Christian Köthe, BSc BSc

PDE-constrained shape optimization for coupled problems using space-time finite elements

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Mathematics

submitted to

Graz University of Technology

Supervisor

Dr. P. Gangl

Institute of Applied Mathematics

Graz, June 2020

Abstract

This thesis deals with shape optimization problems which are motivated by the design optimization of an electrical machine. Mathematically each problem results in an infinite dimensional optimization problem which is constrained by a linear system of coupled parabolic partial differential equations (PDEs). We are facing three different problems: (i) a one-phase problem, (ii) an interface problem, (iii) a linear shape optimization problem for an electric motor. We solve each problem by means of shape optimization techniques based on sensitivity information. The underlying PDEs are solved via a conforming space-time finite element method.

In a first step we solve the coupled system of linear parabolic PDEs of problem (i) and (ii) via a conforming space-time method. The method is based on a Galerkin-Petrov variational formulation and allows continuous ansatz and test spaces. We present numerical results which illustrate the experimental order of convergence (eoc) for the error in the L^2-norm and energy norm.

In a second step we briefly introduce the reader to the theory of shape optimization in which we define the so called shape derivative. We show how this derivative can be embedded in algorithms in order to get improved designs.

Next, we apply the introduced shape optimization techniques to problem (i) and (ii). We compute the shape derivative for each problem and describe the main steps of the algorithm which is used to solve the problems. We state numerical results which demonstrate the correctness of the presented method.

Finally, we focus on (iii) which is a linear shape optimization problem for an electric motor. We describe the mathematical problem as well as the geometrical setting for a motor geometry in 2D. Further we discuss the solvability of the state system and solve (iii) by employing the shape derivative. Since we do not use any parallel implementation and consider the state system to be of linear type the obtained results have to be understood as a proof of concept.

Contents

Introduction

Shape optimization problems arise in many engineering applications in which one wants to find an optimal design of a shape such that a given performance criterion is minimized or maximized. For instance in mechanical engineering one wants to find the optimal structure of a bridge such that the compliance is minimized or in aviation one is interested in finding an optimal layout of a wing. [13]

Following the classification in [12] one basically distinguishes between parametric and non-parametric shape optimization. In parametric shape optimization the design is described by a set of parameters such as dimensions or orientation of an object [12]. This results in a shape optimization problem, where the design space is finite dimensional and has vector space structure. Hence one is able to use standard tools from analysis like the Fréchet-derivative in order to set up gradient-based algorithms. In general the possible designs are restricted due to the finite number of parameters. Further we want to mention (cf. [13]) that parametric shape optimization problems can be solved via evolutionary algorithms. This class of algorithms gets along without the knowledge of a derivative. Hence, they are less prone to getting stuck in local optima and it is more likely that the global optimum can be found [13]. Moreover these algorithms are also useful in the context of multi-objective optimization. One disadvantage of evolutionary algorithms compared to gradient-type algorithms is that they typically need much more computational time.

Non-parametric shape optimization problems allow for a wider variety of optimal designs since the design space is infinite dimensional. Since the design space does not have vector space structure the mathematical tool used in this context is shape sensitivity analysis. [12]

In this thesis we deal with non-parametric shape optimization problems and use shape sensitivity analysis in order to set up gradient-type algorithms.

An important class of shape optimization problems are PDE-constrained shape optimization problems. This means that the function which is minimized or maximized depends on the shape via the solution of a partial differential equation (PDE). In general a PDE-constrained shape optimization problem consists of an objective function $\mathcal{J}(\Omega) = J(\Omega, u_\Omega)$, where $\Omega \in \mathcal{A}$ is an element out of the set of admissible shapes $\mathcal{A}$ and u_Ω satisfies the PDE constraint $E(\Omega, u_\Omega) = 0$ [39]. The goal is to minimize the objective function $\mathcal{J} : \mathcal{A} \to \mathbb{R}$ over the set of admissible shapes, i.e.

$$\min_{\Omega \in \mathcal{A}} J(\Omega, u) \text{ s.t. } u \text{ solves } E(\Omega, u) = 0. \tag{0.1}$$

Note that maximizing the objective function $\mathcal{J}$ can be achieved by minimizing $-\mathcal{J}$ [12]. The PDE constraint E could be a single PDE or a system of PDEs. Elliptic PDE-constrained shape optimization problems are well investigated in literature, see e.g. [7, 34, 39]. They are also considered in practical applications like in the design optimization of an electric motor, see [12, 14]. The literature for parabolic PDE-constrained shape optimization problems is more limited, as it is also remarked in [15]. Theoretical results are presented in [10, 33]. In [15] the numerical solution of a parabolic shape optimization problem using multipole-based space-time boundary elements has been considered. In [32] a parabolic shape optimization problem is solved via a BFGS-quasi Newton method. The underlying parabolic PDE is solved with the vertical method of lines. The space discretization is done via standard linear finite elements and the resulting first-order system is solved using the implicit Euler method.

So far, parabolic shape optimization using space-time finite elements has not been considered to the best of our knowledge. We focus on this approach. Different space-time methods have been considered in literature. In [22, 23] a discontinuous Galerkin method is used in time. Conforming space-time methods have been considered by Steinbach [35] and Toulopoulos [40]. In [20] a space-time method is proposed in which Isogeometric Analysis is used for space-time discretization. This space-time scheme is extended in the master thesis [28] in which a conforming finite element method is used to discretize. We focus on the space-time approach by Steinbach [35] in this thesis, which is a conforming space-time finite element method and allows continuous ansatz and test spaces. The idea is to treat the time variable as an additional spatial coordinate. The discretization is done simultaneously in space and time. Thus one is able to use adaptive mesh refinement strategies in space-time. Further it is possible to parallelize the method not only with respect to space, but also with respect to time. The main disadvantage of the space-time method proposed in [35] is that it leads to a problem in one dimension higher which results in a bigger computational effort.

The objective of this master thesis is to solve parabolic PDE-constrained shape optimization problems via a gradient-based algorithm and using space-time finite elements. The problem we are going to consider is motivated from the design optimization of an electric motor. In general the task is to find an optimal design of the rotor such that a given performance criterion which is measured in terms of an objective function J is as good as possible [12]. For instance J could measure the torque and one is interested in finding a rotor design such that the torque is maximized or in [12, 14] a shape optimization of an electric motor is considered such that the rotation pattern of the rotor becomes smooth. Mathematically, the design optimization of an electric motor can be formulated as PDE-constrained shape optimization problem. The PDE constraint can be of elliptic type if the rotor of the motor is considered to rotate at a constant speed, see [12, 14]. We are also interested in the starting phase of the motor, i.e. when the rotor is accelerated from its resting position. The PDE constraint then turns out to

be a system of parabolic PDEs. This system of PDEs describes the physical model of the motor, which is due to the electromagnetic-thermal coupling in the motor. This coupling is on the one hand given through eddy-currents which have an impact on the temperature and on the other hand given through the electrical conductivity $\sigma = \sigma(\vartheta)$ which depends on the temperature and has an influence on the magnetic induction $\boldsymbol{B}$. Furthermore the properties of the permanent magnets depend on the temperature. The shape optimization problem for a motor geometry $D \subset \mathbb{R}^2$ and $T > 0$ reads:

$$\min_{\Omega \in \mathcal{A}} J(\Omega, A_z, \theta), \text{ such that} \tag{0.2a}$$

$$\sigma(\vartheta)\frac{\partial A_z}{\partial t} - \operatorname{div}\left(\nu \nabla A_z\right) = J_{src} - \operatorname{div}\left(H_0^{\perp}\left(\vartheta\right)\right) \qquad \text{in } Q_D := D \times (0,T), \tag{0.2b}$$

$$c_H \varrho \frac{\partial \vartheta}{\partial t} - \operatorname{div}\left(\lambda \nabla \vartheta\right) = \sigma(\vartheta)\left(\frac{\partial A_z}{\partial t}\right)^2 - J_{src}\frac{\partial A_z}{\partial t} \qquad \text{in } Q_D. \tag{0.2c}$$

Here, A_z denotes the third component of the vector potential, ϑ describes the temperature distribution, H_0 describes the magnetization and J_{src} denotes an impressed current density. The parameter σ describes the electrical conductivity, ν the reluctivity, c_H the specific heat capacity, ϱ the density and λ the heat conductivity. The admissible set of shapes is denoted by $\mathcal{A}$ and the objective function by J. Of course problem (0.2) is not complete for the moment since we have not specified the objective function J, the set $\mathcal{A}$ as well as boundary and initial conditions. These things will be stated later when we consider concrete problems. Further we want to mention that the system of PDEs is of nonlinear type. This is can be seen from the fact that the reluctivity ν is a nonlinear function depending on $|\nabla A_z|$ and the right hand side of (0.2c). Moreover, note that the second term in (0.2b) has to be understood in the distributional sense.

Based on (0.2) we are facing three different shape optimization problems: (i) a one phase-problem, (ii) an interface problem, (iii) a linear version of (0.2). We solve each problem by means of shape optimization techniques based on sensitivity information. Hence, this thesis is structured as follows: In Chapter 1 we derive the physical model (0.2b)-(0.2c) describing the electromagnetic-thermal coupling in an electric motor in more detail. Based on the physical model we state two shape optimization model problems: (i) a one phase-problem, (ii) an interface problem. In Chapter 2 we concentrate on the numerical solution of parabolic PDEs via a conforming space-time method proposed in [35]. We solve the state systems given by (i) and (ii) and determine the experimental order of convergence (eoc) of the L^2-error as well as the energy error. In Chapter 3 we give a brief introduction to the theory of shape optimization. We present standard results on shape calculus and explain the averaged adjoint method which we use to compute the shape derivative of the problems (i)-(iii). Further we describe how one can employ the shape derivative in order to compute a descent direction and state a generic algorithm. In Chapter 4 we apply the introduced methods from Chapter

2-3 to problem (i) and (ii). In Chapter 5 we apply the shape optimization techniques shown in the previous chapter to an electric motor. We consider (iii) on a motor geometry in 2D. We discuss the analysis of the state system, compute the shape derivative of the objective function and give a proof of concept of the numerical method used in Chapter 4.

1 Physical model and model problems

In this chapter we will introduce the reader to the physical model which is used to describe the electromagnetic-thermal coupling in an electrical motor. Therefore we will first give a brief introduction to electrical machines and explain some basic notions. Following that we will derive the physical model which describes the electromagnetic field as well as the heat propagation in an electrical motor. Finally we will state two model problems which are based on the physical model and which will be considered throughout this thesis.

1.1 Introduction to electrical machines

Electrical machines are part of our everyday life. They appear in the sector of mobility like in cars or bikes, in industrial applications like in automated production lines and in household applications like in washing machines. [12] They are also used in the field of energy production and renewable energies like in water power plants or wind power plants. All these different areas of applications of electrical machines lead to the fact that the improvement of the performance of an electrical machine is an active field of research.

Electrical machines can basically be distinguished between generators and motors. A generator converts mechanical energy into electrical energy while a motor converts electrical energy into mechanical energy. Electric motors can further be distinguished in direct current motors (DC-motors) and in alternating current motors (AC-motors). Two important types of AC-motors are the synchronous motor and asynchronous motor. Electric motors generally consist of a fixed part which is called stator and a rotating or moving part which is called rotor. [5, 43]

In this thesis we will deal with an asynchronous motor also called induction motor with a so called squirrel-cage rotor. In this type of motor the rotor rotates at a slower speed than the stator field. This is due to the fact that in an induction motor the torque is built by stator and rotor currents. In the starting phase, i.e., when the rotor is accelerated from its resting position, the magnetic field of the stator, which is caused due to the AC power supply, induces currents in the rotor bars. The induced currents create a magnetic field in the rotor which interacts with the stator field. Due to Lenz's law [42], the rotor will start to rotate in direction of the rotating stator field. The

maximal rotational speed of an induction motor is slightly lower than the rotational speed of the magnetic field in the stator since at synchronous speed no current would be induced in the rotor and therefore the torque would be zero. The difference between actual and synchronous speed is called slip. [41]

For more information on electrical machines we refer the reader to [5].

1.2 Physical model

In this section we will derive the physical model which describes the electromagnetic-thermal coupling in an electrical motor. In a first step we will derive the magneto-quasistatic formulation which is the common physical model to describe the magnetic field in an electrical motor in the starting phase. In a second step we will make use of a classical heat conduction equation in order to model the temperature propagation. The crucial part in this step will be in the choice of the right hand side of the heat equation in order to describe the impact of the magnetic field on the temperature in an appropriate way. Finally we will reduce the 3 dimensional model in space to a 2 dimensional model in space which is commonly used in the context of an electric motor. This section is based on [12, 16, 18, 24].

Starting point for the magnetoquasistatic formulation are Maxwell's equations [16]:

$$\operatorname{curl} \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}, \tag{1.1a}$$

$$\operatorname{curl} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{1.1b}$$

$$\operatorname{div} \boldsymbol{B} = 0, \tag{1.1c}$$

$$\operatorname{div} \boldsymbol{D} = \rho. \tag{1.1d}$$

The quantities which occur in (1.1) have the following meaning:

- $\boldsymbol{H} = \boldsymbol{H}(\boldsymbol{x}, t)$ is called magnetic field intensity,
- $\boldsymbol{E} = \boldsymbol{E}(\boldsymbol{x}, t)$ is called electric field intensity,
- $\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{x}, t)$ is called magnetic induction or magnetic flux density,
- $\boldsymbol{D} = \boldsymbol{D}(\boldsymbol{x}, t)$ is called electric flux density or electric induction,
- $\boldsymbol{J} = \boldsymbol{J}(\boldsymbol{x}, t)$ is called current density,
- $\rho = \rho(\boldsymbol{x}, t)$ electrical charge density.

Note that $\boldsymbol{H}, \boldsymbol{E}, \boldsymbol{B}, \boldsymbol{D}, \boldsymbol{J}$ are vector valued functions mapping form $\mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$. These quantities are connected via the material laws [16]

$$\boldsymbol{D} = \epsilon \boldsymbol{E} + \boldsymbol{P}, \tag{1.2a}$$

$$\boldsymbol{B} = \mu(\boldsymbol{H} + \boldsymbol{H}_0), \tag{1.2b}$$

$$\boldsymbol{J} = \sigma \boldsymbol{E} + \boldsymbol{J}_s, \tag{1.2c}$$

where ϵ denotes the electric permittivity, $\boldsymbol{P}$ the electric polarization, μ the magnetic permeability and σ the electric conductivity. The vector field $\boldsymbol{H}_0$ describes the magnetic field intensity which is for instance due to permanent magnets. One can see that the current density $\boldsymbol{J}$ is split up into a part $\boldsymbol{J}_s$ and a part $\boldsymbol{J}_c = \sigma \boldsymbol{E}$. $\boldsymbol{J}_s$ describes the impressed current density which is usually due to an external source. $\boldsymbol{J}_c$ describes the conduction currents which occur due to the fact that in materials where $\sigma \neq 0$ an electric field causes a current which is described by Ohm's law. An important class of conduction currents are eddy currents which occur due to a changing magnetic field [45].

In general ϵ, μ, σ are tensors of rank 2 which depend on space, time and the field quantities $\boldsymbol{E}$,$\boldsymbol{H}$. We will assume isotropic material throughout this thesis. Hence these quantities become scalar fields [24]. In addition to that we assume that these quantities are constant in time. Of special interest in the context of an electric motor is the magnetic reluctivity which is defined as the reciprocal of the magnetic permeability, i.e. $\nu := \mu^{-1}$. Since we are dealing with ferromagnetic material in an electric motor which is of nonlinear type the reluctivity turns out to be a nonlinear function depending on space and the magnitude of the magnetic flux, i.e. $\nu = \nu(\boldsymbol{x}, |\boldsymbol{B}|)$. For so called linear material the reluctivity is independent of the field intensities, cf. [45].

Furthermore we assume $\boldsymbol{P} = 0$ and we neglect $\frac{\partial \boldsymbol{D}}{\partial t}$ in (1.1a) which is a common assumption for electrical machines since they are low frequency applications [12]. Therefore (1.1d) decouples from (1.1a)-(1.1c) and the resulting system which is called magnetoquasistatic problem or eddy current problem reads

$$\operatorname{curl} \boldsymbol{H} = \boldsymbol{J}, \tag{1.3a}$$

$$\operatorname{curl} \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{1.3b}$$

$$\operatorname{div} \boldsymbol{B} = 0. \tag{1.3c}$$

We will now derive the vector potential formulation of (1.3) which is used to solve the eddy current problem numerically. Due to (1.3c) and the identity $\operatorname{div}\operatorname{curl}(.) = 0$ there exists a vector potential $\tilde{\boldsymbol{A}}$ such that

$$\boldsymbol{B} = \operatorname{curl} \tilde{\boldsymbol{A}}. \tag{1.4}$$

Note that the vector potential $\tilde{\boldsymbol{A}}$ is not unique. We can add a gradient field $\nabla\varphi$ to $\tilde{\boldsymbol{A}}$ which results in a solution to (1.4) as well. Plugging in (1.4) into (1.3b) yields

$$\operatorname{curl}\left(E + \frac{\partial \tilde{\boldsymbol{A}}}{\partial t}\right) = 0.$$

Due to the identity $\operatorname{curl}\nabla(.) = 0$ we conclude that there exists a scalar field φ such that

$$-\nabla\varphi = E + \frac{\partial \tilde{\boldsymbol{A}}}{\partial t}. \tag{1.5}$$

We are now able to introduce the gauging vector potential $\boldsymbol{A}$ by

$$\boldsymbol{A}(\boldsymbol{x},t) = \tilde{\boldsymbol{A}}(\boldsymbol{x},t) + \int_0^t \nabla\varphi(\boldsymbol{x},s)\,\mathrm{d}s \tag{1.6}$$

Note that $\boldsymbol{A}$ fulfills $\boldsymbol{B} = \operatorname{curl}\boldsymbol{A}$ which can be seen by an easy computation. Moreover we get out of the material law (1.2b) that

$$\boldsymbol{H} = \frac{1}{\mu}\boldsymbol{B} - \boldsymbol{H}_0 = \nu\boldsymbol{B} - \boldsymbol{H}_0. \tag{1.7}$$

Plugging in (1.7) into the left hand side of (1.3a) one gets

$$\operatorname{curl}\boldsymbol{H} = \operatorname{curl}(\nu\boldsymbol{B}) - \operatorname{curl}(\boldsymbol{H}_0) = \operatorname{curl}(\nu\operatorname{curl}\boldsymbol{A}) - \operatorname{curl}(\boldsymbol{H}_0) \tag{1.8}$$

The right hand side of (1.3a) turns out to be

$$\begin{aligned}\boldsymbol{J} = \boldsymbol{J}_c + \boldsymbol{J}_s = \sigma\boldsymbol{E} + \boldsymbol{J}_s = \sigma\left(-\nabla\varphi - \frac{\partial\tilde{\boldsymbol{A}}}{\partial t}\right) + \boldsymbol{J}_s \\ = \sigma\left(-\nabla\varphi - \frac{\partial\boldsymbol{A}}{\partial t} + \nabla\varphi\right) + \boldsymbol{J}_s = -\sigma\frac{\partial\boldsymbol{A}}{\partial t} + \boldsymbol{J}_s.\end{aligned} \tag{1.9}$$

Putting (1.8) and (1.9) together yields the vector potential formulation of (1.3) which reads

$$\sigma\frac{\partial\boldsymbol{A}}{\partial t} + \operatorname{curl}(\nu\operatorname{curl}\boldsymbol{A}) = \boldsymbol{J}_s + \operatorname{curl}(\boldsymbol{H}_0). \tag{1.10}$$

Once one has solved (1.10) the magnectic induction as well as the electric field intensity can be computed by $\boldsymbol{B} = \operatorname{curl}\boldsymbol{A}$ and $\boldsymbol{E} = -\frac{\partial\boldsymbol{A}}{\partial t}$, cf. [45].

In order to describe the heat distribution in an electrical motor we use a classical heat conduction equation of the form

$$c_H\varrho\frac{\partial\vartheta}{\partial t} - \operatorname{div}(\lambda\nabla\vartheta) = Q. \tag{1.11}$$

The quantities involved are

- $\vartheta = \vartheta(\boldsymbol{x},t)$ temperature distribution,
- c_H specific heat capacity,
- ϱ mass density,
- λ heat conductivity.

The quantities c_H, ϱ, λ are dependent on space, but constant in time. The term Q represents the Joule losses of the electromagnetic field [18]. We model these losses according to [18] by

$$Q = \boldsymbol{J}\cdot\boldsymbol{E} = (\boldsymbol{J}_c + \boldsymbol{J}_s)\cdot\boldsymbol{E} = \sigma\boldsymbol{E}\cdot\boldsymbol{E} + \boldsymbol{J}_s\cdot\boldsymbol{E} = \sigma\frac{\partial\boldsymbol{A}}{\partial t}\cdot\frac{\partial\boldsymbol{A}}{\partial t} - \boldsymbol{J}_s\cdot\frac{\partial\boldsymbol{A}}{\partial t}. \tag{1.12}$$

The first term in (1.12) is due to eddy currents while the second part is due to the Joule losses [18].

Accounting for the fact that the electrical conductivity σ as well as the properties of permanent magnets (i.e. $\boldsymbol{H}_0$) depend on the temperature ϑ, the physical model to describe the electromagnetic-thermal coupling reads

$$\sigma(\vartheta)\frac{\partial \boldsymbol{A}}{\partial t} + \operatorname{curl}(\nu \operatorname{curl} \boldsymbol{A}) = \boldsymbol{J}_s + \operatorname{curl}(\boldsymbol{H}_0(\vartheta)), \tag{1.13a}$$

$$c_H \varrho \frac{\partial \vartheta}{\partial t} - \operatorname{div}(\lambda \nabla \vartheta) = \sigma(\vartheta)\frac{\partial \boldsymbol{A}}{\partial t} \cdot \frac{\partial \boldsymbol{A}}{\partial t} - \boldsymbol{J}_s \cdot \frac{\partial \boldsymbol{A}}{\partial t}. \tag{1.13b}$$

In order to solve problem (1.13) numerically one has to introduce boundary, initial as well as interface conditions. Let therefore $\Omega \times (0,T)$ with $\Omega \subset \mathbb{R}^3$, $T > 0$ be the computational domain with $\Gamma := \partial\Omega$. Further denote by n the outer unit normal vector to the computational domain. One possible choice of boundary and initial conditions is to choose

$$\boldsymbol{A} \times n = 0 \qquad \text{on } \Gamma \times (0,T), \tag{1.14a}$$

$$\vartheta = \vartheta_0 \qquad \text{on } \Gamma \times (0,T), \tag{1.14b}$$

$$A(.,0) = 0 \qquad \text{in } \Omega, \tag{1.14c}$$

$$\vartheta(.,0) = \vartheta_0 \qquad \text{in } \Omega. \tag{1.14d}$$

Condition (1.14a) implies that $\boldsymbol{B} \cdot n = 0$ on Γ [24]. This means that no magnetic flux can leave the computational domain. This is also called induction boundary condition and is a common choice in the context of an electric motor [12]. The initial condition of $\boldsymbol{A}$ corresponds to the case that at $t = 0$ the current $\boldsymbol{J}_s$ is switched on and the motor is accelerated from its resting position. We assume that the temperature at initial time as well as on the boundary of the motor are constant and given by room temperature. For more information on the different types of boundary conditions on the electromagnetic field we refer the reader to [6],[45].

Since we have different materials in the motor we need interface conditions as well. Let therefore $\Gamma_I \times (0,T)$ represent the material interface where the reluctivity ν and the heat conductivity λ jumps. Further denote by $[\![v]\!]$ the jump of a function v along the interface, i.e.

$$[\![v]\!] = v^+\big|_{\Gamma_I \times (0,T)} - v^-\big|_{\Gamma_I \times (0,T)},$$

where v^+ and v^- are the restriction of v to the corresponding subdomains, cf. [12]. The interface conditions are given by

$$[\![\boldsymbol{B} \cdot n]\!] = 0 \quad \text{on } \Gamma_I \times (0,T), \tag{1.15a}$$

$$[\![\boldsymbol{H} \times n]\!] = 0 \quad \text{on } \Gamma_I \times (0,T), \tag{1.15b}$$

$$[\![\vartheta]\!] = 0 \quad \text{on } \Gamma_I \times (0,T), \tag{1.15c}$$

$$[\![\lambda \nabla \vartheta \cdot n]\!] = 0 \quad \text{on } \Gamma_I \times (0,T). \tag{1.15d}$$

Under certain assumptions it is possible to reduce (1.13) to a two dimensional model in space. The assumptions read [12, 24]:

(i) $\Omega = \tilde{\Omega} \times (-l, l)$ with $l >> \operatorname{diam}(\tilde{\Omega})$,

(ii) $\boldsymbol{J}_s, \boldsymbol{H}_0$ and $\boldsymbol{H}$ are constant with respect to the third spatial coordinate and of the form

$$\boldsymbol{J}_s = \begin{pmatrix} 0 \\ 0 \\ J_3(x_1, x_2, t) \end{pmatrix}, \boldsymbol{H}_0 = \begin{pmatrix} H_{01}(x_1, x_2, t) \\ H_{02}(x_1, x_2, t) \\ 0 \end{pmatrix}, \boldsymbol{H} = \begin{pmatrix} H_1(x_1, x_2, t) \\ H_2(x_1, x_2, t) \\ 0 \end{pmatrix},$$

for $(x_1, x_2, t) \in \tilde{\Omega} \times (0, T)$.

One can immediately see that due the assumptions and the constitutive equation (1.2b) the vector field $\boldsymbol{B}$ is of the same form as $\boldsymbol{H}, \boldsymbol{H}_0$. Since there has to hold $\boldsymbol{B} = \operatorname{curl} \boldsymbol{A}$ this can be achieved by the ansatz

$$\boldsymbol{A} = \begin{pmatrix} 0 \\ 0 \\ u(x_1, x_2, t) \end{pmatrix} \text{ for } (x_1, x_2, t) \in \tilde{\Omega} \times (0, T).$$

Note that this ansatz also ensures the so called Coulomb condition $\operatorname{div} \boldsymbol{A} = 0$. One can show [12], [24] that through this ansatz the physical model (1.13) reduces to the following initial boundary value problem: Find $u, p : \tilde{\Omega} \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned}
\sigma(p)\partial_t u - \operatorname{div}(\nu \nabla u) &= J_3 - \operatorname{div} H_0^{\perp}(p) && \text{in } \tilde{\Omega} \times (0, T), && (1.16a)\\
c_H \varrho \partial_t p - \operatorname{div}(\lambda \nabla p) &= \sigma(p) (\partial_t u)^2 - J_3 \partial_t u && \text{in } \tilde{\Omega} \times (0, T), && (1.16b)\\
u &= 0 && \text{on } \partial\tilde{\Omega} \times (0, T), && (1.16c)\\
p &= p_0 && \text{on } \partial\tilde{\Omega} \times (0, T), && (1.16d)\\
u(., 0) &= 0 && \text{in } \tilde{\Omega}, && (1.16e)\\
p(., 0) &= p_0 && \text{in } \tilde{\Omega}, && (1.16f)\\
[\![u]\!] &= 0 && \text{on } \Gamma_I \times (0, T), && (1.16g)\\
[\![\nu \nabla u \cdot n]\!] &= 0 && \text{on } \Gamma_I \times (0, T), && (1.16h)\\
[\![p]\!] &= 0 && \text{on } \Gamma_I \times (0, T), && (1.16i)\\
[\![\lambda \nabla p \cdot n]\!] &= 0 && \text{on } \Gamma_I \times (0, T), && (1.16j)
\end{aligned}$$

where $\nu = \nu(x, |\nabla u|)$, $p(x, t) := \vartheta(x, t)$ and $H_0^{\perp} := (-H_{02}, H_{01})^T$. Problem (1.16) is a system of coupled nonlinear parabolic PDEs. The nonlinearity can be seen from the fact that the parameter ν is a nonlinear function with respect to $|\nabla u|$ and the right hand side in the second equation.

1.3 Model problems

Based on (1.16) we will formulate two PDE-constrained shape optimization problems which will be considered throughout this thesis. We make the following assumptions:

Assumption 1.1. Let $D \subset \mathbb{R}^d$ with $d \in \{1,2,3\}$ a bounded Lipschitz domain, $\mathcal{A}(D) := \{\Omega \subset D : \ \Omega \text{ open, Lipschitz with uniform Lipschitz constant } L_{\mathcal{A}}\}$ and $T > 0$. Further let $Q_D := D \times (0,T)$ and for $\Omega \in \mathcal{A}(D)$ let $Q_\Omega := \Omega \times (0,T)$. Moreover denote the boundaries of the space-time cylinder with $\Sigma_\Omega := \partial\Omega \times (0,T)$, $\Sigma_{\Omega 0} := \overline{\Omega} \times \{0\}$, $\Sigma_D := \partial D \times (0,T)$ and $\Sigma_{D0} := \overline{D} \times \{0\}$.

The first shape optimization problem reads

$$\min_{\Omega \in \mathcal{A}(D)} J_1(\Omega, u, p) := \int_0^T \int_\Omega |u - u_d|^2 \,\mathrm{d}x\mathrm{d}t + \int_0^T \int_\Omega |p - p_d|^2 \,\mathrm{d}x\mathrm{d}t, \tag{1.17}$$

where (u,p) is solution to

$$\begin{aligned} \partial_t u - \Delta u &= f_1 + p && \text{in } Q_\Omega, && \text{(1.18a)}\\ \partial_t p - \Delta p &= f_2 + u && \text{in } Q_\Omega, && \text{(1.18b)}\\ u &= 0 && \text{on } \Sigma_\Omega \cup \Sigma_{\Omega 0}, && \text{(1.18c)}\\ p &= 0 && \text{on } \Sigma_\Omega \cup \Sigma_{\Omega 0}. && \text{(1.18d)} \end{aligned}$$

and $u_d, p_d, f_1, f_2 \in C^1(\overline{Q_D})$ are given functions. The functional J_1 is of so called tracking type. The goal is to find a shape $\Omega^\star$ such that the states $u = u(\Omega), p = p(\Omega)$ which depend on the shape Ω come as close as possible to given functions u_d, p_d and fulfill (1.18).

In the second shape optimization problem we consider a transmission problem. We assume that the material parameters are piecewise constant functions of the form

$$\nu(x,t) = \begin{cases} \nu_1 & \text{for } (x,t) \in Q_\Omega \\ \nu_2 & \text{for } (x,t) \in Q_D \setminus Q_\Omega \end{cases}, \quad \lambda(x,t) = \begin{cases} \lambda_1 & \text{for } (x,t) \in Q_\Omega \\ \lambda_2 & \text{for } (x,t) \in Q_D \setminus Q_\Omega \end{cases}. \tag{1.19}$$

Thus the material coefficients describe linear material behavior. We denote by $\Gamma_I \times (0,T)$ the interface where the material parameters jump. The problem reads

$$\min_{\Omega \in \mathcal{A}(D)} J_2(u, p) := \int_0^T \int_D |u - u_d|^2 \,\mathrm{d}x\mathrm{d}t + \int_0^T \int_D |p - p_d|^2 \,\mathrm{d}x\mathrm{d}t, \tag{1.20}$$

where (u, p) is solution to

$$\partial_t u - \operatorname{div}(\nu \nabla u) = f_1 + p \qquad \text{in } Q_D, \tag{1.21a}$$
$$\partial_t p - \operatorname{div}(\lambda \nabla p) = f_2 + u \qquad \text{in } Q_D, \tag{1.21b}$$
$$u = 0 \qquad \text{on } \Sigma_D \cup \Sigma_{D0}, \tag{1.21c}$$
$$p = 0 \qquad \text{on } \Sigma_D \cup \Sigma_{D0}, \tag{1.21d}$$
$$[\![u]\!] = 0 \qquad \text{on } \Gamma_I \times (0,T), \tag{1.21e}$$
$$[\![\nu \nabla u \cdot n]\!] = 0 \qquad \text{on } \Gamma_I \times (0,T), \tag{1.21f}$$
$$[\![p]\!] = 0 \qquad \text{on } \Gamma_I \times (0,T), \tag{1.21g}$$
$$[\![\lambda \nabla p \cdot n]\!] = 0 \qquad \text{on } \Gamma_I \times (0,T), \tag{1.21h}$$

where $u_d, p_d, f_1, f_2 \in C^1(\overline{Q_D})$ are given functions. The functional J_2 is again of tracking type. Note the difference between J_1 and J_2. While J_1 also depends explicitly on the shape Ω the dependence of J_2 on the shape Ω is given only implicitly through the states u and p.

2 Parabolic PDE system

In this chapter we focus on the parabolic PDE system as seen in (1.18) and (1.21), respectively. At this point we want to mention that we will not treat any analysis of the coupled system, since this task is more involved and thus is beyond the scope of this master thesis. Instead we will numerically solve the coupled system with a space-time method according to the approach of Steinbach in [35],[37] and determine the experimental order of convergence of the error in the L^2-norm and energy norm. The idea of space-time methods is to treat the time variable as additional space coordinate. In contrast to other solution methods like the vertical method of lines or Rothe's method one has to deal with a problem in one dimension higher, which leads to a higher computational effort. The main advantage of the space-time method described in [35] is that it is possible to use adaptive finite element meshes since the discretization is done simultaneously in space and time. Another advantage lies in the possibility to parallelize the method. Moreover the proposed approach in [35] allows to use continuous ansatz and test spaces.

Throughout this chapter let $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$ be a bounded domain with Lipschitz boundary $\Gamma := \partial\Omega$. Further let $T > 0$ and $Q := \Omega \times (0, T)$ be the space-time cylinder. Moreover let $\Sigma := \Gamma \times (0, T)$, $\Sigma_0 := \overline{\Omega} \times \{0\}$, $\Sigma_T := \overline{\Omega} \times \{T\}$ such that $\partial Q = \Sigma \cup \Sigma_0 \cup \Sigma_T$. For given functions f_1 and f_2 we want to find (u, p) such that

$$\partial_t u - \Delta u = f_1 + p \qquad \text{in } Q, \tag{2.1a}$$

$$\partial_t p - \Delta p = f_2 + u \qquad \text{in } Q, \tag{2.1b}$$

$$u = 0 \qquad \text{on } \Sigma \cup \Sigma_0, \tag{2.1c}$$

$$p = 0 \qquad \text{on } \Sigma \cup \Sigma_0. \tag{2.1d}$$

This chapter is organized as follows: We start with some preliminaries in which we introduce the function spaces and basic definitions which we are going to use throughout this section. Afterwards we will treat the variational formulation of problem (2.1). Following that we will discuss the discretization and last but not least we will consider some numerical examples.

2.1 Preliminaries

In this section we are going to introduce the so called Bochner spaces which we will need to state the variational formulation. In addition to that we will present some basic definitions which we will need to discretize (2.1).

2.1.1 Bochner spaces

The approach in [35] uses Bochner spaces. We will give a brief introduction to them. The basic definitions and results are taken from [26], [29], [46].

Definition 2.1 ([29, Def. 2.1, p. 25])**.** Let $(X, \|.\|_X)$ be a Banach space, $p \in [1, \infty]$ and $T > 0$. The space of p-integrable functions is denoted by

$$\mathcal{L}^p(0,T;X) := \left\{ f : (0,T) \to X \text{ measurable}, \ \int_0^T \|f(t)\|_X^p \, \mathrm{d}t < \infty \right\}, \ p \in [1,\infty),$$

$$\mathcal{L}^\infty(0,T;X) := \left\{ f : (0,T) \to X \text{ measurable}, \ \operatorname*{ess\,sup}_{t\in(0,T)} \|f(t)\|_X < \infty \right\}.$$

Moreover we call

$$L^p(0,T;X) := \mathcal{L}^p(0,T;X)/_\sim, \text{ with } f \sim g :\Leftrightarrow f = g \text{ almost everywhere (a.e.)},$$

Bochner space of p-integrable functions.

One can show [26, Satz 1.22], that $L^p(0,T;X)$ is a Banach space with respect to the norm

$$\|f\|_{L^p(0,T;X)} := \left(\int_0^T \|f(t)\|_X^p \, \mathrm{d}t \right)^{\frac{1}{p}}, \ p \in [1,\infty),$$

$$\|f\|_{L^\infty(0,T;X)} := \operatorname*{ess\,sup}_{t\in(0,T)} \|f(t)\|_X.$$

We will also need a characterization of the dual space $(L^p(0,T;X))'$. One can show [29, Thm. 3.13] that for $p \in [1,\infty)$, $\frac{1}{p} + \frac{1}{q} = 1$ and a reflexive Banach space X the mapping

$$T : L^q(0,T;X') \to (L^p(0,T;X))', \ u \mapsto Tu$$

$$\langle Tu, v \rangle := \int_0^T \langle u(t), v(t) \rangle_{X',X} \, \mathrm{d}t, \ v \in L^p(0,T;X),$$

is an isometric isomorphism. Thus we can identify $(L^p(0,T;X))'$ with $L^q(0,T;X')$.

In order to state the variational formulation we further need so called generalized Sobolev spaces. For this reason we first introduce the concept of a generalized derivative for a function $u \in L^p(0,T;X)$.

Definition 2.2 ([29, Def. 4.17, p. 54])**.** Let $(X, \|.\|_X)$ be a Banach space, $T > 0$, and $u \in L^1_{loc}(0,T;X)$. The function u obtains a generalized n-th derivative if there exists a function $v \in L^1_{loc}(0,T;X)$ such that

$$\int_0^T v(t)\varphi(t) \, \mathrm{d}t = (-1)^n \int_0^T u(t)\varphi^{(n)}(t) \, \mathrm{d}t \quad \forall \varphi \in C_0^\infty(0,T). \tag{2.2}$$

We write $\frac{\mathrm{d}^n u}{\mathrm{d}t^n} := v$ and call $\frac{\mathrm{d}^n u}{\mathrm{d}t^n}$ n-th generalized derivative of u.

We are now able to state the definition of generalized Sobolev spaces.

Definition 2.3 ([29, Def. 5.1, p. 63])**.** Let $(X, \|.\|_X)$ be a Banach space and $p \geq 1$. The space

$$W^{1,p}(0,T;X) := \left\{ u \in L^p(0,T;X) : \ \frac{\mathrm{d}u}{\mathrm{d}t} \in L^p(0,T;X) \right\} \tag{2.3}$$

is called generalized Sobolev space where $\frac{\mathrm{d}u}{\mathrm{d}t}$ has to be understood in the sense of Definition 2.2. For $p = 2$ we also use the notation

$$H^1(0,T;X) := W^{1,2}(0,T;X).$$

One can show [29, Thm.5.2] that the space $W^{1,p}(0,T;X)$ for $p \geq 1$ is a Banach space with respect to the norm

$$\|u\|_{W^{1,p}(0,T;X)} := \|u\|_{L^p(0,T;X)} + \left\| \frac{\mathrm{d}u}{\mathrm{d}t} \right\|_{L^p(0,T;X)}.$$

For more information on the generalized derivative as well as generalized Sobolev spaces we refer the reader to [29, Chapter 4,5], [46, Chapter 23].

2.1.2 Basic notion of finite element discretization

In this section we will recall some basic notions of finite element discretization. For the rest of this section we consider $\Omega \subset \mathbb{R}^d$ to be a bounded Lipschitz domain. The definitions are taken from [8],[11]. We start with the definition of a simplex.

Definition 2.4. Let $s \in \{0, 1, ..., d\}$ and $z_1, ..., z_{s+1} \in \mathbb{R}^d$ be given, such that the vectors $(z_j - z_1)_{j=2,...,s+1}$ are linearly independent. Then $T = \operatorname{conv}\{z_1, ..., z_{s+1}\}$ is called a non-degenerate s-dimensional simplex in $\mathbb{R}^d$. The corner points $z_1, ..., z_{s+1}$ are called nodes of T. For $z'_1, ..., z'_{r+1} \in \{z_1, ..., z_{s+1}\}$ and $r \in \{0, ..., s-1\}$ the convolution $T' = \operatorname{conv}\{z'_1, ..., z'_{r+1}\}$ is called r-dimensional face of the simplex T.

In $\mathbb{R}^2$ a 2-dimensional simplex is called triangle and the 1-dimensional face of the triangle are called edges. In $\mathbb{R}^3$ a 3-dimensional simplex is called tetrahedron. The 2-dimensional faces of the tetrahedron are called surfaces, the 1-dimensional faces are called edges. The next definition states the notion of triangulation, which is a basic concept when dealing with finite elements.

Definition 2.5. The set $\mathcal{T} := \{T_1, ..., T_M\}$ is called triangulation or mesh of $\Omega \subset \mathbb{R}^d$, if the following conditions hold:

(i) For each $i \in \{1, ..., M\}$: T_i is a d-dimensional simplex,

(ii) $\overline{\Omega} = \bigcup_{T \in \mathcal{T}} T$,

(iii) $\forall T, T' \in \mathcal{T}$ with $T \neq T'$ holds: $|T \cap T'| = 0$.

The triangulation $\mathcal{T}$ is called admissible, if for all $T, T' \in \mathcal{T}$ with $T \neq T'$ one of the following statements hold:

- $T \cap T' = \emptyset$,
- $T \cap T' = S$, where S is a r-dimensional face of T and T', where $r \in \{0, ..., d-1\}$.

In numerical simulations we use shape regular triangulations which are defined in the following way:

Definition 2.6 ([11, Def. 1.107])**.** A family of admissible triangulations $(\mathcal{T}_h)_{h \in I}$ is called shape regular if there exists a constant $\sigma_0 > 0$ such that

$$\forall h \in I: \ \sigma(\mathcal{T}_h) := \max_{T \in \mathcal{T}_h} \frac{h(T)}{\rho(T)} < \sigma_0,$$

where $h(T)$ and $\rho(T)$ are defined by

$$h(T) := \max_{x,y \in T} \|x - y\|_2, \ \rho(T) := 2 \max_{x \in T} \left\{ r \in \mathbb{R}_+ : \ \overline{B_r(x)} \subset T \right\}.$$

The finite element space which we are going to use throughout this master thesis is the space of globally continuous and piecewise polynomial functions of order k defined as

$$S^k(\mathcal{T}) := \left\{ v \in C\left(\overline{\Omega}\right) : \ v|_T \in P_k(T) \ \forall T \in \mathcal{T} \right\}, \text{ where}$$

$$P_k(T) := \left\{ v : T \to \mathbb{R} : \ v(x) = \sum_{\substack{\alpha \in \mathbb{N}_0^d \\ |\alpha| \leq k}} c_\alpha x^\alpha, \ c_\alpha \in \mathbb{R} \right\}.$$

2.2 Variational formulation

We derive a weak formulation of (2.1) in the usual manner by multiplying with a test function, integrating over the space-time cylinder and by applying Green's formula. By introducing the spaces $X := L^2(0,T; H_0^1(\Omega)) \cap H^1(0,T; H^{-1}(\Omega))$, $X_0 := \left\{ v \in X : \ v|_{\Sigma_0} = 0 \right\}$ and $Y := L^2(0,T; H_0^1(\Omega))$ the variational formulation of (2.1) is to find $(u,p) \in X_0 \times X_0$ such that

$$\int_Q \partial_t u v \, \mathrm{d}x\mathrm{d}t + \int_Q \nabla_x u \cdot \nabla_x v \, \mathrm{d}x\mathrm{d}t = \int_Q f_1 v \, \mathrm{d}x\mathrm{d}t + \int_Q p v \, \mathrm{d}x\mathrm{d}t \quad \forall v \in Y, \tag{2.4a}$$

$$\int_Q \partial_t p q \, \mathrm{d}x\mathrm{d}t + \int_Q \nabla_x p \cdot \nabla_x q \, \mathrm{d}x\mathrm{d}t = \int_Q f_2 q \, \mathrm{d}x\mathrm{d}t + \int_Q u q \, \mathrm{d}x\mathrm{d}t \quad \forall q \in Y. \tag{2.4b}$$

By introducing the bilinear forms

$$a(u,v) := \int_Q \partial_t uv \,\mathrm{d}x\mathrm{d}t + \int_Q \nabla_x u \cdot \nabla_x v \,\mathrm{d}x\mathrm{d}t$$
$$b(p,v) := \int_Q pv \,\mathrm{d}x\mathrm{d}t$$

the variational formulation (2.4) reads: Find $(u,p) \in X_0 \times X_0$ such that

$$a(u,v) - b(p,v) = \langle f_1, v\rangle_Q \quad \forall v \in Y, \tag{2.5a}$$
$$a(p,q) - b(u,q) = \langle f_2, q\rangle_Q \quad \forall q \in Y. \tag{2.5b}$$

Note that (2.5) is a Petrov-Galerkin variational formulation meaning that ansatz and test spaces are different. Hence it is necessary to establish an $\inf - \sup$ condition in order to show existence and uniqueness of a solution to (2.5). In [35] and [37] a stability estimate is derived for a problem of the form

$$c_H \partial_t u - \operatorname{div}(A\nabla u) = f \qquad \text{in } Q, \tag{2.6a}$$
$$u = 0 \qquad \text{on } \Sigma, \tag{2.6b}$$
$$u = u_0 \qquad \text{on } \Sigma_0, \tag{2.6c}$$

where $c_H > 0$ is a given heat capacity, A a symmetric and positive definite coefficient matrix and $u_0 \in H_0^1(\Omega)$ some initial data. It is however not clear if the proof of this estimate can be generalized to a coupled problem like (2.5).

2.3 Discretization

Let $(\mathcal{T}_h)_{h\in I}$ be a shape regular family of triangulations of the space-time cylinder Q. Further let $X_{0h} \subset X$, $Y_h \subset Y$ and as in the continuous case $X_0 \subset Y$ we assume the inclusion $X_{0h} \subset Y_h$ [35]. The discrete version of the Galerkin-Petrov variational formulation (2.5) is to find $(u_h, p_h) \in X_{0h} \times X_{0h}$ such that

$$a(u_h, v_h) - b(p_h, v_h) = \langle f_1, v_h\rangle_Q \quad \forall v_h \in Y_h, \tag{2.7a}$$
$$a(p_h, q_h) - b(u_h, q_h) = \langle f_2, q_h\rangle_Q \quad \forall q_h \in Y_h. \tag{2.7b}$$

2.4 Numerical examples

In this section we will numerically solve the coupled problem (2.1) which also comes up in the shape optimization problem (1.17)-(1.18). Further we will consider the

interface problem which arises in the shape optimization problem (1.20)-(1.21). We will present numerical solutions to (2.1) as well as (1.21) in one (d=1) and two (d=2) space dimensions, respectively.

In all numerical tests we determine the experimental order of convergence in the L^2-norm as well as in the energy norm. In order to obtain these error rates we have to compute the error of the coupled problems in total. Hence, we define

$$Z := \begin{bmatrix} u \\ p \end{bmatrix}, \; Z_h := \begin{bmatrix} u_h \\ p_h \end{bmatrix}$$

and the total errors by

$$\|Z - Z_h\|_{L^2(Q)} := \sqrt{\|u - u_h\|^2_{L^2(Q)} + \|p - p_h\|^2_{L^2(Q)}},$$
$$\|Z - Z_h\|_{L^2(0,T;H^1_0(\Omega))} := \sqrt{\|u - u_h\|^2_{L^2(0,T;H^1_0(\Omega))} + \|p - p_h\|^2_{L^2(0,T;H^1_0(\Omega))}}.$$

The mesh generation as well as the numerical simulation is done in the finite element software Netgen/NGSolve [31]. In all numerical tests the resulting linear system is built using a piecewise linear and globally continuous finite element space ($k = 1$) where it is sufficient to consider $Y_h = X_h$ [35]. The linear system is solved using the package UMFPACK.

2.4.1 Examples in 1d-1d

Example 1

In this example we consider (2.1) on the unit square, i.e. $d = 1, T = 1, Q = (0,1)^2$. We choose the exact solutions to be

$$u(x,t) = \sin(\pi x)\sin(\pi t) \quad \text{for } (x,t) \in Q,$$
$$p(x,t) = x(1-x)t(1-t) \quad \text{for } (x,t) \in Q.$$

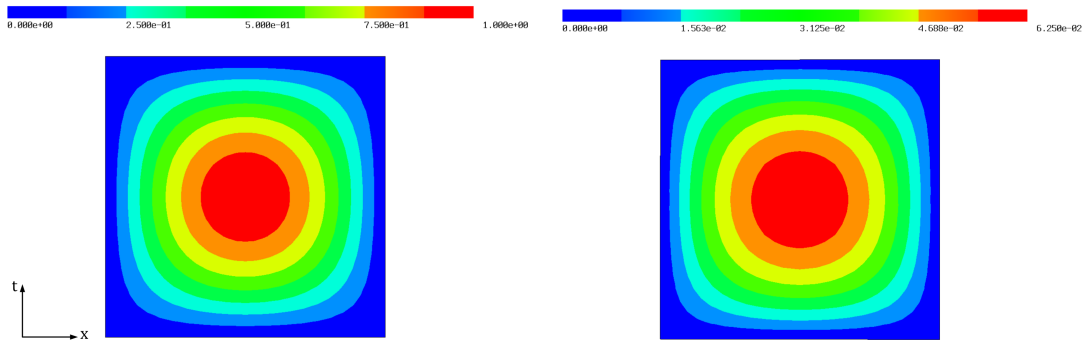
The functions f_1, f_2 are given by

$$f_1(x,t) = -p(x,t) + \pi\cos(\pi t)\sin(\pi x) + \pi^2\sin(\pi t)\sin(\pi x) \quad \text{for } (x,t) \in Q, \tag{2.8a}$$
$$f_2(x,t) = -u(x,t) + x(1-x)(1-2t) + 2t(1-t) \quad \text{for } (x,t) \in Q. \tag{2.8b}$$

In Table 2.1 one can see the errors in the $L^2(Q)$-norm and energy norm as well as the experimental order of convergence for different refinement levels. The order of convergence is quadratic in the $L^2(Q)$-norm and linear in the energy norm. In Figure 2.1 one can see the numerical solutions u_h, p_h for the refinement level $L = 4$.

L	N	$\|Z - Z_h\|_{L^2(Q)}$	eoc	$\|Z - Z_h\|_{L^2(0,T;H_0^1(\Omega))}$	eoc
0	8	3.645e-01	0.000	1.253e+00	0.000
1	21	1.007e-01	1.855	6.277e-01	0.998
2	65	2.537e-02	1.989	3.204e-01	0.970
3	225	6.235e-03	2.025	1.609e-01	0.993
4	833	1.521e-03	2.035	8.063e-02	0.997
5	3201	3.747e-04	2.021	4.034e-02	0.999
6	12545	9.318e-05	2.008	2.017e-02	1.000
7	49665	2.325e-05	2.003	1.009e-02	1.000

Table 2.1: Error table for problem (2.1) with $d = 1$ and $k = 1$



(a) u_h (b) p_h

Figure 2.1: Numerical solutions u_h, p_h for (2.1) with $d = 1$ and $k = 1$ on refinement level $L = 4$

Example 2

In this example we consider the transmission problem (1.21) for $d = 1$. Before we discuss the settings for the numerical test in more detail let us recall problem (1.21).

Let $D \subset \mathbb{R}^d$ be a bounded Lipschitz domain and $\Omega \subset D$ open with Lipschitz boundary. Denote $Q_D := D \times (0,T), Q_\Omega := \Omega \times (0,T), \Sigma_D = \partial D \times (0,T), \Sigma_{D0} = \overline{D} \times \{0\}$. We consider piecewise constant material coefficients which are given by

$$\nu(x,t) = \begin{cases} \nu_1 & \text{for } (x,t) \in Q_\Omega \\ \nu_2 & \text{for } (x,t) \in Q_D \setminus Q_\Omega \end{cases}, \quad \lambda(x,t) = \begin{cases} \lambda_1 & \text{for } (x,t) \in Q_\Omega \\ \lambda_2 & \text{for } (x,t) \in Q_D \setminus Q_\Omega \end{cases}. \tag{2.9}$$

Note that we assume the material parameters to be constant in time. We denote by $\Gamma_I \times (0,T)$ the material interface where the parameters jump. The interface problem

reads: For given functions f_1, f_2 find functions u, p such that

$$\partial_t u - \operatorname{div}(\nu \nabla u) = f_1 + p \quad \text{in } Q_D, \tag{2.10a}$$
$$\partial_t p - \operatorname{div}(\lambda \nabla p) = f_2 + u \quad \text{in } Q_D, \tag{2.10b}$$
$$u = 0 \quad \text{on } \Sigma_D \cup \Sigma_{D0}, \tag{2.10c}$$
$$p = 0 \quad \text{on } \Sigma_D \cup \Sigma_{D0}, \tag{2.10d}$$
$$[\![u]\!] = 0 \quad \text{on } \Gamma_I \times (0,T), \tag{2.10e}$$
$$[\![\nu \nabla u \cdot n]\!] = 0 \quad \text{on } \Gamma_I \times (0,T), \tag{2.10f}$$
$$[\![p]\!] = 0 \quad \text{on } \Gamma_I \times (0,T), \tag{2.10g}$$
$$[\![\lambda \nabla p \cdot n]\!] = 0 \quad \text{on } \Gamma_I \times (0,T). \tag{2.10h}$$

For the numerical simulation we choose the computational domain $Q_D = (0,1)^2$ and $Q_\Omega = (\frac{1}{4}, \frac{3}{4}) \times (0,1)$. We assume that the domain Q_Ω corresponds to iron (ferromagnetic) and that $Q_D \setminus Q_\Omega$ corresponds to air (non-ferromagnetic). In view of this assumption we choose the parameters $\nu_1 = \frac{1}{\mu_0 \cdot \mu_{r1}}, \nu_2 = \frac{1}{\mu_0 \cdot \mu_{r2}}$ where $\mu_0 = \frac{4\pi}{10^7}$ is the magnetic permeability of vacuum, $\mu_{r1} = 5100$ the relative permeability of the ferromagnetic material and $\mu_{r2} = 1$ the relative permeability of air. The values for the parameter λ read $\lambda_1 = 80.2$ and $\lambda_2 = 0.0262$. The computational domain Q_D is depicted in Figure 2.2 in which blue corresponds to iron and red corresponds to air. We choose the functions f_1, f_2 as in (2.8). Note that we do not know the exact solu-

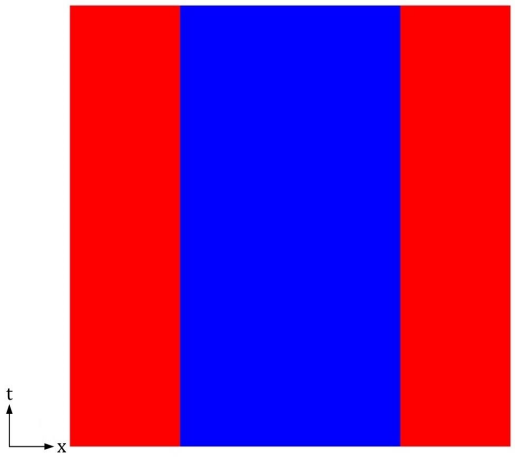


Figure 2.2: Computational domain Q_D, blue: iron, red: air

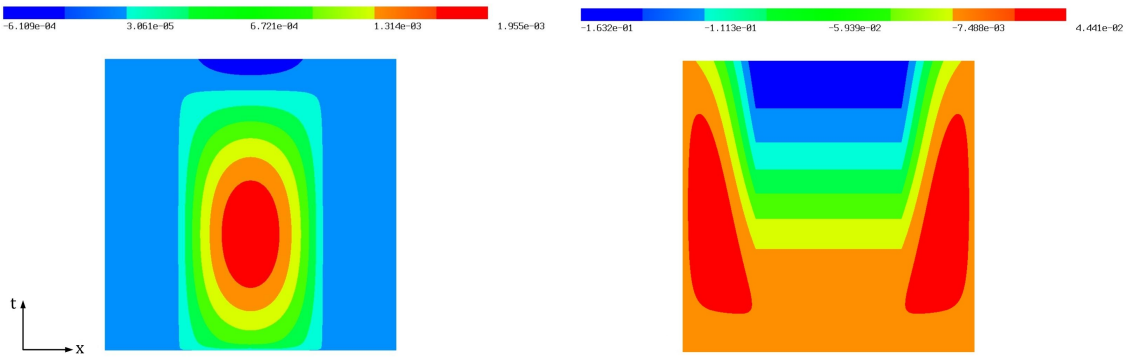
tion for this problem. One possible way to work around this problem is to compute the numerical solutions u_{href}, p_{href} on a very fine mesh and consider them as exact solution, i.e. $u := u_{href}, p := p_{href}$. Hence it is possible to obtain error rates, but one has to treat the obtained rates with care since the numerical solutions u_h and p_h converge in a natural way to u_{href}, p_{href} as the mesh size gets smaller. We compute

the reference solutions on a mesh with $N = 307823$ nodes (dofs). The obtained error rates are shown in Table 2.2. We observe that the rates increase in every refinement step. This is perhaps due to the fact that we do not compute the error against the exact solution. The rates increase to a value of 2.1 for the error in the $L^2(Q_D)$-norm and a value of 1.5 in the energy norm which is better than expected. Figure 2.3 shows

L	N	$\|Z - Z_h\|_{L^2(Q_D)}$	eoc	$\|Z - Z_h\|_{L^2(0,T;H_0^1(D))}$	eoc
0	24	5.238e-02	0.000	3.718e-01	0.000
1	77	3.500e-02	0.582	2.447e-01	0.603
2	273	2.879e-02	0.282	1.527e-01	0.681
3	1025	1.600e-02	0.847	8.129e-02	0.909
4	3969	5.581e-03	1.519	3.185e-02	1.352
5	15617	1.225e-03	2.187	1.098e-02	1.537

Table 2.2: Error rates for problem (2.10) with $d = 1$ and $k = 1$

the numerical solutions u_h, p_h on refinement level 5.



(a) u_h (b) p_h

Figure 2.3: Numerical solutions u_h, p_h for (2.10) with $d = 1$, $k = 1$ on refinement level $L = 5$

2.4.2 Examples in 2d-1d

Example 1

In this example we consider (2.1) on the unit cube, i.e. $d = 2, T = 1, Q = (0,1)^3$. We choose the exact solutions to be

$$\begin{aligned} u(x,y,t) &= \sin(\pi x)\sin(\pi y)\sin(\pi t) \quad \text{for } (x,y,t) \in Q, \\ p(x,y,t) &= x(1-x)y(1-y)t(1-t) \quad \text{for } (x,y,t) \in Q. \end{aligned}$$

The functions f_1, f_2 are given by

$$\begin{aligned} f_1(x,y,t) &= -p(x,y,t) + \pi \sin(\pi x)\sin(\pi y)\cos(\pi t) + 2\pi^2 u(x,y,t) \\ &\quad \text{for } (x,y,t) \in Q, \qquad (2.11a) \\ f_2(x,y,t) &= -u(x,y,t) + x(1-x)y(1-y)(1-2t) + 2y(1-y)t(1-t) \\ &\quad + 2x(1-x)t(1-t) \quad \text{for } (x,y,t) \in Q. \qquad (2.11b) \end{aligned}$$

The numerical results can be seen in Table 2.3. We observe an experimental order of convergence of up to 1.7 for the error in the $L^2(Q)$-norm and a linear order of convergence in the energy norm.

L	N	$\|Z - Z_h\|_{L^2(Q)}$	eoc	$\|Z - Z_h\|_{L^2(0,T;H_0^1(\Omega))}$	eoc
0	21	6.200e-02	0.000	6.977e-01	0.000
1	87	7.045e-02	-0.184	6.610e-01	0.078
2	469	3.805e-02	0.888	4.591e-01	0.526
3	3017	1.245e-02	1.612	2.513e-01	0.869
4	21521	4.043e-03	1.622	1.287e-01	0.966
5	162337	1.231e-03	1.716	6.476e-02	0.991

Table 2.3: Error table for problem (2.1) with $d = 2$ and $k = 1$

Example 2

In this example we consider the interface problem (2.10) in two space dimension, i.e. $d = 2$. We choose $D = (0,1)^2$, $T = 1$, $x_0 = (0.5, 0.5)^T$ and Ω to be a circle with center x_0 and radius 0.3, i.e. $\Omega = \{x \in D : \|x - x_0\|_2 < 0.3\}$. The computational domain Q_D is shown in Figure 2.4, where the domain Q_Ω is highlighted in blue and the domain $Q_D \setminus Q_\Omega$ is highlighted in green. We assume that the material properties of domain Q_Ω corresponds to air and that the material properties of domain $Q_D \setminus Q_\Omega$ corresponds to iron. The values for the material coefficients ν,λ are chosen similar to example 2 in Section 2.4.1. The functions f_1, f_2 are chosen according to (2.11). In order to obtain

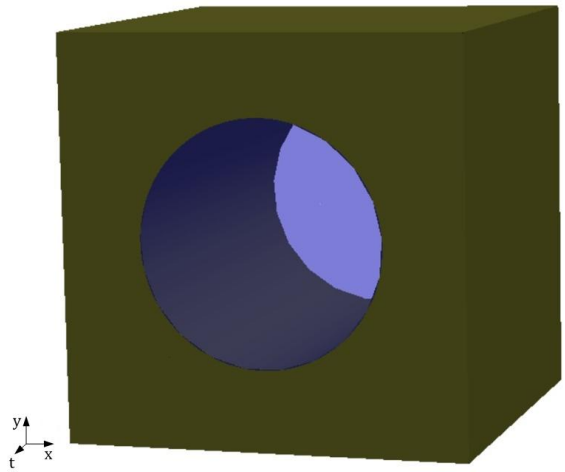


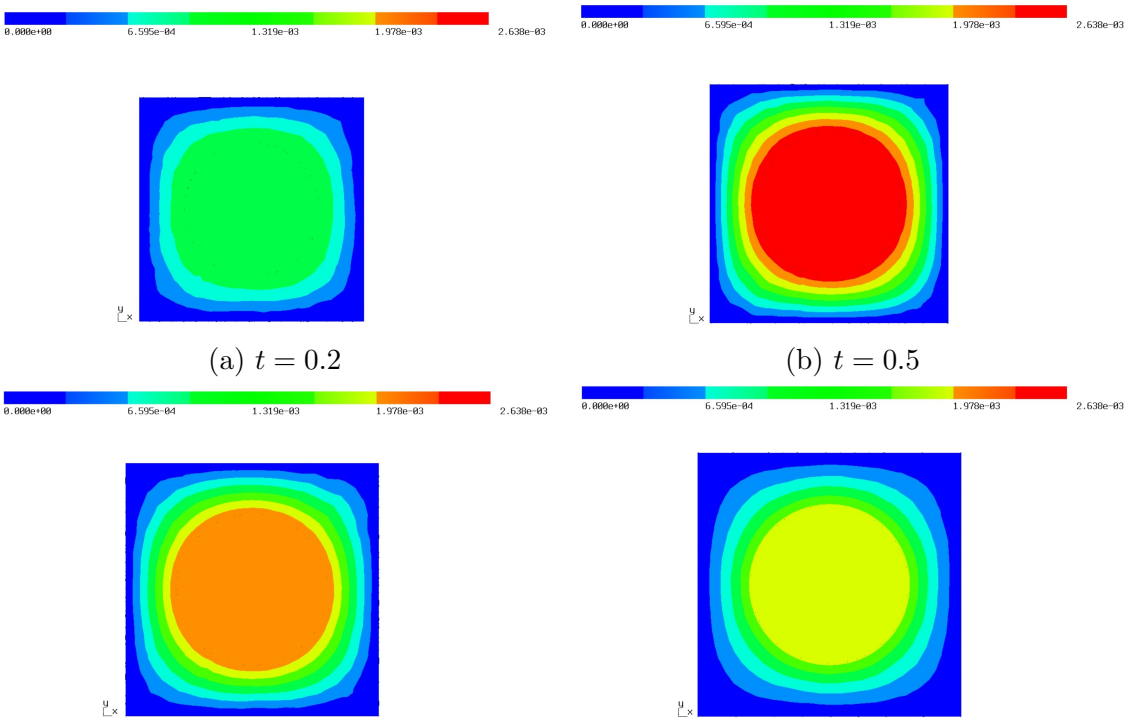
Figure 2.4: Computational domain Q_D, green: iron, blue: air

convergence rates we compute a reference solution on a mesh with $N = 34851$ dofs and consider this solution as "exact" solution. The error rates can be seen in Table 2.4. One observes a rate of 2 for the $L^2(Q_D)$-error and a rate of about 1.3 for the error in the energy norm, which is better than expected. In Figure 2.5 the numerical solution

L	N	$\|Z - Z_h\|_{L^2(Q_D)}$	eoc	$\|Z - Z_h\|_{L^2(0,T;H_0^1(D))}$	eoc
0	256	9.880e-03	0.000	1.978e-01	0.000
1	1706	2.671e-03	1.887	8.721e-02	1.182
2	12467	6.520e-04	2.035	3.450e-02	1.338

Table 2.4: Error rates for problem (2.10) with $d = 2$ and $k = 1$

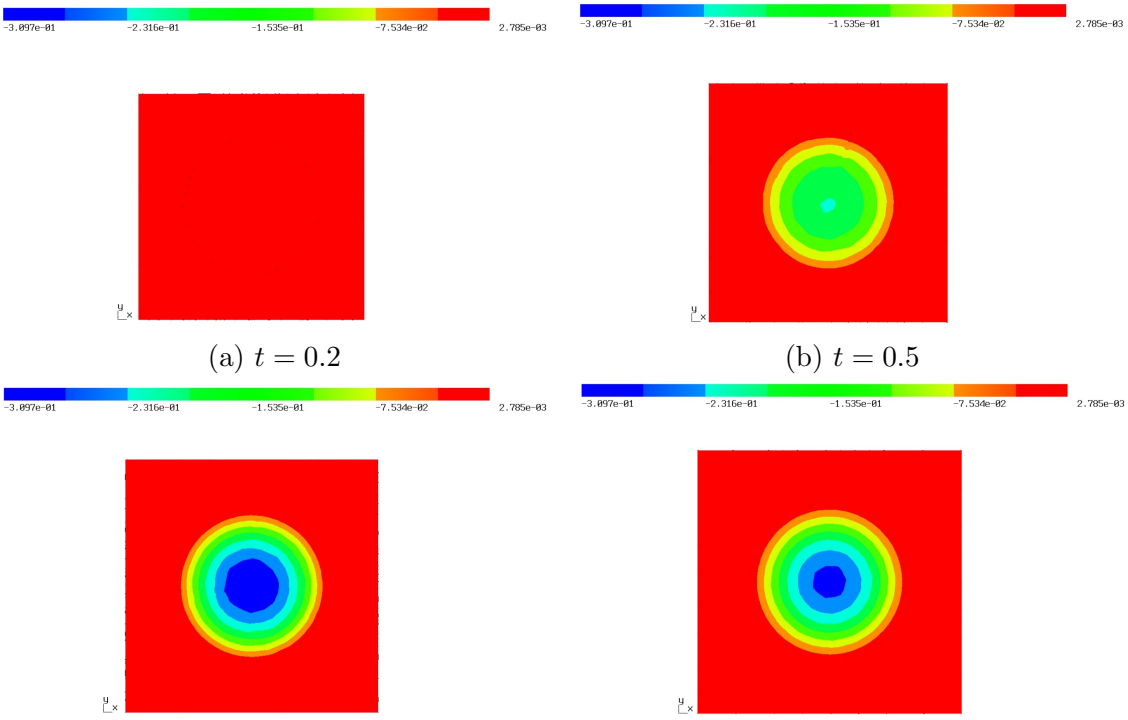
u_h computed on the refinement level $L = 2$ is depicted for different values of t. The numerical solution p_h on refinement level $L = 2$ and for different values of t can be seen in Figure 2.6.



(a) $t = 0.2$ (b) $t = 0.5$

(c) $t = 0.8$ (d) $t = 1$

Figure 2.5: Numerical solution u_h to (2.10) with $d = 2$, $k = 1$ on refinement level $L = 2$



(a) $t = 0.2$ (b) $t = 0.5$

(c) $t = 0.8$ (d) $t = 1$

Figure 2.6: Numerical solution p_h to (2.10) with $d = 2$, $k = 1$ on refinement level $L = 2$

3 Introduction to shape optimization

In this chapter we give a brief introduction to the theory of shape optimization. We want to mention that we focus on those parts of the theory which are essential to set up an algorithm and to solve the model problems from a numerical point of view. Therefore we put a strong emphasis in characterizing the shape derivative for PDE-constrained shape optimization problems and in explaining the numerical procedure in this chapter. While the existence of optimal shapes to shape optimization problems is certainly an interesting task we do not consider this issue here. The following is motivated by [13, 25].

In general a shape optimization problem reads

$$\min_{\Omega \in \mathcal{A}(D)} \mathcal{J}(\Omega) \tag{3.1}$$

in which $\mathcal{A}(D)$ is some admissible set of shapes and $\mathcal{J} : \mathcal{A}(D) \to \mathbb{R}$ a mapping called shape functional. Note that the model problems can be written in the sense of (3.1) too by introducing the reduced functionals $\mathcal{J}_1(\Omega) := J_1(\Omega, u_\Omega, p_\Omega)$, where (u_Ω, p_Ω) is solution to (1.18) and $\mathcal{J}_2(\Omega) := J_2(u_\Omega, p_\Omega)$, where (u_Ω, p_Ω) is solution to (1.21). The problems then read

$$\min_{\Omega \in \mathcal{A}(D)} \mathcal{J}_1(\Omega), \tag{3.2}$$

$$\min_{\Omega \in \mathcal{A}(D)} \mathcal{J}_2(\Omega). \tag{3.3}$$

For this reason it is sufficient to study problem (3.1). In order to get an idea of how a shape optimization algorithm could look like let us first recall the main ingredients of a gradient based optimization algorithm in a Banach space to solve the problem

$$\min_{x \in X} F(x), \tag{3.4}$$

where X is a Banach space and $F : X \to \mathbb{R}$ some functional. The goal is to find a minimizer x^* of the functional F, i.e. $F(x^*) \leq F(x)\ \forall x \in X$. The main steps of the algorithm read:

- Compute a descent direction, i.e. for given $x_k \in X$ find $y_k \in X$ such that $\mathrm{D}F(x_k)y_k < 0$. Here the mapping $\mathrm{D}F : X \to \mathcal{L}(X, \mathbb{R})$ denotes the Fréchet derivative.

- Choose a step size $t_k \in \mathbb{R}$ and calculate $x_{k+1} := x_k + t_k y_k$ such that $F(x_{k+1}) < F(x_k)$.

We want to use a similar approach in the context of shape optimization. Since we minimize over a set of admissible shapes which does not have vector space structure we cannot use the Fréchet derivative. Therefore, in a first step we have to think of a suitable notion of derivative of a function $\mathcal{J}(\Omega)$ with respect to Ω. In literature this derivative is known as shape derivative. In a second step we have to think of how to extract a descent direction out of the shape derivative and finally we have to discuss how to perform the optimization step.

This chapter is structured as follows: First we will introduce basic definitions and results in which we define the shape derivative. Following that we will discuss how to calculate the shape derivative via the averaged adjoint method. Furthermore we will have a look on how to get descent directions and on domain evolution. Last but not least we will state a gradient type shape optimization algorithm.

3.1 Basic definitions and results

In this section we state basic definitions and results which will be used later on. The definitions and results are taken from [7],[12],[25],[30] and [39]. To begin with we start with the definition of a shape function.

Definition 3.1 ([7])**.** Let $\emptyset \neq D \subseteq \mathbb{R}^d$ open, bounded and $\mathcal{P}(D) := \{\Omega : \Omega \subset D\}$ the set of subsets of D. The set D will be referred to as the underlying hold all domain. A shape function(al) $\mathcal{J}$ is a map

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

from some admissible set $\mathcal{A}(D) \subset \mathcal{P}(D)$ into $\mathbb{R}$.

In this thesis we consider the hold all domain D to be a bounded Lipschitz domain or a C^k domain ($k \geq 1$) meaning that D is a bounded domain and its boundary ∂D is of class C^k in the sense of [7, p.68, Def.3.1.]. Before we can define the notion of derivative for a shape function we have to introduce domain pertubations. There are different kinds of pertubations like the pertubation of identity or the velocity method. We focus on the latter approach. The velocity method is a method to define a mapping T_τ used to transform an initial shape Ω. It is important to notice that the shape Ω we are going to transform is an element out of the admissible set, i.e. $\Omega \in \mathcal{A}(D)$. One feature of the transformation T_τ should be that it leads to an element in $\mathcal{A}(D)$, i.e. $T_\tau(\Omega) \in \mathcal{A}(D)$. More precisely we are looking for a transformation which on the one hand preserves the topology and on the other hand preserves the regularity of the boundary of the set Ω. Mappings which guarantee that are so called diffeomorphic mappings. Before we state the definition of the transformation T_τ let us recall the definition of the tangent cone of a set $D \subset \mathbb{R}^d$.

Definition 3.2 ([39, Def. 2.15, Rem. 2.17])**.** The tangent cone of a set $D \subset \mathbb{R}^d$ at $x \in \overline{D}$ is defined as

$$\mathcal{T}_D(x) := \left\{ h \in \mathbb{R}^d : \ \liminf_{t \searrow 0} \frac{d(x+th, D)}{t} = 0 \right\},$$

where $d(x, D) := \inf_{y \in D} |x - y|$.

Note that $d(x, D) = d(x, \overline{D})$ and hence $\mathcal{T}_D(x) = \mathcal{T}_{\overline{D}}(x)$. We are now able to state the definition of the mapping T_τ by the velocity method.

Definition 3.3. Let $V : [0, \tilde{\tau}] \times \overline{D} \to \mathbb{R}^d$ be a non-autonomous vector field satisfying the conditions

$$\forall x \in \overline{D} : \ V(., x) \in C([0, \tilde{\tau}], \mathbb{R}^d), \tag{3.5a}$$

$$\exists L > 0, \text{ s.t. } \forall x, y \in \overline{D} : \ \|V(., x) - V(., y)\|_{C([0,\tilde{\tau}], \mathbb{R}^d)} \leq L |x - y|, \tag{3.5b}$$

$$\forall x \in \overline{D} \ \forall \tau \in [0, \tilde{\tau}] : \ \pm V(\tau, x) \in \mathcal{T}_{\overline{D}}(x). \tag{3.5c}$$

The transformation $T : [0, \tilde{\tau}] \times \overline{D} \to \mathbb{R}^d$ (also called flow) associated with the vector field V is defined by

$$T(\tau, X) := T_\tau(X) := x(\tau)$$

where $x(\tau)$ is solution of the differential equation

$$x'(\tau) = V(\tau, x(\tau)), \ \tau > 0, \ x(0) = X$$

We denote the transformation of a shape $\Omega \subset D \subset \mathbb{R}^d$ by the mapping T_τ as

$$\Omega_\tau := T_\tau(\Omega) = \{T_\tau(X) : \ X \in \Omega\}.$$

Of course the mapping T_τ depends on the vector field V meaning that $T_\tau = T_\tau^V$, but we will not indicate this dependence for ease of notation. The conditions (3.5) are sufficient to conclude [7, pp.194, Thm. 5.1, Remark 5.1.] that for all $\tau \in [0, \tilde{\tau}]$ the mapping $T_\tau : \overline{D} \to \overline{D}$ is a homeomorphism, i.e. T_τ and its inverse are continuous. In particular this means that interior points are mapped to interior points and boundary points are mapped to boundary points. Throughout this thesis we assume the vector field V to be autonomous, i.e. $V = V(x)$. The conditions (3.5) then reduces to [39, p. 14]

$$V \in C^{0,1}(\overline{D}, \mathbb{R}^d) \text{ and } \forall x \in \overline{D} : \ \pm V(x) \in \mathcal{T}_{\overline{D}}(x). \tag{3.6}$$

In addition to that we will assume that $V \in C_c^k(D, \mathbb{R}^d)$ for some $k \geq 1$. Hence V clearly fulfills (3.6) by extending it with zero to $\overline{D}$ [39, p. 15]. As a consequence the flow $T_\tau : \overline{D} \to \overline{D}$ is a homeomorphism and further it holds that $T_\tau(D) = D$, $T_\tau(\partial D) = \partial D$

and for $\Omega \subset D:\ T_\tau(\Omega) \subset D$. Moreover the mappings T_τ and T_τ^{-1} belong to $C^k(\overline{D}, \mathbb{R}^d)$ [34, p.51, Thm.2.16.] which guarantee that the regularity of the boundary [39, p.13, Remark 2.8.] as well as the topology is preserved.

Summing up we can conclude that for $\Omega \in \mathcal{A}(D)$ with

$$\mathcal{A}(D) := \{\Omega \subset D:\ \Omega \text{ open, Lipschitz with uniform Lipschitz constant } L_{\mathcal{A}}\}$$

given like in the model problems, the transformed set $\Omega_\tau = T_\tau(\Omega)$ is an element of $\mathcal{A}(D)$ for sufficient small τ.

We are now able to define the notion of derivative for a shape function.

Definition 3.4 ([38], Eulerian semi-derivative)**.** Let $V \in C_c^k(D, \mathbb{R}^d)$ for $k \geq 1$ and for $\Omega \in \mathcal{A}(D)$ let $\Omega_\tau := T_\tau(\Omega)$ according to Definition 3.3. Moreover let τ sufficiently small such that $\Omega_\tau \in \mathcal{A}(D)$. The Eulerian semi-derivative of a shape function $\mathcal{J} : \mathcal{A}(D) \to \mathbb{R}$ at Ω in direction V is defined as

$$\mathrm{d}\mathcal{J}(\Omega; V) := \lim_{\tau \searrow 0} \frac{\mathcal{J}(\Omega_\tau) - \mathcal{J}(\Omega)}{\tau} \tag{3.7}$$

if the limit exists and is finite.

The Eulerian semi-derivative basically gives us information about the sensitivity of the shape function $\mathcal{J}$ with respect to a perturbation of the domain Ω by a deformation map T_τ [12].

Definition 3.5 ([9])**.** A shape function $\mathcal{J}$ is said to be shape differentiable at Ω if for some $k \geq 1$ the Eulerian semi-derivative $\mathrm{d}\mathcal{J}(\Omega; V)$ exists for all $V \in C_c^k(D, \mathbb{R}^d)$ and the mapping $V \mapsto \mathrm{d}\mathcal{J}(\Omega; V)$ is linear and continuous from $C_c^k(D, \mathbb{R}^d)$ to $\mathbb{R}$. Moreover the smallest integer $k \geq 0$ for which $V \mapsto \mathrm{d}\mathcal{J}(\Omega; V)$ is continuous with respect to the $C^k(D, \mathbb{R})$-topology is called order of $\mathrm{d}\mathcal{J}(\Omega; .)$

In the following we state the structure theorem of Hadamard-Zolésio. Basically, this theorem says that under certain smoothness assumptions on the boundary $\partial\Omega$ and V the shape derivative only depends on the normal component $V \cdot n$ [13, 9].

Theorem 3.1 (Structure theorem of Hadamard-Zolésio)**.** *Let $\mathcal{J}$ be a shape functional. Assume that $\mathcal{J}$ is shape differentiable at $\Omega \in \mathcal{A}(D)$ and that $\partial\Omega$ is of class C^{k+1}, $k \geq 0$. Then there exists a unique outward unit normal vector field $n \in C^k(\partial\Omega, \mathbb{R}^d)$ and a scalar distribution $g \in C^k(\partial\Omega)'$ such that for all $V \in C_c^k(D, \mathbb{R}^d)$*

$$\mathrm{d}\mathcal{J}(\Omega; V) = \langle g, \gamma_{\partial\Omega}(V) \cdot n \rangle_{C^k(\partial\Omega)' \times C^k(\partial\Omega)}. \tag{3.8}$$

When $g \in L^1(\partial\Omega)$, we may write

$$\mathrm{d}\mathcal{J}(\Omega; V) = \int_{\partial\Omega} gV \cdot n \, \mathrm{d}s. \tag{3.9}$$

Proof. A proof can be found in [7, pp.479-481, Thm. 3.6 and Cor. 1] □

Formula (3.9) is also known as Hadamard formula. It can be seen from (3.9) that the shape derivative can be represented as an integral over the boundary if the domain Ω is regular enough. Another way to represent the shape derivative is in volume form, i.e.

$$\mathrm{d}\mathcal{J}(\Omega; V) = \int_D g(V, \mathrm{D}V)\,\mathrm{d}x \tag{3.10}$$

for some function g. The formula (3.10) is also known as distributed shape derivative.

Let us briefly comment on the two representation formulas. One advantage of having the shape derivative in boundary form is that a descent direction is given by choosing $V = -gn$, see Section 3.3. However, this descent direction is only defined on the boundary $\partial\Omega$ and sometimes has to be regularized as it might not be regular enough. [12, Chapter 6.1.2]

The distributed shape derivative is more general, meaning that for shapes with lower regularity the distributed shape derivative may be well defined while the shape derivative in Hadamard form is not. One drawback of this representation form is that in order to obtain a descent direction one has to solve an auxiliary boundary value problem, see Section 3.3. [12, Chapter 6.1.2]

Further we want to mention that in many numerical procedures it is necessary to have the shape derivative defined on the whole domain and not only on the boundary. This is for instance in numerical procedures the case in which the domain D is discretized by means of a finite element mesh and one moves every node a certain distance in direction V.[12, Chapter 6.1.2]

In this master thesis we focus on the shape derivative in volume form since in our numerical procedure it is beneficial to have a descent direction defined on the whole computational domain D.

3.2 Shape derivative via the averaged adjoint method

In this section, based on [21, 38, 39], we discuss the calculation of the shape derivative if the shape function is constrained by a PDE. There are several methods to calculate the shape derivative for a PDE constrained shape function like the material derivative method which is also known as chain rule approach, the minimax formulation introduced in [7] or the (formal) method of Céa. We concentrate on the approach based in [38]. This is the calculation of the shape derivative via the averaged adjoint method. This method is a Lagrangian-type method which allows to compute the shape derivative of a shape functional depending on the solution of a PDE without the need to compute the material derivative of the PDE [21]. The following is motivated by [21].

The idea behind Lagrangian methods is that in order to compute the derivative of a shape functional it is sufficient to compute derivative of the Lagrangian with respect to τ. For better understanding let $E(\Omega), F(\Omega)$ be vector spaces, $\tilde{\tau} > 0$ and $\Omega_\tau := T_\tau(\Omega)$ a parametrization for $\tau \in [0, \tilde{\tau}]$. We further assume that the spaces E, F fulfill the following property

$$\hat{\varphi} \in E(\Omega_\tau) \Leftrightarrow \hat{\varphi} \circ T_\tau \in E(\Omega),$$
$$\hat{\psi} \in F(\Omega_\tau) \Leftrightarrow \hat{\psi} \circ T_\tau \in F(\Omega).$$

In order to compute the derivative of $\mathcal{J}(\Omega_\tau)$ we first write the functional in terms of the Lagrangian, i.e.

$$\mathcal{J}(\Omega_\tau) = \mathcal{L}(\Omega_\tau, u_\tau, \hat{\psi})$$

where $u_\tau \in E(\Omega_\tau)$ is the solution to the perturbed state equation and $\hat{\psi} \in F(\Omega_\tau)$ is arbitrary. The Lagrangian $\mathcal{L}(\Omega_\tau, \hat{\varphi}, \hat{\psi})$, where $\hat{\varphi} \in E(\Omega_\tau)$, $\hat{\psi} \in F(\Omega_\tau)$ usually consists of integrals on Ω_τ. Using a change of variables yields to integrals on a fixed domain Ω and the τ-dependence is shifted to the integrands. Hence this leads to integrands of the form $\hat{\varphi} \circ T_\tau \in E(\Omega)$ and $\hat{\psi} \circ T_\tau \in F(\Omega)$ which are not straightforward to differentiate since $\hat{\varphi}, \hat{\psi}$ are functions on the moving domain Ω_τ. To find a remedy for this problem one defines

$$G(\tau, \varphi, \psi) := \mathcal{L}(\Omega_\tau, \varphi \circ T_\tau^{-1}, \psi \circ T_\tau^{-1}),$$

where $\varphi \in E(\Omega)$, $\psi \in F(\Omega)$, which is a re-parametrization of the Lagrangian. The re-parametrization G leads after a change of variables to functions φ, ψ on the fixed domain Ω. Summing up we end up with

$$\mathrm{d}\mathcal{J}(\Omega; V) = \frac{\mathrm{d}}{\mathrm{d}\tau}\, \mathcal{J}(\Omega_\tau)|_{\tau=0} = \frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{L}(\Omega_\tau, u_\tau, \hat{\psi}) = \frac{\mathrm{d}}{\mathrm{d}\tau}\, G\left(\tau, u^\tau, \psi\right)|_{\tau=0}$$

where $u^\tau := u_\tau \circ T_\tau$ is the pull-back of the solution to the perturbed state equation and $\psi := \hat{\psi} \circ T_\tau$. Hence we will investigate the differentiability of G with respect to τ in the following. We start with introducing some notation and assumptions and then state the main theorem. We follow the lines in [38, Section 3.1].

Let E, F Banach spaces, $\tilde{\tau} > 0$. Consider

$$G : [0, \tilde{\tau}] \times E \times F \to \mathbb{R}$$
$$(\tau, \varphi, \psi) \mapsto G(\tau, \varphi, \psi),$$

such that $\psi \mapsto G(\tau, \varphi, \psi)$ is affine for all $(t, \varphi) \in [0, \tilde{\tau}] \times E$. We introduce the set of solutions to the state equation on the perturbed domain $\Omega_\tau := T_\tau(\Omega)$ brought back to the fixed domain Ω by

$$E(\tau) := \left\{ u \in E : \ \mathrm{d}_\psi G(\tau, u, 0; \hat{\psi}) = 0 \ \forall \hat{\psi} \in F \right\}. \tag{3.11}$$

Furthermore we need the following assumptions.

Assumption 3.1. (H0)

(i) $E(\tau) = \{u^\tau\}$ is single-valued for all $\tau \in [0, \tilde{\tau}]$.

(ii) For all $\tau \in [0, \tilde{\tau}]$ and $\tilde{p} \in F$, the mapping

$$[0,1] \to \mathbb{R},\ s \mapsto G(\tau, su^\tau + (1-s)u^0, \tilde{p})$$

is absolutely continuous.

(iii) For all $\tau \in [0, \tilde{\tau}]$, $\hat{\varphi} \in E$, $\tilde{p} \in F$

$$s \mapsto \mathrm{d}_\varphi G(\tau, su^\tau + (1-s)u^0, \tilde{p}; \hat{\varphi})$$

is well-defined and belongs to $L^1(0,1)$.

Remark 3.1. Let us briefly discuss the consequences of assumption (H0).

- Point (i) of (H0) corresponds to the unique solvability of the perturbed state equation.
- Point (ii) of (H0) implies that for almost all $s \in [0,1]$ the derivative

$$d_\varphi G(\tau, su^\tau + (1-s)u^0, \tilde{p}; u^\tau - u^0)$$

exists and together with (iii) this yields

$$G(\tau, u^\tau, \tilde{p}) - G(\tau, u^0, \tilde{p}) = \int_0^1 d_\varphi G(\tau, su^\tau + (1-s)u^0, \tilde{p}; u^\tau - u^0)\,\mathrm{d}s.$$

For $\tau \in [0, \tilde{\tau}]$, $u^\tau \in E(\tau)$, $u^0 \in E(0)$, let

$$Y(\tau, u^\tau, u^0) := \left\{ q \in F \,|\, \forall \hat{\varphi} \in E : \int_0^1 \mathrm{d}_\varphi G(\tau, su^\tau + (1-s)u^0, q; \hat{\varphi})\,\mathrm{d}s = 0 \right\} \tag{3.12}$$

the set of solutions to the averaged adjoint equation. For $\tau = 0$ the set $Y(0, u^0) := Y(0, u^0, u^0)$ coincides with the solution set of the usual adjoint state equation

$$Y(0, u^0) = \left\{ q \in F \,|\, \forall \hat{\varphi} \in E : \mathrm{d}_\varphi G(\tau, u^0, q; \hat{\varphi})\,\mathrm{d}s = 0 \right\}. \tag{3.13}$$

Theorem 3.2. *Let E, F linear vector spaces and $\tilde{\tau} > 0$. Suppose that the function $G : [0, \tilde{\tau}] \times E \times F \to \mathbb{R}$, $(\tau, \varphi, \psi) \mapsto G(\tau, \varphi, \psi)$ is affine in the last argument. Let assumption (H0) and the following conditions be satisfied.*

(H1) For all $\tau \in [0, \tilde{\tau}]$ and all $(u, p) \in E(0) \times F$, the derivative $\partial_\tau G(\tau, u, p)$ exists.

(H2) For all $\tau \in [0, \tilde{\tau}]$ the set $Y(\tau, u^\tau, u^0)$ is non-empty and $Y(0, u^0)$ is single-valued.

(H3) Let $p^0 \in Y(0, u^0)$. For any sequence $(\tau_n)_n \searrow 0$ there exists a subsequence $(\tau_{n_k})_k$ and $(p^{\tau_{n_k}})_k$ with $p^{\tau_{n_k}} \in Y(\tau_{n_k}, u^{\tau_{n_k}}, u^0)$ such that

$$\lim_{\substack{k \to \infty \\ s \searrow 0}} \partial_\tau G(s, u^0, p^{\tau_{n_k}}) = \partial_\tau G(0, u^0, p^0).$$

Then for any $\psi \in F$:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(G(\tau, u^\tau, \psi)\right)|_{\tau=0} = \partial_\tau G(0, u^0, p^0). \tag{3.14}$$

Proof. A proof can be found in [38, Thm.3.1.]. □

Remark 3.2 ([38]). In concrete applications the conditions (H0)-(H3) have the following meaning:

(i) By assumption (H0) we can apply the fundamental theorem of calculus to G with respect to u. Note that this condition is milder than Fréchet differentiability.

(ii) Condition (H1) allows an application of the mean value theorem with respect to t.

(iii) Condition (H2) ensures that the averaged adjoint equation is solvable and that the adjoint equation has a unique solution.

(iv) Condition (H3) can be verified by showing that p^τ converges weakly to p^0 and that $(\tau, \psi) \mapsto G(\tau, u^0, \psi)$ is weakly continuous. Note that there is no assumption on the convergence of u^τ to u^0, but in applications we need the convergence $u^\tau \to u^0$ to prove $p^\tau \to p$ with respect to the appropriate topologies.

Using Theorem 3.2 the shape derivative can be computed by

$$\mathrm{d}\mathcal{J}(\Omega; V) = \frac{\mathrm{d}}{\mathrm{d}\tau}\,\mathcal{J}(\Omega_\tau)|_{\tau=0} = \frac{\mathrm{d}}{\mathrm{d}\tau}\,G(\tau, u^\tau, \psi)|_{\tau=0} = \partial_\tau G(0, u^0, p^0).$$

Finally we want to mention that the averaged adjoint method is in particular well-suited for problems involving nonlinear PDE constraints [13].

3.3 Descent directions, domain evolution and generic algorithm

In this section we are going to state a generic gradient based shape optimization algorithm. Therefore we first introduce descent directions and discuss how to extract them out of the shape derivative. Furthermore we explain how to update the domain and finally we state the generic algorithm. This section is based on [13].

We start with the definition of a descent direction.

Definition 3.6 (descent direction,[21])**.** The vector field $V \in C_c^{0,1}(D, \mathbb{R}^d)$ is called a descent direction for $\mathcal{J}$ at Ω if there exists an $\epsilon > 0$ such that

$$\mathcal{J}(\Omega_\tau) < \mathcal{J}(\Omega) \quad \forall \tau \in (0, \epsilon)\,. \tag{3.15}$$

The following lemma gives us a characterization of a descent direction by the shape derivative.

Lemma 3.3. *Let $\mathcal{J} : \mathcal{A}(D) \to \mathbb{R}$ be a shape functional and assume that $\mathcal{J}$ is shape differentiable at $\Omega \in \mathcal{A}(D)$ with shape derivative $\mathrm{d}\mathcal{J}(\Omega, .) : C_c^k(D, \mathbb{R}^d) \to \mathbb{R},\ V \mapsto \mathrm{d}\mathcal{J}(\Omega, V)$ for some $k \geq 1$. Furthermore let $V \in C_c^k(D, \mathbb{R}^d)$ such that*

$$\mathrm{d}\mathcal{J}(\Omega; V) < 0. \tag{3.16}$$

Then V is a descent direction.

Proof. Since $\mathcal{J}$ is shape differentiable at Ω we can use the definition of the Eulerian-semi derivative to get

$$0 > \mathrm{d}\mathcal{J}(\Omega; V) = \lim_{\tau \searrow 0} \frac{\mathcal{J}(\Omega_\tau) - \mathcal{J}(\Omega)}{\tau}.$$

By definition of the limit there is a $\epsilon > 0$ such that for all $\tau \in (0, \epsilon)$

$$\frac{\mathcal{J}(\Omega_\tau) - \mathcal{J}(\Omega)}{\tau} < 0.$$

Therefore it follows that V is a descent direction. □

If $\mathrm{d}\mathcal{J}(\Omega; V)$ is given in Hadamard form (3.9) then choosing $V = -gn$ yields a descent direction since

$$\mathrm{d}\mathcal{J}(\Omega; -gn) = \int_{\partial\Omega} g(-gn) \cdot n \,\mathrm{d}s = -\int_{\partial\Omega} g^2 \,\mathrm{d}s < 0.$$

If $\mathrm{d}\mathcal{J}(\Omega; V)$ is given in volume form one common approach to compute a descent direction is by solving an auxiliary boundary value problem. To make things clear let X be a Hilbert space and assume that $\mathrm{d}\mathcal{J}(\Omega, .)$ is linear and bounded on X. Further let $b : X \times X \to \mathbb{R}$ be a positive definite and bounded bilinear form. Consider the auxiliary boundary value problem

$$\text{Find } V \in X : \quad b(V, W) = -\mathrm{d}\mathcal{J}(\Omega, W) \quad \forall W \in X. \tag{3.17}$$

If we solve (3.17) the solution V is a descent direction since

$$\mathrm{d}\mathcal{J}(\Omega, V) = -b(V, V) < 0.$$

Note that we can obtain different descent directions by choosing different bilinear forms. Possible choices are

- $b(V, W) = \int_D \alpha(x)\partial V : \partial W + \beta(x) V \cdot W \,\mathrm{d}x$, with $\alpha, \beta \in L^\infty(D)$ positive,
- $b(V, W) = \int_D E\epsilon(u) : \epsilon(v) \,\mathrm{d}x$, the bilinear form of linearized elasticity.

The next thing we are going to discuss is how to update the domain Ω_n in the n-th iteration if one has computed a descent direction V_n. There are two approaches based on the explicit or implicit representation of the domain. We focus on the first approach. If one has an explicit representation of the domain one can simply move every point in Ω_n a certain distance τ_n in direction V_n, i.e.

$$\Omega_{n+1} = \{x + \tau_n V_n(x) : \ x \in \Omega_n\} = (id + \tau_n V_n)(\Omega_n).$$

The distance τ_n is called step size. There are different opportunities for the choice of τ_n. Popular choices are

- τ_n constant and small,
- "back-tracking": choose $\tau_n = \max\left\{1, \frac{1}{2}, \frac{1}{4}, ...\right\}$ such that $\mathcal{J}(\Omega_{n+1}) < \mathcal{J}(\Omega_n)$,
- Armijo rule.

We are now able to state a generic gradient based shape optimization algorithm.

Algorithm 3.4. *Choose an initial design Ω_0.*
For $n = 0, 1, 2, ...$ until converge,

(1) Find a vector field V_n such that $\mathrm{d}\mathcal{J}(\Omega_n; V_n) < 0$

(2) Choose a step size τ_n according to the step size rule

(3) Set $\Omega_{n+1} = (id + \tau_n V_n)(\Omega_n)$

4 Shape optimization of model problem 1 and 2

In this chapter we focus on the parabolic PDE-constrained shape optimization model problems. So far we have discussed how to solve the forward problem using space-time finite elements and explained shape optimization techniques based on the shape derivative. We will now apply the results from the previous chapters in order to set up an algorithm and to solve the model problems described in (1.17)-(1.18) and (1.20)-(1.21) numerically. We start with some preliminaries in which we collect some basic properties of Sobolev functions composed with flows and state a modification of Theorem 3.2 to the multi-valued case. Following that we will compute the shape derivative of $\mathcal{J}_1$ and $\mathcal{J}_2$ via the averaged adjoint method. Moreover, we will outline a gradient-type shape optimization algorithm used to solve the problems. Finally we will present some numerical tests which demonstrate that the introduced algorithm is feasible.

4.1 Preliminaries

In this section we present some basic results on Sobolev functions composed with flows and state an extension of Theorem 3.2. This section is mainly based on [39, Chapter 2.3.3, Chapter 4.1.4]

In the following let $V \in C_c^1(D, \mathbb{R}^d)$ be a given vector field and T_τ its associated flow [39]. We will make use of the following abbreviations

$$\xi(\tau) := \det(\partial T_\tau), \quad A(\tau) := \det(\partial T_\tau)\partial T_\tau^{-1}\partial T_\tau^{-T}, \quad B(\tau) := \partial T_\tau^{-T}. \tag{4.1}$$

Here, $\det : \mathbb{R}^{d\times d} \to \mathbb{R}$ denotes the determinant and ∂T_τ and ∂T_τ^{-1} denote the Jacobians of T_τ and T_τ^{-1}, respectively. We will now state some properties of the functions $\xi(\tau), A(\tau), B(\tau)$.

Lemma 4.1. *Let $V \in C_c^1(D, \mathbb{R}^d)$ and T_τ the flow associated with V via the velocity method. The mappings $\tau \mapsto \xi(\tau)$, $\tau \mapsto B(\tau)$, $\tau \mapsto A(\tau)$ given by (4.1) are differentiable on $[0, \tilde{\tau}]$, i.e. they are differentiable on $(0, \tilde{\tau})$, the right-sided and left-sided*

derivatives exist in 0 *and* $\tilde{\tau}$*,respectively and these functions satisfy*

$$\frac{\mathrm{d}}{\mathrm{d}\tau}B(\tau) = -B(\tau)(\partial V^\tau)^T B(\tau), \tag{4.2}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\xi(\tau) = \mathrm{tr}(\partial V^\tau B^T(\tau))\xi(\tau), \tag{4.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}A(\tau) = \mathrm{tr}(\partial V^\tau B^T(\tau))A(\tau) - B^T(\tau)\partial V^\tau A(\tau) - (B^T(\tau)\partial V^\tau A(\tau))^T, \tag{4.4}$$

where $V^\tau(x) := V(T_\tau(x))$*.*

Proof. A proof can be found in [39, Lemma 2.14] □

An immediate consequence of Lemma 4.1 together with the fact that $\xi(0) = 1$, $A(0) = B(0) = I$ is

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\xi(\tau)\right)\big|_{\tau=0} = \mathrm{div}\, V, \tag{4.5}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(B(\tau)\right)\big|_{\tau=0} = -\partial V^T, \tag{4.6}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(A(\tau)\right)\big|_{\tau=0} = \mathrm{div}\, V - \partial V - \partial V^T. \tag{4.7}$$

The next lemma states that the mappings $\tau \mapsto \xi(\tau)$, $\tau \mapsto A(\tau)$ are bounded.

Lemma 4.2 ([39])**.** *Let the mappings* $A \in C([0,\tilde{\tau}]\,; C(\overline{D}, \mathbb{R}^{d\times d}))$ *and* $\xi \in C([0,\tilde{\tau}]\,; C(\overline{D}))$ *be given and assume that* $A(0) = I$ *and* $\xi(0) = 1$*. Then there are constants* $\gamma_1, \gamma_2, \delta_1, \delta_2 > 0$ *and* $\hat{\tau} > 0$ *such that for all* $\eta \in \mathbb{R}^d$ *and for all* $\tau \in [0,\hat{\tau}]$*:*

$$\gamma_1 |\eta|^2 \le \eta \cdot A(\tau)\eta \le \gamma_2 |\eta|^2,$$
$$\delta_1 \le \xi(\tau) \le \delta_2.$$

One common approach in shape calculus is to use function space parametrization in order to avoid differentiating functions which live on the moving Ω_τ. Hence for a Banach space $E(\Omega)$ we need the property

$$\hat{\varphi} \in E(\Omega_\tau) \iff \hat{\varphi} \circ T_\tau \in E(\Omega).$$

This condition is fulfilled if $E(\Omega) = W^{1,p}(\Omega)$, which can be seen in the next theorem.

Theorem 4.3. *Let* $p \ge 1$*. Suppose that* $T : \mathbb{R}^d \to \mathbb{R}^d$ *is a bi-Lipschitz mapping. Suppose that* U *is an open subset of* $\mathbb{R}^d$ *and set* $W := T^{-1}(U)$*. Then we have*

$$u \in W^{1,p}(U) \iff u \circ T \in W^{1,p}(W).$$

Proof. A proof can be found in [47, Thm.2.2.2] □

The theorem basically says that for a transformation $\Omega_\tau = T_\tau(\Omega)$ we can write

$$H_0^1(\Omega_\tau) = \left\{\varphi \circ T_\tau^{-1} : \ \varphi \in H_0^1(\Omega)\right\}.$$

We can state a similar result as in Theorem 4.3 for the space $L^2(0,T;H^1(\Omega))$ too.

Theorem 4.4. *Suppose that $T : \mathbb{R}^d \to \mathbb{R}^d$ is a bi-Lipschitz mapping. Suppose that U is an open subset of $\mathbb{R}^d$ and set $W := T^{-1}(U)$. For $u \in L^2(0,T;H^1(U))$ we have $u \circ T \in L^2(0,T;H^1(W))$ and*

$$D_j\left(u \circ T\right) = \sum_{i=1}^{d} \left((D_i u) \circ T\right) \cdot (D_j T_i), \ j \in \{1, ...d\}, \tag{4.8}$$

where D_j denotes the weak derivative with respect to j-th spatial coordinate.

Proof. We have to show that

(i) $u \circ T \in L^2(0,T;L^2(W))$,

(ii) $u \circ T$ admits a weak derivative in $L^2(0,T;L^2(W))$ which is given by (4.8).

Before we start with the proof of (i) and (ii) note that due to Rademacher's theorem the mappings T and T^{-1} are differentiable almost everywhere and there holds for $i,j \in \{1,...,d\}$:

$$|D_j T_i(x)| \leq L, \text{ for } a.e.\ x \in R^d \qquad \left|D_j T_i^{-1}(y)\right| \leq \hat{L}, \text{ for } a.e.\ y \in \mathbb{R}^d,$$

where L,$\hat{L}$ are the Lipschitz constants of T and T^{-1}, respectively. Hence, we conclude that for the determinant of the Jacobian ∂T^{-1} there exists a $C > 0$ such that

$$\left|\det\left(\partial T^{-1}(y)\right)\right| \leq C \text{ for } a.e.\ y \in \mathbb{R}^d. \tag{4.9}$$

This results from the fact that the determinant of ∂T^{-1} corresponds to a finite sum of the bounded entries of the Jacobian.

Now, we proceed with the proof of (i). Using a change of variables $x = T^{-1}(y)$, estimate (4.9) and $u \in L^2(0,T;L^2(U))$ we get

$$\begin{aligned}
\|u \circ T\|^2_{L^2(0,T;L^2(W))} &= \int_0^T \int_W |u(t,T(x))|^2 \,\mathrm{d}x\mathrm{d}t \\
&= \int_0^T \int_U |u(t,y)|^2 \left|\det\left(\partial T^{-1}(y)\right)\right| \mathrm{d}y\mathrm{d}t \\
&\leq C \int_0^T \int_U |u(t,y)|^2 \,\mathrm{d}y\mathrm{d}t = C\|u\|_{L^2(0,T;L^2(U))} < \infty
\end{aligned} \tag{4.10}$$

We continue with the proof of (ii). We define for $j \in \{1, ..., d\}$:

$$f_j(t,x) := \sum_{i=1}^{d} \left((D_i u)\left(t, T(x)\right)\right) \cdot (D_j T_i(x)), \text{ for } a.e.\ (t,x) \in (0,T) \times W.$$

We show that

(a) $f_j \in L^2(0,T; L^2(W))$, $j = 1, ..., d$,

(b) $(u \circ T, D_j \varphi)_{L^2(0,T;L^2(W))} = -\left(f_j, \varphi\right)_{L^2(0,T;L^2(W))} \ \forall \varphi \in L^2(0,T; C_0^\infty(W))$.

We start with (a). Let $j \in \{1, ..., d\}$. Then we have

$$\begin{aligned} |f_j(t,x)| &\leq L \sum_{i=1}^{d} |(D_i u)(t, T(x))| \leq L \left(\sum_{i=1}^{d} 1^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{d} |(D_i u)(t, T(x))|^2 \right)^{\frac{1}{2}} \\ &= L d^{\frac{1}{2}} \left(\sum_{i=1}^{d} |(D_i u)(t, T(x))|^2 \right)^{\frac{1}{2}}, \text{ for } a.e.\ (t,x) \in (0,T) \times W. \end{aligned}$$

Using this estimate, a change of variables $x = T^{-1}(y)$, estimate (4.9) and that by assumption $D_i u \in L^2(0,T; L^2(U))$ for $i \in \{1, ..., d\}$ we conclude

$$\begin{aligned} \|f_j\|^2_{L^2(0,T;L^2(W))} &= \int_0^T \int_W |f_j(t,x)|^2 \, \mathrm{d}x\mathrm{d}t \leq L^2 d \int_0^T \int_W \sum_{i=1}^{d} |(D_i u)(t, T(x))|^2 \, \mathrm{d}x\mathrm{d}t \\ &= L^2 d \sum_{i=1}^{d} \int_0^T \int_U |D_i u(t,y)|^2 \left| \det\left(\partial T^{-1}(y)\right) \right| \mathrm{d}y\mathrm{d}t \\ &\leq L^2 d C \sum_{i=1}^{d} \|D_i u\|^2_{L^2(0,T;L^2(U))} < \infty. \end{aligned}$$

We proceed with (b). The idea is to approximate $u \circ T$ with smooth functions such that one can use integration by parts with respect to x. For $u \in L^2(0,T; H^1(U))$ let $(u_n)_n$ be a sequence such that $u_n(t) \in C_0^\infty(\mathbb{R}^d)$ for a.e. $t \in (0,T)$ and $u_n \to u$ in $L^2(0,T; H^1(U'))$ for all $U' \Subset U$, see Remark 4.1. Hence, we conclude from (4.10) that for all $W' \Subset W$

$$\|u_n \circ T - u \circ T\|^2_{L^2(0,T;L^2(W'))} \leq C \|u_n - u\|^2_{L^2(0,T;L^2(U'))} \to 0, \ n \to \infty, \tag{4.11}$$

and that

$$((D_i u_n) \circ T) \cdot D_j T_i \to ((D_i u) \circ T) \cdot D_j T_i \text{ in } L^2(0,T; L^2(W')) \tag{4.12}$$

which can be seen by

$$\begin{aligned}
&\|((D_i u_n) \circ T)\, D_j T_i - ((D_i u) \circ T)\, D_j T_i\|^2_{L^2(0,T;L^2(W'))} \\
&= \int_0^T \int_{W'} |(D_i(u_n - u)) \circ T \cdot D_j T_i|^2 \,\mathrm{d}x\mathrm{d}t \\
&\leq L^2 \int_0^T \int_{W'} |(D_i(u_n - u))(t, T(x))|^2 \,\mathrm{d}x\mathrm{d}t \\
&= L^2 \int_0^T \int_{U'} |D_i(u_n - u)(t,y)|^2 \left|\det\left(\partial T^{-1}(y)\right)\right| \mathrm{d}y\mathrm{d}t \\
&\leq CL^2 \|D_i(u_n - u)\|^2_{L^2(0,T;L^2(U'))} \to 0,\ n \to \infty
\end{aligned}$$

Now, let $\varphi \in L^2(0,T;C_0^\infty(W))$. Then there exists a $W' \Subset W$: $\operatorname{supp}(\varphi(t)) \subset W' \Subset W$. Then we have

$$\begin{aligned}
(u \circ T, D_j\varphi)_{L^2(0,T;L^2(W))} &= (u \circ T, D_j\varphi)_{L^2(0,T;L^2(W'))} \\
&\overset{(1)}{=} \lim_{n\to\infty} (u_n \circ T, D_j\varphi)_{L^2(0,T;L^2(W'))} \\
&\overset{(2)}{=} -\lim_{n\to\infty} (D_j(u_n \circ T), \varphi)_{L^2(0,T;L^2(W'))} \\
&\overset{(3)}{=} -\lim_{n\to\infty} \left(\left(\sum_{i=1}^d (D_i u_n) \circ T \right) D_j T_i, \varphi \right)_{L^2(0,T;L^2(W'))} \\
&\overset{(4)}{=} -(f_j, \varphi)_{L^2(0,T;L^2(W))}\,.
\end{aligned}$$

In (1) we used (4.11) and the continuity of the inner product. In (2) we applied integration by parts. This is possible since $u_n \circ T$ is as a composition of two Lipschitz functions again Lipschitz continuous with respect to the spatial coordinates and hence by Rademacher's theorem differentiable almost everywhere. (3) follows by the chain rule and in (4) we used continuity of the inner product and (4.12).

Finally, by definition of the weak derivative we conclude that $D_j(u \circ T) = f_j \in L^2(0,T;L^2(W))$ and therefore $u \circ T \in L^2(0,T;H^1(W))$. □

Remark 4.1. Let us briefly comment on the sequence $(u_n)_n$ in the proof of Theorem 4.4 The idea is to approximate the function $u \in L^2(0,T;H^1(U))$ with smooth functions with respect to x. This will be done with standard regularization theory. Let $(\varrho_n)_n$ be a sequence of mollifiers, [1, Def. 6.2] and $(\eta_n)_n$ a sequence of cut-off functions according to [1, p.171]. We denote with $\tilde{u}$ the extension of u by zero with respect to x, i.e.

$$\tilde{u}(t,x) = \begin{cases} u(t,x) & (t,x) \in (0,T) \times U, \\ 0 & (t,x) \in (0,T) \times \mathbb{R}^d \setminus U \end{cases}.$$

We define for a.e. $t \in (0,T)$ and for all $x \in \mathbb{R}^d$ the functions

$$u_n(t,x) := \eta_n(x) \int_{\mathbb{R}^d} \varrho_n(x-y)\tilde{u}(t,y)\,\mathrm{d}y.$$

For the sequence $(u_n)_n$ there holds that $u_n(t) \in C_0^\infty(\mathbb{R}^d)$ for all $n \in \mathbb{N}$ and that $\|u_n(t) - u(t)\|_{H^1(U')} \to 0$, $n \to \infty$ for a.e. $t \in (0,T)$ and for all $U' \Subset U$, cf. [1, Satz 6.16], [47, Lemma 2.1.3]. Further, we have by using the Cauchy-Schwarz inequality and $\int_{\mathbb{R}^d} \varrho(x)\,\mathrm{d}x = 1$ that

$$\begin{aligned} |u_n(t,x)| &\le \int_{\mathbb{R}^d} \varrho_n(x-y)^{\frac{1}{2}} \varrho_n(x-y)^{\frac{1}{2}} \,|\tilde{u}(t,y)|\,\mathrm{d}y \\ &\le \left(\int_{\mathbb{R}^d} \varrho_n(x-y)\mathrm{d}y\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d} \varrho_n(x-y)\,|\tilde{u}(t,y)|^2\,\mathrm{d}y\right)^{\frac{1}{2}} \\ &= \left(\int_{\mathbb{R}^d} \varrho_n(x-y)\,|\tilde{u}(t,y)|^2\,\mathrm{d}y\right)^{\frac{1}{2}}, \text{ for a.e. } t \in (0,T),\ \forall x \in \mathbb{R}^d. \end{aligned} \tag{4.13}$$

Using (4.13) and Fubini's Theorem we get

$$\begin{aligned} \|u_n(t)\|^2_{L^2(U')} &\le \int_{U'} \int_{\mathbb{R}^d} \varrho_n(x-y)\,|\tilde{u}(t,y)|^2\,\mathrm{d}y\mathrm{d}x \\ &\le \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \varrho_n(x-y)\mathrm{d}x\right) |\tilde{u}(t,y)|^2\,\mathrm{d}y \\ &= \|u(t)\|^2_{L^2(U)}, \text{ for a.e. } t \in (0,T). \end{aligned} \tag{4.14}$$

For sufficiently large n and for almost all $t \in (0,T)$ we have, cf. [1, Satz 6.16]

$$D_j(u_n(t)) = \varrho_n * \widetilde{D_j(u(t))} \quad \text{in } U'. \tag{4.15}$$

Using (4.15) and similar estimation techniques as in (4.13) we obtain

$$|D_j u_n(t,x)| \le \left(\int_{\mathbb{R}^d} \varrho_n(x-y) \left|\widetilde{D_j u}(t,y)\right|^2 \mathrm{d}y\right)^{\frac{1}{2}}, \text{ for a.e. } t \in (0,T),\ \forall x \in U'. \tag{4.16}$$

By Fubini's theorem we conclude for sufficiently large n that

$$\begin{aligned} \|D_j u_n(t)\|^2_{L^2(U')} &\le \int_{U'} \int_{\mathbb{R}^d} \varrho_n(x-y) \left|\widetilde{D_j u}(t,y)\right|^2 \mathrm{d}y\mathrm{d}x \\ &\le \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \varrho_n(x-y)\mathrm{d}x\right) \left|\widetilde{D_j u}(t,y)\right|^2 \mathrm{d}y \\ &= \|D_j u(t)\|^2_{L^2(U)}, \text{ for a.e. } t \in (0,T) \end{aligned} \tag{4.17}$$

Now, extracting a subsequence $(u_{n_k})_k \subset (u_n)_n$ one can assume by (4.14), (4.17) that $\|u_{n_k}(t)\|_{H^1(U')} \leq \|u(t)\|_{H^1(U)}$ for all $k \in \mathbb{N}$ and for almost every $t \in (0,T)$. Hence, we can conclude by Lebesgue's theorem that

$$\|u_{n_k} - u\|^2_{L^2(0,T;H^1(U'))} = \int_0^T \|u_{n_k}(t) - u(t)\|^2_{H^1(U')} \,\mathrm{d}t \to 0, \ k \to \infty. \tag{4.18}$$

Moreover, we have the following result.

Lemma 4.5. *Let $D \subset \mathbb{R}^d$ be a bounded Lipschitz domain and $p > 1$. Denote by T_τ the flow of $V \in C^1_c(D, \mathbb{R}^d)$. For any $f \in L^p(D)$, we have*

$$\lim_{t \searrow 0} \|f \circ T_\tau - f\|_{L^p(D)} = 0 \text{ and } \lim_{t \searrow 0} \left\|f \circ T_\tau^{-1} - f\right\|_{L^p(D)} = 0.$$

Proof. See [39, Lemma 2.16] and the reference therein. □

In order to be able to calculate the shape derivative for an optimization problem which is constrained by two PDEs we need an extension of the single-valued case in Theorem 3.2. We follow the lines in [39, Chapter 4.1.4]. Let E_1, E_2, F_1, F_2 be Banach spaces and $\tilde{\tau} > 0$. Consider

$$\begin{aligned} G : [0, \tilde{\tau}] \times E_1 \times E_2 \times F_1 \times F_2 &\to \mathbb{R} \\ (\tau, \varphi, \eta, \psi, \zeta) &\mapsto G(\tau, \varphi, \eta, \psi, \zeta). \end{aligned}$$

We make the following assumptions on the function G.

Assumption 4.1. (D0)

(i) For all $\tau \in [0, \tilde{\tau}]$,$u, \tilde{u} \in E_1, p, \tilde{p} \in E_2$, $w \in F_1, z \in F_2$ the mappings

$$\begin{aligned} &[0,1] \to \mathbb{R},\ s \mapsto G(\tau, u + s\tilde{u}, p, w, z), \\ &[0,1] \to \mathbb{R},\ s \mapsto G(\tau, u, p + s\tilde{p}, w, z) \end{aligned}$$

are absolutely continuous which implies

$$\begin{aligned} G(\tau, \tilde{u}, p, w, z) - G(\tau, u, p, w, z) &= \int_0^1 \mathrm{d}_\varphi G(\tau, s\tilde{u} + (1-s)u, p, w, z; \tilde{u} - u) \,\mathrm{d}s, \\ G(\tau, u, \tilde{p}, w, z) - G(\tau, u, p, w, z) &= \int_0^1 \mathrm{d}_\eta G(\tau, u, s\tilde{p} + (1-s)p, w, z; \tilde{p} - p) \,\mathrm{d}s. \end{aligned}$$

(ii) For all $\tau \in [0, \tilde{\tau}]$, $u, \tilde{u}, \hat{\varphi} \in E_1$, $p, \tilde{p}, \hat{\eta} \in E_2$, $w \in F_1$, $z \in F_2$

$$s \mapsto \mathrm{d}_\varphi G(\tau, u + s\tilde{u}, p, w, z; \hat{\varphi}) \text{ and } s \mapsto \mathrm{d}_\eta G(\tau, u, p + s\tilde{p}, w, z; \hat{\eta})$$

are well-defined and belong to $L^1(0,1)$.

(iii) the mapping $\psi \mapsto G(\tau, \varphi, \eta, \psi, \zeta)$ is affine linear for all $(\tau, \varphi, \eta, \zeta) \in [0, \tilde{\tau}] \times E_1 \times E_2 \times F_2$ and $\zeta \mapsto G(\tau, \varphi, \eta, \psi, \zeta)$ is affine linear for all $(\tau, \varphi, \eta, \psi) \in [0, \tilde{\tau}] \times E_1 \times E_2 \times F_1$.

For any $\tau \in [0, \tilde{\tau}]$ the (perturbed) system of state equations is given by

$$\mathrm{d}_\psi G(\tau, u, p, w, z; \hat{\psi}) = 0 \ \forall \hat{\psi} \in F_1 \tag{4.19a}$$

$$\mathrm{d}_\zeta G(\tau, u, p, w, z; \hat{\zeta}) = 0 \ \forall \hat{\zeta} \in F_2. \tag{4.19b}$$

We introduce the set of solution to (4.19) by

$$\boldsymbol{E}(\tau) := \{(u, p) \in E_1 \times E_2 : \ (u, p) \text{ solves } (4.19)\}.$$

For any $\tau \in [0, \tilde{\tau}]$, $\boldsymbol{q}^\tau := (u^\tau, p^\tau) \in \boldsymbol{E}(\tau)$, $\boldsymbol{q}^0 := (u^0, p^0) \in \boldsymbol{E}(0)$ the averaged adjoint system reads

$$\int_0^1 \mathrm{d}_\varphi G(\tau, su^\tau + (1-s)u^0, p^\tau, w^\tau, z^\tau; \hat{\varphi}) \, \mathrm{d}s = 0 \ \forall \hat{\varphi} \in E_1, \tag{4.20a}$$

$$\int_0^1 \mathrm{d}_\eta G(\tau, u^0, sp^\tau + (1-s)p^0, w^\tau, z^\tau; \hat{\eta}) \, \mathrm{d}s = 0 \ \forall \hat{\eta} \in E_2. \tag{4.20b}$$

We denote the set of solutions to system (4.20) by

$$\boldsymbol{Y}(\tau, \boldsymbol{q}^\tau, \boldsymbol{q}^0) = \{(w^\tau, z^\tau) \in F_1 \times F_2 : (w^\tau, z^\tau) \text{ solves } (4.20)\}$$

Note that for $\tau = 0$ the set $\boldsymbol{Y}(0, \boldsymbol{q}^0) := \boldsymbol{Y}(0, \boldsymbol{q}^0, \boldsymbol{q}^0)$ corresponds to the solution set of the usual adjoint system. That is $(w, z) \in F_1 \times F_2$ solves

$$\mathrm{d}_\varphi G(0, u, p, w, z; \hat{\varphi}) = 0 \ \forall \hat{\varphi} \in E_1, \tag{4.21a}$$

$$\mathrm{d}_\eta G(0, u, p, w, z; \hat{\eta}) = 0 \ \forall \hat{\eta} \in E_2, \tag{4.21b}$$

where $(u, p) := (u^0, p^0) \in \boldsymbol{E}(0)$. Now we are able to state a modification of Theorem 3.2

Theorem 4.6 ([39, Thm. 4.5])**.** *Let E_1, E_2, F_1, F_2 be Banach spaces, $\tilde{\tau} > 0$ and a function*

$$\begin{aligned} G : [0, \tilde{\tau}] \times E_1 \times E_2 \times F_1 \times F_2 &\to \mathbb{R} \\ (\tau, \varphi, \eta, \psi, \zeta) &\mapsto G(\tau, \varphi, \eta, \psi, \zeta), \end{aligned}$$

be given. Let Assumption (D0) be satisfied and assume that the following conditions hold:

(D1) For all $\varphi \in E_1$, $\eta \in E_2$, $\psi \in F_1$, $\zeta \in F_2$ the mapping

$$[0, \tilde{\tau}] \to \mathbb{R} : \tau \mapsto G(\tau, \varphi, \eta, \psi, \zeta)$$

is differentiable.

(D2) *For all* $\tau \in [0, \tilde{\tau}]$ *let* $\boldsymbol{E}(\tau)$ *be non-empty and single valued. Moreover, for all* $\tau \in [0, \tilde{\tau}]$, $\boldsymbol{q}^\tau \in \boldsymbol{E}(\tau)$, $\boldsymbol{q}^0 \in \boldsymbol{E}(0)$ *let* $\boldsymbol{Y}(\tau, \boldsymbol{q}^\tau, \boldsymbol{q}^0)$ *be non-empty and single-valued.*

(D3) *Let* $\boldsymbol{q}^0 \in E(0)$ *and* $\boldsymbol{p}^0 \in \boldsymbol{Y}(0, \boldsymbol{q}^0)$. *For any sequence of non-negative real numbers* $(\tau_n)_{n\mathbb{N}}$ *converging to zero, there exists a subsequence* $(\tau_{n_k})_{k\in\mathbb{N}}$, *elements* $\boldsymbol{q}^{\tau_{n_k}} \in \boldsymbol{E}(\tau_{n_k})$ *and* $\boldsymbol{p}^{\tau_{n_k}} \in \boldsymbol{Y}(\tau_{n_k}, \boldsymbol{q}^{\tau_{n_k}}, \boldsymbol{q}^0)$ *such that*

$$\lim_{\substack{k\to\infty \\ s\searrow 0}} \partial_\tau G(s, \boldsymbol{q}^0, \boldsymbol{p}^{\tau_{n_k}}) = \partial_\tau G(0, \boldsymbol{q}^0, \boldsymbol{p}^0).$$

Then for any $\tilde{\boldsymbol{\psi}} := (\psi, \zeta) \in F_1 \times F_2$:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(G(\tau, \boldsymbol{q}^\tau, \tilde{\boldsymbol{\psi}})\right)\Big|_{\tau=0} = \partial_\tau G(0, \boldsymbol{q}^0, \boldsymbol{p}^0).$$

4.2 Shape derivative of $\mathcal{J}_1$ and $\mathcal{J}_2$

In this section we will compute the shape derivative of model problem 1 and 2 via the averaged adjoint method. At this point we want to mention that the obtained derivatives $\mathrm{d}\mathcal{J}_1(\Omega; V)$, $\mathrm{d}\mathcal{J}_2(\Omega; V)$ have to be taken with care, since for the moment we are not able to prove all assumptions in Theorem 4.6. We will comment on the crucial steps in more detail when deriving the shape derivatives for $\mathcal{J}_1$, $\mathcal{J}_2$.

Throughout this section we assume $V \in C_c^1(D, \mathbb{R}^d)$, i.e. $V = V(x)$.

4.2.1 Shape derivative of $\mathcal{J}_1$

Recall the problem: For given $u_d, p_d, f_1, f_2 \in C^1(\overline{Q_D})$ we are looking for a solution of the problem

$$\min_{\Omega\in\mathcal{A}(D)} J_1(\Omega, u, p) := \int_0^T \int_\Omega |u - u_d|^2 \,\mathrm{d}x\mathrm{d}t + \int_0^T \int_\Omega |p - p_d|^2 \,\mathrm{d}x\mathrm{d}t, \tag{4.22}$$

where $(u, p) \in X_0(\Omega) \times X_0(\Omega)$ is solution to

$$a(\Omega, u, v) - b(\Omega, p, v) = F_1(\Omega, v) \quad \forall v \in Y(\Omega), \tag{4.23a}$$

$$a(\Omega, p, q) - b(\Omega, u, q) = F_2(\Omega, q) \quad \forall q \in Y(\Omega), \tag{4.23b}$$

with

$$a(\Omega, u, v) := \int_{Q_\Omega} \partial_t uv \,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} \nabla_x u \cdot \nabla_x v \,\mathrm{d}x\mathrm{d}t,$$

$$b(\Omega, p, v) := \int_{Q_\Omega} pv \,\mathrm{d}x\mathrm{d}t,$$

$$F_1(\Omega, v) := \int_{Q_\Omega} f_1 v \,\mathrm{d}x\mathrm{d}t,$$

$$F_2(\Omega, q) := \int_{Q_\Omega} f_2 q \,\mathrm{d}x\mathrm{d}t,$$

and the spaces

$$X_0(\Omega) := \left\{ v \in X(\Omega) : \ v|_{\Sigma_{\Omega 0}} = 0 \right\},$$
$$X(\Omega) := L^2(0,T;H_0^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega)),\ Y(\Omega) := L^2(0,T;H_0^1(\Omega)).$$

In the following we use the Greek letters $\varphi, \eta, \psi, \zeta$ as variables while the roman letters (u,p), (w,z) are used for the solution of the state system and adjoint state system, respectively, cf. [21]. In order to set up the re-parametrized Lagrangian G_1 we first consider the Lagrangian $\mathcal{L}_1$ which reads

$$\begin{aligned}\mathcal{L}_1(\Omega,\varphi,\eta,\psi,\zeta) := J_1(\Omega,\varphi,\eta) + a(\Omega,\varphi,\psi) - b(\Omega,\eta,\psi) - F_1(\Omega,\psi)\\ + a(\Omega,\eta,\zeta) - b(\Omega,\varphi,\zeta) - F_2(\Omega,\zeta).\end{aligned}$$

Note that the reduced functional $\mathcal{J}_1(\Omega)$ can be written in terms of $\mathcal{L}_1$ since

$$\mathcal{J}_1(\Omega) := J_1(\Omega,u,p) = \mathcal{L}_1(\Omega,u,p,\psi,\zeta).$$

In order to obtain the shape derivative we need to compute

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\, \mathcal{J}_1(\Omega_\tau)|_{\tau=0} = \frac{\mathrm{d}}{\mathrm{d}\tau}\, \mathcal{L}_1(\Omega_\tau, u_\tau, p_\tau, \hat{\psi}, \hat{\zeta})\Big|_{\tau=0},$$

where $(u_\tau, p_\tau) \in X_0(\Omega_\tau) \times X_0(\Omega_\tau)$ solves the perturbed state system

$$a(\Omega_\tau, u_\tau, \hat{v}) - b(\Omega_\tau, p_\tau, \hat{v}) = F_1(\Omega_\tau, \hat{v}) \quad \forall \hat{v} \in Y(\Omega_\tau), \tag{4.24a}$$
$$a(\Omega_\tau, p_\tau, \hat{q}) - b(\Omega_\tau, u_\tau, \hat{q}) = F_2(\Omega_\tau, \hat{q}) \quad \forall \hat{q} \in Y(\Omega_\tau). \tag{4.24b}$$

However, the derivative of $\mathcal{L}_1$ can not be computed straightforwardly since after a change of variables one has to deal with functions $\hat{\varphi} \circ T_\tau$, where $\hat{\varphi} \in X_0(\Omega_\tau)$ is a function on the moving domain Ω_τ, cf. [21]. To get around this difficulty one defines a re-parametrized Lagrangian

$$\begin{aligned}&G_1 : [0,\tilde{\tau}] \times X_0(\Omega) \times X_0(\Omega) \times X^0(\Omega) \times X^0(\Omega) \to \mathbb{R},\\ &\quad (\tau,\varphi,\eta,\psi,\zeta) \mapsto G_1(\tau,\varphi,\eta,\psi,\zeta),\\ &\quad G_1(\tau,\varphi,\eta,\psi,\zeta) := \mathcal{L}_1(\Omega_\tau, \varphi\circ T_\tau^{-1}, \eta\circ T_\tau^{-1}, \psi\circ T_\tau^{-1}, \zeta\circ T_\tau^{-1})\end{aligned} \tag{4.25}$$

where $X^0(\Omega) := \{w \in X(\Omega) : \ w(x,T) = 0,\ x \in \Omega\}$.

Remark 4.2. In order to define the re-parametrized Lagrangian we assume that

$$\varphi \in X_0(\Omega) \iff \varphi \circ T_\tau^{-1} \in X_0(\Omega_\tau),$$
$$\psi \in X^0(\Omega) \iff \psi \circ T_\tau^{-1} \in X^0(\Omega_\tau).$$

If the spaces X_0 and X^0 correspond to $L^2(\Omega)$ or $W^{1,p}(\Omega)$, $p \geq 1$ this assumption is fulfilled, see Theorem 4.3 and [17, Remark 3.4]. Also for the space $L^2(0,T;H_0^1(\Omega))$ this property is clear by Theorem 4.4. However, for the space $H^1(0,T;H^{-1}(\Omega))$ the re-parametrization property is an open question. We want to mention that this property was already used by Sokołowski in [33].

The re-parametrized Lagrangian G_1 reads after a change of variables

$$\begin{aligned} G_1(\tau,\varphi,\eta,\psi,\zeta) = &\int_{Q_\Omega} \xi(\tau)\,|\varphi - u_d^\tau|^2\,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} \xi(\tau)\,|\eta - p_d^\tau|^2\,\mathrm{d}x\mathrm{d}t \\ &+ \int_{Q_\Omega} \xi(\tau)\partial_t\varphi\psi\,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} A(\tau)\nabla_x\varphi\cdot\nabla_x\psi\,\mathrm{d}x\mathrm{d}t \\ &- \int_{Q_\Omega} \xi(\tau)\eta\psi\,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} \xi(\tau)f_1^\tau\psi\,\mathrm{d}x\mathrm{d}t \\ &+ \int_{Q_\Omega} \xi(\tau)\partial_t\eta\zeta\,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} A(\tau)\nabla_x\eta\cdot\nabla_x\zeta\,\mathrm{d}x\mathrm{d}t \\ &- \int_{Q_\Omega} \xi(\tau)\varphi\zeta\,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} \xi(\tau)f_2^\tau\zeta\,\mathrm{d}x\mathrm{d}t, \end{aligned} \tag{4.26}$$

where $u_d^\tau := u_d \circ T_\tau$, $p_d^\tau := p_d \circ T_\tau$, $f_1^\tau := f_1 \circ T_\tau$, $f_2^\tau := f_2 \circ T_\tau$ and

$$\xi(\tau) := \det(\partial T_\tau) = |\det(\partial T_\tau)| \text{ for } \tau \geq 0 \text{ small}, \tag{4.27}$$

$$A(\tau) := \det(\partial T_\tau)\partial T_\tau^{-1}\partial T_\tau^{-T}. \tag{4.28}$$

Remark 4.3. Note that we assume $V \in C_c^1(D,\mathbb{R}^d)$, i.e. $V = V(x)$. Therefore the transformation $\Omega_\tau = T_\tau(\Omega)$ only acts on the space variables. This means that the time derivatives in (4.26) do not transform after a change of variables. Further the shortcut $f^\tau := f \circ T_\tau$ for a function $f : Q_D \to \mathbb{R}$ means that $f^\tau(t,y) = f(t,T_\tau(x))$, where $y \in \Omega_\tau$ and $x \in \Omega$.

The function G_1 can be interpreted as perturbed Lagrangian brought back to the original domain Ω [13].

The shape derivative in the context of the re-parametrized Lagrangian G_1 reads

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\,\mathcal{J}_1(\Omega_\tau)|_{\tau=0} = \frac{\mathrm{d}}{\mathrm{d}\tau}\,(G_1(\tau,u^\tau,p^\tau,\psi,\zeta))|_{\tau=0}\,, \tag{4.29}$$

where $(u^\tau,p^\tau) \in X_0(\Omega)\times X_0(\Omega)$ solves

$$\begin{aligned} \int_{Q_\Omega} \xi(\tau)\partial_t u^\tau\psi\,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} A(\tau)\nabla_x u^\tau\cdot\nabla_x\psi\,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} \xi(\tau)p^\tau\psi\,\mathrm{d}x\mathrm{d}t \\ = \int_{Q_\Omega} \xi(\tau)f_1^\tau\psi\,\mathrm{d}x\mathrm{d}t \quad \forall\psi\in Y(\Omega), \end{aligned} \tag{4.30a}$$

$$\begin{aligned} \int_{Q_\Omega} \xi(\tau)\partial_t p^\tau\zeta\,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} A(\tau)\nabla_x p^\tau\cdot\nabla_x\zeta\,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} \xi(\tau)u^\tau\zeta\,\mathrm{d}x\mathrm{d}t \\ = \int_{Q_\Omega} \xi(\tau)f_2^\tau\zeta\,\mathrm{d}x\mathrm{d}t \quad \forall\zeta\in Y(\Omega). \end{aligned} \tag{4.30b}$$

Note that (4.30) corresponds to the perturbed state system (4.24) after a change of variables $\Omega_\tau = T_\tau(\Omega)$. Therefore if (u^τ,p^τ) solves (4.30) and (u_τ,p_τ) solves (4.24)

these solutions can be written in terms of each other via $u_\tau = u^\tau \circ T_\tau^{-1}$ and $p_\tau = p^\tau \circ T_\tau^{-1}$.

In order to compute the shape derivative we want to apply Theorem 4.6 to (4.29). Hence, one has to check assumptions (D0)-(D3) in Theorem 4.6. Let us briefly comment on these assumptions. The ideas are taken from [13] and are adapted to our setting.

(D0)(i) can be verified by showing that the maps

$$s \mapsto G_1(\tau, u + s\tilde{u}, \eta, \psi, \zeta), \;\; s \mapsto G_1(\tau, \varphi, p + s\tilde{p}, \psi, \zeta)$$

are continuous differentiable. This can be done by using the theorem of Lebesgue and the fact that the integrand is continuously differentiable with respect to s.

(D0)(ii) can be shown by direct computation using the boundedness of the mappings $\tau \mapsto A(\tau)$, $\tau \mapsto \xi(\tau)$ described in Theorem 4.2.

(D0)(iii) follows by construction of G_1. This means in order to see that G_1 is affine linear with respect to ψ write

$$G_1(\tau, \varphi, \eta, \psi, \zeta) = C(\tau, \varphi, \eta, \psi) + D(\tau, \varphi, \eta, \zeta) \tag{4.31}$$

where $C(\tau, \varphi, \eta, \psi)$ is linear and $D(\tau, \varphi, \eta, \zeta)$ is constant with respect to ψ. The same can be done to see that G_1 is affine linear with respect to ζ.

Assumption $(D1)$ can be shown by Lebesgue's theorem using the fact that the integrand is differentiable with respect to τ since the mappings $\tau \mapsto A(\tau)$ and $\tau \mapsto \xi(\tau)$ are differentiable according to Lemma 4.1.

Assumption (D2) can not be verified for the moment since we have not proved any existence and uniqueness results for the state and adjoint state system in this master thesis. This task is more involved and is postponed to future work.

In order to verify condition (D3) we assume that assumption (D2) is true and that there exist constants $c_1, c_2 > 0$ such that

$$\forall \tau \in [0, \tilde{\tau}] : \;\; \|(u^\tau, p^\tau)\|_{X \times X} \leq c_1, \;\; \|(w^\tau, z^\tau)\|_{X \times X} \leq c_2, \tag{4.32}$$

where $(u^\tau, p^\tau) =: \boldsymbol{q}^\tau \in \boldsymbol{E}(\tau)$, $(w^\tau, z^\tau) =: \boldsymbol{p}^\tau \in \boldsymbol{Y}(\tau, \boldsymbol{q}^\tau, \boldsymbol{q}^0)$. Note that (4.32) is usually a consequence of (D2). We show that for all sequences $(\tau_n)_n \searrow 0$ there exists a subsequence $(\tau_{n_k})_k \subset (\tau_n)_n$ such that

(i) $\exists\, \boldsymbol{p}^{\tau_{n_k}} \in \boldsymbol{Y}\left(\tau_{n_k}, \boldsymbol{q}^{\tau_{n_k}}, \boldsymbol{q}^0\right) : \;\; \boldsymbol{p}^{\tau_{n_k}} \rightharpoonup \boldsymbol{p}^0$ in $X \times X$,

(ii) $(\tau, \psi, \zeta) \mapsto G_1(\tau, u, p, \psi, \zeta)$ is weakly continuous.

We start with (i). Let $(\tau_n)_n \searrow 0$. By (4.32) there exist a constant $c_1 > 0$ such that for all $n \in \mathbb{N}$

$$\|(u^{\tau_n}, p^{\tau_n})\|_{X \times X} \leq c_1.$$

Hence, $\left(u^{\tau_n}, p^{\tau_n}\right)_n \subset X \times X$ is a bounded sequence in a reflexive Banach space. Therefore there exists a subsequence $\left(u^{\tau_{n_k}}, p^{\tau_{n_k}}\right)_k \subset \left(u^{\tau_n}, p^{\tau_n}\right)_n$ and $z_1, z_2 \in X$ such that

$(u^{\tau_{n_k}}, p^{\tau_{n_k}}) \rightharpoonup (z_1, z_2)$ in $X \times X$. Since $(u^{\tau_{n_k}}, p^{\tau_{n_k}}) \in \boldsymbol{E}(\tau_{n_k})$ we have $(u^{\tau_{n_k}}, p^{\tau_{n_k}}) \in X_0 \times X_0$ and

$$\begin{aligned}\int_{Q_\Omega} \xi(\tau_{n_k}) \partial_t u^{\tau_{n_k}} \hat{\psi} \,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} A(\tau_{n_k}) \nabla_x u^{\tau_{n_k}} \cdot \nabla_x \hat{\psi} \,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} \xi(\tau_{n_k}) p^{\tau_{n_k}} \hat{\psi} \,\mathrm{d}x\mathrm{d}t \\ = \int_{Q_\Omega} \xi(\tau_{n_k}) f_1^{\tau_{n_k}} \hat{\psi} \,\mathrm{d}x\mathrm{d}t \quad \forall \hat{\psi} \in X^0,\end{aligned} \tag{4.33}$$

$$\begin{aligned}\int_{Q_\Omega} \xi(\tau_{n_k}) \partial_t p^{\tau_{n_k}} \hat{\zeta} \,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} A(\tau_{n_k}) \nabla_x p^{\tau_{n_k}} \cdot \nabla_x \hat{\zeta} \,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} \xi(\tau_{n_k}) u^{\tau_{n_k}} \hat{\zeta} \,\mathrm{d}x\mathrm{d}t \\ = \int_{Q_\Omega} \xi(\tau_{n_k}) f_2^{\tau_{n_k}} \hat{\zeta} \,\mathrm{d}x\mathrm{d}t \quad \forall \hat{\zeta} \in X^0.\end{aligned} \tag{4.34}$$

Using the weak convergence $(u^{\tau_{n_k}}, p^{\tau_{n_k}}) \rightharpoonup (z_1, z_2)$, the continuity of the mappings $\tau \mapsto \xi(\tau)$, $\tau \mapsto A(\tau)$ and that according to Lemma 4.5 $f_1^\tau \to f_1$ in $L^2(Q_\Omega)$, $f_2^\tau \to f_2$ in $L^2(Q_\Omega)$, we may pass to the limit $k \to \infty$ in (4.33), (4.34) and obtain due to $\boldsymbol{E}(0) = \{(u, p)\}$ that $z_1 = u$, $z_2 = p$. Hence, we conclude $(u^{\tau_{n_k}}, p^{\tau_{n_k}}) \rightharpoonup (u, p)$ in $X \times X$. By assumption (4.32) the sequence $(w^{\tau_{n_k}}, z^{\tau_{n_k}})_k$ is bounded. Therefore we can extract a subsequence which is again denoted by $(w^{\tau_{n_k}}, z^{\tau_{n_k}})_k$ and there exist $q_1, q_2 \in X$ such that $(w^{\tau_{n_k}}, z^{\tau_{n_k}}) \rightharpoonup (q_1, q_2)$ in $X \times X$. Since $(w^{\tau_{n_k}}, z^{\tau_{n_k}}) \in \boldsymbol{Y}(\tau_{n_k}, \boldsymbol{q}^{\tau_{n_k}}, \boldsymbol{q}^0)$ there holds

$$\begin{aligned}&- \int_{Q_\Omega} \xi(\tau_{n_k}) \partial_t w^{\tau_{n_k}} \hat{\varphi} \,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} A(\tau_{n_k}) \nabla_x w^{\tau_{n_k}} \cdot \nabla_x \hat{\varphi} \,\mathrm{d}x\mathrm{d}t \\ &- \int_{Q_\Omega} \xi(\tau_{n_k}) z^{\tau_{n_k}} \hat{\varphi} \,\mathrm{d}x\mathrm{d}t = - \int_{Q_\Omega} \xi(\tau_{n_k}) (u^0 + u^{\tau_{n_k}} - 2u_d^{\tau_{n_k}}) \hat{\varphi} \,\mathrm{d}x\mathrm{d}t \quad \forall \hat{\varphi} \in X_0,\end{aligned} \tag{4.35}$$

$$\begin{aligned}&- \int_{Q_\Omega} \xi(\tau_{n_k}) \partial_t z^{\tau_{n_k}} \hat{\eta} \,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} A(\tau_{n_k}) \nabla_x z^{\tau_{n_k}} \cdot \nabla_x \hat{\eta} \,\mathrm{d}x\mathrm{d}t \\ &- \int_{Q_\Omega} \xi(\tau_{n_k}) w^{\tau_{n_k}} \hat{\eta} \,\mathrm{d}x\mathrm{d}t = - \int_{Q_\Omega} \xi(\tau_{n_k}) (p^0 + p^{\tau_{n_k}} - 2p_d^{\tau_{n_k}}) \hat{\eta} \,\mathrm{d}x\mathrm{d}t \quad \forall \hat{\eta} \in X_0,\end{aligned} \tag{4.36}$$

Using the weak convergence of the sequences $(w^{\tau_{n_k}}, z^{\tau_{n_k}}) \rightharpoonup (q_1, q_2)$, $(u^{\tau_{n_k}}, p^{\tau_{n_k}}) \rightharpoonup (u, p)$, the continuity of the mappings $\tau \mapsto \xi(\tau)$, $\tau \mapsto A(\tau)$ and that according to Lemma 4.5 $u_d^{\tau_{n_k}} \to u_d$ in $L^2(Q_\Omega)$, $p_d^{\tau_{n_k}} \to p_d$ in $L^2(Q_\Omega)$ we may pass to the limit $k \to \infty$ in (4.35), (4.36) and obtain due to $\boldsymbol{Y}(0, \boldsymbol{q}^0) = \{(w, z)\}$ that $q_1 = w$,$q_2 = z$. Hence, assertion (i) follows. Finally, note that (ii) is fulfilled and we conclude condition (D3).

Even though we can not verify all conditions of Theorem 4.6, we will use it to compute the shape derivative. A formal computation yields

$$\begin{aligned}
\mathrm{d}\mathcal{J}_1(\Omega;V) = \partial_\tau G_1(0,u,p,w,z) &= \int_{Q_\Omega} \operatorname{div} V\,|u-u_d|^2\,\mathrm{d}x\mathrm{d}t \\
&- \int_{Q_\Omega} 2(u-u_d)\nabla_x u_d \cdot V\,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} \operatorname{div} V\,|p-p_d|^2\,\mathrm{d}x\mathrm{d}t \\
&- \int_{Q_\Omega} 2(p-p_d)\nabla_x p_d \cdot V\,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} \operatorname{div} V\,\partial_t u w\,\mathrm{d}x\mathrm{d}t \\
&+ \int_{Q_\Omega} (\operatorname{div} V - \partial V - \partial V^T)\nabla_x u \cdot \nabla_x w\,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} \operatorname{div} V p w\,\mathrm{d}x\mathrm{d}t \\
&- \int_{Q_\Omega} \operatorname{div} V f_1 w\,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} \nabla_x f_1 \cdot V w\,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} \operatorname{div} V \partial_t p z\,\mathrm{d}x\mathrm{d}t \\
&+ \int_{Q_\Omega} (\operatorname{div} V - \partial V - \partial V^T)\nabla_x p \cdot \nabla_x z\,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} \operatorname{div} V u z\,\mathrm{d}x\mathrm{d}t \\
&- \int_{Q_\Omega} \operatorname{div} V f_2 z\,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} \nabla_x f_2 \cdot V z\,\mathrm{d}x\mathrm{d}t,
\end{aligned} \tag{4.37}$$

where $(u,p) = (u^0,p^0)$ is solution to the state system (4.23) and $(w,z) = (w^0,z^0)$ is solution to the adjoint system which reads: Find (w,z) such that

$$-\partial_t w - \Delta w - z = -2(u-u_d) \qquad \text{in } Q_\Omega, \tag{4.38a}$$
$$-\partial_t z - \Delta z - w = -2(p-p_d) \qquad \text{in } Q_\Omega, \tag{4.38b}$$
$$w = 0 \qquad \text{on } \Sigma_\Omega, \tag{4.38c}$$
$$z = 0 \qquad \text{on } \Sigma_\Omega, \tag{4.38d}$$
$$w(T) = 0 \qquad \text{in } \Omega, \tag{4.38e}$$
$$z(T) = 0 \qquad \text{in } \Omega. \tag{4.38f}$$

The adjoint system can be derived out of the equations

$$\mathrm{d}_\varphi G_1(0,u,p,w,z)(v) = 0 \quad \forall v \in X_0, \tag{4.39}$$
$$\mathrm{d}_\eta G_1(0,u,p,w,z)(q) = 0 \quad \forall q \in X_0. \tag{4.40}$$

In the following we will show how to do that by deriving the adjoint equation for the state u out of (4.39). Expression (4.39) is equivalent to

$$\int_{Q_\Omega} \partial_t v w\,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} \nabla_x v \cdot \nabla_x w\,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} v z\,\mathrm{d}x\mathrm{d}t = -\int_{Q_\Omega} 2(u-u_d)v\,\mathrm{d}x\mathrm{d}t \quad \forall v \in X_0.$$

Integration by parts with respect to the variable t yields

$$\begin{aligned}
&\int_\Omega v(x,T)w(x,T)\,\mathrm{d}x - \int_\Omega v(x,0)w(x,0)\,\mathrm{d}x - \int_{Q_\Omega} \partial_t w v\,\mathrm{d}x\mathrm{d}t + \\
&\int_{Q_\Omega} \nabla_x w \cdot \nabla_x v\,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} v z\,\mathrm{d}x\mathrm{d}t = -\int_{Q_\Omega} 2(u-u_d)v\,\mathrm{d}x\mathrm{d}t \quad \forall v \in X_0.
\end{aligned} \tag{4.41}$$

Since $v \in X_0$ and $w \in X^0$ the first and second integral of (4.41) vanishes and we end up with

$$\begin{aligned} -\int_{Q_\Omega} \partial_t w v \,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} \nabla_x w \cdot \nabla_x v \,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} vz \,\mathrm{d}x\mathrm{d}t = \\ - \int_{Q_\Omega} 2(u - u_d) v \,\mathrm{d}x\mathrm{d}t \quad \forall v \in X_0. \end{aligned} \tag{4.42}$$

Equation (4.42) is even well defined for $v \in Y$. Therefore a solution to the variational formulation: Find $w \in X^0$ such that

$$\begin{aligned} -\int_{Q_\Omega} \partial_t w v \,\mathrm{d}x\mathrm{d}t + \int_{Q_\Omega} \nabla_x w \cdot \nabla_x v \,\mathrm{d}x\mathrm{d}t - \int_{Q_\Omega} vz \,\mathrm{d}x\mathrm{d}t = \\ - \int_{Q_\Omega} 2(u - u_d) v \,\mathrm{d}x\mathrm{d}t \quad v \in Y, \end{aligned} \tag{4.43}$$

is also a solution to (4.42) and hence $\mathrm{d}_\varphi G_1(0, u, p, w, z)(v) = 0 \quad \forall v \in X_0$. We will use variational formulation (4.43) for numerical implementation as it suits our setting described in Chapter 2 better. Note that (4.43) is the weak formulation of

$$-\partial_t w - \Delta w - z = -2(u - u_d) \qquad \text{in } Q_\Omega, \tag{4.44a}$$
$$w = 0 \qquad \text{on } \Sigma_\Omega \tag{4.44b}$$
$$w(T) = 0 \qquad \text{in } \Omega. \tag{4.44c}$$

The adjoint equation for z can be derived in the same way out of (4.40).

4.2.2 Shape derivative of $\mathcal{J}_2$

Recall model problem 2: For given functions $u_d, p_d, f_1, f_2 \in C^1(\overline{Q_D})$ we consider the shape optimization problem

$$\min_{\Omega \in \mathcal{A}(D)} J_2(u, p) := \int_0^T \int_D |u - u_d|^2 \,\mathrm{d}x\mathrm{d}t + \int_0^T \int_D |p - p_d|^2 \,\mathrm{d}x\mathrm{d}t, \tag{4.45}$$

where $(u, p) \in X_0 \times X_0$ is solution to

$$a_1(\Omega, u, v) - b(p, v) = F_1(v) \quad \forall v \in Y, \tag{4.46a}$$
$$a_2(\Omega, p, q) - b(u, q) = F_2(q) \quad \forall q \in Y, \tag{4.46b}$$

with

$$
\begin{aligned}
a_1(\Omega, u, v) &:= \int_{Q_D} \partial_t uv \,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \nu_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x\mathrm{d}t,\\
a_2(\Omega, p, q) &:= \int_{Q_D} \partial_t pq \,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \lambda_\Omega \nabla p \cdot \nabla q \,\mathrm{d}x\mathrm{d}t\\
b(p, v) &:= \int_{Q_D} pv \,\mathrm{d}x\mathrm{d}t,\\
F_1(v) &:= \int_{Q_D} f_1 v \,\mathrm{d}x\mathrm{d}t,\\
F_2(q) &:= \int_{Q_D} f_2 q \,\mathrm{d}x\mathrm{d}t,
\end{aligned}
$$

where $X_0 := \left\{v \in X : \ v|_{\Sigma_{D0}} = 0\right\}$, $X = L^2(0,T;H_0^1(D)) \cap H^1(0,T;H^{-1}(D))$ and $Y = L^2(0,T;H_0^1(D))$. The material parameters are given as $\nu_\Omega := \nu_1\chi_{Q_\Omega} + \nu_2(1-\chi_{Q_\Omega})$ and $\lambda_\Omega := \lambda_1\chi_{Q_\Omega} + \lambda_2(1-\chi_{Q_\Omega})$, where $\nu_1, \nu_2, \lambda_1, \lambda_2 > 0$ and χ_{Q_Ω} denotes the characteristic function. Note that in contrast to problem 1 the dependence of the functional J_2 on Ω is given only implicitly via the states u, p. We can now follow the same procedure as in Section 4.2.1 to build the Lagrangian and to compute the shape derivative. We will use again the Greek letters $(\varphi, \eta, \psi, \zeta)$ as variables while the roman letters (u, p, w, z) are used for the solution of the state and adjoint state system. The Lagrangian $\mathcal{L}_2$ of problem 2 reads

$$
\begin{aligned}
\mathcal{L}_2(\Omega, \varphi, \eta, \psi, \zeta) := J_2(\varphi, \eta) + a_1(\Omega, \varphi, \psi) + b(\eta, \psi) - F_1(\psi)\\
+ a_2(\Omega, \eta, \zeta) + b(\varphi, \zeta) - F_2(\zeta).
\end{aligned}
$$

In order to compute the shape derivative we introduce the re-parametrized Lagrangian

$$
\begin{aligned}
&G_2 : [0, \tilde{\tau}] \times X_0 \times X_0 \times X^0 \times X^0 \to \mathbb{R},\\
&\quad (\tau, \varphi, \eta, \psi, \zeta) \mapsto G_2(\tau, \varphi, \eta, \psi, \zeta),\\
&\quad G_2(\tau, \varphi, \eta, \psi, \zeta) := \mathcal{L}_2(\Omega_\tau, \varphi \circ T_\tau^{-1}, \eta \circ T_\tau^{-1}, \psi \circ T_\tau^{-1}, \zeta \circ T_\tau^{-1})
\end{aligned}
\tag{4.47}
$$

where $X^0 := \{v \in X : \ v(x,T) = 0, \ x \in D\}$.

Remark 4.4. In order to build the re-parametrized Lagrangian we assume

$$
\begin{aligned}
\varphi \in X_0 &\iff \varphi \circ T_\tau^{-1} \in X_0\\
\psi \in X^0 &\iff \psi \circ T_\tau^{-1} \in X^0.
\end{aligned}
$$

The re-parametrized Lagrangian G_2 reads after a change of variables

$$\begin{aligned} G_2(\tau,\varphi,\eta,\psi,\zeta) &= \int_{Q_D} \xi(\tau)\,|\varphi - u_d^\tau|^2\,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \xi(\tau)\,|\eta - p_d^\tau|^2\,\mathrm{d}x\mathrm{d}t \\ &+ \int_{Q_D} \xi(\tau)\partial_t\varphi\psi\,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \nu_\Omega A(\tau)\nabla_x\varphi\cdot\nabla_x\psi\,\mathrm{d}x\mathrm{d}t \\ &- \int_{Q_D} \xi(\tau)\eta\psi\,\mathrm{d}x\mathrm{d}t - \int_{Q_D} \xi(\tau) f_1^\tau\psi\,\mathrm{d}x\mathrm{d}t \\ &+ \int_{Q_D} \xi(\tau)\partial_t\eta\zeta\,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \lambda_\Omega A(\tau)\nabla_x\eta\cdot\nabla_x\zeta\,\mathrm{d}x\mathrm{d}t \\ &- \int_{Q_D} \xi(\tau)\varphi\zeta\,\mathrm{d}x\mathrm{d}t - \int_{Q_D} \xi(\tau) f_2^\tau\zeta\,\mathrm{d}x\mathrm{d}t, \end{aligned} \tag{4.48}$$

Note that the transformation T_τ does not change the hold all domain, meaning that $T_\tau(D) = D$, $T_\tau(\partial D) = \partial D$. Hence the domain of integration stays the same before and after the change of variables. The shape derivative in terms of the Lagrangian G_2 reads

$$\mathrm{d}\mathcal{J}_2(\Omega;V) = \frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{J}_2(\Omega_\tau) = \frac{\mathrm{d}}{\mathrm{d}\tau}\, G_2(\tau,u^\tau,p^\tau,\psi,\zeta)|_{\tau=0}\,, \tag{4.49}$$

where (u^τ, p^τ) solves

$$\begin{aligned} &\int_{Q_D} \xi(\tau)\partial_t u^\tau\psi\,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \nu_\Omega A(\tau)\nabla_x u^\tau\cdot\nabla_x\psi\,\mathrm{d}x\mathrm{d}t - \int_{Q_D} \xi(\tau)p^\tau\psi\,\mathrm{d}x\mathrm{d}t \\ &\qquad = \int_{Q_D} \xi(\tau) f_1^\tau\psi\,\mathrm{d}x\mathrm{d}t \quad \forall\psi\in Y, \end{aligned} \tag{4.50a}$$

$$\begin{aligned} &\int_{Q_D} \xi(\tau)\partial_t p^\tau\zeta\,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \lambda_\Omega A(\tau)\nabla_x p^\tau\cdot\nabla_x\zeta\,\mathrm{d}x\mathrm{d}t - \int_{Q_D} \xi(\tau)u^\tau\zeta\,\mathrm{d}x\mathrm{d}t \\ &\qquad = \int_{Q_D} \xi(\tau) f_2^\tau\zeta\,\mathrm{d}x\mathrm{d}t \quad \forall\zeta\in Y. \end{aligned} \tag{4.50b}$$

In order to compute the shape derivative we again want to make use of Theorem 4.6. Assumptions (D0),(D1),(D3) can be shown analogously as in section 4.2.1 described. However, assumption (D2) can not be verified for the moment since we have not considered any existence and uniqueness results to the state and adjoint state system,

respectively. Nevertheless we apply Theorem 4.6 to (4.49) and get

$$\begin{aligned}
\mathrm{d}\mathcal{J}_2(\Omega;V) = \partial_\tau G_2(0,u,p,w,z) = &\int_{Q_D} \operatorname{div} V\,|u-u_d|^2\,\mathrm{d}x\mathrm{d}t \\
&- \int_{Q_D} 2(u-u_d)\nabla_x u_d \cdot V\,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \operatorname{div} V\,|p-p_d|^2\,\mathrm{d}x\mathrm{d}t \\
&- \int_{Q_D} 2(p-p_d)\nabla_x p_d \cdot V\,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \operatorname{div} V \partial_t u w\,\mathrm{d}x\mathrm{d}t \\
&+ \int_{Q_D} (\operatorname{div} V - \partial V - \partial V^T)\nu_\Omega \nabla_x u \cdot \nabla_x w\,\mathrm{d}x\mathrm{d}t - \int_{Q_D} \operatorname{div} V p w\,\mathrm{d}x\mathrm{d}t \\
&- \int_{Q_D} \operatorname{div} V f_1 w\,\mathrm{d}x\mathrm{d}t - \int_{Q_D} \nabla_x f_1 \cdot V w\,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \operatorname{div} V \partial_t p z\,\mathrm{d}x\mathrm{d}t \\
&+ \int_{Q_D} (\operatorname{div} V - \partial V - \partial V^T)\lambda_\Omega \nabla_x p \cdot \nabla_x z\,\mathrm{d}x\mathrm{d}t - \int_{Q_D} \operatorname{div} V u z\,\mathrm{d}x\mathrm{d}t \\
&- \int_{Q_D} \operatorname{div} V f_2 z\,\mathrm{d}x\mathrm{d}t - \int_{Q_D} \nabla_x f_2 \cdot V z\,\mathrm{d}x\mathrm{d}t,
\end{aligned} \tag{4.51}$$

where (u,p) solves (4.46) and (w,z) is solution to the adjoint system which reads

$$-\partial_t w - \operatorname{div}(\nu_\Omega \nabla w) - z = -2(u-u_d) \qquad \text{in } Q_D, \tag{4.52a}$$
$$-\partial_t z - \operatorname{div}(\lambda_\Omega \nabla z) - w = -2(p-p_d) \qquad \text{in } Q_D, \tag{4.52b}$$
$$w = 0 \qquad \text{on } \Sigma_D, \tag{4.52c}$$
$$z = 0 \qquad \text{on } \Sigma_D, \tag{4.52d}$$
$$w(T) = 0 \qquad \text{in } D, \tag{4.52e}$$
$$z(T) = 0 \qquad \text{in } D. \tag{4.52f}$$

4.3 Numerical algorithm

In this section we focus on the numerical method which is used to solve model problem 1 and 2. First of all we summarize the main steps of the algorithm and describe the procedure. Afterwards we discuss the computation of a descent direction in more detail and demonstrate how one can obtain it in an advantageous manner in a parabolic setting.

Let us briefly summarize the main ideas of the algorithm. The presented algorithm is of gradient-type and makes use of the shape derivatives $\mathrm{d}\mathcal{J}_1(\Omega;V)$ and $\mathrm{d}\mathcal{J}_2(\Omega;V)$ computed in the previous section in order to generate descent directions. First of all we have to solve the state equations (4.23), (4.46) also known as forward problem to get the states u,p. We solve this problem by using space-time finite elements as

explained in Chapter 2. Next we have to solve the adjoint problems (4.38),(4.52) also known as backward problem to get the adjoint states w, z. This is also done by using the approach explained in Chapter 2. Afterwards we compute a descent direction V via an auxiliary boundary value problem as explained in (3.17). In the last step we choose a step size τ and move every node of the finite element mesh according to the step size in direction V. This procedure results in the following algorithm:

Algorithm 4.7. *Choose an initial design* Ω_0.
For $n = 0, 1, 2, ...$ *until converge,*

(1) Solve (4.23) *to obtain* u_n, p_n

(2) Solve (4.38) *to obtain* w_n, z_n

(3) Find a vector field V_n *such that* $\mathrm{d}\mathcal{J}_1(\Omega_n; V_n) < 0$

(4) Choose a step size τ_n *according to the step size rule*

(5) Set $\Omega_{n+1} = (id + \tau_n V_n)(\Omega_n)$

Remark 4.5. Algorithm 4.7 is written in terms of model problem 1. If one wants to solve model problem 2 replace (4.23),(4.38),$\mathrm{d}\mathcal{J}_1(\Omega_n; V_n)$ with (4.46),(4.52),$\mathrm{d}\mathcal{J}_2(\Omega_n; V_n)$.

Let us comment on step 3 of the algorithm. We have to compute a descent direction which only acts on the space variables and lets the time variable unaffected. This is different to a shape optimization problem in an elliptic setting, but we will see that it can be done in a similar way. We explain the computation of a descent direction in terms of model problem 1.

In order to determine a descent direction we solve the auxiliary boundary value problem: Find $V \in C^1(\overline{\Omega}, \mathbb{R}^d)$ such that

$$\int_\Omega \partial_x V : \partial_x W + V \cdot W \mathrm{d}x\mathrm{d}t = -\mathrm{d}\mathcal{J}_1(\Omega, W) \quad \forall W \in C^1(\overline{\Omega}, \mathbb{R}^d). \tag{4.53}$$

This can be done in the following way: First we compute a vector field $\tilde{V} \in C^1(\overline{Q_\Omega}, \mathbb{R}^d)$ as solution to the boundary value problem

$$\int_{Q_\Omega} \partial_x \tilde{V} : \partial_x \tilde{W} + \tilde{V} \cdot \tilde{W} \mathrm{d}x = -\mathrm{d}\mathcal{J}_1(\Omega; \tilde{W}) \quad \forall \tilde{W} \in C^1(\overline{Q_\Omega}, \mathbb{R}^d). \tag{4.54}$$

Note that the solution $\tilde{V}$ to (4.54) is time dependent. This is not what we want to have since a time dependent vector field would lead to geometries which change in time. To overcome this issue we compute in a second step a time independent vector field V by averaging the vector field $\tilde{V}$ with respect to the time variable, i.e.

$$V(x) = \int_0^T \tilde{V}(x, s) \mathrm{d}s. \tag{4.55}$$

The vector field $V \in C^1(\overline{\Omega}, \mathbb{R}^d)$ then solves (4.53). This can be seen by testing (4.54) with $W \in C^1(\overline{\Omega}, \mathbb{R}^d)$ which yields

$$\begin{aligned} -\mathrm{d}\mathcal{J}_1(\Omega; W) &= \int_{Q_\Omega} \partial_x \tilde{V} : \partial_x W + \tilde{V} \cdot W \mathrm{d}x\mathrm{d}t \\ &= \int_\Omega \partial_x V : \partial_x W + V \cdot W \mathrm{d}x \quad \forall W \in C^1(\overline{\Omega}, \mathbb{R}^d), \end{aligned}$$

where we used the fact that

$$\begin{aligned} \int_{Q_\Omega} \tilde{V} \cdot W \mathrm{d}x\mathrm{d}t &= \int_\Omega \int_0^T \sum_{j=1}^d \tilde{V}_j(x,t) W_j(x) \mathrm{d}t\mathrm{d}x \\ &= \int_\Omega \sum_{j=1}^d \int_0^T \tilde{V}_j(x,t) W_j(x) \mathrm{d}t\mathrm{d}x = \int_\Omega \sum_{j=1}^d V_j(x) W_j(x) \mathrm{d}x \\ &= \int_\Omega V \cdot W \mathrm{d}x, \end{aligned}$$

and

$$\begin{aligned} \int_{Q_\Omega} \partial_x \tilde{V} : \partial_x W \mathrm{d}x\mathrm{d}t &= \int_\Omega \int_0^T \sum_{i,j=1}^d \partial_j \tilde{V}_i(x,t) \partial_j W_i(x) \mathrm{d}t\mathrm{d}x \\ &= \int_\Omega \sum_{i,j=1}^d \partial_j W_i(x) \partial_j \int_0^T \tilde{V}(x,t) \mathrm{d}t\mathrm{d}x = \int_\Omega \partial_x V : \partial_x W \mathrm{d}x. \end{aligned}$$

Remark 4.6. A descent direction for model problem 2 can be computed in the same way as described above. One just has to replace Ω by D, the space C^1 by C_c^1 and $\mathrm{d}\mathcal{J}_1$ by $\mathrm{d}\mathcal{J}_2$.

4.4 Results for model problem 1

In this section we apply Algorithm 4.7 to model problem 1. We present numerical examples in one and two space variables, respectively. The implementation is done in the finite element software Netgen/NGSolve [31].

4.4.1 1d-1d

In this example we choose $\Omega_{ref} = (0,1)$ and $T = 1$. Thus, $Q_{\Omega_{ref}} = (0,1)^2$, which is the unit square. We choose the functions u_d, p_d to be

$$\begin{aligned} u_d(x,t) &= \sin(\pi x)\sin(\pi t) \quad \text{for } (x,t) \in Q_{ref}, \\ p_d(x,t) &= x(1-x)t(1-t) \quad \text{for } (x,t) \in Q_{ref}. \end{aligned}$$

iteration	$\mathcal{J}_1$	Ω
0	1.043e-01	$(0.2, 0.8)$
3	4.477e-02	$(0.109, 0.89)$
7	2.227e-04	$(0.007, 0.993)$
15	3.487e-09	$(0, 1)$

Table 4.1: Evolution of cost $\mathcal{J}_1$ and domain Ω for problem (4.22)-(4.23) with $d = 1$

and determine f_1, f_2 such that u_d, p_d is solution to (1.18) on $Q_{\Omega_{ref}}$. Hence, the functions f_1, f_2 read

$$\begin{aligned} f_1(x,t) &= -p_d(x,t) + \pi\cos(\pi t)\sin(\pi x) + \pi^2\sin(\pi t)\sin(\pi x) \quad \text{for } (x,t) \in Q_{\Omega_{ref}}, \\ f_2(x,t) &= -u_d(x,t) + x(1-x)(1-2t) + 2t(1-t) \qquad\qquad\qquad\quad \text{for } (x,t) \in Q_{\Omega_{ref}}. \end{aligned}$$

We choose as initial domain $\Omega = (0.2, 0.8)$ which result in $Q_\Omega = (0.2, 0.8) \times (0, 1)$. Due to the chosen functions we expect that the initial shape Ω converges to Ω_{ref}. In the numerical test we solved all arising PDEs with piecewise quadratic and globally continuous finite elements and choose the step size to be $\tau = 0.3$ in every iteration.

In Table 4.1 and Figure 4.1 one can see the numerical results. Table 4.1 shows the evolution of the cost $\mathcal{J}_1$ and the domain Ω while in Figure 4.1 one can see the corresponding computational domains Q_Ω. We observe that after 15 iterations the domain Ω is basically identical to the reference domain Ω_{ref} with cost $\mathcal{J}_1$=3.487e-09. Moreover, we see that the mesh quality over the optimization process is satisfying.

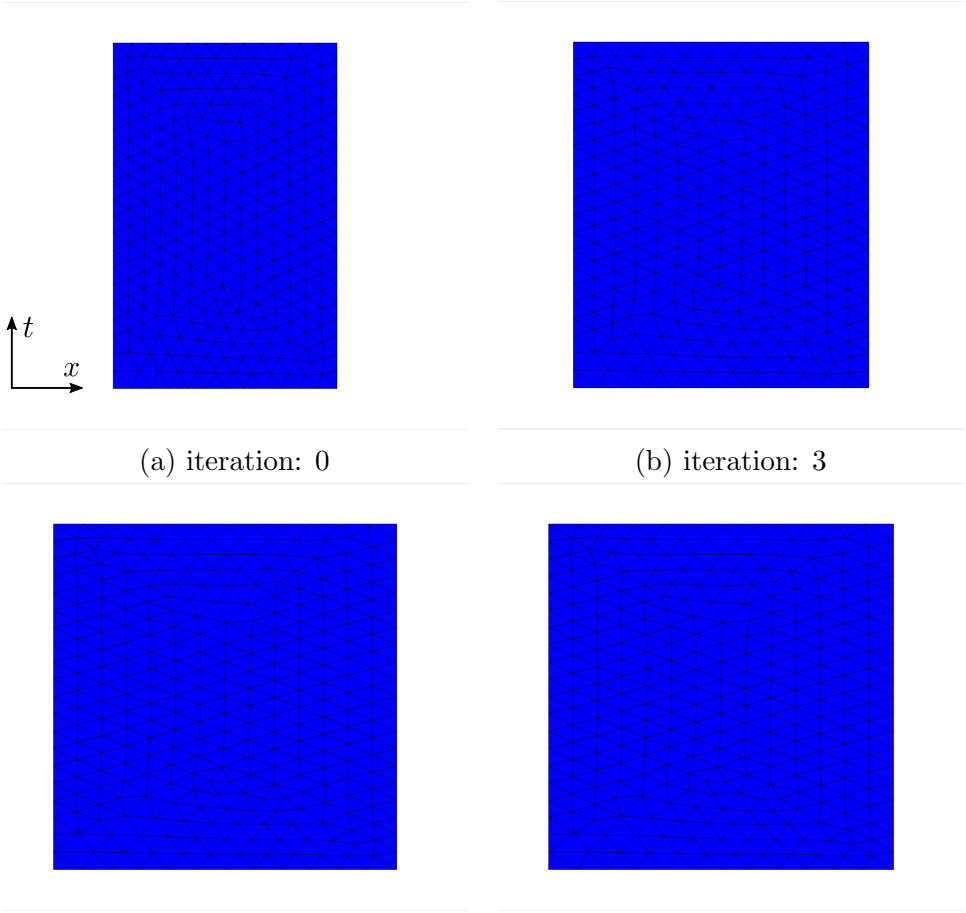
4.4.2 2d-1d

In this example we consider model problem 1 in two spatial coordinates, i.e. $d = 2$. In the numerical test we choose the reference domain $\Omega_{ref} := (0, 1)^2$ and $T = 1$. The reference space-time cylinder is then simply the unit cube, i.e. $Q_{ref} = (0, 1)^3$. We choose the functions u_d, p_d to be

$$\begin{aligned} u_d(x,y,t) &= \sin(\pi x)\sin(\pi y)\sin(\pi t) \qquad \text{for } (x,y,t) \in Q_{ref}, \\ p_d(x,y,t) &= x(1-x)y(1-y)t(1-t) \quad \text{for } (x,y,t) \in Q_{ref}, \end{aligned}$$

and determine f_1, f_2 such that u_d, p_d is solution to (1.18) on the reference space-time cylinder. The functions f_1, f_2 are therefore given by

$$\begin{aligned} f_1(x,y,t) &= -p_d(x,y,t) + \pi\sin(\pi x)\sin(\pi y)\cos(\pi t) \\ &\quad + 2\pi^2 u_d(x,y,t) \qquad \text{for } (x,y,t) \in Q_{ref}, \\ f_2(x,y,t) &= -u_d(x,y,t) + x(1-x)y(1-y)(1-2t) + 2y(1-y)t(1-t) \\ &\quad + 2x(1-x)t(1-t) \qquad \text{for } (x,y,t) \in Q_{ref}. \end{aligned}$$



(a) iteration: 0 (b) iteration: 3

(c) iteration: 7 (d) iteration: 15

Figure 4.1: Evolution of the computational domain Q_Ω for problem (4.22)-(4.23) with $d = 1$

iteration	$\mathcal{J}_1$
0	5.078e-02
5	4.104e-04
10	2.632e-04
20	1.237e-04
80	1.362e-05
200	3.565e-06

Table 4.2: Evolution of cost function $\mathcal{J}_1$ for problem (4.22)-(4.23) with $d = 2$

We choose our initial domain to be $\Omega := (0.2, 0.8)^2$ and thus the space-time cylinder reads $Q_\Omega = (0.2, 0.8)^2 \times (0, 1)$. Moreover, all arising PDEs were solved with piecewise quadratic and globally continuous finite elements. Since the functional (1.17) is of tracking type we expect that the initial domain Ω converge to Ω_{ref}. In the numerical test we used a constant step size of $\tau = 1.3$ in every iteration.

The results of the simulation are shown in Table 4.2 and Figure 4.2. Table 4.2 shows the values of the objective function $\mathcal{J}_1$ while Figure 4.2 depicts the evolution of shape Ω. We observe that the values of the cost function decrease from 5.078e-02 in iteration 0 to 3.565e-06 in iteration 200. Moreover, we see that the initial domain converges as expected to the reference domain. Finally, we want to mention that after 200 iterations we still observed a decrease of the objective functional in every iteration step.

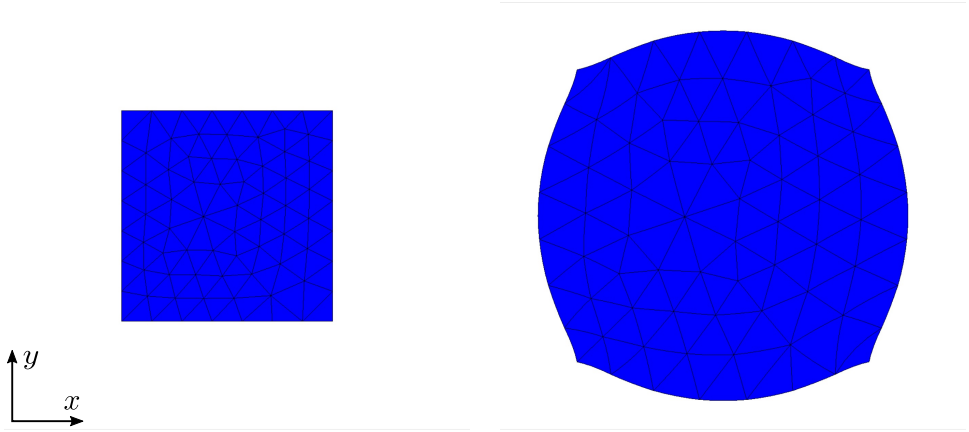
4.5 Results for model problem 2

In this section we apply Algorithm 4.7 to model problem 2. We present numerical examples in one and two space variables, respectively. The implementation is done in the finite element software Netgen/NGSolve [31].

4.5.1 1d-1d

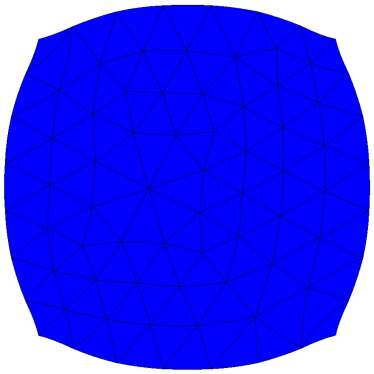
We choose $D = (0, 1)$, $\Omega = (0.2, 0.8)$, $\Omega_{ref} = (0.4, 0.6)$ and $T = 1$. We assume that the material in the domain Q_Ω corresponds to iron while the domain $Q_D \setminus Q_\Omega$ corresponds to air. The particular choice of the material leads to the material coefficients

$$\nu_\Omega = \begin{cases} \frac{1}{\mu_0 \mu_{r_1}} & \text{in } Q_\Omega \\ \frac{1}{\mu_0 \mu_{r_2}} & \text{in } Q_D \setminus Q_\Omega \end{cases}, \quad \lambda_\Omega = \begin{cases} 80.2 & \text{in } Q_\Omega \\ 0.0262 & \text{in } Q_D \setminus Q_\Omega \end{cases},$$

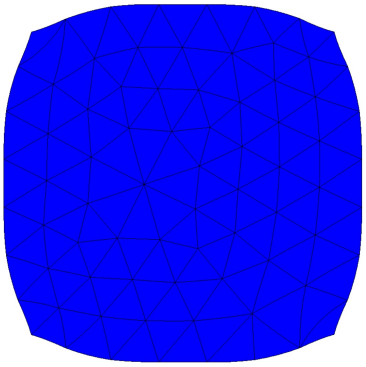


(a) iteration: 0

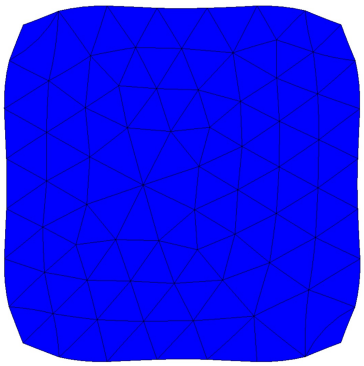
(b) iteration: 5



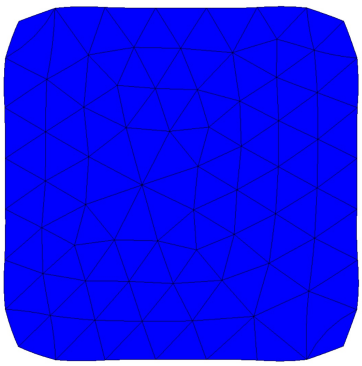
(c) iteration: 10



(d) iteration: 20



(e) iteration: 80



(f) iteration: 200

Figure 4.2: Evolution of the domain Ω for problem (4.22)-(4.23) with $d = 2$

iteration	$\mathcal{J}_2$	Ω
0	1.107e-03	$(0.2, 0.8)$
2	1.134e-04	$(0.3182, 0.6818)$
5	2.273e-05	$(0.3581, 0.6419)$
10	4.907e-06	$(0.3789, 0.6211)$
20	6.259e-07	$(0.3921, 0.6078)$
50	2.417e-08	$(0.3991, 0.6009)$

Table 4.3: Evolution of the objective function $\mathcal{J}_2$ and domain Ω for problem (4.45)-(4.46) with $d = 1$

where $\mu_0 = \frac{4\pi}{10^7}$, $\mu_{r_1} = 5100$, $\mu_{r_2} = 1$. We choose the functions f_1, f_2 to be

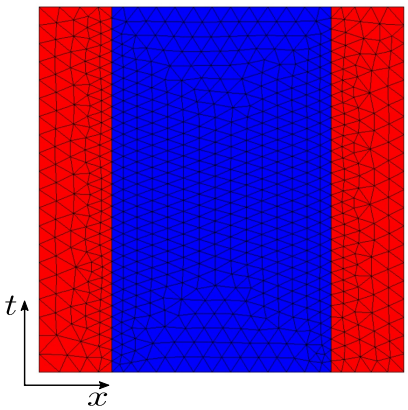
$$f_1(x,t) = -x(1-x)t(1-t) + \pi\cos(\pi t)\sin(\pi x) + \pi^2\sin(\pi t)\sin(\pi x) \quad \text{for } (x,t) \in Q_D,$$
$$f_2(x,t) = -\sin(\pi x)\sin(\pi t) + x(1-x)(1-2t) + 2t(1-t) \quad \text{for } (x,t) \in Q_D.$$

The functions u_d, p_d are chosen as numerical solution to (1.21) on Q_D where we used the material parameters $\nu_{\Omega_{ref}}$, $\lambda_{\Omega_{ref}}$, i.e. we assume our reference geometry to consist of iron in Ω_{ref} and air in $D \setminus \Omega_{ref}$. With these settings we expect that the initial domain, which consists of iron in Ω and air in $D \setminus \Omega$, converges to the reference geometry. We solved all arising PDEs with piecewise quadratic and globally continuous finite elements and used a constant step size of $\tau = 100$ in every iteration.

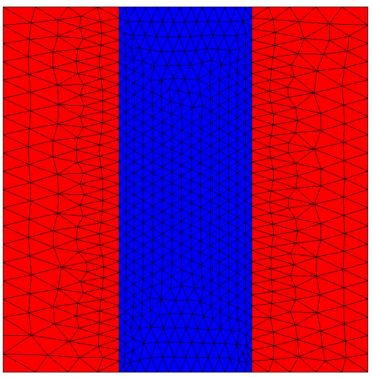
The numerical results can be seen in Table 4.3 and Figure 4.3. Table 4.3 shows the objective value $\mathcal{J}_2$ and the interval Ω in different iterations while Figure 4.3 depicts the evolution of the computational domain Q_D. The blue area in Figure 4.3 indicates iron while the area in red corresponds to air. We observe that the objective value decreases from a start value of 1.107e-03 to 2.417e-08 in iteration 50. Further we see that the interval converge from $(0.2, 0.8)$ to $(0.3991, 0.6009)$ which is almost Ω_{ref}. Hence the simulation fulfills our expectations. Finally we want to mention that also the mesh quality is satisfying during the whole optimization process as one can see in Figure 4.3.

4.5.2 2d-1d

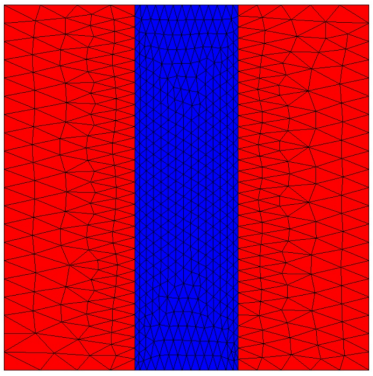
In this example we consider model problem 2 in two spatial coordinates, i.e. $d = 2$. We choose $D = (0,1)^2$, $T = 1$, $\Omega = (0.35, 0.65)^2$, $x_0 = (0.5, 0.5)^T$ and $\Omega_{ref} = \{x \in D : \|x - x_0\|_2 < 0.3\}$. We assume that the material in the domains Ω, Ω_{ref} corresponds to air while $D \setminus \Omega$, $D \setminus \Omega_{ref}$ has the material properties of iron. Hence the reference computational domain can be imagined as an iron cube with a hole of radius 0.3. The initial computational domain can be viewed as an iron cube with a



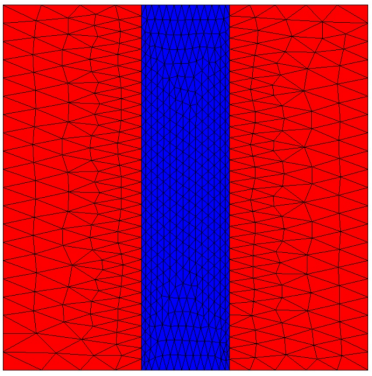
(a) iteration: 0



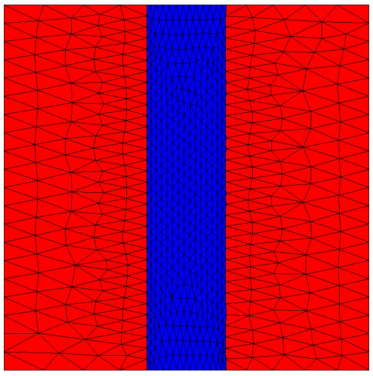
(b) iteration: 2



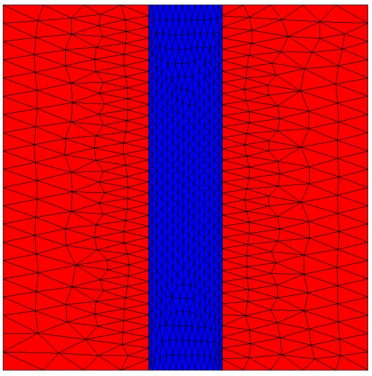
(c) iteration: 5



(d) iteration: 10



(e) iteration: 20



(f) iteration: 50

Figure 4.3: Evolution of the computational domain Q_D for problem (4.45)-(4.46) with $d = 1$, blue: iron, red: air

iteration	$\mathcal{J}_2$
0	2.815e-03
2	1.802e-03
5	5.502e-04
10	4.088e-05
50	1.155e-06
100	9.608e-07

Table 4.4: Evolution of the objective function $\mathcal{J}_2$ for problem (4.45)-(4.46) with $d = 2$

square hole. We choose the material parameters

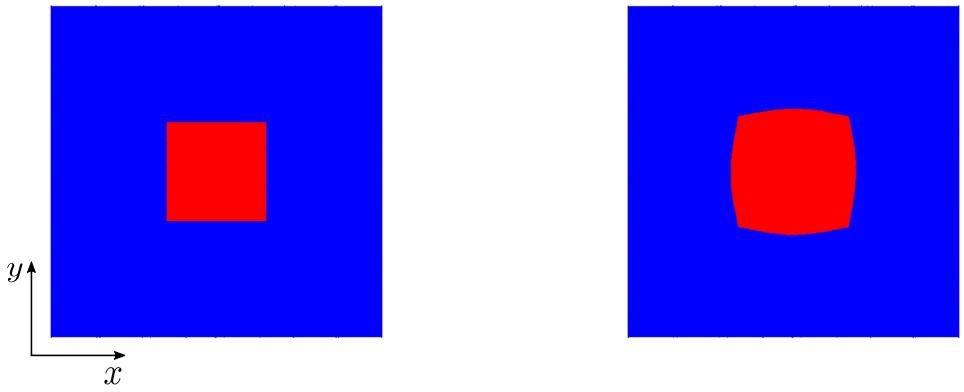
$$\nu_\Omega = \begin{cases} \frac{1}{\mu_0 \mu_{r_1}} & \text{in } Q_\Omega \\ \frac{1}{\mu_0 \mu_{r_2}} & \text{in } Q_D \setminus Q_\Omega \end{cases}, \quad \lambda_\Omega = \begin{cases} 0.0262 & \text{in } Q_\Omega \\ 80.2 & \text{in } Q_D \setminus Q_\Omega \end{cases},$$

where $\mu_0 = \frac{4\pi}{10^7}$, $\mu_{r_1} = 1$, $\mu_{r_2} = 5100$. The functions f_1, f_2 read

$$\begin{aligned} f_1(x,y,t) &= -x(1-x)y(1-y)t(1-t) + \pi \sin(\pi x)\sin(\pi y)\cos(\pi t) \\ &\quad + 2\pi^2 u_d(x,y,t) \qquad \text{for } (x,y,t) \in Q_D, \\ f_2(x,y,t) &= -\sin(\pi x)\sin(\pi y)\sin(\pi t) + x(1-x)y(1-y)(1-2t) + 2y(1-y)t(1-t) \\ &\quad + 2x(1-x)t(1-t) \qquad \text{for } (x,y,t) \in Q_D. \end{aligned}$$

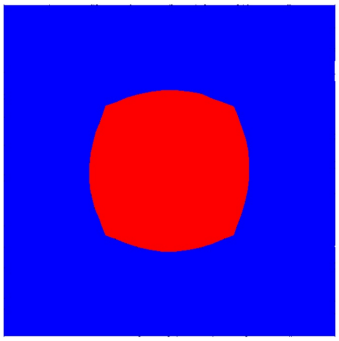
The functions u_d, p_d are chosen as numerical solution to (1.21) on the reference geometry, i.e. we used $\nu_{\Omega_{ref}}$ and $\lambda_{\Omega_{ref}}$ as material coefficients. Through these settings we expect that the initial geometry converges to the reference geometry. We solved all PDEs in the algorithm with piecewise linear and globally continuous finite elements. The step size was chosen with $\tau = 10$ in every iteration.

In Table 4.4 and Figure 4.4 one can see the numerical results. Table 4.4 shows the objective value $\mathcal{J}_2$ in different iterations. We observe that the value of the cost function decreases from 2.815e-03 in iteration 0 to 9.608e-07 in iteration 100. Figure 4.4 depicts the evolution of the domain D. The blue area indicates the material corresponding to iron while the red area characterizes air. One can clearly see that the initial shape converges to the reference shape. Thus the numerical simulation fulfills our expectations.

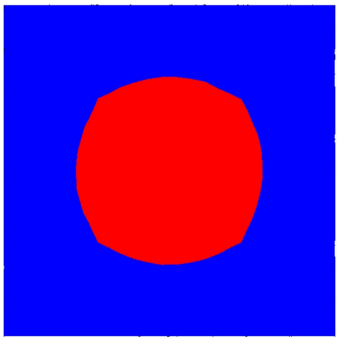


(a) iteration: 0

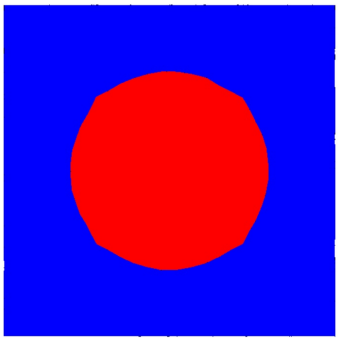
(b) iteration: 2



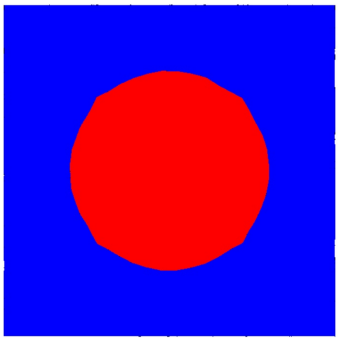
(c) iteration: 5



(d) iteration: 10



(e) iteration: 50



(f) iteration: 100

Figure 4.4: Evolution of the domain D for problem (4.45)-(4.46) with $d = 2$, blue: iron, red: air

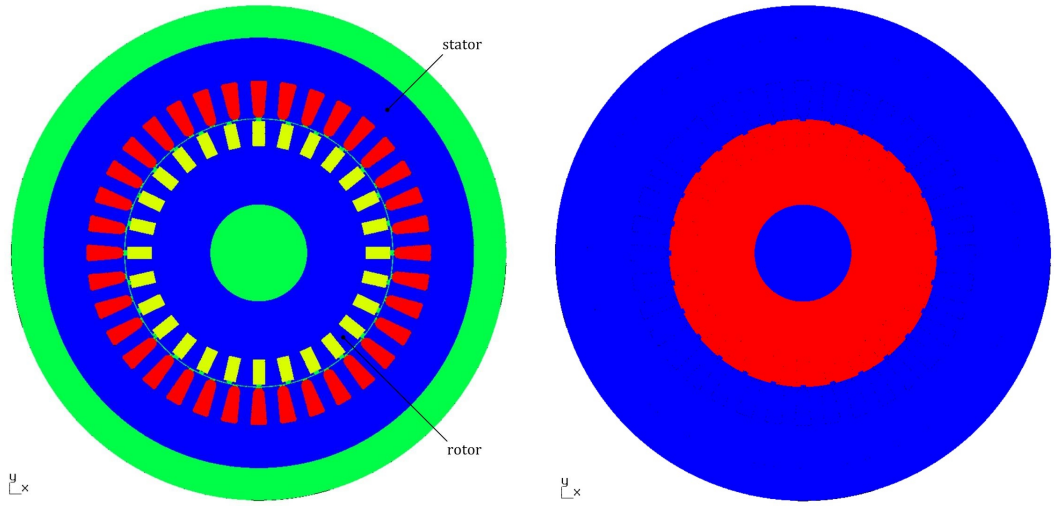
5 Shape optimization of an asynchronous motor

In this chapter we apply the presented shape optimization techniques to an asynchronous motor in order to determine an optimal rotor design with respect to a given performance criterion. We want to mention that we do not use any parallel implementation of the shape optimization algorithm. Hence, we are restricted in the time scale of the geometric model in order to keep the dimensions of the resulting linear system low. Further, we restrict ourselves to a linear model of (1.16) in order to keep things simple. Moreover, we consider a fixed position of the rotor in our simulations. For the above reasons the obtained results have to be understood as proof of concept, meaning that parabolic shape optimization using space-time finite elements can work for this real world application. In order to obtain more practically relevant results one has to take parallel implementation and the nonlinearity in the physical model into account. Since this is beyond the scope of this master thesis we postpone these tasks to future work.

This chapter is organized as follows: First we explain the shape optimization problem and specify problem (1.16) to our concrete application. Next, we discuss the solvability of the state system. Further we derive the shape derivative of the shape optimization problem which we need to set up a gradient-based algorithm. Finally, we apply the algorithm and present some first numerical results.

5.1 Problem formulation

We consider an asynchronous electric motor with a squirrel cage rotor. A schematic construction can be seen in Figure 5.1(a). The motor consists of a fixed part called stator and a rotating part called rotor which are separated by a thin air gap, cf. [12]. The stator consists of coils made of copper which are highlighted in red. The rotor contains aluminum bars, marked in yellow, which are set into grooves and connected at both ends via shorting rings forming a cage-like shape, cf. [44]. We refer to the geometry depicted in Figure 5.1(a) as computational domain $D \subset \mathbb{R}^2$. The colors indicate the areas in the motor with different material. Blue corresponds to iron/steel, red to copper, yellow to aluminum and green to air. We denote the associated domains with Ω_{Fe}, Ω_{Cu}, Ω_{Al} and Ω_{air}, respectively.



(a) (b)

Figure 5.1: (a) Computational domain D, blue: iron, red: copper, green: air, yellow: aluminum, (b) Design domain Ω_d (red)

Remark 5.1 (cf. [12]). For the simulation of electrical machines the computational domain D typically includes a layer of air around the stator. The reason for this is that one assumes a homogeneous Dirichlet boundary condition on the boundary ∂D, which implies that $\boldsymbol{B} \cdot n|_{\partial D} = 0$. This boundary condition in the context of an electric motor is more realistic if the computational domain D includes a layer of air outside the stator.

The design of the rotor grooves and hence the aluminum bars have an impact on the speed-torque characteristics of the induction motor, cf. [44]. We refer the reader to [5, Chapter 6.3] for different designs of the bars and their influence on the speed-torque characteristics. We are interested in finding an optimal design of the rotor bars such that the overall Joule losses in the rotor due to the harmonics are minimized. In order to reach that goal we start from the layout in Figure 5.1(a) and apply shape optimization techniques based on the shape derivative to find an optimal shape of aluminum inside the rotor. Therefore we first need a suitable formulation as a shape optimization problem. We state the problem in the following.

In order to describe the electromagnetic-thermal coupling in the induction motor we use model (1.16). Note that (1.16) is a model for a motor geometry in two space dimensions. We already mentioned that for simulating electrical machines it is common to use a two dimensional geometry in space since the axial dimension of the motor is large compared to its diameter, cf. [12]. Moreover we restrict (1.16) to the linear case, i.e. we assume linear material behavior meaning that ν, λ, σ only depend on

space and we neglect the nonlinear term on the right side of (1.16b). Further the magnetization term on the right hand side of (1.16a) vanishes since we do not have any permanent magnets in the induction motor. In order to state the mathematical model we introduce the set $\Omega_d \subset D$, which denotes the design domain. The design domain can be seen in Figure 5.1(b) which corresponds to the rotor of the motor. Note that the design domain stays fixed during the whole optimization process. We denote by $\Omega \subset \Omega_d$ those parts in the design domain which are currently occupied with aluminum. The current configuration of iron in the motor is then given by

$$\Omega_f := (\Omega_d \setminus \Omega) \cup \Omega_{Fe,s},$$

where $(\Omega_d \setminus \Omega)$ corresponds to the iron in the rotor and $\Omega_{Fe,s}$ denotes the iron in the stator. We denote by $\Gamma_{I,\nu}$ and $\Gamma_{I,\lambda}$ the interface, where the corresponding material parameters ν, λ jump, respectively. The problem reads

$$\min_{\Omega \in \mathcal{A}} J(p) := \int_0^T \int_{\Omega_d} |p|^2 \,\mathrm{d}x\mathrm{d}t \tag{5.1}$$

such that

$$\begin{aligned}
\sigma_\Omega \partial_t u - \operatorname{div}(\nu_\Omega \nabla u) &= J_3 && \text{in } D \times (0,T), && \text{(5.2a)}\\
c_{H,\Omega} \varrho_\Omega \partial_t p - \operatorname{div}(\lambda_\Omega \nabla p) &= -J_3 \partial_t u && \text{in } D \times (0,T), && \text{(5.2b)}\\
u &= 0 && \text{on } \partial D \times (0,T), && \text{(5.2c)}\\
p &= g && \text{on } \partial D \times (0,T), && \text{(5.2d)}\\
u(.,0) &= 0 && \text{in } D, && \text{(5.2e)}\\
p(.,0) &= p_0 && \text{in } D, && \text{(5.2f)}\\
[\![u]\!] &= 0 && \text{on } \Gamma_{I,\nu} \times (0,T), && \text{(5.2g)}\\
[\![\nu_\Omega \nabla u \cdot n]\!] &= 0 && \text{on } \Gamma_{I,\nu} \times (0,T), && \text{(5.2h)}\\
[\![p]\!] &= 0 && \text{on } \Gamma_{I,\lambda} \times (0,T), && \text{(5.2i)}\\
[\![\lambda_\Omega \nabla p \cdot n]\!] &= 0 && \text{on } \Gamma_{I,\lambda} \times (0,T), && \text{(5.2j)}
\end{aligned}$$

where the set of admissible shapes $\mathcal{A}$ is given by

$$\mathcal{A} := \{\Omega \subset \Omega_d : \ \Omega \text{ open, Lipschitz with uniform Lipschitz constant } L_{\mathcal{A}}\}. \tag{5.3}$$

The boundary condition $p = g$ on $\partial D \times (0,T)$ denotes the temperature on the boundary of the air layer outside the stator in the simulation interval $(0,T)$ and we choose $g = 295.15$ Kelvin which corresponds to room temperature. The initial condition $p(.,0) = p_0$ in D denotes the temperature inside the motor at initial time $t = 0$. We choose $p_0 = 295.15$ Kelvin. The material parameters read

$$\nu_\Omega = \begin{cases} \frac{1}{\mu_0} & \text{in } \Omega \cup \Omega_{Cu} \cup \Omega_{air} \\ \frac{1}{4000\mu_0} & \text{in } \Omega_f \end{cases}, \quad \sigma_\Omega = \begin{cases} 37.7 \cdot 10^6 & \text{in } \Omega \\ 58 \cdot 10^6 & \text{in } \Omega_{Cu} \\ 0 & \text{in } \Omega_{air} \\ 10^7 & \text{in } \Omega_f \end{cases}$$

$$\lambda_\Omega = \begin{cases} 236 & \text{in } \Omega \\ 0.1 & \text{in } \Omega_{Cu} \\ 0.026 & \text{in } \Omega_{air} \\ 40 & \text{in } \Omega_f \end{cases}, \quad c_{H,\Omega} = \begin{cases} 897 & \text{in } \Omega \\ 400 & \text{in } \Omega_{Cu} \\ 1005 & \text{in } \Omega_{air} \\ 465 & \text{in } \Omega_f \end{cases}, \quad \varrho_\Omega = \begin{cases} 2.6989 \cdot 10^3 & \text{in } \Omega \\ 8.94 \cdot 10^3 & \text{in } \Omega_{Cu} \\ 1.293 & \text{in } \Omega_{air} \\ 7.86 \cdot 10^3 & \text{in } \Omega_f \end{cases},$$

where $\mu_0 = \frac{4\pi}{10^7}$. Note that the objective function J in (5.1) depends implicitly on the shape Ω via the state $p = p(\Omega)$. The states $u = u(\Omega)$, $p = p(\Omega)$ depend on Ω via the material parameters. Further we see that the reluctivity ν_Ω of copper and aluminum are the same as for air meaning that the ferromagnetic behavior of these materials are equal.

5.2 Analysis of the state system

In this section we discuss the solvability of (5.2). The key observation is that the system is decoupled meaning that the boundary value problem associated to (5.2a) is independent of the temperature p. Hence, we can show existence and uniqueness of a solution (u, p) to (5.2) in the following way: In a first step we show that there exists a unique solution u to the eddy current problem, that is (5.2a) together with the corresponding boundary, initial and interface conditions. In a second step we use this solution in order to show that there exists a unique solution p to the heat conduction problem, that is (5.2b) together with the corresponding boundary, initial and interface conditions. For the analysis of the eddy current problem we focus on the approach in [2], [3], [4], [19], while for the analysis of the heat conduction problem we focus on the space-time approach by Steinbach [35], [36] and [37].

5.2.1 Preliminaries

In this section we state the theorems which we will need to state existence and uniqueness results of the state system (5.2). The definitions and results are taken from [11, 46].

We start with the definition of an evolution triple.

Definition 5.1 ([46, Def. 23.11])**.** The triple (V, H, V^*) is called evolution triple if

(i) $(V, \|.\|_V)$ is a real, separable and reflexive Banach space,

(ii) $(H, (.,.)_H)$ is a real, separable Hilbert space,

(iii) the embedding $V \hookrightarrow H$ is dense and continuous, i.e. $V \subseteq H$, $\overline{V}^{\|.\|_H} = H$ and there exists a $c > 0$ such that $\|x\|_H \leq c\|x\|_V$ for all $x \in V$.

The following theorem states existence and uniqueness of an abstract evolution equation. We will need this theorem in order to show that the eddy current problem has a unique solution.

Theorem 5.1. *Let (V, H, V^*) be an evolution triple. Let $A : L^2(0,T;V) \to L^2(0,T;V^*)$ be linear. Further assume that A is bounded and coercive, i.e. there exist constants $c_1^A, c_2^A > 0$ such that*

$$\begin{aligned} |\langle Au, v\rangle| &\leq c_2^A \|u\|_{L^2(0,T;V)} \|v\|_{L^2(0,T;V)} \quad \forall u, v \in L^2(0,T;V) \text{ and} \\ \langle Au, u\rangle &\geq c_1^A \|u\|^2_{L^2(0,T;V)} \qquad\qquad \forall u \in L^2(0,T;V), \end{aligned}$$

where $\langle ., . \rangle$ denotes the duality pairing. Suppose furthermore that $F \in L^2(0,T;V^*)$ and $u_0 \in H$. Then the initial value problem

$$\begin{aligned} \partial_t u + Au &= F \quad \text{in } L^2(0,T;V^*), \\ u(0) &= u_0, \end{aligned}$$

has a unique solution $u \in L^2(0,T;V)$ with weak derivative $\partial_t u \in L^2(0,T;V^*)$.

Proof. A proof can be found in [46, Thm. 23.A, Cor. 23.24] □

Next, we state a theorem which we will need to show that the heat conduction problem has a unique solution.

Theorem 5.2 (Banach-Nečas-Babuška)**.** *Let X be a Banach space and let Y be a reflexive Banach space. Let $a \in \mathcal{L}(X \times Y, \mathbb{R})$ and $f \in Y'$. The variational problem:*

$$\text{Find } u \in X : \ a(u,v) = f(v) \quad \forall v \in Y$$

has a unique solution if and only if the following conditions hold:

(i)

$$\exists \alpha > 0 \ \inf_{u \in X} \sup_{v \in Y} \frac{a(u,v)}{\|u\|_X \|v\|_Y} \geq \alpha,$$

(ii)

$$\forall v \in Y, \ (\forall u \in X, \ a(u,v) = 0) \Rightarrow v = 0.$$

Moreover, the following a priori estimate holds:

$$\|u\|_X \leq \frac{1}{\alpha} \|f\|_{Y'}.$$

Proof. A proof can be found in [11, Theorem 2.6] □

5.2.2 Eddy current problem

In this section we discuss the existence and uniqueness of a solution to the initial boundary value problem associated to (5.2a). This initial boundary value problem corresponds to a two dimensional eddy current problem. One feature of eddy current problems in general is that they are different for conducting ($\sigma_\Omega > 0$) and non-conducting ($\sigma_\Omega = 0$) regions. In conducting regions they turn out to be of parabolic type, whereas in non-conducting regions they reduce to an elliptic problem [19]. The strategy for showing existence and uniqueness results is based on the observation that the solution in the non-conducting domain is uniquely determined by the solution in the conducting domain [3]. Hence, one is able to reformulate the variational formulation in terms of the solution to the conducting region and can apply standard tools to show existence and uniqueness of a solution, cf. [19]. We will explain the crucial steps in more detail in the following. We follow the lines in [3, 19] and adapt them to our setting. We introduce the domains

$$\begin{aligned}\Omega_1 &:= \Omega \cup \Omega_f \cup \Omega_{Cu},\\ \Omega_2 &:= \Omega_{air}.\end{aligned}$$

The domain Ω_1 consists of those parts of the computational domain D where the conductivity is greater zero and the domain Ω_2 corresponds to those parts of D where the conductivity is equal to zero. The eddy current problem in more detail reads: Find $u : D \times (0,T) \to \mathbb{R}$ such that

$$\begin{aligned}
\sigma_1 \partial_t u - \operatorname{div}(\nu_1 \nabla u) &= J_3 && \text{in } \Omega_1 \times (0,T), && (5.4a)\\
-\operatorname{div}(\nu_2 \nabla u) &= 0 && \text{in } \Omega_2 \times (0,T), && (5.4b)\\
u &= 0 && \text{on } \partial D \times (0,T), && (5.4c)\\
u(.,0) &= 0 && \text{in } \Omega_1, && (5.4d)\\
[\![u]\!] &= 0 && \text{on } \Gamma_{I,\nu} \times (0,T), && (5.4e)\\
[\![\nu_\Omega \nabla u \cdot n]\!] &= 0 && \text{on } \Gamma_{I,\nu} \times (0,T), && (5.4f)
\end{aligned}$$

where $\sigma_1 := \sigma_\Omega|_{\Omega_1}$, $\nu_i := \nu_\Omega|_{\Omega_i}$ for $i = 1, 2$. Note that the source current J_3 only acts in the coils of the motor, i.e. $\operatorname{supp}(J_3) \subset \Omega_{Cu}$. Therefore we have no source term in Ω_2. Further note that ν_1 is a piecewise constant function while ν_2 is a constant. In the following we will make use of the notation $u_i := u|_{\Omega_i}$ for $i = 1, 2$.

We show that there exists a unique solution to (5.4) in the following way: First, we show that there exists a unique solution u_2 in the non-conducting domain, which depends on u_1. In a second step we exploit this fact to show that there exists a unique solution u_1 in the conducting domain. For this reason we start with the observation that the solution in the conducting domain u_2 is given by the solution of the boundary

value problem: For given $u_1 \in L^2(0,T;H^1(\Omega_1))$ find u_2 such that

$$
\begin{aligned}
-\operatorname{div}(\nu_2 \nabla u_2) &= 0 && \text{in } \Omega_2 \times (0,T), && \text{(5.5a)}\\
u_2 &= 0 && \text{on } \partial D \times (0,T), && \text{(5.5b)}\\
u_2 &= u_1 && \text{on } \Gamma_{I,c} \times (0,T), && \text{(5.5c)}
\end{aligned}
$$

where $\Gamma_{I,c}$ denotes the material interface where the conductivity σ_Ω jumps from $\sigma_\Omega = 0$ to $\sigma_\Omega > 0$. One can see that u_1 together with the condition (5.4c) provide the boundary conditions for the boundary value problem in (5.5). It is well known that (5.5) obtains a unique solution by the lemma of Lax-Milgram. Hence, one can define the operator

$$
\mathcal{H} : L^2(0,T;H^1(\Omega_1)) \to L^2(0,T;H^1(\Omega_2)),\ u_1 \mapsto \mathcal{H}(u_1) = u_2, \tag{5.6}
$$

which maps a given $u_1 \in L^2(0,T;H^1(\Omega_1))$ to the unique solution $u_2 \in L^2(0,T;H^1(\Omega_2))$ of (5.5). Note that this operator is linear and bounded. To be more precise there exists a constant $c_2^{\mathcal{H}} > 0$ such that

$$
\|\mathcal{H}(u_1)\|_{L^2(0,T;H^1(\Omega_2))} \le c_2^{\mathcal{H}} \|u_1\|_{L^2(0,T;H^1(\Omega_1))} \quad \forall u_1 \in L^2(0,T;H^1(\Omega_1)). \tag{5.7}
$$

Estimation (5.7) is a consequence of the a priori estimate from the lemma of Lax-Milgram and the boundedness of the trace operator. Introducing the space

$$
\mathcal{Z} = \left\{ u \in L^2(0,T;H^1(D)) : \begin{array}{l} u_1 \in L^2(0,T;H^1_*(\Omega_1)),\\ u_2 = \mathcal{H}(u_1),\\ u = 0 \text{ on } \partial D \times (0,T) \end{array} \right\}, \tag{5.8}
$$

where

$$
H^1_*(\Omega_1) := \left\{ v \in H^1(\Omega_1) : \int_{\Omega_1} v \,\mathrm{d}x = 0 \right\},
$$

the variational formulation of (5.4) reads: Find $u \in \mathcal{Z}$ with $\partial_t u \in \mathcal{Z}^*$ and $u(.,0) = 0$ in Ω_1 such that

$$
\begin{aligned}
\int_0^T \int_{\Omega_1} \sigma_1 \partial_t u v \,\mathrm{d}x\mathrm{d}t + \int_0^T \int_{\Omega_1} \nu_1 \nabla u \cdot \nabla v \,\mathrm{d}x\mathrm{d}t + \int_0^T \int_{\Omega_2} \nu_2 \nabla u \cdot \nabla v \,\mathrm{d}x\mathrm{d}t \\
= \int_0^T \int_{\Omega_1} J_3 v \,\mathrm{d}x\mathrm{d}t \quad \forall v \in \mathcal{Z}.
\end{aligned} \tag{5.9}
$$

The variational formulation (5.9) can be rewritten in terms of the function u_1 which reads: Find $u_1 \in L^2(0,T;H^1_*(\Omega_1))$ with $\partial_t u_1 \in L^2(0,T;(H^1_*(\Omega_1))^*)$ and $u_1(0) = 0$ such that

$$
\begin{aligned}
\int_0^T \int_{\Omega_1} \sigma_1 \partial_t u_1 v_1 \,\mathrm{d}x\mathrm{d}t + \int_0^T \int_{\Omega_1} \nu_1 \nabla u_1 \cdot \nabla v_1 \,\mathrm{d}x\mathrm{d}t \\
+ \int_0^T \int_{\Omega_2} \nu_2 \nabla \mathcal{H}(u_1) \cdot \nabla \mathcal{H}(v_1) \,\mathrm{d}x\mathrm{d}t = \int_0^T \int_{\Omega_1} J_3 v_1 \,\mathrm{d}x\mathrm{d}t,
\end{aligned} \tag{5.10}
$$

for all $v_1 \in L^2(0,T;H^1_*(\Omega_1))$.

Theorem 5.3. *Let $J_3 \in L^2(0,T;(H^1_*(\Omega_1))^*)$ and let the operator $\mathcal{H}$ be defined as in (5.6). Then there exists a unique solution $u_1 \in L^2(0,T;H^1_*(\Omega_1))$ with $\partial_t u_1 \in L^2(0,T;(H^1_*(\Omega_1))^*)$ to the variational formulation (5.10).*

Proof. Defining the operator $A : L^2(0,T;H^1_*(\Omega_1)) \to L^2(0,T;(H^1_*(\Omega_1))^*)$

$$\langle Au_1, v_1\rangle := \int_0^T \int_{\Omega_1} \nu_1 \nabla u_1 \cdot \nabla v_1 \,\mathrm{d}x\mathrm{d}t + \int_0^T \int_{\Omega_2} \nu_2 \nabla \mathcal{H}(u_1) \cdot \nabla \mathcal{H}(v_1) \,\mathrm{d}x\mathrm{d}t, \tag{5.11}$$

where $\langle .,. \rangle$ denotes the duality pairing and $u_1, v_1 \in L^2(0,T;H^1_*(\Omega_1))$, the variational formulation (5.10) is equivalent to the operator equation

$$\sigma_1 \partial_t u_1 + Au_1 = J_3 \qquad \text{in } L^2(0,T;(H^1_*(\Omega_1))^*), \tag{5.12a}$$

$$u_1(0) = 0. \tag{5.12b}$$

We apply Theorem 5.1 to (5.12). First note that by setting $V = H^1_*(\Omega_1)$, $H = L^2(\Omega_1)$ the triple (V, H, V^*) is an evolution triple. The operator A is linear since $\mathcal{H}$ is linear. Using the Cauchy-Schwarz inequality, estimate (5.7) and that the material parameter ν_Ω is uniformly bounded, i.e.

$$\exists\ \underline{\nu},\ \overline{\nu} > 0 :\ \underline{\nu} \le \nu_\Omega(x,t) \le \overline{\nu} \quad \forall (x,t) \in D \times (0,T)$$

we conclude that for all $u_1, v_1 \in L^2(0,T;H^1_*(\Omega_1))$ there holds

$$\begin{aligned}
|\langle Au_1, v_1\rangle| &\le \bar{\nu}\|\nabla u_1\|_{L^2(Q_{\Omega_1})}\|\nabla v_1\|_{L^2(Q_{\Omega_1})} + \bar{\nu}\|\nabla\mathcal{H}(u_1)\|_{L^2(Q_{\Omega_2})}\|\nabla\mathcal{H}(v_1)\|_{L^2(Q_{\Omega_2})} \\
&\le \bar{\nu}\|u_1\|_{L^2(0,T;H^1(\Omega_1))}\|v_1\|_{L^2(0,T;H^1(\Omega_1))} + \bar{\nu}\|\mathcal{H}u_1\|_{L^2(0,T;H^1(\Omega_2))}\|\mathcal{H}v_1\|_{L^2(0,T;H^1(\Omega_2))} \\
&= \bar{\nu}\left(1 + \left(c_2^{\mathcal{H}}\right)^2\right)\|u_1\|_{L^2(0,T;H^1(\Omega_1))}\|v_1\|_{L^2(0,T;H^1(\Omega_1))},
\end{aligned}$$

where $Q_{\Omega_i} := \Omega_i \times (0,T)$ for $i = 1,2$. Hence, the operator A is bounded. Since $\|\nabla w\|_{L^2(Q_{\Omega_1})}$ defines an equivalent norm in $L^2(0,T;H^1_*(\Omega_1))$ there exists a constant $c_1 > 0$ such that

$$c_1\|w\|_{L^2(0,T;H^1(\Omega_1))} \le \|\nabla w\|_{L^2(Q_{\Omega_1})} \quad \forall w \in L^2(0,T;H^1_*(\Omega_1)). \tag{5.13}$$

Using the boundedness of the material parameter ν_Ω, the estimate $\|\nabla\mathcal{H}(u_1)\|_{L^2(Q_{\Omega_2})} \ge 0$ and estimate (5.13) we conclude

$$\begin{aligned}
\langle Au_1, u_1\rangle &\ge \underline{\nu}\|\nabla u_1\|^2_{L^2(Q_{\Omega_1})} + \underline{\nu}\|\nabla\mathcal{H}(u_1)\|^2_{L^2(Q_{\Omega_2})} \\
&\ge \underline{\nu}\|\nabla u_1\|^2_{L^2(Q_{\Omega_1})} \\
&\ge \underline{\nu}c_1^2\|u_1\|^2_{L^2(0,T;H^1(\Omega_1))},
\end{aligned}$$

which gives coercivity of the operator A. By Theorem 5.1 there exists a unique solution $u_1 \in L^2(0,T;H^1_*(\Omega_1))$ with $\partial_t u_1 \in L^2(0,T;(H^1_*(\Omega_1))^*)$. □

5.2.3 Heat conduction problem

In this section we show the existence and uniqueness of a solution to the initial boundary value problem associated to (5.2b) using the fact that there exists a unique solution u to the eddy current problem. This section is based on [35].

Let $T > 0$ and $Q_D := D \times (0,T)$ be the space time cylinder of the computational domain D. Further let $\Sigma_D := \partial D \times (0,T)$ and $\Sigma_{D0} := \overline{D} \times \{0\}$. We introduce the spaces

$$X := L^2(0,T;H_0^1(D)) \cap H^1(0,T;H^{-1}(D)), \tag{5.14a}$$

$$X_0 := \left\{ v \in X : \ v|_{\Sigma_{D0}} = 0 \right\}, \tag{5.14b}$$

$$Y := L^2(0,T;H_0^1(D)). \tag{5.14c}$$

Since the boundary as well as the initial data on the state p can be seen as Dirichlet data in the space time cylinder (cf. [35]) we choose the ansatz $p = \bar{p} + \bar{p}_0$, where $\bar{p} \in X_0$ and $\bar{p}_0 \in L^2(0,T;H^1(D)) \cap H^1(0,T;H^{-1}(D))$ denotes an arbitrary but fixed extension of the given boundary and initial data. The Petrov-Galerkin variational formulation of (5.2b) is to find $\bar{p} \in X_0$ such that

$$a(\bar{p},q) = F(u,q) - a(\bar{p}_0,q) \quad \forall q \in Y, \tag{5.15}$$

where

$$a(\bar{p},q) := \int_{Q_D} c_{H,\Omega}\varrho_\Omega \partial_t \bar{p} q \,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \lambda_\Omega \nabla_x \bar{p} \cdot \nabla_x q \,\mathrm{d}x\mathrm{d}t, \tag{5.16}$$

$$F(u,q) := \int_{Q_D} -J_3 \partial_t u q \,\mathrm{d}x\mathrm{d}t. \tag{5.17}$$

We are now able to state the main theorem of this section.

Theorem 5.4. *Let the spaces X, X_0, Y be given as in* (5.14). *Let $J_3 \in L^\infty(Q_D)$ and $\bar{p}_0 \in L^2(0,T;H^1(D)) \cap H^1(0,T;H^{-1}(D))$ be some given extension of the initial and boundary data. Suppose furthermore that $u \in \mathcal{Z}$ with $\partial_t \in \mathcal{Z}^*$ is the unique solution to* (5.2a). *Then the variational formulation* (5.15) *admits a unique solution.*

Proof. We use Theorem 5.2 in order to conclude that (5.15) admits a unique solution $\bar{p} \in X_0$. Hence we show that $a(.,.)$ is bounded and fulfills conditions (i) and (ii) in Theorem 5.2.

First note that the coefficients $c_{H,\Omega}$, ϱ_Ω, λ_Ω are positive and uniform bounded. Hence we conclude that

$$\|p\|_{Y,\lambda_\Omega} := \sqrt{\int_0^T \int_D \lambda_\Omega \nabla_x p \cdot \nabla_x p \,\mathrm{d}x\mathrm{d}t} \tag{5.18}$$

defines an equivalent norm in Y and

$$\|p\|_{X,\lambda_\Omega} := \sqrt{\|c_{H,\Omega}\varrho_\Omega \partial_t p\|^2_{L^2(0,T;H^{-1}(D))} + \|p\|^2_{Y,\lambda_\Omega}} \tag{5.19}$$

defines an equivalent norm in X. We make use of the norms (5.19), (5.18) in order to show the boundedness and the stability condition on the bilinear form $a(.,.)$. The bilinear form $a(.,.)$ is continuous since for all $\bar{p} \in X_0$ and $q \in Y$ there holds by using duality and the Cauchy-Schwarz inequality

$$\begin{aligned}
|a(\bar{p},q)| &\leq |\langle c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}, q\rangle| + \left| \int_0^T \int_D \lambda_\Omega \nabla_x \bar{p} \cdot \nabla_x q \,\mathrm{d}x\mathrm{d}t \right| \\
&\leq \|c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\|_{L^2(0,T;H^{-1}(D))} \|q\|_{Y,\lambda_\Omega} + \|\bar{p}\|_{Y,\lambda_\Omega} \|q\|_{Y,\lambda_\Omega} \\
&\leq \sqrt{2} \sqrt{\|c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\|^2_{L^2(0,T;H^{-1}(D))} + \|\bar{p}\|^2_{Y,\lambda_\Omega}} \|q\|_{Y,\lambda_\Omega} \\
&= \sqrt{2} \|\bar{p}\|_{X,\lambda_\Omega} \|q\|_{Y,\lambda_\Omega}.
\end{aligned}$$

The key in showing a stability estimate for the bilinear form a lies in the special choice of an element $w \in Y$ and the particular choice $q = \bar{p} + w$. We start with the trivial estimate

$$\sup_{0 \neq q \in Y} \frac{a(\bar{p},q)}{\|q\|_{Y,\lambda_\Omega}} \geq \frac{a(\bar{p},\bar{p}+w)}{\|\bar{p}+w\|_{Y,\lambda_\Omega}} \tag{5.20}$$

For $\bar{p} \in X_0$ we choose $w \in Y$ as unique solution to the variational formulation

$$\int_{Q_D} \lambda_\Omega \nabla_x w \cdot \nabla_x v \mathrm{d}x\mathrm{d}t = \int_{Q_D} c_{H,\Omega}\varrho_\Omega \partial_t \bar{p} v \,\mathrm{d}x\mathrm{d}t \quad \forall v \in Y. \tag{5.21}$$

By inserting $v = w$ in (5.21) we immediately obtain

$$\begin{aligned}
\|w\|^2_{Y,\lambda_\Omega} &= \int_{Q_D} \lambda_\Omega \nabla_x w \cdot \nabla_x w \,\mathrm{d}x\mathrm{d}t = \int_{Q_D} c_{H,\Omega}\varrho_\Omega \partial_t \bar{p} w \,\mathrm{d}x\mathrm{d}t \\
&\leq \|c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\|_{L^2(0,T;H^{-1}(D))} \|w\|_{Y,\lambda_\Omega}.
\end{aligned}$$

On the other hand we have

$$\begin{aligned}
\|c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\|_{L^2(0,T;H^{-1}(D))} &= \sup_{0 \neq v \in Y} \frac{\langle c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}, v\rangle}{\|v\|_{Y,\lambda_\Omega}} \\
&= \sup_{0 \neq v \in Y} \frac{\langle w, v\rangle_{Y,\lambda_\Omega}}{\|v\|_{Y,\lambda_\Omega}} \leq \|w\|_{Y,\lambda_\Omega}.
\end{aligned}$$

Summing up we get

$$\|w\|_{Y,\lambda_\Omega} = \|c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\|_{Y,\lambda_\Omega}, \tag{5.22a}$$

$$\|c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\|^2_{L^2(0,T;H^{-1}(D))} = \|w\|^2_{Y,\lambda_\Omega} = \int_{Q_D} c_{H,\Omega}\varrho_\Omega \partial_t \bar{p} w \,\mathrm{d}x\mathrm{d}t = \langle c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}, w\rangle. \tag{5.22b}$$

The properties (5.22) on the element $w \in Y$ will help us to conclude the stability estimate as we will see. Considering (5.20) again we can estimate

$$\begin{aligned} a(\bar{p},\bar{p}) &= \int_0^T \int_D c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\bar{p}\,\mathrm{d}x\mathrm{d}t + \int_0^T \int_D \lambda_\Omega \nabla_x \bar{p}\cdot\nabla_x \bar{p}\,\mathrm{d}x\mathrm{d}t \\ &= \frac{1}{2}\int_0^T \int_D c_{H,\Omega}\varrho_\Omega \frac{\mathrm{d}}{\mathrm{d}t}\bar{p}^2\,\mathrm{d}x\mathrm{d}t + \|\bar{p}\|^2_{Y,\lambda_\Omega} \\ &\geq \frac{1}{2}\underline{c}\,\underline{\varrho}\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\|\bar{p}(t)\|^2_{L^2(D)}\,\mathrm{d}t + \|\bar{p}\|^2_{Y,\lambda_\Omega} \\ &= \frac{1}{2}\underline{c}\,\underline{\varrho}\|\bar{p}(T)\|^2_{L^2(D)} + \|\bar{p}\|^2_{Y,\lambda_\Omega} \geq \|\bar{p}\|^2_{Y,\lambda_\Omega}, \end{aligned}$$

where $\underline{c}$ and $\underline{\varrho}$ denote the corresponding lower bounds of $c_{H,\Omega}$ and ϱ_Ω, respectively and we used that $\bar{p}(0) = 0$ for $\bar{p} \in X_0$. Further, using (5.22b), the Cauchy-Schwarz inequality, Young's inequality and (5.22a) we have

$$\begin{aligned} a(\bar{p},w) &= \langle c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}, w\rangle + \int_0^T \int_D \lambda_\Omega \nabla_x \bar{p}\cdot\nabla_x w\,\mathrm{d}x\mathrm{d}t \\ &\geq \|c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\|^2_{L^2(0,T;H^{-1}(D))} - \|\bar{p}\|_{Y,\lambda_\Omega}\|w\|_{Y,\lambda_\Omega} \\ &\geq \|c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\|^2_{L^2(0,T;H^{-1}(D))} - \frac{1}{2}\left(\|\bar{p}\|^2_{Y,\lambda_\Omega} + \|w\|^2_{Y,\lambda_\Omega}\right) \\ &= \frac{1}{2}\|c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\|^2_{L^2(0,T;H^{-1}(D))} - \frac{1}{2}\|\bar{p}\|^2_{Y,\lambda_\Omega}. \end{aligned}$$

Hence we get

$$\begin{aligned} a(\bar{p},\bar{p}+w) &\geq \frac{1}{2}\left(\|c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\|^2_{L^2(0,T;H^{-1}(D))} + \|\bar{p}\|^2_{Y,\lambda_\Omega}\right) \\ &= \frac{1}{2}\|\bar{p}\|^2_{X,\lambda_\Omega} \end{aligned}$$

and together with

$$\begin{aligned} \|\bar{p}+w\|^2_{Y,\lambda_\Omega} &\leq 2\left(\|\bar{p}\|^2_{Y,\lambda_\Omega} + \|w\|^2_{Y,\lambda_\Omega}\right) \\ &= 2\|\bar{p}\|^2_{X,\lambda_\Omega} \end{aligned}$$

we conclude the stability estimate

$$\begin{aligned} \sup_{0\neq q\in Y} \frac{a(\bar{p},q)}{\|q\|_{Y,\lambda_\Omega}} &\geq \frac{a(\bar{p},\bar{p}+w)}{\|\bar{p}+w\|_{Y,\lambda_\Omega}} \\ &\geq \frac{\frac{1}{2}\|\bar{p}\|^2_{X,\lambda_\Omega}}{\sqrt{2}\|\bar{p}\|_{X,\lambda_\Omega}} \\ &= \frac{1}{2\sqrt{2}}\|\bar{p}\|_{X,\lambda_\Omega} \quad \forall \bar{p} \in X_0. \end{aligned} \tag{5.23}$$

In order to show condition (ii) in Theorem 5.2 we show that for all $q \in Y \setminus \{0\}$ there exists a $\bar{p} \in X_0$ such that $a(\bar{p}, q) > 0$. For $q \in Y \setminus \{0\}$ we define

$$\bar{p}(x,t) := \int_0^t q(x,s)\,\mathrm{d}s \quad \text{for } x \in D, t \in [0,T]\,.$$

By definition we have $\bar{p} \in X_0$ and there holds

$$\begin{aligned} a(\bar{p},q) &= \int_0^T \int_D \left[c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}(x,t) q(x,t) + \nabla_x \bar{p}(x,t) \cdot \nabla_x q(x,t)\right] \mathrm{d}x\mathrm{d}t \\ &= \int_0^T \int_D \left[c_{H,\Omega}\varrho_\Omega \left[\partial_t \bar{p}(x,t)\right]^2 + \nabla_x \bar{p}(x,t) \cdot \nabla_x \partial_t \bar{p}(x,t)\right] \mathrm{d}x\mathrm{d}t \\ &= \|c_{H,\Omega}\varrho_\Omega \partial_t \bar{p}\|^2_{L^2(Q_D)} + \frac{1}{2}\|\nabla_x \bar{p}(T)\|^2_{L^2(D)} > 0. \end{aligned} \tag{5.24}$$

Finally, note that the right hand side in (5.15) defines a linear and bounded functional on Y. This follows from $J_3 \in L^\infty(Q_D)$, $\partial_t u \in \mathcal{Z}^*$ and the fact that the bilinear form $a(.,.)$ is bounded for functions $\bar{p}_0 \in L^2(0,T;H^1(D)) \cap H^1(0,T;H^{-1}(D))$. Thus by using Theorem 5.2 the variational formulation (5.15) has a unique solution $\bar{p} \in X_0$. □

5.3 Shape derivative of the objective function

In this section we derive the shape derivative of the objective function via the averaged adjoint method. Let $\Omega \subset \Omega_d$ be the current configuration of aluminum in the design domain. We choose $V \in C_c^1(\Omega_d, \mathbb{R}^2)$ and assume that the vector field V is extended by zero to the computational domain $\overline{D}$. This leads to a transformation $T_\tau : \overline{D} \to \overline{D}$, which only acts on the design domain Ω_d and lets the stator of the motor as well as the air gap unchanged. To be more precise we have for τ small that $\Omega_\tau = T_\tau(\Omega) \subset \Omega_d$, $T_\tau(\Omega_d) = \Omega_d$, $T_\tau(\Omega_{Cu}) = \Omega_{Cu}$, $T_\tau(\Omega_{air}) = \Omega_{air}$, $T_\tau(\Omega_{Fe,s}) = \Omega_{Fe,s}$ and $T_\tau(D) = D$.

In order to compute the shape derivative we have to set up the Lagrangian of problem (5.1)-(5.2). We make use of the spaces

$$\mathscr{Z}_0 := \{u \in \mathcal{Z} : \partial_t u \in \mathcal{Z}^*,\ u(.,0) = 0 \text{ in } \Omega_1\}\,, \tag{5.25}$$

$$\mathscr{Z}^0 := \{u \in \mathcal{Z} : \partial_t u \in \mathcal{Z}^*,\ u(.,T) = 0 \text{ in } \Omega_1\}\,, \tag{5.26}$$

$$\mathcal{X} := L^2(0,T;H^1(D)) \cap H^1(0,T;H^{-1}(D)), \tag{5.27}$$

$$X^0 := \{v \in X :\ v(T) = 0\}\,, \tag{5.28}$$

where $\mathcal{Z}$ is given as in (5.8) and X is given as in (5.14). Recall the variational formulation (5.9), which reads in more compact form: Find $u \in \mathscr{Z}_0$ such that

$$a_1(u,v) = F_1(v) \quad \forall v \in \mathcal{Z}, \tag{5.29}$$

where

$$a_1(u,v) := \int_0^T \int_D \sigma_\Omega \partial_t u v \,\mathrm{d}x\mathrm{d}t + \int_0^T \int_D \nu_\Omega \nabla_x u \cdot \nabla_x v \,\mathrm{d}x\mathrm{d}t,$$
$$F_1(v) := \int_0^T \int_D J_3 v \,\mathrm{d}x\mathrm{d}t.$$

Instead of the variational formulation (5.15) we use the equivalent formulation: Find $p \in \mathcal{X}$ such that

$$a_2(p,q) = F_2(u,q) \quad \forall q \in Y, \tag{5.30a}$$
$$p|_{\Sigma_D} = g, \tag{5.30b}$$
$$p|_{\Sigma_{D0}} = p_0, \tag{5.30c}$$

where a_2 is given as in (5.16) and F_2 is given as in (5.17). The Lagrangian then reads

$$\begin{aligned}\mathcal{L}(\Omega,\varphi,\eta,\psi,\zeta) := J(\eta) + a_1(\varphi,\psi) - F_1(\psi) + a_2(\eta,\zeta) - F_2(\varphi,\zeta) \\ + \int_0^T \int_{\partial D} (\eta - g)\,\zeta \mathrm{d}s\mathrm{d}t + \int_D c_{H,\Omega}\varrho_\Omega\,(\eta(0) - p_0)\,\zeta(0)\mathrm{d}x.\end{aligned} \tag{5.31}$$

The adjoint system can be derived from the equations

$$\mathrm{d}_\varphi \mathcal{L}(\Omega,u,p,w,z)(v) = 0 \quad \forall v \in \mathscr{Z}_0,$$
$$\mathrm{d}_\eta \mathcal{L}(\Omega,u,p,w,z)(q) = 0 \quad \forall q \in \mathcal{X},$$

and reads in strong form

$$-\sigma_\Omega \partial_t w - \operatorname{div}(\nu_\Omega \nabla w) - J_3 \partial_t z = 0 \quad \text{in } Q_D, \tag{5.32a}$$
$$-c_{H,\Omega}\varrho_\Omega \partial_t z - \operatorname{div}(\lambda_\Omega \nabla z) = -2p\chi_{Q_{\Omega_d}} \quad \text{in } Q_D, \tag{5.32b}$$
$$w = 0 \quad \text{on } \Sigma_D, \tag{5.32c}$$
$$z = 0 \quad \text{on } \Sigma_D, \tag{5.32d}$$
$$w(T) = 0 \quad \text{in } D, \tag{5.32e}$$
$$z(T) = 0 \quad \text{in } D, \tag{5.32f}$$

where $Q_{\Omega_d} := \Omega_d \times (0,T)$. Note that the adjoint system is decoupled. Hence we can conclude unique solvability to (5.32) in the same way as for the state system described in Section 5.2. The re-parametrized Lagrangian

$$G : [0,\tilde{\tau}] \times \mathscr{Z}_0 \times \mathcal{X} \times \mathscr{Z}^0 \times X^0 \to \mathbb{R}$$
$$G(\tau,\varphi,\eta,\psi,\zeta) := \mathcal{L}(\Omega_\tau, \varphi \circ T_\tau^{-1}, \eta \circ T_\tau^{-1}, \psi \circ T_\tau^{-1}, \zeta \circ T_\tau^{-1})$$

then reads

$$\begin{aligned} G(\tau,\varphi,\eta,\psi,\zeta) = &\int_{Q_D} \xi(\tau)\,|\eta|^2\,\chi_{Q_{\Omega_d}}\,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \xi(\tau)\sigma_\Omega \partial_t\varphi\psi\,\mathrm{d}x\mathrm{d}t \\ &+ \int_{Q_D} A(\tau)\nu_\Omega \nabla_x\varphi\cdot\nabla_x\psi\,\mathrm{d}x\mathrm{d}t - \int_{Q_D} J_3^\tau\psi\,\mathrm{d}x\mathrm{d}t \\ &+ \int_{Q_D} \xi(\tau)c_{H,\Omega}\varrho_\Omega\partial_t\eta\zeta\,\mathrm{d}x\mathrm{d}t + \int_{Q_D} A(\tau)\lambda_\Omega\nabla_x\eta\cdot\nabla_x\zeta\,\mathrm{d}x\mathrm{d}t \\ &+ \int_{Q_D} \xi(\tau)J_3^\tau\partial_t\varphi\zeta\,\mathrm{d}x\mathrm{d}t + \int_0^T\int_{\partial D} \xi(\tau)\,(\eta-g^\tau)\,\zeta\,\mathrm{d}s\mathrm{d}t \\ &+ \int_D \xi(\tau)c_{H,\Omega}\varrho_\Omega\,(\eta(0)-p_0^\tau)\,\zeta(0)\,\mathrm{d}x, \end{aligned} \tag{5.33}$$

where $\xi(\tau) = |\det(\partial T_\tau)| = \det(\partial T_\tau)$ for τ small and $A(\tau) = \xi(\tau)\partial T_\tau^{-1}\partial T_\tau^{-T}$. In order to compute the shape derivative we make use of Theorem 4.6. The hypotheses (D0),(D1) can be shown analogously as described in Section 4.2. Assumption (D2) is satisfied since the state system, the adjoint system and the averaged adjoint system are uniquely solvable. The verification of (D3) can be done with similar techniques as shown in Section 4.2. We apply Theorem 4.6 and conclude that $\mathrm{d}J(\Omega;V) = \partial_\tau G(0,u,p,w,z)$, where $(u,p)\in\mathscr{Z}_0\times\mathcal{X}$ solves the state system (5.2) and $(w,z)\in\mathscr{Z}^0\times X^0$ solves the adjoint system. For computing the shape derivative note that $\operatorname{supp}(V)\subset\Omega_d$, $\operatorname{supp}(J_3)\subset\Omega_{Cu}$ and that $\Omega_d\cap\Omega_{Cu}=\emptyset$. Further notice that in our case the functions g and p_0 are constants. Therefore their derivatives vanish and the shape derivative reads

$$\begin{aligned} \mathrm{d}J(\Omega;V) = &\int_{Q_D} \operatorname{div}V\,|p|^2\,\chi_{Q_{\Omega_d}}\,\mathrm{d}x\mathrm{d}t + \int_{Q_D} \operatorname{div}V\sigma_\Omega\partial_t uw\,\mathrm{d}x\mathrm{d}t \\ &+ \int_{Q_D} \left(\operatorname{div}V - \partial V - \partial V^T\right)\nu_\Omega\nabla_x u\cdot\nabla_x w\,\mathrm{d}x\mathrm{d}t \\ &+ \int_{Q_D} \operatorname{div}V c_{H,\Omega}\varrho_\Omega\partial_t pz\,\mathrm{d}x\mathrm{d}t \\ &+ \int_{Q_D} \left(\operatorname{div}V - \partial V - \partial V^T\right)\lambda_\Omega\nabla_x p\cdot\nabla_x z\,\mathrm{d}x\mathrm{d}t. \end{aligned} \tag{5.34}$$

5.4 Numerical results

In this section we apply a gradient based shape optimization algorithm to problem (5.1)-(5.2) and present the numerical results. The implementation of the algorithm is done in the finite element software Netgen/NGSolve [31].

We want to mention that we do not use any parallel implementation. Thus we are restricted in the size of the space-time geometric model in order to keep the dimensions of the resulting linear system low. For this reason we only simulate the PDEs for $T = 0.02$ seconds in time and neglect the layer of air outside the stator in the computational domain D in our simulations. Due to the short simulation time we induce a very high current density J_3 in order to see a deformation of the shape of the aluminum bars in the rotor.

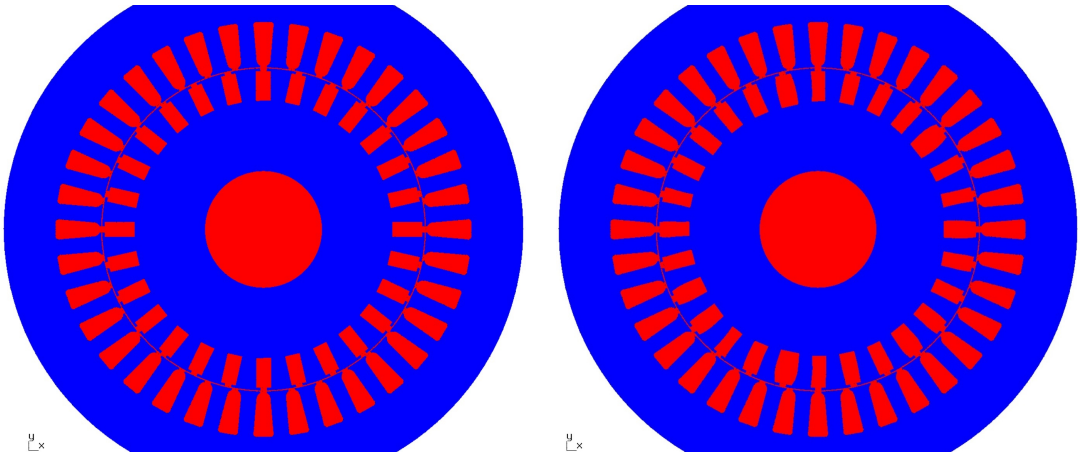
In our numerical test we use the initial design depicted in Figure 5.2(a). We choose a rotor design in which the rotor consists of 28 aluminum bars in rectangular shape. The numerical algorithm we apply in the test is the same as described in Algorithm 4.7. In the algorithm one has to use for the state system (5.2), for the adjoint system (5.32) and for the shape derivative (5.34). The computation of a descent direction is done via an auxiliary boundary value problem which reads: Find $V \in H_0^1(\Omega_d, \mathbb{R}^2)$ such that

$$\int_{\Omega_d} \partial_x V : \partial_x W + V \cdot W = -\mathrm{d}\mathcal{J}(\Omega; W) \quad \forall W \in H_0^1(\Omega_d, \mathbb{R}^2). \tag{5.35}$$

The numerical results can be seen in Table 5.1 and Figure 5.2(b) for a constant step size of $\tau = 0.001$ in every iteration. Table 5.1 shows the evolution of the objective value for 9 iterations. We observe an initial value of 6.729 which decreases to a value of 6.648 in iteration 9. Further we see that the decrease of the objective value is very small. This is maybe due to the short period of the simulation time. Figure 5.2(b) shows the design of the rotor after 9 iterations. We observe that the change of the shape of the aluminum bars is not uniform in the rotor. Some bars tend to get bigger while others tend to get smaller. Still others seem to not change their shape over the optimization process. Moreover, we have to say that the change of the shape of most aluminum bars is small. This observation correlates with the small decrease of the objective value.

iteration	$\mathcal{J}$
0	6.729
1	6.718
2	6.708
3	6.698
4	6.688
5	6.680
6	6.671
7	6.663
8	6.655
9	6.648

Table 5.1: Values of the objective function $\mathcal{J}$



(a) (b)

Figure 5.2: (a) Initial design of the rotor, (b) Design of the rotor after 9 iterations, blue: ferromagnetic material, red: non-ferromagnetic material

Conclusion

In this thesis we considered three parabolic shape optimization problems: (i) a one-phase problem, (ii) an interface problem, (iii) a linear shape optimization problem for an electric motor. We solved each problem by means of shape optimization techniques based on sensitivity information. The underlying PDEs were solved via a conforming space-time finite element method.

First, we concentrated on the parabolic PDE systems which occurred in problem (i) and (ii). We applied a conforming space-time method to both systems and determined the experimental order of convergence for the L^2-error as well as energy error for piecewise linear finite elements. Numerical tests showed for (i) and (ii) a quadratic convergence behavior in the L^2-norm and a linear convergence rate in the energy norm. This coincides with the well-known result for elliptic problems.

Second, we focused on the essential parts in the theory of shape optimization used to set up an algorithm. We introduced the shape derivative and discussed its computation for PDE-constrained problems within a Lagrangian framework. Further, we explained how to exploit the shape derivative to set up gradient-type algorithms.

Next, we considered the shape optimization problems (i) and (ii). We formally computed the shape derivatives for both problems and pointed out issues when proving them in a rigorous way. We embedded these derivatives in a gradient-type algorithm using space-time finite elements. Numerical results showed the correctness of the proposed algorithm.

Finally, we applied the proposed algorithm to (iii). In order to run the algorithm we had to use a small simulation time since a parallel implementation was not used. Due to the small simulation time only minor changes in the design of the rotor could be seen in the numerical test. Nevertheless, the results demonstrated that the algorithm also might work in the framework of an electric motor.

Open questions and future work

The work presented in this thesis can be extended in the following directions:

A detailed analysis of the coupled parabolic system presented in Chapter 2 is an open question. It would be interesting to see if the existence of a unique solution to this coupled system could be shown by similar techniques as in [35] or if another approach like the theory of compact perturbations [27, Chapter 8] would be necessary.

Assuming that one is able to prove that the state system has a unique solution the next step would be to prove the shape derivatives of (i) and (ii) in a rigorous way. Here a further crucial point lies in showing the function space parametrization, see Remark 4.2. We want to mention that this property was already used by Sokołowski in [33].

Problem (iii) demonstrated that a parallel implementation of the space-time method is mandatory in order to solve parabolic shape optimization problems in the context of electric motors. It would be interesting to see how the design of the rotor in (iii) would change if a bigger simulation interval and a parallel implementation of the algorithm was used.

We only considered a linear shape optimization problem for an electric motor. It would be more realistic to consider the nonlinear shape optimization problem given in (0.2). Here, one problem from the modeling point of view is to find the correct relation between the electrical conductivity σ and the temperature ϑ. In general this mapping is not given analytically and one has to approximate it from measured data. A further challenging task will be the analysis of the nonlinear state system. The discretization of the nonlinear state system with a conforming space-time method will lead to a system of nonlinear equations which could be solved by Newton's method.

Bibliography

[1] W. Arendt and K. Urban. *Partielle Differenzialgleichungen, Eine Einführung in analytische und numerische Methoden*. Spektrum Akademischer Verlag Heidelberg, 2010.

[2] F. Bachinger. Multigrid solvers for 3D multiharmonic nonlinear magnetic field computations. Master's thesis, Johannes Kepler University Linz, 2003.

[3] F. Bachinger, U. Langer, and J. Schöberl. Numerical analysis of nonlinear multiharmonic eddy current problems. Technical Report 1, Johannes Kepler University Linz, SFB "Numerical and Symbolic Scientific Computing", 2004.

[4] F. Bachinger, U. Langer, and J. Schöberl. Numerical analysis of nonlinear multiharmonic eddy current problems. *Numerische Mathematik*, 100(4):593–616, 2005.

[5] A. Binder. *Elektrische Maschinen und Antriebe: Grundlagen, Betriebsverhalten*. Springer, 2017.

[6] O. Bíró. *Numerische Aspekte von Potentialformulierungen in der Elektrodynamik, Habilitationsschrift*. TU Graz, 1992.

[7] M. C. Delfour and J.-P. Zolésio. *Shapes and Geometries*, volume 22 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2011. Metrics, analysis, differential calculus, and optimization.

[8] G. Dziuk. *Theorie und Numerik partieller Differentialgleichungen*. Walter de Gruyter, Berlin, 2010.

[9] M. Eigel and K. Sturm. Reproducing kernel Hilbert spaces and variable metric algorithms in PDE-constrained shape optimization. *Optimization Methods and Software*, 33(2):268–296, 2018.

[10] S. El Yacoubi and J. Sokołowski. Domain optimization problems for parabolic control systems. *Applied Mathematics and Computer Science*, 6:277–290, 1996.

[11] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*, volume 159. Springer, 2004.

[12] P. Gangl. *Sensitivity-Based Topology and Shape Optimization with Application to Electrical Machines*. PhD thesis, Johannes Kepler University Linz, 2016.

[13] P. Gangl. Lecture on Shape and Topology Optimization, 2019. Lecturenotes.

[14] P. Gangl, U. Langer, A. Laurain, H. Meftahi, and K. Sturm. Shape optimization of an electric motor subject to nonlinear magnetostatics. *SIAM Journal on Scientific Computing*, 37(6):B1002–B1025, 2015.

[15] H. Harbrecht and J. Tausch. On the numerical solution of a shape optimization problem for the heat equation. *SIAM Journal on Scientific Computing*, 35(1):A104–A121, 2013.

[16] M. Jung and U. Langer. *Methode der finiten Elemente für Ingenieure, Eine Einführung in die numerischen Grundlagen und Computersimulation*. Springer, 2013.

[17] D. Kalise, K. Kunisch, and K. Sturm. Optimal actuator design based on shape calculus. *Mathematical Models and Methods in Applied Sciences*, 28(13):2667–2717, 2018.

[18] C. Kaufmann, M. Günther, D. Klagges, M. Knorrenschild, M. Richwin, S. Schöps, and E. J. Ter Maten. Efficient frequency-transient co-simulation of coupled heat-electromagnetic problems. *Journal of Mathematics in Industry*, 4(1):1, 2014.

[19] M. Kolmbauer. Existence and uniqueness of eddy current problems in bounded and unbounded domains. Technical Report 06, Johannes Kepler Universtity, Doctoral Program "Computational Mathematics", 2011.

[20] U. Langer, S. E. Moore, and M. Neumüller. Space-time isogeometric analysis of parabolic evolution problems. *Computer Methods in Applied Mechanics and Engineering*, 306:342–363, 2016.

[21] A. Laurain and K. Sturm. Distributed shape derivative via averaged adjoint method and applications. *ESAIM: Mathematical Modelling and Numerical Analysis*, 50(4):1241–1267, 2016.

[22] M. Neumüller. Space-Time Methods. Fast Solvers and Applications. *Monographic Series TU Graz: Computation in Engineering and Science*, 20, 2013.

[23] M. Neumüller and O. Steinbach. Refinement of flexible space–time finite element meshes and discontinuous Galerkin methods. *Computing and Visualization in Science*, 14:189–205, 2011.

[24] C. Pechstein. Multigrid-Newton-methods for nonlinear magnetostatic problems. Master's thesis, Johannes Kepler University Linz, 2004.

[25] H.-G. Raumer. Shape Optimization for Interface Problems using unfitted Finite Elements. Master's thesis, University of Göttingen, 2018.

[26] M. Růžička. *Nichtlineare Funktionalanalysis*. Springer, 2004.

[27] F. J. Sayas, T. S. Brown, and M. E. Hassell. *Variational Techniques for Elliptic Partial Differential Equations: Theoretical Tools and Advanced Applications*. CRC Press, 2019.

[28] A. Schafelner. Space-time finite element methods for parabolic initial-boundary problems. Master's thesis, Johannes Kepler University Linz, 2017.

[29] F. Schmitz, P. Nägele, and J. Daube. Bochner-Räume. `https://aam.uni-freiburg.de/agru/lehre/ws15/bochner/bochner.pdf`, 2016. [Online; accessed 26-May-2020].

[30] R. Schneckenleitner. Isogeometrical Analysis based Shape Optimization. Master's thesis, Johannes Kepler University Linz, 2017.

[31] J. Schöberl. C++11 Implementation of Finite Elements in NGSolve. Technical Report 30, Institute for Analysis and Scientific Computing, Vienna University of Technology, 2014.

[32] V. H. Schulz, M. Siebenborn, and K. Welker. Structured inverse modeling in parabolic diffusion problems. *SIAM Journal on Control and Optimization*, 53(6):3319–3338, 2015.

[33] J. Sokołowski. Shape sensitivity analysis of boundary optimal control problems for parabolic systems. *SIAM Journal on Control and Optimization*, 26(4):763–787, 1988.

[34] J. Sokołowski and J.-P. Zolésio. *Introduction to Shape Optimization. Shape Sensitivity Analysis*, volume 16. Springer-Verlag, 1992.

[35] O. Steinbach. Space-time finite element methods for parabolic problems. *Computational Methods in Applied Mathematics*, 15(4):551–566, 2015.

[36] O. Steinbach and H. Yang. Space-time finite element methods for parabolic evolution equations: discretization, a posteriori error estimation, adaptivity and solution. *Berichte aus dem Institut für Angewandte Mathematik, Bericht*, 9, 2018.

[37] O. Steinbach and M. Zank. Coercive space-time finite element methods for initial boundary value problems. *ETNA - Electronic Transactions on Numerical Analysis*, 52:154–194, 01 2020.

[38] K. Sturm. Minimax Lagrangian approach to the differentiability of nonlinear PDE constrained shape functions without saddle point assumption. *SIAM Journal on Control and Optimization*, 53(4):2017–2039, 2015.

[39] K. Sturm. *On shape optimization with non-linear partial differential equations.* PhD thesis, TU Berlin, 2015.

[40] I. Toulopoulos. Space-time finite element methods stabilized using bubble function spaces. *Applicable Analysis*, 99(7):1153–1170, 2020.

[41] Wikipedia contributors. Induction motor — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Induction_motor&oldid=924768711`, 2019. [Online; accessed 7-November-2019].

[42] Wikipedia contributors. Lenz's law — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=Lenz%27s_law&oldid=924899471`, 2019. [Online; accessed 7-November-2019].

[43] Wikipedia contributors. Electric machine — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/wiki/Electric_machine`, 2020. [Online; accessed 28-May-2020].

[44] Wikipedia contributors. Squirrel-cage rotor — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/wiki/Squirrel-cage_rotor`, 2020. [Online; accessed 02-June-2020].

[45] S. Zaglmayr. *High Order Finite Element Methods for Electromagnetic Field Computation*. PhD thesis, Johannes Kepler University Linz, 2006.

[46] E. Zeidler. *Nonlinear Functional Analysis and Its Applications II/A: Linear Monotone Operators*. Springer-Verlag New York, 1990.

[47] W. P. Ziemer. *Weakly differentiable functions: Sobolev spaces and functions of bounded variation*, volume 120. Springer, New York, 1989.

Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitely indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Date	Signature