gradients. In Subsection 5.3.2, the BFGS Algorithm 2 is tested for the parabolic shape interface optimization problem defined in Subsection 4.2.2. For this problem, the shape derivative and the (standard) shape gradient can be found in Theorem 4.32. Moreover, in the Riemannian setting (B_e, g^1), the Riemannian shape gradient is given by (4.85).

5.3 Numerical results

This section discusses numerical results for the Lagrange-Newton approach (Subsection 5.3.1) and the quasi-Newton approach (Subsection 5.3.2) established in the previous sections.

5.3.1 Lagrange-Newton approach

We solve the optimal control problem (5.32)-(5.35) by employing a conjugate gradient iteration (CG-iteration) for the reduced problem (5.38), i.e., we iterate over the variable w. In each step, the CG-iteration needs a residual of equation (5.38) from w^k. First, we compute the state variable z^k from the state equation (5.33)-(5.35). Afterwards, we compute the adjoint variable q^k from the adjoint equation (5.36)-(5.37). This enables us to evaluate the residual

$$r^k = -\left[\!\left[f\right]\!\right]\left(p + \kappa p w^k + q^k\right) + \mu\kappa - \mu\frac{\partial^2 w^k}{\partial \tau^2} \tag{5.41}$$

from the design equation (5.38). In this way, we create an iterative solution technique very similar to SQP techniques known from linear spaces.

The particular values for the parameters are chosen as $f_1 = 1\,000$ and $f_2 = 1$. Moreover, the regularization parameter is $\mu = 10$. The data $\overline{y}$ are generated from a solution of the state equation (4.17)-(4.18) with u being the straight line connection of the points $(0.5, 0)$, $(0.5, 1)$. The starting point of the iterations is described by a B-spline defined by the two control points $(0.6, 0.7)$, $(0.4, 0.3)$.

Similar to [84], the retraction chosen for the shape is just the addition of $w^k n$ to the current shape. We build a coarse unstructured triangular grid Ω_h^1 with roughly 6 000 triangles as shown in the left picture of Figure 5.1. We also perform computations on uniformly refined grids Ω_h^2 and Ω_h^3 with roughly 24 000 and 98 000 triangles. In

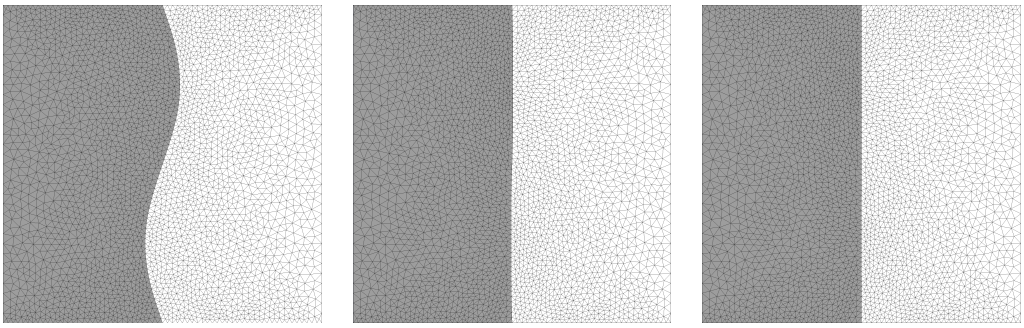


Figure 5.1: Iterations 0, 1 and 2 (left to right) together with deformations of the coarsest mesh Ω_h^1.

each iteration, the volume mesh is deformed according to the elasticity equation. In Figure 5.1, there are also shown the next two iterations on the coarsest grid.

In each iteration, the distance of a shape u^k to the optimal solution u^* is approximated by

$$\operatorname{dist}(u^k, u^*) = \int_{u^*} \left| \left\langle u^k, e_1 \right\rangle - \frac{1}{2} \right| ds,$$

where $e_1 = (1,0)^T$ denotes the first unit vector. The following table, which gives these distances, demonstrates that indeed quadratic convergence can be observed on the finest mesh, but also that the mesh resolution has a strong influence on the convergence properties:

It.-No.	Ω_h^1	Ω_h^2	Ω_h^3
0	0.0705945	0.070637	0.0706476
1	0.0043115	0.004104	0.0040465
2	0.0003941	0.000104	0.0000645

A standard shape calculus steepest method is based on the (reduced) shape derivative given in Theorem 4.25. The major advantage of the Newton method over such a standard method is the natural scaling of the step. When first experimenting with a steepest descent method, we found by trial and error that one needs a scaling around 10 000 in order to obtain sufficient progress.

5.3.2 Quasi-Newton approach

We test the algorithms developed in Section 5.2 with the problem (4.50)-(4.56) in the domain $X = (-1,1)^2$. This domain contains a compact and closed subset X_2 with smooth boundary Γ_{int}. The parameter valid in the exterior $X_1 = X \setminus X_2$ is chosen to be $k_1 = 1$, the parameter valid in the interior X_2 is chosen to be $k_2 = 0.001$ and the regularization parameter is chosen to be $\mu = 0.0001$. The final time of the

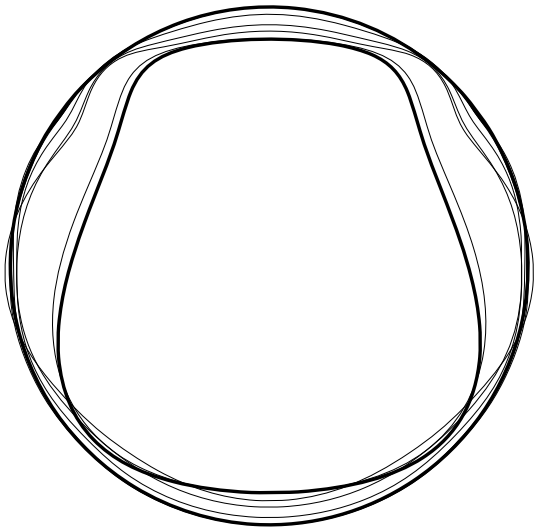


Figure 5.2: Initial and final shape geometry together with the iterations of the BFGS method in B_e.

simulation is $T = 20$. For our test case, the initial condition is $y_0(x) = 0$ for all $x \in \Omega$. Furthermore, we set $f(x,t) = 0$ in $(x,t) \in \Omega \times (0,T]$. First, we build artificial data $\overline{y}$ by solving the state equation (4.51)-(4.56) for the setting $\widetilde{X}_2 = \{x\colon\ |x| < r\}$ with $r = 0.5$. Afterwards, we choose another domain X_2. Figure 5.2 illustrates the interior boundary Γ_{int} around the initial domain X_2 and the target domain $\widetilde{X}_2$.

Remark 5.3. We obtain $\overline{y} \in L^2(0,T;H^1(X))$ as assumed in Subsection 4.2.2 by choosing the measurements $\overline{y}$ as the solution of model equation (4.51)-(4.56) .

In order to solve the boundary value problem (4.51)-(4.56), its weak form (4.57)-(4.61) is discretized in space using standard linear finite elements. The parameter k is approximated in an element-wise constant space. Furthermore, we choose the implicit Euler method for the temporal discretization. The interval $[0,T]$ is divided into 30 equidistantly distributed time steps. We can solve the adjoint equation (4.65)-(4.68) by applying the same spatial and temporal discretization as for the primal one. Finally, the resulting linear systems are solved using the conjugate gradient method.

An essential part of Algorithm 2 is a discrete version of the Laplace-Beltrami operator, which, on the one hand, is used to get a feasible representation of the shape gradient and, on the other hand, is needed for the scalar products in the BFGS method. Therefore, we implement the formulas given in [61], which describe an operator that can be used as the Laplace-Beltrami operator as well as to compute the discrete mean curvature. However, this approach is tailored to two dimensional triangulated surfaces. Thus, we have to extend the polygonal line in our test case in the third coordinate direction such that a surface is spanned and then triangulated.

We investigate the convergence behaviour of the steepest descent method and the limited memory BFGS quasi-Newton method. We choose a constant metric parameter $A = 0.001$ and a fixed step-size 1 in both optimization strategies.

Remark 5.4. We do not apply a line search strategy because of the computational cost. Each descent test in the line search requires the solution of the parabolic PDE in time and, in addition, the computation of the mesh deformation, which includes also a PDE. Since the resulting step lengths in both, the gradient and BFGS method, are feasible for this particular setting, a line search is not obligatory.

The necessary operations between the tangent spaces and the manifold are chosen essentially as the identity operator, i.e., for $\eta \in T_c B_e$ we define

$$\begin{aligned} \mathcal{R}_c(\eta)(s) &:= s + \eta(s) \quad \forall s \in c, \\ \mathcal{T}_\eta v(s) &:= v(s - \eta(s)) \quad \forall s \in \mathcal{R}_c(\eta). \end{aligned}$$

This setting corresponds to one explicit Euler step for the exponential map and the parallel transport in the case of the choice $A = 0$ in the metric g^1. From the point of view of implementation, this is the most convenient choice. Computing an explicit Euler step for the exponential map and a parallel transport for $A > 0$ requires the solution of yet another solution of an elliptic equation on the surface. However, numerical experiments have shown that the convergence properties of the resulting iterations are not changed and, thus, the additional numerical effort does not pay off in comparison with the inexpensive retraction above.

A main problem, which arises in the discrete case using linear finite elements, is that both, the representation of the shape gradient, which is determined by (4.76) or (4.77), and the normal vector field, are discontinuous across element interfaces. Thus, they cannot be applied directly as a deformation to the shape. Therefore, for all linear test-functions v on Γ_{int}, we solve the following L^2-projection to obtain a representation in piece-wise linear basis functions:

$$\int_{\Gamma_{\text{int}}} uv \, ds = \int_{\Gamma_{\text{int}}} \left(\int_0^T [\![k]\!] \, \nabla y_1^T \nabla p_2 \, dt \right) nv \, ds. \tag{5.42}$$

Then the resulting element-wise linear function u can be applied as a Dirichlet boundary condition in a linear elasticity equation. A second Dirichlet condition is chosen to be zero on the outer boundary of X such that the domain keeps its outer shape. Solving this PDE gives a deformation field, which can be evaluated in each mesh node, and a triangulation of the optimized shape without the need of re-meshing the domain X.

As specified in [11], the measurements of convergence rates ideally have to be performed in terms of the geodesic distance defined in (2.20). However, this requires the computation of the full geodesic connecting the current iterate with the solution, which is a highly expensive operation. Because of the local rigidity condition of retractions (cf. Definition 2.34), an approximation of the geodesic distance is

$$\operatorname{dist}(c_j, \hat{c}) := \|\eta\|_{g^1_{\hat{c}}}, \tag{5.43}$$

where $\hat{c}$ denotes the optimal solution and $\eta \in T_{\hat{c}}B_e$ is defined by $c_j = \mathcal{R}_{\hat{c}}(\eta)$. Therefore, in the discrete setting, we compute the shortest distance to $\hat{c}$ in normal direction for each node of the iterated shape c_j. Then we form the L^2-norm of this distance field over $\hat{c}$, which is used to measure the convergence. The cost of this operation is quadratic with respect to the number of nodes on the surface. Starting in one node on c_j in normal direction, the determination of a point of intersection with $\hat{c}$ requires to check all boundary segments. This is the reason why we restrict our numerical results to two-dimensional computations.

Figure 5.3a visualizes the convergence history of different BFGS strategies compared to a pure gradient method for the problem (4.50)-(4.56). It can clearly be seen that the BFGS methods are superior to the gradient based method. Furthermore, we observe superlinear convergence in the BFGS case. It is yet surprising that, in this particular test case, there is hardly any difference between the number of stored gradients in the limited memory BFGS methods. This changes for the pure elliptic case of (4.50)-(4.54). Leaving out the the time dependence in (4.50)-(4.54) leads to the following PDE constrained shape optimization problem:

$$\min_{u \in B_e} J(u) = \int_{X(u)} (y - \overline{y})^2 dx + \mu \int_u 1 \, ds \tag{5.44}$$

$$\text{s.t.} \quad -\operatorname{div}(k \nabla y) = f \quad \text{in } X(u) \tag{5.45}$$

$$y = 1 \quad \text{on } \Gamma_{\text{top}} \tag{5.46}$$

$$y = 0 \quad \text{on } \Gamma_{\text{bottom}} \tag{5.47}$$

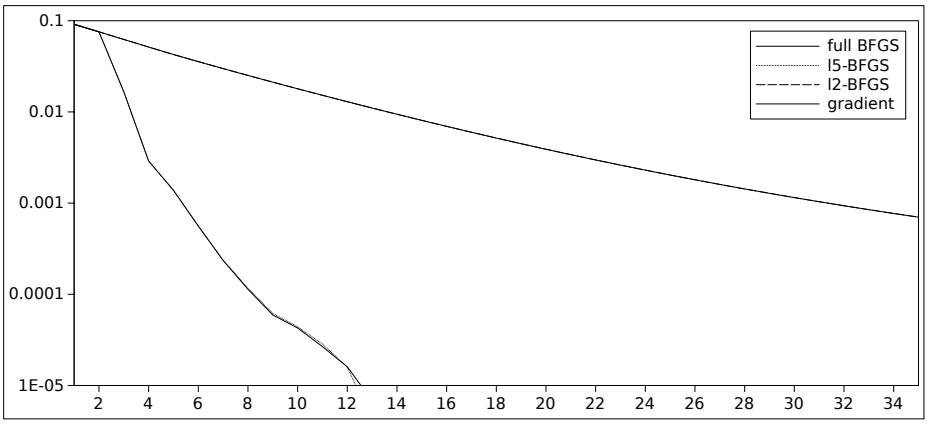
$$\frac{\partial y}{\partial n} = 0 \quad \text{on } \Gamma_{\text{bottom}} \cup \Gamma_{\text{left}} \cup \Gamma_{\text{right}} \tag{5.48}$$

In this elliptic case, we observe small improvements in the convergence while enlarging the memory width for the BFGS method, which is visualized in Figure 5.3b.

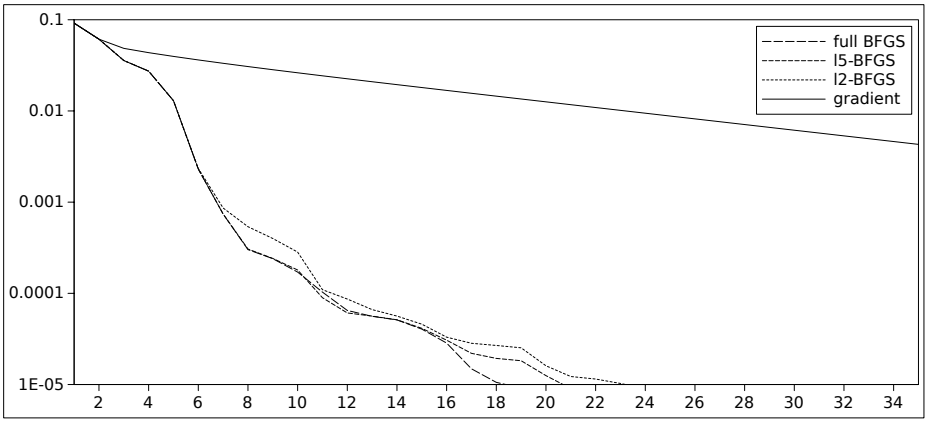
Remark 5.5. Note that the boundary conditions of (5.44)-(5.48) are changed compared to the parabolic model (4.50)-(4.54) since these conditions lead to a homogeneous steady state distribution of y. The shape gradient for the problem (5.44)-(5.48) can be found in [46].

Back in the parabolic case, we also investigate the influence of the grid on the convergence, which is depicted in Figure 5.4. Two grids are tested. A coarse one with approximately 25 000 cells and a much finer grid with about 100 000 cells. It can be seen that the convergence is almost grid independent for both, the gradient and the BFGS method. This also visualizes the discretization error.

In a final test run, we investigate the convergence under noisy measurements. For this purpose, we add white noise $\omega(t, x)$ to the measurements $\overline{y}(t, x)$ with an amplitude of 5% of the maximum value of $\overline{y}$, which is 1.0 due to the boundary conditions. We perform 100 runs of Algorithm 2 in the above-described setting. Due to the disturbed measurements, we obtain slightly different optimal shapes. In order to estimate the difference between these shapes, the maximum point-wise distance



(a) Parabolic problem.



(b) Elliptic problem.

Figure 5.3: Limited memory BFGS methods compared to a pure gradient method.

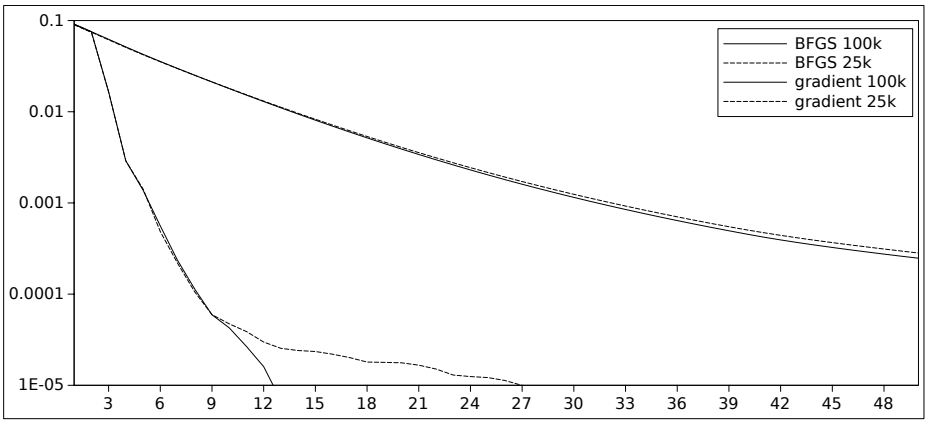
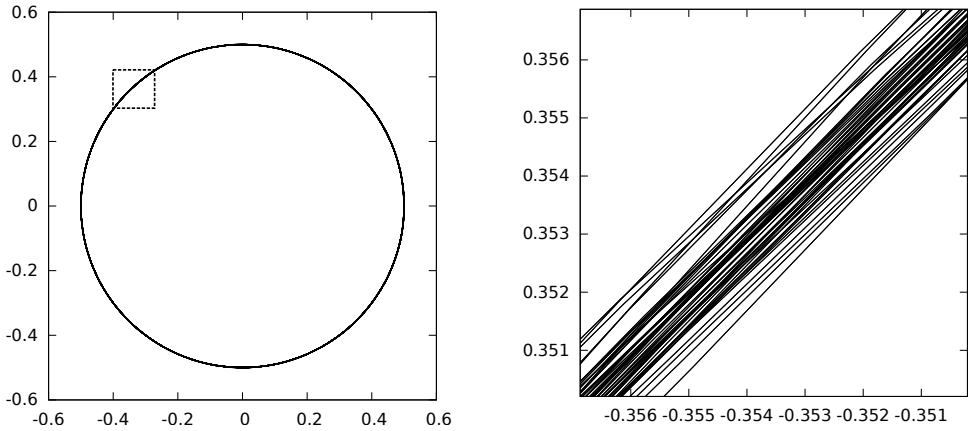


Figure 5.4: BFGS method compared to pure gradient method on different grids.



(a) Optimal shapes. (b) Zoom into the dashed frame.

Figure 5.5: 100 optimized shapes with 5% noise in the measurements $\overline{y}$.

is evaluated. We observe that this distance is only 0.21% of the mean diameter of all converged shapes, which is relatively small compared to the noise added to the measurements. In Figure 5.5a, all 100 converged shapes are visualized. From this point of view there are hardly any differences noticeable. Figure 5.5b shows a zoom into the region framed with dashed lines with approximately 200x magnification. Furthermore, in these experiments, we observe that also a regularization parameter $\mu = 0$ can be used without a noticeable difference.

Chapter 6

Optimization based on Steklov-Poincaré metrics

As seen in Chapter 4, it is a very tedious process to derive boundary formulations of shape derivatives. Along the way, domain formulations in the form of integrals over the entire domain appear as an intermediate step. In the literature, e.g., [23, 93], major effort in shape calculus has been devoted towards expressions for shape derivatives in the Hadamard-form, i.e., in the boundary integral form. It is known that the second order shape derivative, formerly coined as the shape Hessian, is nonsymmetric in general. This has been an obstacle for the algorithmic developments in shape optimization in the fashion of non-linear programming for a long time. In the previous chapters and in [84], shape optimization has been considered as optimization on Riemannian shape manifolds. This enables the design and analysis of NLP-like algorithms including one-shot sequential quadratic programming and gives a theoretical insight into the structure of the second order shape derivative in comparison to the Riemannian shape Hessian. Coercivity results for shape Hessians for elliptic problems can be found in [22] and well-posedness results for those are given in [29].

On the other hand, it has been shown that the above-mentioned intermediate formulations have numerical advantages, see for instance [16, 34, 44, 73]. In [58], also practical advantages of the domain shape formulations have been presented. For example, they require less smoothness assumptions. Furthermore, the derivation as well as the implementation of domain integral formulations require less manual and programming work. Thus, there arises the natural aim to combine domain integral formulations of shape derivatives with NLP-type optimization strategies on shape spaces, which seem, so far, tightly coupled with boundary integral formulations of shape derivatives. This chapter aims at demonstrating that this coupling is indeed possible and that it naturally leads to Steklov-Poincaré metrics on shape spaces. In contrast to Chapter 5, this chapter avoids consciously surface formulations of shape derivatives in order to provide more handy optimization algorithms. Another aim

of this chapter is a comparison – under algorithmic aspects – between the approach described in Section 5.2 and and the novel approach provided in this chapter. Recall that the approach described in Section 5.2 is based on surface expressions of shape derivatives and contains the Riemannian shape gradient built with respect to the Sobolev metric g^1. In contrast, the novel approach provided in this chapter involves volume formulations of shape derivatives and a corresponding metric.

This chapter is organized in the following way. In Section 6.1, we discuss generalized Steklov-Poincaré operators as basis for scalar products on shape spaces. Section 6.2 rephrases the NLP-like optimization algorithms on shape spaces within the framework of domain integral formulations of shape derivatives and in the context of the Steklov-Poincaré metrics introduced in Section 6.1. Finally, Section 6.3 discusses not just algorithmic and implementation details, but also numerical results for the parabolic transmission shape optimization problem introduced in Subsection 4.2.2. Moreover, it compares the two above-mentioned approaches with each other from a computational point of view.

6.1 The Steklov-Poincaré metric $\mathbf{g^S}$

We first consider two-dimensional shapes. Recall that we look at connected and compact subsets $X_2 \subset X \subset \mathbb{R}^2$ with $X_2 \neq \emptyset$ and $\mathcal{C}^\infty$-boundary $\Gamma_{\text{int}} = \partial X_2$, where X denotes a bounded domain with Lipschitz-boundary (cf. right picture of Figure 4.1). As in Section 5.2, we identify the variable boundary Γ_{int} with a simple closed smooth curve $c\colon S^1 \to \mathbb{R}^2$. The set of smooth boundary curves c is characterized by the shape space B_e defined in (3.1). Its tangent space T_cB_e is given in (3.2). Recall that the Sobolev metric g^1 given in (3.6) has previously been considered in this thesis. However, it requires a shape derivative in Hadamard-form as an efficient means to solve linear systems involving the Laplace-Beltrami operator in surfaces. All of this is certainly not impossible, but requires computational overhead, which we can get rid of by using the metrics discussed below. The algorithmic aspects of the both approaches are compared in Section 6.3. This section is devoted to the introduction of metrics dovetailed to shape optimization and based on domain formulations of shape derivatives.

The ideal Riemannian metric for shape manifolds in the context of PDE constrained shape optimization problems is to be derived from a symmetric representation of the second shape derivative in the solution of the optimization problems. Often, this operator can be related to the Dirichlet-to-Neumann map, aka Steklov-Poincaré operator, or the Laplace-Beltrami operator, see for instance [82]. If one aims at mesh independent convergence properties, one of these two is appropriate in most cases. It can be observed that the Laplace-Beltrami operator is spectrally equivalent to the square of the Steklov-Poincaré operator. We recall that two unitary operators $F, \widetilde{F}$ on separable Hilbert spaces $H, \widetilde{H}$ are said to be *spectrally equivalent* if there exists an isometric isomorphism $W\colon H \to \widetilde{H}$ such that $WF = \widetilde{F}W$. The latter operator, the Steklov-Poincaré operator, seems to be more fundamental and

we focus on it as basis for the scalar product on $T_c B_e$. Another advantage of this operator is that it blends well with a corresponding mesh deformation strategy.

Most often, the Dirichlet-to-Neumann map is associated with the Laplace operator. However, as pointed out in [3, 49], more general elliptic operators can be involved. For the purpose of mesh deformation, an elasticity operator may be the ideal choice. In numerical computations, its inverse, the Neumann-to-Dirichlet map, aka Poincaré-Steklov operator, is also of importance. Therefore, we first define these operators. In the following, we restrict no longer to the two-dimensional case. Note that (B_e, g^1) can be extended to higher dimensions d (cf. Section 3.3). In the sequel, we use the continuous generalized trace map

$$\begin{aligned} \gamma\colon H_0^1(X,\mathbb{R}^d) &\to H^{1/2}(\Gamma_{\text{int}},\mathbb{R}^d)\times H^{-1/2}(\Gamma_{\text{int}},\mathbb{R}^d), \\ U &\mapsto \begin{pmatrix}\gamma_0 U\\ \gamma_1 U\end{pmatrix} \coloneqq \begin{pmatrix} U|_{\Gamma_{\text{int}}} \\ \frac{\partial}{\partial n} U|_{\Gamma_{\text{int}}}\end{pmatrix}. \end{aligned} \tag{6.1}$$

Analogously to [49], we define the Neumann and Dirichlet solution operator.

Definition 6.1 (Neumann solution operator, Dirichlet solution operator)**.** *Let X, Γ_{int} and γ be as above and let $U, V \in H_0^1(X,\mathbb{R}^d)$ be two vector fields. Moreover, let*

$$a\colon H_0^1(X,\mathbb{R}^d)\times H_0^1(X,\mathbb{R}^d)\to\mathbb{R} \tag{6.2}$$

denote a symmetric and coercive bilinear form. The Neumann solution operator for the inner boundary Γ_{int} is given by

$$E_N\colon H^{-1/2}(\Gamma_{\text{int}},\mathbb{R}^d)\to H_0^1(X,\mathbb{R}^d),\ u\mapsto U \tag{6.3}$$

with U defined as the solution of the variational problem

$$a(U,V)=\int_{\Gamma_{\text{int}}} u^T(\gamma_0 V)\ ds \quad \forall V\in H_0^1(X,\mathbb{R}^d). \tag{6.4}$$

The Dirichlet solution operator for the inner boundary Γ_{int} is given by

$$E_D\colon H^{1/2}(\Gamma_{\text{int}},\mathbb{R}^d)\to H_0^1(X,\mathbb{R}^d),\ u\mapsto U \tag{6.5}$$

with U defined as the solution of the variational problem

$$a(U,V)=0 \quad \forall V\in H_0^1(X,\mathbb{R}^d);\quad U|_{\Gamma_{\text{int}}}=u. \tag{6.6}$$

Remark 6.2. Note that both, the Neumann and Dirichlet solution operator, are derived from a symmetric and coercive bilinear form (cf. (6.2)). Furthermore, the integral in the right-hand side of equation (6.4) is to be understood as duality pairing.

Now, we can define the Dirichlet-to-Neumann map and the Neumann-to-Dirichlet map.

Definition 6.3 (Steklov-Poincaré operator, Poincaré-Steklov operator)**.** *Let γ_0 and γ_1 be defined as in (6.1). In the setting of Definition 6.1, the Dirichlet-to-Neumann map T and the Neumann-to-Dirichlet map S are defined by*

$$T \coloneqq \gamma_1 \circ E_D \colon H^{1/2}(\Gamma_{\text{int}}, \mathbb{R}^d) \to H^{-1/2}(\Gamma_{\text{int}}, \mathbb{R}^d), \tag{6.7}$$

$$S \coloneqq \gamma_0 \circ E_N \colon H^{-1/2}(\Gamma_{\text{int}}, \mathbb{R}^d) \to H^{1/2}(\Gamma_{\text{int}}, \mathbb{R}^d). \tag{6.8}$$

The Dirichlet-to-Neumann map is also called the Steklov-Poincaré operator and the Neumann-to-Dirichlet map is also known as the Poincaré-Steklov operator.

The following theorem is a generalization of Theorem 2.3.1 in [49] from scalar fields to vector fields. It collects basic properties of the Steklov-Poincaré and Poincaré-Steklov operator. For the proof we refer to [49].

Theorem 6.4. *Let T and S be the Steklov-Poincaré and Poincaré-Steklov operators. Then the following statements hold:*

(i) $T = S^{-1}$.

(ii) The operator S is continuous and symmetric with respect to the standard L^2-duality pairing between $H^{-1/2}(\Gamma_{\text{int}}, \mathbb{R}^d)$ and $H^{1/2}(\Gamma_{\text{int}}, \mathbb{R}^d)$.

(iii) Both operators, S and T, are coercive.

For the purpose of defining an appropriate scalar product on the tangent space of shape spaces, we define the following mappings.

Definition 6.5 (Projected Steklov-Poincaré operator, projected Poincaré-Steklov operator)**.** *Let $H \in \{H^{-1/2}, H^{1/2}\}$ and let T and S be the Steklov-Poincaré and Poincaré-Steklov operators. We define*

$$\eta \colon H(\Gamma_{\text{int}}) \to H(\Gamma_{\text{int}}, \mathbb{R}^d),\ \alpha \mapsto \alpha \cdot n,$$

$$\eta^T \colon H(\Gamma_{\text{int}}, \mathbb{R}^d) \to H(\Gamma_{\text{int}}),\ U \mapsto n^T U.$$

The projected Steklov-Poincaré operator T^{pr} and the projected Poincaré-Steklov operator S^{pr} are given by

$$T^{pr} \coloneqq \eta^T \circ T \circ \eta \colon H^{1/2}(\Gamma_{\text{int}}) \to H^{-1/2}(\Gamma_{\text{int}}), \tag{6.9}$$

$$S^{pr} \coloneqq \eta^T \circ S \circ \eta \colon H^{-1/2}(\Gamma_{\text{int}}) \to H^{1/2}(\Gamma_{\text{int}}). \tag{6.10}$$

Both operators, T^{pr} and S^{pr}, inherit symmetry, coercivity, continuity and invertibility from the operators T and S. However, in general, we observe $T^{pr} \neq (S^{pr})^{-1}$. Both operators can be used for the definition of a scalar product on the tangent space of a shape space. In line with the discussion of Sobolev metrics in [63], we prefer a scalar product with a smoothing effect like the projected Dirichlet-to-Neumann

map T^{pr}. However, we need its inverse in numerical computations, which is usually not S^{pr}, although spectrally equivalent. We can limit the computational burden if we use directly $(S^{pr})^{-1}$ as a metric on the tangent space, having a similar smoothing effect, but also the advantage of the straightforward inverse S^{pr}. Let us explicitly formulate the operator S^{pr}. Due to Definition 6.1, Definition 6.3 and Definition 6.5, the projected Poincaré-Steklov operator is given by

$$S^{pr}\colon H^{-1/2}(\Gamma_{\text{int}}) \to H^{1/2}(\Gamma_{\text{int}}),\ \alpha \mapsto (\gamma_0 U)^T n, \tag{6.11}$$

where $U \in H_0^1(X, \mathbb{R}^d)$ solves the Neumann problem

$$a(U,V) = \int_{\Gamma_{\text{int}}} \alpha \cdot (\gamma_0 V)^T n \ ds \quad \forall V \in H_0^1(X, \mathbb{R}^d), \tag{6.12}$$

which corresponds to an elliptic problem with fixed outer boundary and forces $\alpha \cdot n$ at the inner boundary Γ_{int}. After careful consideration of all opportunities, we propose to use the scalar products g^S defined below.

Definition 6.6 (Steklov-Poincaré metric)**.** *In the setting above, we define the scalar products g^S on $H^{1/2}(\Gamma_{\text{int}})$ by*

$$\begin{aligned} g^S\colon H^{1/2}(\Gamma_{\text{int}}) \times H^{1/2}(\Gamma_{\text{int}}) &\to \mathbb{R}, \\ (\alpha, \beta) &\mapsto \langle \alpha, (S^{pr})^{-1}\beta \rangle = \int_{\Gamma_{\text{int}}} \alpha(s) \cdot [(S^{pr})^{-1}\beta](s) \ ds. \end{aligned} \tag{6.13}$$

These scalar products are called the Steklov-Poincaré metrics.

Remark 6.7. Note that a Steklov-Poincaré metric depends on the choice of the bilinear form (6.2). Thus, different bilinear forms lead to various Steklov-Poincaré metrics.

6.2 Quasi-Newton methods

As already mentioned, the shape derivative can be expressed as a boundary integral (cf. (4.4)) due to the Hadamard Structure Theorem 4.7. Recall that the shape derivative can be written more concisely as

$$DJ_{\Gamma_{\text{int}}}[V] = \int_{\Gamma_{\text{int}}} \alpha r \ ds$$

if $V|_{\Gamma_{\text{int}}} = \alpha n$, where r is at least in $L^1(\Gamma_{\text{int}})$ (cf. Section 4.3). Due to this handy expression of the shape derivative and isomorphism (3.2), we can state the connection of the shape space B_e with respect to the Steklov-Poincaré metric g^S to shape calculus. A representation $h \in T_{\Gamma_{\text{int}}} B_e \cong \mathcal{C}^\infty(\Gamma_{\text{int}})$ of the shape gradient in terms of g^S is determined by

$$g^S(\phi, h) = (r, \phi)_{L^2(\Gamma_{\text{int}})} \quad \forall \phi \in \mathcal{C}^\infty(\Gamma_{\text{int}}), \tag{6.14}$$

which is equivalent to

$$\int_{\Gamma_{\text{int}}} \phi(s) \cdot [(S^{pr})^{-1} h](s)\ ds = \int_{\Gamma_{\text{int}}} r(s)\phi(s)\ ds \quad \forall \phi \in \mathcal{C}^\infty(\Gamma_{\text{int}}). \tag{6.15}$$

From this we get $h = S^{pr} r = (\gamma_0 U)^T n$, where $U \in H_0^1(X, \mathbb{R}^d)$ solves

$$a(U,V) = \int_{\Gamma_{\text{int}}} r \cdot (\gamma_0 V)^T n\ ds = DJ_{\Gamma_{\text{int}}}[V] = DJ_X[V] \quad \forall V \in H_0^1(X, \mathbb{R}^d). \tag{6.16}$$

This means that a representation of the domain integral formulation in terms of the elliptic form (6.2) as used in [34] can – projected to the normal component on the interface Γ_{int} – be interpreted as the representation of the boundary integral formulation in terms of $(S^{pr})^{-1}$. However, in physical terms, the information of the shape derivative is used as a force in the domain or on the boundary. We obtain a vector field U as an (intermediate) result, which can serve as a deformation of the computational mesh identical to a Dirichlet deformation.

Remark 6.8. In general, $h = S^{pr} r = (\gamma_0 U)^\top n$ is not necessarily an element of $T_{\Gamma_{\text{int}}} B_e$ because it is not ensured that $U \in H_0^1(X, \mathbb{R}^d)$ is $\mathcal{C}^\infty$. Under special assumptions depending on the coefficients of a second-order partial differential operator and the right-hand side of a PDE, a weak solution U which is at least H_0^1-regular is $\mathcal{C}^\infty$ by Theorem 2.14.

We rephrase the limited memory BFGS quasi-Newton method for shape optimization (cf. Section 5.2) in terms of g^S and in generalization to domain formulations of shape derivatives. We note that a complete deformation of a shape optimization algorithm is just the (linear) sum of all iterations. This means that the BFGS update formulas can be rephrased directly in terms of the deformation vector field, rather than only as boundary deformations to be transferred to the domain mesh in each iteration.

BFGS update formulas need the evaluation of scalar products, where at least one argument is a gradient-type vector. This is a vector which indeed arises as a gradient on the variable boundary Γ_{int}. In contrast, a deformation-type vector describes an arbitrary boundary deformation in normal direction. According to the Steklov-Poincaré metric introduced in Section 6.1, we can assume that a gradient-type vector $u \in T_c B_e$ can be written as

$$u = (\gamma_0 U)^T n \tag{6.17}$$

for some vector field $U \in H_0^1(X, \mathbb{R}^d)$. The other argument v is either of gradient-type or of deformation-type, which can also be assumed to be of the form (6.17), i.e., $v = (\gamma_0 V)^T n$ for some $V \in H_0^1(X, \mathbb{R}^d)$. If u is a gradient of a shape objective function J, we observe

$$g^S(u,v) = DJ_{\Gamma_{\text{int}}}[V] = DJ_X[V] = a(U,V). \tag{6.18}$$

This observation can be used to reformulate the scalar product $g^S(\cdot,\cdot)$ on the boundary equivalently as $a(\cdot,\cdot)$ for domain representations. In the sequel, we only consider domain representations $U_j \in H_0^1(X,\mathbb{R}^d)$ of $\mathrm{grad}J(c_j) \in H^{1/2}(\Gamma_{\mathrm{int}})$, mesh deformations $S_j \in H_0^1(X,\mathbb{R}^d)$ and differences $Y_j \coloneqq U_{j+1} - \mathcal{T}_{S_j} U_j \in H_0^1(X,\mathbb{R}^d)$, where $\mathcal{T}_{S_j}$ denotes the vector transport as in Section 5.2. With this notation we formulate the double-loop of an inverse limited memory BFGS quasi-Newton method in Algorithm 3. The resulting vector $q = G_j^{-1}\mathrm{grad}J(c_j) \in H_0^1(X,\mathbb{R}^d)$ is simultaneously a shape deformation as well as a deformation of the domain mesh.

Algorithm 3. Inverse limited memory BFGS update in terms of g^S.

$\rho_j \leftarrow g^S\left((\gamma_0 Y_j)^T n, (\gamma_0 S_j)^T n\right)^{-1} = a(Y_j,S_j)^{-1}$
$q \leftarrow U_j$
for $i = j-1,\dots,j-m$ **do**
  $S_i \leftarrow \mathcal{T}_q S_i$
  $Y_i \leftarrow \mathcal{T}_q Y_i$
  $\alpha_i \leftarrow \rho_i\, g^S\left((\gamma_0 S_i)^T n, (\gamma_0 q)^T n\right) = \rho_i\, a(S_i,q)$
  $q \leftarrow q - \alpha_i Y_i$
end for
$q \leftarrow \frac{g^S\left((\gamma_0 Y_{j-1})^T n, (\gamma_0 S_{j-1})^T n\right)}{g^S\left((\gamma_0 Y_{j-1})^T n, (\gamma_0 Y_{j-1})^T n\right)}\, U_j = \frac{a(Y_{j-1},S_{j-1})}{a(Y_{j-1},Y_{j-1})}\, U_j$
for $i = j-m,\dots,j-1$ **do**
  $\beta_i \leftarrow \rho_i\, g^S\left((\gamma_0 Y_i)^T n, (\gamma_0 z)^T n\right) = \rho_i\, a(Y_i,z)$
  $q \leftarrow q + (\alpha_i - \beta_i) Y_i$
end for
return q

6.3 Numerical results

This section discusses algorithmic and implementation details as well as numerical results for the parabolic shape interface problem introduced in Subsection 4.2.2. Moreover, the limited memory BFGS shape optimization algorithms of Section 5.2 based on surface expressions of shape derivatives and the Sobolev metric g^1 are compared with the analogous Algorithm 3 based on volume representations of shape derivatives and the Steklov-Poincaré metric g^S.

Implementation and algorithmic details

As in Subsection 5.3.2, we consider the domain $X = (-1,1)^2$, which contains a compact and closed subset X_2 with smooth boundary Γ_{int}, and we choose $k_1 = 1$, $k_2 = 0.001$, $T = 20$, $y_0(x) = 0$ for all $x \in X$ and $f(x,t) = 0$ in $(x,t) \in X \times (0,T]$. Furthermore, we build the artificial data $\overline{y}$ exactly as described in Subsection 5.3.2. In contrast to Subsection 5.3.2, where the perimeter regularization is weighted by

$\mu = 10^{-4}$, the results shown in this section are computed under a mild perimeter regularization with $\mu = 10^{-6}$.

The numerical solution of the parabolic boundary value problem (4.51)-(4.56) is obtained by discretizing its weak formulation (4.57)-(4.61) with linear finite elements in space and an implicit Euler scheme in time. For the time discretization, 30 equidistantly distributed time steps are chosen. The diffusion parameter k is discretized as a piecewise constant function in contrast to the continuous trial and test functions. The corresponding adjoint problem (4.65)-(4.70) can be discretized in the same way. More precisely, it is not necessary to assemble different linear operators, which is attractive in terms of computational effort. All arising linear systems are then solved using the preconditioned conjugate gradient method.

An essential part of a shape optimization algorithm is to update the finite element mesh after each iteration. For this purpose, we use a solution of the linear elasticity equation

$$\operatorname{div}(\sigma) = f^{\text{elas}} \quad \text{in} \quad X \tag{6.19}$$

$$U = 0 \quad \text{on} \quad \Gamma_{\text{out}} \tag{6.20}$$

with

$$\sigma \coloneqq \lambda \operatorname{tr}(\epsilon) I + 2\mu\epsilon, \tag{6.21}$$

$$\epsilon \coloneqq \frac{1}{2}\left(\nabla U + \nabla U^T\right), \tag{6.22}$$

where σ is the stress tensor and ϵ is the strain tensor. Here λ and μ denote the Lamé parameters, which can be expressed in terms of Young's modulus E and Poisson's ratio ν (cf. (4.104) and (4.105)). The solution U is then added to the coordinates of the finite element nodes. Note that the Lamé parameters do not need to have a physical meaning here. It is rather essential to understand their effect on the mesh deformation. Young's modulus E states the stiffness of the material, which enables to control the step size for the shape update, and Poisson's ratio ν gives the ratio controlling how much the mesh expands in the remaining coordinate directions when compressed in one particular direction. The numerical results are obtained using $\nu = 0.01$ and $E = 0.1$.

Remark 6.9. Equations (6.19)-(6.22) are modified according to the optimization approach under consideration.

Comparison between Algorithm 2 and 3

Our investigations focus on the comparison between two approaches. More precisely, we compare Algorithm 2, based on *surface expressions of the shape derivatives*, versus Algorithm 3, based on *volume formulations of shape derivatives*. For convenience we summarize the main aspects of these two approaches:

- **First approach (Algorithm 2):**
  This approach is based on surface expressions of shape derivatives as described intensively in Section 5.2. Here a representation of the shape gradient on Γ_{int} with respect to the Sobolev metric g^1 is computed and applied as a Dirichlet boundary condition in the linear elasticity mesh deformation. This involves two operations, which are non-standard in finite element tools and, thus, lead to additional coding effort:

  (i) Since we are dealing with linear finite elements, the gradient expressions of the state y and the adjoint p in (4.76) or, equivalently, (4.77) are piecewise constant and cannot be applied directly to the mesh as deformations. Thus, we have to implement a kind of L^2-projection on Γ_{int} as described in Subsection 5.3.2 bringing back the sensitivity information into the space of continuous and linear functions.

  (ii) We need a discrete version of the Laplace-Beltrami operator for the Sobolev metric g^1. We implement the formulas given in [61] as described in Subsection 5.3.2. Since the approach presented in this publication is tailored to two-dimensional surfaces, we have to extend our two-dimensional grid in the third coordinate direction.

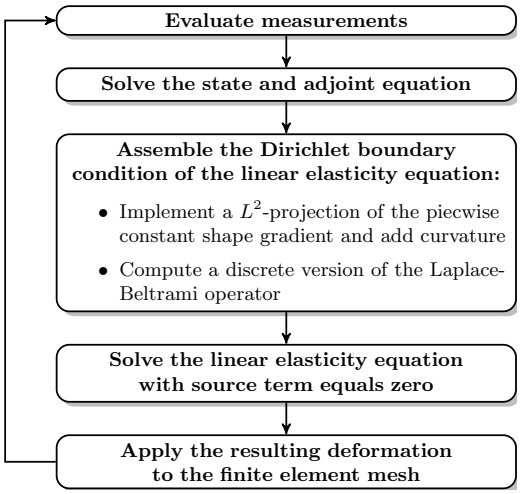
- **Second approach (Algorithm 3):**
  Algorithm 3 involves volume formulations of shape derivatives and a corresponding metric, the Steklov-Poincaré metric g^S, which is very attractive from a computational point of view. The computation of a representation of the shape gradient with respect to the chosen inner product of the tangent space is moved into the mesh deformation itself. The elliptic operator (6.2) – in our setting, the linear elasticity – is used as both, an inner product and a mesh deformation, leading to only one linear system, which has to be solved.

Both approaches follow roughly the same steps. In Figure 6.1, the complete optimization algorithms are summarized for the choice $m = 0$ in Algorithm 2 and 3. Note that Algorithm 2 and 3 boil down to steepest descent methods by choosing $m = 0$. Figure 6.1b summarizes the entire optimization algorithm for the second approach. In contrast to this, Figure 6.1a gives the complete optimization algorithm in the case of surface shape derivative expressions.

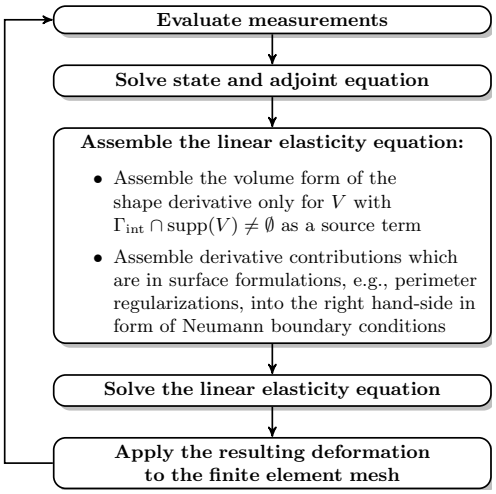
As already mentioned in Remark 6.9, equations (6.19)-(6.22) are modified according to the optimization approach under consideration. If we use a surface formulation of the shape derivative, which is given in (4.76) or, equivalently, (4.77) for the parabolic model problem, the Dirichlet boundary condition

$$U = U^{\text{surf}} \quad \text{on} \quad \Gamma_{\text{int}}$$

is added. Here U^{surf} is a representation of the shape gradient with respect to the Sobolev metric g^1. In this case, the source term f^{elas} is set to zero. Otherwise, if the mesh deformation operator is also used as shape metric, f^{elas} is assembled



(a) Surface shape derivative expressions.



(b) Volume shape derivative expressions.

Figure 6.1: Complete optimization algorithms.

according to the domain expression (4.32) and there is no Dirichlet condition on U. This covers only the portion of the shape derivative for which a volume formulation is available. Parts of the objective function leading to surface expressions only, such as, for instance, the perimeter regularization j_{reg}, are incorporated in Neumann boundary conditions. More precisely, for our model problem given in Subsection 4.2.2 we have to solve the following equation in the context of a domain formulation of the shape derivative and its representation in terms of g^S:

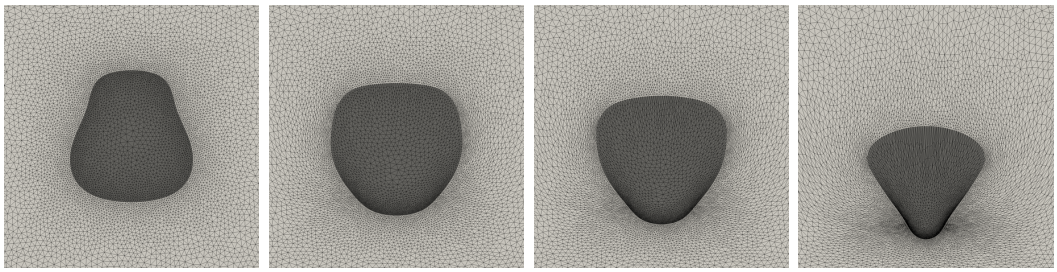
$$a(U,V) = Dj_X[V] + Dj_{\text{reg}}(X)[V] \quad \forall V \in H_0^1(X,\mathbb{R}^d), \tag{6.23}$$

where $Dj_X[V]$ is given in (4.73) and $Dj_{\text{reg}}(X)[V]$ is given in (4.13). We set the bilinear form (6.2) as the weak form of the linear elasticity equation leading to

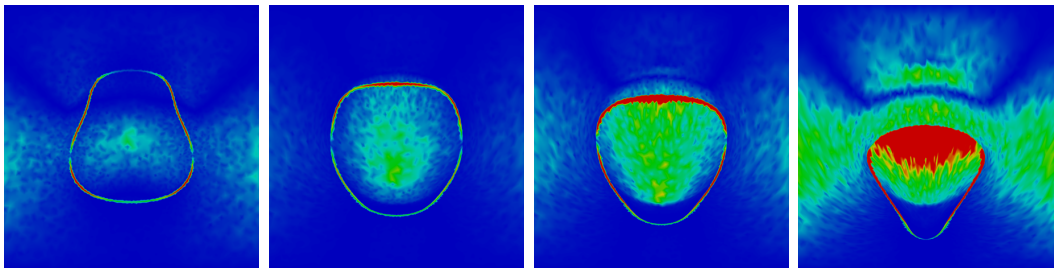
$$a(U,V) = \int_X \sigma(U) : \epsilon(V)\ dx. \tag{6.24}$$

Remark 6.10. Note that (6.23) is justified by the main result (6.16) of the previous sections stating that the connection between a volume formulation of a shape derivative and a bilinear form leads to a representation of the shape gradient with respect to g^S. By choosing the bilinear form (6.2) as the weak form of the linear elasticity equation (cf. (6.24)), equation (6.23) is the weak form of (6.19)-(6.22) involving the incorporated Neumann boundary condition. As mentioned above, this boundary condition arises from the perimeter regularization j_{reg}.

The right-hand side of the discretized weak form of (6.23) is assembled only for test functions whose support includes Γ_{int}. The behaviour of the algorithm with full assembly for all test functions is illustrated in Figure 6.2. The magnitude of the



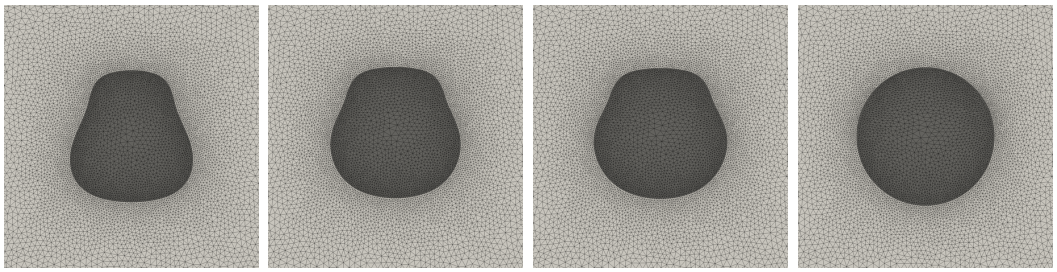
(a) BFGS iterates with unmodified approximation of the volume shape gradient.



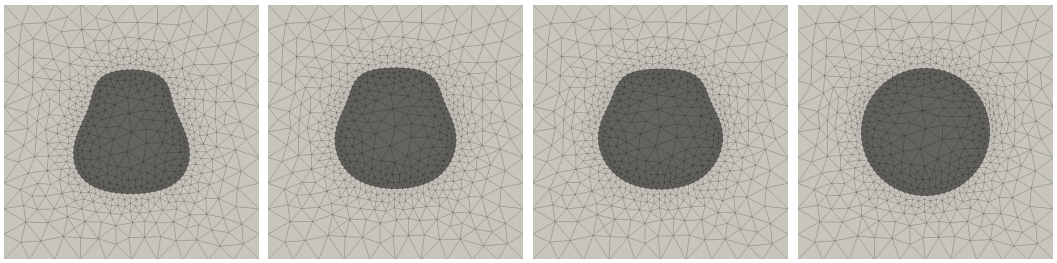
(b) Magnitude of unmodified volume source term.

Figure 6.2: Wrong mesh deformations and source term due to discretization errors in the unmodified right-hand side of (6.23).

unmodified discretization of the source term is visualized (cf. Figure 6.2b), which shows not only non-zero values outside of Γ_{int} due to discretization errors, but leads also to detrimental mesh deformations (cf. Figure 6.2a). Assembling the right-hand side of (6.23) only for test functions whose support intersect with Γ_{int} is due to the following reasoning. In exact integration, the integral $Dj_X[V]$ should be zero for all test functions V which do not have Γ_{int} within their support. Thus, non-zero integral contributions are caused by discretization noise. On the other hand, its effect on the optimization algorithm can be understood from a perturbation point of view. We may assume that the Riemannian shape Hessian, whose action in the optimal solution coincides with the action of the (standard) shape Hessian, i.e., $g^S(\nabla \text{grad} J[V], U) = D(DJ[V])[U]$, is coercive on the boundary, i.e., for projections $\eta^T V|_{\Gamma_{\text{int}}}$. This guarantees a well-posed problem. However, the Hessian operator approximated in the BFGS update strategy described in Algorithm 3 deals with a Hessian defined on the whole mesh, which possess a huge kernel, determined by all vector fields with zero normal component on the boundary. Thus, the space $H_0^1(\Omega, \mathbb{R}^d)$ of all admissible deformations has a decomposition $H_0^1(\Omega, \mathbb{R}^d) = H_{\Gamma_{\text{int}}} \oplus H_{\Gamma_{\text{int}}}^{\perp}$, where $H_{\Gamma_{\text{int}}} \coloneqq \{E_N(\alpha n) \colon \alpha \in H^{-1/2}(\Gamma_{\text{int}})\}$ and $H_{\Gamma_{\text{int}}}^{\perp}$ denotes its orthogonal complement in the chosen bilinear form (6.2) – in our setting, (6.24). Shape gradients and increments in $H_0^1(\Omega, \mathbb{R}^d)$ lie in $H_{\Gamma_{\text{int}}}$ only. It is obvious that limited memory BFGS update formulas produce steps which lie again in $H_{\Gamma_{\text{int}}}$ only. This means that the



(a) Smooth deformations and convergence to optimal shape due to modified source term.



(b) The volume-based optimization approach enables the use of much coarser spatial discretizations, approximately 1 000 cells.

Figure 6.3: BFGS iterates with corrected source term $Dj_X[V]$ indicating mesh independent convergence.

optimization algorithm in function spaces always acts on the coercive shape Hessian. However, the discretized version is a perturbation of this Hessian. Thus, perturbed coercive operators stay coercive if the perturbation is not too large. However, positive semidefinite operators with a non-trivial kernel get directions of negative curvature if they are perturbed. These directions of negative curvature are chosen if we allow non-zero components in the right-hand side of the discretized mesh deformation equation (6.23) in the interior of the domain. On the other hand, if we do not allow zero components there, the algorithm acts in the subspace of the discretization of $H_{\Gamma_{\text{int}}}$ only, where the projected Hessian is a perturbation of the shape Hessian and, thus, coercive if the perturbation is not too large.

We conclude this section with a brief discussion of the numerical results. Figure 6.2-6.3 show the initial configuration and the iterations 2, 4 and 20 of the full BFGS algorithm as just described. In Figure 6.2, the algorithm is shown for the unmodified assembly of the right-hand side in (6.23) leading to divergence, whereas Figure 6.3 shows a selection of BFGS iterates for the modified source term. Figure 6.3b demonstrates that the optimization algorithm based on domain shape derivative expressions can be applied to very coarse meshes. This is due to the fact that there is no dependence on normal vectors like in the case of surface shape gradients. Finally, Figure 6.4 shows the convergence of a full BFGS method, a limited memory BFGS method with three gradients in memory and a pure gradient method for the surface

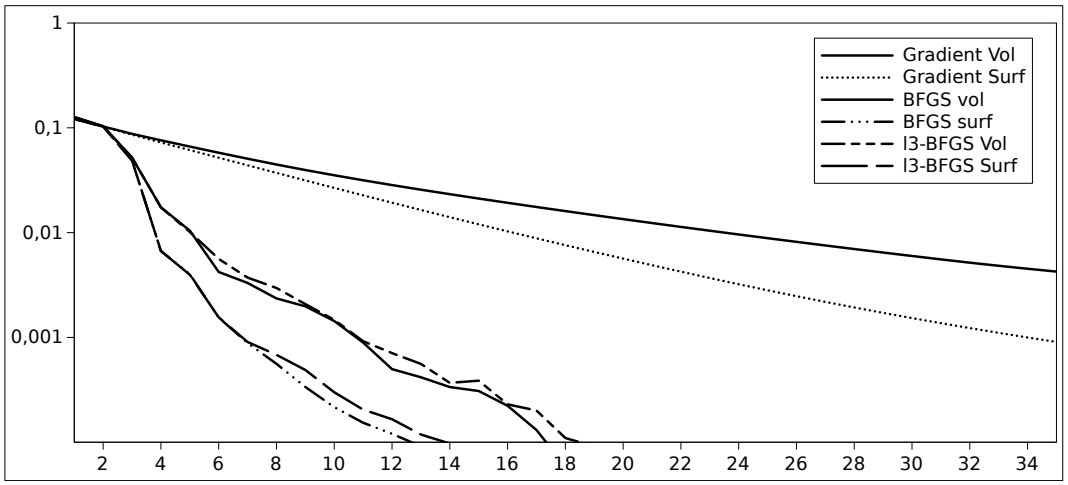


Figure 6.4: Convergence history of BFGS and gradient methods for volume versus surface shape derivative expressions on a grid with approximately 100 000 cells. Convergence is measured as an approximation of the geodesic distance in the shape space.

and volume shape derivative formulation. In our tests, the convergence with the Laplace-Beltrami representation of the shape gradient seems to require fewer iterations compared to the domain-based formulation. Yet, the domain-based form is computationally more attractive since it also works for much coarser discretizations. This can be seen in Figure 6.3. Figure 6.3a shows the necessary fineness of the mesh for the surface gradient to lead to a reasonable convergence. However, the coarse grid in Figure 6.3b works only for the domain-based formulation.

Remark 6.11. This chapter deals with elements of B_e only. In the next chapter, a novel shape space is introduced and the optimization Algorithm 3 is formulated on it. In its numerical part (Section 7.2), there are an illustration and a further discussion about the mesh quality difference by using the first and second approach (cf. Figure 7.3).

Chapter 7

Towards a novel shape space

In general, finding a shape space and an associated metric is a challenging task and different approaches lead to various models. There exists no common shape space or shape metric suitable for all applications. The specific situation is essential for the suitability of a certain approach.

The scalar product introduced in Definition 6.6 correlates shape gradients with H^1-deformations. Under special assumptions, these deformations give shapes of class $H^{1/2}$, which are defined below. Now, we get to the point mentioned in Section 4.3 (cf. Remark 4.34), where the shape space B_e is no longer suitable for our investigations because it unnecessarily limits the application of the methods established in the previous chapters. In the setting of B_e, shapes can be considered as the images of embeddings from the unit circle into the plane. From now on we have to think of shapes as boundary contours of deforming objects. Therefore, the definition of another shape space is required. An advantage over the shape space B_e is that the novel shape space to be constructed in the sequel is not limited to objects with $\mathcal{C}^\infty$-boundary.

Gradient flows and consequently geodesics depend on the metric which is chosen on a shape space. Thus, this chapter is not only devoted to a novel shape space definition, but also to the problem of quantifying differences between shapes. In [32], a Riemannian metric on a space of shape contours motivated by linear elasticity is proposed. This metric can be interpreted as the rate of physical dissipation during the deformation of a viscous liquid object (cf. [106]). In [104, 105, 106, 108], shapes are represented by level set functions, which allows topological transitions along geodesic paths whose computations are based on a variational time discretization. This establishes a link between a pairwise elastic shape matching and a Riemannian flow perspective on paths in shape spaces. All this is summarized in [104]. However, we are only interested in shapes sharing the same topology. The approach in [32] allows general topologies of shapes, but it requires the topology to stay the same during the evolution.

This chapter is organized as follows. Section 7.1 is devoted to the set of all shapes in

the context of the Steklov-Poincaré metric g^S, i.e., to a novel shape space definition. Moreover, its connection to shape calculus is given in order to formulate the shape quasi-Newton methods of Chapter 6 (Algorithm 3) on this shape space. Section 7.2 demonstrates how Algorithm 3 works with such shapes. Finally, Section 7.3 links geodesics discussed in [32] with geodesics in the novel shape space $\mathcal{B}^{1/2}$ with respect to the Steklov-Poincaré metric g^S in order to obtain a distance measure for shapes of $\mathcal{B}^{1/2}$ with respect to g^S.

7.1 The shape space $\mathcal{B}^{1/2}$

In this section, we extend the definition of $\mathcal{C}^\infty$-shapes, which are elements of the shape space of Peter W. Michor and David Mumford introduced in Chapter 3, to *shapes of class* $H^{1/2}$. In the following, it is clarified what we mean by $H^{1/2}$-shapes.

We would like to recall once again that a shape in the sense of the shape space of Peter W. Michor and David Mumford is given by the image of an embedding from the unit sphere S^{d-1} into the Euclidean space $\mathbb{R}^d$. In view of our generalization, it has technical advantages to consider a prior shape Γ_0 as the boundary $\Gamma_0 = \partial\mathcal{X}_0$ of a connected and compact subset $\mathcal{X}_0 \subset X \subset \mathbb{R}^d$ with $\mathcal{X}_0 \neq \emptyset$, where X denotes a bounded Lipschitz domain (cf. Figure 4.1). Let the prior set $\mathcal{X}_0$ be a Lipschitz domain, i.e., Γ_0 is a Lipschitz boundary. An example of a prior shape is the cube. It is the union of six faces, where each is a portion of a plane, i.e., a smooth surface. General shapes – in our novel terminology – arise from H^1-deformations of such a prior set $\mathcal{X}_0$. These H^1-deformations, evaluated at a prior shape $\Gamma_0 = \partial\mathcal{X}_0$, give deformed shapes Γ_{int} if the deformations are injective and continuous. We call these shapes of class $H^{1/2}$ and define the set

$$\begin{aligned}\mathcal{H}^{1/2}(\Gamma_0, \mathbb{R}^d) \coloneqq \{w\colon \Gamma_0 \to X\colon \exists W \in H^1(X, X) \text{ s.t.} \\ W|_{\Gamma_0} \text{ injective, continuous, } W|_{\Gamma_0} = w\}.\end{aligned} \tag{7.1}$$

However, in order to have a unique representation for each shape, we have to factor out the homeomorphisms from the prior shape Γ_0 into itself which are compatible with the set (7.1). Thus, we characterize the following shape space:

Definition 7.1 (Shape space $\mathcal{B}^{1/2}$)**.** *Let X, $\mathcal{X}_0$ and Γ_0 be as above. The space of all $H^{1/2}$-shapes is given by*

$$\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d) \coloneqq \mathcal{H}^{1/2}(\Gamma_0, \mathbb{R}^d) \big/ \text{Homeo}^{1/2}(\Gamma_0), \tag{7.2}$$

where $\mathcal{H}^{1/2}(\Gamma_0, \mathbb{R}^d)$ is given in (7.1) and $\text{Homeo}^{1/2}(\Gamma_0)$ is defined by

$$\text{Homeo}^{1/2}(\Gamma_0) \coloneqq \{w\colon w \in \mathcal{H}^{1/2}(\Gamma_0, \mathbb{R}^d),\ w\colon \Gamma_0 \to \Gamma_0 \text{ homeomorphism}\}. \tag{7.3}$$

Remark 7.2. Of course, the properties of the shape space $\mathcal{B}^{1/2}\left(\Gamma_0, \mathbb{R}^d\right)$ have to be investigated. For example the *independence of the prior shape Γ_0 in the shape space*

definition is an open question. If it is independent, we can choose, for example, the unit sphere S^{d-1} as prior shape. Another important question is whether the *shape space has a manifold structure*. Note that this question is very hard and a lot of effort has to be put into it to find the answer. From a theoretical point of view there are several other open questions. However, this goes beyond the scope of this thesis and is a topic of subsequent work.

Remark 7.3. In the following, we assume that $\mathcal{B}^{1/2}\left(\Gamma_0, \mathbb{R}^d\right)$ has a manifold structure. If necessary, we can refine the space $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$, e.g., by restriction to an explicit deformation field W. In our setting, it arises from the linear elasticity equation and the request of the existence of an arbitrary one is perhaps too strong. This way, we can replace $\mathcal{H}^{1/2}(\Gamma_0, \mathbb{R}^d)$ by a linear space, which is in particular a manifold. However, this conceivable limitation leaves the following theory untouched.

If $\Gamma \in \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ is smooth enough to admit a normal vector field n, the following isomorphisms arise from definition (7.1):

$$\begin{aligned} T_\Gamma \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d) &\cong \{h \colon h = \phi n \text{ a.e., } \phi \in H^{1/2}(\Gamma) \text{ continuous}\} \\ &\cong \{\phi \colon \phi \in H^{1/2}(\Gamma) \text{ continuous}\} \end{aligned} \tag{7.4}$$

Now, we can formulate the shape quasi-Newton methods of Section 6.2 on the shape space $\mathcal{B}^{1/2}\left(\Gamma_0, \mathbb{R}^d\right)$ with respect to g^S. Before we can do that, we have to state its connection to shape calculus.

Connection with shape calculus

As often mentioned in the previous chapters, the shape derivative of a shape functional can be expressed as boundary integral (cf. (4.4)) due to the Hadamard Structure Theorem 4.7. Moreover, it can be expressed as

$$DJ_{\Gamma_{\text{int}}}[V] = \int_{\Gamma_{\text{int}}} \alpha\, r \, ds \tag{7.5}$$

if $V\big|_{\Gamma_{\text{int}}} = \alpha n$, where r is at least in $L^1(\Gamma_{\text{int}})$ (cf. Section 4.3). Due to the handy expression (7.5) and isomorphism (7.4), we can state the connection of $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ with respect to the Steklov-Poincaré metric g^S to shape calculus. A representation $h \in T_{\Gamma_{\text{int}}}\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d) \cong \{h \colon h \in H^{1/2}(\Gamma) \text{ continuous}\}$ of the shape gradient in terms of g^S is determined by

$$g^S(\phi, h) = (r, \phi)_{L^2(\Gamma_{\text{int}})} \tag{7.6}$$

for all continuous $\phi \in H^{1/2}(\Gamma_{\text{int}})$, which is equivalent to

$$\int_{\Gamma_{\text{int}}} \phi(s) \cdot [(S^{pr})^{-1} h](s) \, ds = \int_{\Gamma_{\text{int}}} r(s)\phi(s) \, ds \tag{7.7}$$

for all continuous $\phi \in H^{1/2}(\Gamma_{\text{int}})$.

Quasi-Newton methods

Based on the connection (7.6) we can formulate the quasi-Newton methods of Section 6.2 also on $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ with respect to g^S. From (7.7) we get

$$h = S^{pr} r = (\gamma_0 U)^T n,$$

where $U \in H_0^1(X, \mathbb{R}^d)$ solves

$$a(U, V) = \int_{\Gamma_{\text{int}}} r \cdot (\gamma_0 V)^T n \, ds = DJ_{\Gamma_{\text{int}}}[V] = DJ_X[V] \quad \forall V \in H_0^1(X, \mathbb{R}^d). \tag{7.8}$$

In general, $h = S^{pr} r = (\gamma_0 U)^\top n$ is not necessarily an element of $T_{\Gamma_{\text{int}}} \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ because it is not ensured that $U \in H_0^1(X, \mathbb{R}^d)$ is continuous. Under special assumptions depending on the coefficients of a second-order partial differential operator, the right-hand side of a PDE, the domain X on which a PDE is defined and the dimension of X, the continuity of a weak solution of a PDE is guaranteed by Theorem 2.12 combined with the Sobolev embedding Theorem 2.8. More precisely, in the setting above, if the conditions of Theorem 2.12 are fulfilled, we obtain that the H_0^1-regular solution U is H^{m+2}-regular, where $m \in \mathbb{N}$. Moreover, we see that U is a $\mathcal{C}_b^m$-function in the two-dimensional case by choosing $r = 0$ in Theorem 2.8.

7.2 Numerical results

This section considers two model problems:

In Subsection 7.2.1, we consider the parabolic shape interface problem introduced in Subsection 4.2.2. We solve it by an application of the quasi-Newton method using the volume formulation (4.73) of the shape derivative as described in the previous chapter (cf. second approach in Section 6.3). In contrast to Section 6.3, the boundary of the initial shape is not smooth, but it has four kinks as illustrated in the left picture of Figure 7.1. The aim of this subsection is to illustrate how Algorithm 3 works with two-dimensional shapes of $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^2)$.

Subsection 7.2.2 is devoted to a test problem whose optimal solution cannot be achieved in the shape space B_e and with the classical approaches which are based on surface expressions of shape derivatives and on the Sobolev metric g^1. More precisely, we compute a shape embedded in a Stokes-flow, which minimizes drag and satisfies geometric constraints. The optimal solution is well-known as the so-called Haack ogive [39, 76]. This is a shape with two kinks as visualized as the final shape in Figure 7.2. The aim of this subsection is to illustrate the difference of the mesh quality by using the limited memory BFGS shape optimization Algorithm 2 in B_e and the analogous Algorithm 3 in $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^2)$. It has to be mentioned that this comparison is done in [85] and the results are briefly reproduced in this subsection because they endorse and illustrate the theoretical results of Chapter 6 and Section 7.1.

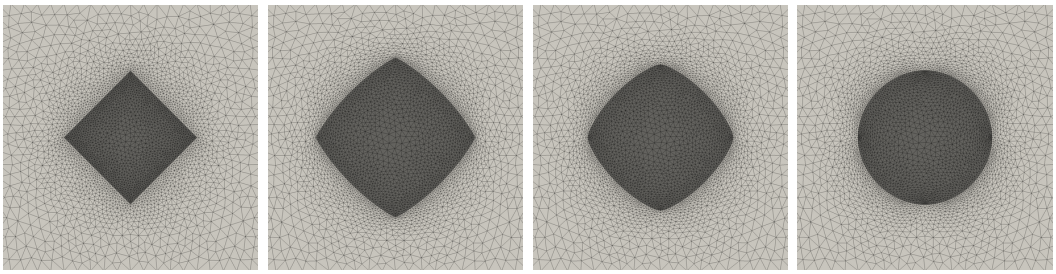


Figure 7.1: Smooth mesh deformations with kinks in the initial configuration.

7.2.1 Parabolic shape interface problem

We consider the parabolic shape interface problem introduced in Subsection 4.2.2 and use the same implementation details as in Section 6.3. More precisely, the numerical solution of the parabolic problem (4.51)-(4.56) is obtained by discretizing its weak formulation (4.57)-(4.61) with linear finite elements in space and an implicit Euler scheme in time. For discretization details we refer to Section 6.3.

We consider the domain $X = (-1,1)^2$, which contains a compact and closed subdomain X_2 with boundary Γ, and we set $k_1 = 1$, $k_2 = 0.001$, $T = 20$, $y_0(x) = 0$ for all $x \in X$ and $f(x,t) = 0$ in $(x,t) \in X \times (0,T]$. In contrast to Section 6.3, the boundary Γ of the initial shape is not smooth, but it has four kinks as illustrated in the left picture of Figure 7.1. The artificial data $\overline{y}$ are exactly as described in Subsection 5.3.2. In contrast to Subsection 5.3.2, where the perimeter regularization is weighted by $\mu = 10^{-4}$, and in contrast to Section 6.3, where a mild perimeter regularization $\mu = 10^{-6}$ is chosen, we use a stronger regularization for the non-smooth initial configuration shown in the left picture of Figure 7.1. It is chosen in the first iterations as $\mu_{\text{init}} = 0.01$. In this particular case, the regularization is controlled by a decreasing sequence from μ_{init} to $\mu = 10^{-6}$.

As already mentioned in Section 6.3, an essential part of the shape optimization algorithm is to update the finite element mesh after each iteration. For this purpose, we use a solution of the linear elasticity equation (6.19)-(6.22), where the stiffness E and the ratio ν are chosen as in Section 6.3. For the parabolic model problem, we have to solve equation (6.23) in the context of a domain formulation of the shape derivative and its representation in terms of g^S. Moreover, we have to assemble the right-hand side of (6.23) as described in Section 6.3 to prevent divergence as shown in Figure 6.2. For more details we refer to Section 6.3.

Figure 7.1 illustrates the initial configuration and the iterations 2, 4 and 20 of the full BFGS Algorithm 3. We see that this algorithm works also for shape geometries with kinks in the boundary.

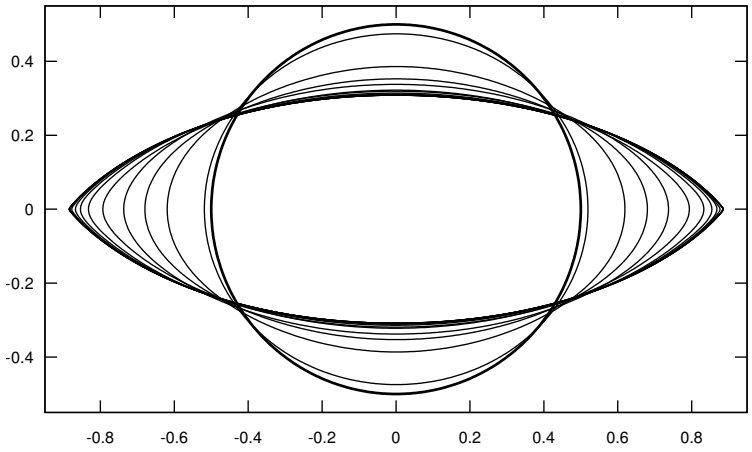


Figure 7.2: Initial and final shape geometry together with the iterations of the BFGS method in $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^2)$. Source: [85].

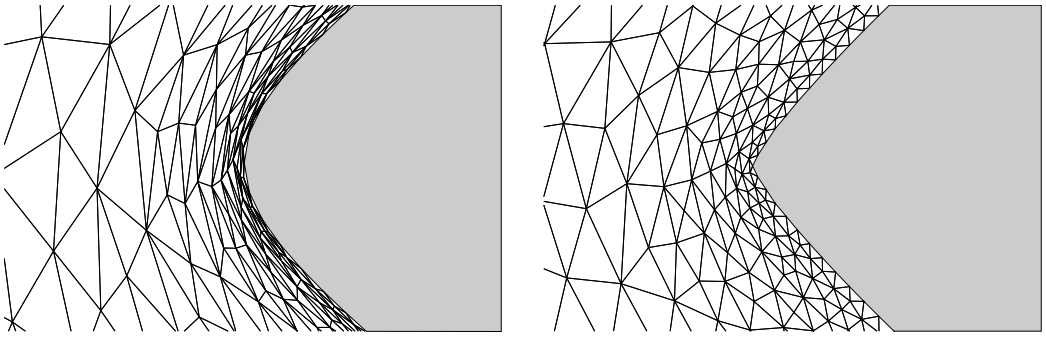
7.2.2 Minimization of energy dissipation in Stokes flow

This subsection reproduces the results of [85] because they endorse and illustrate the theoretical results of Chapter 6 and Section 7.1.

The test problem in this subsection is the minimization of energy dissipation of a body in a Stokes flow. We consider an incompressible flow which is dominated by viscous forces around an obstacle. A bounded domain X is partitioned into two disjoint subdomains $X_1, X_2 \subset X$ such that $X = X_1 \cup \Gamma \cup X_2$ with $\Gamma = \partial X_2$ as illustrated in Figure 4.1. The interior domain X_2 is the obstacle with variable boundary Γ and the exterior domain X_1 denotes the flow field. In the implementations, we use finite elements, where X_2 is a hole in the discretization mesh. The aim is to shape the two-dimensional shape X_2 such that the energy dissipation of the system is minimized under certain geometrical constraints. The volume $\mathrm{vol}(X_2)$ and the barycenter $\mathrm{bc}(X_2)$ of the shape X_2 are required to be constant for the optimization to be reasonable. These geometric constraints are necessary in order to obtain non-trivial solutions. The shape shrinks to a straight line without the volume constraint and floats out of the computational domain without the barycenter constraints.

The detailed problem formulation is given in [85, Section 2]. It includes an objective functional together with a system PDE – the Stokes equation, where the viscosity is normalized to 1 – and the above-mentioned geometric constraints. In [85], the quasi-Newton methods established in Section 5.2 for g^1 (cf. Algorithm 2) and in Section 6.2 for g^S (cf. Algorithm 3) together with an augmented Lagrangian framework is set up to solve this PDE constrained minimization problem.

The aim of this subsection is to visualize the difference of the mesh quality by using surface versus volume expressions of shape derivatives. Note that in Section 6.3 the implementation differences of these two approaches are pointed out (cf. first versus second approach). Figure 6.1 illustrates the complete optimization algorithms of



(a) g^1 after 65 gradient steps. (b) g^S after 16 BFGS steps.

Figure 7.3: Comparison of shape metrics with respect to the mesh quality. Source: [85].

these two approaches in the case of the steepest descent method. The algorithmic details of the augmented Lagrangian approach can be found in [85]. The discussion of the mesh quality focusses on the limited memory BFGS shape optimization Algorithm 2 and the analogous Algorithm 3. Note that the augmented Lagrangian framework is only necessary to handle the geometric constraints.

The computational domain is chosen to be $X = (-3, 6) \times (-2, 2)$. The initial shape X_2 is given by $X_2 = \{x \colon |x| < 0.5\}$ leading to $\mathrm{vol}(X_2) = \frac{\pi}{4}$. Thus, the prior shape Γ_0, which is necessary to define the shape space $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^2)$, is a circle with barycenter $\mathrm{bc}(X_2) = (0, 0)$, radius $r = 0.5$ and an element of B_e. The computational grid consists of 10 150 triangles. The PDE constrained minimization problem described above is solved by applying the limited memory BFGS shape optimization Algorithm 2 in B_e and the analogous Algorithm 3 in $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^2)$. Figure 7.2 shows the initial and final shape geometry, which are the bold shapes, and the corresponding 16 iterated shapes which arise when Algorithm 3 is applied in $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^2)$ and 3 gradients are stored. The final shape is an element of $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^2)$, where the prior shape is the above-mentioned circle centred at $(0, 0)$ with radius 0.5. An important point to note is that this optimal solution cannot be achieved by applying the limited memory BFGS shape optimization Algorithm 2. This algorithm – or, in general, a classical approach based on surface expressions of shape derivatives – only take deformations normal to the shape boundary into account. This reflects the Hadamard Structure Theorem 4.7 stating that only the normal component of deformations affects the objective function. The iterations lead to discretization meshes that do not allow further computations due to degenerated cells as visualized in Figure 7.3a. This figure shows a part of the shape and the corresponding mesh after 65 gradient steps. However, using the Steklov-Poincaré metric allows also surface nodes to slide along the shape, which strongly influences the mesh quality. This can be seen in Figure 7.3b, which illustrates a part of the shape and the corresponding mesh after 16 BFGS steps. For implementation details we refer to [85, Section 5].

7.3 Geodesics related to continuum mechanics

The aim of this section is to investigate the relation between the novel shape space $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ with respect to the Steklov-Poincaré metric g^S and the shape distance measure proposed in [32]. With this relation, we obtain a distance measure for shapes of $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ with respect to g^S. We see that the geodesics in $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ with respect to g^S are the geodesics discussed in [32]. The corresponding distance measure is called the *elastic deformation distance*. The authors use some concepts from continuum mechanics to propose a Riemannian metric motivated by linear elasticity on a space of shape contours.

Remark 7.4. The metric considered in [32] is interpreted as the rate of physical dissipation during the deformation of a viscous object in [105, 106] and is elaborated in [108]. Geodesics are connected with continuum mechanics of a viscous fluid transport. All this is summarized in [104]. Therefore, we refer to [104] for more details. Moreover, an implementation approach of geodesic paths, which is based on a variational time discretization, can be found in [104]. Note that in [104, 105, 106, 108], shapes are represented by level set functions, which allows topological transitions along the geodesic paths. In contrast to this, the approach in [32] allows general topologies, but it requires to maintain the topology during the evolution.

To measure distances in the novel shape space with respect to the Steklov-Poincaré metric, we use some Riemannian geometry concepts. We recall the following differential geometric statements from Section 2.3:
A Riemannian metric is an inner product on the tangent bundle of a manifold. On manifolds, the distance between two arbitrary points is defined as the infimum of the lengths of all differentiable paths connecting the two points under consideration. Note that a differentiable local shortest path between two points of a manifold is a geodesic. To compute the length of a path, we have to integrate its velocity, i.e., the absolute value of its first derivative along the path. To integrate the velocity of a path on the manifold, the length of tangent vectors have to be measured. This is usually done by utilizing the norm induced by the Riemannian metric.

The elastic deformation energy given in [32] and defined in the sequel is induced by the Steklov-Poincaré metric g^S due to the following reasoning:
We get a representation of the shape gradient in terms of g^S by solving (7.8). In particular, we obtain a tangent vector arising from the solution of (7.8), i.e., a deformation inducing shape morphings. As already clarified in Section 6.2, the scalar product $g^S(\cdot, \cdot)$ on the surface can be reformulated equivalently as a symmetric and coercive bilinear form $a(\cdot, \cdot)$ for volume representations (cf. (6.18)). Thus, it is sufficient to consider such an inner product to measure distances between shapes in $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ with respect to g^S. If we choose the inner product as the weak form of the linear elasticity equation leading to (6.24), we get an elastic deformation energy and a corresponding distance measure as in [32].

Remark 7.5. The elastic deformation energy can be defined via a variational formulation. It can be shown that minimizers of the corresponding variational problem exist and that they are essentially unique (cf. [32]). The computation of shortest paths goes beyond the scope of this thesis. Thus, we refer to [32, 104, 105, 106, 108]. In the numerical part of [32], shortest paths with respect to the elastic deformation energy are computed. The corresponding figures can be found in [32, Section 5.3].

Now, we consider the shape space $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$, more precisely, elements of it. We recall once again that – in this setting – shapes are given by the images of injective and continuous $H^{1/2}$-deformations of prior shapes Γ_0. These $H^{1/2}$-deformations come from H^1-deformations of open and bounded domains X containing Γ_0. Roughly speaking, the elements of $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ arise from H^1-deformations acting on prior sets $\mathcal{X}_0 \subset X$ with Lipschitz boundaries Γ_0. Thus, a variation Γ of a prior shape Γ_0 is associated with a mapping

$$U\colon [0,1] \to H_0^1(X, \mathbb{R}^d),\ t \mapsto U(t), \tag{7.9}$$

where the variable $t \in [0,1]$ represents geometrically the coordinate along a path of transport fields $U(t) \in H_0^1(X, \mathbb{R}^d)$. In this setting, $\mathcal{X}(t) \subset X(t)$ describes the deformed object and $\Gamma(t)$ characterizes the deformed shape at time $t \in [0,1]$. Therefore, a smooth path

$$\gamma\colon [0,1] \to \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d),\ t \mapsto u(t) \tag{7.10}$$

in this shape space is associated with a family $(U(t))_{t\in[0,1]}$ of deformations, where $u(t) = U(t)|_{\Gamma_0}$.

Remark 7.6. Note that a path (7.10) is given by a curve of injective and continuous deformation fields, i.e.,

$$\gamma\colon [0,1] \to \big(\Gamma_0 \to \mathbb{R}^d\big),\ t \mapsto \big(\theta \mapsto u(t,\theta)\big). \tag{7.11}$$

More precisely, a shape variation $\Gamma\,(= \Gamma(t))$ of a prior shape Γ_0 is given by the image of an injective and continuous deformation u of Γ_0 at time $t \in [0,1]$, i.e., $\Gamma = u(t, \Gamma_0)$. Since Γ_0 is fixed, we write $\Gamma = u(t)\,(= \gamma(t))$ to be in line with (7.10).

In our setting, such deformations U arise from (7.8) as already stated above. Thus, we can choose the inner product

$$a(U,V) = \int_X \sigma(U) : \epsilon(V)\ dx, \tag{7.12}$$

where $U, V \in H_0^1(X, \mathbb{R}^d)$, to measure the distance between two shapes. In (7.12), σ denotes the stress tensor and ϵ is the strain tensor in linear elasticity, which are defined in (6.21)-(6.22).

In a first step to a definition of a suitable shape distance measure, we give the definition of a *volume elastic deformation energy* (cf. [32]).

Definition 7.7 (Volume elastic deformation energy)**.** *Let $\lambda \geq 0$ and $\mu > 0$ denote the Lamé parameters involved in the stress and strain tensor in linear elasticity (cf. (6.21)-(6.22)). Moreover, let $U \in H_0^1(X, \mathbb{R}^d)$ denote a vector field defined on X. The elastic energy of the deformation U on X is given by*

$$E(U) = a(U, U) = \int_X \sigma(U) : \epsilon(U)\ dx, \tag{7.13}$$

where σ and ϵ are defined in (6.21)-(6.22).

In the following, we consider deformation energies on the shape itself. Due to (7.4), these energies should deform a shape $\Gamma \in \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ in normal direction. This means that we have to define an elastic energy of a surface deformation $\phi \in H^{1/2}(\Gamma)$. For this purpose, we define a special kind of trace operator. Let Γ be smooth enough to admit a normal vector field n. Moreover, let $\mathrm{tr}\colon H_0^1(X, \mathbb{R}^d) \to H^{1/2}(\Gamma, \mathbb{R}^d)$ be the trace operator on the Sobolev spaces for vector valued functions restricted to Γ. The trace operator

$$\mathrm{tr}_n\colon H_0^1(X, \mathbb{R}^d) \to H^{1/2}(\Gamma),\ U \mapsto \langle \mathrm{tr}(U), n\rangle \tag{7.14}$$

defined in [32] is continuous and surjective. This means that for every $\phi \in H^{1/2}(\Gamma)$ there exists an $U \in H_0^1(X, \mathbb{R}^d)$ such that $\mathrm{tr}_n(U) = \phi$, as shown in [32].

Now, we are able to define the *surface elastic deformation energy* as the infimum of all domain elastic energies (7.13), which deform a shape Γ in normal direction. Note that X depends on Γ in our setting. To clarify this, we write X_Γ in the following.

Definition 7.8 (Surface elastic deformation energy)**.** *Let $\Gamma \in \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$. The elastic energy of a surface deformation $\phi \in H^{1/2}(\Gamma)$ is defined as*

$$E_\Gamma(\phi) \coloneqq \inf_{\substack{U \in H_0^1(X_\Gamma, \mathbb{R}^d) \\ \mathrm{tr}_n U = \phi}} E(U). \tag{7.15}$$

With this surface elastic deformation energy, which is firstly defined in [32], we obtain an elastic shape distance on $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ with respect to the Steklov-Poincaré metric g^S. For a shape variation along a path $\gamma\colon [0,1] \to \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$, which is induced by the deformation family $(U(t))_{t\in[0,1]}$, the path length L is given by

$$L(\gamma) \coloneqq \int_0^1 \sqrt{E_{\gamma(t)}(\dot\gamma(t))}\ dt \tag{7.16}$$

and the energy E is given by

$$E(\gamma) \coloneqq \int_0^1 E_{\gamma(t)}(\dot\gamma(t))\ dt, \tag{7.17}$$

where $\dot\gamma = \frac{\partial\gamma}{\partial t}$ denotes the temporal variation of γ at time $t \in [0,1]$. The temporal variation $\dot\gamma(t)$ is the velocity of $\gamma(t)$ normal to $\gamma(t)$. This is a tangent vector at $\gamma(t) \in \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ and, thus, a continuous surface deformation $\phi(t) \in H^{1/2}(\Gamma(t))$ due to (7.4).

With all this, we can define a geodesic path.

Definition 7.9 (Geodesic)**.** *A geodesic path between two shapes* $\Upsilon, \tilde{\Upsilon} \in \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ *is a curve* $\gamma\colon [0,1] \to \mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ *with* $\gamma(0) = \Upsilon$ *and* $\gamma(1) = \tilde{\Upsilon}$ *which locally minimizes the length* L *or, equivalently, the energy* E *defined in (7.16)-(7.17).*

Remark 7.10. Note that

$$\sigma(U) : \epsilon(V) = \frac{\lambda}{2} \mathrm{tr}\epsilon(U)\mathrm{tr}\epsilon(V) + \mu\, \mathrm{tr}(\epsilon(U)^T \epsilon(V)) \tag{7.18}$$

holds. As stated in [106], in an isotropic Newtonian fluid, the integrand in (7.13) is the local rate of viscous dissipation, i.e.,

$$\mathrm{diss}(U) = \sigma(U) : \epsilon(U), \tag{7.19}$$

where $U \in H_0^1(X, \mathbb{R}^d)$. It describes the rate at which the mechanical energy is locally converted into heat due to friction. More precisely, $(\mathrm{tr}\epsilon(U))^2$ measures the local change of volume and $\mathrm{tr}(\epsilon(U)^2)$ measures the local change of length induced by U. By using the notation of [106], the energy E can be expressed as the infimum over the dissipation

$$\mathrm{Diss}\left((U(t), X_\Gamma(t))_{t\in[0,1]}\right) = \int_0^1 \int_{X_\Gamma(t)} \mathrm{diss}(U(t))\; dx\, dt. \tag{7.20}$$

A geodesic defined with respect to this energy mimics the energetically optimal way to continuously deform a fluid volume as stated in [105, 106, 108].

With Definition 7.9 we reach the aim of this section. We establish the relation between the geodesics in $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ with respect to g^S and the geodesics discussed in [32] and, thus, we obtain a distance measure between shapes. The authors propose an energy of infinitesimal deformation (cf. surface elastic deformation energy) of shapes in $B_e(S^d, \mathbb{R}^{d+1})$. This energy is based on the elastic deformation energy (cf. volume elastic deformation energy), i.e., on linear elasticity. A shape metric is derived from it. Since geodesics are locally minimizers of the length or, equivalently, of the energy, they also obtain a definition of geodesics in defining the energy of infinitesimal deformation of shapes. However, they do not formulate such a definition, but it should be mentioned that it would be almost the same as Definition 7.9. The difference is that we work with the shape space $\mathcal{B}^{1/2}(\Gamma_0, \mathbb{R}^d)$ and we consider the Steklov-Poincaré metric g^S. This leads us to the above definition. In contrast, in [32], the shape space $B_e(S^d, \mathbb{R}^{d+1})$ is considered and such a definition would be obtained in defining these energies without considering a special metric. However, they give an alternative interpretation of the elastic deformation energy based on differential geometry. Moreover, the authors show that the perturbation of the metric on a shape, which is deformed by an infinitesimal deformation, is a special case of the elastic deformation energy. In our case, the above energies are justified by (6.18), which allows to reformulate the scalar product $g^S(\cdot,\cdot)$ on the surface equivalently as a symmetric and coercive bilinear form $a(\cdot,\cdot)$ for volume representations. If we choose

the inner product as the weak form of the linear elasticity equation, we obtain the elastic deformation energies and a corresponding distance measure as in [32]. Note that we also get a connection to [104, 105, 106, 108] by establishing the correlation to [32] (cf. Remark 7.4 and 7.10).

Finally, it should be mentioned that there are several open questions from a theoretical point of view. As already stated in [32], the existence of minimizers of the length (or the energy), i.e., geodesics, is an open problem. In a finite dimensional setting, the theorem of Hopf and Rinow ensures the existence of geodesics on manifolds which are complete as metric spaces. However, in our infinite dimensional setting, this theorem cannot be applied.

Chapter 8

Conclusion and outlook

This final chapter addresses the questions *"What have we done in this thesis?"* (Section 8.1) and *"Which questions remain open for further research?"* (Section 8.2).

8.1 Summary

In this thesis, we have done the following:

Chapter 2 introduces notations and background knowledge required in this thesis. All used symbols and function spaces are listed in Section 2.1. Section 2.2 is devoted to second-order PDEs and regularity results of second-order elliptic PDEs, which are relevant in Chapter 6 and 7. Finally, Section 2.3 provides some definitions and results from differential topology and geometry.

Shape spaces are the topic of **Chapter 3**. Among all shape space concepts, the space B_e of two-dimensional shapes introduced by Peter W. Michor and David Mumford is essential in this thesis. In order to prepare the discussions and investigations in Chapter 5 to 7, Section 3.1 defines this essential shape space, which is even a shape manifold. Section 3.2 is devoted to Riemannian metrics on this shape space. Special attention is paid to the first Sobolev metric, which is mainly considered in Section 3.2. In particular, an expression of the *covariant derivative associated with the first Sobolev metric* is provided in Theorem 3.3. The covariant derivative is needed later on in order to give an expression of the Riemannian shape Hessian (cf. (4.86)). For the sake of completeness, the short Section 3.3 generalizes the shape space B_e and its properties to higher dimensions.

Chapter 4 deals with shape derivatives. In Section 4.1, notations and definitions from shape calculus are provided. Special attention is paid to the material and shape derivative, which are needed throughout this thesis, and to the theorem of Correa and Seeger. In Section 4.2, we consider problems of finding interfaces between two subdomains. An elliptic (Subsection 4.2.1) and a parabolic shape interface optimization problem (Subsection 4.2.2) are introduced. Moreover, their *shape derivatives*

expressed as domain and boundary integrals are deduced. These expressions are essential in Chapter 5 to 7 because they exemplify and support the theoretical approaches established in these chapters. Section 4.3 focuses on shape optimization in the context of shape calculus. In particular, the *connection of Riemannian geometry on the shape space B_e to shape optimization* is analyzed by providing *expressions of the Riemannian shape gradient and the Riemannian shape Hessian with respect to the first Sobolev metric*. Finally, Section 4.4 provides a *volume shape derivative formula of a general shape interface problem* (cf. Theorem 4.39). In this context, a *volume shape derivative of an optimization problem constrained by the linear elasticity equation* – more precisely, shape derivative (4.107) of (4.99)-(4.101) – is given.

In **Chapter 5**, the *Riemannian geometrical point of view on shape optimization established in [84] is extended to shape optimization problems constrained by PDEs*. Section 5.1 presents a *sequential quadratic programming approach*, which is based on the idea that Riemannian shape Hessians do not differ from classical shape Hessians in the solution of a shape optimization problem and that Newton methods still converge locally quadratically if Hessian terms which are zero at the solution anyway are neglected. Subsection 5.1.1 provides a *Riemannian vector bundle framework*, which is the main tool for the development of respective Lagrange-Newton methods. This vector bundle framework is applied to the elliptic shape interface optimization problem of Subsection 4.2.1. Furthermore, *shape variants of quasi-Newton methods* are provided in Section 5.2. They are exemplified by the parabolic shape interface optimization problem of Subsection 4.2.2. Numerical results, which are presented in Section 5.3, endorse the results of the first two sections. In Subsection 5.3.1, *quadratic convergence rates of the Lagrange-Newton approach* are demonstrated. Finally, Subsection 5.3.2 shows *superlinear convergence rates of the quasi-Newton methods*.

In Chapter 4, volume integrals are converted into surface integrals. We see that a lot of effort has to be put into these conversions. Moreover, it is not always clear how a surface formulation looks like and which additional assumptions have to be made in its derivation. In Chapter 5, we see that a lot of coding work is obsolete using surface formulations of shape derivatives. Another point, which pushes us to our limits, is that the shape deformation sometimes leads to shapes, where normal vectors can no longer be reliably evaluated. Thus, **Chapter 6** is intended to enable the *usage of domain integral forms of shape derivatives in optimization strategies*. Section 6.1 discusses *generalized Steklov-Poincaré operators as basis for scalar products on shape spaces*. Section 6.2 rephrases optimization algorithms on shape spaces within the framework of volume integral formulations of shape derivatives and in the context of Steklov-Poincaré metrics, which are introduced in Section 6.1. In this manner, *quasi-Newton methods based on volume expressions of shape derivatives and Steklov-Poincaré metrics* are provided. In particular, an optimization algorithm, Algorithm 3, is formulated. Section 6.3 discusses not only algorithmic and implementation details, but also numerical results for the parabolic shape interface optimization problem given in Subsection 4.2.2. Moreover, the approach of Chapter 5 based on surface expressions of shape derivatives is compared with the approach of this chap-

ter based on volume formulations of shape derivatives from a computational point of view.

All chapters mentioned so far are concerned with $\mathcal{C}^\infty$-shapes, which limits the application of the methods proposed in the previous chapters to a certain extent. As a remedy, **Chapter 7** proposes a novel shape space $\mathcal{B}^{1/2}$ in order to extend these methods to shapes of it with respect to the shape metric g^S introduced in Section 6.1. Section 7.1 is devoted to the *shape space definition*. Moreover, its *connection to shape calculus* is given in order to formulate the shape quasi-Newton methods of Chapter 6 on this shape space. Section 7.2 gives numerical results. Finally, Section 7.3 *establishes the connection of the geodesics in* $\mathcal{B}^{1/2}$ *with respect to the Steklov-Poincaré metric* g^S *to the geodesics discussed in [32]* in order to obtain a *distance measure for shapes of* $\mathcal{B}^{1/2}$ *with respect to* g^S.

To sum up, the novelties of this thesis are as follows:

- In Chapter 3, an expression of the covariant derivative associated with the first Sobolev metric is provided (Theorem 3.3).
- In Chapter 4, not only novel shape derivatives for interface problems are developed (volume shape derivatives: Theorem 4.23, Theorem 4.30 and Theorem 4.39; surface shape derivatives: Theorem 4.24 and Theorem 4.31), but also the connection of Riemannian geometry on the shape space B_e to shape optimization is stated by providing expressions of the Riemannian shape gradient (Definition 4.35) and the Riemannian shape Hessian (Definition 4.36) with respect to the first Sobolev metric.
- In Chapter 5, the Riemannian shape calculus framework established in [84] is generalized to Lagrange-Newton and quasi-Newton approaches for PDE constrained shape optimization problems. Lagrange-Newton and quasi-Newton methods are applied in shape spaces. Two optimization algorithms, Algorithm 1 and 2, are formulated. Finally, in Section 5.3, it is shown that these approaches are viable and lead to computational methods with superior convergence properties when compared to only linearly converging standard steepest descent methods.
- In Chapter 6, the joint work of surface and volume based shape derivative expressions is enabled. As outlined in Section 6.3, this leads to a novel shape optimization algorithm, Algorithm 3, with several computational and analytic advantages.
- In Chapter 7, the definition of a novel shape space (Definition 7.1) which is not limited to objects with $\mathcal{C}^\infty$-boundary is given.

8.2 Future work

The results of this thesis leave space for several issues which have to be addressed in the future, like:

- **Analysis of the properties of $\mathcal{B}^{1/2}$:**
  In Chapter 7, the shape space $\mathcal{B}^{1/2}$ is introduced. From a theoretical point of view there are several open questions about this shape space, e.g., the *independence of the prior shape Γ_0 in the shape space definition.* Another important question is whether the *shape space has a manifold structure.* Note that this question is very hard and a lot of effort has to be put into it to find the answer. In future work, such kind of questions have to be addressed to analyze the properties of this shape space.
- **Computation of geodesics and geodesic distances:**
  The usage of the exponential map within optimization algorithms is an expensive operation. Therefore, in all optimization algorithms, we implement retractions instead of the exponential map. However, we should keep in mind that the retractions implemented in this thesis may leave the shape space B_e for large shape deformations because intersections and kinks may appear in the shape. In general, this does not happen for geodesics or the exponential map. Thus, at least, *more refined retractions* have to be developed for large shape deformations. In this context, there is another topic which has to be addressed in future work. In general, measurements of convergence rates ideally has to be performed in terms of the geodesic distance. In this thesis, we approximate the geodesic distance. An efficient way to *compute geodesics* and in this manner to *measure the convergence rates in terms of the geodesic distance* has to be established in future work.
- **Optimizing variational inequalities on shape manifolds:**
  Some PDEs arise from simplified variational inequalities (VIs). Thus, it is a natural question to ask *how we can optimize shape optimization problems constrained by VIs* of the following form:

$$\min_{\Omega} \; J(\Omega, y) \tag{8.1}$$

$$\text{s.t. } a(y, v-y) + \phi_\Omega(v) - \phi_\Omega(y) \geq (f_\Omega, v-y)_{V_\Omega' \times V_\Omega} \quad \forall v \in V_\Omega. \tag{8.2}$$

  Here the set $\Omega \subset \mathbb{R}^d$, $d = 2$ or $d = 3$, is open and connected, the state $y \in V_\Omega$ is an element of a Hilbert space V_Ω with dual space $V_\Omega' \ni f_\Omega$ and $(\cdot,\cdot)_{V_\Omega' \times V_\Omega}$ denotes the duality pairing. The bilinear form $a\colon V_\Omega \times V_\Omega \to \mathbb{R}$ is assumed to be symmetric and continuous, the function $\phi_\Omega\colon V_\Omega \to \overline{\mathbb{R}}$ is a proper convex function and the objective function J is a real valued shape functional. For a parabolic problem, a time derivative term has to be added in (8.2) in analogy to [46, Chapter 10]. The aim of future work is to *treat VI constrained shape optimization problems of the form (8.1)-(8.2) and its parabolic generalization*

from an analytical and numerical point of view by an approach aiming at *semi-smooth Newton methods on shape vector bundles*. It is not a very far step from the optimization approaches on shape spaces established in this thesis to semi-smooth Newton methods. However, note that the problem class (8.1)-(8.2) is very challenging because of the necessity to operate in inherently non-linear and non-convex shape spaces. In classical VIs, there is no explicit dependence on the domain, which adds an unavoidable source of non-linearity and non-convexity due to the non-linear and non-convex nature of shape spaces.

List of figures

Bibliography

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

[2] A. Adem, J. Leida, and Y. Ruan. *Orbifolds and Stringy Topology*. Number 171. Cambridge University Press, 2007.

[3] V.I. Agoshkov and V.I. Lebedev. Poincaré-Steklov operators and domain decomposition methods in variational problems. In G.I. Marchuk, editor, *Vychislitel'nye Protsessy i Sistemy*, volume 2, pages 173–227. Nauka, 1985.

[4] C. Allaire. *Conception Optimale de Structures*, volume 58 of *Mathématiques and Applications*. Springer, 2007.

[5] L. Ambrosio, N. Gigli, and G. Savaré. Gradient flows with metric and differentiable structures, and applications to the Wasserstein space. *Rendiconti Lincei – Matematica e Applicazioni*, 15(3-4):327–343, 2004.

[6] H.B. Ameur, M. Burger, and B. Hackl. Level set methods for geometric inverse problems in linear elasticity. *Inverse Problems*, 20(3):673–696, 2004.

[7] E. Arian. Analysis of the Hessian for aeroelastic optimization. Technical Report No. NAS1-19480, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, 1995.

[8] E. Arian and S. Ta'asan. Analysis of the Hessian for aerodynamic optimization: Inviscid flow. *Computers & Fluids*, 28(7):853–877, 1999.

[9] E. Arian and V.N. Vatsa. A preconditioning method for shape optimization governed by the Euler equations. *International Journal of Computational Fluid Dynamics*, 12(1):17–27, 1999.

[10] M. Bauer. *Almost Local Metrics on Shape Space*. PhD thesis, Universität Wien, 2010.

[11] M. Bauer, P. Harms, and P.M. Michor. Sobolev metrics on shape space of surfaces. *Journal of Geometric Mechanics*, 3(4):389–438, 2011.

[12] M. Bauer, P. Harms, and P.W. Michor. Sobolev metrics on shape space II: Weighted Sobolev metrics and almost local metrics. *Journal of Geometric Mechanics*, 4(4):365–383, 2012.

[13] M.F. Beg, M.I. Miller, A. Trouvé, and L. Younes. Computing large deformation diffeomorphic metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, 2005.

[14] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerical Mathematics*, 84(3):375–393, 2000.

[15] J.-D. Benamou, Y. Brenier, and K. Guittet. The Monge-Kantorovitch mass transfer and its computational fluid mechanics formulation. *International Journal for Numerical Methods in Fluids*, 40(1-2):21–30, 2002.

[16] M. Berggren. A unified discrete-continuous sensitivity analysis method for shape optimization. In W. Fitzgibbon et al., editors, *Applied and Numerical Partial Differential Equations*, volume 15 of *Computational Methods in Applied Sciences*, pages 25–39. Springer, 2010.

[17] F.L. Bookstein. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, 1997.

[18] J. Céa. Conception optimale ou identification de formes calcul rapide de la dérivée directionelle de la fonction coût. *RAIRO Modelisation mathématique et analyse numérique*, 20(3):371–402, 1986.

[19] V. Cervera, F. Mascaró, and P.W. Michor. The action of the diffeomorphism group on the space of immersions. *Differential Geometry and its Applications*, 1(4):391–401, 1991.

[20] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Active shape models – their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.

[21] R. Correa and A. Seeger. Directional derivative of a minmax function. *Nonlinear Analysis: Theory, Methods & Applications*, 9(1):13–22, 1985.

[22] M. Dambrine and M. Pierre. About stability of equilibrium shapes. *ESAIM: Mathematical Modelling and Numerical Analysis*, 34(4):811–834, 2000.

[23] M.C. Delfour and J.-P. Zolésio. *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization*, volume 22 of *Advances in Design and Control*. SIAM, 2nd edition, 2001.

[24] F. Demengel, G. Demengel, and R. Erné. *Functional Spaces for the Theory of Elliptic Partial Differential Equations*. Springer, 2012.

[25] M. Droske and M. Rumpf. Multi scale joint segmentation and registration of image morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12):2181–2194, 2007.

[26] S. Durrleman, X. Pennec, A. Trouvé, and N. Ayache. Statistical models of sets of curves and surfaces based on currents. *Medical Image Analysis*, 13(5):793–808, 2009.

[27] S. Durrleman, X. Pennec, A. Trouvé, P. Thompson, and N. Ayache. Inferring brain variability from diffeomorphic deformations of currents: An integrative approach. *Medical Image Analysis*, 12(5):626–637, 2008.

[28] K. Eppler and H. Harbrecht. A regularized Newton method in electrical impedance tomography using shape Hessian information. *Control and Cybernetics*, 34(1):203–225, 2005.

[29] K. Eppler, H. Harbrecht, and R. Schneider. On convergence in elliptic shape optimization. *SIAM Journal on Control and Optimization*, 46(1):61–83, 2007.

[30] K. Eppler, S. Schmidt, V.H. Schulz, and C. Ilic. Preconditioning the pressure tracking in fluid dynamics by shape Hessian information. *Journal of Optimization Theory and Applications*, 141(3):513–531, 2009.

[31] L.C. Evans. *Partial Differential Equations*. American Mathematical Society, 1993.

[32] M. Fuchs, B. Jüttler, O. Scherzer, and H. Yang. Shape metrics based on elastic deformations. *Journal of Mathematical Imaging and Vision*, 35(1):86–102, 2009.

[33] D. Gabay. Minimizing a differentiable function over a differential manifold. *Journal of Optimization Theory and Applications*, 37(2):177–219, 1982.

[34] P. Gangl, A. Laurain, H. Meftahi, and K. Sturm. Shape optimization of an electric motor subject to nonlinear magnetostatics. *SIAM Journal on Scientific Computing*, 37(6):B1002–B1025, 2015.

[35] N.R. Gauger, C. Ilic, S. Schmidt, and V.H. Schulz. Non-parametric aerodynamic shape optimization. In G. Leugering et al., editors, *Constrained Optimization and Optimal Control for Partial Differential Equations*, volume 160 of *Numerical Mathematics*, pages 289–300. Birkhäuser, 2012.

[36] J. Glaunès, A. Qui, M.I. Miller, and L. Younes. Large deformation diffeomorphic metric curve mapping. *International Journal of Computer Vision*, 80(3):317–336, 2008.

[37] S. Gross and A. Reusken. *Numerical Methods for Two-Phase Incompressible Flows*, volume 40 of *Computational Mathematics*. Springer, 2010.

[38] A. Günther, H. Lannecker, and M. Weiser. Direct LDDMM of discrete currents with adaptive finite elements. In X. Pennec, S. Joshi, and M. Nielsen, editors,

Proceedings of the Third International Workshop on Mathematical Foundations of Computational Anatomy – Geometrical and Statistical Methods for Modelling Biological Shape Variability, pages 1–14, 2011.

[39] W. Haack. Geschoßformen kleinsten Wellenwiderstandes. *Bericht der Lilienthal-Gesellschaft*, 136(1):14–28, 1941.

[40] B. Hafner, S. Zachariah, and J. Sanders. Characterization of three-dimensional anatomic shapes using principal components: Application to the proximal tibia. *Medical and Biological Engeneering and Computing*, 38(1):9–16, 2000.

[41] J. Haslinger and R.A.E. Mäkinen. *Introduction to Shape Optimization: Theory, Approximation, and Computation*, volume 7 of *Advances in Design and Control*. SIAM, 2003.

[42] M. Hintermüller and L. Laurain. Optimal shape design subject to elliptic variational inequalities. *SIAM Journal on Control and Optimization*, 49(3):1015–1047, 2011.

[43] M. Hintermüller and W. Ring. A second order shape optimization approach for image segmentation. *SIAM Journal on Applied Mathematics*, 64(2):442–467, 2004.

[44] R. Hiptmair and A. Paganini. Shape optimization by pursuing diffeomorphisms. *Computational Methods in Applied Mathematics*, 15(3):291–305, 2015.

[45] D. Holm, A. Trouvé, and L. Younes. The Euler-Poincaré theory of metamorphosis. *Quarterly of Applied Mathematics*, 67(4):661–685, 2009.

[46] K. Ito and K. Kunisch. *Lagrange Multiplier Approach to Variational Problems and Applications*, volume 15 of *Advances in Design and Control*. SIAM, 2008.

[47] K. Ito, K. Kunisch, and G.H. Peichl. Variational approach to shape derivatives. *ESAIM: Control, Optimisation and Calculus of Variations*, 14(3):517–539, 2008.

[48] D.G. Kendall. Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2):81–121, 1984.

[49] B.N. Khoromskij and G. Wittum. *Numerical Solution of Elliptic Differential Equations by Reduction to the Interface*, volume 36 of *Lecture Notes in Computational Science and Engineering*. Springer, 2004.

[50] M. Kilian, N.J. Mitra, and H. Pottmann. Geometric modelling in shape space. *ACM Transactions on Graphics*, 26(64):1–8, 2007.

[51] V. Komkov, E.J. Haug, and K.K. Choi. *Design Sensitivity Analysis of Structural Systems*, volume 177 of *Mathematics in Science and Engineering*. Academic Press, 1986.

[52] A. Kriegl and P.W. Michor. *The Convient Setting of Global Analysis*, volume 53 of *Mathematical Surveys and Monographs*. American Mathematical Society, 1997.

[53] W. Kühnel. *Differentialgeometrie: Kurven, Flächen und Mannigfaltigkeiten*. Vieweg, 4th edition, 2008.

[54] S. Kurtek, E. Klassen, Z. Ding, and A. Srivastava. A novel Riemannian framework for shape analysis of 3D objects. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1625–1632, 2010.

[55] S. Kushnarev. Teichons: Solitonlike geodesics on universal Teichmüller space. *Experimental Mathematics*, 18(3):325–336, 2009.

[56] O.A. Ladyzhenskaya, V.A. Solonnikov, and N.N. Ural'ceva. *Linear and Quasilinear Equations of Parabolic Type*. American Mathematical Society, 1968.

[57] S. Lang. *Fundamentals in Differential Geometry*, volume 191 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 2001.

[58] A. Laurain and K. Sturm. Domain expression of the shape derivative and application to electrical impedance tomography. Technical Report No. 1863, Weierstraß-Institut für angewandte Analysis und Stochastik, Berlin, 2013.

[59] J.M. Lee. *Manifolds and Differential Geometry*, volume 107 of *Graduate Studies in Mathematics*. American Mathematical Society, 2009.

[60] H. Ling and D.W. Jacobs. Shape classification using the inner-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(2):286–299, 2007.

[61] M. Meyer, M. Desbrun, P. Schröder, and A.H. Barr. Discrete differential-geometry operators for triangulated 2-manifolds. In H.-C. Hege and K. Polthier, editors, *Visualization and Mathematics III*, pages 35–57. Springer, 2003.

[62] P.M. Michor and D. Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Documenta Mathematica*, 10:217–245, 2005.

[63] P.M. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *Journal of the European Mathematical Society*, 8(1):1–48, 2006.

[64] P.M. Michor and D. Mumford. An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23(1):74–113, 2007.

[65] P.W. Michor, D. Mumford, J. Shah, and L. Younes. A metric on shape space with explicit geodesics. *Rendiconti Lincei – Matematica e Applicazioni*, 9:25–57, 2007.

[66] W. Mio, A. Srivastava, and S. Joshi. On shape of plane elastic curves. *International Journal of Computer Vision*, 73(3):307–324, 2007.

[67] B. Mohammadi and O. Pironneau. *Applied Shape Optimization for Fluids*. Oxford University Press, 2001.

[68] A. Nägel, V.H. Schulz, M. Siebenborn, and G. Wittum. Scalable shape optimization methods for structured inverse modeling in 3D diffusive processes. *Computing and Visualization in Science*, 17(2):79–88, 2015.

[69] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, 2nd edition, 2006.

[70] A. Novotny and J. Sokolowski. *Topological Derivatives in Shape Optimization*. Springer, 2013.

[71] A. Novruzi and M. Pierre. Structure of shape derivatives. *Journal of Evolution Equations*, 2(3):365–382, 2002.

[72] A. Novruzi and J.R. Roche. Newton's method in shape optimisation: A three-dimensional case. *BIT Numerical Mathematics*, 40(1):102–120, 2000.

[73] A. Paganini. Approximative shape gradients for interface problems. In A. Pratelli and G. Leugering, editors, *New Trends in Shape Optimization*, volume 166 of *International Series of Numerical Mathematics*, pages 217–227. Springer, 2015.

[74] O. Pantz. Sensibilitè de l'èquation de la chaleur aux sauts de conductivitè. *Comptes Rendus Mathematique de l'Académie des Sciences*, 341(5):333–337, 2005.

[75] D. Perperidis, R. Mohiaddin, and D. Rueckert. Construction of a 4D statistical atlas of the cardiac anatomy and its use in classification. In J. Duncan and G. Gerig, editors, *Medical Image Computing and Computer Assisted Intervention*, volume 3750 of *Lecture Notes in Computer Science*, pages 402–410. Springer, 2015.

[76] O. Pironneau. On optimum profiles in Stokes flow. *Journal of Fluid Mechanics*, 59(1):117–128, 1973.

[77] W. Ring and B. Wirth. Optimization methods on Riemannian manifolds and their application to shape space. *SIAM Journal on Control and Optimization*, 22(2):596–627, 2012.

[78] L. Risser, F.X. Vialard, M. Murgasova, D. Holm, and D. Rueckert. Large deformation diffeomorphic registration using fine and coarse strategies. In B. Fischer, B. Dawant, and C. Lorenz, editors, *Biomedical Image Registration*, volume 6204 of *Lecture Notes in Computer Science*, pages 186–197. Springer, 2010.

[79] C. Schillings, S. Schmidt, and V.H. Schulz. Efficient shape optimization for certain and uncertain aerodyamic design. *Computers & Fluids*, 46(1):78–87, 2011.

[80] S. Schmidt. *Efficient Large Scale Aerodynamic Design Based on Shape Calculus*. PhD thesis, Universiät Trier, 2010.

[81] S. Schmidt, C. Ilic, V.H. Schulz, and N.R. Gauger. Three-dimensional large-scale aerodynamic shape optimization based on shape calculus. *AIAA Journal*, 51(11):2615–2627, 2013.

[82] S. Schmidt and V.H. Schulz. Impulse response approximations of discrete shape Hessians with application in CFD. *SIAM Journal on Control and Optimization*, 48(4):2562–2580, 2009.

[83] S. Schmidt and V.H. Schulz. Shape derivatives for general objective functions and the incompressible Navier-Stokes equations. *Control and Cybernetics*, 39(3):677–713, 2010.

[84] V.H. Schulz. A Riemannian view on shape optimization. *Foundations of Computational Mathematics*, 14(3):483–501, 2014.

[85] V.H. Schulz and M. Siebenborn. Computational comparison of surface metrics for PDE constrained shape optimization. *Computational Methods in Applied Mathematics*, 16(3):485–496, 2016.

[86] V.H. Schulz, M. Siebenborn, and K. Welker. Efficient PDE constrained shape optimization based on Steklov-Poincaré type metrics. *SIAM Journal on Optimization (accepted)*, 2015.

[87] V.H. Schulz, M. Siebenborn, and K. Welker. PDE constrained shape optimization as optimization on shape manifolds. In F. Nielsen and F. Barbaresco, editors, *Geometric Science of Information*, volume 9389 of *Lecture Notes in Computer Science*, pages 499–508. Springer, 2015.

[88] V.H. Schulz, M. Siebenborn, and K. Welker. Structured inverse modeling in parabolic diffusion problems. *SIAM Journal on Control and Optimization*, 53(6):3319–3338, 2015.

[89] V.H. Schulz, M. Siebenborn, and K. Welker. Towards a Lagrange-Newton approach for PDE constrained shape optimization. In A. Pratelli and G. Leugering, editors, *New Trends in Shape Optimization*, volume 166 of *International Series of Numerical Mathematics*, pages 229–249. Springer, 2015.

[90] M.J. Sewell. *Maximum and Minimum Principles: A Unified Approach with Applications*. Cambridge University Press, 1987.

[91] E. Sharon and D. Mumford. 2D-shape analysis using conformal mapping. *International Journal of Computer Vision*, 70(1):55–75, 2006.

[92] M. Söhn, M. Birkner, D. Yan, and M. Alber. Modelling individual geometric variation based on dominant eigenmodes of organ deformation: Implementation and evaluation. *Physics in Medicine and Biology*, 50(24):5893–5908, 2005.

[93] J. Sokolowski and J.-P. Zolésio. *Introduction to Shape Optimization*, volume 16 of *Computational Mathematics*. Springer, 1992.

[94] E. Stavropoulou, M. Hojjat, and K.-U. Bletzinger. In-plane mesh regularization for node-based shape optimization problems. *Computer Methods in Applied Mechanics and Engineering*, 275:39–54, 2014.

[95] R. Stoffel. *Structural Optimization of Coupled Problems*. PhD thesis, Universiät Trier, 2013.

[96] K. Sturm. Lagrange method in shape optimization for non-linear partial differential equations: A material derivative free approach. Technical Report No. 1817, Weierstraß-Institut für angewandte Analysis und Stochastik, Berlin, 2013.

[97] K. Sturm. *On Shape Optimization with Non-Linear Partial Differential Equations*. PhD thesis, Technische Universität Berlin, 2015.

[98] K. Sturm. Shape differentiability under non-linear PDE constraints. In A. Pratelli and G. Leugering, editors, *New Trends in Shape Optimization*, volume 166 of *International Series of Numerical Mathematics*, pages 271–300. Springer, 2015.

[99] L. Tartar. *An Introduction to Sobolev Spaces and Interpolation Spaces*, volume 3 of *Lecture Notes of the Unione Matematica Italiana*. Springer, 2000.

[100] F. Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods, and Applications*, volume 112 of *Applied Mathematics*. American Mathematical Society, 2010.

[101] A. Trouvé and L. Younes. Metamorphoses through Lie group action. *Foundations of Computational Mathematics*, 5(2):173–198, 2005.

[102] R. Udawalpola and M. Berggren. Optimization of an acoustic horn with respect to efficiency and directivity. *International Journal for Numerical Methods in Engineering*, 73(11):1571–1606, 2008.

[103] M. Vaillant and J. Glaunès. Surface matching via currents. In G.E. Christensen and M. Sonka, editors, *Information Processing in Medical Imaging*, volume 3565 of *Lecture Notes in Computer Science*, pages 381–392. Springer, 2005.

[104] B. Wirth. *Variational Methods in Shape Spaces.* PhD thesis, Friedrich-Wilhelms-Universität Bonn, 2009.

[105] B. Wirth, L. Bar, M. Rumpf, and G. Sapiro. Geodesics in shape space via variational time discretization. In M. Figueiredo, J. Zerubia, and A. Jain, editors, *Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 5681 of *Lecture Notes in Computer Science*, pages 288–302. Springer, 2009.

[106] B. Wirth, L. Bar, M. Rumpf, and G. Sapiro. A continuum mechanical approach to geodesics in shape space. *International Journal of Computer Vision*, 93(3):293–318, 2011.

[107] B. Wirth and M. Rumpf. A nonlinear elastic shape averaging approach. *SIAM Journal on Imaging Sciences*, 2(3):800–833, 2009.

[108] B. Wirth and M. Rumpf. Variational methods in shape analysis. In O. Scherzer, editor, *Handbook of Mathematical Methods in Imaging*, pages 1363–1401. Springer, 2011.

[109] J. Wloka. *Partielle Differentialgleichungen.* Vieweg & Teubner, 1982.

[110] E. Zeidler. *Nonliniear Functional Analysis and its Application II/B: Nonlinear Monotone Operators.* Springer, 1990.

[111] J.-P. Zolésio. Control of moving domains, shape stabilization and variational tube formulations. *International Series of Numerical Mathematics*, 155:329–382, 2007.