JOHANNES KEPLER
UNIVERSITÄT LINZ
Netzwerk für Forschung, Lehre und Praxis

Sensitivity Analysis: The Direct and Adjoint Method

Masterarbeit

zur Erlangung des akademischen Grades

Diplomingenieur

in der Studienrichtung

Industriemathematik

Angefertigt am *Institut für Numerische Mathematik*

Betreuung:

A. Univ.-Prof. Dipl.-Ing. Dr. Walter Zulehner

Eingereicht von:

Markus Kollmann

Linz, Februar 2010

Abstract

Shape optimization is widely used in practice. The typical problem is to find the optimal shape which minimizes/maximizes a certain cost functional and satisfies some given constraints. Usually shape optimization problems are solved numerically, by some iterative method which also requires some gradient information.
There are two approaches to provide such information: the direct approach and the adjoint approach.
In this thesis the different approaches, for getting gradient information of the functional, will be presented and then the focus is to compare them. It will be shown that the adjoint approach has a great advantage, but that there also exist examples, where the adjoint technique does not work in the sense we will introduce it.

Starting with the explanation of shape optimization, the direct and adjoint approach will be introduced. Then an introduction in shape derivatives is given and some necessary examples for them are shown.
On the basis of a two-dimensional magnetostatic field problem, where the mathematical model is derived using Maxwell's equations, and is the variational formulation of a boundary value problem, the process of getting gradient information will be shown.
Finally, numerical results illustrate the two approaches.

Zusammenfassung

Formoptimierung ist ein weit verbreitetes Gebiet der Optimierung. Das Problem besteht darin, eine optimale Form (Geometrie) zu finden, die ein gegebenes Zielfunktional minimiert/maximiert und vorgegebene Nebenbedingungen erfüllt. Solche Probleme werden meistens mit iterativen Verfahren numerisch gelöst und diese Verfahren benötigen Gradienteninformation.
Um diese zu bewerkstelligen gibt es zwei Methoden: die direkte Methode und die adjungierte Methode.
In dieser Arbeit werden die beiden Zugänge, um Information über den Gradienten des Funktionals zu bekommen, aufgezeigt und dann liegt der Fokus darin, die Methoden zu vergleichen. Es wird sich herausstellen, dass der adjungierte Zugang einen großen Vorteil besitzt, es aber auch Beispiele existieren, wo er nicht anwendbar ist.

Zuerst folgt eine Erklärung der Idee der Formoptimierung und die direkte und adjungierte Methode werden erläutert. Dann wird eine Einführung in die Theorie der shape derivatives gegeben und einige wichtige Beispiele werden präsentiert.
Auf Basis eines zweidimensionalen magnetostatischen Feldproblems, bei dem das mathematische Modell mithilfe der Maxwell Gleichungen hergeleitet wird (Variationsformulierung eines Randwertproblems), werden die Methoden gezeigt.
Zum Schluss werden numerische Resultate die beiden Methoden veranschaulichen.

Acknowledgement

First of all, I would like to express my thanks to my supervisor Prof. W. Zulehner, for giving me the chance to write this thesis, for his guidance and suggestions on this work; especially, thanks for the providing me with the knowledge on shape derivatives.

I also want to thank my colleagues, especially Stefan Mühlböck and Lisa Frank for the valuable contributions.

Last but not least I want to thank my father for giving me the possibility to study and my girlfriend Sylvia for the mental assistance.

This work has been carried out at the Institute of Computational Mathematics, JKU Linz.

Contents

Chapter 1

Introduction

The thesis deals with the usage of the so-called shape derivatives in shape optimization. The whole thesis is mainly based on the books from Haslinger and Mäkinen [1], Sokolowski and Zolesio [2] and Delfour and Zolesio [3] (some parts can also be found in the book from Le Tallec and Laporte [4]).
In this books one can find the mathematical approach for introducing shape derivatives and working with it. In this thesis the concentration is about on how to calculate and work with shape derivatives and not about all the assumptions which have to be satisfied. It is assumed that everything is smooth enough, well defined and exists. So in this thesis it is shown how someone can work with shape derivatives.

The tasks of this thesis:

In this thesis we want to get an idea of the behaviour of two different approaches (direct and adjoint method) for shape optimization by analysing a two-dimensional problem. This happens in the following way:

- Use the direct method for getting gradient information of some cost functional.
- Introduce the adjoint technique and use it for the task of providing gradients.
- Compare the results of these two methods and draw conclusions.

The organization of this thesis:

- Chapter 1 is the introduction.
- In chapter 2 the idea of shape optimization is introduced and the term shape optimization problem is defined. It will turn out that gradients

have to be calculated. For this, two possible approaches are presented: the direct approach and the adjoint approach. The details of these methods will be worked out for the discrete and the continuous case.

- In chapter 3 some basic knowledge about shape derivatives is introduced, especially the definition of the change of the geometry of some bounded Lipschitz domain $\Omega \subseteq \mathbb{R}^N$. Then the term Eulerian derivative (directional derivative of a functional) is defined and some examples for this follow. So with this it will be shown how to compute the derivative with respect to the change of the geometry of domain- and boundary-integrals.

- In chapter 4 from the physical model for magnetostatics, i.e. Maxwell's equations, the mathematical model is derived, which is then the variational formulation of a boundary value problem for a partial differential equation. Then the cost functional, which describes the torque, is derived and will depend on the solution of the boundary value problem.

- In chapter 5 a variational formulation for the shape derivative is derived out of the given variational formulation of the boundary value problem from chapter 4 using the techniques introduced in chapter 3. Then the direct and adjoint approach are figured out for the (in chapter 4) derived problem.

- In chapter 6 we give a short introduction to the finite element method (FEM). After applying it to the test problem, we shortly introduce the conjugate gradient method (CG method), with which we solve the resulting linear systems.

- In chapter 7 the numerical results for the direct and adjoint method are presented and compared.

- In chapter 8 the results are summarized and conclusions about possible continuations are drawn.

Chapter 2

The Idea of Shape Optimization

In this chapter the whole idea and principle of shape optimization is introduced (as one can find it in [1] and [5]). This will lead us then to two approaches which are presented: the direct method and the adjoint method (see also [6]). There the so-called *shape derivatives* appear, which are then formally introduced in chapter 3.

2.1 Motivation

The typical aim of a shape optimization problem is to modify the shape of an object in such a way that the resulting object is optimal with respect to a certain criterion. This criterion is usually some sort of cost functional which then has to be minimized or maximized. Since any minimization problem can be easily written as a maximization problem and vice versa, only maximization problems are considered.
In most of the cases the functional itself is depending on a variable which is the solution of a state problem and given through equations.

2.2 An Optimal Shape Design Problem

Definition 2.1 Let $\mathcal{O}$ be a family of admissible domains $\Omega \subseteq \mathbb{R}^N$, Ω the domain for the state problem and $u(\Omega)$ the solution of the state problem. Then the *graph* of the optimal shape design problem is defined as the set:

$$G := \{(\Omega, u(\Omega)) | \Omega \in \mathcal{O}\}.$$

Now let $I(\Omega, u(\Omega))$ with $I : G \to \mathbb{R}$ be the cost functional and set $J(\Omega) = I(\Omega, u(\Omega))$. Then the optimal shape design problem can be written as:

$$\text{Find } \Omega^* \in \mathcal{O} : \quad J(\Omega^*) \geq J(\Omega) \text{ for all } \Omega \in \mathcal{O}. \tag{2.1}$$

If the shape is given through a parametrization we have the following definition:

Definition 2.2 Let $\mathcal{P}$ be a family of admissible parameters, $\mathcal{P} \subseteq \mathbb{R}^k$, $\Omega = \Omega(p)$ the domain for the state problem with $p \in \mathcal{P}$ and $u(p) = u(\Omega(p))$ the solution of the state problem. Then the *graph* of the optimal shape design problem for parameterized shapes is defined as the set:

$$G := \{(p, u(p)) | p \in \mathcal{P}\}.$$

Now let again $I(p, u(p))$ with $I : G \to \mathbb{R}$ be the cost functional and set $J(p) = I(p, u(p))$. Then the optimal shape design problem for parameterized shapes can be written as:

$$\text{Find } p^* \in \mathcal{P} : \quad J(p^*) \geq J(p) \text{ for all } p \in \mathcal{P}. \tag{2.2}$$

The kind and number of design parameters is depending on the problem.

In the next sections the direct and adjoint methods will be introduced in the discrete and continuous setting but the finite element method for discretization will be discussed in detail in chapter 6. So for now only some results from the finite element method are used which will be discussed in a later chapter.

2.3 The Direct and Adjoint Method in the Discrete Setting

Assume that the state problem is already discretized and reads as follows:

$$\text{Find } u \in \mathbb{R}^{N_p} : \quad K(p)u(p) = f(p), \tag{2.3}$$

*u is the stress field

where K is a sparse matrix.
Additionally, consider the shape design problem (2.2) with constraints $c_l(p, u(p)) \leq 0$, where $l = 1, ..., M$ and M is the number of constraints.

So in short the following problem has to be solved (with the setting $\hat{c}_l(p) = c_l(p, u(p))$):

$$\begin{aligned} J(p) &\to \max_{p \in \mathbb{R}^k} \\ \hat{c}_l(p) &\leq 0. \end{aligned} \tag{2.4}$$

In optimization routines the gradients of J and $\hat{c}$ are needed and they look as follows:

$$\begin{aligned} \frac{dJ}{dp_i} &= \frac{\partial I}{\partial p_i} + \frac{\partial I}{\partial u} \cdot \frac{\partial u}{\partial p_i}, \\ \frac{d\hat{c}_l}{dp_i} &= \frac{\partial c_l}{\partial p_i} + \frac{\partial c_l}{\partial u} \cdot \frac{\partial u}{\partial p_i}, \end{aligned} \tag{2.5}$$

where $i = 1, ..., k$ and $l = 1, ..., M$.

The quantities $\partial u / \partial p_i$ are called *sensitivities*.

For calculating these gradients the special structure of the state equation ($Ku = f$) is exploited and leads to two approaches: the direct method and the adjoint method.

2.3.1 The Direct Method

Differentiation of the state equation with respect to p_i gives:

$$\frac{\partial (Ku)}{\partial p_i} = \frac{\partial f}{\partial p_i}$$

which is equivalent to

$$\frac{\partial K}{\partial p_i} u + K \frac{\partial u}{\partial p_i} = \frac{\partial f}{\partial p_i}$$

or, equivalently

$$K \frac{\partial u}{\partial p_i} = \frac{\partial f}{\partial p_i} - \frac{\partial K}{\partial p_i} u. \tag{2.6}$$

The idea is now to solve (2.6) with respect to $\partial u / \partial p_i$ and use this solution for the calculation of the gradients of J and $\hat{c}_l$ in (2.5).

Therefore, for each design parameter p_i one solution of problem (2.6) is needed.

2.3.2 The Adjoint Method

We have

$$\frac{\partial u}{\partial p_i} = K^{-1} \Big(\frac{\partial f}{\partial p_i} - \frac{\partial K}{\partial p_i} u \Big) \tag{2.7}$$

Then for the calculation of the gradients of J and $\hat{c}_l$ we obtain:

$$\begin{aligned}
\frac{dJ}{dp_i} &= \frac{\partial I}{\partial p_i} + \frac{\partial I}{\partial u} \cdot K^{-1}\Big(\frac{\partial f}{\partial p_i} - \frac{\partial K}{\partial p_i}u\Big) \\
&= \frac{\partial I}{\partial p_i} + K^{-T}\frac{\partial I}{\partial u} \cdot \Big(\frac{\partial f}{\partial p_i} - \frac{\partial K}{\partial p_i}u\Big), \\
\frac{d\hat{c}_l}{dp_i} &= \frac{\partial c_l}{\partial p_i} + \frac{\partial c_l}{\partial u} \cdot K^{-1}\Big(\frac{\partial f}{\partial p_i} - \frac{\partial K}{\partial p_i}u\Big) \\
&= \frac{\partial c_l}{\partial p_i} + K^{-T}\frac{\partial c_l}{\partial u} \cdot \Big(\frac{\partial f}{\partial p_i} - \frac{\partial K}{\partial p_i}u\Big),
\end{aligned}$$

where $i = 1, ..., k$ and $l = 1, ..., M$.

Now *adjoint variables* $\lambda = (\lambda_1, ..., \lambda_{M+1})^T$ are defined by

$$\begin{aligned}
K^T\lambda_1 &= \frac{\partial I}{\partial u}, \\
K^T\lambda_{l+1} &= \frac{\partial c_l}{\partial u}.
\end{aligned} \tag{2.8}$$

The problem (2.8) is called the *adjoint state problem.*
Plugging this into the formulas for calculating the gradients of J and $\hat{c}_l$ leads to

$$\begin{aligned}
\frac{dJ}{dp_i} &= \frac{\partial I}{\partial p_i} + \lambda_1 \cdot \Big(\frac{\partial f}{\partial p_i} - \frac{\partial K}{\partial p_i}u\Big), \\
\frac{d\hat{c}_l}{dp_i} &= \frac{\partial c_l}{\partial p_i} + \lambda_{l+1} \cdot \Big(\frac{\partial f}{\partial p_i} - \frac{\partial K}{\partial p_i}u\Big).
\end{aligned} \tag{2.9}$$

Compared to the direct approach, where we need one solution of the differentiated state problem (2.6) for each design parameter p_i, here only one solution of the adjoint state problem (2.8) is needed for the objective and each restriction.

In the next section it is shown how to set up the approaches in the continuous setting.

2.4 The Direct and Adjoint Method in the Continuous Setting

Assume that the state problem is of the following form:

$$\text{Find } u \in V: \quad a(u, v) = \langle F, v \rangle, \quad \forall v \in V, \tag{2.10}$$

with an appropriate Hilbert space V, a bilinear form $a(.,.)$ and a linear form $\langle F, . \rangle = F(.)$. Or, in operator notation

$$\text{Find } u \in V: \quad Au = F. \tag{2.11}$$

Additionally, consider again problem (2.4) but for simplicity we assume that the functional and the constraints are not explicitly depending on the parameter p, i.e. $J(p) = I(u(p))$ and $\hat{c}_l(p) = c_l(u(p))$, and that there is only one design parameter:

$$\begin{aligned} J(p) &\to \max_{p \in \mathbb{R}} \\ \hat{c}_l(p) &\leq 0. \end{aligned}$$

Again the calculation of the gradients of the objective functional and the constraints is needed. In chapter 3 we will see that these gradients are linear functionals of the following form:

$$\begin{aligned} dJ &= \langle f_1, u' \rangle, \\ d\hat{c}_l &= \langle f_{l+1}, u' \rangle, \end{aligned} \tag{2.12}$$

where $l = 1, ..., M$ and u' denotes the so-called shape derivative.

2.4.1 The Direct Method

By important techniques involving shape derivatives (see chapter 3) it is possible to derive an equation for the unknown shape derivative u' out of the equation for the state problem (2.10):

$$\text{Find } u' \in V: \quad a'(u', v) = \langle F', v \rangle, \quad \forall v \in V, \tag{2.13}$$

where the bilinear form $a'(.,.)$ is derived by differentiating the bilinear form in (2.10) and the linear form $\langle F', . \rangle$ is derived by differentiating the linear form in (2.10) (see chapter 3) . Again in operator notation

$$\text{Find } u' \in V: \quad A'u' = F'. \tag{2.14}$$

The idea is now to solve (2.13) and use this solution for the calculation of the gradients of J and $\hat{c}_l$ in (2.12).

2.4.2 The Adjoint Method

The shape derivative u' is determined as the solution of the following problem:

$$\text{Find } u' \in V: \quad \langle A'u', v \rangle = \langle F', v \rangle, \quad \forall v \in V.$$

Now the adjoint variables $\lambda = (\lambda_1, ..., \lambda_{M+1})^T$ are defined straight forward by the following adjoint problem

$$\begin{aligned} \text{Find } \lambda \in (V)^{M+1}: \quad \langle (A')^* \lambda_1, w \rangle &= \langle f_1, w \rangle, \quad \forall w \in V, \\ \langle (A')^* \lambda_{l+1}, w \rangle &= \langle f_{l+1}, w \rangle, \quad \forall w \in V. \end{aligned} \tag{2.15}$$

where $(A')^*$ denotes the adjoint of A'.

As u' is an element in V we have

$$\begin{aligned} \langle (A')^* \lambda_1, u' \rangle = \langle \lambda_1, A' u' \rangle = \langle F', \lambda_1 \rangle, \\ \langle (A')^* \lambda_{l+1}, u' \rangle = \langle \lambda_{l+1}, A' u' \rangle = \langle F', \lambda_{l+1} \rangle \end{aligned}$$

and therefore

$$\begin{aligned} dJ &= \langle F', \lambda_1 \rangle, \\ d\hat{c}_l &= \langle F', \lambda_{l+1} \rangle. \end{aligned} \tag{2.16}$$

where $l = 1, ..., M$.

Both approaches can be extended to problems with more than one design parameter. Then, in the direct approach, for each parameter we have to solve one differentiated state problem for the shape derivative. In the adjoint approach only one solution of the adjoint state problem is needed for the objective and each restriction.

Chapter 3

Introduction to Shape Derivatives

In this chapter the basic knowledge in shape derivatives is presented based on the *material derivative* idea of continuum mechanics (see also [7] and [8]). This presentation will be formal, meaning that it is correct provided that all data we need are sufficiently smooth.
As seen in chapter 2, shape optimization deals with computations of derivatives of solutions to state problems as well as cost functionals with respect to shape variations. Here it will be shown, how to compute the derivative of a functional (*Eulerian derivative*) using shape derivatives.

3.1 The Geometry

We want to study the geometric change of a bounded domain $\Omega \subseteq \mathbb{R}^N$, $N \in \mathbb{N}$, with Lipschitz boundary $\partial\Omega$. Ω is thought to be a collection of material particles changing their position in time. The space occupied by them at time t will determine a new configuration Ω_t. The change in the geometry of Ω will be given by a process which deforms the initial configuration Ω.

To formalize this mathematically, let

$$T_t : \Omega \to \mathbb{R}^N, \quad t \in [0, \varepsilon)$$

be a family of transformations, which describes the motion of each material particle in the domain Ω:

$$X \in \Omega \mapsto x = T_t(X) \equiv x(t, X)$$

such that

$$T_0(\Omega) = \Omega$$

i.e., $T_0 = \text{identity}$.

The transformed geometry is given by:

$$\Omega_t = T_t(\Omega)$$

i.e., Ω_t is the image of Ω with respect to T_t.

Like in continuum mechanics we assume T_t to be a one-to-one transformation of Ω onto Ω_t, such that

$$T_t\left(\text{int}\Omega\right) = \text{int}\Omega_t, \quad T_t(\partial\Omega) = \partial\Omega_t.$$

The point X may be thought of as the *Lagrangian coordinates* while x are the *Eulerian coordinates* of a particle.

Remark 3.1 In our purposes the parameter t will not be time but describes the amount of change of the geometry.

Definition 3.2 The *Eulerian velocity field* $V(t,x)$ at the point $x(t)$ is given by:

$$V(t,x) = \frac{\partial x}{\partial t}(t, T_t^{-1}(x)).$$

By this definition one can see, that $x(t,X)$ satisfies the following initial value problem

$$\begin{aligned} \frac{d}{dt}x(t,X) &= V(t,x(t,X)) \\ x(0,X) &= X \end{aligned} \tag{3.1}$$

and, vice versa, for each $V(t,x)$ the family of transformations $T_t(V)(X)$ can be determined as solution of the problem (3.1).

3.2 Shape Functionals

In this section the so-called *Eulerian derivative* is introduced, calculated for domain and boundary integrals and the term *shape derivative* is defined.

Definition 3.3 Let $\Omega \subseteq \mathbb{R}^N$ be a bounded Lipschitz domain and let J be a functional with $\Omega \mapsto J(\Omega)$. Then the *Eulerian derivative* of the functional J at Ω in the direction of a vector field V is given by:

$$dJ(\Omega;V) = \frac{d}{dt}J(\Omega_t)\bigg|_{t=0} = \lim_{t\downarrow 0}\frac{1}{t}(J(\Omega_t) - J(\Omega)) \tag{3.2}$$

with $\Omega_t = T_t(V)(\Omega)$.

The Eulerian derivative is a directional derivative for the, in this sense often called, shape functional J.

In the next subsections the Eulerian derivative of some important examples of functionals is calculated. Let $\Gamma = \partial\Omega$.

3.2.1 Domain Integrals

Consider

$$J(\Omega) = \int_\Omega F(x, y(\Omega)(x), \nabla y(\Omega)(x))\ dx,$$

where $F : \Omega \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$, $(x, y, p) \mapsto F(x, y, p)$.

Then

$$J(\Omega_t) = \int_{\Omega_t} F(x, y(\Omega_t)(x), \nabla y(\Omega_t)(x))\ dx.$$

*change of variable

Transforming the integral to an integral over Ω (substitution rule) leads to

$$J(\Omega_t) = \int_\Omega F(T_t(x), (y(\Omega_t) \circ T_t)(x), (DT_t^{-T}(x)\nabla(y(\Omega_t) \circ T_t)(x)))\det(DT_t(x))\ dx$$

with $DT_t = (\frac{\partial T_{t,i}(V)}{\partial X_j}(X))_{i,j=1,\dots,N}$.

With (3.2) we have

$$\begin{aligned} dJ(\Omega; V) &= \frac{d}{dt} J(\Omega_t)\Big|_{t=0} \\ &= \frac{d}{dt} \int_\Omega F(T_t(x), (y(\Omega_t) \circ T_t)(x), (DT_t^{-T}(x)\nabla(y(\Omega_t) \circ T_t)(x)))\gamma(t)(x)\ dx\Big|_{t=0} \end{aligned}$$

with $\gamma(t)(x) := \det(DT_t(x))$.
It can be easily shown that $\gamma(0)(x) = 1$.

Since the domain of integration does not depend on t anymore we can interchange the order of differentiation and integration and get

$$\begin{aligned} &dJ(\Omega; V) \\ &= \int_\Omega \frac{d}{dt}\Big[F(T_t(x), (y(\Omega_t) \circ T_t)(x), (DT_t^{-T}(x)\nabla(y(\Omega_t) \circ T_t)(x)))\gamma(t)(x)\Big]\Big|_{t=0} dx. \end{aligned}$$

Product rule and chain rule for differentiation leads to

$$\begin{aligned} dJ(\Omega;V) &= \int_\Omega F(x,y(\Omega)(x),\nabla y(\Omega)(x))\gamma'(0)(x)\ dx \\ &+ \int_\Omega \nabla_x F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot V(0,x)\ dx \\ &+ \int_\Omega D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))\dot{y}(\Omega;V)(x)\ dx \\ &+ \int_\Omega D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot(\dot{\nabla y})(\Omega;V)(x)\ dx \\ &+ \int_\Omega D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\ \frac{d}{dt}\Big[DT_t^{-T}(x)\Big]\Big|_{t=0}\nabla y(\Omega)(x)\ dx \end{aligned}$$

with

$$V(0,x) := \frac{dT_t(x)}{dt}\Big|_{t=0}, \tag{3.3}$$

the material derivative

$$\dot{y}(\Omega;V)(x) := \frac{d}{dt}\Big[(y(\Omega_t)\circ T_t)(x)\Big]\Big|_{t=0} \tag{3.4}$$

and where $v\cdot w$ denotes the Euclidean inner product in $\mathbb{R}^N$. D_y (resp. D_p) denotes the derivative with respect to y (resp. p).

It can be shown that (proof: see [2, page 76])

$$\gamma'(0)(x) = \mathrm{div}V(0,x) \tag{3.5}$$

and

$$\frac{d}{dt}\Big[DT_t^{-T}(x)\Big]\Big|_{t=0} = -DV^T(0,x). \tag{3.6}$$

Under certain smoothness assumptions the material and spatial derivatives commute:

$$(\dot{\nabla y})(\Omega;V) = \nabla\dot{y}(\Omega;V).$$

With this we have

$$\begin{aligned} dJ(\Omega;V) &= \int_\Omega F(x,y(\Omega)(x),\nabla y(\Omega)(x))\mathrm{div}V(0,x)\ dx \\ &\quad + \int_\Omega \nabla_x F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot V(0,x)\ dx \\ &\quad + \int_\Omega D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))\dot{y}(\Omega;V)(x)\ dx \\ &\quad + \int_\Omega D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot\nabla\dot{y}(\Omega;V)(x)\ dx \\ &\quad - \int_\Omega D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))DV^T(0,x)\nabla y(\Omega)(x)\ dx. \end{aligned}$$

Now we introduce the so called shape derivative $y'(\Omega;V)$:

Definition 3.4 The *shape derivative* of $y(\Omega)$ in the direction V is given by:

$$y'(\Omega;V)(x) := \dot{y}(\Omega;V)(x) - \nabla y(\Omega)(x)\cdot V(0,x). \tag{3.7}$$

Using this definition we have

$$\begin{aligned} \nabla\dot{y}(\Omega;V)(x) &= \nabla y'(\Omega;V)(x) + \nabla\Big[\nabla y(\Omega)(x)\cdot V(0,x)\Big] \\ &= \nabla y'(\Omega;V)(x) + DV^T(0,x)\nabla y(\Omega)(x) + (V(0,x)\cdot\nabla)\nabla y(\Omega)(x). \end{aligned}$$

Therefore,

$$\nabla\dot{y}(\Omega;V)(x) - DV^T(0,x)\nabla y(\Omega)(x) = \nabla y'(\Omega;V)(x) + (V(0,x)\cdot\nabla)\nabla y(\Omega)(x)$$

and

$$\begin{aligned} dJ(\Omega;V) &= \int_\Omega F(x,y(\Omega)(x),\nabla y(\Omega)(x))\mathrm{div}V(0,x)\ dx \\ &\quad + \int_\Omega \nabla_x F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot V(0,x)\ dx \\ &\quad + \int_\Omega D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))y'(\Omega;V)(x)\ dx \\ &\quad + \int_\Omega D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))(V(0,x)\cdot\nabla)y(\Omega)(x)\ dx \\ &\quad + \int_\Omega D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot\nabla y'(\Omega;V)(x)\ dx \\ &\quad + \int_\Omega D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot(V(0,x)\cdot\nabla)\nabla y(\Omega)(x)\ dx. \end{aligned}$$

Since

$$\begin{aligned}
&\mathrm{div}\Big[F(x,y(x),p(x))V(x)\Big]\\
&= F(x,y(x),p(x))\mathrm{div}V(x) + D_x\Big[F(x,y(x),p(x))\Big]\cdot V(x)\\
&= F(x,y(x),p(x))\mathrm{div}V(x) + \nabla_x F(x,y(x),p(x))\cdot V(x)\\
&\quad + D_y F(x,y(x),p(x))\nabla_x y(x)\cdot V(x) + D_p F(x,y(x),p(x))D_x p(x)V(x)\\
&= F(x,y(x),p(x))\mathrm{div}V(x) + \nabla_x F(x,y(x),p(x))\cdot V(x)\\
&\quad + D_y F(x,y(x),p(x))(V(x)\cdot\nabla)y(x) + D_p F(x,y(x),p(x))\cdot(V(x)\cdot\nabla)p(x),
\end{aligned}$$

we finally obtain

$$\begin{aligned}
dJ(\Omega;V) &= \int_\Omega D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))y'(\Omega;V)(x)\ dx\\
&\quad + \int_\Omega D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot\nabla y'(\Omega;V)(x)\ dx\\
&\quad + \int_\Omega \mathrm{div}\Big[F(x,y(\Omega)(x),\nabla y(\Omega)(x))V(0,x)\Big]\ dx,
\end{aligned}$$

or, by using Gauss' theorem

$$\begin{aligned}
dJ(\Omega;V) &= \int_\Omega D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))y'(\Omega;V)(x)\ dx \qquad (3.8)\\
&\quad + \int_\Omega D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot\nabla y'(\Omega;V)(x)\ dx\\
&\quad + \int_\Gamma F(x,y(\Omega)(x),\nabla y(\Omega)(x))V(0,x)\cdot n(x)\ ds.
\end{aligned}$$

3.2.2 Boundary Integrals

Consider

$$J(\Gamma) = \int_\Gamma F(x,y(\Omega)(x),\nabla y(\Omega)(x))\ ds,$$

where $F:\Omega\times\mathbb{R}\times\mathbb{R}^N\to\mathbb{R}$, $(x,y,p)\mapsto F(x,y,p)$.

Then

$$J(\Gamma_t) = \int_{\Gamma_t} F(x,y(\Omega_t)(x),\nabla y(\Omega_t)(x))\ ds.$$

Transforming the boundary integral to an integral over Γ (substitution rule) leads to

$$J(\Gamma_t) = \int_\Gamma F(T_t(x),(y(\Omega_t)\circ T_t)(x),(DT_t^{-T}(x)\nabla(y(\Omega_t)\circ T_t)(x)))\omega(t)(x)\ ds$$

with

$$\omega(t)(x) = \gamma(t)(x)\|DT_t^{-T}(x)n(x)\|_{l^2}.$$

We have (interchanging differentiation and integration)

$$\begin{aligned}&\frac{d}{dt}J(\Gamma_t)\Big|_{t=0}\\ &= \int_\Gamma \frac{d}{dt}\Big[F(T_t(x),(y(\Omega_t)\circ T_t)(x),(DT_t^{-T}(x)\nabla(y(\Omega_t)\circ T_t)(x)))\omega(t)(x)\Big]\Big|_{t=0}\ ds.\end{aligned}$$

It can be shown that (proof: see [2, page 80])

$$\omega'(0)(x) = \mathrm{div}V(0,x) - DV(0,x)n(x)\cdot n(x). \tag{3.9}$$

Definition 3.5 Let Ω be a bounded Lipschitz domain with the boundary Γ and V a vector field. Then the *tangential divergence* is defined as

$$\mathrm{div}_\Gamma V := \mathrm{div}V - DVn\cdot n.$$

With this we have

$$\begin{aligned}\frac{d}{dt}J(\Gamma_t)\Big|_{t=0} &= \int_\Gamma F(x,y(\Omega)(x),\nabla y(\Omega)(x))\mathrm{div}_\Gamma V(0,x)\ ds\\ &+ \int_\Gamma \nabla_x F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot V(0,x)\ ds\\ &+ \int_\Gamma D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))y'(\Omega;V)(x)\ ds\\ &+ \int_\Gamma D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))V(0,x)\cdot\nabla y(\Omega)(x)\ ds\\ &+ \int_\Gamma D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot\nabla y'(\Omega;V)(x)\ ds\\ &+ \int_\Gamma D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot(V(0,x)\cdot\nabla)\nabla y(\Omega)(x)\ ds.\end{aligned}$$

Definition 3.6 The tangential gradient of y in a point of Γ is defined as:

$$\nabla_\Gamma y = \nabla y - \frac{\partial y}{\partial n}n. \tag{3.10}$$

Remark 3.7 For the tangential gradient of y the following holds:

$$\begin{aligned}\int_\Gamma \nabla_\Gamma y\cdot W\ d\Gamma &= -\int_\Gamma y\mathrm{div}_\Gamma W\ d\Gamma\\ \forall W \quad &\text{with} \quad W\cdot n = 0 \quad \text{on} \quad \Gamma.\end{aligned} \tag{3.11}$$

Using (3.10) we get

$$\begin{aligned}
\frac{d}{dt}J(\Gamma_t)\Big|_{t=0} &= \int_\Gamma F(x,y(\Omega)(x),\nabla y(\Omega)(x))\mathrm{div}_\Gamma V(0,x)\ ds \\
&+ \int_\Gamma \nabla_x F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot n(x)V(0,x)\cdot n(x)\ ds \\
&+ \int_\Gamma \nabla_\Gamma F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot V(0,x)\ ds \\
&+ \int_\Gamma D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))y'(\Omega;V)(x)\ ds \\
&+ \int_\Gamma D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))\frac{\partial y}{\partial n}(\Omega)(x)V(0,x)\cdot n(x)\ ds \\
&+ \int_\Gamma D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))V(0,x)\cdot\nabla_\Gamma y(\Omega)(x)\ ds \\
&+ \int_\Gamma D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot\nabla y'(\Omega;V)(x)\ ds \\
&+ \int_\Gamma D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot(V(0,x)\cdot n(x))\frac{\partial}{\partial n}(\nabla y(\Omega)(x))\ ds \\
&+ \int_\Gamma D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot(V(0,x)\cdot\nabla_\Gamma)\nabla y(\Omega)(x)\ ds.
\end{aligned}$$

Since

$$\begin{aligned}
&\mathrm{div}_\Gamma\Big[F(x,y(x),p(x))V(x)\Big] \\
&= \nabla_\Gamma(F(x,y(x),p(x)))\cdot V(x) + F(x,y(x),p(x))\mathrm{div}_\Gamma V(x) \\
&= \nabla_\Gamma F(x,y(x),p(x))\cdot V(x) + D_y F(x,y(x),p(x))V(x)\cdot\nabla_\Gamma y(x) \\
&\quad + D_p F(x,y(x),p(x))\cdot(V(x)\cdot\nabla_\Gamma)p(x) + F(x,y(x),p(x))\mathrm{div}_\Gamma V(x).
\end{aligned}$$

we get

$$\begin{aligned}
\frac{d}{dt}J(\Gamma_t)\Big|_{t=0} &= \int_\Gamma \nabla_x F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot n(x)V(0,x)\cdot n(x)\ ds \\
&+ \int_\Gamma D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))y'(\Omega;V)(x)\ ds \\
&+ \int_\Gamma D_y F(x,y(\Omega)(x),\nabla y(\Omega)(x))\frac{\partial y}{\partial n}(\Omega)(x)V(0,x)\cdot n(x)\ ds \\
&+ \int_\Gamma D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot\nabla y'(\Omega;V)(x)\ ds \\
&+ \int_\Gamma D_p F(x,y(\Omega)(x),\nabla y(\Omega)(x))\cdot(V(0,x)\cdot n(x))\frac{\partial}{\partial n}(\nabla y(\Omega)(x))\ ds \\
&+ \int_\Gamma \mathrm{div}_\Gamma\Big[F(x,y(\Omega)(x),\nabla y(\Omega)(x))V(0,x)\Big]\ ds.
\end{aligned}$$

Observe that (see [2, page 115])

$$\int_\Gamma \mathrm{div}_\Gamma \Big[y(\Omega)(x) V(0,x) \Big] \; ds = \int_\Gamma y(\Omega)(x) \kappa V(0,x) \cdot n(x) \; ds$$

with the mean curvature on the manifold Γ

$$\kappa = \mathrm{div}_\Gamma n(x).$$

So finally we get

$$\begin{aligned} \frac{d}{dt} J(\Gamma_t)\Big|_{t=0} \;=\; & \int_\Gamma \nabla_x F(x, y(\Omega)(x), \nabla y(\Omega)(x)) \cdot n(x) V(0,x) \cdot n(x) \; ds \\ & + \int_\Gamma D_y F(x, y(\Omega)(x), \nabla y(\Omega)(x)) y'(\Omega; V)(x) \; ds \\ & + \int_\Gamma D_y F(x, y(\Omega)(x), \nabla y(\Omega)(x)) \frac{\partial y}{\partial n}(\Omega)(x) V(0,x) \cdot n(x) \; ds \\ & + \int_\Gamma D_p F(x, y(\Omega)(x), \nabla y(\Omega)(x)) \cdot \nabla y'(\Omega; V)(x) \; ds \\ & + \int_\Gamma D_p F(x, y(\Omega)(x), \nabla y(\Omega)(x)) \cdot (V(0,x) \cdot n(x)) \frac{\partial}{\partial n} (\nabla y(\Omega)(x)) \; ds \\ & + \int_\Gamma F(x, y(\Omega)(x), \nabla y(\Omega)(x)) \kappa V(0,x) \cdot n(x) \; ds. \end{aligned} \tag{3.12}$$

Example

Consider

$$J(\Gamma) = \int_\Gamma y(\Omega)(x) \; ds$$

with $y(\Omega) : \Omega \to \mathbb{R}$.
Using (3.12) we get

$$\begin{aligned} \frac{d}{dt} J(\Gamma_t)\Big|_{t=0} \;=\; & \int_\Gamma y'(\Omega; V)(x) \\ & + \int_\Gamma \frac{\partial y}{\partial n}(\Omega)(x) V(0,x) \cdot n(x) \; ds \\ & + \int_\Gamma \kappa y(\Omega)(x) V(0,x) \cdot n(x) \; ds. \end{aligned}$$

With the example above it is shown, how to compute the gradient of functionals (Eulerian derivative) which consist of integrals.
In the next chapter the physical model will be discussed and the state equation and the cost functional will be derived.

Chapter 4

Derivation of the Physical Problem

In this chapter the mathematical model is derived from the physical model for magnetostatics, i.e. Maxwell's equations. The mathematical model is the variational formulation of a boundary value problem for a partial differential equation (PDE) (cf. [9]). The derivation is done as in [10] and [11].
Then the cost functional, which describes the torque, is derived on the basis of [12]. This functional will depend on the solution of the boundary value problem.

4.1 Physical Model

4.1.1 Maxwell's Equations

Electromagnetic phenomena can be described by Maxwell's equations (4.1) - (4.4):

$$\text{curl}\mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{4.1}$$

$$\text{curl}\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{4.2}$$

$$\text{div}\mathbf{B} = 0 \tag{4.3}$$

$$\text{div}\mathbf{D} = \rho \tag{4.4}$$

where the quantities involved are

$\mathbf{H}$	magnetic field strength
$\mathbf{E}$	electric field strength
$\mathbf{B}$	magnetic flux density
$\mathbf{D}$	electric induction density
$\mathbf{J}$	electric current density
ρ	electric charge density.

The boldface letters denote three-dimensional vector fields. All these quantities depend on the position in space $\mathbf{x} = (x_1, x_2, x_3)$ and on time t.

For the magnetic flux density $\mathbf{B}$ and the magnetic field strength $\mathbf{H}$ there exists the following relation (constitutive laws):

$$\mathbf{B} = \mu_o \mu_r (\mathbf{H} + \mathbf{H}_0) \tag{4.5}$$

with

μ_0	permeability of vacuum ($\mu_0 := 4\pi \cdot 10^{-7} \frac{Vs}{Am}$)
μ_r	relative permeability

and $\mathbf{H}_0$ given.
By neglecting the facts of hysteresis in the sense that μ_r can be represented as a function of $|\mathbf{B}|$ we get

$$\mathbf{B} = \mu_o \mu_r (|\mathbf{B}|)(\mathbf{H} + \mathbf{H}_0). \tag{4.6}$$

where $|.|$ denotes the Euclidean norm in $\mathbb{R}^3$.

Therefore, we have

$$\nu(|\mathbf{B}|)\mathbf{B} = \mathbf{H} + \mathbf{H}_0 \tag{4.7}$$

with the so-called *reluctivity*:

$$\nu(|\mathbf{B}|) \ := \frac{1}{\mu_o \mu_r (|\mathbf{B}|)}. \tag{4.8}$$

Because of (4.3) a vector potential $\mathbf{A}$ can be found such that

$$\mathbf{B} = \mathrm{curl}\mathbf{A}. \tag{4.9}$$

$\mathbf{A}$ is unique up to a gradient field.

Considering the low-frequency (quasi-static) case of electromagnetism, displacement currents are negligible in comparison with the impressed currents:

$$\left|\frac{\partial \mathbf{D}}{\partial t}\right| \ll |\mathbf{J}|.$$

If the fields $\mathbf{H}, \mathbf{B}, \mathbf{J}$ and consequently $\mathbf{A}$ are assumed to be independent of t, we obtain the following reduced set of equations, the magnetostatic formulation:

$$\begin{aligned} \text{curl}\mathbf{H} &= \mathbf{J} \\ \text{div}\mathbf{B} &= 0 \\ \nu(|\mathbf{B}|)\mathbf{B} &= \mathbf{H} + \mathbf{H}_0. \end{aligned} \tag{4.10}$$

By expressing $\mathbf{B}$ and $\mathbf{H}$ in terms of $\mathbf{A}$ we arrive at the magnetostatic vector potential formulation

$$\text{curl}[\nu(|\text{curl}\mathbf{A}|)\text{curl}\mathbf{A}] = \mathbf{J} + \text{curl}\mathbf{H}_0. \tag{4.11}$$

The material influence appears in form of the reluctivity ν. We distinguish between the case of linear materials, where ν is constant, and the more general case of nonlinear materials.
We will consider the case of linear materials and the process of shape derivatives will only be worked out for such kind of materials.

4.2 Mathematical Model

4.2.1 Reduction to 2D

We now restrict the equations to a bounded domain $\Omega \subset \mathbb{R}^2$. So we consider the magnetic field problem only in the x_1-x_2-plane. Also appropriate boundary conditions are needed: we will force the normal component of $\mathbf{B}$ to vanish (cf. [11]).

Therefore, let $\Omega \subset \mathbb{R}^2$ be a bounded domain with sufficiently smooth boundary $\Gamma = \partial\Omega$ and let $n = n(x_1, x_2)$ denote the outer unit normal vector on Γ.
We assume that the electric current density $\mathbf{J}$ is perpendicular to the magnetic field strength $\mathbf{H}$, which should lie in the x_1-x_2-plane, i.e.,

$$\mathbf{J} = \begin{pmatrix} 0 \\ 0 \\ J_3(x_1, x_2) \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} H_1(x_1, x_2) \\ H_2(x_1, x_2) \\ 0 \end{pmatrix}. \tag{4.12}$$

We additionally assume that

$$\mathbf{H}_0 = \begin{pmatrix} H_{01}(x_1, x_2) \\ H_{02}(x_1, x_2) \\ 0 \end{pmatrix}$$

and, therefore,

$$\mathrm{curl}\mathbf{H}_0 = \begin{pmatrix} 0 \\ 0 \\ \dfrac{\partial H_{02}}{\partial x_1} - \dfrac{\partial H_{01}}{\partial x_2} \end{pmatrix}. \tag{4.13}$$

Because of (4.5) $\mathbf{B}$ has the following form:

$$\mathbf{B} = \begin{pmatrix} B_1(x_1, x_2) \\ B_2(x_1, x_2) \\ 0 \end{pmatrix}.$$

On the other hand we have

$$\mathbf{B} = \begin{pmatrix} \dfrac{\partial A_3}{\partial x_2} - \dfrac{\partial A_2}{\partial x_3} \\ \dfrac{\partial A_1}{\partial x_3} - \dfrac{\partial A_3}{\partial x_1} \\ \dfrac{\partial A_2}{\partial x_1} - \dfrac{\partial A_1}{\partial x_2} \end{pmatrix}$$

since $\mathbf{B} = \mathrm{curl}\mathbf{A}$.

By comparison it follows that

$$\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2} = 0$$

and this leads to the following form for the vector potential $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} 0 \\ 0 \\ A_3(x_1, x_2) \end{pmatrix} \tag{4.14}$$

and, therefore,

$$\mathbf{B} = \mathrm{curl}\mathbf{A} = \begin{pmatrix} \dfrac{\partial A_3}{\partial x_2} \\ -\dfrac{\partial A_3}{\partial x_1} \\ 0 \end{pmatrix}. \tag{4.15}$$

In the following, A_3 will be denoted by u viewed as a function from $\Omega \subset \mathbb{R}^2 \to \mathbb{R}$:

$$u := A_3.$$

So we get

$$\begin{aligned} |\mathrm{curl}\mathbf{A}| &= |\nabla u|, \\ \mathrm{curl}[\nu(|\mathrm{curl}\mathbf{A}|)\mathrm{curl}\mathbf{A}] &= \begin{pmatrix} 0 \\ 0 \\ -\mathrm{div}[\nu(|\nabla u|)\nabla u] \end{pmatrix}. \end{aligned}$$

where $|.|$ denotes the Euclidean norm in $\mathbb{R}^2$ or $\mathbb{R}^3$.

As already mentioned we force the normal component of $\mathbf{B}$ to vanish on the whole boundary Γ. It can be shown that (cf. [11]) it is possible to replace this boundary condition by the condition

$$n \times \mathbf{A} = 0$$

which is equivalent to

$$\begin{pmatrix} n_2 A_3 \\ -n_1 A_3 \\ 0 \end{pmatrix} = 0$$

and, therefore,

$$u = 0.$$

Finally the problem is completely reduced to 2D and by summarizing we have

$$\begin{aligned} -\mathrm{div}[\nu(|\nabla u|)\nabla u] &= J_3 + \frac{\partial H_{02}}{\partial x_1} - \frac{\partial H_{01}}{\partial x_2} \quad \text{in } \Omega \subset \mathbb{R}^2, \\ u &= 0 \quad \text{on } \Gamma. \end{aligned} \tag{4.16}$$

or, equivalently

$$\begin{aligned} -\mathrm{div}[\nu(|\nabla u|)\nabla u - M_\perp] &= J_3 \quad \text{in } \Omega \subset \mathbb{R}^2, \\ u &= 0 \quad \text{on } \Gamma \end{aligned} \tag{4.17}$$

with

$$M_\perp = \begin{pmatrix} -H_{02} \\ H_{01} \end{pmatrix}.$$

4.2.2 Variational Formulation

Equation (4.17) is the classical formulation of our problem.

We now consider m different materials in the domain Ω such that between two of them there is an interface. This means that the domain is split into m non overlapping subdomains Ω_1,...,Ω_m with $\overline{\Omega} = \bigcup_{i=1}^{m} \overline{\Omega_i}$.
As already mentioned we only consider linear materials, i.e., in every subdomain Ω_i it is assumed that the corresponding ν_i is a positive constant.
The interfaces are defined the following way:

$$\Gamma_{i,j} := \partial\Omega_i \cap \partial\Omega_j, \quad i,j = 1,...,m, i \neq j$$

where Ω_i and Ω_j are the two adjacent subdomains. Note that $\Gamma_{i,j} = \Gamma_{j,i}$. Let $n_i = n_i(x_1, x_2)$ denote the outer unit normal vector to Ω_i on $\Gamma_{i,j}$ and $n_j = n_j(x_1, x_2)$ denote the outer unit normal vector to Ω_j on $\Gamma_{i,j}$.

Now let $M_\perp^{(i)}$ be the restriction of $M_\perp$ to Ω_i and $J_3^{(i)}$ the restriction of J_3 to Ω_i respectively. Then we can write the boundary value problem (4.17) in the following form:

$$\begin{aligned} -\mathrm{div}(\nu_i \nabla u - M_\perp^{(i)}) &= J_3^{(i)} \quad \text{in } \Omega_i, \quad i = 1,...,m \\ u &= 0 \quad \text{on } \Gamma. \end{aligned} \tag{4.18}$$

One can show that from the three-dimensional physical interface conditions

$$\mathbf{H}^{(i)} \times n_i + \mathbf{H}^{(j)} \times n_j = 0 \quad \text{on } \Gamma_{i,j}$$

the following interface conditions can be derived:

$$\left[\nu_i \nabla u - M_\perp^{(i)}\right] \cdot n_i + \left[\nu_j \nabla u - M_\perp^{(j)}\right] \cdot n_j = 0 \quad \text{on } \Gamma_{i,j}. \tag{4.19}$$

Multiplying the PDE from (4.18) by an arbitrary test function v with $v = 0\Big|_\Gamma$ and integrating over the domain leads to

$$\sum_{i=1}^{m} \int_{\Omega_i} \left[-\mathrm{div}(\nu_i \nabla u - M_\perp^{(i)})\right] v \; dx = \sum_{i=1}^{m} \int_{\Omega_i} J_3^{(i)} v \; dx.$$

Integration by parts on the left hand side gives

$$\sum_{i=1}^{m}\int_{\Omega_i}\big[-\operatorname{div}(\nu_i\nabla u - M_\perp^{(i)})\big]v\ dx = \sum_{i=1}^{m}\Big[\int_{\Omega_i}(\nu_i\nabla u - M_\perp^{(i)})\cdot\nabla v\ dx$$
$$-\int_{\Gamma\cap\partial\Omega_i}((\nu_i\nabla u - M_\perp^{(i)})\cdot n)v\ ds\Big] - \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{i-1}\Big[\int_{\Gamma_{i,j}}((\nu_i\nabla u - M_\perp^{(i)})\cdot n_i)v\ ds$$
$$+\int_{\Gamma_{i,j}}((\nu_j\nabla u - M_\perp^{(j)})\cdot n_j)v\ ds\Big]$$

By using the interface conditions (4.19) and the fact that $v = 0\Big|_\Gamma$ we get

$$\sum_{i=1}^{m}\int_{\Omega_i}\big[-\operatorname{div}(\nu_i\nabla u - M_\perp^{(i)})\big]v\ dx = \sum_{i=1}^{m}\int_{\Omega_i}(\nu_i\nabla u - M_\perp^{(i)})\cdot\nabla v\ dx.$$

Therefore, the linear variational formulation reads

$$\begin{aligned}\text{Find } u\in V_0 = H_0^1(\Omega) &:= \{v\in H^1(\Omega) : v = 0 \text{ on } \Gamma := \partial\Omega\}: \quad (4.20)\\ a(u,v) &= \langle F, v\rangle, \quad \forall v\in V_0,\end{aligned}$$

with

$$a(u,v) := \sum_{i=1}^{m}\int_{\Omega_i}(\nu_i\nabla u - M_\perp^{(i)})\cdot\nabla v\ dx, \tag{4.21}$$

$$\langle F, v\rangle := \sum_{i=1}^{m}\int_{\Omega_i} J_3^{(i)} v\ dx. \tag{4.22}$$

The Sobolev space[1] V_0 is equipped with the following norm

$$||u||^2_{H^1(\Omega)} := ||u||^2_{L_2(\Omega)} + |u|^2_{H^1(\Omega)}$$

where

$$\begin{aligned}||u||_{L_2(\Omega)} &= \left(\int_\Omega u^2\ dx\right)^{\frac{1}{2}}\\ |u|_{H^1(\Omega)} &= \left(\int_\Omega |\nabla u|^2\ dx\right)^{\frac{1}{2}}.\end{aligned}$$

The variational equation (4.20) can be written as an operator equation in the dual space V_0^*

$$Au = F \tag{4.23}$$

[1]For a comprehensive introduction on Sobolev spaces see [13].

with the linear operator $A : V_0 \to V_0^*$ defined by

$$\langle Au, v\rangle = a(u, v). \tag{4.24}$$

In order to ensure that the integrals appearing in (4.21) and (4.22) are well-defined, we need

- $J_3 \in L_2(\Omega)$ and
- $M_\perp \in L_2(\Omega)^2$.

4.2.3 Existence and Uniqueness

Existence and uniqueness of the solution of linear variational formulations is guaranteed by the following theorem by Lax and Milgram (cf. [14]):

Theorem 4.1 (Lax-Milgram)
Let V be a Hilbert space, $F \in V^*$ and $a(.,.) : V \times V \to \mathbb{R}$ be a bilinear form with the following properties:

1) V-elliptic, i.e.,

$$\exists \mu_1 = \text{const} > 0 : \mu_1 ||v||_V^2 \leq a(v, v), \quad \forall v \in V,$$

2) V-bounded, i.e.,

$$\exists \mu_2 = \text{const} > 0 : |a(w, v)| \leq \mu_2 ||w||_V ||v||_V, \quad \forall v, w \in V.$$

Then there exists a unique solution $u \in V$ of the variational problem

$$a(u, v) = \langle F, v\rangle, \quad \forall v \in V, \tag{4.25}$$

and we have the following two-side estimate

$$\frac{1}{\mu_2}||F||_{V^*} \leq ||u||_V \leq \frac{1}{\mu_1}||F||_{V^*}. \tag{4.26}$$

For problem (4.20) one can show that the assumptions of theorem 4.1 are satisfied (the $M_\perp$ - term is put to the right hand side) :

- $F \in V_0^*$:

$$\begin{aligned} |\langle F, v\rangle| &= \left| \sum_{i=1}^{m} \left[\int_{\Omega_i} J_3^{(i)} v \ dx + \int_{\Omega_i} M_\perp^{(i)} \cdot \nabla v \ dx \right] \right| \\ &= \left| \int_\Omega J_3 v \ dx + \int_\Omega M_\perp \cdot \nabla v \ dx \right| \\ &\leq \left| \int_\Omega J_3 v \ dx \right| + \left| \int_\Omega M_\perp \cdot \nabla v \ dx \right| \\ &\leq ||J_3||_{L_2} ||v||_{L_2} + ||M_\perp||_{L_2^2} |v|_{H^1} \\ &\leq \left[||J_3||_{L_2} + ||M_\perp||_{L_2^2} \right] ||v||_{H^1} \end{aligned}$$

- V_0-boundedness:

$$\begin{aligned} |a(w,v)| &= \left| \sum_{i=1}^{m} \int_{\Omega_i} \nu_i \nabla w \cdot \nabla v \ dx \right| \leq \max_{i=1,\ldots,m} \nu_i \left| \sum_{i=1}^{m} \int_{\Omega_i} \nabla w \cdot \nabla v \ dx \right| \\ &= \max_{i=1,\ldots,m} \nu_i \left| \int_\Omega \nabla w \cdot \nabla v \ dx \right| \\ &\leq \max_{i=1,\ldots,m} \nu_i |w|_{H^1} |v|_{H^1} \leq \max_{i=1,\ldots,m} \nu_i ||w||_{H^1} ||v||_{H^1} \end{aligned}$$

- V_0-ellipticity:

$$\begin{aligned} a(v,v) &= \sum_{i=1}^{m} \int_{\Omega_i} \nu_i \nabla v \cdot \nabla v \ dx = \sum_{i=1}^{m} \nu_i \int_{\Omega_i} |\nabla v|^2 \ dx \\ &\geq \min_{i=1,\ldots,m} \nu_i \sum_{i=1}^{m} \int_{\Omega_i} |\nabla v|^2 \ dx = \min_{i=1,\ldots,m} \nu_i \int_\Omega |\nabla v|^2 \ dx \\ &= \min_{i=1,\ldots,m} \nu_i |v|_{H^1}^2. \end{aligned}$$

  By using Friedrichs' inequality (4.27):

$$||u||_{L_2}^2 \leq c_F^2 |u|_{H^1}^2, \quad \forall v \in H_0^1(\Omega) \tag{4.27}$$

  or, equivalently

$$||u||_{H^1}^2 \leq (1 + c_F^2) |u|_{H^1}^2$$

  we get

$$a(v,v) \geq \frac{\min\limits_{i=1,\ldots,m} \nu_i}{1 + c_F^2} ||u||_{H^1}^2$$

Therefore existence and uniqueness of the solution of our problem is guaranteed.

4.3 The Cost Functional

The derivation of the functional is based on [12].

The force acting on a body in a magnetic field can be described using the Maxwell stress tensor as defined in [12, page 1]:

$$\sigma = \mu \begin{pmatrix} H_1^2 - \frac{1}{2}H^2 & H_1H_2 & H_1H_3 \\ H_2H_1 & H_2^2 - \frac{1}{2}H^2 & H_2H_3 \\ H_3H_1 & H_3H_2 & H_3^2 - \frac{1}{2}H^2 \end{pmatrix}$$

with the magnetic field strength $\mathbf{H}$, the magnetic permeability μ and $H = \sqrt{H_1^2 + H_2^2 + H_3^2}$.

Now the force acting on the surface of a body inside a closed area A can be calculated via the integral

$$F = \int_{\partial A} \sigma n \ ds$$

where n denotes the outer unit normal vector on the surface of A.

We are interested in the effect of the force acting on the rotor of an electrical motor, therefore, we consider a cross section of A (x_1-x_2-plane), which is a circular disk with radius r and boundary Γ_s.

For such a planar problem the normal stress σ_n and the tangential stress σ_t can be calculated as follows ([12, page 1-2])

$$\begin{pmatrix} \sigma_n \\ \sigma_t \end{pmatrix} = \frac{1}{2\mu_0} \begin{pmatrix} B_n^2 - B_t^2 \\ 2B_nB_t \end{pmatrix}$$

where B_n denotes the normal component, B_t the tangential component of the magnetic flux density and μ_0 is the permeability of vacuum.

For the calculation of the torque only the force in tangential direction is of importance and so the torque M is given by:

$$\begin{aligned} M &= F_t r \\ &= \int_{\Gamma_s} r\sigma_t \ ds = \frac{1}{\mu_0} \int_{\Gamma_s} rB_n(\Gamma_s)B_t(\Gamma_s) \ ds \end{aligned}$$

where $r = \sqrt{x_1^2 + x_2^2}$.

Because of

$$\mathbf{B} = \text{curl}\mathbf{A} = \begin{pmatrix} \frac{\partial u}{\partial x_2} \\ -\frac{\partial u}{\partial x_1} \\ 0 \end{pmatrix}$$

B_n, B_t can be calculated with $n = \frac{1}{\sqrt{x_1^2 + x_2^2}} \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$:

$$\begin{aligned} B_n &= \mathbf{B} \cdot n = \frac{1}{\sqrt{x_1^2 + x_2^2}} (\frac{\partial u}{\partial x_2} x_1 - \frac{\partial u}{\partial x_1} x_2) \\ B_t &= \mathbf{B} \cdot n^{\perp} = \frac{1}{\sqrt{x_1^2 + x_2^2}} (\frac{\partial u}{\partial x_2} x_2 + \frac{\partial u}{\partial x_1} x_1), \end{aligned}$$

where $n^{\perp}$ is the vector perpendicular to n, i.e., $n^{\perp} = \frac{1}{\sqrt{x_1^2 + x_2^2}} \begin{pmatrix} x_2 \\ -x_1 \\ 0 \end{pmatrix}$.

So the functional (torque) becomes

$$M = J(\Gamma) = \frac{1}{\mu_0} \int_{\Gamma_s} \nabla u(\Omega)(x)^T Q(x) \nabla u(\Omega)(x) \; ds \tag{4.28}$$

with

$$Q(x) = \frac{1}{\sqrt{x_1^2 + x_2^2}} \begin{pmatrix} -x_1 x_2 & \frac{x_1^2 - x_2^2}{2} \\ \frac{x_1^2 - x_2^2}{2} & x_1 x_2 \end{pmatrix}. \tag{4.29}$$

Remark 4.2 The functional describing the torque involves the matrix $Q(x)$ as defined in (4.29) . It is possible to replace this special matrix by any other matrix, e.g.: $Q(x) = I$, but the resulting functional will then not describe the torque anymore.

One can see that the functional depends on the solution of the boundary value problem (BVP) (4.18).

As mentioned above, we assume that the domain for the state problem is split into m non overlapping subdomains, which leads to the already discussed interfaces. So Γ_s refers to the boundary of some circular subdomain inside.

This functional has to be maximized.

Chapter 5

Shape Derivative Analysis for the Physical Problem

For doing optimization, the Eulerian derivative of the functional (4.28) has to be computed (like in chapter 3). Since it depends on the solution of the BVP, we already saw that the Eulerian derivative will then depend also on the solution of the BVP and on its shape derivative. So we need to derive a BVP for the shape derivative. How this is done is shown in this chapter.

5.1 The Eulerian Derivative of the Cost Functional

As discussed in the previous chapter the cost functional, which describes the torque, has the following form:

$$J(\Gamma) = \frac{1}{\mu_0} \int_{\Gamma_s} \nabla u(\Omega)(x)^T Q(x) \nabla u(\Omega)(x) \; ds$$

with

$$Q(x) = \frac{1}{\sqrt{x_1^2 + x_2^2}} \begin{pmatrix} -x_1 x_2 & \frac{x_1^2 - x_2^2}{2} \\ \frac{x_1^2 - x_2^2}{2} & x_1 x_2 \end{pmatrix}$$

and where u solves the following variational problem:

$$\text{Find } u \in V_0 = H_0^1(\Omega): \quad a(u,v) = \langle F, v \rangle, \quad \forall v \in V_0,$$

with

$$\begin{aligned} a(u,v) &= \sum_{i=1}^{m} \int_{\Omega_i} (\nu_i \nabla u - M_\perp^{(i)}) \cdot \nabla v \; dx, \\ \langle F, v \rangle &= \sum_{i=1}^{m} \int_{\Omega_i} J_3^{(i)} v \; dx. \end{aligned}$$

Remark 5.1 The gradient of an H^1- function is not well defined at boundaries, so the functional would not be well defined in H^1. Therefore we have to assume that the solution u is more smooth, e.g. in H^2, then the functional is well defined.

The functional can be written as

$$J(\Gamma) = \frac{1}{\mu_0} \int_{\Gamma_s} F(x, \nabla u(\Omega)(x)) \ ds$$

with

$$F(x, p) = p^T Q(x) p.$$

Now the functional is of the form of the boundary integral in subsection 3.2.2. So we can use equality (3.12) for calculating the Eulerian derivative and this yields

$$\begin{aligned} \frac{d}{dt} J(\Gamma_t)\Big|_{t=0} &= \int_{\Gamma_s} \nabla_x F(x, \nabla u(\Omega)(x)) \cdot n(x) V(0,x) \cdot n(x) \ ds \\ &+ \int_{\Gamma_s} D_p F(x, \nabla u(\Omega)(x)) \cdot \nabla u'(\Omega; V)(x) \ ds \\ &+ \int_{\Gamma_s} D_p F(x, \nabla u(\Omega)(x)) \cdot (V(0,x) \cdot n(x)) \frac{\partial}{\partial n} (\nabla u(\Omega)(x)) \ ds \\ &+ \int_{\Gamma_s} F(x, \nabla u(\Omega)(x)) \kappa V(0,x) \cdot n(x) \ ds. \end{aligned}$$

We assume that the integration curve Γ_s is constant with respect to t. Therefore, $V(0,x) \cdot n(x)$, which describes the velocity of particles on Γ_s in normal direction due to variation of design parameters, equals zero.
So finally we get

$$\begin{aligned} \frac{d}{dt} J(\Gamma_t)\Big|_{t=0} &= \frac{1}{\mu_0} \Big[\int_{\Gamma_s} \nabla u'(\Omega; V)(x)^T Q(x) \nabla u(\Omega)(x) \ ds \\ &+ \int_{\Gamma_s} \nabla u(\Omega)(x)^T Q(x) \nabla u'(\Omega; V)(x) \ ds \Big]. \end{aligned} \tag{5.1}$$

One can see that the derivative of the functional depends on the solution of the BVP (4.18) and its shape derivative.

For test purposes we consider a second functional of the following form:

$$J(\Gamma) = \int_{\Gamma_s} u^2(\Omega)(x) \ ds \tag{5.2}$$

Again we can use equality (3.12) for calculating the Eulerian derivative and this yields

$$\frac{d}{dt}J(\Gamma_t)\bigg|_{t=0} = 2\int_{\Gamma_s} u'(\Omega;V)(x)u(\Omega)(x)\ ds. \tag{5.3}$$

5.2 The Boundary Value Problem for the Shape Derivative

The variational formulation of the boundary value problem (4.18) reads

$$\text{Find } u \in V_0 = H_0^1(\Omega): \quad a(u,v) = \langle F, v\rangle, \quad \forall v \in V_0,$$

with

$$\begin{aligned} a(u,v) &= \sum_{i=1}^{m}\int_{\Omega_i} (\nu_i \nabla u - M_\perp^{(i)})\cdot\nabla v\ dx, \\ \langle F, v\rangle &= \sum_{i=1}^{m}\int_{\Omega_i} J_3^{(i)} v\ dx. \end{aligned}$$

With the setting

$$y_i(\Omega) = \begin{pmatrix} y_{i,1}(\Omega) \\ y_{i,2}(\Omega) \\ y_{i,3}(\Omega) \end{pmatrix} = \begin{pmatrix} u(\Omega) \\ v(\Omega) \\ M_\perp^{(i)}(\Omega) \end{pmatrix}$$

we can write

$$a(u,v) = \sum_{i=1}^{m}\int_{\Omega_i} F_i(y_i(\Omega)(x), \nabla y_i(\Omega)(x)) \tag{5.4}$$

with

$$F_i(y_i, p_i) = (\nu_i p_{i,1} - y_{i,3})\cdot p_{i,2}. \tag{5.5}$$

Now $a(u,v)$ is of the form of the domain integral in subsection 3.2.1. So we can use equality (3.8) for differentiating $a(u,v)$ and this yields

$$
\begin{aligned}
da(u,v) \;=\; & \sum_{i=1}^{m} \Big[\int_{\Omega_i} D_{y_{i,3}} F_i(y_i(\Omega)(x), \nabla y_i(\Omega)(x)) \cdot M_{\perp}^{(i)\prime}(\Omega;V)(x) \; dx \\
& + \int_{\Omega_i} D_{p_{i,1}} F_i(y_i(\Omega)(x), \nabla y_i(\Omega)(x)) \cdot \nabla u'(\Omega;V)(x) \; dx \\
& + \int_{\Omega_i} D_{p_{i,2}} F_i(y_i(\Omega)(x), \nabla y_i(\Omega)(x)) \cdot \nabla v'(\Omega;V)(x) \; dx \\
& + \int_{\Gamma \cap \partial\Omega_i} F_i(y_i(\Omega)(x), \nabla y_i(\Omega)(x)) V(0,x) \cdot n(x) \; ds \Big] \\
& + \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j\neq i}}^{i-1} \Big[\int_{\Gamma_{i,j}} F_i(y_i(\Omega)(x), \nabla y_i(\Omega)(x)) V(0,x) \cdot n_i(x) \; ds \\
& + \int_{\Gamma_{i,j}} F_j(y_j(\Omega)(x), \nabla y_j(\Omega)(x)) V(0,x) \cdot n_j(x) \; ds \Big].
\end{aligned}
$$

v is an arbitrary test function from $H_0^1(\Omega)$ and we take v_t in the form $v_t = v \circ T_t^{-1} \in H_0^1(\Omega_t)$. Then the material derivative $\dot{v}$ vanishes in Ω and therefore we have for the shape derivative v'

$$
v'(\Omega;V)(x) = -\nabla v(\Omega)(x) \cdot V(0,x) \qquad \text{in } \Omega.
$$

Continuing with $da(u,v)$ we have

$$
\begin{aligned}
da(u,v) = & \sum_{i=1}^{m} \Big[- \int_{\Omega_i} \nabla v(\Omega)(x) \cdot M_{\perp}^{(i)\prime}(\Omega;V)(x) \; dx \\
& + \int_{\Omega_i} \nu_i \nabla u'(\Omega;V)(x) \cdot \nabla v(\Omega)(x) \; dx \\
& - \int_{\Omega_i} \Big[\nu_i \nabla u(\Omega)(x) - M_{\perp}^{(i)}(\Omega)(x) \Big] \cdot \nabla (\nabla v(\Omega)(x) \cdot V(0,x)) \; dx \\
& + \int_{\Gamma \cap \partial\Omega_i} \Big(\Big[\nu_i \nabla u(\Omega)(x) - M_{\perp}^{(i)}(\Omega)(x) \Big] \cdot \nabla v(\Omega)(x) \Big) (V(0,x) \cdot n(x)) \; ds \Big] \\
& + \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j\neq i}}^{i-1} \Big[\int_{\Gamma_{i,j}} \Big(\Big[\nu_i \nabla u(\Omega)(x) - M_{\perp}^{(i)}(\Omega)(x) \Big] \cdot \nabla v(\Omega)(x) \Big) (V(0,x) \cdot n_i(x)) \; ds \\
& + \int_{\Gamma_{i,j}} \Big(\Big[\nu_j \nabla u(\Omega)(x) - M_{\perp}^{(j)}(\Omega)(x) \Big] \cdot \nabla v(\Omega)(x) \Big) (V(0,x) \cdot n_j(x)) \; ds \Big].
\end{aligned}
$$

By introducing the flux

$$
\Phi_i(\nabla u)(\Omega)(x) = \nu_i \nabla u(\Omega)(x) - M_{\perp}^{(i)}(\Omega)(x)
$$

we can further write

$$
\begin{aligned}
da(u,v) &= \sum_{i=1}^{m}\Bigg[-\int_{\Omega_i} \nabla v(\Omega)(x)\cdot M_{\perp}^{(i)\prime}(\Omega;V)(x)\ dx \\
&\quad + \int_{\Omega_i} \nu_i \nabla u'(\Omega;V)(x)\cdot\nabla v(\Omega)(x)\ dx \\
&\quad - \int_{\Omega_i} \Phi_i(\nabla u)(\Omega)(x)\cdot\nabla(\nabla v(\Omega)(x)\cdot V(0,x))\ dx \\
&\quad + \int_{\Gamma\cap\partial\Omega_i} (\Phi_i(\nabla u)(\Omega)(x)\cdot\nabla v(\Omega)(x))\,(V(0,x)\cdot n(x))\ ds\Bigg] \\
&\quad + \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{i-1}\Bigg[\int_{\Gamma_{i,j}} (\Phi_i(\nabla u)(\Omega)(x)\cdot\nabla v(\Omega)(x))\,(V(0,x)\cdot n_i(x))\ ds \\
&\quad + \int_{\Gamma_{i,j}} (\Phi_j(\nabla u)(\Omega)(x)\cdot\nabla v(\Omega)(x))\,(V(0,x)\cdot n_j(x))\ ds\Bigg].
\end{aligned}
$$

By doing integration by parts in the third Ω_i-integral we get

$$
\begin{aligned}
da(u,v) &= \sum_{i=1}^{m}\Bigg[\int_{\Omega_i}\Big[\nu_i\nabla u'(\Omega;V)(x) - M_{\perp}^{(i)\prime}(\Omega;V)(x)\Big]\cdot\nabla v(\Omega)(x)\ dx \\
&\quad + \int_{\Omega_i} \mathrm{div}\Phi_i(\nabla u)(\Omega)(x)(\nabla v(\Omega)(x)\cdot V(0,x))\ dx + \int_{\Gamma\cap\partial\Omega_i} I_i\ ds\Bigg] \\
&\quad + \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{i-1}\Bigg[\int_{\Gamma_{i,j}} I_{i,j}\ ds + \int_{\Gamma_{i,j}} I_{j,i}\ ds\Bigg].
\end{aligned}
$$

where

$$
\begin{aligned}
I_i &:= (\Phi_i(\nabla u)(\Omega)(x)\cdot\nabla v(\Omega)(x))\,(V(0,x)\cdot n(x)) \\
&\quad -(V(0,x)\cdot\nabla v(\Omega)(x))\,(\Phi_i(\nabla u)(\Omega)(x)\cdot n(x))\,, \\
I_{i,j} &:= (\Phi_i(\nabla u)(\Omega)(x)\cdot\nabla v(\Omega)(x))\,(V(0,x)\cdot n_i(x)) \\
&\quad -(V(0,x)\cdot\nabla v(\Omega)(x))\,(\Phi_i(\nabla u)(\Omega)(x)\cdot n_i(x))\,, \\
I_{j,i} &:= (\Phi_j(\nabla u)(\Omega)(x)\cdot\nabla v(\Omega)(x))\,(V(0,x)\cdot n_j(x)) \\
&\quad -(V(0,x)\cdot\nabla v(\Omega)(x))\,(\Phi_j(\nabla u)(\Omega)(x)\cdot n_j(x))\,.
\end{aligned}
$$

Also the derivative with respect to t of the linear form can be computed using equation (3.8):

$$
\begin{aligned}
d\langle F, v\rangle &= \sum_{i=1}^{m}\Bigg[\int_{\Omega_i} J_3^{(i)\prime}(\Omega;V)(x)v(\Omega)(x)\ dx + \int_{\Omega_i} J_3^{(i)}(\Omega)(x)v'(\Omega;V)(x)\ dx \\
&\quad + \int_{\Gamma\cap\partial\Omega_i} J_3^{(i)}(\Omega)(x)v(\Omega)(x)(V(0,x)\cdot n(x))\ ds\Bigg] \\
&\quad + \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{i-1}\Bigg[\int_{\Gamma_{i,j}} J_3^{(i)}(\Omega)(x)v(\Omega)(x)(V(0,x)\cdot n_i(x))\ ds \\
&\quad + \int_{\Gamma_{i,j}} J_3^{(j)}(\Omega)(x)v(\Omega)(x)(V(0,x)\cdot n_j(x))\ ds\Bigg] \\
&= \sum_{i=1}^{m}\Bigg[\int_{\Omega_i}\left[J_3^{(i)\prime}(\Omega;V)(x)v(\Omega)(x) - J_3^{(i)}(\Omega)(x)(\nabla v(\Omega)(x)\cdot V(0,x))\right] dx \\
&\quad + \int_{\Gamma\cap\partial\Omega_i} J_3^{(i)}(\Omega)(x)v(\Omega)(x)(V(0,x)\cdot n(x))\ ds\Bigg] \\
&\quad + \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{i-1}\Bigg[\int_{\Gamma_{i,j}} J_3^{(i)}(\Omega)(x)v(\Omega)(x)(V(0,x)\cdot n_i(x))\ ds \\
&\quad + \int_{\Gamma_{i,j}} J_3^{(j)}(\Omega)(x)v(\Omega)(x)(V(0,x)\cdot n_j(x))\ ds\Bigg].
\end{aligned}
$$

This leads to the following variational problem for the shape derivative $u'(\Omega)(x)$:

$$
\begin{aligned}
&\text{Find } u' \in V_0 = H_0^1(\Omega) : \\
&\sum_{i=1}^{m} \Bigg[\int_{\Omega_i} \Big[\nu_i \nabla u'(\Omega; V)(x) - M_{\perp}^{(i)\prime}(\Omega; V)(x) \Big] \cdot \nabla v(\Omega)(x) \; dx \\
&+ \int_{\Omega_i} \mathrm{div} \Phi_i(\nabla u)(\Omega)(x) (\nabla v(\Omega)(x) \cdot V(0,x)) \; dx + \int_{\Gamma \cap \partial\Omega_i} I_i \; ds \Bigg] \\
&+ \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq i}}^{i-1} \Bigg[\int_{\Gamma_{i,j}} I_{i,j} \; ds + \int_{\Gamma_{i,j}} I_{j,i} \; ds \Bigg] \\
&= \sum_{i=1}^{m} \Bigg[\int_{\Omega_i} \Big[J_3^{(i)\prime}(\Omega; V)(x) v(\Omega)(x) - J_3^{(i)}(\Omega)(x) (\nabla v(\Omega)(x) \cdot V(0,x)) \Big] \; dx \\
&+ \int_{\Gamma \cap \partial\Omega_i} J_3^{(i)}(\Omega)(x) v(\Omega)(x) (V(0,x) \cdot n(x)) \; ds \Bigg] \\
&+ \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq i}}^{i-1} \Bigg[\int_{\Gamma_{i,j}} J_3^{(i)}(\Omega)(x) v(\Omega)(x) (V(0,x) \cdot n_i(x)) \; ds \\
&+ \int_{\Gamma_{i,j}} J_3^{(j)}(\Omega)(x) v(\Omega)(x) (V(0,x) \cdot n_j(x)) \; ds \Bigg], \quad \forall v \in V_0.
\end{aligned}
$$

We can further simplify this variational equation:

Since (cf. (4.18))

$$
-\mathrm{div} \Phi_i(\nabla u) = J_3^{(i)} \qquad \text{in } \Omega_i, \quad i = 1, ..., m
$$

and $v = 0\Big|_{\Gamma}$ we obtain

$$
\begin{aligned}
&\sum_{i=1}^{m} \int_{\Omega_i} \Big[\nu_i \nabla u'(\Omega; V)(x) - M_{\perp}^{(i)\prime}(\Omega; V)(x) \Big] \cdot \nabla v(\Omega)(x) \; dx \\
&= \sum_{i=1}^{m} \Bigg[\int_{\Omega_i} J_3^{(i)\prime}(\Omega; V)(x) v(\Omega)(x) \; dx - \int_{\Gamma \cap \partial\Omega_i} I_i \; ds \Bigg] \\
&+ \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq i}}^{i-1} \Bigg[\int_{\Gamma_{i,j}} J_3^{(i)}(\Omega)(x) v(\Omega)(x) (V(0,x) \cdot n_i(x)) \; ds \\
&+ \int_{\Gamma_{i,j}} J_3^{(j)}(\Omega)(x) v(\Omega)(x) (V(0,x) \cdot n_j(x)) \; ds - \int_{\Gamma_{i,j}} I_{i,j} \; ds - \int_{\Gamma_{i,j}} I_{j,i} \; ds \Bigg].
\end{aligned}
$$

Now we further discuss the I_i - term:

$$\begin{aligned}
&\int_{\Gamma\cap\partial\Omega_i} I_i \ ds \\
&= \int_{\Gamma\cap\partial\Omega_i} \left(\Phi_i(\nabla u)(\Omega)(x)\cdot\nabla v(\Omega)(x)\right)\left(V(0,x)\cdot n(x)\right) \ ds \\
&- \int_{\Gamma\cap\partial\Omega_i} \left(V(0,x)\cdot\nabla v(\Omega)(x)\right)\left(\Phi_i(\nabla u)(\Omega)(x)\cdot n(x)\right) \ ds,
\end{aligned}$$

and by using the definition of the tangential gradient (3.10)

$$\begin{aligned}
&\int_{\Gamma\cap\partial\Omega_i} I_i \ ds \\
&= \int_{\Gamma\cap\partial\Omega_i} \left(\Phi_i(\nabla u)(\Omega)(x)\cdot\nabla_\Gamma v(\Omega)(x)\right)\left(V(0,x)\cdot n(x)\right) \ ds \\
&- \int_{\Gamma\cap\partial\Omega_i} \left(V(0,x)\cdot\nabla_\Gamma v(\Omega)(x)\right)\left(\Phi_i(\nabla u)(\Omega)(x)\cdot n(x)\right) \ ds \\
&+ \int_{\Gamma\cap\partial\Omega_i} \left(\Phi_i(\nabla u)(\Omega)(x)\cdot\frac{\partial v}{\partial n}n(x)\right)\left(V(0,x)\cdot n(x)\right) \ ds \\
&- \int_{\Gamma\cap\partial\Omega_i} \left(V(0,x)\cdot\frac{\partial v}{\partial n}n(x)\right)\left(\Phi_i(\nabla u)(\Omega)(x)\cdot n(x)\right) \ ds \\
&= \int_{\Gamma\cap\partial\Omega_i} \left(\Phi_i(\nabla u)(\Omega)(x)\cdot\nabla_\Gamma v(\Omega)(x)\right)\left(V(0,x)\cdot n(x)\right) \ ds \\
&- \int_{\Gamma\cap\partial\Omega_i} \left(V(0,x)\cdot\nabla_\Gamma v(\Omega)(x)\right)\left(\Phi_i(\nabla u)(\Omega)(x)\cdot n(x)\right) \ ds \\
&= \int_{\Gamma\cap\partial\Omega_i} \Big[\left(V(0,x)\cdot n(x)\right)\Phi_i(\nabla u)(\Omega)(x) \\
&- \left(\Phi_i(\nabla u)(\Omega)(x)\cdot n(x)\right)V(0,x)\Big]\cdot\nabla_\Gamma v(\Omega)(x) \ ds.
\end{aligned}$$

Additionally we can do integration by parts on Γ (cf. (3.11)) and this yields

$$\begin{aligned}
&\int_{\Gamma\cap\partial\Omega_i} I_i \ ds \\
&= -\int_{\Gamma\cap\partial\Omega_i} \mathrm{div}_\Gamma\Big[\left(V(0,x)\cdot n(x)\right)\Phi_i(\nabla u)(\Omega)(x) \\
&- \left(\Phi_i(\nabla u)(\Omega)(x)\cdot n(x)\right)V(0,x)\Big]v(\Omega)(x) \ ds = 0.
\end{aligned}$$

The last equality is again due to $v \in H_0^1(\Omega)$.

Now we can do similar computations for the $I_{i,j}$ - and $I_{j,i}$ - term and get

$$\begin{aligned}
&\int_{\Gamma_{i,j}} I_{i,j} \; ds + \int_{\Gamma_{i,j}} I_{j,i} \; ds \\
&= \int_{\Gamma_{i,j}} \Big[(V(0,x) \cdot n_i(x)) \Phi_i(\nabla u)(\Omega)(x) \\
&- (\Phi_i(\nabla u)(\Omega)(x) \cdot n_i(x)) \, V(0,x) \Big] \cdot \nabla_\Gamma v(\Omega)(x) \; ds \\
&+ \int_{\Gamma_{i,j}} \Big[(V(0,x) \cdot n_j(x)) \Phi_j(\nabla u)(\Omega)(x) \\
&- (\Phi_j(\nabla u)(\Omega)(x) \cdot n_j(x)) \, V(0,x) \Big] \cdot \nabla_\Gamma v(\Omega)(x) \; ds.
\end{aligned}$$

By using the interface conditions (4.19) this simplifies to

$$\begin{aligned}
&\int_{\Gamma_{i,j}} I_{i,j} \; ds + \int_{\Gamma_{i,j}} I_{j,i} \; ds \\
&= \int_{\Gamma_{i,j}} (V(0,x) \cdot n_i(x)) \Phi_i(\nabla u)(\Omega)(x) \cdot \nabla_\Gamma v(\Omega)(x) \; ds \\
&+ \int_{\Gamma_{i,j}} (V(0,x) \cdot n_j(x)) \Phi_j(\nabla u)(\Omega)(x) \cdot \nabla_\Gamma v(\Omega)(x) \; ds \\
&= \int_{\Gamma_{i,j}} (V(0,x) \cdot n_i(x)) \Big[\nu_i \nabla u(\Omega)(x) - M_\perp^{(i)}(\Omega)(x) \Big] \cdot \nabla_\Gamma v(\Omega)(x) \; ds \\
&+ \int_{\Gamma_{i,j}} (V(0,x) \cdot n_j(x)) \Big[\nu_j \nabla u(\Omega)(x) - M_\perp^{(j)}(\Omega)(x) \Big] \cdot \nabla_\Gamma v(\Omega)(x) \; ds.
\end{aligned}$$

Again we use the definition of the tangential gradient and obtain

$$\begin{aligned}
&\int_{\Gamma_{i,j}} I_{i,j} \; ds + \int_{\Gamma_{i,j}} I_{j,i} \; ds \\
&= \int_{\Gamma_{i,j}} (V(0,x) \cdot n_i(x)) \Big[\nu_i \nabla_\Gamma u(\Omega)(x) - M_\perp^{(i)}(\Omega)(x) \Big] \cdot \nabla_\Gamma v(\Omega)(x) \; ds \\
&+ \int_{\Gamma_{i,j}} (V(0,x) \cdot n_j(x)) \Big[\nu_j \nabla_\Gamma u(\Omega)(x) - M_\perp^{(j)}(\Omega)(x) \Big] \cdot \nabla_\Gamma v(\Omega)(x) \; ds.
\end{aligned}$$

Note that the term with the normal component of ∇u vanishes.

Summarizing the variational problem for the shape derivative reads:

$$\text{Find } u' \in V_0 = H_0^1(\Omega): \quad a'(u', v) = \langle F', v \rangle, \quad \forall v \in V_0, \tag{5.6}$$

with

$$
\begin{aligned}
a'(u',v) &:= \sum_{i=1}^{m} \int_{\Omega_i} \nu_i \nabla u'(\Omega;V)(x) \cdot \nabla v(\Omega)(x)\ dx, \\
\langle F', v\rangle &:= \sum_{i=1}^{m} \int_{\Omega_i} \Big[J_3^{(i)\prime}(\Omega;V)(x) v(\Omega)(x) + M_\perp^{(i)\prime}(\Omega;V)(x) \cdot \nabla v(\Omega)(x) \Big]\ dx \\
&\quad + \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j\neq i}}^{i-1} \Big[\int_{\Gamma_{i,j}} J_3^{(i)}(\Omega)(x) v(\Omega)(x) (V(0,x)\cdot n_i(x))\ ds \\
&\quad + \int_{\Gamma_{i,j}} J_3^{(j)}(\Omega)(x) v(\Omega)(x) (V(0,x)\cdot n_j(x))\ ds \\
&\quad - \int_{\Gamma_{i,j}} (V(0,x)\cdot n_i(x)) \Big[\nu_i \nabla_\Gamma u(\Omega)(x) - M_\perp^{(i)}(\Omega)(x) \Big] \cdot \nabla_\Gamma v(\Omega)(x)\ ds \\
&\quad - \int_{\Gamma_{i,j}} (V(0,x)\cdot n_j(x)) \Big[\nu_j \nabla_\Gamma u(\Omega)(x) - M_\perp^{(j)}(\Omega)(x) \Big] \cdot \nabla_\Gamma v(\Omega)(x)\ ds \Big].
\end{aligned}
$$

We discussed two types of functionals: the one which describes the torque (4.28) and the one which was introduced for test purposes (5.2). In the following subsections it is shown, how the direct and adjoint method work for this two functionals.

5.2.1 The Direct Approach

In chapter 2 we already saw how the direct method works in the continuous setting:

First the solution of the variational problem (4.20) for u is computed. Then the solution of the variational problem (5.6) for the shape derivative u' is computed. Now the Eulerian derivative of the two functionals (5.1) and (5.3) can be evaluated.

5.2.2 The Adjoint Approach

In chapter 2 we also saw how the adjoint method works in the continuous setting:

First the solution of the variational problem (4.20) for u is computed. Then we have to set up one adjoint problem for each of the two functionals:

- $J(\Gamma) = \dfrac{1}{\mu_0} \displaystyle\int_{\Gamma_s} \nabla u(\Omega)(x)^T Q(x) \nabla u(\Omega)(x)\ ds$:

We define the adjoint variable λ as solution of the following problem (the bilinear form is symmetric):

$$\text{Find } \lambda \in V_0 = H_0^1(\Omega): \quad a'(\lambda, w) = \langle G, w \rangle, \quad \forall w \in V_0,$$

with

$$\begin{aligned} a'(\lambda, w) &= \sum_{i=1}^{m} \int_{\Omega_i} \nu_i \nabla\lambda(\Omega)(x) \cdot \nabla w(\Omega)(x) \; dx, \\ \langle G, w \rangle &:= \frac{1}{\mu_0} \Big[\int_{\Gamma_s} \nabla w(\Omega)(x)^T Q(x) \nabla u(\Omega)(x) \; ds \\ &\quad + \int_{\Gamma_s} \nabla u(\Omega)(x)^T Q(x) \nabla w(\Omega)(x) \; ds \Big]. \end{aligned}$$

where $\langle G, w \rangle$ is the Eulerian derivative of the cost functional with w instead of u'.

Remark 5.2 Since w is a test function in H^1, the right hand side of the variational equation is not well defined in H^1 and this means that the adjoint method **does not make sense** in H^1 for such kind of functionals.

Therefore, we cannot compare the direct and adjoint method when we use this type of functional.

- $J(\Gamma) = \int_{\Gamma_s} u^2(\Omega)(x) \; ds$:

  We define the adjoint variable λ as solution of the following problem (the bilinear form is symmetric):

$$\text{Find } \lambda \in V_0 = H_0^1(\Omega): \quad a'(\lambda, w) = \langle G, w \rangle, \quad \forall w \in V_0, \tag{5.7}$$

  with

$$\begin{aligned} a'(\lambda, w) &= \sum_{i=1}^{m} \int_{\Omega_i} \nu_i \nabla\lambda(\Omega)(x) \cdot \nabla w(\Omega)(x) \; dx, \\ \langle G, w \rangle &:= 2 \int_{\Gamma_s} w(\Omega)(x) u(\Omega)(x) \; ds. \end{aligned}$$

  where $\langle G, w \rangle$ is now the Eulerian derivative of the second functional with w instead of u'.

  This variational problem is well defined in H^1 and we can compare the direct and adjoint method when we use this type of functional.

Then we have to solve the adjoint problem (5.7) with respect to λ. Since $u' \in V_0$ and (5.7) holds for all functions in V_0 we have for the Eulerian derivative

$$\frac{d}{dt} J(\Gamma_t)\Big|_{t=0} = \langle G, u' \rangle = a'(\lambda, u') = \langle F', \lambda \rangle \,.$$

For comparing the two different approaches, the following test problem was considered:

5.3 The Test Problem

The domain Ω is split into two subdomains Ω_1 and Ω_2 as follows:

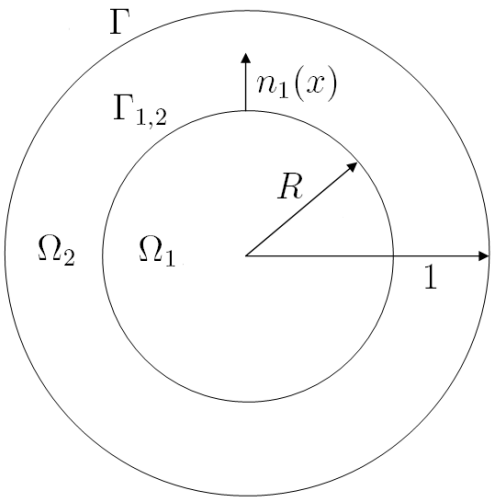

Figure 5.1: The test problem

where $n_1(x)$ is the outer unit normal vector to Ω_1 at the interface $\Gamma_{1,2}$ and $-n_1(x)$ is the outer unit normal vector to Ω_2 at the interface $\Gamma_{1,2}$. Further we set $\nu_1 = \nu_2 = 1$ and the restrictions of the electric current density to the subdomains $J_3^{(1)} = 0$, $J_3^{(2)} = 1$ and, therefore, $J_3^{(1)\prime} = J_3^{(2)\prime} = 0$. We additionally assume that the magnetic field $\mathbf{H}_0$ is 0 and, therefore, $M_\perp = 0$. So the boundary value problem (4.18) reads:

$$\begin{aligned} -\mathrm{div}(\nabla u) &= J_3^{(i)} \quad \text{in } \Omega_i, \quad i = 1, ..., m \\ u &= 0 \quad \text{on } \Gamma \end{aligned} \tag{5.8}$$

and the corresponding variational formulation

$$\text{Find } u \in V_0 = H_0^1(\Omega): \quad a(u, v) = \langle F, v \rangle \,, \quad \forall v \in V_0, \tag{5.9}$$

with

$$\begin{aligned} a(u,v) &:= \sum_{i=1}^{2} \int_{\Omega_i} \nabla u(\Omega)(x) \cdot \nabla v(\Omega)(x) \; dx, \\ \langle F, v \rangle &:= \int_{\Omega_2} v(\Omega)(x) \; dx. \end{aligned}$$

The analytical solution of this PDE can be computed and is given by:

$$u(x_1, x_2) = \begin{cases} \frac{1}{4}(1 - R^2) + \frac{1}{2} R^2 \ln(R) & \text{if } 0 \leq r \leq R \\ \frac{1}{4}(1 - r^2) + \frac{1}{2} R^2 \ln(r) & \text{if } R < r \leq 1 \end{cases} \tag{5.10}$$

with $r = \sqrt{x_1^2 + x_2^2}$.

Next we define the deformation of the domain using the formalism introduced in chapter 3:

We assume to have only one parameter with respect to which shape variations are considered: the radius of subdomain Ω_1.
Therefore, the deformation T_t is given by:

$$T_t : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{cases} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{t}{R} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \text{if } 0 \leq r \leq R \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{t(1-r)}{r(1-R)} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \text{if } R < r \leq 1 \end{cases} \tag{5.11}$$

with $r = \sqrt{x_1^2 + x_2^2}$.

The variational problem for the shape derivative now reads

$$\text{Find } u' \in V_0 : \quad a'(u', v) = \langle F', v \rangle, \quad \forall v \in V_0,$$

with

$$\begin{aligned} a'(u',v) &:= \sum_{i=1}^{2} \int_{\Omega_i} \nabla u'(\Omega;V)(x) \cdot \nabla v(\Omega)(x) \; dx, \\ \langle F', v \rangle &:= -\int_{\Gamma_{1,2}} v(\Omega)(x)(V(0,x) \cdot n_1(x)) \; ds \\ &\quad -\int_{\Gamma_{1,2}} (V(0,x) \cdot n_1(x)) \nabla_\Gamma u(\Omega)(x) \cdot \nabla_\Gamma v(\Omega)(x) \; ds \\ &\quad +\int_{\Gamma_{1,2}} (V(0,x) \cdot n_1(x)) \nabla_\Gamma u(\Omega)(x) \cdot \nabla_\Gamma v(\Omega)(x) \; ds \\ &= -\int_{\Gamma_{1,2}} v(\Omega)(x)(V(0,x) \cdot n_1(x)) \; ds. \end{aligned}$$

By equation (3.3) we have

$$\begin{aligned} V(0,x) &:= \left. \frac{dT_t(x)}{dt} \right|_{t=0} \\ &= \begin{cases} \frac{1}{R} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \text{if } 0 \le r \le R \\ \frac{(1-r)}{r(1-R)} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \text{if } R < r \le 1 \end{cases} \end{aligned}$$

with $r = \sqrt{x_1^2 + x_2^2}$.

Therefore,

$$V(0,x) \cdot n_1(x) = 1 \qquad \text{on } \Gamma_{1,2}$$

and the variational formulation for the shape derivative is simplified to

$$\text{Find } u' \in V_0 : \quad a'(u',v) = \langle F', v \rangle, \quad \forall v \in V_0, \tag{5.12}$$

with

$$\begin{aligned} a'(u',v) &:= \sum_{i=1}^{2} \int_{\Omega_i} \nabla u'(\Omega;V)(x) \cdot \nabla v(\Omega)(x) \; dx, \\ \langle F', v \rangle &:= -\int_{\Gamma_{1,2}} v(\Omega)(x) \; ds. \end{aligned}$$

The analytical solution of this problem can be computed and is given by:

$$u'(\Omega;V)(x_1,x_2) = \begin{cases} R\ln(R) & \text{if } 0 \le r \le R \\ R\ln(r) & \text{if } R < r \le 1 \end{cases} \tag{5.13}$$

with $r = \sqrt{x_1^2 + x_2^2}$.
Note that the shape derivative u' equals the derivative of u with respect to R, which is reasonable since the radius of subdomain Ω_1 is the design parameter.

One can show that for the variational problem (5.12) the assumptions of theorem 4.1 are satisfied by using trace theorems (cf. [13]) and therefore uniqueness of the solution is guaranteed.

The cost functionals of consideration were already discussed and they had the following form

$$J(\Gamma) = \frac{1}{\mu_0}\int_{\Gamma_s} \nabla u(\Omega)(x)^T Q(x) \nabla u(\Omega)(x)\ ds$$

and

$$J(\Gamma) = \int_{\Gamma_s} u^2(\Omega)(x)\ ds.$$

Remark 5.3 As already mentioned, the first functional is only well defined if u is smooth enough, e.g. in H^2. This is indeed true for this test problem.

For the test problem we set $\mu_0 = 1$ and, since taking the matrix $Q(x)$ as in (4.29) would lead to trivial solutions, $Q(x) = I$. So the first functional is simplified to

$$J(\Gamma) = \int_{\Gamma_s} \nabla u(\Omega)(x) \cdot \nabla u(\Omega)(x)\ ds \tag{5.14}$$

and for its Eulerian derivative we have

$$\frac{d}{dt} J(\Gamma_t)\Big|_{t=0} \quad = \quad 2\int_{\Gamma_s} \nabla u'(\Omega;V)(x) \cdot \nabla u(\Omega)(x)\ ds. \tag{5.15}$$

The first functional will be evaluated at the boundary of some circle with radius $r_e > R$, the second functional will be evaluated at the interface $\Gamma_{1,2}$.

The Direct Approach

First the solution of the variational problem (5.9) for u is computed. Then the solution of the variational problem (5.12) for the shape derivative u' is computed. Now the Eulerian derivative of the functionals (5.15) and (5.3) can be evaluated.

The Adjoint Approach for the Second Functional

First the solution of the variational problem (5.9) for u is computed. Then we define the adjoint variable λ as solution of the following problem:

$$\text{Find } \lambda \in V_0 : \quad a'(\lambda, w) = \langle G, w \rangle , \quad \forall w \in V_0, \tag{5.16}$$

with

$$\begin{aligned} a'(\lambda, w) &= \sum_{i=1}^{2} \int_{\Omega_i} \nabla\lambda(\Omega)(x) \cdot \nabla w(\Omega)(x) \; dx, \\ \langle G, w \rangle &:= 2 \int_{\Gamma_{1,2}} u(\Omega)(x) w(\Omega)(x) \; ds. \end{aligned}$$

Since $u' \in V_0$ and (5.16) holds for all functions in V_0 we have

$$\frac{d}{dt} J(\Gamma_t)\bigg|_{t=0} = \langle G, u' \rangle = a'(\lambda, u') = \langle F', \lambda \rangle = -\int_{\Gamma_{1,2}} \lambda(\Omega)(x) \; ds.$$

The analytical solution of this adjoint problem can be computed and is given by:

$$\lambda(x_1, x_2) = \begin{cases} -2 \cdot R(\frac{1}{4}(1 - R^2) + \frac{1}{2} R^2 \ln(R)) \ln(R) & \text{if } 0 \leq r \leq R \\ -2 \cdot R(\frac{1}{4}(1 - R^2) + \frac{1}{2} R^2 \ln(R)) \ln(r) & \text{if } R < r \leq 1 \end{cases} \tag{5.17}$$

with $r = \sqrt{x_1^2 + x_2^2}$.

One can show that for the variational problem (5.16) the assumptions of theorem 4.1 are satisfied by using trace theorems and therefore uniqueness of the solution is guaranteed.

So far we discussed how gradients of functionals can be computed using the direct and adjoint method. We saw that the adjoint method is not well defined for the functional, which describes the torque.
Then we considered a particular test problem, applied the direct and adjoint method there and this resulted in 3 variational problems: the problem for u, the problem for the shape derivative u' and the adjoint problem for λ. So far, everything was formulated in the continuous setting. For solving these variational problems, we need to discretize them. This is done by using the finite element method, which is introduced in the next chapter.

Chapter 6

Discretization

The aim of this chapter is to describe the discretization by using the finite element method. First we give an introduction to the Galerkin-Ritz method with special ansatz functions and then we apply it to our linear problems (the original problem for u, the problem for the shape derivative u' and the adjoint problem for λ). The resulting systems of linear equations will be solved by the conjugate gradient (CG) method.

6.1 Introduction to the Finite Element Method

A comprehensive introduction to the finite element method (FEM) can be found in [15], [16] and [17].

As starting point we consider a linear variational problem

$$\text{Find } u \in V: \quad a(u,v) = \langle F, v\rangle, \quad \forall v \in V, \tag{6.1}$$

in a Hilbert space V, whereby $a(.,.)$ and $\langle F, v\rangle$ should satisfy the assumptions of the Lax-Milgram theorem (cf. Theorem 4.1), which ensures the existence and uniqueness of a solution u^* of the problem (6.1).

The Galerkin-Ritz Method

Let us assume that we have a sequence of finite-dimensional subspaces $V_h \subset V$, and we associate the following *discrete problems* to the variational problem (6.1):

$$\text{Find } u_h \in V_h \subset V: \quad a(u_h, v_h) = \langle F, v_h\rangle, \quad \forall v_h \in V_h, \tag{6.2}$$

where h indicates the discretization parameter and in fact we want to have that for $h \to 0$ the solution of the discrete problem u_h^* converges to the exact

solution u^*. Since $V_h \subset V$, the Lax-Milgram theorem guarantees the existence and uniqueness of a solution u_h^* of the discrete problem (6.2).

The next step is to choose appropriate basis functions for the finite-dimensional space V_h, such that $V_h = \text{span}\{p^{(i)}(x) : i = 1, ..., N_h\}$, where $p^{(i)}$ denote the basis functions and N_h the dimension of the space. Since $u_h \in V_h$, we can use the representation

$$u_h(x) = \sum_{i=1}^{N_h} u_h^{(i)} p^{(i)}(x) \tag{6.3}$$

with the coefficients $u_h^{(i)} \in \mathbb{R}$. Therefore, we get the *Galerkin system*

$$\text{Find } \underline{u}_h \in \mathbb{R}^{N_h} : \quad K_h \underline{u}_h = \underline{f}_h, \tag{6.4}$$

with

$$\begin{aligned} \underline{u}_h &:= \left(u_h^{(i)}\right)_{i=1}^{N_h}, \\ K_h &:= \left(a(p^{(j)}, p^{(i)})\right)_{i,j=1}^{N_h}, \\ \underline{f}_h &:= \left(\left\langle F, p^{(i)} \right\rangle\right)_{i=1}^{N_h}. \end{aligned}$$

The linear system (6.4) is equivalent to the discrete problem (6.2). This equivalence defines the so called *Ritz isomorphism*:

$$u_h \in V_h \longleftrightarrow \underline{u}_h \in \mathbb{R}^{N_h}. \tag{6.5}$$

Construction of V_h using the Courant Element

Now we want to construct finite element subspaces of certain Hilbert spaces and corresponding basis functions. First we introduce the notion of triangulation:

Definition 6.1 Let $\Omega \in \mathbb{R}^d$ be an open, bounded set with sufficiently smooth boundary. Then $\mathcal{T}_h := \{\delta_r : r \in \mathbb{R}_h\}$ is called triangulation of Ω if:

1) $\overline{\Omega} = \bigcup_{r \in \mathbb{R}_h} \overline{\delta_r}$

2) $\delta_r \cap \delta_s = 0$.

Additionally we force the triangulation to be *admissible*, which means in 2D:

Definition 6.2 A triangulation $\mathcal{T}_h$ is admissible if:

$$\overline{\delta_r} \cap \overline{\delta_s} = \begin{cases} 0 \\ \text{common vertex} \\ \text{common edge.} \end{cases}$$

In the Galerkin-Ritz FEM we have to face three main aspects:

- A triangulation $\mathcal{T}_h$ of the domain Ω.
- The finite element space V_h is constructed such that the functions are continuous and for each $\delta_r \in \mathcal{T}_h$ the restriction of $v_h \in V_h$ is a polynomial.
- The basis of V_h is constructed such that the basis functions have local support.

In 2D the widely used element is the so called *Courant element*. The domain is subdivided into triangles, on which the functions v_h should be linear. Therefore we have for the finite element space

$$V_h := \{v_h \in C(\overline{\Omega}) |\ v_h|_{\delta_r} \in P_1 \ \ \forall \delta_r \in \mathcal{T}_h\}. \tag{6.6}$$

The construction of this space is as follows:

- Let $x^{(j)}$, $j \in \overline{\omega}_h$ denote all vertices (called *nodes*) of the triangulation and $\overline{\omega}_h$ the index set of the nodes.
- We set

  $$p^{(i)}(x^{(j)}) := \delta_{ij}, \quad \forall i, j \in \overline{\omega}_h,$$

  with

  $$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

This basis is called *nodal basis*, since the unknowns are the function values at the nodes: $u_h^{(j)} = u_h(x^{(j)})$. Therefore the basis function $p^{(i)}$ forms a pyramid function with height 1 at the node $x^{(i)}$. Additionally this basis functions have local support, therefore, since $a(p^{(i)}, p^{(j)}) = 0$ if the supports are not overlapping, the corresponding *stiffness matrix* K_h is sparse.

We can write

$$V_h := \Big\{ v_h(x) = \sum_{j \in \overline{\omega}_h} v^{(j)} p^{(j)}(x) \Big| v^{(j)} \in \mathbb{R} \Big\}. \tag{6.7}$$

with $V_h \subset H^1$ for an admissible subdivision into triangles.

The Reference Triangle

The stiffness matrix and the load vector (right hand side) are assembled by summing over the elements $\delta_r \in \mathcal{T}_h$. The appearing integrals over the elements can be evaluated efficiently using the concept of the reference triangle. For Courant triangles we define the reference triangle as $\Delta := \left\{[\xi_1, \xi_2] \in [0,1]^2 \middle| \xi_1 + \xi_2 \leq 1\right\}$ with the following reference basis functions:

$$\begin{aligned} p^{(1)}(\xi) &:= 1 - \xi_1 - \xi_2 \\ p^{(2)}(\xi) &:= \xi_1 \\ p^{(3)}(\xi) &:= \xi_2. \end{aligned}$$

Now every δ_r can be written as an affine linear transformation of Δ and, consequently, the basis functions on every δ_r can be written in terms of the reference basis functions. So the integrals can be easily transformed to the reference triangle using the three element nodes $x^{(i,1)}, x^{(i,2)}, x^{(i,3)}$. For efficiently evaluating these integrals we used numerical integration (midpoint rule).

6.2 The Finite Element Method applied to the Test Problem

In our test problem we have to solve three variational problems:

- 1) The VP for u:

  $$\text{Find } u \in V_0 = H_0^1(\Omega): \quad a(u,v) = \langle F, v\rangle, \quad \forall v \in V_0,$$

  with

  $$\begin{aligned} a(u,v) &= \sum_{i=1}^{2} \int_{\Omega_i} \nabla u(\Omega)(x) \cdot \nabla v(\Omega)(x) \; dx, \\ \langle F, v\rangle &= \int_{\Omega_2} v(\Omega)(x) \; dx. \end{aligned}$$

- 2) The VP for u':

  $$\text{Find } u' \in V_0: \quad a'(u',v) = \langle F', v\rangle, \quad \forall v \in V_0,$$

  with

  $$\begin{aligned} a'(u',v) &= \sum_{i=1}^{2} \int_{\Omega_i} \nabla u'(\Omega; V)(x) \cdot \nabla v(\Omega)(x) \; dx, \\ \langle F', v\rangle &= -\int_{\Gamma_{1,2}} v(\Omega)(x) \; ds. \end{aligned}$$

- 3) The VP for λ:

$$\text{Find } \lambda \in V_0 : \quad a'(\lambda, w) = \langle G, w \rangle , \quad \forall w \in V_0,$$

with

$$\begin{aligned} a'(\lambda, w) &= \sum_{i=1}^{2} \int_{\Omega_i} \nabla\lambda(\Omega)(x) \cdot \nabla w(\Omega)(x) \ dx, \\ \langle G, w \rangle &= 2 \int_{\Gamma_{1,2}} u(\Omega)(x) w(\Omega)(x) \ ds. \end{aligned}$$

In all the problems we have the same Hilbert space H_0^1 and the same bilinear form $a(u, v)$. One can easily see that

$$V_{0h} := \left\{ v_h(x) = \sum_{j \in \omega_h} v^{(j)} p^{(j)}(x) \middle| v^{(j)} \in \mathbb{R} \right\} \subset H_0^1$$

where ω_h is the set of all indices whose corresponding nodes do not lie on Γ.

The construction of K_h starts with initializing all the entries with 0. Then we have

$$K_{h(i,j)} = a(p^{(j)}, p^{(i)}) = \int_{\Omega} \nabla p^{(j)} \cdot \nabla p^{(i)}(x) \ dx = \sum_{r \in B_{i,j}} \int_{\delta_r} \nabla p^{(j)} \cdot \nabla p^{(i)}(x) \ dx$$

where $B_{i,j} := B_i \cap B_j$ and $B_i := \{r \in \mathbb{R}_h : x^{(i)} \in \overline{\delta}_r\}$.
Therefore we iterate over all elements δ_r, compute the integrals using the reference triangle in each step and add this value to $K_{h(i,j)}$.
Because of the Ritz isomorphism the properties of the bilinear form carry over to the stiffness matrix and therefore the stiffness matrix in the three problems is symmetric and positive definite.

For calculating the load vectors, the same principle is used.

The implementation of the interface conditions works through iterating over all edges, which belong to the interface. As a reference element the line from 0 to 1 is taken. The rest is almost the same as above.

The implementation of the homogeneous Dirichlet boundary condition works as follows:

- set $u_h^{(j)} = 0$ for $j \in \gamma_h$ with $\overline{\omega}_h = \omega_h \cup \gamma_h$
- set $K_{h(i,j)} = K_{h(j,i)} = \delta_{ij}$
- set $f_h^{(j)} = 0$ for $j \in \gamma_h$.

6.3 The CG - Method

If the matrix K in the linear system

$$Ku = f, \tag{6.8}$$

is symmetric and positive definite, problem (6.8) is equivalent to the following minimization problem:

$$J(u) := \frac{1}{2}(Ku, u) - (f, u) \to \min_u, \tag{6.9}$$

because

$$\begin{aligned} \nabla J(u) &= Ku - f \stackrel{!}{=} 0, \\ \nabla^2 J(u) &= K = K^T > 0. \end{aligned}$$

J is called the energy functional.

The CG - algorithm reads as follows:

- Initialization: choose initial guess u^0

$$\begin{aligned} d^0 &= f - Ku^0 \\ p^0 &= d^0 \end{aligned}$$

- Iteration: for $n = 0, 1, 2, ...$ until convergence do

$$\begin{aligned} u^{n+1} &= u^n + \alpha^{(n)} p^n \quad \text{with} \quad \alpha^{(n)} = \frac{(d^n, p^n)}{(Kp^n, p^n)}, \\ d^{n+1} &= d^n - \alpha^{(n)} Kp^n, \\ p^{n+1} &= d^{n+1} - \beta^{(n)} p^n \quad \text{with} \quad \beta^{(n)} = \frac{(d^{n+1}, Kp^n)}{(Kp^n, p^n)}. \end{aligned}$$

  end.

For the CG - method we have the following convergence result:

Theorem 6.3 Let $K \in \mathbb{R}^{n \times n}$ be symmetric and positive definite.
Then the CG - method is convergent and the following error estimate holds:

$$||u^n - u||_K \leq \frac{2q^n}{1 + q^{2n}} ||u^0 - u||_K \leq 2q^n ||u^0 - u||_K, \tag{6.10}$$

with

$$0 \leq q = \frac{\sqrt{\kappa(K)} - 1}{\sqrt{\kappa(K)} + 1} < 1, \tag{6.11}$$

where $\kappa(K)$ is the condition number of the matrix K.

Since the stiffness matrix in our three problems satisfies these assumptions, the CG - method can be used for solving them.
In the next chapter the numerical results are presented.

Chapter 7

Numerical Results

In this chapter the numerical results for the direct and adjoint method applied to the test problem from chapter 5 are presented. As already mentioned, for comparing the two approaches we use the functional (5.2), since the adjoint method is not well defined for the functional (5.14). But we also present the results of the direct method used with the functional (5.14).

For the following results the radius of subdomain Ω_1 was set to $R = 0.4$.

7.1 The Analytical Solutions

The analytical solutions of (5.9), (5.12) and (5.16) are:

$$u(x_1, x_2) = \begin{cases} \frac{1}{4}(1 - 0.4^2) + \frac{1}{2}0.4^2 \ln(0.4) \approx 0.137 & \text{if } 0 \leq r \leq R \\ \frac{1}{4}(1 - r^2) + \frac{1}{2}0.4^2 \ln(r) & \text{if } R < r \leq 1, \end{cases}$$

$$u'(\Omega; V)(x_1, x_2) = \begin{cases} 0.4 \ln(0.4) \approx -0.367 & \text{if } 0 \leq r \leq R \\ 0.4 \ln(r) & \text{if } R < r \leq 1 \end{cases}$$

and

$$\lambda(x_1, x_2) = \begin{cases} -2 \cdot 0.4(\frac{1}{4}(1 - 0.4^2) + \frac{1}{2}0.4^2 \ln(0.4)) \ln(0.4) \approx 0.1002 & \text{if } 0 \leq r \leq R \\ -2 \cdot 0.4(\frac{1}{4}(1 - 0.4^2) + \frac{1}{2}0.4^2 \ln(0.4)) \ln(r) & \text{if } R < r \leq 1 \end{cases}$$

with $r = \sqrt{x_1^2 + x_2^2}$.

Figure 7.1 shows $u(x_1, x_2)$, $u'(\Omega; V)(x_1, x_2)$ and $\lambda(x_1, x_2)$:

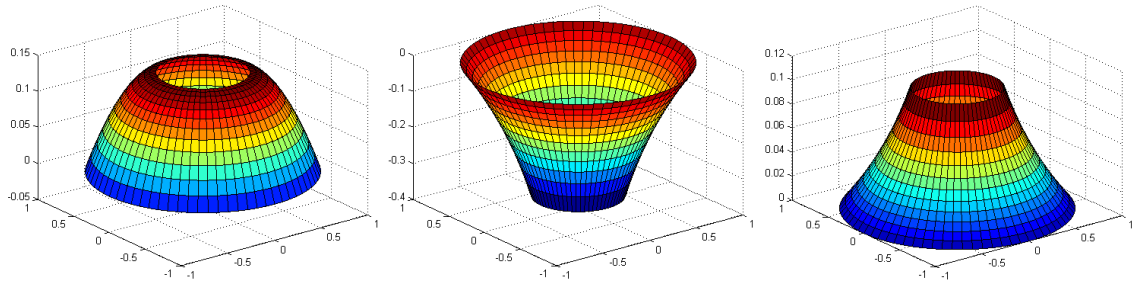

Figure 7.1: The analytical solutions $u(x_1, x_2)$, $u'(\Omega; V)(x_1, x_2)$ and $\lambda(x_1, x_2)$

In the next section we present the results for the functional (5.2) with the direct and adjoint method and the results for the functional (5.14) with the direct method.

For the FEM the MATLAB mesh generator was used.

7.2 Comparison of the Direct and Adjoint Method

First we consider the functional (5.2):

$$\begin{aligned} J(\Gamma) &= \int_{\Gamma_s} u^2(\Omega)(x)\ ds = \int_{R=0.4} u^2(\Omega)(x)\ ds \\ &= \left(\frac{1}{4}(1-0.4^2) + \frac{1}{2}0.4^2\ln(0.4)\right)^2 2\cdot 0.4\cdot\pi \\ &\approx 0.046963 \end{aligned} \tag{7.1}$$

and its derivative

$$\begin{aligned} \frac{d}{dt}J(\Gamma_t)\Big|_{t=0} &= 2\int_{\Gamma_s} u'(\Omega; V)(x)u(\Omega)(x)\ ds = 2\int_{R=0.4} u'(\Omega; V)(x)u(\Omega)(x)\ ds \\ &= 2\cdot 0.4\ln(0.4)\left(\frac{1}{4}(1-0.4^2) + \frac{1}{2}0.4^2\ln(0.4)\right)2\cdot 0.4\cdot\pi \\ &\approx -0.25184. \end{aligned} \tag{7.2}$$

Therefore we can compare the numerical results with the exact solutions.

For the FEM we used the following meshes:

mesh number	number of unknowns (nodes)	number of triangles
1	152	268
2	571	1072
3	2213	4288
4	8713	17152
5	34577	68608

The meshes were generated using the "Refinement" - option in Matlab.

In figure 7.2 one can see the generated meshes. Figures 7.3 - 7.7 show the FE solutions of the problems (5.9), (5.12) and (5.16) for u, u' and λ.

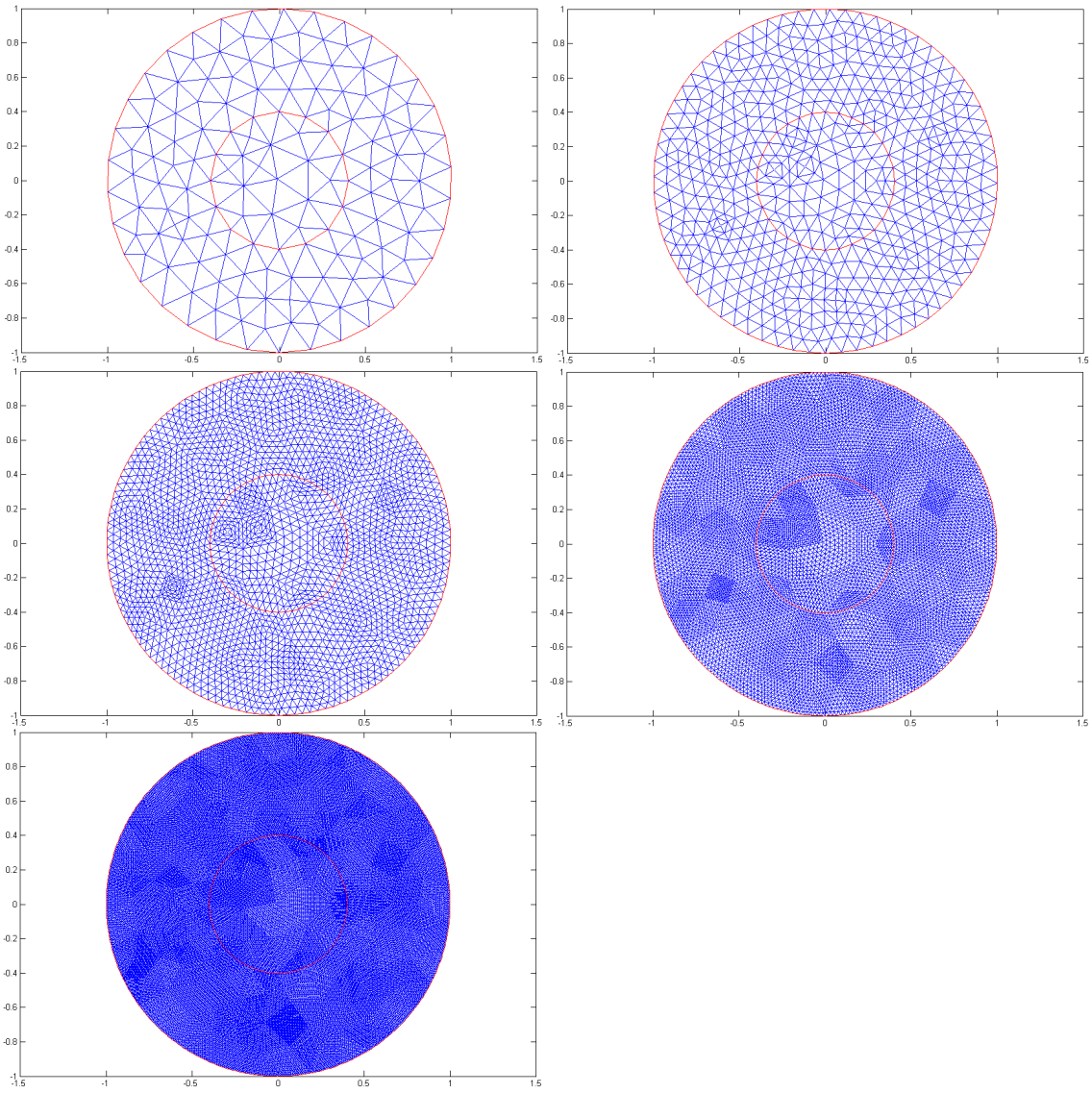

Figure 7.2: The used meshes

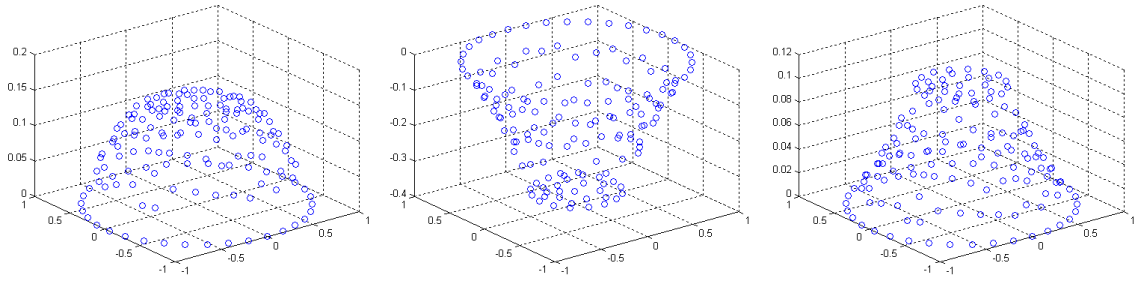

Figure 7.3: The solutions u, u' and λ with mesh 1

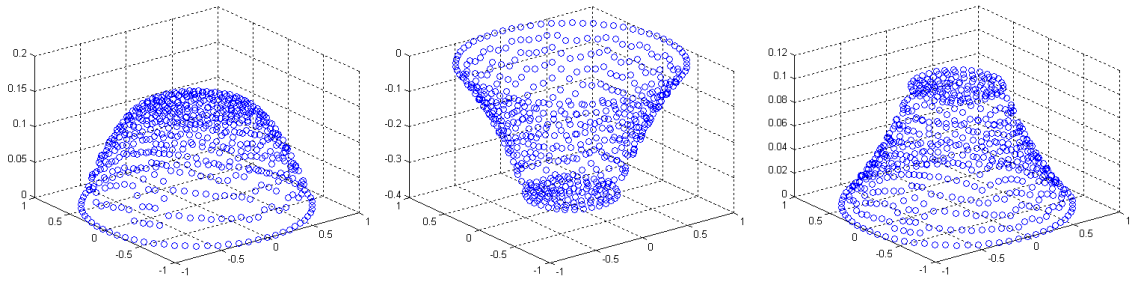

Figure 7.4: The solutions u, u' and λ with mesh 2

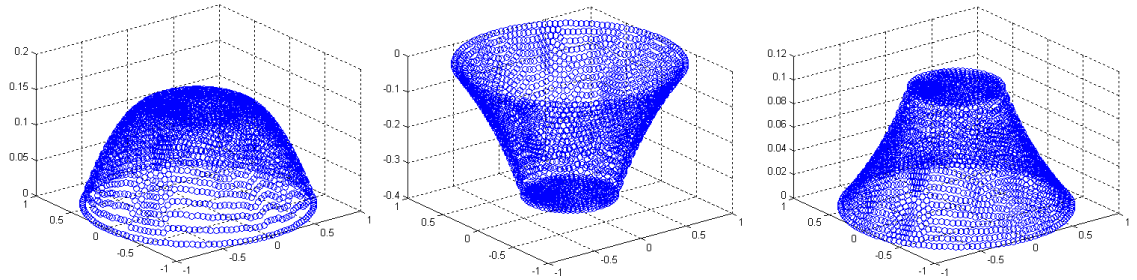

Figure 7.5: The solutions u, u' and λ with mesh 3

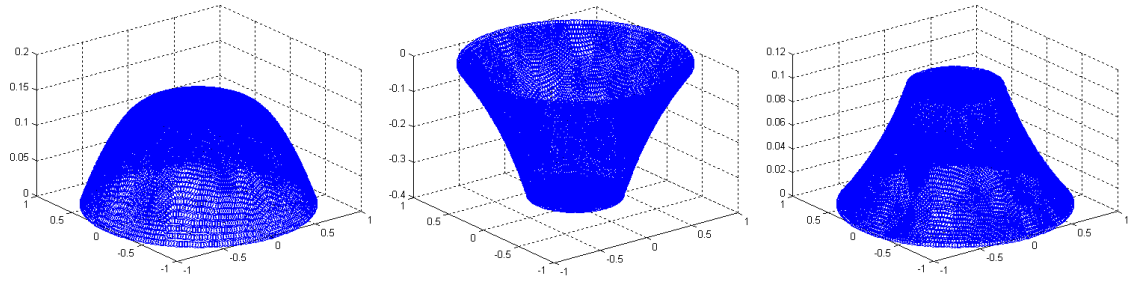

Figure 7.6: The solutions u, u' and λ with mesh 4

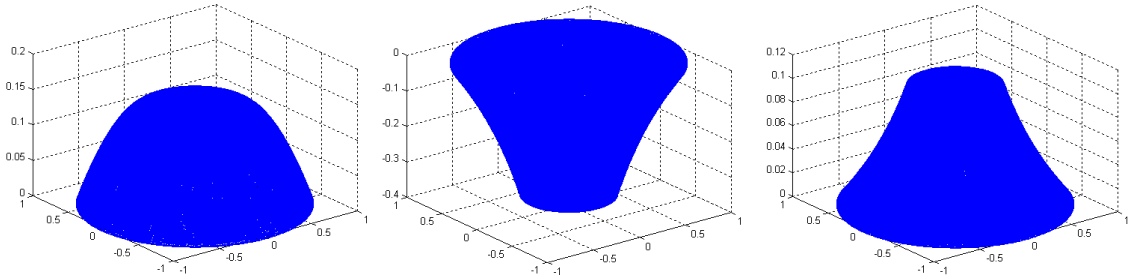

Figure 7.7: The solutions u, u' and λ with mesh 5

With these FE solutions, we evaluated the functional and its derivative (using numerical integration, in particular: midpoint rule) and got the following results:

mesh number	functional	derivative direct	derivative adjoint
1	0.0478194	-0.251122	-0.251122
2	0.0471791	-0.251666	-0.251666
3	0.0470173	-0.251796	-0.251796
4	0.0469766	-0.251828	-0.251828
5	0.0469664	-0.251835	-0.251835

If we compare these results with the exact values (7.1) and (7.2) we immediately see convergence of the discrete solution. We also see that the direct and the adjoint approach deliver the same result for the derivative of the functional and this confirms the theory, which was done in the previous chapters.

For completeness we now consider functional (5.14) and provide gradient

information using the direct method only (the adjoint one is not well defined for this type of functional).
The functional (5.14) looks as follows:

$$J(\Gamma) = \int_{\Gamma_s} \nabla u(\Omega)(x) \cdot \nabla u(\Omega)(x) \; ds$$

with the derivative:

$$\frac{d}{dt} J(\Gamma_t)\Big|_{t=0} \;=\; 2 \int_{\Gamma_s} \nabla u'(\Omega; V)(x) \cdot \nabla u(\Omega)(x) \; ds.$$

As already mentioned, this functional is evaluated at the boundary of some circle with radius $r_e > 0.4$. We take $r_e = 0.5$. Therefore,

$$\begin{aligned} J(\Gamma) &= \int_{r_e=0.5} \nabla u(\Omega)(x) \cdot \nabla u(\Omega)(x) \; ds \\ &= \int_{r_e=0.5} \begin{pmatrix} \frac{0.4^2 x_1}{2 \cdot 0.5^2} - \frac{x_1}{2} \\ \frac{0.4^2 x_2}{2 \cdot 0.5^2} - \frac{x_2}{2} \end{pmatrix} \cdot \begin{pmatrix} \frac{0.4^2 x_1}{2 \cdot 0.5^2} - \frac{x_1}{2} \\ \frac{0.4^2 x_2}{2 \cdot 0.5^2} - \frac{x_2}{2} \end{pmatrix} \; ds \\ &= \left(\frac{0.08^2}{0.5^2} - 0.08 + \frac{0.5^2}{4} \right) 2 \cdot 0.5 \cdot \pi \\ &\approx 0.0254469 \end{aligned} \tag{7.3}$$

and

$$\begin{aligned} \frac{d}{dt} J(\Gamma_t)\Big|_{t=0} &= 2 \int_{r_e=0.5} \nabla u'(\Omega; V)(x) \cdot \nabla u(\Omega)(x) \; ds \\ &= 2 \int_{r_e=0.5} \begin{pmatrix} \frac{0.4 x_1}{0.5^2} \\ \frac{0.4 x_2}{0.5^2} \end{pmatrix} \cdot \begin{pmatrix} \frac{0.4^2 x_1}{2 \cdot 0.5^2} - \frac{x_1}{2} \\ \frac{0.4^2 x_2}{2 \cdot 0.5^2} - \frac{x_2}{2} \end{pmatrix} \; ds \\ &= 2 \cdot \left(\frac{0.4^3}{2 \cdot 0.5^2} - 0.2 \right) 2 \cdot 0.5 \cdot \pi \\ &\approx -0.4523893. \end{aligned} \tag{7.4}$$

For computing this functional and the derivative using the FE solutions, we have to guarantee that there exist nodes at the radius $r_e = 0.5$. Therefore we used different meshes than before:

mesh number	number of unknowns (nodes)	number of triangles
1	248	460
2	955	1840
3	3749	7360
4	14857	29440
5	59153	117760

In figure 7.8 one can see the generated meshes.

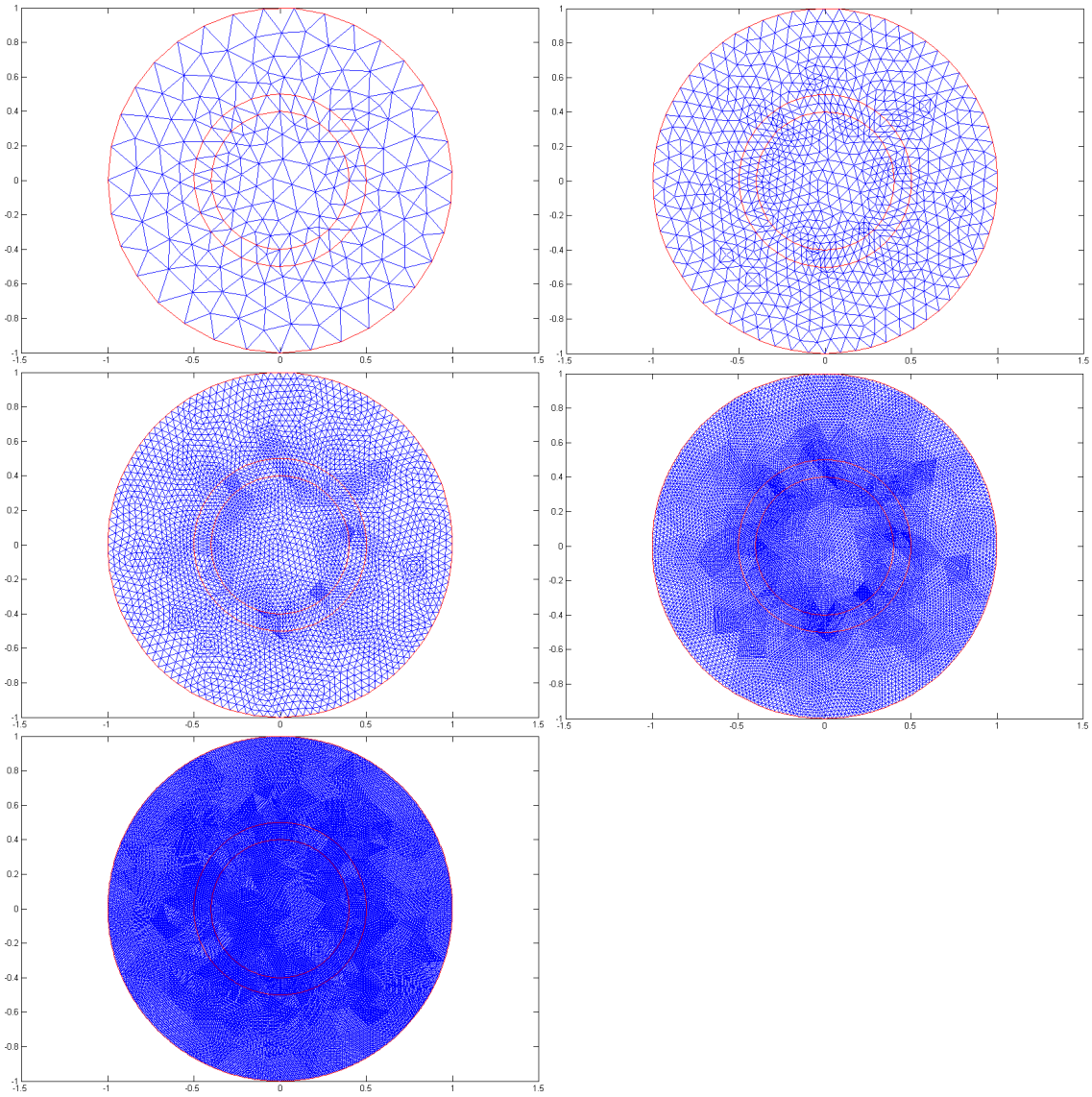

Figure 7.8: The used meshes

For evaluating the functional and the derivative, we need to calculate gradients out of the FE solutions. Since we used P_1-functions for the FEM, the gradient is constant on each finite element, and therefore will not be continuous across the edge of two neighbouring elements. This means that we have the following possibilities:

- use the information about the gradient in elements which touch the circle with radius $r_e = 0.5$ **inside**
- use the information about the gradient in elements which touch the circle with radius $r_e = 0.5$ **outside**
- take a **mean value** of the inside and outside gradient information.

How this is meant is illustrated in figure 7.9

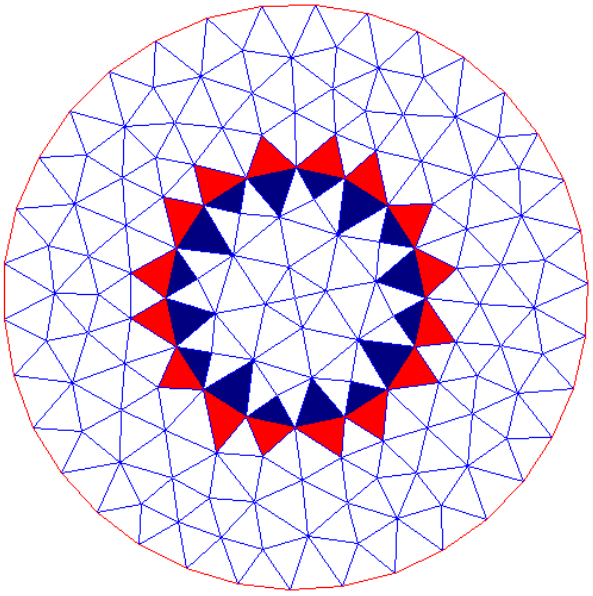

Figure 7.9: The touching elements

The blue triangles are the elements which touch the circle with radius $r_e = 0.5$ inside and the red ones are the elements which the touch this circle outside. For the calculation of the gradient of the FE solution according to the first possibility, only the blue triangles are used. For the calculation of the gradient of the FE solution according to the second possibility, only the red triangles are used. For the calculation of the gradient of the FE solution according to the third possibility, the gradient is calculated in the red and blue triangles and then a mean value of the gradient of the neighbouring red and blue triangles is taken.

For the different possibilities we got the following results:

- gradient inside

mesh number	functional	derivative
1	0.00849013	-0.290104
2	0.0156795	-0.375572
3	0.0202496	-0.415243
4	0.0227719	-0.434146
5	0.0240905	-0.443351

- gradient outside

mesh number	functional	derivative
1	0.048333	-0.552596
2	0.036016	-0.507536
3	0.0304912	-0.481164
4	0.0279072	-0.467068
5	0.0266615	-0.459801

- mean value of gradient inside and outside

mesh number	functional	derivative
1	0.0242782	-0.442411
2	0.0247937	-0.446863
3	0.0251064	-0.449526
4	0.0252736	-0.450936
5	0.0253595	-0.451658

If we compare these results with the exact values (7.3) and (7.4) we see, that the solutions in all cases converge. But we also see that the one with the mean values is the fastest converging solution.

For interest, consider the adjoint problem with this type of functional:

$$\text{Find } \lambda \in V_0 = H_0^1(\Omega): \quad a'(\lambda, w) = \langle G, w \rangle, \quad \forall w \in V_0,$$

with

$$\begin{aligned} a'(\lambda, w) &= \sum_{i=1}^{2} \int_{\Omega_i} \nabla\lambda(\Omega)(x) \cdot \nabla w(\Omega)(x) \; dx, \\ \langle G, w \rangle &= 2 \int_{\Gamma_s} \nabla u(\Omega)(x) \cdot \nabla w(\Omega)(x) \; ds. \end{aligned}$$

As already mentioned, w is a test function from H^1, but the gradient of a H^1 - function is not well defined at boundaries and, therefore, the variational problem for the adjoint variable λ is not well defined in H^1.

Nevertheless, one could ignore this fact and try to discretize and implement this variational problem using the FEM with Courant elements, like we did for the previous (well defined) problems. But one has to think about how to implement it. Again we have the three possibilities like above:

- use the information about the gradient of u in elements which touch the circle with radius $r_e = 0.5$ **inside**
- use the information about the gradient of u in elements which touch the circle with radius $r_e = 0.5$ **outside**
- take a **mean value** of the inside and outside gradient information.

The following tables show the results for the different possibilities compared with the results from the direct approach:

- gradient inside

mesh number	derivative direct	derivative adjoint
1	-0.290104	-0.631029
2	-0.375572	-1.67432
3	-0.415243	-3.74837
4	-0.434146	-7.88531
5	-0.443351	-16.1521

- gradient outside

mesh number	derivative direct	derivative adjoint
1	-0.552596	-1.55389
2	-0.507536	-2.7924
3	-0.481164	-5.24236
4	-0.467068	-10.1309
5	-0.459801	-19.903

- mean value of gradient inside and outside

mesh number	derivative direct	derivative adjoint
1	-0.442411	-0.264351
2	-0.446863	-0.536184
3	-0.449526	-1.07902
4	-0.450936	-2.16381
5	-0.451658	-4.33281

One immediately sees, that the solutions of the adjoint approach in all cases do not converge and, therefore, the results are, as expected, useless.

This can be seen as a warning example in the sense that before starting discretizing it is very important that one first makes sure that the formulated problems are well defined.

Chapter 8

Conclusions and Outlook

8.1 Conclusion

In this thesis the entire mathematical treatment of using the shape derivative approach in shape optimization is embraced. We introduced two different methods for providing gradient information of functionals, which consist of domain and/or boundary integrals and depend on the solution of some state problem: the direct and the adjoint technique. We have seen that the adjoint approach has a great advantage compared to the direct one: the adjoint state equation has to be solved only once.
We considered a magnetostatic problem, where we have derived a mathematical formulation from the governing physical equations (Maxwell). Then we applied the two approaches and this resulted in three linear variational problems. Solvability of the problems has been proven under the assumptions of the Lax-Milgram theorem.

We have discretized the continuous problems with the finite element method. In order to solve the resulting systems of equations, where the involved matrices are symmetric and positive definite, we used the conjugate gradient method.

Numerical results have shown that indeed the direct and the adjoint approach lead to the same results for the test problem. But we also have seen that there exist examples where the adjoint technique is not useful in the sense of providing suitable results.

8.2 Outlook

This work can be continued in the following directions:

- The nonlinear problem:

  When deriving the mathematical formulation we assumed the involved materials to be linear and therefore the variational problem got linear. But in most of the problems in physics there is a nonlinear part involved. Therefore the adjoint method has to be set up in the nonlinear case.

- The adjoint method for "general" functionals:

  We have shown an example of a functional where the adjoint problem, which was derived in the usual way (straight forward), is not well defined and therefore the adjoint approach is inapplicable in the sense we introduced it. For this one has to think about how to reformulate the adjoint problem.

Bibliography

[1] J.HASLINGER and R.A.E.MÄKINEN. *Introduction to Shape Optimization: Theory, Approximation and Computation.* SIAM, 2003.

[2] JAN SOKOLOWSKI and JEAN-P. ZOLESIO. *Introduction to Shape Optimization: Shape sensitivity analysis.* Springer Verlag Berlin, 1992.

[3] JEAN-P. ZOLESIO and M.C. DELFOUR. *Shapes and Geometries: Analysis, Differential Calculus and Optimization.* SIAM, 2001.

[4] P.LE TALLEC and E.LAPORTE. *Numerical Methods in Sensitivity and Shape Optimization.* BIRKHÄUSER, 2002.

[5] W.MÜHLHUBER. *Spezialforschungsbereich F013: Various Aspects of Shape Optimization.* Johannes Kepler University Linz, Presentation.

[6] E.J.HAUG K.K.CHOI and V.KOMKOV. *Mathematics in Science and Engineering, Volume 177: Design Sensitivity Analysis of Structural Systems.* ACADEMIC PRESS, INC, 1986.

[7] H.ENGL. *Lecture Notes for the Course Mathematical Methods in Continuum Mechanics.* Johannes Kepler University Linz, Winter Semester 2008/09.

[8] W.ZULEHNER. *Lecture Notes for the Course Numerical Methods in Continuum Mechanics.* Johannes Kepler University Linz, Summer Semester 2009.

[9] B.HEISE. *Mehrgitter-Newton-Verfahren zur Berechnung nichtlinearer magnetischer Felder.* Technische Universität Chemnitz, 1991.

[10] S.MÜHLBÖCK and A.LECHNER. *Numerical solution of nonlinear stationary magnetic field problems.* Johannes Kepler University Linz, Bachelor Thesis, 2008.

[11] C.PECHSTEIN. *Multigrid-Newton-Methods for Nonlinear Magnetostatic Problems.* Johannes Kepler University Linz, Master Thesis, 2004.

[12] W.GRUBER W.AMRHEIN S.SILBER H.GRABNER and M.REISINGER. *Theoretical analysis of force and torque calculation in magnetic bearing systems with circular airgap.* Institute of Electrical Drives and Power Electronics JKU Linz.

[13] R.ADAMS. *Sobolev Spaces.* Academic Press, New York, 1975.

[14] U.LANGER. *Lecture Notes for the Course Numerical Methods for Partial Differential Equations.* Johannes Kepler University Linz, Winter Semester 2007/08.

[15] P.G. CIARLET. *The Finite Element Method for Elliptic Problems.* Classics in Applied Mathematics, Vol. 40, SIAM, Philadelphia, 2002.

[16] U.LANGER and M.JUNG. *Methode der finiten Elemente für Ingenieure.* Teubner-Verlag, Stuttgart, Leipzig, Wiesbaden, 2001.

[17] W.ZULEHNER. *Lecture Notes for the Course Numerical Methods for Elliptic Problems.* Johannes Kepler University Linz, Summer Semester 2008.

Eidesstattliche Erklärung

Ich, Markus Kollmann, erkläre an Eides statt, dass ich die vorliegende Masterarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Leonding, Februar 2010 Markus Kollmann

Curriculum Vitae

Name: Markus Kollmann

Nationality: Austria

Date of Birth: 1^{st} of October, 1984

Place of Birth: Linz, Austria

Education:

1991 - 1995	Elementary School Keferfeld, Linz
1995 - 1999	Secondary Comprehensive School Landwiedstraße, Linz
1999 - 2004	Polytechnic Leonding
2005 - 2008	Studies of Technical Mathematics (Bachelor) Johannes Kepler University, Linz
2008 - 2010	Studies of Industrial Mathematics (Master) Johannes Kepler University, Linz

Special Activities:

August 2009	23^{rd} ECMI Modelling Week, Wroclaw, Poland