Lorenz Johannes John

Optimal Boundary Control in Energy Spaces

Monographic Series TU Graz

Computation in Engineering and Science

Series Editors

Monographic Series TU Graz

Computation in Engineering and Science Volume 24

Lorenz Johannes John

Optimal Boundary Control in Energy Spaces

Preconditioning and Applications

This work is based on the dissertation *"Optimal boundary control in energy spaces. Preconditioning and applications"*, presented by Lorenz Johannes John at Graz University of Technology, Institute of Computational Mathematics in March 2014.
Supervisor: O. Steinbach (Graz University of Technology)
Reviewer: R. H. W. Hoppe (University of Augsburg), K. Kunisch (University of Graz)

Cover photo	Vier-Spezies-Rechenmaschine by courtesy of the Gottfried Wilhelm Leibniz Bibliothek – Niedersächsische Landesbibliothek Hannover
Layout	Wolfgang Karl, TU Graz / Universitätsbibliothek Christina Fraueneder, TU Graz / Büro des Rektorates
Printed	by TU Graz / Büroservice

Verlag der Technischen Universität Graz

www.ub.tugraz.at/Verlag

Print:

ISBN: 978-3-85125-373-3

E-Book:

ISBN: 978-3-85125-374-0

DOI: 10.3217/978-3-85125-373-3

Abstract

In this thesis we study optimal boundary control problems in energy spaces, their construction of robust preconditioners and applications to arterial blood flow.

More precisely we first consider the unconstrained optimal Dirichlet and Neumann boundary control problems for the Poisson equation as a model problem. In both cases it turns out that the control can be eliminated and thus a variational formulation in saddle point structure is obtained. The existence and uniqueness of a solution is investigated and for the finite element discretization optimal error estimates are shown. In the particular case of the Laplace equation as a constraint we are able to prove that the primal states of Dirichlet and Neumann boundary control problem coincide.

Further, the construction of corresponding robust preconditioners for optimal boundary control problems is investigated. We observe that the optimal boundary control problems are related to biharmonic equation of first kind. For the preconditioner we consider either a preconditioner motivated from boundary element methods, resulting in an optimal condition number, or a multilevel preconditioner of BPX type, where the condition number depends on a logarithmic factor of the mesh size. For both, the related spectral equivalent estimates are proven. Several numerical examples illustrate the obtained theoretical results.

Moreover, we study the application of the optimal Dirichlet boundary control problem to arterial blood flow. In particular, we are interested in the optimal inflow profile into an arterial system, motivated for instance by an artificial heart pump. Also, we investigate on hemodynamic indicators, for showing potential risk factors for aneurysms. Here a comparison of two commonly used approaches is considered, where it is shown by several numerical simulations that these can lead to significant differences in the solution. This model problem motivates also the optimization of hemodynamic indicators. Finally, several numerical examples are presented.

Zusammenfassung

In der vorliegenden Dissertation werden optimale Steuerungsprobleme mit Randkontrolle in Energieräumen behandelt, sowie die Konstruktion von zugehörigen robusten Vorkonditionierern. Eine Anwendung ist die numerische Simulation von Blutströmungen, auf die in weiterer Folge eingegangen wird.

Wir betrachten zunächst das optimale Steuerungsproblem für die Poisson Gleichung mit Dirichlet und Neumann Randkontrolle. In beiden Fällen stellt sich heraus, dass die Kontrolle eliminiert und die Variationsformulierung somit in klassischer Sattelpunkt Struktur dargestellt werden kann. Es wird die Existenz und Eindeutigkeit einer Lösung untersucht und für eine Finite Elemente Diskretisierung optimale Fehlerabschätzungen bewiesen. Wird die Nebenbedingung des optimalen Steuerungsproblems durch die Laplace Gleichung formuliert, so stellt sich heraus, dass die primalen Zustände von Dirichlet und Neumann Steuerungsproblem übereinstimmen.

Des Weiteren wird ein robuster Vorkonditionierer für das optimale Steuerungsproblem mit Randkontrolle konstruiert. Hier stellt sich heraus, dass das optimale Steuerungsproblem in Beziehung zur biharmonischen Gleichung erster Art steht. Der Vorkonditioner wird entweder durch einen diskreten Operator aus der Randelementmethode realisiert, welcher eine optimale Konditionszahl aufweist, oder durch einen Multilevel Vorkonditionierer vom BPX–Typ mit logarithmischer Abhängigkeit der Konditionszahl von der Maschenweite. Zahlreiche numerische Beispiele illustrieren die erhaltenen theoretischen Resultate.

Der Anwendungsteil dieser Arbeit diskutiert den Einsatz des optimalen Steuerungsproblems mit Dirichlet Randkontrolle in der Simulation von Blutströmungen in Arterien. Im Speziellen sind wir an der Konstruktion von optimalen Einströmrändern interessiert, welche zum Beispiel beim Einsatz von künstlichen Herzpumpen relevant sind. Des Weiteren werden hämodynamische Indikatoren untersucht, welche potenzielle Risikobereiche für Aneurysmen darstellen. Hier werden zwei verschieden Formulierungen der Indikatoren untersucht, welche zu signifikanten Unterschieden in der Simulation führen können. Zahlreiche numerische Beispiele werden präsentiert.

CONTENTS

Introduction

Mathematical models, computational methods and their analysis are an essential part in the simulation of problems arising in physics, biology, chemistry, etc. One of the possible ways to describe such problems is their formulation in the framework of continuum mechanics. Here, most of the models are formulated by partial differential equations, or by a system of coupled partial differential equations. Beside many real-world problems, let us mention for instance arterial blood flow, on which we focus as an application in this work, see, e.g., [23, 28, 78]. Nowadays, cardiovascular diseases, such as the rupture of aneurysm and strokes, are widely spread. It is well known that the size or shape of an aneurysm is in general not an indicator for a potential risk factor. It is rather the shear stress acting on the arterial wall, who can lead to a predictive diagnosis. Due to these reasons, the so-called hemodynamic indicators have been derived, which shall give an indication for a potential rupture of an aneurysm. Further, it is well known that the vortex formation of the blood can cause blood clots, so-called thrombi, which can consequently block the artery and may lead to a stroke. It is now interesting to ask how such vortex formations can be reduced or minimized. This motivates to consider optimal control problems for arterial blood flow. In particular we are interested in the inflow control, for instance by an artificial heart pump, into a bypass, with respect to vortex minimization.

In order to simulate arterial blood flow and optimize certain quantities a realistic model, which describes the flow, is needed. Even though blood is a mixture of different biological substances it is mostly described as a single constituent fluid. A model which is known to be a good approximation, especially for the case of large arteries, is the system of the Navier–Stokes equations. In general, the analytical solution of this system is not known, which motivates numerical approximation schemes such as the finite element method. We shall present a stabilized finite element method of lowest order for solving the Navier–Stokes equations.

For the simulation of these model problems it is important that the computations can be done in a reasonable time. By this we mean that efficient solution techniques for the corresponding algebraic systems are needed. This motivates iterative solution strategies with preconditioning. The crucial point is the construction of a robust preconditioner, where the corresponding inverse can be efficiently realized. For this purpose multilevel and multigrid methods, see, e.g., [10, 34], have been developed. But the robustness of the preconditioner with respect to the discretization is for numerous problems not sufficient, other parameters, such as the viscosity constant or the cost coefficient of the optimal control problem have to be additionally taken into account. This is one of the main advantages of multilevel preconditioners of BPX type, since it is additive, and thus these

parameters can be rather easily handled. We refer for a general overview on preconditioners to [10, 11, 22, 34, 73].

The notation of this thesis is strongly based on the book [73]. In the case that additional information is needed, we refer the reader to this book. Further, we assume the basic concepts and results for the numerical analysis of partial differential equations, for an overview we refer to [11, 12, 30, 64, 73]. As we will see, especially for the numerical analysis of preconditioners, the consideration of certain Sobolev spaces is needed. Therefore we refer for a general overview on Sobolev spaces, to [1, 30, 33] and in particular for fractional Sobolev spaces on the boundary to [38, 53, 73].

Structure and summary of the main results

In the following we discuss the structure of this work and summarize the main results of the individual chapters. In principle this work investigates optimal boundary control problems in energy spaces, where the related Dirichlet and Neumann boundary control problem are considered. In the case of the Poisson equation as constraint a unified analysis and numerical analysis is presented and the relation of both problems are discussed. Further the construction of robust preconditioners is studied. Here it turns out that the problem is related to the biharmonic equation of first kind. Thus we study the preconditioning of the biharmonic equation and apply the obtained ideas afterwards to the optimal control problem. We shall present corresponding spectral equivalence estimates and numerical examples. In the third part of the thesis, applications to arterial blood flow are discussed. Here we apply the Dirichlet boundary control approach in the energy space to a blood flow related model problem. Further, the simulation of hemodynamic indicators is discussed. For both, we present several numerical examples.

In the following we shall summarize the content of this work and give an overview on the main results.

Chapter 1: Modeling and discretization

In the first chapter we recall the constitutive equations of fluid dynamics and introduce two models for the description of arterial blood flow. These differ in the form of the extra stress tensor and have either a constant or generalized viscosity of Carreau type. Further, we discuss the non-dimensionalization of the equations and the meaning of corresponding boundary and initial conditions. Finally we present the (generalized) Navier–Stokes equations as a model problem. Moreover we present an overview on the existence and uniqueness results for the solution of the Navier–Stokes equations.

In the second part of the chapter we consider a finite element discretization of the (generalized) Navier–Stokes equations. In particular we investigate a stabilized finite element

method, which is free of any stabilization parameter. The advantages of such an approach shall be discussed.

Chapter 2: Hemodynamics for arterial blood flow

In the second chapter we consider the simulation of arterial blood flow in hemodynamics. In particular we consider hemodynamic indicators, such as the wall shear stress and the oscillatory shear index. It turns out that the calculation of these indicators is not always done by the same formula. This means, from the boundary stress vector, from which the indicators are calculated, an important term is often neglected. We present by several numerical simulations, that this can lead to a significant difference in the hemodynamic indicators.

Chapter 3: Optimal boundary control problems in energy spaces

This chapter, on optimal boundary control problems in energy spaces, is divided into two parts. First, we consider optimal Dirichlet boundary control problems, where the constraint is given by the Poisson equation as a model problem and the control is considered in the energy space. We derive the first order necessary optimality conditions. It turns out that the control can be eliminated and thus a variational formulation in classical saddle point structure is obtained. Consequently, we can apply standard results for these systems and prove the existence and uniqueness of a solution. It turns out that the optimal control problem is related to the biharmonic equation of first kind. Finally, a lowest order discretization is introduced and corresponding error estimates are given. These are illustrated by several numerical examples.

In the second part of this chapter we apply the idea of the control in the energy space to the related Neumann boundary control problem. We start with the Yukawa equation as a constraint, which is easier to treat since the boundary value problem is always uniquely solvable. Following the ideas of the first part we derive the corresponding first order necessary optimality conditions. As second case, we consider as a constraint the Poisson equation, which is more involved, since we have to introduce a certain scaling condition. It turns out that this is nevertheless a special case of the Yukawa equation. Further, the existence and uniqueness of the solution of the corresponding optimality system is investigated, as well as a mixed finite element discretization. Here, we observe that the Schur complement equation with respect to the control is for the Neumann boundary control of the same structure as in the Dirichlet boundary control case. Further, we prove that the primal states of Dirichlet and Neumann boundary control problems coincide in the case of the Laplace equation as a constraint.

Chapter 4: An optimal control problem for arterial blood flow

As it was motivated in the outline, the minimization of vortices in the aneurysm is an important task. Within this chapter we consider the optimal control of the inflow velocity (inflow profile) into a bypass of an arterial system. Such problems can be solved by considering a Dirichlet boundary control problem, where the constraint is in our case the Navier–Stokes system with a constant viscosity for the description of the blood. In particular, we apply the energy control approach for this model problem. We present numerical examples and show that this approach leads to numerical solutions of physical relevance.

Chapter 5: The biharmonic equation

In this chapter, we start with the derivation of a mixed finite element formulation for the biharmonic equation of first kind. It is well known that this formulation is advantageous, since we are able to consider less regular solution spaces and thus can use a standard finite element discretization. The existence and uniqueness of a solution is investigated and corresponding error estimates are discussed. At the end several numerical examples are presented.

Chapter 6: Preconditioning strategies for the biharmonic equation

We consider the construction of a robust preconditioner for the mixed finite element discretization of the biharmonic equation of first kind, introduced in Chapter 5. The preconditioner for the Schur complement equation with respect to the boundary is treated first. Therefore certain fractional Sobolev spaces on the boundary have to be introduced. In particular, we define Sobolev spaces $\widetilde{H}^{1/2}_{\mathrm{pw}}(\Gamma)$ on open parts of the boundary. We prove for the Schur complement equation corresponding spectral equivalence estimates within the fractional Sobolev space $\widetilde{H}^{1/2}_{\mathrm{pw}}(\Gamma)$ and present two possible discrete operators for the realization of this norm. The first, a local single layer boundary integral operator, motivated from boundary element methods, and the second one a multilevel representation of BPX type. Several numerical examples illustrate the obtained theoretical results. Further, two preconditioners for the global system are proposed. Again, numerical results show the effectiveness of these preconditioners.

Chapter 7: Preconditioning strategies for optimal boundary control problems

In this chapter we apply the ideas of the preconditioners as discussed in Chapter 6 to the optimal boundary control problems, analyzed in Chapter 3. We prove the spectral equivalence of the Schur complement system with respect to the control on the boundary. The preconditioner is then either realized by a boundary element approach, via a combination

of the local single layer boundary integral operator and the hypersingular boundary integral operator, or by a multilevel representation of BPX type. Several numerical examples illustrate the obtained theoretical results.

At the end, we present some concluding remarks and several interesting open problems as a possible future work.

1 MODELING AND DISCRETIZATION

In this chapter we introduce the systems of partial differential equations and the corresponding finite element discretization which are studied in this thesis. By this we mean the constitutive equations of fluid dynamics, describing for example blood flow. First, we give a short introduction and review of hemodynamic models for arterial blood flow. These models are given by the form of the so-called Cauchy stress tensor which, in this work, has always an explicit form. More precisely, we consider models where the extra stress part of the Cauchy stress tensor is either linear or non-linear (of polynomial growth) with respect to the deformation. In the case that the function of proportionality (viscosity) is constant, the resulting equations are the Navier–Stokes equations. In the second case they are the generalized Navier–Stokes equations. Moreover we present an overview on the existence and uniqueness results for the solution of the Navier–Stokes equations. As a last part of this chapter we study a stabilized finite element method for the numerical solution of these equations. The advantages of such an approach shall be discussed.

Even though blood is a mixture of biological substances, namely blood cells suspended in a blood plasma medium consisting of water, macromolecules, ions, etc., we consider it, on the macroscopic scale, as a single constituent incompressible, homogeneous and isotropic fluid. Thus, we shall describe its flow in the framework of continuum mechanics.

Let us consider a bounded Lipschitz domain $\Omega \subset \mathbb{R}^3$, which represents for instance the artery of interest, and a time interval $(0,\bar{t})$ with $\bar{t} > 0$. We describe the flow of the blood in terms of the velocity field $\underline{u}(t,x)$ and the pressure $p(t,x)$, which we denote from now on in short by $\underline{u}$ and p. Since we assume the fluid to be homogeneous and incompressible the density ρ is constant in space and time, respectively. We may also consider a bulk force $\underline{f}$ which is acting on the fluid, for instance the gravitation.

The flow of the fluid is then described by the balance of momentum

$$\partial_t(\rho\underline{u}) - \operatorname{div} T + \operatorname{div}(\rho\underline{u}\otimes\underline{u}) = \rho\underline{f} \qquad \text{in } \Omega\times(0,\bar{t}), \tag{1.1}$$

and the conservation of mass

$$\operatorname{div}(\rho\underline{u}) = 0 \qquad \text{in } \Omega\times(0,\bar{t}), \tag{1.2}$$

where $(\underline{u}\otimes\underline{v})_{ij} = u_i v_j$, $i,j = 1,\dots,3$, denotes the tensor product. In the particular case as considered above, the fluid model is given by the Cauchy stress tensor

$$T = -pI + S, \tag{1.3}$$

where pI is the mean normal stress, with p denoting the hydrodynamical pressure, and S is the extra stress tensor which needs to be specified by a suitable constitutive equation reflecting the rheological nature of the considered fluid.

For the needs of computational simplicity, blood is very commonly considered as a Newtonian fluid, that means, its rheological behavior is described by a single parameter, called viscosity, being a constant of proportionality between the shear stress and the shear rate during a simple shear. Such an approximation can be validated for blood flow in vessels with large diameters. This, on the other hand, should not be presumed in the case of blood flow in a vessel with aneurysm where its typical non-Newtonian phenomena occur. In the following we describe these different models.

1.1 Hemodynamic models

In a physiological environment, the non-Newtonian character of blood manifests mainly in its ability to thin the shear and the stress relaxation. The shear-thinning behavior and its connection to the red blood cell deformation and rouleau aggregation was originally recognized already in the 1970's in [14,15]. Shortly after, in [79] the property of red blood cells is described to store energy via the rouleau network deformation and consequently measured the viscoelastic nature of blood. Such a behavior, related to the rouleau network deformation, must be thus shear-rate dependent as it is the formation of such a structure, see [80]. Until now, only few viscoelastic models (describing among others the mentioned stress relaxation) have been proposed: a linear Maxwell model proposed in [79], a generalized Oldroyd-B model with a non-linear apparent viscosity of shear-thinning proposed in [84], and a thermodynamically consistent model in [2] describing blood as a mixture of shear-thinning viscoelastic and Newtonian fluids, created in the framework of maximization of the rate of dissipation corresponding to the material natural (stress-free) configuration.

For our aim, we consider first a standard Newtonian model, where the extra stress tensor is given by

$$S = 2\mu D, \tag{1.4}$$

and D denotes the symmetric part of the velocity gradient, that means

$$D = \frac{1}{2}\left(\nabla \underline{u} + (\nabla \underline{u})^{\top}\right),$$

and $\mu > 0$ denotes the constant dynamic viscosity. As a second model, we shall consider a generalized Newtonian model describing the shear-thinning property of blood, namely

$$S = 2\mu(|D|^2)D, \tag{1.5}$$

where the generalized viscosity μ is shear rate dependent, having the form of a power-law-like Carreau model, see, e.g., [28],

$$\mu(|D|^2) = \eta_\infty + (\eta_0 - \eta_\infty)(1 + \kappa|D|^2)^r.$$

Here η_0, η_∞, κ and r are material parameters. While $\kappa > 0$ and $r \in (-\frac{1}{2}, 0)$ are parameters of shear-thinning, η_0 and η_∞ are asymptotic apparent viscosities of blood for the shear rates, in a simple shear flow. From this, it is clear that η_0, η_∞ are (in theory) independent of the particular shear-thinning model, while κ and r need to be specified from the specific model that they fit the experimental data. Note that it is very common that the shear-thinning of blood is described by a more general Carreau–Yasuda model, see, e.g., [28], having in comparison with the Carreau model an additional material parameter. Nevertheless, in the case of blood, both models give the same quantitative and qualitative fits. Thus we use the simpler one, i.e. (1.5). In this work, we use the values of material parameters, determined by experiments, as given in [28, Chapter II], namely $\eta_0 = 65.7 \times 10^{-3}$ Pas, $\eta_\infty = 4.45 \times 10^{-3}$ Pas, $\kappa = 212.2$ s^2, and $r = -0.325$. In the case of the Newtonian model we use $\mu = \eta_\infty$. It is important to mention that these values differ through the literature, since the blood viscosity is in general depending on many factors like hematocrit, pH, age, gender, etc.

As one can see, we completely neglect possible pathological influences on the blood rheology which can occur in the case of blood flow in an aneurysm sack, like degeneration of the blood cells, thrombus formation etc., see, e.g., [78]. Nevertheless, such biochemical questions should be discussed in future, and in order of a better description of the blood flow nature in the aneurysm, more advanced hemodynamic models with biochemical part should be considered. A consideration of more rigorous models of blood, as well as the interaction of the blood and the vessel wall (fluid structure interaction), can be seen as a future work.

1.2 Non-dimensionalization and boundary conditions

As it was mentioned before, the velocity $\underline{u}$ and the pressure p are governed by the (non-) stationary incompressible (generalized) Navier–Stokes equations (1.1) and (1.2) with the different viscosity models as given in (1.4) and (1.5). For the consecutive numerical computations, it will be useful to recast the governing equations in terms of dimensionless variables, i.e.

$$x \to \frac{x}{L^*}, \quad \underline{u} \to \frac{\underline{u}}{U^*}, \quad p \to \frac{p}{P^*}, \quad \mu \to \frac{\mu}{M^*}, \tag{1.6}$$

where L^* and U^* are the characteristic length and velocity, respectively, M^* is the characteristic dynamic viscosity and P^* is the scaling pressure. All the values, which are denoted by $*$ are suitably chosen for a particular computational setting in order to describe the character of the specific flow problem. For consistency, we choose

$$P^* = \rho(U^*)^2, \quad M^* = \eta_\infty, \tag{1.7}$$

where ρ denotes the constant density of the fluid. Then the time is naturally non-dimensionalized with respect to L^*/U^* and the same holds true for the extra stress tensor S with respect to M^*U^*/L^*.

Next, we shall discuss the different type of boundary conditions. Let us denote by $\Gamma = \partial\Omega$ the boundary of the given domain Ω. In the case of an artery the boundary is decomposed into three mutually different parts, i.e.

$$\overline{\Gamma} = \overline{\Gamma}_{\mathrm{w}} \cup \overline{\Gamma}_{\mathrm{in}} \cup \overline{\Gamma}_{\mathrm{out}},$$

where Γ_{w} denotes the wall, Γ_{in} the inflow and Γ_{out} the outflow boundary, all of positive measure. On these individual parts of the boundary Γ we prescribe the following boundary conditions of mixed type

$$\underline{u} = \underline{0} \text{ on } \Gamma_{\mathrm{w}}, \qquad \underline{u} = \underline{g} \text{ on } \Gamma_{\mathrm{in}}, \qquad T\underline{n} = \underline{0} \text{ on } \Gamma_{\mathrm{out}}.$$

By that, we impose the wall to be non-penetrable on which the fluid perfectly adheres (no-slip), on the outflow we prescribe a physical zero stress, sometimes called "do nothing" boundary condition. On the inflow boundary we prescribe either a physiological inflow condition or a constant (artificial) inflow, where $\underline{g}$ is suitable chosen. This will be discussed in more detail in Chapter 2. From a mathematical point of view, the arterial wall Γ_{w} and the inflow boundary Γ_{in} can be considered as a Dirichlet boundary. The outflow boundary Γ_{out} can be seen as Neumann boundary. In the case of a non-stationary fluid flow we have to additionally impose a suitable initial condition, i.e.

$$\underline{u}(0,x) = \underline{u}^0(x) \text{ in } \Omega,$$

this will be more specified in Chapter 2.

For the given domain Ω and time interval $(0,\bar{t})$ with $\bar{t} > 0$, the system of governing equations (1.1)–(1.2) is then transformed into

$$\begin{aligned}
\partial_t \underline{u} - \frac{2}{Re}\operatorname{div}\left(\mu(|D|^2)D\right) + (\nabla\underline{u})\underline{u} + \nabla p &= \underline{f} && \text{in } \Omega\times(0,\bar{t}),\\
\operatorname{div}\underline{u} &= 0 && \text{in } \Omega\times(0,\bar{t}),\\
\underline{u} &= \underline{0} && \text{on } \Gamma_{\mathrm{w}}\times(0,\bar{t}),\\
\underline{u} &= \underline{g} && \text{on } \Gamma_{\mathrm{in}}\times(0,\bar{t}),\\
T\underline{n} &= \underline{0} && \text{on } \Gamma_{\mathrm{out}}\times(0,\bar{t}),\\
\underline{u} &= \underline{u}^0 && \text{in } \Omega\times\{0\},
\end{aligned} \tag{1.8}$$

where we use the notation for the reduced Reynolds number $Re = \rho L^*U^*/\mu_\infty$. Note, that sometimes the Navier–Stokes equations are expressed in terms of the kinematic viscosity $\nu = \mu_\infty/\rho$. Nevertheless, in the case of non-dimensionalization, this issue is irrelevant due to the Reynolds number being then of the form $Re = L^*U^*/\nu$.

1.3 Variational formulation and discretization

In this section we study the variational formulation and the discretization of the (generalized) Navier–Stokes equations (1.8). We also comment on existence and uniqueness results of weak solutions of the Navier–Stokes equations. In the following we consider the bulk force $\underline{f} \in \widetilde{H}^{-1}(\Omega)^3 = [H^1(\Omega)^3]^*$. Note that for the Dirichlet boundary value problem we may consider the space $H^{-1}(\Omega)^3 = [H_0^1(\Omega)^3]^*$ instead. But for most applications the more regular assumption $\underline{f} \in L^2(\Omega)^3$ is justified. For the individual boundary parts Γ_{w}, Γ_{in} and Γ_{out} of the boundary Γ we assume that

$$\overline{\Gamma}_{\mathrm{in}} \cap \overline{\Gamma}_{\mathrm{w}} \neq \emptyset, \qquad \overline{\Gamma}_{\mathrm{in}} \cap \overline{\Gamma}_{\mathrm{out}} = \emptyset,$$

which is not restrictive for our applications in mind. Consequently the given function $\underline{g}$ on Γ_{in}, describing the inflow, has to satisfy certain properties, i.e. it has to be an element of the space

$$\widetilde{H}^{1/2}(\Gamma_{\mathrm{in}})^3 = \left\{ \underline{v} = \underline{\widetilde{v}}_{|\Gamma_{\mathrm{in}}} \ : \ \underline{\widetilde{v}} \in H^{1/2}(\Gamma)^3, \ \mathrm{supp}\ \underline{\widetilde{v}} \subset \overline{\Gamma}_{\mathrm{in}} \right\}.$$

We shall reformulate the boundary conditions in terms of Dirichlet and Neumann boundary conditions. Thus we introduce $\Gamma_{\mathrm{D}} = \mathrm{int}(\overline{\Gamma}_{\mathrm{in}} \cup \overline{\Gamma}_{\mathrm{w}})$ and $\Gamma_{\mathrm{N}} = \Gamma_{\mathrm{out}}$. This means on the Dirichlet boundary Γ_{D} we introduce the Dirichlet datum $\underline{\widetilde{g}} \in \widetilde{H}^{1/2}(\Gamma_{\mathrm{D}})^3$ as

$$\underline{\widetilde{g}} = \begin{cases} \underline{0} & \text{on } \Gamma_{\mathrm{w}}, \\ \underline{g} & \text{on } \Gamma_{\mathrm{in}}. \end{cases}$$

Let us recall, the system of partial differential equations, i.e. the (generalized) Navier–Stokes equations (1.8), in terms of Dirichlet and Neumann boundary conditions,

$$\begin{aligned} \partial_t \underline{u} - \frac{2}{Re}\mathrm{div}\left(\mu(|D|^2)D\right) + (\nabla \underline{u})\underline{u} + \nabla p &= \underline{f} && \text{in } \Omega \times (0,\bar{t}), \\ \mathrm{div}\ \underline{u} &= 0 && \text{in } \Omega \times (0,\bar{t}), \\ \underline{u} &= \underline{\widetilde{g}} && \text{on } \Gamma_{\mathrm{D}} \times (0,\bar{t}), \\ T\underline{n} &= \underline{0} && \text{on } \Gamma_{\mathrm{N}} \times (0,\bar{t}), \\ \underline{u} &= \underline{u}^0 && \text{in } \Omega \times \{0\}. \end{aligned} \tag{1.9}$$

1.3.1 Remarks on the existence and uniqueness of solutions

In the following we comment on some known existence and uniqueness results of weak solutions of the Navier–Stokes equations. Although the analysis of the system started already in the 1930's, there are still many open problems. Those arise mainly from the nonlinear convective term and from the non-symmetric coupling between velocity and pressure. Because of these, the structure of the system behaves differently in the two- and three

dimensional case, i.e. we consider separately $\Omega \subset \mathbb{R}^n$ with $n = 2,3$. While the two dimensional problem is possible to treat with standard techniques, and thus one directly obtains existence, uniqueness and regularity of the weak solution, the three dimensional case requires more advanced methods and additional assumptions. Let us demonstrate this for the homogeneous Dirichlet boundary value problem of the Navier-Stokes equations, which is in the literature mostly studied. We will comment on the generalization to mixed boundary conditions and generalized viscosity at the end of this section. Namely, we consider for now the problem

$$\begin{aligned} \partial_t \underline{u} - \nu \Delta \underline{u} + (\nabla \underline{u})\underline{u} + \nabla p &= \underline{f} && \text{in } \Omega \times (0,\bar{t}), \\ \operatorname{div} \underline{u} &= 0 && \text{in } \Omega \times (0,\bar{t}), \\ \underline{u} &= \underline{0} && \text{on } \Gamma \times (0,\bar{t}), \\ \underline{u} &= \underline{u}^0 && \text{in } \Omega \times \{0\}, \end{aligned} \tag{1.10}$$

where $\nu = 1/Re$ denotes the viscosity constant. The following overview is based on [27,76], where one can also find a detailed description of the methods used to prove the existence, uniqueness and regularity results. Let us start with the definition of the following spaces, which are suitable for the incompressible Navier–Stokes equations,

$$H^1_{0,\operatorname{div}}(\Omega)^n = \left\{ \underline{v} \in H^1_0(\Omega)^n : \operatorname{div} \underline{v} = 0 \right\},$$

and

$$L^2_{0,\operatorname{div}}(\Omega)^n = \left\{ \underline{v} \in L^2(\Omega)^n : \underline{v} = \nabla q,\ q \in H^1(\Omega) \right\}^{\perp}.$$

Also, let us define what we mean by the term weak solution. We assume in the following $\underline{f} \in L^2(0,\bar{t}; [H^1_{0,\operatorname{div}}(\Omega)^n]^*)$ and $\underline{u}^0 \in L^2_{0,\operatorname{div}}(\Omega)^n$. Then

$$\underline{u} \in L^2(0,\bar{t}; H^1_{0,\operatorname{div}}(\Omega)^n) \cap L^\infty(0,\bar{t}; L^2(\Omega)^n)$$

with

$$\partial_t \underline{u} \in L^1(0,\bar{t}; [H^1_{0,\operatorname{div}}(\Omega)^n]^*)$$

is a weak solution of the Navier–Stokes equations (1.10), iff

$$\langle \partial_t \underline{u}, \underline{v} \rangle_\Omega + \nu \langle \nabla \underline{u}, \nabla \underline{v} \rangle_{L^2(\Omega)} + \langle (\nabla \underline{u})\underline{u}, \underline{v} \rangle_{L^2(\Omega)} = \langle \underline{f}, \underline{v} \rangle_\Omega,$$

for all $\underline{v} \in H^1_{0,\operatorname{div}}(\Omega)^n$ and a.a. $t \in (0,\bar{t})$. If in addition the energy inequality

$$\|\underline{u}(t)\|^2_{L^2(\Omega)} + 2\nu \int_0^t |\underline{u}(\tau)|^2_{H^1(\Omega)} \, d\tau \le \|\underline{u}^0\|^2_{L^2(\Omega)} + 2 \int_0^t \langle \underline{f}(\tau), \underline{u}(\tau) \rangle_\Omega \, d\tau,$$

for a.a. $t \in (0,\bar{t})$, is satisfied, the weak solution is called Leray–Hopf weak solution. Note, that the definition of the weak solution can vary, especially with the method one uses for

the consequent proofs. The reason for introducing the Leray-Hopf solution comes from the necessity to define a suitable solution class where we can obtain uniqueness. Also, this is a reasonable class from physical point of view, since the kinetic energy of the fluid is then controlled by the input data of the problem. The existence results of the above system is then summarized in the following theorem, see [27, Theorem 3.1].

Theorem 1.1. *Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain, $\underline{f} \in L^2(0,\bar{t};[H^1_{0,\mathrm{div}}(\Omega)^n]^*)$ and $\underline{u}^0 \in L^2_{0,\mathrm{div}}(\Omega)^n$, then there holds:*

- *For $n = 2$, there exists a unique weak solution, which is also a Leray–Hopf weak solution, satisfying*

$$\lim_{t\to 0^+} \|\underline{u}(t) - \underline{u}^0\|_{L^2(\Omega)} = 0, \qquad \underline{u} \in C([0,\bar{t}];L^2_{0,\mathrm{div}}(\Omega)^n).$$

- *For $n = 3$, there exists at least one Leray–Hopf weak solution. This solution satisfies the initial condition in the following sense,*

$$\lim_{t\to 0^+} \|\underline{u}(t) - \underline{u}^0\|_{L^2(\Omega)} = 0.$$

The uniqueness for the two dimensional model problem, $n = 2$, is stated in the theorem above. For the three dimensional case, $n = 3$, the following results is valid, see [27, p. 30 ff.].

Theorem 1.2. *Let the assumptions of* Theorem 1.1 *be satisfied and $n = 3$. If the conditions (Prodi–Serrin conditions),*

$$\underline{u} \in L^r(0,\bar{t};L^s(\Omega)^3), \qquad \frac{2}{r} + \frac{3}{s} \le 1, \qquad s \ge 3,$$

are satisfied, the Leray–Hopf weak solution is unique.

It remains to answer the question concerning the reconstruction of the pressure p. As it is shown in [71], this is in general not possible for a right-hand side

$$\underline{f} \in L^2(0,\bar{t};[H^1_{0,\mathrm{div}}(\Omega)^n]^*).$$

Nevertheless, if the domain Ω is of class C^2, $\underline{f} \in L^2(0,\bar{t};H^{-1}(\Omega)^n)$ and $\underline{u}^0$ is smooth enough we obtain

$$\nabla p \in L^r(0,\bar{t};L^s(\Omega)^n), \qquad \frac{2}{r} + \frac{n}{s} = n+1, \qquad 1 < s \le \frac{n}{n-1},$$

see, e.g., [44]. Moreover, if we scale the pressure by $\langle p(t), 1\rangle_\Omega = 0$ for a.a. $t \in (0,\bar{t})$, the pressure is unique.

For the analysis of the Navier–Stokes equations with mixed boundary conditions we refer the reader to, e.g., [5]. The main difficulty of this problem is the violence of the skew-symmetricity of the convective term and thus also of the energy inequality. On the other hand, for the generalized Navier–Stokes equations one needs to handle an extra nonlinearity in the diffusive term. This influences the choice of the solution spaces and requires more advanced techniques, which is studied, e.g., in [52]. As last, we would like to mention that results for the stationary model problem can be found, e.g., in [30, 76].

1.3.2 Discretization

Within this section we discuss the discretization of the initial boundary value problem (1.9). Therefore a time stepping scheme and a stabilized finite element method for the spatial discretization are applied. Further we discuss the application of the Newton method to handle the nonlinear term. As a first step we discretize the first equation of (1.9) in time by using a standard Crank–Nicholson method, see, e.g., [82], where $\Delta t = \bar{t}/n_t$ denotes the time step, for some $n_t \in \mathbb{N}$. This means that we obtain for all $k = 0, \dots, n_t - 1$

$$\begin{aligned}\underline{u}^{k+1} - \frac{\Delta t}{2}\left[\frac{2}{Re}\operatorname{div}\left(\mu(|D^{k+1}|^2)D^{k+1}\right) - (\nabla \underline{u}^{k+1})\underline{u}^{k+1}\right] + \Delta t\, \nabla p^{k+1} \\ = \frac{\Delta t}{2}\left[\underline{f}^{k+1} + \underline{f}^{k}\right] + \underline{u}^{k} - \frac{\Delta t}{2}\left[\frac{2}{Re}\operatorname{div}\left(\mu(|D^{k}|^2)D^{k}\right) - (\nabla \underline{u}^{k})\underline{u}^{k}\right],\end{aligned}$$

where $\underline{f}^k = \underline{f}(k\Delta t)$. Additionally we multiply the above equations with a test function $\underline{v} \in H_0^1(\Omega, \Gamma_D)^3$, and apply integration by parts. This leads to the following variational formulation. At each time step $k = 0, \dots, n_t - 1$ find $(\underline{u}^{k+1}, p^{k+1}) \in H^1(\Omega)^3 \times L^2(\Omega)$ with $\underline{u}^{k+1} = \underline{\tilde{g}}^{k+1}$ on Γ_D such that

$$\begin{aligned}\langle \underline{u}^{k+1}, \underline{v}\rangle_{L^2(\Omega)} + \frac{\Delta t}{2} a(\underline{u}^{k+1}, \underline{v}) - \Delta t\, b(\underline{v}, p^{k+1}) \quad &= \frac{\Delta t}{2}\langle \underline{f}^{k+1} + \underline{f}^{k}, \underline{v}\rangle_{\Omega} \\ &\quad + \langle \underline{u}^{k}, \underline{v}\rangle_{L^2(\Omega)} + \frac{\Delta t}{2} a(\underline{u}^{k}, \underline{v}), \qquad (1.11) \\ b(\underline{u}^{k+1}, q) \quad &= 0,\end{aligned}$$

for all $(\underline{v}, q) \in H_0^1(\Omega, \Gamma_D)^3 \times L^2(\Omega)$, where the corresponding forms are given by

$$\begin{aligned} a(\underline{u}, \underline{v}) &= \frac{2}{Re}\langle \mu(|D(\underline{u})|^2)D(\underline{u}), D(\underline{v})\rangle_{L^2(\Omega)} + \langle (\nabla \underline{u})\underline{u}, \underline{v}\rangle_{L^2(\Omega)}, \\ b(\underline{v}, \underline{p}) &= \langle \operatorname{div} \underline{v}, p\rangle_{L^2(\Omega)}. \end{aligned}$$

An important property for the existence and uniqueness analysis of the saddle point problem (1.11) is the inf–sup condition. The following result can be found in [65, Proposition 5.3.2, p. 157].

Lemma 1.1. *For all $q \in L^2(\Omega)$, the following inf–sup condition is valid*

$$c_S \|q\|_{L^2(\Omega)} \leq \sup_{\underline{0} \neq \underline{v} \in H_0^1(\Omega,\Gamma_D)^n} \frac{b(\underline{v},q)}{\|\underline{v}\|_{H^1(\Omega)}}. \tag{1.12}$$

The variational formulation (1.11) involves still the non-linear terms of the generalized viscosity and the convective term. In particular we can express the variational formulation (1.11) by

$$\mathcal{F}(\underline{x}) = 0,$$

where $\underline{x} = (\underline{u}^{k+1}, p^{k+1})$. In order to solve this nonlinear equation a full Newton method, see, e.g., [20], is applied. This means we obtain

$$D_{\underline{x}_\ell} \mathcal{F}(\underline{x}_\ell)(\underline{x}_\ell - \underline{x}_{\ell+1}) = \mathcal{F}(\underline{x}_\ell),$$

where the left-hand side denotes the Fréchet derivative of $\mathcal{F}(\underline{x}_\ell)$, applied to $\underline{x}_\ell - \underline{x}_{\ell+1}$ for all $\ell \in \mathbb{N}$. The Newton iteration is stopped when a certain relative accuracy $\varepsilon > 0$ for the difference $|\underline{x}_{\ell+1} - \underline{x}_\ell|$ is obtained. For the initial guess in the Newton method we consider the solution of the previous time step. If the initial guess is good enough, i.e. close enough to the solution $\underline{x}$, the Newton method convergences quadratic.

For the spatial discretization we consider an admissible, shape-regular and globally quasi-uniform triangulation $\mathcal{T}_h$ of the domain Ω into tetrahedra. We introduce the finite element spaces

$$\mathcal{V}_h = \operatorname{span}\{\underline{\varphi}_i^1\}_{i=1}^{n_V} \subset H^1(\Omega)^3, \quad \mathcal{Q}_h = \operatorname{span}\{\varphi_i^1\}_{i=1}^{n_Q} \subset L^2(\Omega), \tag{1.13}$$

i.e. of piecewise linear and globally continuous basis functions for both, the velocity and the pressure. Further we need the finite element space

$$\mathcal{V}_{h,0} = \operatorname{span}\{\underline{\varphi}_i^1\}_{i=1}^{n_{V_0}} \subset H_0^1(\Omega,\Gamma_D)^3,$$

with a homogeneous Dirichlet boundary condition. The finite element paring $(\mathcal{V}_h, \mathcal{Q}_h)$ is also known as $\mathcal{P}_1$–$\mathcal{P}_1$ approximation. It is important to mention that this method, as a low order approximation, has an advantage in computation of large systems. This is due to the lower number of degrees of freedom in comparison to a Taylor–Hood approximation, where piecewise quadratic and globally continuous finite elements for the velocity are considered. Nevertheless, the pairing $(\mathcal{V}_h, \mathcal{Q}_h)$ does not satisfy the discrete inf–sup condition

$$\widetilde{c}_S \|q_h\|_{L^2(\Omega)} \leq \sup_{\underline{0} \neq \underline{v}_h \in \mathcal{V}_{h,0}} \frac{b(\underline{v}_h, q_h)}{\|\underline{v}_h\|_{H^1(\Omega)}},$$

for all $q_h \in \mathcal{Q}_h$. Consequently, it results in an unstable method, see, e.g., [12, 30, 39]. In order to overcome this problem, stabilized finite element methods have been developed, see,

e.g., [24, 25, 77]. All of them with the advantage that the finite element spaces (1.13) can be used. The disadvantage of these methods is the fact that often a suitable choice for the stabilization parameter is needed, which is in general difficult to calculate. This problem is captured in a more recent work, see [6], by the Bochev–Dohrmann stabilization, where such a parameter is no longer needed. It introduces an additional pressure penalizing term having the form

$$c(q_h, p_h) := Re\langle p_h - Q_h p_h, q_h - Q_h q_h\rangle_{L^2(\Omega)}, \tag{1.14}$$

where $Q_h : L^2(\Omega) \to \mathcal{Q}_h^0$ is the $L^2(\Omega)$ projection onto the space of piecewise constants. Beside the fact that no stabilization parameter is needed, this method has the advantage that it only acts on the pressure level and it can be easily realized in the case of linear finite elements, as considered here. By this we mean that the projection on the piecewise constants can be simply computed via the representation

$$Q_h q_h|_T = \frac{1}{|T|}\langle q_{h|T}, 1\rangle_{L^2(T)},$$

for all $T \in \mathcal{T}_h$ and $q_h \in \mathcal{Q}_h$. In the following we shall always use the Bochev–Dohrmann stabilization (1.14), due to the advantages mentioned above. Let us denote by I_h a standard interpolation operator on $\mathcal{V}_h|_{\Gamma_D}$. The discrete variational formulation, including the Newton method and the stabilization, is then given as follows: Find at each time step $k = 0, \dots, n_t - 1$ and Newton step $\ell \in \mathbb{N}_0$ the pair $(\underline{\widetilde{u}}_{h,\ell+1}^{k+1}, \widetilde{p}_{h,\ell+1}^{k+1}) \in \mathcal{V}_h \times \mathcal{Q}_h$ with $\underline{\widetilde{u}}_{h,\ell+1}^{k+1} = I_h \underline{\widetilde{g}}^{k+1}$ on Γ_D such that

$$\begin{aligned}
&\langle \underline{\widetilde{u}}_{h,\ell+1}^{k+1}, \underline{v}_h\rangle_{L^2(\Omega)} + \frac{\Delta t}{Re}\langle \mu'(|D(\underline{\widetilde{u}}_{h,\ell}^{k+1})|^2) 4[D(\underline{\widetilde{u}}_{h,\ell}^{k+1}) : D(\underline{\widetilde{u}}_{h,\ell+1}^{k+1})] D(\underline{\widetilde{u}}_{h,\ell}^{k+1}), D(\underline{v}_h)\rangle_{L^2(\Omega)} \\
&\quad + \frac{\Delta t}{Re}\langle \mu(|D(\underline{\widetilde{u}}_{h,\ell}^{k+1})|^2) D(\underline{\widetilde{u}}_{h,\ell+1}^{k+1}), D(\underline{v}_h)\rangle_{L^2(\Omega)} \\
&\qquad + \frac{\Delta t}{2}\langle (\nabla \underline{\widetilde{u}}_{h,\ell+1}^{k+1})\underline{\widetilde{u}}_{h,\ell}^{k+1} + (\nabla \underline{\widetilde{u}}_{h,\ell}^{k+1})\underline{\widetilde{u}}_{h,\ell+1}^{k+1}, \underline{v}_h\rangle_{L^2(\Omega)} - \Delta t\, b(\underline{v}_h, \widetilde{p}_{h,\ell+1}^{k+1}) \\
&= \frac{\Delta t}{Re}\langle \mu'(|D(\underline{\widetilde{u}}_{h,\ell}^{k+1})|^2) 4[D(\underline{\widetilde{u}}_{h,\ell}^{k+1}) : D(\underline{\widetilde{u}}_{h,\ell}^{k+1})] D(\underline{\widetilde{u}}_{h,\ell}^{k+1}), D(\underline{v}_h)\rangle_{L^2(\Omega)} \\
&\quad + \frac{\Delta t}{2}\langle (\nabla \underline{\widetilde{u}}_{h,\ell}^{k+1})\underline{\widetilde{u}}_{h,\ell}^{k+1}, \underline{v}_h\rangle_{L^2(\Omega)} + \frac{\Delta t}{2}\langle \underline{f}^{k+1} + \underline{f}^{k}, \underline{v}_h\rangle_{\Omega} + \langle \underline{\widetilde{u}}_h^k, \underline{v}_h\rangle_{L^2(\Omega)} \\
&\qquad + \frac{\Delta t}{Re}\langle \mu(|D(\underline{\widetilde{u}}_h^k)|^2) D(\underline{\widetilde{u}}_h^k), D(\underline{v}_h)\rangle_{L^2(\Omega)} + \frac{\Delta t}{2}\langle (\nabla \underline{\widetilde{u}}_h^k)\underline{\widetilde{u}}_h^k, \underline{v}_h\rangle_{L^2(\Omega)}, \\
&b(\underline{\widetilde{u}}_{h,\ell+1}^{k+1}, q_h) \qquad\qquad + c(q_h, \widetilde{p}_{h,\ell+1}^{k+1}) = 0,
\end{aligned}$$

for all $(\underline{v}_h, q_h) \in \mathcal{V}_{h,0} \times \mathcal{Q}_h$. Note that $(\underline{\widetilde{u}}_{h,\ell+1}^{k+1}, \widetilde{p}_{h,\ell+1}^{k+1})$ denotes here the finite element solution of the perturbed problem which is due to the interpolation of the Dirichlet datum $\underline{\widetilde{g}}^{k+1}$. Further $\underline{\widetilde{u}}_h^k$ denotes the solution of the previous time step which is obtained at the last Newton iteration.

For error estimates of the Bochev–Dohrmann stabilization only results for the Stokes equations with homogeneous Dirichlet boundary conditions are known. This means that a constant viscosity is assumed and the nonlinear convective term is neglected. In this particular case the following result is valid, see [6].

Theorem 1.3. *Let us consider the finite element spaces $\mathcal{V}_h$ and $\mathcal{Q}_h$, defined in* (1.13)*. For the Stokes equations with the Bochev–Dohrmann stabilization* (1.14) *the following error estimate is valid,*

$$\|\underline{u}-\underline{u}_h\|_{L^2(\Omega)}+h\|p-p_h\|_{L^2(\Omega)} \leq ch^s\left(|\underline{u}|_{H^s(\Omega)}+|p|_{H^s(\Omega)}\right),$$

where $(\underline{u},p)\in H_0^1(\Omega)^3\cap H^s(\Omega)^3\times H^s(\Omega)$ is the exact solution for some $s\in[1,2]$.

It is important to mention that in the stabilization term $c(q_h,p_h)$ the pressure p_h is projected onto the piecewise constants. As a consequence we can not expect full second order of convergence in $L^2(\Omega)$, in general. In particular we obtain a reduced order of convergence with a factor of minus one.

1.4 Concluding remarks

In this chapter we have discussed different hemodynamic models for arterial blood flow, which are prescribed via the Cauchy stress tensor (1.3). The corresponding extra stress tensor involves either a constant viscosity (1.4), or a non-constant viscosity of a power-law-like Carreau model (1.5). We have described corresponding boundary conditions, which model the inflow, outflow and the no-slip condition on the arterial wall, and presented the (generalized) Navier–Stokes equations as a model problem. We have presented an overview on the existence and uniqueness results for the solution of the Navier–Stokes equations with a constant viscosity. Moreover we have introduced a discretization of the equations, using a Crank–Nicholson scheme in time and a stabilized finite element method in space. The advantages of such an approach have been discussed.

In the upcoming chapter we discuss hemodynamic indicators which are calculated form the system above and present several numerical simulations.

2 HEMODYNAMICS FOR ARTERIAL BLOOD FLOW

In this chapter we simulate the blood flow in arteries with aneurysms and identify important factors of the pathological flow with respect to the geometry. These factors, as we will see later, will be the wall shear stresses on the arterial wall and the complex vortices within the aneurysm. It is known, that especially a vortex in an aneurysm can lead to a blood clot (trombus) formation, which can consequently lead to its rupture. This motivates the consideration and simulation of hemodynamic indicators which are an important measure for the initiation, evolution and rupture of arteries, in particular aneurysms. In the past, the following two indicators became commonly used, namely the wall shear stress (WSS) and the oscillatory shear index (OSI). It turns out that the computation of these indicators is not always done by the same formula. By this we mean, that sometimes in the boundary stress vector, from which the indicators are calculated, a specific term is neglected. This can result in main differences in WSS and OSI, mainly at the critical points of the artery where aneurysms or arterial plaque appear. We present the differences of these indicators by several numerical simulations, for the different blood models, as described in Chapter 1. The influence of the correct choice, from a rheological point of view, appears to be significant, see also [40].

2.1 Hemodynamic indicators

The initiation, evolution and rupture of aneurysm result, as most degenerative cardiovascular diseases, from a combination of hemodynamics, vessel wall mechanics, and physical and biochemical processes within and between them. In the past decades, the hypothesis of strong correlation between the blood flow induced mechanical stresses and the arterial wall functionality, degenerative chemical processes within and around it has been several times verified [18, 43, 46, 51, 67, 70, 83]. More precisely, the endothelial cells of the vessel wall are mechanosensitive to the local shear stresses [50, 83], transferring the abnormal wall shear stress into specific biochemical signals which modulate the cellular structure of the wall. This results in the wall thinning and in an increase of the lipid and adhesion molecules permeability through the wall, see e.g. [18], which is connected to aneurysm plaque or thrombus formation. On the other hand, the physiological level of the shear stress at the wall is protective. As specified for example in [50], the range, its non-dimensionalized physical quantities shall be specified later, of wall shear stress is about $15 - 70$ dyne cm^{-2}. From above it is clear that identification of local wall shear stress (WSS) plays an important role in the study of aneurysm evolution *in silico*, and thus, it shall be a key factor in the characterization of aneurysms.

Across the literature one can find different approaches to the WSS computation. By a rheological definition, WSS is the shear traction caused by the blood flow acting on the endothelial cell surface. In terms of the Cauchy stress tensor $\boldsymbol{T}$, defined in (1.3), this means

$$\mathrm{WSS} := |(\boldsymbol{T}\underline{n}) \cdot \underline{t}_{\mathrm{blood}}|, \tag{2.1}$$

where $\underline{n}$ is the outer normal unit vector of the tangential plane to the vessel wall and $\underline{t}_{\mathrm{blood}}$ is the tangential unit vector living in such a plane, having the same direction as the velocity vector of blood. Note that $\underline{t}_{\mathrm{blood}}$ is orthogonal on $\underline{n}$. For a two-dimensional case of unidirectional flow, this characterization is intuitive due to the unique identification of $\underline{t}_{\mathrm{blood}}$. Similarly, such a characterization of wall shear stress is meaningful also for simple shear flow in the three-dimensional case. In order for a better understanding of the problematics in the choice of $\underline{t}_{\mathrm{blood}}$ we derive WSS for simple shear flow, described in Cartesian coordinates in the x_1-axis. Then, the velocity vector and the corresponding normal and tangential unit vectors are

$$\underline{u} = (u_1(x_2), 0, 0)^\top, \quad \underline{n} = (0,1,0)^\top, \quad \underline{t}_{\mathrm{blood}} = (1,0,0)^\top.$$

Correspondingly we have

$$\begin{aligned}(\boldsymbol{T}\underline{n}) \cdot \underline{t}_{\mathrm{blood}} &= -p\underline{n} \cdot \underline{t}_{\mathrm{blood}} + 2\mu(\cdot)(\boldsymbol{D}\underline{n}) \cdot \underline{t}_{\mathrm{blood}} \\ &= \mu(\cdot) \begin{pmatrix} 0 & \partial_{x_2} u_1(x_2) & 0 \\ \partial_{x_2} u_1(x_2) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \mu(\cdot)\partial_{x_2} u_1(x_2),\end{aligned}$$

and

$$2\mu(\cdot)\boldsymbol{D}\underline{n} = \mu(\cdot) \begin{pmatrix} \partial_{x_2} u_1(x_2) \\ 0 \\ 0 \end{pmatrix}.$$

As a consequence we observe that

$$|(\boldsymbol{T}\underline{n}) \cdot \underline{t}_{\mathrm{blood}}| = |2\mu(\cdot)\boldsymbol{D}\underline{n}|.$$

Obviously, for simple shear flow for which the flow is purely laminar, the wall shear stress vector $\underline{\tau}_{\mathrm{w}}$, for which we have $\mathrm{WSS} = |\underline{\tau}_{\mathrm{w}}|$, and the (scalar) wall shear stress are then given as

$$\underline{\tau}_{\mathrm{w},1} = 2\mu(\cdot)\boldsymbol{D}\underline{n}, \qquad \mathrm{WSS}_1 = |2\mu(\cdot)\boldsymbol{D}\underline{n}|. \tag{2.2}$$

Such a derivation then tempt, that these formulae are also valid in a general three-dimensional setting, see, e.g., [45, 46, 50] and many others. However, for complicated geometries, like arteries with aneurysms, the flow at the arterial wall can not be approximated by simple shear, since a non-negligible part of $\boldsymbol{D}\underline{n}$ will not act in the shear direction.

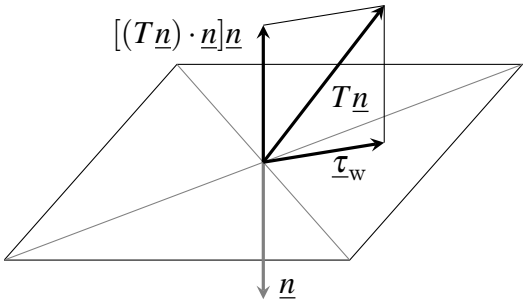


Figure 2.1: Stress decomposition at the infinitesimal plane into its normal and tangential part. In the case that $\underline{n}$ coincides with the outer normal of the vessel wall, we call the tangential part $\underline{\tau}_{\mathrm{w}}$ the wall shear stress vector.

Formula (2.1) is then for those problems not usable, since $\underline{t}_{\mathrm{blood}}$ is given by two a priori not known tangential vectors, and thus, together with the zero Dirichlet boundary condition for the velocity one can not directly derive it in this way.

Instead, one can use a full decomposition approach. This means, from the unique decomposition

$$T\underline{n} = [(T\underline{n})\cdot\underline{n}]\underline{n} + [(T\underline{n})\cdot\underline{t}_{\mathrm{blood}}]\underline{t}_{\mathrm{blood}},$$

we conclude

$$|(T\underline{n})\cdot\underline{t}_{\mathrm{blood}}| = |T\underline{n} - [(T\underline{n})\cdot\underline{n}]\underline{n}|.$$

Consequently, the shear stress vector is then derived from the wall traction $T\underline{n}$ by subtracting the normal stress vector as depicted in Figure 2.1. This motivates the following definition of the wall shear stress vector

$$\begin{aligned}\underline{\tau}_{\mathrm{w},2} &= T\underline{n} - [(T\underline{n})\cdot\underline{n}]\underline{n}\\ &= -p\underline{n} + S\underline{n} + [(p\underline{n}\cdot\underline{n})]\underline{n} - [(S\underline{n})\cdot\underline{n}]\underline{n}\\ &= S\underline{n} - [(S\underline{n})\cdot\underline{n}]\underline{n},\end{aligned}$$

and specifically for the (generalized) Newtonian case we obtain

$$\begin{aligned}\underline{\tau}_{\mathrm{w},2} &= S\underline{n} - [(S\underline{n})\cdot\underline{n}]\underline{n}\\ &= 2\mu(\cdot)(D\underline{n} - [(D\underline{n})\cdot\underline{n}]\underline{n}).\end{aligned} \tag{2.3}$$

As a consequence we have for the wall shear stress (2.1) the following representation

$$\mathrm{WSS}_2 = |\underline{\tau}_{\mathrm{w},2}|. \tag{2.4}$$

Here one should notice that the normal stress is not identified with the mean normal stress $-p\underline{n}$ but it is superposed together with $[(S\underline{n})\cdot\underline{n}]\underline{n}$. This normal part of the extra stress is

very often neglected, reasoned by the flow conditions close to the simple shear. Later we shall see that this term is nevertheless not of small order mainly at the critical points of the domain where an aneurysm or arterial plaque appear and thus it can not be neglected for the simulation of hemodynamic indicators.

Whatever definition of WSS we use, it is still a local physical quantity expressed at a given time. Thus, it is preferable to consider this indicator over a certain time period, either the time of observation or the period of the cardiac cycle. For this, we introduce a time-averaged wall shear stress (AWSS) as proposed in [46], characterizing the areas of low shear stresses at the vessel wall during the time interval $(0,\bar{t})$

$$\mathrm{AWSS} := \frac{1}{\bar{t}} \int_0^{\bar{t}} |\underline{\tau}_{\mathrm{w}}(t)|\, dt,$$

where $\underline{\tau}_w$ is the wall shear stress vector. Thus, as we are interested in the differences arising from the choice of the form of $\underline{\tau}_w$, we define

$$\mathrm{AWSS}_1 := \frac{1}{\bar{t}} \int_0^{\bar{t}} \left|\underline{\tau}_{\mathrm{w},1}(t)\right| dt, \quad \mathrm{AWSS}_2 := \frac{1}{\bar{t}} \int_0^{\bar{t}} \left|\underline{\tau}_{\mathrm{w},2}(t)\right| dt. \tag{2.5}$$

However, in the case of a pulsative flow, some pathological flow patterns at or near the wall can develop, such as stagnation points or wall shear stresses with an oscillating character, for which a quantity such as AWSS can be high as well. This is due to the fact that AWSS is computed from the magnitude of the shear force and thus it is free from the information about the oscillatory character. Hence we introduce, see [46], an additional hemodynamic indicator, the oscillatory shear index (OSI)

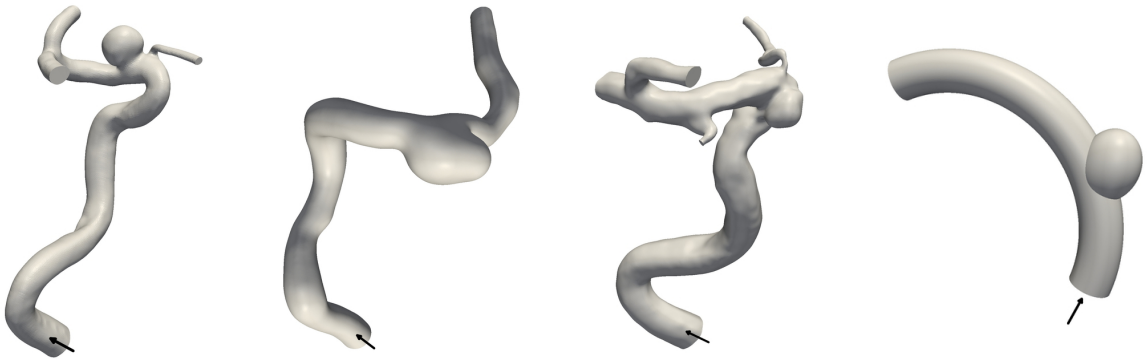
$$\mathrm{OSI} := \frac{1}{2}\left(1 - \frac{\left|\int_0^{\bar{t}} \underline{\tau}_{\mathrm{w}}(t)\, dt\right|}{\int_0^{\bar{t}} |\underline{\tau}_{\mathrm{w}}(t)|\, dt}\right), \tag{2.6}$$

to provide a characterization of the deviation of the WSS vector from its averaged direction, in other words a measure of WSS oscillations where AWSS is not predictive. Here, $\int_0^{\bar{t}} \underline{a}(t)\, dt$ stands for a vector with components computed as integrals of corresponding components of the vector $\underline{a}$. In the case that the denominator in (2.6) is zero, we set $\mathrm{OSI} = 0$, since in that case the nominator adopts zero value as well. The values of OSI are in the range $[0, 0.5]$, where $\mathrm{OSI} = 0$ corresponds to a unidirectional flow (protective) and $\mathrm{OSI} = 0.5$ to a purely oscillating flow (pathological). Again, in the same fashion as (2.5), we define

$$\mathrm{OSI}_1 := \frac{1}{2}\left(1 - \frac{\left|\int_0^{\bar{t}} \underline{\tau}_{\mathrm{w},1}(t)\, dt\right|}{\int_0^{\bar{t}} \left|\underline{\tau}_{\mathrm{w},1}(t)\right| dt}\right), \quad \mathrm{OSI}_2 := \frac{1}{2}\left(1 - \frac{\left|\int_0^{\bar{t}} \underline{\tau}_{\mathrm{w},2}(t)\, dt\right|}{\int_0^{\bar{t}} \left|\underline{\tau}_{\mathrm{w},2}(t)\right| dt}\right). \tag{2.7}$$

From what has been described above, both indicators need to be investigated simultaneously, with a focus on the regions where either AWSS and OSI are small, or, OSI is high regardless of the AWSS value. Since we are interested in relative differences between AWSS_1 and AWSS_2, OSI_1 and OSI_2, we will present the numerical results also in non-dimensionalized form. The corresponding values in the physical units can be obtained by a simple calculation due to (1.6)–(1.7) and (2.8).

2.2 Geometry and data



2.2.1: GEOM_1 2.2.2: GEOM_2 2.2.3: GEOM_3 2.2.4: GEOM_4

Figure 2.2: Studied geometries. Black arrows denote the inflow boundary.

We consider four different geometries of cerebral arteries with aneurysm, three realistic, reconstructed from CTA imaging, and one artificial (symmetric), see Figure 2.2. Geometry GEOM_1 was obtained as open source mesh from the CISTIB lab at the Universitat Pompeu Fabra of Barcelona, GEOM_2 was provided from [29], and GEOM_3 from [66]. We always denote the domain of interest by Ω and its boundary by Γ which is decomposed into three parts, namely the boundary of the wall and the parts of inflow and outflow, i.e. $\overline{\Gamma} = \overline{\Gamma}_{\mathrm{w}} \cup \overline{\Gamma}_{\mathrm{in}} \cup \overline{\Gamma}_{\mathrm{out}}$, all of positive measure, as described in Chapter 1.

Under normal conditions, the diameter of a cerebral artery is approximately $5\,\mathrm{mm}$, and, as it will be specified below, the velocity inflow is in the range $10-50\,\mathrm{cm\,s^{-1}}$. This, together with the viscosity $\eta_\infty = 4.45\times10^{-3}\,\mathrm{Pa\,s}$, gives us the characteristic units of the problem under consideration,

$$L^* = 1\,\mathrm{cm}, \quad U^* = 10\,\mathrm{cm\,s^{-1}}, \quad \frac{M^*}{\varrho} = 4.2\times10^{-2}\,\mathrm{cm^2 s^{-1}}, \tag{2.8}$$

which scales the time by $\bar{t}^* = 0.1\,\mathrm{s}$ and the stress by $S^* = 0.445\,\mathrm{dynes\,cm^{-2}}$. The corresponding Reynolds number is then $\mathrm{Re} \approx 240$. The mesh geometries are scaled into the characteristic units as well, that means the diameter of all computational arteries is approximately 0.5 of non-dimensional units.

Boundary conditions

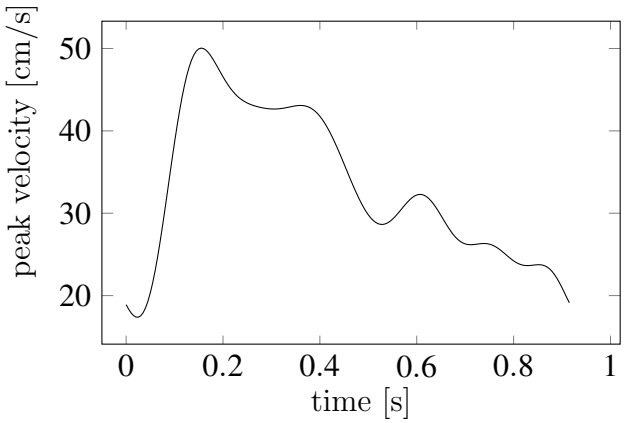


Figure 2.3: Archetypal waveform of the peak velocity (in units) in the internal carotid artery over one cardiac cycle, created by fitting to experimental data, from [78].

As it was mentioned in Chapter 1 we impose, the wall Γ_{w} to be non-penetrable on which the fluid perfectly adheres (no-slip), and on the outflow Γ_{out} we prescribe physical zero stress boundary condition. Moreover, on the inflow boundary Γ_{in} we prescribe either a physiological inflow condition or a constant (artificial) inflow. For both cases, the inflow is governed by a given velocity function

$$\underline{g}(t,x) = g_t(t)\phi(t)\underline{g}_x(x).$$

Here $\underline{g}_x(x)$ represents a parabolic profile of the inflow, with $0 \leq |\underline{g}_x(x)| \leq 1$, scaled by $g_t(t)$, representing its periodic change over one cardiac cycle. In general, the vessel cross-section may not posses a circular profile. In that case, the prescribed parabolic function needs to be properly scaled or has to have a suitable decay at the boundary of such a cross-section. Additionally, $\phi(t)$ stands for initial damping.

In arteries, the velocity profile of the blood flow is generated by the heart beat, nevertheless, the magnitude and the shape of pulses change at different parts of the arterial system, mainly due to the branching, wall deformation, and the complex curvature of the cardiovascular system. We use a profile experimentally determined for an internal carotid artery (artery of our interest) by [78], see Figure 2.3, with a period of the cardiac cycle of 0.917 s. Such a multi-harmonic function can be decomposed into Fourier series, where in this particular case, 7 summands of the series are approximating the waveform accurately enough, i.e.

$$g_{t1}(t) = \frac{a_0}{2} + \sum_{k=1}^{7} \left\{ a_k \cos(k\omega t) + b_k \sin(k\omega t) \right\}, \tag{2.9}$$

where $\omega = 2\pi/\bar{t}_f$ denotes the frequency of oscillation and a_k, b_k are the Fourier coefficients given by a fitting of g_t to experimental data, which are taken from [78]. As a second profile

of the inflow velocity we consider a constant profile, computed as an average of the above described wave, namely

$$g_{t2}(t) = \frac{1}{\bar{t}_f} \int_0^{\bar{t}_f} g_{t1}(t)\, dt \approx 3.33. \tag{2.10}$$

For computational reasons we damp the wave g_t at the initial period by

$$\phi(t) = \begin{cases} \frac{1}{2}\left(1+\cos(\pi(\frac{t}{\bar{t}_f}-1))\right) & \text{for } t < \bar{t}_f, \\ 1 & \text{else,} \end{cases} \tag{2.11}$$

to obtain a smooth evolution of the flow from the initial rest state $\underline{u}(0,x) = \underline{0}$, $x \in \Omega$.

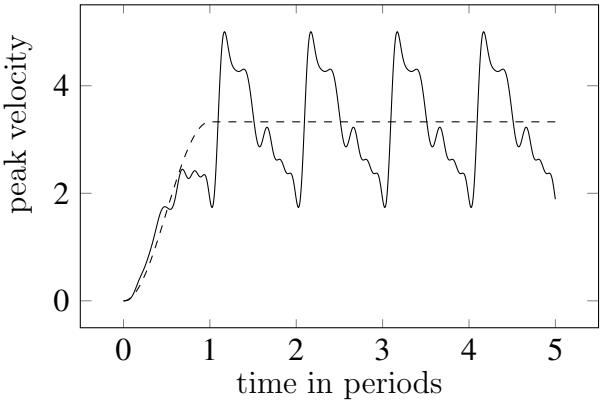


Figure 2.4: Two different inflow profiles of the function $g_t(t)$ with a scaling $\phi(t)$ from the rest state, over the computational time interval $(0, 5\bar{t}_f)$; Full line: $g_{t1}(t)\phi(t)$ - physiological profile, as in Figure 2.4, dashed line: $g_{t2}(t)\phi(t)$ - averaged (over one period) profile.

Both inflow time profiles are depicted in Figure 2.4.

For the Crank–Nicholson scheme in time we use as a time step $\Delta t = \bar{t}_f/50$ in the simulation time interval $(0, 5\bar{t}_f)$. For the spatial discretization we apply a stabilized finite element method as described in Chapter 1. The number of degrees of freedom and the number of spatial elements are presented for each geometry in Table 2.1. For the Newton method we consider a relative accuracy of $1e-8$. The resulting linear system of equations is solved by the direct solver Pardiso, see [7].

2.3 Numerical results

As a first computational result we present the streamlines (at the same time) of the velocity in the aneurysm for all considered geometries in Figure 2.5. We observe the typical vortex formation in the aneurysms.

	GEOM_1	GEOM_2	GEOM_3	GEOM_4
space elements	568 050	1 388 238	1 376 085	1 002 972
DoFs	352 905	900 693	903 104	648 567

Table 2.1: Numbers of spatial elements and degrees of freedom (DoFs) for all considered geometries.

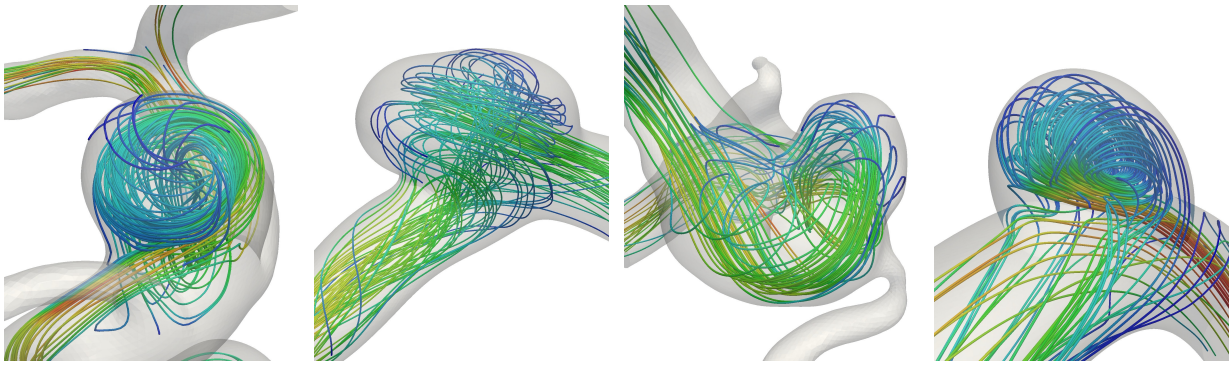


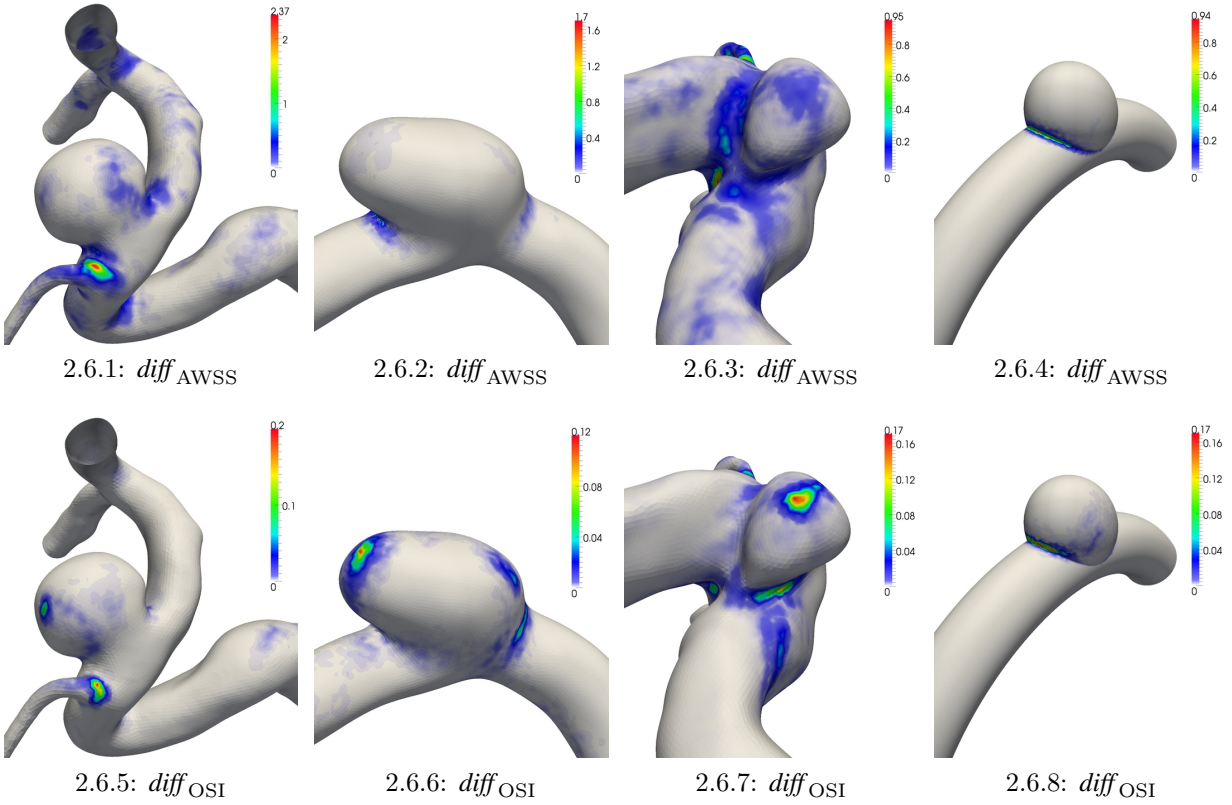
Figure 2.5: Typical vortex formation in the aneurysm for all considered geometries. Computational setting: generalized viscosity, periodic inflow.

Moreover, as it was outlined in the introduction, we focus on the determination of the difference between the hemodynamic indicators (both AWSS and OSI) with respect to the used formulae of the wall shear stress vector $\underline{\tau}_{\mathrm{w}}$, see (2.2) and (2.3). This means, we aim to demonstrate absolute differences

$$\mathit{diff}_{\mathrm{AWSS}} := |\mathrm{AWSS}_1 - \mathrm{AWSS}_2|, \quad \text{and} \quad \mathit{diff}_{\mathrm{OSI}} := |\mathrm{OSI}_1 - \mathrm{OSI}_2|, \tag{2.12}$$

for all four geometries, where AWSS_1, AWSS_2, OSI_1 and OSI_2 are defined in (2.5) and (2.7). The computational results for the periodic inflow (2.9) and for the shear rate dependent viscosity model (1.5) are presented in Figure 2.6, with zooms on the aneurysms. The results clearly show that the differences are strongly dependent on the complexity of the geometry. This correlates with the fact that near the "smooth" parts of the boundary the characteristic of the flow is close to the simple shear. Hence, both wall shear stress vectors $\underline{\tau}_{\mathrm{w},1}$ and $\underline{\tau}_{\mathrm{w},2}$ are identical, and thus OSI and AWSS are the same for both approaches. These areas are represented in Figure 2.6 without color. The parts of the walls where the simple shear approximation fails are either those with higher curvature (like at bifurcations, sharp curves, necks of aneurysms), or those which are near the flow vortices (like the heads of aneurysms). Even though those parts of the boundary are minor, they are exactly the critical regions where aneurysms evolve, and thus, to obtain more precise results, the whole decomposition of wall traction should be assumed.

To compare the magnitude of the differences with the actual magnitudes of computed AWSS and OSI, we include Figure 2.7 of geometry GEOM_3. For this particular computational setting, the maximal differences are approximately of 20% in AWSS and 40% in OSI

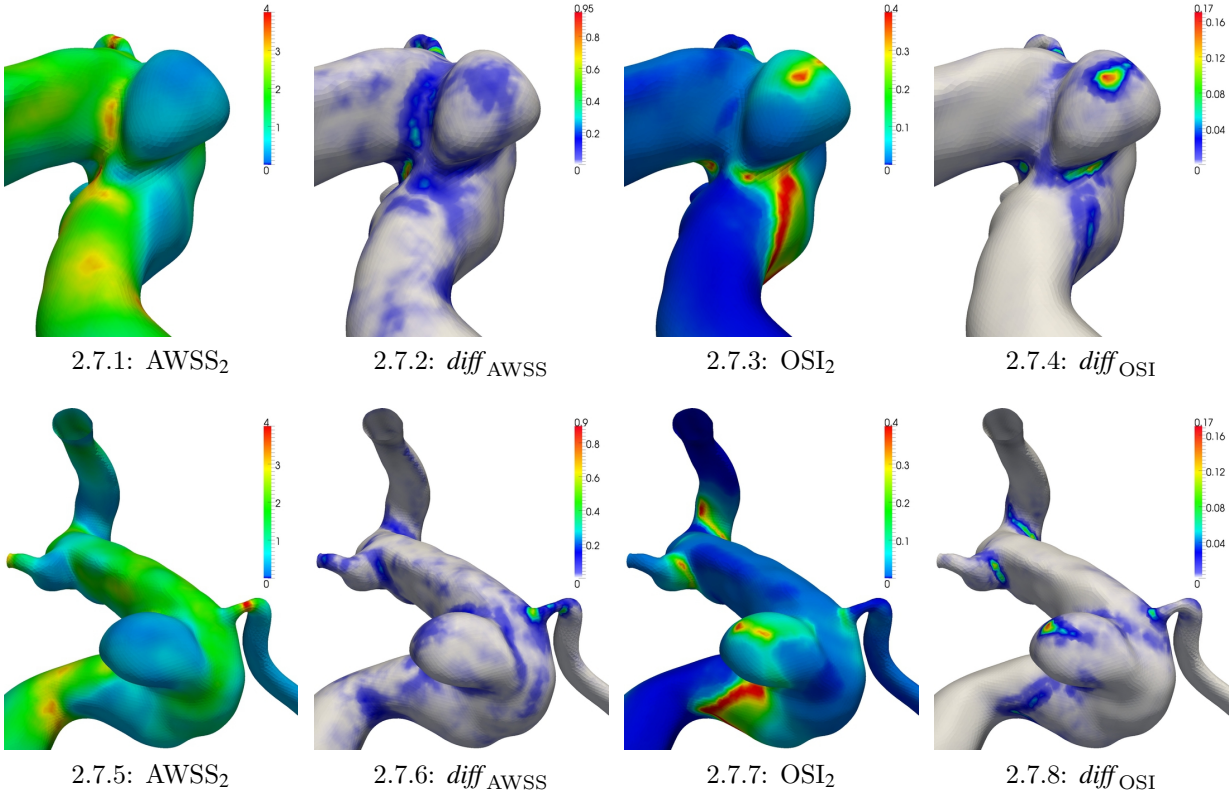


2.6.1: $diff_{\mathrm{AWSS}}$ 2.6.2: $diff_{\mathrm{AWSS}}$ 2.6.3: $diff_{\mathrm{AWSS}}$ 2.6.4: $diff_{\mathrm{AWSS}}$

2.6.5: $diff_{\mathrm{OSI}}$ 2.6.6: $diff_{\mathrm{OSI}}$ 2.6.7: $diff_{\mathrm{OSI}}$ 2.6.8: $diff_{\mathrm{OSI}}$

Figure 2.6: Absolute differences between AWSS and OSI computations for all four geometries. On parts of the boundary without color are the indicators identical, i.e. $\mathrm{AWSS}_1 = \mathrm{AWSS}_2$ and $\mathrm{OSI}_1 = \mathrm{OSI}_2$. Computational setting: generalized viscosity, periodic inflow.

with respect to the result obtained by the full decomposition. Nevertheless, the parts of the wall where the AWSS and OSI differ mostly are not identical. AWSS_1 differs from AWSS_2 mainly at the sharp curve of the main vessel, while the OSI differences are mainly located at the head of aneurysm and bifurcations. The areas where $diff_{\mathrm{AWSS}}$ are highest reflect regions where the vector $\underline{\tau}_{\mathrm{w}}$ changes mainly in magnitude, while for the $diff_{\mathrm{OSI}}$ they are determined by the $\underline{\tau}_{\mathrm{w}}$ having different directions but possibly of similar magnitude during the flow period. This then causes that regions of highest AWSS/OSI differences are not always identical. For better illustration of this fact, we include line cut profiles in Figure 2.8. For both cases, AWSS_1 and AWSS_2, respectively OSI_1 and OSI_2, exhibit similar characteristics but the values at critical regions differ. The cut 1 is represented in Figure 2.8.1, and as it goes from front to the back, the cut parametrization in Figures 2.8.2–2.8.3 goes from 0 to 1. For this cut, we observe nearly no differences in AWSS but profound differences in OSI. In the case of cut 2 (Figure 2.8.4), the remarkable differences in AWSS are positioned at the inner curve of the main vessel, i.e. the left part of the representation of cut 2. For clearer comparison reasons, the plotted values are neither smoothened nor interpolated.



2.7.1: AWSS_2 2.7.2: $diff_{\mathrm{AWSS}}$ 2.7.3: OSI_2 2.7.4: $diff_{\mathrm{OSI}}$

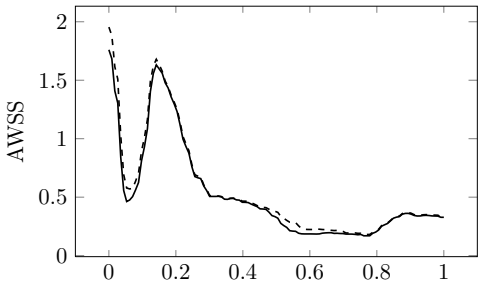
2.7.5: AWSS_2 2.7.6: $diff_{\mathrm{AWSS}}$ 2.7.7: OSI_2 2.7.8: $diff_{\mathrm{OSI}}$

Figure 2.7: Comparison of the relative differences for the geometry GEOM_3 in two zoomed views. For this case the maximal difference takes up to approximately 20% in AWSS and 40% in OSI with respect to the result obtained by full decomposition. Computational setting: generalized viscosity, periodic inflow.

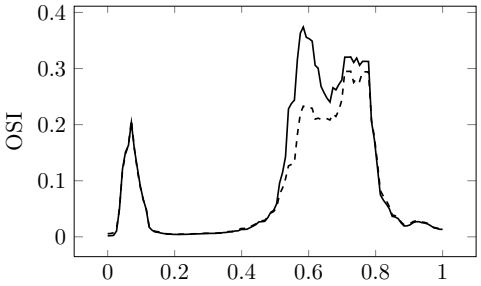
As last, we present in Figure 2.9 simulation results with a focus on the influence of the computational setting on the value of OSI (here are the differences mostly distinguishable) and the corresponding differences between the full and partial decomposition approaches in the computation of $\underline{\tau}_{\mathrm{w}}$. The pictures in Figure 2.9 are of geometry GEOM_2 in two mutually opposite zooms on the aneurysm head. First, we can notice remarkable influence of the used viscosity model on the OSI distribution, i.e. the difference in computation with Newtonian (constant) viscosity and generalized (shear-rate dependent) viscosity, compare first and third row of the figure. These both cases are results for a periodic inflow of the velocity. On the other hand, the difference between the distributions of OSI for the case of periodic and constant inflow is not of such a magnitude. This is due to the fact that the flow in the aneurysm head is slowed down and it does not exhibit such a strong periodic character as in the vessel itself. Nevertheless, on the vessel wall are the characteristics much more distinguishable (partially notable in the zoom views as well). In this set of pictures, the differences are defined as in (2.12).

We have considered only two averaged hemodynamic indicators. Nevertheless, there are

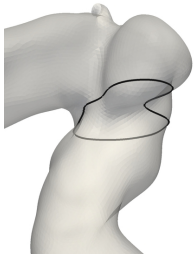
2.8.1: cut 1



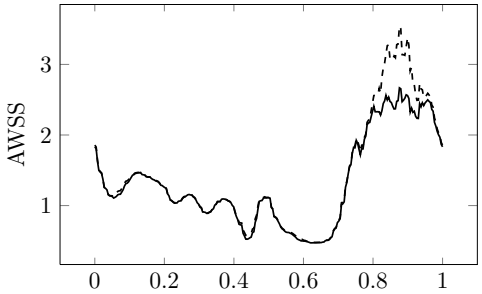
2.8.2: AWSS along cut 1



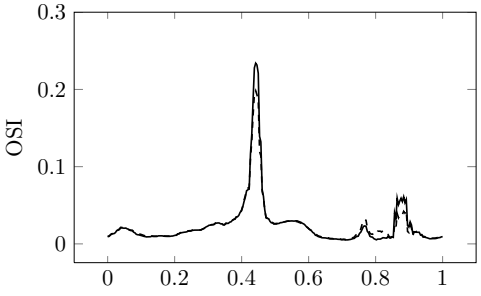
2.8.3: OSI along cut 1



2.8.4: cut 2



2.8.5: AWSS along cut 2



2.8.6: OSI along cut 2

Figure 2.8: Direct comparison of the indicators on the line cuts as schematically depicted in Figures 2.8.1 and 2.8.4. The line cut through the boundary is parametrized to the interval $(0,1)$, used as a characterization of a position on the cut for plots in Figure 2.8.2, 2.8.3, 2.8.5 and 2.8.6. As one can see, the differences of AWSS are placed away from the aneurysm head, while for the OSI they occur on the top of the head. Full line: full decomposition (indicators computed from $\underline{\tau}_{\mathrm{w},2}$), dashed line: partial decomposition (indicators computed from $\underline{\tau}_{\mathrm{w},1}$).

also other indicators which play an important role in physiological research of the vessel wall. The focus was not to specify the whole scale of the relevant indicators since most of them are directly derived from the wall shear stress vector $\underline{\tau}_{\mathrm{w}}$. We rather wanted to give a direct example that the way of how the wall shear stress vector is computed leads to significant differences, illustrated on the two most used indicators. Obviously, differences in the indicators will also have an impact on the interpretation of the results from a physiological point of view.

2.4 Concluding remarks

In this chapter we have been focused on the blood flow simulation in aneurysms. Mainly on the illustration of the importance of the formula for the use of the wall shear stress vector in the computation of hemodynamic indicators. A full decomposition approach is

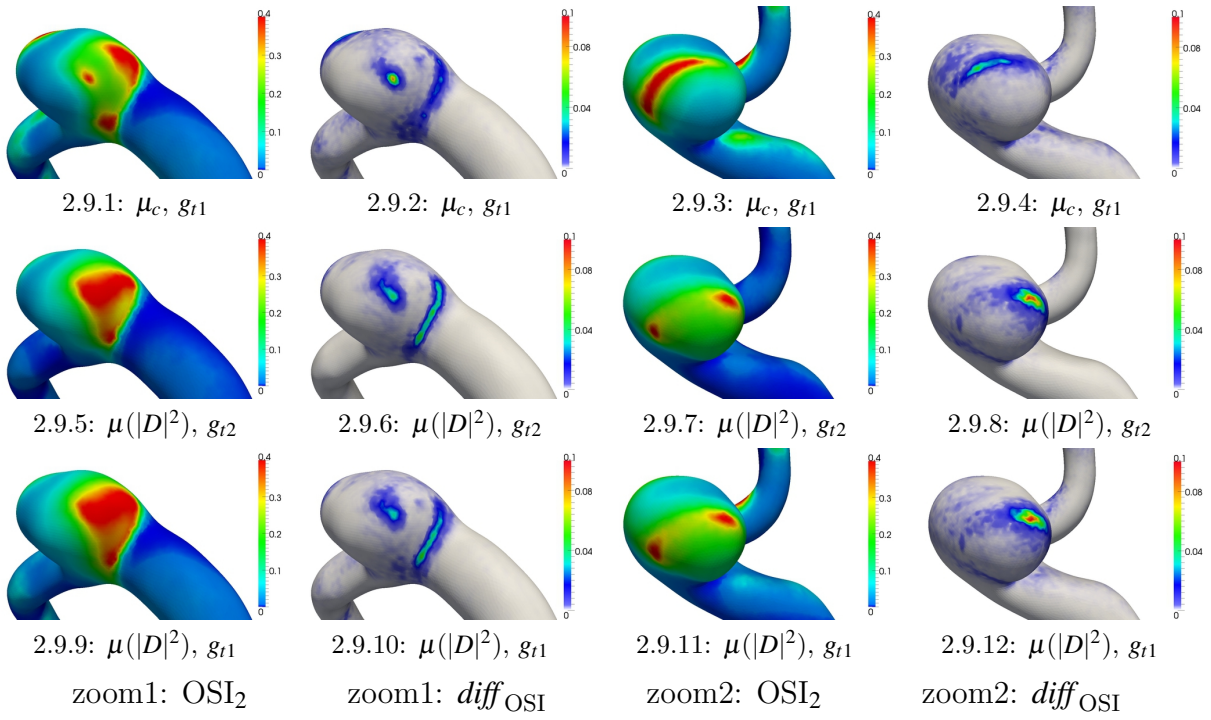


Figure 2.9: Comparison of OSI computed by full decomposition (from $\underline{\tau}_{w,2}$) and of the absolute differences between the two approaches, with focus on the influence of different computational setting. Top row: Newtonian model (constant viscosity), middle row: generalized Newtonian model (non-constant viscosity), constant inflow (described by function g_{t2}), bottom row: generalized Newtonian model (non-constant viscosity).

a good starting point for the characterization of critical areas of artery walls with respect to the formation and progression of aneurysms. Nevertheless, in further work the models should be improved by inclusion of the most significant aspects which can influence those indicators as well. From our perspective, this includes the following. First, a more realistic blood model which can, in a reasonable range, capture the pathological behavior of blood near the critical areas and/or its non-Newtonian properties. And as a second, the influence of the wall deformation caused by the blood flow circulation. For the considered numerical method this means to include the fluid-structure interaction, and, for the modeling part, a reasonable solid-like deformation model.

Another important point is the question about the optimization of the wall shear stress vector, and as a consequence the optimization of the hemodynamic indicators. By this we mean the optimal control of the inflow velocity into an arterial system, motivated for instance by an artificial heart pump. The same question can be asked about the vortex minimization in the aneurysm with respect to the inflow velocity, see, e.g., [47]. In such cases the corresponding model problem can be described by the minimization of a tracking

type cost functional, constrained by the partial differential equation describing blood flow. Let us mention the following example, which is given as: Minimize the cost functional

$$\mathcal{J}(\underline{u},\underline{z}) := \frac{1}{2}\|\underline{u}-\overline{\underline{u}}\|^2_{L^2(\Omega)} + \frac{1}{2}\varrho|\underline{z}|^2_{H^{1/2}(\Gamma)},$$

subject to the constraint, the Navier–Stokes equations. Here $\overline{\underline{u}}$ describes the desired state, which could be for instance a constant flow field or the solution of the Stokes equations, etc. Moreover we consider the cost or regularization coefficient $\varrho > 0$ and the unknown control $\underline{z}$, which describes the unknown inflow velocity.

In the following we shall first present a unified numerical analysis for such a model problem in case of the Poisson equation as a constraint. Moreover we present numerical results for the Navier–Stokes equation, where the focus is on the vortex minimization for an arterial blood flow application.

3 OPTIMAL BOUNDARY CONTROL PROBLEMS IN ENERGY SPACES

In this chapter we consider optimal boundary control problems in energy spaces for the Poisson equation which are broadly used in many applications, such as inflow control. The idea to use the energy space as a control or regularization space was already introduced in [49], see also [59], for the case of Dirichlet boundary control problems. The aim of this chapter is to present a unified analysis of the Dirichlet and the Neumann boundary control problem for the Poisson equation without box constraints. Further we shall study a finite element discretization of lowest order. As we will see in this chapter, it turns out that such an approach can be advantageous compared to classical formulations, where the control is realized in $L^2(\Gamma)$, see also [39, 59, 63]. For a detailed discussion on the mathematical and numerical analysis of optimal control problems and their applications we refer, e.g., to [37, 49, 74, 81].

This chapter is organized as follows: In Section 3.1 we consider the Dirichlet boundary control problem for the Poisson equation. First we introduce the Steklov–Poincaré operator for the realization of the semi-norm in the energy space and derive the first order necessary optimality conditions, given by the optimality system. For the variational formulation an elimination of the control is possible and thus a standard structure of saddle point type is obtained. Existence and uniqueness of a corresponding solution is then proven. Further we introduce a corresponding finite element discretization of lowest order for which we prove optimal error estimates. At the end of this section we present several numerical examples which illustrate the obtained theoretical results.

In Section 3.2 we discuss the related model problem for the Neumann boundary control. This case is nevertheless more involved, since for the Poisson equation a suitable scaling has to be introduced in order to guarantee uniqueness of the solution. Due to this reason we consider first the Neumann boundary control problem for the Yukawa equation, since the constraint has a unique solution. Afterwards we study the optimal Neumann boundary control for the Poisson equation. Therefore certain Sobolev spaces with additional constraints on the boundary have to be taken into consideration. It turns out that the Neumann boundary control problem for the Poisson equation is already included in the more general case of the Yukawa equation. The existence and uniqueness for both constraints is investigated. Moreover a finite element discretization of lowest order is introduced. In the case of the Laplace equation it turns out that the primal state, and the Schur complement equation with respect to the boundary data coincide with the Dirichlet boundary control problem, see also [42].

3.1 Optimal Dirichlet boundary control

Let $\Omega \subset \mathbb{R}^n$ $(n = 2,3)$ be a bounded Lipschitz domain with a piecewise smooth boundary $\Gamma = \partial\Omega$. We assume that the desired state $\overline{u} \in L^2(\Omega)$ and the right-hand side $f \in H^{-1}(\Omega)$ are given. As a model problem, we consider an optimal Dirichlet boundary control problem for the Poisson equation in the energy space, which is given as follows: Minimize the cost functional

$$\mathcal{J}(u,z) := \frac{1}{2}\|u-\overline{u}\|^2_{L^2(\Omega)} + \frac{1}{2}\varrho|z|^2_{H^{1/2}(\Gamma)}, \tag{3.1}$$

subject to the constraint

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega, \\ u &= z \qquad \text{on } \Gamma. \end{aligned} \tag{3.2}$$

Additionally we may require that the control satisfies the box constraints

$$z_a \le z \le z_b \qquad \text{a.e. on } \Gamma, \tag{3.3}$$

for given constraints $z_a, z_b \in H^{1/2}(\Gamma)$.

The analysis and corresponding finite element approximation of the optimal control problem (3.1)–(3.3) was investigated in [59]. As it was motived in the introduction we are interested in the construction of robust preconditioners for the above model problem. Due to this reason we only consider the unconstrained model problem.

In the following subsections we explain how the semi-norm in $H^{1/2}(\Gamma)$ is realized and derive the first order necessary optimality conditions. For the corresponding variational formulation it turns out that a formal elimination of the control can be done. We prove the existence and uniqueness of a solution. Afterwards, a lowest order finite element discretization is introduced, for which we prove corresponding optimal error estimates. These results are illustrated by several numerical examples.

3.1.1 Optimality system

Before we derive the optimality system, we need to discuss the realization of the semi-norm in $H^{1/2}(\Gamma)$. In this work we use the representation via the so-called Steklov–Poincaré operator. Other possibilities are the consideration of the Sobolev–Slobodeckii norm or the hypersingular boundary integral operator, see, e.g., [53, 73].

For the derivation of the Steklov–Poincaré operator we split the state into $u = u_f + u_z$ with $u_f \in H^1_0(\Omega)$ and $u_z \in H^1(\Omega)$, being the unique solutions of

$$\begin{aligned} -\Delta u_f &= f \qquad \text{in } \Omega, \\ u_f &= 0 \qquad \text{on } \Gamma, \end{aligned} \tag{3.4}$$

and

$$\begin{aligned} -\Delta u_z &= 0 \qquad \text{in } \Omega, \\ u_z &= z \qquad \text{on } \Gamma. \end{aligned} \tag{3.5}$$

For the boundary value problem (3.5) we obtain the following variational formulation: Find $u_z \in H^1(\Omega)$ with $u_z = z$ on Γ such that

$$\langle \nabla u_z, \nabla q \rangle_{L^2(\Omega)} = 0,$$

for all $q \in H_0^1(\Omega)$. Note this problem has a unique solution $u_z \in H^1(\Omega)$, which is the harmonic extension of the Dirichlet datum $z \in H^{1/2}(\Gamma)$. Green's first formula

$$0 = \langle -\Delta u_z, v \rangle_\Omega = \langle \nabla u_z, \nabla v \rangle_{L^2(\Omega)} - \langle \partial_n u_z, v_{|\Gamma} \rangle_\Gamma,$$

for all $v \in H^1(\Omega)$, then motivates the following definition of the semi-norm

$$|z|^2_{H^{1/2}(\Gamma)} := \langle \partial_n u_z, z \rangle_\Gamma = |u_z|^2_{H^1(\Omega)},$$

for all $z \in H^{1/2}(\Gamma)$. Now, we introduce the Steklov–Poincaré operator S, as a mapping $S : H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$,

$$Sz := \partial_n u_z, \tag{3.6}$$

which realizes the Dirichlet to Neumann map of the boundary value problem (3.5). Consequently, we obtain for the semi-norm, using the Steklov–Poincaré operator (3.6), the representation

$$|z|^2_{H^{1/2}(\Gamma)} = \langle Sz, z \rangle_\Gamma, \tag{3.7}$$

for all $z \in H^{1/2}(\Gamma)$. The properties of the Steklov–Poincaré operator are summarized in the following proposition, see also [59].

Proposition 3.1. *The Steklov–Poincaré operator, defined in* (3.6)*, is self-adjoint, bounded and semi-elliptic in* $H^{1/2}(\Gamma)$*.*

Now, we are in the position to apply the standard theory of optimal control, see, e.g., [37,81], for the derivation of the first order necessary optimality conditions as an equivalent formulation of the optimal control problem (3.1)–(3.2). We derive these conditions in the following. For the boundary value problem (3.5) we introduce the solution operator $\mathcal{H}$, with $u_z = \mathcal{H}z$ for all $z \in H^{1/2}(\Gamma)$. Due to the compact embedding $H^1(\Omega) \hookrightarrow L^2(\Omega)$, the solution operator is a mapping $\mathcal{H} : H^{1/2}(\Gamma) \to L^2(\Omega)$. This gives us the possibility to introduce the reduced cost functional, depending only on the control z, by

$$\widetilde{\mathcal{J}}(z) := \frac{1}{2} \| \mathcal{H}z + u_f - \overline{u} \|^2_{L^2(\Omega)} + \frac{1}{2} \varrho \langle Sz, z \rangle_\Gamma.$$

Thus the optimal control problem (3.1)–(3.2) can be written in the following form: Find the optimal control $\hat{z} \in H^{1/2}(\Gamma)$ such that

$$\widetilde{\mathcal{J}}(\hat{z}) = \min_{z \in H^{1/2}(\Gamma)} \widetilde{\mathcal{J}}(z) = \min_{z \in H^{1/2}(\Gamma)} \left\{ \frac{1}{2} \|\mathcal{H}z + u_f - \overline{u}\|^2_{L^2(\Omega)} + \frac{1}{2} \varrho \langle Sz, z \rangle_\Gamma \right\}. \tag{3.8}$$

Since the Steklov–Poincaré operator S is a bounded and self-adjoint operator we can apply a standard result in optimal control theory, see, e.g., [81, Theorem 2.22], which states that the above reduced minimization problem (3.8) is equivalent to the following operator equation

$$\mathcal{H}^*(\mathcal{H}z + u_f - \overline{u}) + \varrho Sz = 0, \tag{3.9}$$

in the sense of $H^{-1/2}(\Gamma)$, with the adjoint solution operator $\mathcal{H}^* : L^2(\Omega) \to H^{-1/2}(\Gamma)$. Note that in the case of box constraints we would obtain a variational inequality, see, e.g., [59,81]. In order to solve equation (3.9) we need the adjoint solution operator, which is characterized by the following theorem. The proof can be found, e.g., in [59].

Theorem 3.1. *Let $\psi \in L^2(\Omega)$ be arbitrary but fixed. Moreover let $p \in H^1_0(\Omega)$ be the unique solution of the problem*

$$\begin{aligned} -\Delta p &= \psi \qquad \text{in } \Omega, \\ p &= 0 \qquad \text{on } \Gamma. \end{aligned}$$

Then for the adjoint solution operator there holds $\mathcal{H}^*\psi = -\partial_n p$.

Using the relation $u = \mathcal{H}z + u_f$ we obtain

$$\mathcal{H}^*(\mathcal{H}z + u_f - \overline{u}) = \mathcal{H}^*(u - \overline{u}) = -\partial_n p.$$

Moreover we get from (3.9) the relation

$$-\partial_n p + \varrho Sz = 0,$$

in the sense of $H^{-1/2}(\Gamma)$.

As a consequence we obtain the first order necessary optimality conditions which are equivalent to the optimal control problem (3.1)–(3.2). They are given by the following system of coupled partial differential equations, the so-called optimality system

$$\begin{array}{lll} \text{Primal problem} & \text{Adjoint problem} & \text{Optimality condition} \\ -\Delta u = f \quad \text{in } \Omega, & -\Delta p = u - \overline{u} \quad \text{in } \Omega, & \\ u = z \quad \text{on } \Gamma, & p = 0 \qquad \text{on } \Gamma, & -\partial_n p + \varrho Sz = 0 \quad \text{on } \Gamma. \end{array} \tag{3.10}$$

This means, that in the following we consider instead of the optimal control problem (3.1)–(3.2), the optimality system (3.10).

Remark 3.1. *From the optimality system* (3.10) *we obtain the relation*

$$\Delta^2 p = -\Delta u + \Delta \overline{u} = f + \Delta \overline{u}.$$

In the limit case, i.e. when the cost coefficient $\varrho \to 0$*, we then conclude the biharmonic equation of first kind for the adjoint state* p*,*

$$\begin{aligned} \Delta^2 p &= f + \Delta \overline{u} && \text{in } \Omega, \\ p = \partial_n p &= 0 && \text{on } \Gamma. \end{aligned} \tag{3.11}$$

This is a first indication that the optimal Dirichlet boundary control problem is related to the biharmonic equation of first kind. Note that the primal state can be calculated then by $u = \overline{u} - \Delta p$*.*

Remark 3.2. *For the optimality system* (3.10) *it is possible to eliminate the control* z*. This can be done by splitting the primal state* $u = u_f + u_z$ *and using the relation* $Sz = \partial_n u_z$*. Moreover we replace the boundary condition* $u_z = z$ *by the optimality condition* $-\partial_n p + \varrho \partial_n u_z = 0$*, which leads to the optimality system*

$$\begin{aligned} -\Delta u_f &= f && \text{in } \Omega, & -\Delta u_z &= 0 && \text{in } \Omega, & -\Delta p &= u_f + u_z - \overline{u} && \text{in } \Omega, \\ u_f &= 0 && \text{on } \Gamma, & \varrho \partial_n u_z &= \partial_n p && \text{on } \Gamma, & p &= 0 && \text{on } \Gamma. \end{aligned} \tag{3.12}$$

Note that in a post processing step we then find $z = u_{z|\Gamma}$*.*

3.1.2 Variational formulation

Within this subsection we derive the variational formulation for the optimality system (3.12). The reason for choosing (3.12) and not (3.10) is the elimination of the control z. We prove the existence and uniqueness of a solution of the optimality system (3.12) and present a stability estimate.

Based on the optimality system (3.12) we obtain for all test functions $v \in H^1(\Omega)$

$$\langle u_f + u_z - \overline{u}, v \rangle_{L^2(\Omega)} = \langle -\Delta p, v \rangle_\Omega = \langle \nabla p, \nabla v \rangle_{L^2(\Omega)} - \langle \partial_n p, v \rangle_\Gamma,$$

and

$$0 = \varrho \langle -\Delta u_z, v \rangle_\Omega = \varrho \langle \nabla u_z, \nabla v \rangle_{L^2(\Omega)} - \varrho \langle \partial_n u_z, v \rangle_\Gamma = \varrho \langle \nabla u_z, \nabla v \rangle_{L^2(\Omega)} - \langle \partial_n p, v \rangle_\Gamma,$$

where we used the boundary condition $\varrho \partial_n u_z = \partial_n p$. Combining these two equations leads to

$$\langle u_f + u_z, v \rangle_{L^2(\Omega)} + \varrho \langle \nabla u_z, \nabla v \rangle_{L^2(\Omega)} - \langle \nabla p, \nabla v \rangle_{L^2(\Omega)} = \langle \overline{u}, v \rangle_{L^2(\Omega)},$$

and additionally we add the term $\varrho\langle\nabla u_f,\nabla v\rangle_{L^2(\Omega)}$ on the left- and right-hand side, which gives

$$\begin{aligned}\langle u_f+u_z,v\rangle_{L^2(\Omega)}+\varrho\langle\nabla(u_f+u_z),\nabla v\rangle_{L^2(\Omega)}-\langle\nabla p,\nabla v\rangle_{L^2(\Omega)}\\ =\langle\overline{u},v\rangle_{L^2(\Omega)}+\varrho\langle\nabla u_f,\nabla v\rangle_{L^2(\Omega)},\end{aligned}\tag{3.13}$$

for all $v\in H^1(\Omega)$. Now we consider a test function $q\in H_0^1(\Omega)$ for which we obtain

$$\langle f,q\rangle_\Omega=\langle-\Delta u_f,q\rangle_\Omega=\langle\nabla u_f,\nabla q\rangle_{L^2(\Omega)},$$

and

$$0=\langle-\Delta u_z,q\rangle_\Omega=\langle\nabla u_z,\nabla q\rangle_{L^2(\Omega)}.$$

We add these two equations and obtain

$$\langle\nabla(u_f+u_z),\nabla q\rangle_{L^2(\Omega)}=\langle f,q\rangle_\Omega,\tag{3.14}$$

for all $q\in H_0^1(\Omega)$. Again we use for the primal state the relation $u=u_f+u_z$ and obtain with the equations (3.13)–(3.14) the following variational formulation for the optimality system (3.12) in saddle point form. Find $(u,p)\in H^1(\Omega)\times H_0^1(\Omega)$ such that

$$\begin{aligned}a(u,v)-b(v,p)&=\langle\overline{u},v\rangle_{L^2(\Omega)}+\varrho\langle\nabla u_f,\nabla v\rangle_{L^2(\Omega)},\\ b(u,q)&=\langle f,q\rangle_\Omega,\end{aligned}\tag{3.15}$$

for all $(v,q)\in H^1(\Omega)\times H_0^1(\Omega)$. The corresponding bilinear forms are given by

$$a(u,v)=\langle u,v\rangle_{L^2(\Omega)}+\varrho\langle\nabla u,\nabla v\rangle_{L^2(\Omega)},\quad b(v,q)=\langle\nabla v,\nabla q\rangle_{L^2(\Omega)}.\tag{3.16}$$

Next, we prove the existence and uniqueness of a solution of the saddle point formulation (3.15). Therefore we define the kernel of the bilinear form $b(\cdot,\cdot)$ by

$$\operatorname{Ker}B:=\left\{v\in H^1(\Omega):b(v,q)=0\quad\text{for all }q\in H_0^1(\Omega)\right\}\subset H^1(\Omega).$$

Theorem 3.2. *Let $\overline{u}\in L^2(\Omega)$, $f\in H^{-1}(\Omega)$. Then for the variational* (3.15) *there exists a unique solution $(u,p)\in H^1(\Omega)\times H_0^1(\Omega)$ and there holds the stability estimate*

$$\|u\|_{H^1(\Omega)}+\|p\|_{H^1(\Omega)}\le c\left(\|\overline{u}\|_{L^2(\Omega)}+\|f\|_{H^{-1}(\Omega)}\right).\tag{3.17}$$

Proof. The boundedness of the bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ are easy to prove. Further we obtain by the definition of $a(\cdot,\cdot)$ the estimate

$$a(v,v)=\|v\|_{L^2(\Omega)}^2+\varrho|v|_{H^1(\Omega)}^2\ge\min\{1,\varrho\}\|v\|_{H^1(\Omega)}^2=c_1^A\|v\|_{H^1(\Omega)}^2,$$

for all $v \in H^1(\Omega)$. Since Ker $B \subset H^1(\Omega)$ we conclude the Ker B–ellipticity of the bilinear form $a(\cdot,\cdot)$. It remains to prove the inf–sup condition, which is a consequence of Friedrichs inequality, see, e.g., [11], thus

$$\sup_{0\neq v\in H^1(\Omega)} \frac{b(v,q)}{\|v\|_{H^1(\Omega)}} \geq \frac{|q|^2_{H^1(\Omega)}}{\|q\|_{H^1(\Omega)}} \geq (1+c_F^{-2})^{-1}\|q\|_{H^1(\Omega)} = c_S\|q\|_{H^1(\Omega)},$$

for all $q \in H^1_0(\Omega)$. Consequently we can conclude by the standard theory of saddle point problems the existence and uniqueness of the solution $(u,p) \in H^1(\Omega)\times H^1_0(\Omega)$. The stability estimate follows from the estimates above, for the details we refer for instance to [12, Proposition 1.3] or [73, Theorem 3.11]. □

3.1.3 Discretization and error estimates

In the following subsection we introduce a finite element discretization for the variational formulation (3.15). Therefore we denote by $\mathcal{T}_h$ an admissible, shape-regular and globally quasi-uniform triangulation into triangles or tetrahedra of the bounded Lipschitz domain Ω. The elements of $\mathcal{T}_h$ are denoted by T. Further, we introduce the finite dimensional subspaces

$$\mathcal{V}_h = \operatorname{span}\{\varphi_i^1\}_{i=1}^{n_I+n_C} \subset H^1(\Omega), \quad \mathcal{Q}_h = \operatorname{span}\{\varphi_i^1\}_{i=1}^{n_I} \subset H^1_0(\Omega), \tag{3.18}$$

both of piecewise linear and globally continuous shape functions φ_i^1. Note that $n_I = \dim\, \mathcal{Q}_h$ is the number of interior degrees of freedom, and, n_C is the number of degrees of freedom on the boundary with $\dim\, \mathcal{V}_h = n_I + n_C$.

Based on the continuous variational formulation (3.15) we obtain the following discrete problem. Find $(u_h, p_h) \in \mathcal{V}_h \times \mathcal{Q}_h$ such that

$$\begin{aligned} a(u_h, v_h) - b(v_h, p_h) &= \langle \overline{u}, v_h\rangle_{L^2(\Omega)} + \varrho\langle \nabla u_f, \nabla v_h\rangle_{L^2(\Omega)}, \\ b(u_h, q_h) &= \langle f, q_h\rangle_\Omega, \end{aligned} \tag{3.19}$$

for all $(v_h, q_h) \in \mathcal{V}_h \times \mathcal{Q}_h$. For the existence and uniqueness of a discrete solution the following theorem is valid.

Theorem 3.3. *Let $\overline{u} \in L^2(\Omega)$, $f \in H^{-1}(\Omega)$. Then for the discrete variational formulation (3.19) there exists a unique solution $(u_h, p_h) \in \mathcal{V}_h \times \mathcal{Q}_h$ and there holds the stability estimate*

$$\|u_h\|_{H^1(\Omega)} + \|p_h\|_{H^1(\Omega)} \leq c\left(\|\overline{u}\|_{L^2(\Omega)} + \|f\|_{H^{-1}(\Omega)}\right).$$

Further, the quasi-optimal error estimate

$$\|u-u_h\|_{H^1(\Omega)} + \|p-p_h\|_{H^1(\Omega)} \leq c\left(\inf_{v_h\in\mathcal{V}_h}\|u-v_h\|_{H^1(\Omega)} + \inf_{q_h\in\mathcal{Q}_h}\|p-q_h\|_{H^1(\Omega)}\right),$$

is satisfied, where $(u,p) \in H^1(\Omega)\times H^1_0(\Omega)$ is the exact solution of (3.15).

Proof. The proof of existence and uniqueness, as well as the stability estimate is similar to the proof of Theorem 3.2. Note that the discrete inf–sup condition

$$\sup_{0\neq v_h\in\mathcal{V}_h}\frac{b(v_h,q_h)}{\|v_h\|_{H^1(\Omega)}}\geq\widetilde{c}_S\|q_h\|_{H^1(\Omega)},$$

for all $q_h\in\mathcal{Q}_h$ can be shown by the same arguments as in the continuous case, due to the inclusion $\mathcal{Q}_h\subset\mathcal{V}_h$. For the error estimate we obtain from (3.15) and (3.19) the Galerkin orthogonality

$$\begin{aligned} a(u-u_h,v_h)-b(v_h,p-p_h) &= 0,\\ b(u-u_h,q_h) &= 0, \end{aligned} \tag{3.20}$$

for all $(v_h,q_h)\in\mathcal{V}_h\times\mathcal{Q}_h$. Let us consider arbitrary $(\widetilde{u}_h,\widetilde{p}_h)\in\mathcal{V}_h\times\mathcal{Q}_h$, for which we obtain

$$\begin{aligned} a(\widetilde{u}_h-u_h,v_h)-b(v_h,\widetilde{p}_h-p_h) &= a(\widetilde{u}_h-u,v_h)-b(v_h,\widetilde{p}_h-p),\\ b(\widetilde{u}_h-u_h,q_h) &= b(\widetilde{u}_h-u,q_h), \end{aligned}$$

for all $(v_h,q_h)\in\mathcal{V}_h\times\mathcal{Q}_h$. For this problem we can apply the stability estimate, c.f. [12] and [73, Theorem 8.7], and, obtain the estimate

$$\|\widetilde{u}_h-u_h\|_{H^1(\Omega)}+\|\widetilde{p}_h-p_h\|_{H^1(\Omega)}\leq c\left(\|\widetilde{u}_h-u\|_{H^1(\Omega)}+\|\widetilde{p}_h-p\|_{H^1(\Omega)}\right).$$

The desired error estimate follows then by applying the triangle inequality and corresponding infima. □

As a direct consequence of the above error estimate we obtain with the standard approximation property, see, e.g., [73, Theorem 9.10], the following error estimate

$$\|u-u_h\|_{H^1(\Omega)}+\|p-p_h\|_{H^1(\Omega)}\leq ch^{s-1}\left(|u|_{H^s(\Omega)}+|p|_{H^s(\Omega)}\right), \tag{3.21}$$

where $(u,p)\in H^s(\Omega)\times H^1_0(\Omega)\cap H^s(\Omega)$ is the exact solution for some $s\in[1,2]$.

Next, we prove by a duality argument (Aubin–Nitsche trick) an error estimate in the $L^2(\Omega)$ norm. Therefore we consider the following adjoint problem: Find $(w,r)\in H^1(\Omega)\times H^1_0(\Omega)$ such that

$$\begin{aligned} a(w,v)+b(v,r) &= \langle u-u_h,v\rangle_{L^2(\Omega)},\\ -b(w,q) &= \langle p-p_h,q\rangle_{L^2(\Omega)}, \end{aligned} \tag{3.22}$$

for all $(v,q)\in H^1(\Omega)\times H^1_0(\Omega)$.

Theorem 3.4. *Let the assumptions of Theorem* 3.3 *be satisfied and let us assume in addition that for the problem* (3.22) *the estimate*

$$\|w\|_{H^2(\Omega)}+\|r\|_{H^2(\Omega)}\leq c\left(\|u-u_h\|_{L^2(\Omega)}+\|p-p_h\|_{L^2(\Omega)}\right),$$

is valid, which holds if the domain is either convex or has a smooth boundary. Then there holds the error estimate

$$\|u-u_h\|_{L^2(\Omega)} + \|p-p_h\|_{L^2(\Omega)} \leq ch^s \left(|u|_{H^s(\Omega)} + |p|_{H^s(\Omega)} \right), \tag{3.23}$$

where $(u,p) \in H^s(\Omega) \times H_0^1(\Omega) \cap H^s(\Omega)$ *is the exact solution and* $s \in [1,2]$.

Proof. Let us denote by $Q^1_{\mathcal{V}_h}$ and $Q^1_{\mathcal{Q}_h}$ the standard $H^1(\Omega)$ projection onto the finite element space $\mathcal{V}_h$ and $\mathcal{Q}_h$, respectively. Applying the Galerkin orthogonality (3.20) and the definition of the auxiliary problem (3.22) leads to the estimate

$$\begin{aligned}
\|u-u_h\|^2_{L^2(\Omega)} + \|p-p_h\|^2_{L^2(\Omega)} &= a(w,u-u_h) + b(u-u_h,r) - b(w,p-p_h) \\
&= a(u-u_h, w-Q^1_{\mathcal{V}_h} w) - b(w-Q^1_{\mathcal{V}_h} w, p-p_h) + b(u-u_h, r-Q^1_{\mathcal{Q}_h} r) \\
&\leq c_2^A \|u-u_h\|_{H^1(\Omega)} \|w-Q^1_{\mathcal{V}_h} w\|_{H^1(\Omega)} + \|w-Q^1_{\mathcal{V}_h} w\|_{H^1(\Omega)} \|p-p_h\|_{H^1(\Omega)} \\
&\quad + \|u-u_h\|_{H^1(\Omega)} \|r-Q^1_{\mathcal{Q}_h} r\|_{H^1(\Omega)} \\
&\leq ch \left(|w|_{H^2(\Omega)} + |r|_{H^2(\Omega)} \right) \left(\|u-u_h\|_{H^1(\Omega)} + \|p-p_h\|_{H^1(\Omega)} \right) \\
&\leq ch \left(\|u-u_h\|_{L^2(\Omega)} + \|p-p_h\|_{L^2(\Omega)} \right) \left(\|u-u_h\|_{H^1(\Omega)} + \|p-p_h\|_{H^1(\Omega)} \right).
\end{aligned}$$

The assertion then follows by applying the error estimate (3.21). □

Remark 3.3. *It is important to mention that we do not have an exact representation for the right-hand side in* (3.19)*, since it involves the solution* u_f*. In particular we consider* $u_{f,h} \in \mathcal{Q}_h$ *which denotes the finite element approximation of* $u_f \in H_0^1(\Omega)$*. Thus a perturbed problem has to be analyzed. By using Strang lemmata, see, e.g.* [73, Theorem 8.2, 8.3] *we can prove optimal error estimates, i.e. the above error estimates remain valid.*

It remains to answer the question about the order of convergence of the control z on the boundary. As a direct consequence we obtain from the trace theorem and (3.21) the following error estimate for the control,

$$\|z-z_h\|_{H^{1/2}(\Gamma)} \leq ch^{s-1} \left(|u|_{H^s(\Omega)} + |p|_{H^s(\Omega)} \right),$$

assuming the exact solution (u,p) is regular enough, and $s \in [1,2]$. It is known that, in the case of more regular solutions, the estimate above is not optimal and one can gain in the convergence rate an additional factor of up to $1/2$. The same argument is valid for the $L^2(\Gamma)$ error of the control on the boundary.

In order to prove an error estimate for the control in the $L^2(\Gamma)$ with an order of $h^{s-1/2}$ is quite simple and shown in the following lemma.

Lemma 3.1. *Let the assumptions of Theorem 3.4 be satisfied, then there holds the error estimate*

$$\|z - z_h\|_{L^2(\Gamma)} \leq ch^{s-1/2} \left(|u|_{H^s(\Omega)} + |p|_{H^s(\Omega)} \right) \tag{3.24}$$

where $(u,p) \in H^s(\Omega) \times H_0^1(\Omega) \cap H^s(\Omega)$ *is the exact solution for some* $s \in [1,2]$.

Proof. For any $v \in H^1(\Omega)$ we have, see, e.g., [11, Theorem 1.6.6], the estimate

$$\|v_{|\Gamma}\|_{L^2(\Gamma)} \leq c\|v\|_{L^2(\Omega)}^{1/2} \|v\|_{H^1(\Omega)}^{1/2},$$

for all $v \in H^1(\Omega)$. This result we can apply to the error of the control $z - z_h = (u - u_h)_{|\Gamma}$, which leads together with the error estimates (3.21) and (3.23) to

$$\|z - z_h\|_{L^2(\Gamma)} \leq c\|u - u_h\|_{L^2(\Omega)}^{1/2} \|u - u_h\|_{H^1(\Omega)}^{1/2} \leq ch^{s-1/2} \left(|u|_{H^s(\Omega)} + |p|_{H^s(\Omega)} \right),$$

which concludes the proof. □

As it was mentioned, both error estimates in $H^{1/2}(\Gamma)$ and $L^2(\Gamma)$, can be improved when the solution is regular enough by a factor of up to $1/2$. Recently, in [3, 55] error estimates in the $L^2(\Gamma)$ norm in the context of a Neumann boundary value problem and Lagrange multipliers on the boundary were shown, which are optimal up to a logarithmic factor, in the case of piecewise linear and globally continuous finite elements. We might use these ideas for the proof of optimal error estimates for the control z. This can be seen as an interesting and challenging future work.

Finally, we would like to comment on the linear system for the discrete variational formulation (3.19) of the optimal Dirichlet boundary control problem. As mentioned before we consider the piecewise linear and globally continuous finite element spaces $\mathcal{V}_h$ and $\mathcal{Q}_h$, defined in (3.18), with $\dim \mathcal{V}_h = n_I + n_C$ and $\dim \mathcal{Q}_h = n_I$, respectively. We introduce mass, stiffness matrices and right-hand sides by

$$M_h[j,i] = \langle \varphi_i^1, \varphi_j^1 \rangle_{L^2(\Omega)},\ A_h[j,i] = \langle \nabla \varphi_i^1, \nabla \varphi_j^1 \rangle_{L^2(\Omega)},\ \underline{\overline{u}}[i] = \langle \overline{u}, \varphi_i^1 \rangle_{L^2(\Omega)},\ f_I[\ell] = \langle f, \varphi_\ell^1 \rangle_\Omega,$$

for all $i,j = 1,\ldots,n_I + n_C$ and $\ell = 1,\ldots,n_I$. Note that $M_h, A_h \in \mathbb{R}^{(n_I+n_C)\times(n_I+n_C)}$, $\underline{\overline{u}} \in \mathbb{R}^{n_I+n_C}$ and $\underline{f}_I \in \mathbb{R}^{n_I}$. By separation of interior and boundary degrees of freedom, we can write the stiffness matrix as

$$A_h = \begin{pmatrix} A_{II} & A_{IC} \\ A_{CI} & A_{CC} \end{pmatrix},$$

with $A_{II} \in \mathbb{R}^{n_I \times n_I}$, $A_{IC} = A_{CI}^\top \in \mathbb{R}^{n_I \times n_C}$ and $A_{CC} \in \mathbb{R}^{n_C \times n_C}$. Further, we introduce

$$A_{IA} = A_{AI}^\top = \begin{pmatrix} A_{II} & A_{IC} \end{pmatrix}.$$

The discrete solution vector $\underline{u}_f \in \mathbb{R}^{n_I}$ of the homogeneous Dirichlet (3.4) problem is then given by $\underline{u}_f = A_{II}^{-1}\underline{f}_I$ and consequently is discrete vector of the first equation of (3.19) given by

$$\widetilde{\underline{f}} = \underline{\overline{u}} + \varrho A_{AI}\underline{u}_f = \underline{\overline{u}} + \varrho A_{AI}A_{II}^{-1}\underline{f}_I,$$

with $\widetilde{\underline{f}} \in \mathbb{R}^{n_I+n_C}$. The equivalent linear system for the discrete variational formulation (3.19) reads then

$$\begin{pmatrix} M_h + \varrho A_h & -A_{AI} \\ A_{IA} & \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p}_I \end{pmatrix} = \begin{pmatrix} \widetilde{\underline{f}} \\ \underline{f}_I \end{pmatrix}, \tag{3.25}$$

where the individual blocks are of the following dimensions, $M_h + \varrho A_h \in \mathbb{R}^{(n_I+n_C)\times(n_I+n_C)}$ and $A_{IA} = A_{AI}^\top \in \mathbb{R}^{n_I\times(n_I+n_C)}$.

Since the unknown $\underline{u}$ has its degrees of freedom corresponding to interior and boundary degrees of freedom, we can separate them and rewrite $\underline{u} = (\underline{u}_I, \underline{u}_C)^\top$ with $\underline{u}_I \in \mathbb{R}^{n_I}$ and $\underline{u}_C \in \mathbb{R}^{n_C}$. Note that $\underline{u}_C$ is the solution vector for the control z_h. Splitting the linear system into interior and boundary degrees of freedom leads to

$$\begin{pmatrix} M_{II} + \varrho A_{II} & M_{IC} + \varrho A_{IC} & -A_{II} \\ M_{CI} + \varrho A_{CI} & M_{CC} + \varrho A_{CC} & -A_{CI} \\ A_{II} & A_{IC} & \end{pmatrix} \begin{pmatrix} \underline{u}_I \\ \underline{u}_C \\ \underline{p}_I \end{pmatrix} = \begin{pmatrix} \widetilde{\underline{f}}_I \\ \widetilde{\underline{f}}_C \\ \underline{f}_I \end{pmatrix},$$

or, by a simple reordering of the variables, as

$$\begin{pmatrix} M_{II} + \varrho A_{II} & -A_{II} & M_{IC} + \varrho A_{IC} \\ A_{II} & & A_{IC} \\ M_{CI} + \varrho A_{CI} & -A_{CI} & M_{CC} + \varrho A_{CC} \end{pmatrix} \begin{pmatrix} \underline{u}_I \\ \underline{p}_I \\ \underline{u}_C \end{pmatrix} = \begin{pmatrix} \widetilde{\underline{f}}_I \\ \underline{f}_I \\ \widetilde{\underline{f}}_C \end{pmatrix}. \tag{3.26}$$

In the following we derive the Schur complement system with respect to the control, i.e. $\underline{u}_C$. From the linear system (3.26) we obtain

$$\underline{u}_I = A_{II}^{-1}\left[\underline{f}_I - A_{IC}\underline{u}_C\right],$$

and

$$\begin{aligned} \underline{p}_I &= A_{II}^{-1}\left[A_{II}^{-1}\left[M_{II}\underline{u}_I + M_{IC}\underline{u}_C\right] + \varrho\left[A_{II}\underline{u}_I + A_{IC}\underline{u}_C\right] - \widetilde{\underline{f}}_I\right] \\ &= A_{II}^{-1}\left[M_{II}A_{II}^{-1}\underline{f}_I - M_{II}A_{II}^{-1}A_{IC}\underline{u}_C + M_{IC}\underline{u}_C + \varrho\underline{f}_I - \widetilde{\underline{f}}_I\right], \end{aligned}$$

which results in the Schur complement system

$$\begin{aligned} \Big[\big[M_{CC} - M_{CI}A_{II}^{-1}A_{IC} - A_{CI}A_{II}^{-1}M_{IC} + A_{CI}A_{II}^{-1}M_{II}A_{II}^{-1}A_{IC}\big] \\ + \varrho\big[A_{CC} - A_{CI}A_{II}^{-1}A_{IC}\big]\Big]\underline{u}_C \\ = \big[A_{CI}A_{II}^{-1}M_{II} - M_{CI}\big]A_{II}^{-1}\underline{f}_I + \widetilde{\underline{f}}_C - A_{CI}A_{II}^{-1}\widetilde{\underline{f}}_I. \end{aligned} \tag{3.27}$$

The corresponding Schur complement matrix is then given by

$$\begin{aligned} T_h + \varrho S_h = M_{CC} - M_{CI}A_{II}^{-1}A_{IC} - A_{CI}A_{II}^{-1}A_{IC} + A_{CI}A_{II}^{-1}M_{II}A_{II}^{-1}A_{IC} \\ + \varrho\big[A_{CC} - A_{CI}A_{II}^{-1}A_{IC}\big]. \end{aligned} \tag{3.28}$$

As we will see in the forthcoming chapters this Schur complement consists of the Schur complement T_h of the biharmonic equation and $S_h = A_{CC} - A_{CI}A_{II}^{-1}A_{IC}$, which is the Schur complement of the stiffness matrix of the Laplace equation, or in other words, the discrete Galerkin matrix of the Steklov–Poincaré operator, see [39, 59]. Note that the Schur complement system (3.27) with the unknown $\underline{u}_C$ is an equation for the discrete control z_h, since we have the isomorphism $z_h \leftrightarrow \underline{u}_C \in \mathbb{R}^{n_C}$.

3.1.4 Numerical results

In this subsection we emphasize on numerical examples for the energy space approach for the optimal Dirichlet boundary control problem (3.1)–(3.2). We consider therefore the mixed discrete variational formulation (3.19) with piecewise linear and globally continuous finite elements, being equivalent to the linear system (3.25). As a computational domain we consider the cube $\Omega = (0, \frac{1}{2})^n$, for both, $n = 2, 3$. The number of elements for different refinement levels are given by, $N = 4^{L+1}$ for $n = 2$ and by $N = 12 \cdot 8^L$ elements for $n = 3$, where L denotes the refinement level. Note that we refine the mesh by the midpoints of the element edges. As given data we consider

$$\varrho = 1, \qquad f = 0, \qquad \overline{u} = \left(\sum_{i=1}^{n} (x_i(x_i - 1/2) + 1)^2 \right)^{1/2}. \tag{3.29}$$

We consider the numerical solution on refinement level $L = 9$ and $L = 6$, denoted by $(u_{h_9}, p_{h_9}, z_{h_9})$ and $(u_{h_6}, p_{h_6}, z_{h_6})$, for $n = 2, 3$ as the reference solution for the computation of the errors. Note that the estimated order of convergence (eoc) might be slightly higher than predicted by the theory, which is due to the reference solution.

In the following we present errors in the L^2 norm for all unknowns. As proven in Section 3.1 we expect second order of convergence for the primal and adjoint state, (u_h, p_h). For the control we expect at least $3/2$ as order of convergence in the $L^2(\Gamma)$ norm, see error estimate (3.24).

In Table 3.1 and Table 3.2 we present the corresponding errors and estimated order of convergence, for the two- and three-dimensional model problem, respectively. We observe for both examples optimal orders of convergence for the primal and adjoint state, which illustrates the error estimates. In particular we obtain second order for the control on the boundary, which is an order of $1/2$ more as shown in (3.24).

L	Dofs	$\|u_{h_9} - u_h\|_{L^2(\Omega)}$	eoc	$\|p_{h_9} - p_h\|_{L^2(\Omega)}$	eoc	$\|z_{h_9} - z_h\|_{L^2(\Gamma)}$	eoc
0	6	1.99488 e–05	–	2.05235 e–05	–	7.10323 e–05	–
1	18	5.45896 e–06	1.87	2.05682 e–05	0.00	2.29918 e–05	1.63
2	66	4.02234 e–06	0.44	5.25710 e–06	1.97	1.36912 e–05	0.75
3	258	9.96254 e–07	2.01	1.45931 e–06	1.85	3.78661 e–06	1.85
4	1 026	2.39679 e–07	2.06	3.83092 e–07	1.93	9.78068 e–07	1.95
5	4 098	5.91005 e–08	2.02	9.72388 e–08	1.98	2.49145 e–07	1.97
6	16 386	1.46358 e–08	2.01	2.42042 e–08	2.01	6.27297 e–08	1.99
7	65 538	3.53393 e–09	2.05	5.81441 e–09	2.06	1.54177 e–08	2.02
8	262 146	7.70841 e–10	2.20	1.20565 e–09	2.27	3.44813 e–09	2.16
expected			2.00		2.00		1.50

Table 3.1: Errors and eoc for optimal Dirichlet boundary control, $n = 2$.

L	Dofs	$\|u_{h_6} - u_h\|_{L^2(\Omega)}$	eoc	$\|p_{h_6} - p_h\|_{L^2(\Omega)}$	eoc	$\|z_{h_6} - z_h\|_{L^2(\Gamma)}$	eoc
0	8	2.43619 e–05	–	3.43965 e–05	–	1.27627 e–04	–
1	28	1.21087 e–05	1.01	2.24667 e–05	0.61	4.84704 e–05	1.40
2	152	4.70483 e–06	1.36	1.00057 e–05	1.17	2.71253 e–05	0.84
3	1 072	1.47357 e–06	1.67	3.19439 e–06	1.65	8.07966 e–06	1.75
4	8 288	4.10507 e–07	1.84	8.58818 e–07	1.90	2.15955 e–06	1.90
5	65 728	9.29795 e–08	2.14	1.83784 e–07	2.22	4.70650 e–07	2.20
expected			2.00		2.00		1.50

Table 3.2: Errors and eoc for optimal Dirichlet boundary control, $n = 3$.

3.2 Optimal Neumann boundary control

Within this section we apply the energy space control approach to optimal Neumann boundary control problems. First we derive the first order necessary optimality conditions for the Yukawa and the Poisson equation. As we will see, the second constraint is more involved. Afterwards we prove for both constraints the existence and uniqueness of a solution of the corresponding optimality system. In the case of the Laplace equation, it turns out that the Schur complement, with respect to the boundary, is the same as in the case of the optimal Dirichlet boundary control problem, see Section 3.1. Further, we prove that in this particular case the primal states of Dirichlet and Neumann boundary control coincide.

In the following $\Omega \subset \mathbb{R}^n$ $(n = 2,3)$ shall always denote a bounded Lipschitz domain with a piecewise smooth boundary $\Gamma = \partial\Omega$.

3.2.1 Optimality system – Yukawa

Let us consider the desired state $\overline{u} \in L^2(\Omega)$ and the right-hand side $f \in \widetilde{H}^{-1}(\Omega)$. As a model problem we consider an optimal Neumann boundary control problem for the Yukawa equation without box constraints, where we assume the cost coefficient $\varrho > 0$, which is given as follows: Minimize the cost functional

$$\mathcal{J}(u,z) := \frac{1}{2}\|u-\overline{u}\|^2_{L^2(\Omega)} + \frac{1}{2}\varrho\|z\|^2_{H^{-1/2}(\Gamma)}, \tag{3.30}$$

subject to the constraint, with $\kappa > 0$,

$$\begin{aligned} -\Delta u + \kappa u &= f \qquad \text{in } \Omega, \\ \partial_n u &= z \qquad \text{on } \Gamma. \end{aligned} \tag{3.31}$$

Note that the assumption $\kappa > 0$ has to be made in order to ensure the unique solvability of the Neumann boundary value problem (3.31). The limit case, i.e. $\kappa = 0$, will be treated later.

In the following, we discuss the realization of the $H^{-1/2}(\Gamma)$ norm, which is from its idea strongly related to the realization of the semi-norm in $H^{1/2}(\Gamma)$ for the Dirichlet boundary control, see Section 3.1. Further we derive the first order necessary optimality conditions.

We split the state into $u = u_f + u_z$, with $u_f \in H^1(\Omega)$ and $u_z \in H^1(\Omega)$, being the unique solutions of

$$\begin{aligned} -\Delta u_f + \kappa u_f &= f \qquad \text{in } \Omega, \\ \partial_n u_f &= 0 \qquad \text{on } \Gamma, \end{aligned}$$

and

$$\begin{aligned} -\Delta u_z + \kappa u_z &= 0 \qquad \text{in } \Omega, \\ \partial_n u_z &= z \qquad \text{on } \Gamma. \end{aligned} \tag{3.32}$$

Multiplying (3.32) with a test function $v \in H^1(\Omega)$ and applying integration by parts leads to

$$0 = \langle -\Delta u_z + \kappa u_z, v\rangle_\Omega = \langle \nabla u_z, \nabla v\rangle_{L^2(\Omega)} + \kappa\langle u_z, v\rangle_{L^2(\Omega)} - \langle \partial_n u_z, v_{|\Gamma}\rangle_\Gamma.$$

This leads to the following variational formulation. Find $u_z \in H^1(\Omega)$ such that

$$\langle \nabla u_z, \nabla v\rangle_{L^2(\Omega)} + \kappa\langle u_z, v\rangle_{L^2(\Omega)} = \langle z, v_{|\Gamma}\rangle_\Gamma, \tag{3.33}$$

for all $v \in H^1(\Omega)$. For a given $z \in H^{-1/2}(\Gamma)$ this formulation has clearly a unique solution $u_z \in H^1(\Omega)$. We introduce the following weighted norm

$$\|v\|_{H^1(\Omega),\kappa} := \left(|v|^2_{H^1(\Omega)} + \kappa\|v\|^2_{L^2(\Omega)}\right)^{1/2},$$

for all $v \in H^1(\Omega)$. The variational formulation (3.33) then motivates the following definition of the norm

$$\|z\|^2_{H^{-1/2}(\Gamma)} := \langle z, u_{z|\Gamma}\rangle_\Gamma = |u_z|^2_{H^1(\Omega)} + \kappa \|u_z\|^2_{L^2(\Omega)} = \|u_z\|^2_{H^1(\Omega),\kappa},$$

for all $z \in H^{-1/2}(\Gamma)$.

Now, we introduce the inverse Steklov–Poincaré operator, or Poincaré–Steklov operator, as a mapping $S_\kappa^{-1} : H^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$,

$$S_\kappa^{-1} z := u_{z|\Gamma}, \tag{3.34}$$

which realizes the Neumann to Dirichlet map for the boundary value problem (3.32). Consequently we obtain for the norm in $H^{-1/2}(\Gamma)$, using the inverse Steklov–Poincaré operator (3.34), the representation

$$\|z\|^2_{H^{-1/2}(\Gamma)} = \langle S_\kappa^{-1} z, z\rangle_\Gamma. \tag{3.35}$$

We summarize the properties of the inverse Steklov–Poincaré operator in the following remark.

Proposition 3.2. *The inverse Steklov–Poincaré operator S_κ^{-1}, defined in* (3.34)*, is self-adjoint, bounded and elliptic in $H^{-1/2}(\Gamma)$.*

As for the optimal Dirichlet boundary control problem, see Section 3.1, we derive the first order necessary optimality condition as an equivalent formulation for the optimal control problem (3.30)–(3.31), see also [81]. For the boundary value problem (3.32) we introduce the solution operator $\mathcal{H}_\kappa$, with $u_z = \mathcal{H}_\kappa z$ for all $z \in H^{-1/2}(\Gamma)$. Due to the compact embedding of $H^1(\Omega) \hookrightarrow L^2(\Omega)$, the solution operator is then a mapping $\mathcal{H}_\kappa : H^{-1/2}(\Gamma) \to L^2(\Omega)$. This gives us the possibility to introduce the reduced cost functional as

$$\widetilde{\mathcal{J}}(z) := \frac{1}{2}\|\mathcal{H}_\kappa z + u_f - \overline{u}\|^2_{L^2(\Omega)} + \frac{1}{2}\varrho \langle S_\kappa^{-1} z, z\rangle_\Gamma.$$

Thus, the optimal control problem (3.30)–(3.31) can be stated in the following form: Find the optimal control $\hat{z} \in H^{-1/2}(\Gamma)$ which satisfies

$$\widetilde{\mathcal{J}}(\hat{z}) = \min_{z \in H^{-1/2}(\Gamma)} \widetilde{\mathcal{J}}(z) = \min_{z \in H^{-1/2}(\Gamma)} \left\{ \frac{1}{2}\|\mathcal{H}_\kappa z + u_f - \overline{u}\|^2_{L^2(\Omega)} + \frac{1}{2}\varrho \langle S_\kappa^{-1} z, z\rangle_\Gamma \right\}. \tag{3.36}$$

Since the inverse Steklov–Poincaré operator S_κ^{-1} is a bounded and self-adjoint operator we can apply a standard result in optimal control theory, see for e.g. [81, Theorem 2.22], which states that the above reduced minimization problem (3.36) is equivalent to the following operator equation

$$\mathcal{H}_\kappa^*(\mathcal{H}_\kappa z + u_f - \overline{u}) + \varrho S_\kappa^{-1} z = 0, \tag{3.37}$$

in the sense of $H^{1/2}(\Gamma)$ with the adjoint solution operator $\mathcal{H}_\kappa^* : L^2(\Omega) \to H^{1/2}(\Gamma)$. In order to solve this equation we need to characterize the adjoint solution operator for which the following theorem holds.

Theorem 3.5. *Let $\psi \in L^2(\Omega)$ be arbitrary but fixed. Moreover let $p \in H^1(\Omega)$ be the unique solution of the problem*

$$\begin{aligned} -\Delta p + \kappa p &= \psi && \text{in } \Omega, \\ \partial_n p &= 0 && \text{on } \Gamma. \end{aligned}$$

Then for the adjoint solution operator there holds $\mathcal{H}_\kappa^* \psi = p_{|\Gamma}$.

Proof. For the homogeneous problem (3.32) we obtain the following variational formulations: Find $u_z \in H^1(\Omega)$ such that

$$\langle \nabla u_z, \nabla v \rangle_{L^2(\Omega)} + \kappa \langle u_z, v \rangle_{L^2(\Omega)} = \langle z, v \rangle_\Gamma,$$

for all $v \in H^1(\Omega)$. Moreover, we want to find $p \in H^1(\Omega)$ such that

$$\langle \nabla p, \nabla v \rangle_{L^2(\Omega)} + \kappa \langle p, v \rangle_{L^2(\Omega)} = \langle \psi, v \rangle_{L^2(\Omega)},$$

for all $v \in H^1(\Omega)$. Now we set in the first problem $v = p$ and for the second one $v = u_z$. Consequently we obtain, due to symmetry, the equality

$$\langle \mathcal{H}_\kappa z, \psi \rangle_{L^2(\Omega)} = \langle u_z, \psi \rangle_{L^2(\Omega)} = \langle \nabla p, \nabla u_z \rangle_{L^2(\Omega)} + \kappa \langle p, u_z \rangle_{L^2(\Omega)} = \langle z, p \rangle_\Gamma = \langle z, \mathcal{H}_\kappa^* \psi \rangle_\Gamma,$$

for all $\psi \in L^2(\Omega)$ and $z \in H^{-1/2}(\Gamma)$, which concludes the proof. □

Using the relation $u = \mathcal{H}_\kappa z + u_f$ we obtain

$$\mathcal{H}_\kappa^* (\mathcal{H}_\kappa z + u_f - \overline{u}) = \mathcal{H}_\kappa^* (u - \overline{u}) = p_{|\Gamma}.$$

Moreover we obtain from (3.37) the relation

$$p|_\Gamma + \varrho S_\kappa^{-1} z = 0,$$

in the sense of $H^{1/2}(\Gamma)$.

The first order necessary optimality conditions, which are equivalent to the optimal control problem (3.30)–(3.31), are given by the following optimality system,

$$\begin{array}{lll} \text{Primal problem} & \text{Adjoint problem} & \text{Optimality condition} \\ -\Delta u + \kappa u = f \quad \text{in } \Omega, & -\Delta p + \kappa p = u - \overline{u} \quad \text{in } \Omega, & \\ \partial_n u = z \quad \text{on } \Gamma, & \partial_n p = 0 \quad \text{on } \Gamma, & p + \varrho S_\kappa^{-1} z = 0 \quad \text{on } \Gamma. \end{array} \tag{3.38}$$

Remark 3.4. *As in the Dirichlet boundary control case we can eliminate the control z by splitting $u = u_f + u_z$. With the relation $S_\kappa^{-1} z = u_z|_\Gamma$ we can replace the boundary condition $\partial_n u_z = z$ by $p + \varrho u_z = 0$. This leads to the following optimality system*

$$\begin{aligned} -\Delta u_f + \kappa u_f &= f && \text{in } \Omega, & -\Delta u_z + \kappa u_z &= 0 && \text{in } \Omega, \\ \partial_n u_f &= 0 && \text{on } \Gamma, & \varrho u_z &= -p && \text{on } \Gamma, \\ & & & -\Delta p + \kappa p &= u_f + u_z - \overline{u} && \text{in } \Omega, \\ & & & \partial_n p &= 0 && \text{on } \Gamma. \end{aligned} \tag{3.39}$$

Note the control can be found in a post processing step via $z = \partial_n u_z$.

3.2.2 Optimality system – Poisson

In the previous subsection we considered the optimal Neumann boundary control for the Yukawa equation. Within this subsection we consider the optimal Neumann boundary control problem for the Poisson equation. In this particular case we have to be more careful with the function spaces and equivalent norms, since the state is by the constraint only unique up to an additive constant.

Let us consider the desired state $\overline{u} \in L^2(\Omega)$ and a right-hand side $f \in \widetilde{H}_*^{-1}(\Omega)$, where

$$\widetilde{H}_*^{-1}(\Omega) = \big\{ f \in \widetilde{H}^{-1}(\Omega) : \langle f, 1 \rangle_\Omega = 0 \big\}.$$

As a model problem we consider an optimal Neumann boundary control problem for the Poisson equation without box constraints, with a cost coefficient $\varrho > 0$, which is given as follows: Minimize the cost functional

$$\mathcal{J}(u,z) := \frac{1}{2} \|u - \overline{u}\|_{L^2(\Omega)}^2 + \frac{1}{2} \varrho \|z\|_{H^{-1/2}(\Gamma)}^2, \tag{3.40}$$

subject to the constraint

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega, \\ \partial_n u &= z \qquad \text{on } \Gamma. \end{aligned} \tag{3.41}$$

Note, that we have to assume a solvability condition for the control z and use the correct function space for the state u in order to guarantee the uniqueness of a solution of the constraint (3.41) for a given z. More precisely, for the constraint (3.41), we have to ensure the solvability condition

$$\langle z, 1 \rangle_\Gamma + \langle f, 1 \rangle_\Omega = 0,$$

from which we obtain $\langle z, 1 \rangle_\Gamma = 0$, since $f \in \widetilde{H}_*^{-1}(\Omega)$. This motivates the definition of the space

$$H_*^{-1/2}(\Gamma) := \Big\{ \psi \in H^{-1/2}(\Gamma) : \langle \psi, 1 \rangle_\Gamma = 0 \Big\},$$

for the control z. Before we identify the corresponding dual space, a different type of scaling has to be introduced. Note, that this is only for theoretical reasons necessary and not needed for computations, which will be more clear later. Without loss of generality, let us assume that $\operatorname{diam}(\Omega) < 1$ for $n = 2$, which can be realized by simple scaling. Further, we introduce for a given $\psi \in H^{-1/2}(\Gamma)$ the single layer boundary integral operator $V : H^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$ by

$$(V\psi)(x) := \int_\Gamma U^*(x,y) \psi(y) \, ds_y,$$

for $x \in \Gamma$, where

$$U^*(x,y) = \begin{cases} -\dfrac{1}{2\pi}\log|x-y| & \text{for } n=2, \\ \dfrac{1}{4\pi|x-y|} & \text{for } n=3, \end{cases}$$

denotes the fundamental solution of the Laplace operator. The single layer boundary integral operator is bounded and elliptic in $H^{-1/2}(\Gamma)$, see, e.g., [73, Theorem 6.22, 6.23]. Consequently, we can introduce the natural density $w_{\rm eq} \in H^{-1/2}(\Gamma)$, which is the unique solution of the following saddle point problem. Find $(w_{\rm eq}, \lambda) \in H^{-1/2}(\Gamma) \times \mathbb{R}$ such that

$$\begin{aligned} \langle V w_{\rm eq}, \psi \rangle_\Gamma - \lambda \langle 1, \psi \rangle_\Gamma &= 0, \\ \langle w_{\rm eq}, 1 \rangle_\Gamma &= 1, \end{aligned}$$

for all $\psi \in H^{-1/2}(\Gamma)$. It turns out, see [73, p. 144] that the dual space of $H_*^{-1/2}(\Gamma)$ is then

$$H_*^{1/2}(\Gamma) = [H_*^{-1/2}(\Gamma)]^* = \left\{ v \in H^{1/2}(\Gamma) : \langle v, w_{\rm eq} \rangle_\Gamma = 0 \right\},$$

with the corresponding norm

$$\|v\|_{H_*^{1/2}(\Gamma)} = \left(\langle v, w_{\rm eq} \rangle_\Gamma^2 + |v|^2_{H^{1/2}(\Gamma)} \right)^{1/2},$$

which defines an equivalent norm in $H^{1/2}(\Gamma)$ by the norm equivalence theorem of Sobolev, c.f. [73, Theorem 2.6]. Further, we introduce the space

$$H_*^1(\Omega) = \left\{ v \in H^1(\Omega) : \langle v_{|\Gamma}, w_{\rm eq} \rangle_\Gamma = 0 \right\},$$

equipped with the norm

$$\|v\|_{H_*^1(\Omega)} = \left(\langle v_{|\Gamma}, w_{\rm eq} \rangle_\Gamma^2 + |v|^2_{H^1(\Omega)} \right)^{1/2},$$

which defines an equivalent norm in $H^1(\Omega)$, by the same arguments as above.

As in the previous sections we split the primal state $u = u_f + u_z$, where $u_f \in H^1(\Omega)$ and $u_z \in H^1(\Omega)$ are the solutions of the Neumann boundary value problems

$$\begin{aligned} -\Delta u_f &= f && \text{in } \Omega, \\ \partial_n u_f &= 0 && \text{on } \Gamma, \end{aligned} \tag{3.42}$$

and

$$\begin{aligned} -\Delta u_z &= 0 && \text{in } \Omega, \\ \partial_n u_z &= z && \text{on } \Gamma. \end{aligned} \tag{3.43}$$

Since u_z is via the problem (3.43) only uniquely determined up to an additive constant, we introduce the following decomposition

$$u_z = u_0 + \alpha_\Omega, \tag{3.44}$$

with $u_0 \in H^1_*(\Omega)$ and $\alpha_\Omega \in \mathbb{R}$. Consequently, we have

$$\langle u_z, w_{\mathrm{eq}} \rangle_\Gamma = \langle u_0, w_{\mathrm{eq}} \rangle_\Gamma + \alpha_\Omega \langle 1, w_{\mathrm{eq}} \rangle_\Gamma = \alpha_\Omega.$$

We shall note at this point that for the boundary value problem (3.42) only one particular solution u_f is needed.

Now, we multiply the equation (3.43) by a test function and apply integration by parts. This leads to the following variational formulation. Find $u_0 \in H^1_*(\Omega)$ such that

$$\langle \nabla u_0, \nabla v \rangle_{L^2(\Omega)} = \langle z, v_{|\Gamma} \rangle_\Gamma, \tag{3.45}$$

for all $v \in H^1_*(\Omega)$. This problem has clearly a unique solution $u_0 \in H^1_*(\Omega)$. In particular this motivates the following definition of the norm

$$\|z\|^2_{H^{-1/2}_*(\Gamma)} := \langle z, u_{0|\Gamma} \rangle_\Gamma = \langle \nabla u_0, \nabla u_0 \rangle_{L^2(\Omega)} = \|u_0\|^2_{H^1_*(\Omega)},$$

for all $z \in H^{-1/2}_*(\Gamma)$.

Now, we can introduce the stabilized inverse Steklov–Poincaré operator as a mapping $S_* : H^{-1/2}_*(\Gamma) \to H^{1/2}_*(\Gamma)$,

$$S_*^{-1} z := u_{0|\Gamma}, \tag{3.46}$$

which realizes the Neumann to Dirichlet map for the boundary value problem (3.43). This definition implies the following representation of the norm in $H^{-1/2}_*(\Gamma)$,

$$\|z\|^2_{H^{-1/2}_*(\Gamma)} = \langle S_*^{-1} z, z \rangle_\Gamma,$$

for all $z \in H^{-1/2}_*(\Gamma)$. Let us summarize the properties of the stabilized inverse Steklov–Poincaré operator.

Proposition 3.3. *The inverse stabilized Steklov–Poincaré operator S_*^{-1}, defined in* (3.46)*, is self-adjoint, bounded and elliptic in $H^{-1/2}_*(\Gamma)$.*

It remains to determine the unknown parameter $\alpha_\Omega \in \mathbb{R}$ of the decomposition (3.44). Therefore we rewrite the cost functional (3.40) with the definition of the inverse stabilized Steklov–Poincaré operator as

$$\begin{aligned}
\mathcal{J}(u,z) &= \frac{1}{2} \|u_0 + \alpha_\Omega + u_f - \overline{u}\|^2_{L^2(\Omega)} + \frac{1}{2} \varrho \langle S_*^{-1} z, z \rangle_\Gamma \\
&= \frac{1}{2} \|u_0 + u_f - \overline{u}\|^2_{L^2(\Omega)} + \alpha_\Omega \langle u_0 + u_f - \overline{u}, 1 \rangle_{L^2(\Omega)} + \frac{\alpha_\Omega^2}{2} |\Omega| + \frac{1}{2} \varrho \langle u_{0|\Gamma}, z \rangle_\Gamma.
\end{aligned}$$

Minimization with respect to α_Ω leads to

$$\alpha_\Omega = \frac{1}{|\Omega|}\langle \overline{u} - u_0 - u_f, 1\rangle_{L^2(\Omega)},$$

and consequently with the definition of α_Ω follows

$$\langle u - \overline{u}, 1\rangle_{L^2(\Omega)} = \langle u_0 + u_f - \overline{u}, 1\rangle_{L^2(\Omega)} + \alpha_\Omega |\Omega| = 0.$$

Thus, we have $u - \overline{u} \in L^2_*(\Omega)$, where

$$L^2_*(\Omega) = \left\{ v \in L^2(\Omega) : \langle v, 1\rangle_{L^2(\Omega)} = 0 \right\}.$$

Similarly to the optimal Neumann boundary control problem for the Yukawa equation, in Subsection 3.2.1, we derive the first order necessary optimality condition as an equivalent formulation for our problem (3.40)–(3.41), see [81]. Therefore we introduce, for the boundary value problem (3.43) the solution operator $\mathcal{H}_*$, which is for each $z \in H^{1/2}_*(\Gamma)$ defined by

$$\mathcal{H}_* z = u_0 - \frac{1}{|\Omega|}\langle u_0, 1\rangle_{L^2(\Omega)},$$

where $u_0 \in H^1_*(\Omega)$ is the unique solution of (3.45). Then the solution operator is obviously a mapping $\mathcal{H}_* : H^{-1/2}_*(\Gamma) \to L^2_*(\Omega)$. Let us recall that

$$u - \overline{u} = u_z + u_f - \overline{u} = \mathcal{H}_* z + |\Omega|^{-1}\langle u_0, 1\rangle_{L^2(\Omega)} + \alpha_\Omega + u_f - \overline{u},$$

which gives us the possibility to introduce the reduced cost functional as

$$\widetilde{\mathcal{J}}(z) := \frac{1}{2}\|\mathcal{H}_* z + |\Omega|^{-1}\langle u_0, 1\rangle_{L^2(\Omega)} + \alpha_\Omega + u_f - \overline{u}\|^2_{L^2(\Omega)} + \frac{1}{2}\varrho\langle S_*^{-1} z, z\rangle_\Gamma.$$

By this observation, the optimal control problem (3.40)–(3.41) can be formulated in the following way: Find the optimal control $\hat{z} \in H^{-1/2}_*(\Gamma)$ such that

$$\begin{aligned} \widetilde{\mathcal{J}}(\hat{z}) &= \min_{z \in H^{-1/2}_*(\Gamma)} \widetilde{\mathcal{J}}(z) \\ &= \min_{z \in H^{-1/2}_*(\Gamma)} \left\{ \frac{1}{2}\|\mathcal{H}_* z + |\Omega|^{-1}\langle u_0, 1\rangle_{L^2(\Omega)} + \alpha_\Omega + u_f - \overline{u}\|^2_{L^2(\Omega)} + \frac{1}{2}\varrho\langle S_*^{-1} z, z\rangle_\Gamma \right\}, \end{aligned} \tag{3.47}$$

is satisfied. Since the inverse stabilized Steklov–Poincaré operator S_*^{-1} is a bounded and self-adjoint operator, we can apply as before the result [81, Theorem 2.22], which states that the above reduced minimization problem (3.47) is equivalent to the following operator equation

$$\mathcal{H}^*_*(\mathcal{H}_* z + |\Omega|^{-1}\langle u_0, 1\rangle_{L^2(\Omega)} + \alpha_\Omega + u_f - \overline{u}) + \varrho S_*^{-1} z = 0, \tag{3.48}$$

in the sense of $H^{-1/2}_*(\Gamma)$ with the adjoint solution operator $\mathcal{H}^*_* : L^2_*(\Omega) \to H^{1/2}_*(\Gamma)$. To solve this equation we need the adjoint solution operator, which is characterized by the following theorem.

Theorem 3.6. *Let $\psi \in L^2_*(\Omega)$ be arbitrary but fixed. Moreover, let $p \in H^1_*(\Omega)$ be the unique solution of the problem*

$$\begin{aligned} -\Delta p &= \psi \qquad \text{in } \Omega, \\ \partial_n p &= 0 \qquad \text{on } \Gamma. \end{aligned}$$

Then for the adjoint solution operator there holds $\mathcal{H}_*^* \psi = p_{|\Gamma}$.

Proof. We proceed with the proof similarly as the proof of Theorem 3.5. From the decomposition $u_z = u_0 + \alpha_\Omega$ and the variational formulation (3.45) we obtain: Find $u_0 \in H^1_*(\Omega)$ such that

$$\langle \nabla u_0, \nabla v \rangle_{L^2(\Omega)} = \langle z, v_{|\Gamma} \rangle_\Gamma,$$

for all $v \in H^1_*(\Omega)$. Moreover, we want to find $p \in H^1_*(\Omega)$ such that

$$\langle \nabla p, \nabla v \rangle_{L^2(\Omega)} = \langle \psi, v \rangle_{L^2(\Omega)},$$

for all $v \in H^1_*(\Omega)$. As previously, we set in the first problem $v = p$ and for the second problem $v = u_0$, and obtain, due to symmetry, the equality

$$\begin{aligned} \langle \mathcal{H}_* z, \psi \rangle_{L^2(\Omega)} &= \langle u_0, \psi \rangle_{L^2(\Omega)} - |\Omega|^{-1} \langle u_0, 1 \rangle_{L^2(\Omega)} \langle \psi, 1 \rangle_{L^2(\Omega)} = \langle u_0, \psi \rangle_{L^2(\Omega)} \\ &= \langle \nabla p, \nabla u_0 \rangle_{L^2(\Omega)} = \langle z, p_{|\Gamma} \rangle_\Gamma = \langle z, \mathcal{H}_*^* \psi \rangle_\Gamma, \end{aligned}$$

since $\psi \in L^2_*(\Omega)$. This concludes the proof. □

By using $u_z = \mathcal{H}_* z + |\Omega|^{-1} \langle u_0, 1 \rangle_{L^2(\Omega)} + \alpha_\Omega$ we obtain

$$\mathcal{H}_*^* (\mathcal{H}_* z + |\Omega|^{-1} \langle u_0, 1 \rangle_{L^2(\Omega)} + \alpha_\Omega + u_f - \overline{u}) = \mathcal{H}_*^* (u - \overline{u}) = p_{|\Gamma}.$$

In particular, we obtain from (3.48) the relation

$$p_{|\Gamma} + \varrho S_*^{-1} z = 0,$$

in the sense of $H^{1/2}_*(\Gamma)$. As a consequence an arbitrary constant can be added. In particular we can add α_Ω and because of $S_*^{-1} z = u_{0|\Gamma}$ the above equation turns into

$$p_{|\Gamma} + \varrho u_{z|\Gamma} = 0. \tag{3.49}$$

Note that the assumption $u - \overline{u} \in L^2_*(\Omega)$ ensures the solvability of the adjoint equation, while $p_{|\Gamma} \in H^{1/2}_*(\Gamma)$ is the scaling condition for uniqueness.

The first order necessary optimality conditions, which are equivalent to the optimal control problem (3.40)–(3.41), are given by the optimality system

Primal problem	Adjoint problem	Optimality condition
$-\Delta u = f \quad \text{in } \Omega,$	$-\Delta p = u - \overline{u} \quad \text{in } \Omega,$	$p + \varrho S_*^{-1} z = 0 \quad \text{on } \Gamma.$
$\partial_n u = z \quad \text{on } \Gamma,$	$\partial_n p = 0 \quad \text{on } \Gamma,$	

(3.50)

Remark 3.5. *As in the previous subsection for the Yukawa equation we can eliminate the control z by splitting $u = u_f + u_z$. With the relation (3.49) we can replace the boundary condition $\partial_n u_z = z$ by $p + \varrho u_z = 0$. This leads to the following optimality system*

$$\begin{array}{lll} -\Delta u_f = f & \text{in } \Omega, & -\Delta u_z = 0 \quad \text{in } \Omega, \quad -\Delta p = u_f + u_z - \overline{u} \quad \text{in } \Omega, \\ \partial_n u_f = 0 & \text{on } \Gamma, & \varrho u_z = -p \quad \text{on } \Gamma, \quad \partial_n p = 0 \quad \text{on } \Gamma. \end{array} \tag{3.51}$$

Again, the control can be found in a post processing step by $z = \partial_n u_z$.

3.2.3 Variational formulation

In the following we prove the existence and uniqueness of a solution of the Neumann boundary control problem, for both constraints, the Yukawa equation (3.31) and the Poisson equation (3.41). Note that in the latter case we assume the scaling condition

$$\langle f, 1 \rangle_\Omega = 0.$$

For both constraints, i.e. Poisson and Yukawa equations, we can write the corresponding optimality systems, for $\kappa \geq 0$, as

$$\begin{aligned} -\Delta u_f + \kappa u_f &= f && \text{in } \Omega, \\ \partial_n u_f &= 0 && \text{on } \Gamma, \\ -\Delta u_z + \kappa u_z &= 0 && \text{in } \Omega, \\ \varrho u_z &= -p && \text{on } \Gamma, \\ -\Delta p + \kappa p &= u_f + u_z - \overline{u} && \text{in } \Omega, \\ \partial_n p &= 0 && \text{on } \Gamma. \end{aligned} \tag{3.52}$$

From (3.52) we obtain for the first equation the following variational formulation: Find $u_f \in H^1(\Omega)$ such that

$$\langle \nabla u_f, \nabla v \rangle_{L^2(\Omega)} + \kappa \langle u_f, v \rangle_{L^2(\Omega)} = \langle f, v \rangle_\Omega,$$

for all $v \in H^1(\Omega)$. This problem has a unique solution $u_f \in H^1(\Omega)$ for $\kappa > 0$ and a unique solution $u_f \in H^1_*(\Omega)$ in the case $\kappa = 0$. The corresponding variational formulation for u_z and p reads then: Find $(u_z, p) \in H^1(\Omega) \times H^1(\Omega)$ with $\varrho u_z = -p$ on Γ, such that

$$\begin{aligned} \langle u_z, v \rangle_{L^2(\Omega)} - \langle \nabla p, \nabla v \rangle_{L^2(\Omega)} - \kappa \langle p, v \rangle_{L^2(\Omega)} &= \langle \overline{u} - u_f, v \rangle_{L^2(\Omega)}, \\ \langle \nabla u_z, \nabla q \rangle_{L^2(\Omega)} + \kappa \langle u_z, q \rangle_{L^2(\Omega)} &= 0, \end{aligned} \tag{3.53}$$

for all $(v, q) \in H^1(\Omega) \times H^1_0(\Omega)$.

In order to prove the existence and uniqueness of the solution of the variational formulation above, we introduce the following solution operators. Let $\varphi \in H^{1/2}(\Gamma)$ be arbitrary but fixed and let $u_\varphi \in H^1(\Omega)$ with $u_\varphi = \varphi$ on Γ be the unique solution of

$$\langle \nabla u_\varphi, \nabla q \rangle_{L^2(\Omega)} + \kappa \langle u_\varphi, q \rangle_{L^2(\Omega)} = 0,$$

for all $q \in H^1_0(\Omega)$. Then we introduce the operator $\mathcal{H}_\kappa : H^{1/2}(\Gamma) \to H^1(\Omega)$, defined as $u_\varphi = \mathcal{H}_\kappa \varphi$. For the related Neumann datum we introduce, for all $\kappa \geq 0$, the associated Steklov–Poincaré operator $S_\kappa : H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$, defined via

$$S_\kappa \varphi := \partial_n u_\varphi.$$

Note that S_κ is semi-elliptic in $H^{1/2}(\Gamma)$ for all $\kappa \geq 0$, see Proposition 3.1. Further, we introduce the operator $T_{\kappa,\varrho} : H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$, by

$$T_{\kappa,\varrho} := \mathcal{H}_\kappa^* \mathcal{H}_\kappa + \varrho S_\kappa. \tag{3.54}$$

Due to boundedness of the Steklov–Poincaré operator S_κ and the fact that

$$\|u_\varphi\|_{H^1(\Omega)} \leq c \|\varphi\|_{H^{1/2}(\Gamma)},$$

we can conclude the boundedness of $T_{\kappa,\varrho}$,

$$\|T_{\kappa,\varrho}\varphi\|_{H^{-1/2}(\Gamma)} \leq c_2^{T_{\kappa,\varrho}} \|\varphi\|_{H^{1/2}(\Gamma)}.$$

Lemma 3.2. *The operator $T_{\kappa,\varrho} : H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$, defined in (3.54), is elliptic in $H^{1/2}(\Gamma)$, i.e.,*

$$\langle T_{\kappa,\varrho}\varphi, \varphi \rangle_\Gamma \geq c_1^{T_{\kappa,\varrho}} \|\varphi\|^2_{H^{1/2}(\Gamma)},$$

for all $\varphi \in H^{1/2}(\Gamma)$ and $\kappa \geq 0$.

Proof. The ellipticity follows directly from the definitions and properties of $\mathcal{H}_\kappa$ and S_κ. More precisely we have

$$\begin{aligned}
\langle T_{\kappa,\varrho}\varphi, \varphi \rangle_\Gamma &= \langle \mathcal{H}_\kappa \varphi, \mathcal{H}_\kappa \varphi \rangle_{L^2(\Omega)} + \varrho \langle S_\kappa \varphi, \varphi \rangle_\Gamma \\
&= \langle u_\varphi, u_\varphi \rangle_{L^2(\Omega)} + \varrho \langle \nabla u_\varphi, \nabla u_\varphi \rangle_{L^2(\Omega)} + \varrho\kappa \langle u_\varphi, u_\varphi \rangle_{L^2(\Omega)} \\
&\geq \min\{\varrho, 1 + \kappa\varrho\} \|u_\varphi\|^2_{H^1(\Omega)} \geq c_1^{T_{\kappa,\varrho}} \|\varphi\|^2_{H^{1/2}(\Gamma)},
\end{aligned}$$

for all $\varphi \in H^{1/2}(\Gamma)$, which proves the assertion. □

In order to prove the existence and uniqueness of the mixed variational formulation (3.53), we would like to use standard results for saddle point problems, see, e.g., [12, 73]. This is nevertheless more complicated due to the boundary condition $\varrho u_z = -p$. Alternatively, we consider the operator Schur complement equation with respect to the adjoint state $p_{|\Gamma}$ on the boundary.

Theorem 3.7. *Let $\overline{u} \in L^2(\Omega)$, then for all $\kappa \geq 0$ and $\varrho > 0$ there exists a unique solution $(u_z, p) \in H^1(\Omega) \times H^1(\Omega)$ of the variational formulation* (3.53).

Proof. For a given $\psi \in L^2(\Omega)$ and $\kappa \geq 0$ let $\mathcal{N}_\kappa : L^2(\Omega) \to H^1_0(\Omega)$ be defined as $u_\psi = \mathcal{N}_\kappa \psi$, where $u_\psi \in H^1_0(\Omega)$ is the unique solution of the homogeneous Dirichlet boundary value problem

$$\langle \nabla u_\psi, \nabla q \rangle_{L^2(\Omega)} + \kappa \langle u_\psi, q \rangle_{L^2(\Omega)} = \langle \psi, q \rangle_{L^2(\Omega)},$$

for all $q \in H^1_0(\Omega)$. Consequently we obtain for the corresponding Neumann datum

$$\begin{aligned} \langle \partial_n \mathcal{N}_\kappa \psi, \varphi \rangle_\Gamma &= \langle \partial_n u_\psi, u_\varphi \rangle_\Gamma \\ &= \langle \nabla u_\psi, \nabla u_\varphi \rangle_{L^2(\Omega)} + \kappa \langle u_\psi, u_\varphi \rangle_{L^2(\Omega)} - \langle \psi, u_\varphi \rangle_{L^2(\Omega)} \\ &= -\langle \psi, \mathcal{H}_\kappa \varphi \rangle_{L^2(\Omega)} = -\langle \mathcal{H}^*_\kappa \psi, \varphi \rangle_\Gamma, \end{aligned}$$

for all $\varphi \in H^{1/2}(\Gamma)$. By the definition of $\mathcal{H}_\kappa$ we conclude from the primal problem that

$$\varrho u_z = -\mathcal{H}_\kappa p_{|\Gamma}.$$

Further, we have from the adjoint problem the relation

$$\begin{aligned} p &= \mathcal{H}_\kappa p_{|\Gamma} + \mathcal{N}_\kappa (u_z + u_f - \overline{u}) = \mathcal{H}_\kappa p_{|\Gamma} + \mathcal{N}_\kappa (-\varrho^{-1} \mathcal{H}_\kappa p_{|\Gamma} + u_f - \overline{u}) \\ &= \mathcal{H}_\kappa p_{|\Gamma} - \varrho^{-1} \mathcal{N}_\kappa (\mathcal{H}_\kappa p_{|\Gamma} + \varrho(\overline{u} - u_f)), \end{aligned}$$

and consequently for the associated Neumann datum, using the relation above,

$$\partial_n p = S_\kappa p_{|\Gamma} + \varrho^{-1} \mathcal{H}^*_\kappa (\mathcal{H}_\kappa p_{|\Gamma} + \varrho(\overline{u} - u_f)) = 0.$$

This means that we have to solve the variational formulation

$$\langle T_{\kappa,\varrho} p_{|\Gamma}, \varphi \rangle_\Gamma = \varrho \langle \mathcal{H}^*_\kappa (u_f - \overline{u}), \varphi \rangle_\Gamma,$$

for all $\varphi \in H^{1/2}(\Gamma)$, where $T_{\kappa,\varrho} = \mathcal{H}^*_\kappa \mathcal{H}_\kappa + \varrho S_\kappa$. As proven in Lemma 3.2 the operator $T_{\kappa,\varrho}$ is bounded and elliptic for all $\kappa \geq 0$, see also [59]. Consequently the above variational formulation has a unique solution $p_{|\Gamma} \in H^{1/2}(\Gamma)$. Further, we can conclude the existence of a unique solution $(u_z, p) \in H^1(\Omega) \times H^1(\Omega)$ by solving the related Dirichlet and Neumann boundary value problem. □

As stated in Theorem 3.7, the variational formulation (3.53) of the optimality system, has a unique solution also in the particular case $\kappa = 0$. By taking the test function $v = 1$, we conclude from the variational formulation (3.53) the solvability condition

$$\langle u - \overline{u}, 1 \rangle_{L^2(\Omega)} = 0.$$

If the state $u \in H^1(\Omega)$ is known we can compute the control z in a post processing step as

$$z = \partial_n u \quad \text{on } \Gamma,$$

which, by Green's formula, automatically satisfies

$$\langle z, 1\rangle_\Gamma = \langle \partial_n u, 1\rangle_\Gamma = 0,$$

and thus the control z belongs to $H_*^{-1/2}(\Gamma)$. Hence, we conclude that the Neumann boundary control problem for the Laplace equation is a priori included in the more general case of the Neumann control problem for the Yukawa equation, $\kappa \geq 0$.

3.2.4 Discretization

The finite element discretization of the variational formulation (3.53) as well as the related error analysis can be done similarly as in the case of the Dirichlet boundary control problem, see Section 3.1, and also [59]. As before, we consider an admissible, shape-regular and quasi-uniform triangulation $\mathcal{T}_h$ and introduce the finite element spaces

$$\mathcal{V}_h = \operatorname{span}\{\varphi_i^1\}_{i=1}^{n_I+n_C} \subset H^1(\Omega), \quad \mathcal{Q}_h = \operatorname{span}\{\varphi_i^1\}_{i=1}^{n_I} \subset H_0^1(\Omega),$$

of piecewise linear and globally continuous shape functions, defined in (3.18). With the separation of interior and boundary degrees of freedoms we have for the dimensions of the finite element spaces $\dim \mathcal{V}_h = n_I + n_C$ and $\dim \mathcal{Q}_h = n_I$, as introduced in Section 3.1. Let us denote by $u_{f,h} \in \mathcal{V}_h$ the finite element solution of (3.52). The corresponding variational formulation reads then: Find $(u_{z,h}, p_h) \in \mathcal{V}_h \times \mathcal{V}_h$ with $\varrho u_{z,h} = -p_h$ on Γ, such that

$$\begin{aligned} \langle u_{z,h}, v_h\rangle_{L^2(\Omega)} - \langle \nabla p_h, \nabla v_h\rangle_{L^2(\Omega)} - \kappa\langle p_h, v_h\rangle_{L^2(\Omega)} &= \langle \overline{u} - u_{f,h}, v_h\rangle_{L^2(\Omega)}, \\ \langle \nabla u_{z,h}, \nabla q_h\rangle_{L^2(\Omega)} + \kappa\langle u_{z,h}, q_h\rangle_{L^2(\Omega)} &= 0, \end{aligned} \tag{3.55}$$

for all $(v_h, q_h) \in \mathcal{V}_h \times \mathcal{Q}_h$.

We introduce standard finite element mass and stiffness matrices, as well as the right-hand side vectors, by

$$M_h[j,i] = \langle \varphi_i^1, \varphi_j^1\rangle_{L^2(\Omega)}, \quad A_h[j,i] = \langle \nabla\varphi_i^1, \nabla\varphi_j^1\rangle_{L^2(\Omega)}, \quad \overline{u}[i] = \langle \overline{u}, \varphi_i^1\rangle_{L^2(\Omega)}, \quad f[i] = \langle f, \varphi_i^1\rangle_\Omega,$$

for all $i, j = 1, \ldots, n_I + n_C$. Additionally, we consider the weighted combination

$$A_{h,\kappa}[j,i] = \langle \nabla\varphi_i^1, \nabla\varphi_j^1\rangle_{L^2(\Omega)} + \kappa\langle \varphi_i^1, \varphi_j^1\rangle_{L^2(\Omega)},$$

for all $i, j = 1, \ldots, n_I + n_C$. Note that for the case $\kappa = 0$ there holds $A_{h,0} = A_h$. The discrete variational formulation (3.55) is then equivalent to the linear system

$$\begin{pmatrix} M_{II} & -A_{II,\kappa} & -A_{IC,\kappa} - \varrho^{-1} M_{IC} \\ A_{II,\kappa} & & -\varrho^{-1} A_{IC,\kappa} \\ M_{CI} & -A_{CI,\kappa} & -A_{CC,\kappa} - \varrho^{-1} M_{CC} \end{pmatrix} \begin{pmatrix} \underline{u}_I \\ \underline{p}_I \\ \underline{p}_C \end{pmatrix} = \begin{pmatrix} \widetilde{\underline{u}}_I \\ \underline{0} \\ \widetilde{\underline{u}}_C \end{pmatrix}, \tag{3.56}$$

where the modified right-hand side is given by $\widetilde{\underline{u}} = (\widetilde{\underline{u}}_I, \widetilde{\underline{u}}_C)^\top = \overline{\underline{u}} - A_{h,\kappa}^{-1} \underline{f}$. As for the Dirichlet boundary control problem we derive the Schur complement system. By using

$$\underline{u}_I = \varrho^{-1} A_{II,\kappa}^{-1} A_{IC,\kappa} \underline{p}_C,$$

and

$$\underline{p}_I = A_{II,\kappa}^{-1} \big[\varrho^{-1} M_{II} A_{II,\kappa}^{-1} A_{IC,\kappa} \underline{p}_C - A_{IC,\kappa} \underline{p}_C - \varrho^{-1} M_{IC} \underline{p}_C - \widetilde{\underline{u}}_I \big],$$

we can derive the Schur complement system as

$$\begin{aligned} \Big[\big[M_{CC} - M_{CI} A_{II,\kappa}^{-1} A_{IC,\kappa} - A_{CI,\kappa} A_{II,\kappa}^{-1} M_{IC} + A_{CI,\kappa} A_{II,\kappa}^{-1} M_{II} A_{II,\kappa}^{-1} A_{IC,\kappa} \big] \\ + \varrho \big[A_{CC,\kappa} - A_{CI,\kappa} A_{II,\kappa}^{-1} A_{IC,\kappa} \big] \Big] \underline{p}_C \\ = \varrho \big[A_{CI,\kappa} A_{II,\kappa}^{-1} \widetilde{\underline{u}}_I - \widetilde{\underline{u}}_C \big], \end{aligned}$$

with the Schur complement

$$\begin{aligned} T_{h,\kappa} + \varrho S_{h,\kappa} = M_{CC} - M_{CI} A_{II,\kappa}^{-1} A_{IC,\kappa} - A_{CI,\kappa} A_{II,\kappa}^{-1} M_{IC} \\ + A_{CI,\kappa} A_{II,\kappa}^{-1} M_{II} A_{II,\kappa}^{-1} A_{IC,\kappa} \\ + \varrho \big[A_{CC,\kappa} - A_{CI,\kappa} A_{II,\kappa}^{-1} A_{IC,\kappa} \big]. \end{aligned} \tag{3.57}$$

Note that the Schur complement matrix $T_{h,\kappa} + \varrho S_{h,\kappa}$ defined in (3.57), coincides with the Schur complement matrix $T_h + \varrho S_h$ of the Dirichlet boundary control problem defined in (3.28), for the particular case $\kappa = 0$, i.e., $T_{h,0} + \varrho S_{h,0} = T_h + \varrho S_h$. Using the discrete optimality condition

$$\underline{p}_C + \varrho \underline{u}_C = \underline{0},$$

we can rewrite the Schur complement equation in terms of $\underline{u}_C$ and thus obtain

$$(T_{h,\kappa} + \varrho S_{h,\kappa}) \underline{u}_C = \widetilde{\underline{u}}_C - A_{CI,\kappa} A_{II,\kappa}^{-1} \widetilde{\underline{u}}_I. \tag{3.58}$$

This means, that for $\kappa = 0$ and $f = 0$ the discrete Schur complement equation of the Neumann boundary control problem (3.58) coincides with the Schur complement equation of the Dirichlet boundary control problem (3.27). In this particular case the Schur complement equation is given by

$$(T_h + \varrho S_h) \underline{u}_C = \overline{\underline{u}}_C - A_{CI} A_{II}^{-1} \overline{\underline{u}}_I.$$

The linear system (3.58) admits a unique solution $\underline{u}_C \in \mathbb{R}^{n_C}$, from which we can determine $\underline{u}_I \in \mathbb{R}^{n_I}$, and therefore $\underline{u} = (\underline{u}_I, \underline{u}_C)^\top \leftrightarrow u_h \in \mathcal{V}_h$ defines an approximate state. From this we can find a piecewise constant approximation of the control z by computing

$$z_h = \partial_n u_h \quad \text{on } \Gamma. \tag{3.59}$$

Remark 3.6. *If we consider the finite element space $Z_h \subset H^{-1/2}(\Gamma)$ for the discrete control, the discretization of the original optimality system, which includes the unknown z, would require an approximation $z_h \in Z_h$ of the control $z \in H^{-1/2}(\Gamma)$. Correspondingly, the variational formulation of the optimality condition* (3.49) *would be then formulated as*

$$\langle p_h + \varrho u_h, \mu_h \rangle_\Gamma = 0,$$

for all $\mu_h \in Z_h$. Since this corresponds to a mixed finite element discretization scheme, the discrete inf–sup condition,

$$\widetilde{c}_S \|\mu_h\|_{H^{-1/2}(\Gamma)} \leq \sup_{v_h \in \mathcal{V}_h} \frac{\langle \mu_h, v_h \rangle_\Gamma}{\|v_h\|_{H^1(\Omega)}}, \tag{3.60}$$

for all $\mu_h \in Z_h$, is required. Note that this excludes the choice of a piecewise constant approximation z_h of the control which is defined on the boundary mesh of the piecewise linear finite element approximations u_h and p_h. If the mesh size H of the finite element space Z_H on the boundary is nevertheless coarse enough, i.e. $h \leq cH$, the inf–sup condition (3.60) *is valid, see* [73, Theorem 11.5]. *In our situation, when the control is eliminated, the computation of* (3.59) *is a post processing step, i.e. a discrete inf–sup condition such as* (3.60) *is not required.*

For $\kappa = 0$ the optimality condition of the Dirichlet boundary control problem involves the same operator as for the Neumann boundary control problem, and since the system matrices of both Schur complement systems (3.28) and (3.57) coincide, one may rise the question if there is a relation between the solutions of the underlying Dirichlet and Neumann boundary control problems.

Theorem 3.8. *For given $\overline{u} \in L_2(\Omega)$ and $\varrho > 0$ we consider the Dirichlet boundary control problem* (3.10) *with $f = 0$ and the state solution $u_D \in H^1(\Omega)$. Further, we consider the Neumann boundary control problem* (3.50) *with $\kappa = 0$, $f = 0$ and the state solution $u_N \in H^1(\Omega)$. Then the states coincide, i.e. we have $u_D = u_N$ in $H^1(\Omega)$.*

Proof. Let $(u_D, p_D) \in H^1(\Omega) \times H^1_0(\Omega)$ be the unique solution of the optimality system (3.10), i.e.

$$\begin{aligned} -\Delta u_D &= 0 && \text{in } \Omega, & \qquad -\Delta p_D &= u_D - \overline{u} && \text{in } \Omega, \\ \varrho \partial_n u_D &= \partial_n p_D && \text{on } \Gamma, & \qquad p_D &= 0 && \text{on } \Gamma. \end{aligned}$$

Using the representation $\partial_n w = Sw_{|\Gamma} - Nf$ on Γ, where N is the Newton potential, to describe the Dirichlet to Neumann map subject to the solution of the Poisson problem $-\Delta w = f$ in Ω, we obtain

$$\partial_n u_D = Su_{D|\Gamma}, \qquad \partial_n p_D = Sp_{D|\Gamma} - N(u_D - \overline{u}),$$

and therefore

$$N\overline{u} = \varrho Su_{D|\Gamma} - Sp_{D|\Gamma} + Nu_D = \varrho Su_{D|\Gamma} + Nu_D.$$

Analogously, let $(u_N, p_N) \in H^1(\Omega) \times H^1(\Omega)$ be the solution of the optimality system (3.50), i.e.,

$$\begin{aligned} -\Delta u_N &= 0 && \text{in } \Omega, & -\Delta p_N &= u_N - \overline{u} && \text{in } \Omega, \\ \varrho u_N &= -p_N && \text{on } \Gamma, & \partial_n p_N &= 0 && \text{on } \Gamma. \end{aligned}$$

As above we can write

$$0 = \partial_n p_N = Sp_{N|\Gamma} - N(u_N - \overline{u}) = -\varrho Su_{N|\Gamma} - N(u_N - \overline{u}),$$

and hence

$$N\overline{u} = \varrho Su_{N|\Gamma} + Nu_N.$$

In particular, we obtain

$$\varrho S(u_{N|\Gamma} - u_{D|\Gamma}) + N(u_N - u_D) = 0,$$

which means that, $\widetilde{u} = u_N - u_D$ is a solution of the homogeneous Neumann boundary value problem

$$\begin{aligned} -\varrho\Delta\widetilde{u} + \widetilde{u} &= 0 && \text{in } \Omega, \\ \partial_n\widetilde{u} &= 0 && \text{on } \Gamma. \end{aligned}$$

Since this problem admits the unique solution $\widetilde{u} = 0$ we can therefore conclude $u_D = u_N$, which concludes the proof. □

3.3 Concluding remarks

In this chapter we have studied optimal boundary control problems in the energy space.

In the first section, on optimal Dirichlet boundary control, we have started with the derivation of the first order necessary optimality conditions. In particular, it has been possible to eliminate the control and derive a mixed variational formulation for which we have proven the existence and uniqueness of a solution, including a stability estimate. The corresponding finite element discretization has been done by piecewise linear and globally continuous

shape functions for which we have proven optimal error estimates. Corresponding numerical examples have illustrated these theoretical results.

The optimal Neumann boundary control has been treated in the second part of this chapter. It turned out that many ideas from the Dirichlet control case could be applied. First, we have considered as a constraint the Yukawa equation, since this equation has a unique solution. For this model problem we have derived the optimality system. In the second part, we have considered the Neumann boundary control for the Poisson equation, where we had to introduce a stabilized inverse Steklov–Poincaré operator. It nevertheless turned out that this problem is a special case of the Yukawa equation. Existence and uniqueness of the problem has been shown. Further, we proved that in the case of the Laplace equation as a constraint, the primal states of Dirichlet and Neumann boundary control problems, coincide.

In the upcoming chapters we apply the boundary control approach in the energy space to a blood flow related application, as motivated in Chapter 1 and 2. Further, we discuss the construction of robust preconditioners for the problems stated in this chapter. As it was mentioned, the optimality system leads in the limit case, i.e., for $\varrho = 0$, to the biharmonic equation of first kind. Due to this reason this model problem should be studied and corresponding preconditioning strategies discussed. These ideas shall be applied afterwards for the optimal Dirichlet and Neumann boundary control problem.

4 AN OPTIMAL CONTROL PROBLEM FOR ARTERIAL BLOOD FLOW

In Chapter 2 we have discussed the numerical simulation of hemodynamic indicators, such as the wall shear stress and the oscillatory shear index. From the point of optimization, we can ask how to minimize such factors, for instance by controlling the inflow velocity. An important point here is the specification of a reasonable cost functional, which should be done in such a way that those risk factors are minimized. Within this chapter the focus will be laid on vortex minimization for arterial blood flow. In particular we apply the ideas of the optimal Dirichlet boundary control in the energy space, introduced in the previous chapter, see also [39, 59], to a blood flow related problematics. More precisely, we are interested in the control of the inflow velocity (inflow profile) into an artery, subject to the minimization of vortices of the flow. Such simulations can be used for example for the inflow control of artificial heart pumps. For the description of the blood flow we consider the Navier–Stokes equations (1.8) with a constant viscosity, i.e. (1.4). One may also think of an extension to the generalized viscosity (1.5) or even more advanced models, which can be seen as a future work. We shall point out that the two dimensional model problem was already studied in [39].

This chapter is organized as follows: After the introduction of the model problem and a suitable Sobolev space for the control we discuss the realization of the cost functional. Therefore we introduce the vector valued Steklov–Poincaré operator via a Laplace problem, which allows us to rewrite the norm in $\widetilde{H}^{1/2}(\Gamma_C)$. Further, we present the first order necessary optimality conditions in form of the optimality system, for which we introduce a finite element discretization. In particular we consider a stabilized finite element method, as discussed in Chapter 1, which allows a lowest order approximation. Several numerical results illustrate the advantages of the energy control approach. The optimization of the wall shear stress, or other hemodynamic indicators, will be discussed at the end of this chapter.

Let $\Omega \subset \mathbb{R}^n$ $(n = 2,3)$ be a bounded Lipschitz domain with boundary $\Gamma = \partial\Omega$. As a model problem we consider the stationary Navier–Stokes equations with a constant viscosity $\nu > 0$ and mixed boundary conditions on mutually different parts of the boundary, i.e. a Dirichlet boundary Γ_D, a Neumann boundary Γ_N, and a control boundary Γ_C, all of positive measure. As it was mentioned previously in Chapter 1, these parts describe the inflow and outflow boundary, as well as the arterial wall. For the given data we assume the optimal state $\underline{\overline{u}} \in L^2(\Omega)^n$, the force term $\underline{f} \in \widetilde{H}^{-1}(\Omega)^n$ and a given inflow, which is given via $\underline{\widetilde{g}} \in H^{1/2}(\Gamma_D)^n$, as described in Chapter 1.

The corresponding optimal control problem is given as follows: Minimize the cost functional

$$\mathcal{J}(\underline{u},\underline{z}) := \frac{1}{2}\|\underline{u}-\overline{\underline{u}}\|^2_{L^2(\Omega)} + \frac{1}{2}\varrho|\underline{z}|^2_{H^{1/2}(\Gamma_C)} \tag{4.1}$$

subject to the constraint

$$\begin{aligned} -\nu\Delta\underline{u} + (\nabla\underline{u})\underline{u} + \nabla p &= \underline{f} && \text{in } \Omega,\\ \operatorname{div}\underline{u} &= 0 && \text{in } \Omega,\\ \underline{u} &= \widetilde{\underline{g}} && \text{on } \Gamma_D,\\ \nu\partial_n\underline{u} - p\underline{n} &= \underline{0} && \text{on } \Gamma_N,\\ \underline{u} &= \underline{z} && \text{on } \Gamma_C. \end{aligned} \tag{4.2}$$

Note that this optimal control problem is one of the simplest to consider for the construction of optimal inflow profiles, with focus on the minimization of the vortex in the flow field. Other choices for the cost functional are possible. For instance, one can consider the optimal state $\overline{\underline{u}}$ as the solution of the Stokes equations, or replace the argument in the tracking term by curl $\underline{u}$, see, e.g., [47]. Additionally we may also apply box constraints for the control z, see, e.g., [19].

In order to define a suitable Sobolev space for the control z, we need to specify the intersections of the different boundary parts. Namely, for our particular application, the blood flow optimization, we assume for the control boundary

$$\overline{\Gamma}_D \cap \overline{\Gamma}_C \neq \emptyset, \qquad \overline{\Gamma}_N \cap \overline{\Gamma}_C = \emptyset.$$

This then motivates for the control $\underline{z}$ to introduce the space

$$\widetilde{H}^{1/2}(\Gamma_C)^n = \left\{ \underline{v} = \widetilde{\underline{v}}_{|\Gamma_C} : \widetilde{\underline{v}} \in H^{1/2}(\Gamma)^n,\ \operatorname{supp}\widetilde{\underline{v}} \subset \overline{\Gamma}_C \right\}.$$

Note that the dual space is

$$H^{-1/2}(\Gamma_C)^n = [\widetilde{H}^{1/2}(\Gamma_C)^n]^*,$$

see also [73]. The choice of the control space $\widetilde{H}^{1/2}(\Gamma_C)^n$ needs to be reflected in the cost functional (4.1). This means that we replace the semi-norm by $|\underline{z}|_{\widetilde{H}^{1/2}(\Gamma_C)}$, where the realization of this particular semi-norm needs to be discussed. For this purpose we introduce the following vector valued Laplace problem

$$\begin{aligned} -\Delta\underline{u}_z &= \underline{0} && \text{in } \Omega,\\ \underline{u}_z &= \underline{0} && \text{on } \Gamma_D,\\ \partial_n\underline{u}_z &= \underline{0} && \text{on } \Gamma_N,\\ \underline{u}_z &= \underline{z} && \text{on } \Gamma_C. \end{aligned} \tag{4.3}$$

Instead, we may also think of realizing the semi-norm by the harmonic Stokes equations or an Oseen-type problem. These alternatives are not discussed here. Note that this choice can nevertheless be of great importance for the construction of robust preconditioners.

For the boundary value problem (4.3) we obtain the following variational formulation: Find $\underline{u}_z \in H^1_0(\Omega,\Gamma_D)^n$ with $\underline{u}_z = \underline{z}$ on Γ_C such that

$$\langle \nabla \underline{u}_z, \nabla \underline{v} \rangle_{L^2(\Omega)} = 0,$$

for all $\underline{v} \in H^1_0(\Omega,\Gamma_D \cup \Gamma_C)^n$. This problem has a unique solution $\underline{u}_z \in H^1_0(\Omega,\Gamma_D)^n$ for a given $\underline{z} \in \widetilde{H}^{1/2}(\Gamma_C)^n$. Green's first formula

$$0 = \langle -\Delta \underline{u}_z, \underline{v} \rangle_\Omega = \langle \nabla \underline{u}_z, \nabla \underline{v} \rangle_{L^2(\Omega)} - \langle \partial_n \underline{u}_z, \underline{v}_{|\Gamma} \rangle_{\Gamma_C},$$

for all $\underline{v} \in H^1_0(\Omega,\Gamma_D)^n$, then motivates, as in Section 3.1 for the Poisson equation, the following definition of the semi-norm

$$|\underline{z}|^2_{\widetilde{H}^{1/2}(\Gamma_C)} := \langle \partial_n \underline{u}_z, \underline{z} \rangle_{\Gamma_C} = \langle \nabla \underline{u}_z, \nabla \underline{u}_z \rangle_{L^2(\Omega)} = |\underline{u}_z|^2_{H^1(\Omega)},$$

for all $\underline{z} \in \widetilde{H}^{1/2}(\Gamma_C)^n$. Now, we introduce the Steklov–Poincaré operator S, as a mapping $S : \widetilde{H}^{1/2}(\Gamma_C)^n \to H^{-1/2}(\Gamma_C)^n$,

$$S\underline{z} := \partial_n \underline{u}_z, \tag{4.4}$$

which realizes the Dirichlet to Neumann map for the boundary value problem (4.3). As a consequence we obtain for the semi-norm, using the Steklov–Poincaré operator (4.4), the following relation

$$|\underline{z}|^2_{\widetilde{H}^{1/2}(\Gamma_C)} = \langle S\underline{z}, \underline{z} \rangle_{\Gamma_C}, \tag{4.5}$$

for all $\underline{z} \in \widetilde{H}^{1/2}(\Gamma_C)^n$. It is further important to mention that the semi-norm $|\cdot|_{\widetilde{H}^{1/2}(\Gamma_C)}$ is an equivalent norm in $\widetilde{H}^{1/2}(\Gamma_C)^n$. Let us summarize the properties of the Steklov–Poincaré operator.

Proposition 4.1. *The Steklov–Poincaré operator, defined in* (4.4)*, is self-adjoint, bounded and elliptic in* $\widetilde{H}^{1/2}(\Gamma_C)^n$.

In particular the cost functional (4.1), using the relation (4.5), ca be written as

$$\mathcal{J}(\underline{u},\underline{z}) := \frac{1}{2}\|\underline{u} - \overline{\underline{u}}\|^2_{L^2(\Omega)} + \frac{1}{2}\varrho \langle S\underline{z}, \underline{z} \rangle_{\Gamma_C}.$$

4.1 Optimality system

In order to obtain the first order necessary optimality conditions for the optimal control problem (4.1)–(4.2) we apply the formal Lagrange method, see, e.g., [81]. Introducing the adjoint velocity $\underline{w}$ and the adjoint pressure r, these conditions are given by the following optimality system:

Primal problem

$$\begin{aligned}
-\nu\Delta\underline{u} + (\nabla\underline{u})\underline{u} + \nabla p &= \underline{f} && \text{in } \Omega,\\
\operatorname{div}\underline{u} &= 0 && \text{in } \Omega,\\
\underline{u} &= \underline{g} && \text{on } \Gamma_{\mathrm{D}},\\
\nu\partial_n\underline{u} - p\underline{n} &= \underline{0} && \text{on } \Gamma_{\mathrm{N}},\\
\underline{u} &= \underline{z} && \text{on } \Gamma_{\mathrm{C}},
\end{aligned}$$

Adjoint problem (4.6)

$$\begin{aligned}
-\nu\Delta\underline{w} - (\nabla\underline{w})\underline{u} - (\nabla\underline{w})^{\top}\underline{u} - \nabla r &= \underline{u} - \overline{\underline{u}} && \text{in } \Omega,\\
\operatorname{div}\underline{w} &= 0 && \text{in } \Omega,\\
\underline{w} &= \underline{0} && \text{on } \Gamma_{\mathrm{D}}\cup\Gamma_{\mathrm{C}},\\
\nu\partial_n\underline{w} + (\underline{u}\cdot\underline{w})\underline{n} + (\underline{u}\cdot\underline{n})\underline{w} + r\underline{n} &= \underline{0} && \text{on } \Gamma_{\mathrm{N}},
\end{aligned}$$

Optimality condition

$$-\nu\partial_n\underline{w} - (\underline{u}\cdot\underline{w})\underline{n} - (\underline{u}\cdot\underline{n})\underline{w} - r\underline{n} + \varrho S\underline{z} = \underline{0} \qquad \text{on } \Gamma_{\mathrm{C}}.$$

We would like to mention that there exists an alternative formulation of the adjoint equations given by

$$\begin{aligned}
-\nu\Delta\underline{w} - (\nabla\underline{w})\underline{u} + (\nabla\underline{u})^{\top}\underline{w} - \nabla r &= \underline{u} - \overline{\underline{u}} && \text{in } \Omega,\\
\operatorname{div}\underline{w} &= 0 && \text{in } \Omega,\\
\underline{w} &= \underline{0} && \text{on } \Gamma_{\mathrm{D}}\cup\Gamma_{\mathrm{C}},\\
\nu\partial_n\underline{w} + (\underline{u}\cdot\underline{w})\underline{n} - (\underline{w}\cdot\underline{n})\underline{u} + r\underline{n} &= \underline{0} && \text{on } \Gamma_{\mathrm{N}}.
\end{aligned}$$

The form of the adjoint equations depends on the sequence of integration by parts and linearizing the constraint (4.2). It turns out that the first formulation of the adjoint equations in (4.6) is less restrictive in terms of regularity of the velocity $\underline{u}$, which can be advantageous. Due to this reason we use the formulation (4.6).

4.2 Variational formulation and discretization

In the following we discuss the variational formulation of the optimality system (4.6) and introduce a stabilized finite element discretization. In particular, we discuss the finite element approximation of the Steklov–Poincaré operator S, for the details we refer to [39].

For the optimality system (4.6) we obtain the following variational formulation, see also [39]. Find $(\underline{u}, p, \underline{w}, r, \underline{z}) \in H^1(\Omega)^n \times L^2(\Omega) \times H_0^1(\Omega, \Gamma_D \cup \Gamma_C)^n \times L^2(\Omega) \times \widetilde{H}^{1/2}(\Gamma_C)^n$, with $\underline{u} = \underline{g}$ on Γ_D such that

$$\begin{aligned}
-\langle \underline{u}, \underline{\sigma} \rangle_{L^2(\Omega)} \quad & + a(\underline{w}, \underline{\sigma}) + a_1^*(\underline{w}, \underline{u}, \underline{\sigma}) + b(\underline{\sigma}, r) && = -\langle \overline{\underline{u}}, \underline{\sigma} \rangle_{L^2(\Omega)}, \\
& b(\underline{w}, s) && = 0, \\
a(\underline{u}, \underline{v}) + a_1(\underline{u}, \underline{u}, \underline{v}) \quad & - b(\underline{v}, p) && = \langle \underline{f}, \underline{v} \rangle_\Omega, \\
b(\underline{u}, q) & && = 0, \\
\langle \underline{u}, \mathcal{E}\underline{\varphi} \rangle_\Omega - a_1^*(\underline{w}, \underline{u}, \mathcal{E}\underline{\varphi}) \quad & - a(\underline{w}, \mathcal{E}\underline{\varphi}) \quad - b(\mathcal{E}\underline{\varphi}, r) & \\
& + \varrho \langle S\underline{z}, \underline{\varphi} \rangle_{\Gamma_C} = \langle \overline{\underline{u}}, \mathcal{E}\underline{\varphi} \rangle_{L^2(\Omega)}, &
\end{aligned} \tag{4.7}$$

for all $(\underline{v}, q, \underline{\sigma}, s, \underline{\varphi}) \in H_0^1(\Omega, \Gamma_D)^n \times L^2(\Omega) \times H_0^1(\Omega, \Gamma_D \cup \Gamma_C)^n \times L^2(\Omega) \times \widetilde{H}^{1/2}(\Gamma_C)^n$. The corresponding linear forms are given by

$$a(\underline{u}, \underline{v}) = \nu \langle \nabla \underline{u}, \nabla \underline{v} \rangle_{L^2(\Omega)}, \qquad a_1(\underline{w}, \underline{u}, \underline{v}) = \langle (\nabla \underline{u}) \underline{w}, \underline{v} \rangle_{L^2(\Omega)}, \qquad b(\underline{u}, q) = \langle \operatorname{div} \underline{u}, q \rangle_{L^2(\Omega)},$$
$$a_1^*(\underline{w}, \underline{u}, \underline{v}) = \langle (\nabla \underline{v}) \underline{u}, \underline{w} \rangle_{L^2(\Omega)} + \langle (\nabla \underline{u})^\top \underline{w}, \underline{v} \rangle_{L^2(\Omega)} + \langle \underline{w} \cdot \underline{u}, \operatorname{div} \underline{v} \rangle_{L^2(\Omega)}.$$

Note that the uniqueness of the primal and adjoint pressure is guaranteed by the Neumann boundary conditions.

To the nonlinear variational problem of the optimality system (4.6) we apply a standard Newton method, which leads to a linear variational formulation, see Chapter 1 and [39].

In the following we discuss a finite element discretization of the variational formulation (4.7) of the optimality system. As in the previous section, we are interested in a low order discretization of the problem, to lower the number of degrees of freedom. We consider an admissible, shape-regular and quasi-uniform triangulation $\mathcal{T}_h$ of the domain Ω and the finite element spaces (1.13), i.e.,

$$\mathcal{V}_h \subset H^1(\Omega)^n, \quad \mathcal{Q}_h \subset L^2(\Omega),$$

both of piecewise linear and globally continuous finite elements. Since the finite element pairing $(\mathcal{V}_h, \mathcal{Q}_h)$ does not satisfy the discrete inf–sup condition we consider a suitable stabilization. As discussed in Chapter 1, we use the Dohrmann–Bochev stabilization (1.14), see also [6], being advantageous in computations in comparison to other stabilized finite element methods. Applied to the discrete variational formulation of (4.7) the term

$$c(q_h, p_h) = \frac{1}{\nu} \langle p_h - Q_h^0 p_h, q_h - Q_h^0 q_h \rangle_{L^2(\Omega)},$$

is added in the second and fourth equation, where Q_h^0 denotes the standard $L^2(\Omega)$ projection onto the space of piecewise constants, see also [39].

It remains to discuss the discretization of the Steklov–Poincaré operator S. As we have already mentioned in the beginning of this section, its realization is done via the vector

valued Laplace problem (4.3). For the discretization, we consider the finite element space with zero Dirichlet boundary conditions, given by

$$\mathcal{V}_{h,0} \subset H_0^1(\Omega, \Gamma_D)^n.$$

The discrete variational formulation of problem (4.3) is then given as follows. Find $\underline{u}_{z,h} \in \mathcal{V}_{h,0}$ with $\underline{u}_{z,h} = \underline{z}_h$ on Γ_C such that

$$\langle \nabla \underline{u}_{z,h}, \nabla \underline{v}_h \rangle_{L^2(\Omega)} = \langle \partial_n \underline{u}_{z,h}, \underline{v}_h \rangle_{\Gamma_C},$$

for all $\underline{v}_h \in \mathcal{V}_{h,0}$. According to the separation of interior and boundary degrees of freedom, where we denote by index I interior and Neumann degrees of freedom and by C the degrees of freedom of the control on Γ_C, we can split $\underline{u}_z = (\underline{u}_{z,I}, \underline{z}_C)^\top \leftrightarrow \underline{u}_{z,h} \in \mathcal{V}_{h,0}$. Further, we obtain with the isomorphism $\underline{z}_C \leftrightarrow \underline{z}_h \in Z_h = \mathcal{V}_{h|\Gamma_C}$ the following equivalent linear system

$$\begin{pmatrix} A_{II} & A_{IC} \\ A_{CI} & A_{CC} \end{pmatrix} \begin{pmatrix} \underline{u}_{z,I} \\ \underline{z}_C \end{pmatrix} = \begin{pmatrix} \underline{0} \\ S_h \underline{z}_C \end{pmatrix}. \tag{4.8}$$

Note that $S_h \underline{z}_C$ denotes the discrete Neumann data on the control boundary Γ_C. Deriving the Schur complement with respect to the discrete control $\underline{z}_C$ leads to the following Galerkin matrix of the discrete Steklov–Poincaré operator,

$$S_h = A_{CC} - A_{CI} A_{II}^{-1} A_{IC}.$$

The advantage of the approach is that the implementation of any boundary integrals is not necessary, even though a boundary control problem is considered.

At last, we would like to mention that in a practical implementation, we may not realize the discrete Steklov–Poincaré operator due to the inversion of A_{II}^{-1}. We rather prefer the computation of the additional unknown $\underline{u}_{z,I}$ with the representation (4.8) for S_h. The disadvantage is that the number of degrees of freedom increases, which is nevertheless rather small compared to the total number of degrees of freedom.

4.3 Numerical results

As a numerical example we consider the control of the inflow in an artery, see Figure 4.1. The model problem consists of a host artery, the lower left one, where we impose a parabolic inflow boundary condition $\underline{g}$ with $|\underline{g}| \leq 1$. On the upper left arterial part, which models the bypass, we consider the control boundary Γ_C. Further, an aneurysm sack is considered and three outlets on the right part of the artery, where we impose a do-nothing boundary condition, i.e., a Neumann boundary condition. On the arterial wall we consider a no-slip boundary condition, i.e. zero Dirichlet boundary condition. Further, we consider the following given data

$$\underline{f} = \underline{0}, \quad \nu = 0.04, \quad \overline{\underline{u}} = (0,1,0)^\top, \quad \varrho = 1,$$

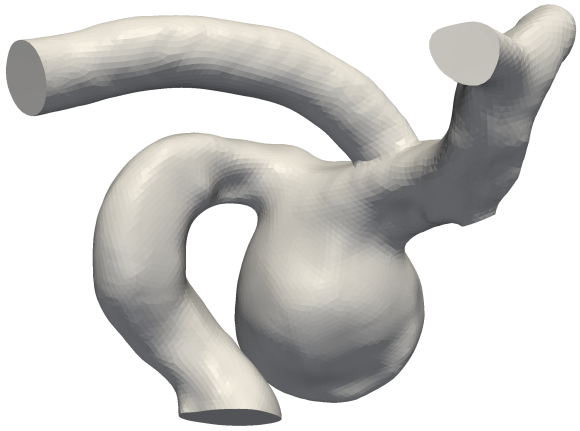


Figure 4.1: Geometry with host artery, bypass and aneurysm sack.

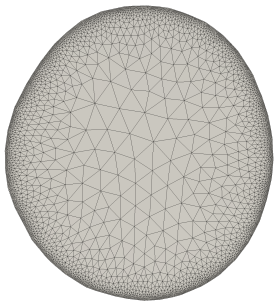
which corresponds to a Reynolds number $Re \approx 100$.

In order to demonstrate the advantages of the method we compare the results with those obtained by the $L^2(\Gamma_C)^n$ control approach. The geometry in Figure 4.1 is triangulated into 385 606 tetrahedra, where an adaptive refinement on the control boundary Γ_C is considered, see Figure 4.2.1. The global linear system is solved by the direct solver Pardiso, see [7]. The total number of degrees of freedom (DoFs) are presented in Table 4.1 for both approaches.

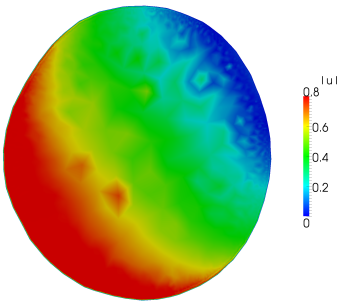
	Total DoFs	Control DoFs
$L^2(\Gamma_C)^n$	500 008	3 632
$\widetilde{H}^{1/2}(\Gamma_C)^n$	667 075	3 632

Table 4.1: Total number of degrees of freedom (DoFs) for the different approaches and degrees of freedom on the control boundary Γ_C.

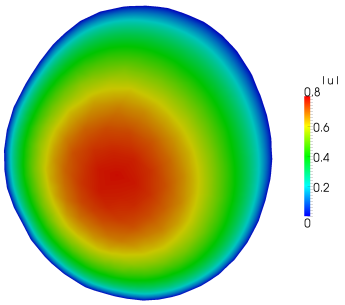
In Figure 4.2, the solution profiles on the control boundary Γ_C are depicted, i.e., the magnitudes of the controls $\underline{z}_h$ for both approaches. In the case of the (classical) $L^2(\Gamma_C)^n$ control approach we obtain, as shown in subfigure 4.2.2, a non-physical inflow profile where, in addition, a boundary layer occurs due to the no-slip condition on the arterial wall. On the other hand, for the $\widetilde{H}^{1/2}(\Gamma_C)^n$ case in subfigure 4.2.3 we obtain a smooth inflow profile of physical relevance. Another important factor, showing that the $\widetilde{H}^{1/2}(\Gamma_C)^n$ gives better results with respect to the optimization, is the vortex formation in the aneurysm. The Figure 4.3 clearly shows that the new approach can significantly minimize the vortex (which is actually the aim), while in the classical approach the vortex of the flow is preserved. The reason for the non-physical solution in the case of an $L^2(\Gamma_C)^n$ control approach can be explained as follows. If we consider a function in $L^2(\Gamma_C)^n$, there exists no continuous extension to $H^1(\Omega)^n$, see the classical trace theorems, stated in [11,30,33,73]. Further, the no-slip boundary condition, i.e., continuity on $I = \overline{\Gamma}_C \cap \overline{\Gamma}_D$ can not be imposed. This is the



4.2.1: Adaptive mesh on Γ_C.

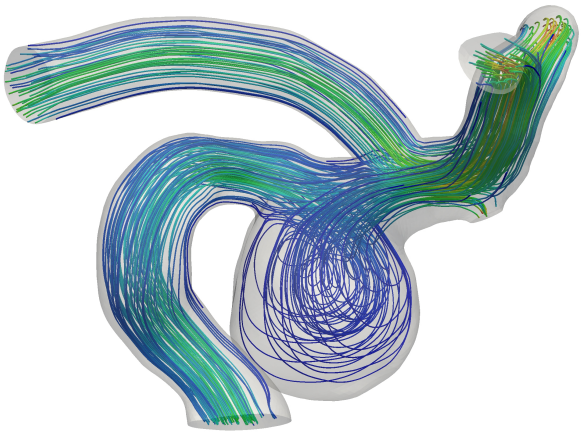


4.2.2: $L^2(\Gamma_C)^n$ control.

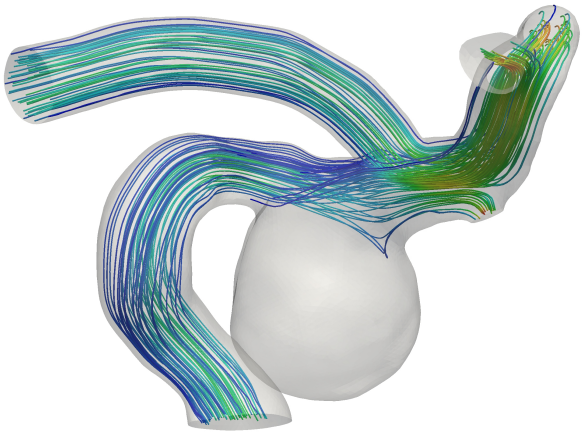


4.2.3: $\widetilde{H}^{1/2}(\Gamma_C)^n$ control.

Figure 4.2: Mesh and solutions at the control boundary Γ_C.



4.3.1: $L^2(\Gamma_C)^n$ control.



4.3.2: $\widetilde{H}^{1/2}(\Gamma_C)^n$ control.

Figure 4.3: Streamlines of the velocity $\underline{u}$ and vortex formation.

advantage of considering the control in $\widetilde{H}^{1/2}(\Gamma_C)^n$, which is the natural trace space if the velocity is considered in $H^1(\Omega)^n$. For a further discussion on the differences between these approaches we refer to [39, 59].

4.4 Concluding remarks

In this section, we have presented a Dirichlet boundary control approach in the energy space for a blood flow related application. In particular, we have controlled the inflow velocity (inflow profile) at a part of the boundary. The numerical examples demonstrate the advantages of this method.

We would like to mention, that there exists a similar approach where the more regular space $H_0^1(\Gamma_C)^n$ for the control is considered. The corresponding cost functional would be then of the form

$$\mathcal{J}(\underline{u},\underline{z}) := \frac{1}{2}\|\underline{u}-\overline{\underline{u}}\|^2_{L^2(\Omega)} + \frac{1}{2}\varrho|\underline{z}|^2_{H^1(\Gamma_C)},$$

where the semi-norm can be simply realized by

$$|\underline{z}|^2_{H^1(\Gamma_C)} = \langle \nabla \underline{z}, \nabla \underline{z}\rangle_{L^2(\Gamma_C)}.$$

Consequently, the introduction of the Steklov–Poincaré operator is not necessary which simplifies the method. On the other hand, this approach requires a priori, a more regular velocity, i.e., $\underline{u} \in H^{3/2}(\Omega)$. This extra regularity is of course questionable for some problems, for instance when the domain has sharp re-entrant corners.

As an outlook we would like to mention the optimization of the wall shear stress (WSS). As it was introduced in Chapter 2, the wall shear is calculated from the wall shear stress vector $\underline{\tau}_w$, and thus is in general in $H^{-1/2}(\Gamma)^n$. Let us consider the desired wall shear stress by $\overline{\underline{\tau}}_w$, which might be for instance a constant (of physiological value). Further, we denote by $\Gamma_O \subset \Gamma$, the observation boundary, which is for instance the wall of the aneurysm. Then the optimal control problem could be given as follows. Minimize the cost functional

$$\mathcal{J}(\underline{u},\underline{z}) := \frac{1}{2}\|\underline{\tau}_w-\overline{\underline{\tau}}_w\|^2_{H^{-1/2}(\Gamma_O)} + \frac{1}{2}\varrho|\underline{z}|^2_{H^{1/2}(\Gamma_C)},$$

subject to the constraint, the Navier–Stokes equations (4.2). We think that this optimal control problem, applied for the case of an aneurysm with bypass, can bring a better understanding of the bypass inflow realization. At this point, it is clear that the correct derivation of the wall shear stress is essential for a reasonable optimal control setting. Moreover, the above model problem is from a mathematical point of view a nice combination of the Dirichlet and Neumann boundary control approaches. In the case that the constraint is the Poisson equation, this problem is also known as the Cauchy problem, with the aim of recovering the unknown boundary condition on a part of the boundary.

As a further extension of this work we would like to mention the analysis and corresponding numerical analysis for the nonlinear optimal control problem (4.1)–(4.2), including error estimates.

Finally, we would like to mention the development of preconditioners for such problems, in order to compute the numerical results for the simulation in a reasonable time. As a first step in this direction, we consider in the upcoming chapters robust preconditioners for the optimal boundary control problem with the Poisson equation as a constraint.

5 THE BIHARMONIC EQUATION

In this chapter we aim to analyze a mixed finite element discretization of the biharmonic equation of first kind, which has applications in fluid/solid mechanics and, as we have seen in Remark 3.1, also in PDE constraint optimization. In particular it describes the optimal solution in the limit case, when the cost coefficient is $\varrho = 0$. This, as we will see in the forthcoming chapters, will be important for the construction of preconditioners.

For a standard discretization of the biharmonic equation of first kind the solution space $H_0^2(\Omega)$ is required, see, e.g., [11,16]. Thus, for a conforming discretization of the formulation in $H_0^2(\Omega)$, special finite elements, such as the Argyris element, are in general needed. In order to reduce the regularity of the solution space, we consider a mixed formulation of the biharmonic equation of first kind, where an additional unknown is introduced. This method was first introduced and analyzed in [13, 17]. Its main advantage lies in the possibility of using lowest order finite elements, i.e. standard piecewise linear and globally continuous ones. Nevertheless, the disadvantage of this formulation is that existence and uniqueness results for the solution are only obtained when the domain is a convex polygon. This issue will be discussed in the following.

The chapter is organized as follows: In the first section we show the existence and uniqueness of a solution for the mixed problem in the continuous setting, under the assumption that the domain is a convex polygon. Afterwards, we introduce a lowest order finite element method by using piecewise linear and globally continuous shape functions. We prove the existence and uniqueness of a corresponding discrete solution and comment on error estimates. Several numerical examples are presented.

5.1 Variational formulation

Let $\Omega \subset \mathbb{R}^n$ $(n = 2,3)$ be a bounded Lipschitz domain with boundary $\Gamma = \partial\Omega$, and let $f \in H^{-2}(\Omega)$ be given. The biharmonic equation of first kind for the unknown p is given by

$$\begin{aligned} \Delta^2 p &= f \qquad \text{in } \Omega, \\ p = \partial_n p &= 0 \qquad \text{on } \Gamma. \end{aligned} \tag{5.1}$$

The standard variational formulation in $H_0^2(\Omega)$ is obtained as follows. We multiply the equation above with a test function $q \in H_0^2(\Omega)$ and apply integration by parts twice, which lead to the following variational formulation. Find $p \in H_0^2(\Omega)$ such that

$$\langle \Delta p, \Delta q \rangle_{L^2(\Omega)} = \langle f, q \rangle_\Omega, \tag{5.2}$$

for all $q \in H_0^2(\Omega)$. By using the fact that $\|\Delta q\|_{L^2(\Omega)}$ defines an equivalent norm in $H_0^2(\Omega)$ we obtain the existence and uniqueness of a solution $p \in H_0^2(\Omega)$ of the variational formulation (5.2), see also [16, 33].

Remark 5.1. *In the particular case that $f \in H^{-1}(\Omega)$ and the domain Ω is a convex polygon we have for the solution of* (5.2) *that $p \in H_0^2(\Omega) \cap H^3(\Omega)$, see, e.g.,* [12, p. 164] *and* [33].

As it was pointed out at the beginning of this chapter, it might be of interest to reduce the regularity of the solution space $H_0^2(\Omega)$, which would give us the possibility to use a lowest order finite element approximation. Therefore we introduce an additional unknown $u = -\Delta p$. Multiplication with a test function $v \in H^1(\Omega)$, applying integration by parts and using the boundary condition $\partial_n p = 0$ on Γ lead to

$$0 = \langle u, v \rangle_{L^2(\Omega)} + \langle \Delta p, v \rangle_\Omega = \langle u, v \rangle_{L^2(\Omega)} - \langle \nabla p, \nabla v \rangle_{L^2(\Omega)}.$$

For the remaining equation, i.e. $-\Delta u = f$, we obtain with a test function $q \in H_0^1(\Omega)$ and applying integration by parts

$$\langle f, q \rangle_\Omega = \langle -\Delta u, q \rangle_{L^2(\Omega)} = \langle \nabla u, \nabla q \rangle_{L^2(\Omega)}.$$

Hence we conclude the following mixed variational formulation of the boundary value problem (5.1) in saddle point form: Find $(u, p) \in H^1(\Omega) \times H_0^1(\Omega)$ such that

$$\begin{aligned} a(u, v) - b(v, p) &= 0, \\ b(u, q) &= \langle f, q \rangle_\Omega, \end{aligned} \tag{5.3}$$

for all $(v, q) \in H^1(\Omega) \times H_0^1(\Omega)$, where the bilinear forms are given by

$$a(u, v) = \langle u, v \rangle_{L^2(\Omega)}, \qquad b(v, q) = \langle \nabla v, \nabla q \rangle_{L^2(\Omega)}.$$

As first, we observe that the bilinear form $a(\cdot, \cdot)$ is not elliptic in $H^1(\Omega)$. We shall note that for the existence and uniqueness of a solution, using the saddle point theory, only the ellipticity on the kernel

$$\operatorname{Ker} B := \left\{ v \in H^1(\Omega) : b(v, q) = 0 \quad \text{for all } q \in H_0^1(\Omega) \right\} \subset H^1(\Omega),$$

is required. This is nevertheless also not valid for the bilinear form $a(\cdot, \cdot)$, see also [13]. As a consequence the classical theory of saddle point problems is not applicable. Note that the kernel $\operatorname{Ker} B$ describes all harmonic functions which are the solutions of the Laplace equation.

In the case that the domain Ω is a convex polygon, the existence and uniqueness of a solution of the formulation (5.3) can be proven, see also [12,17,33]. This result is formulated in the following theorem.

Theorem 5.1. *Let Ω be a convex polygon and $f \in H^{-1}(\Omega)$. Then for the variational formulation* (5.3) *there exists a unique solution $(u,p) \in H^1(\Omega) \times H_0^1(\Omega)$. Moreover we have $p \in H_0^2(\Omega)$, which is the unique solution of* (5.2).

Proof. First, we prove the existence of a solution. From Remark 5.1 we have the existence and uniqueness of $p \in H_0^2(\Omega) \cap H^3(\Omega)$ satisfying $\Delta^2 p = f$ in the sense of $H^{-1}(\Omega)$. Further we define $u = -\Delta p$ and conclude due to the regularity of p that $u \in H^1(\Omega)$. This gives the first equation of (5.3). The remaining term $-\Delta u = f$ in $H^{-1}(\Omega)$ leads to the second equation. As a consequence we have the existence of $(u,p) \in H^1(\Omega) \times H_0^1(\Omega)$ which satisfies (5.3).

Now, let $(u,p) \in H^1(\Omega) \times H_0^1(\Omega)$ be one solution of (5.3). From the first equation of the variational formulation (5.3), follows

$$\langle \nabla p, \nabla q \rangle_{L^2(\Omega)} = \langle u, q \rangle_{L^2(\Omega)},$$

for all $q \in H_0^1(\Omega)$, due to the inclusion $H_0^1(\Omega) \subset H^1(\Omega)$. This problem has for each given $u \in H^1(\Omega)$ clearly a unique solution $p \in H_0^1(\Omega)$, which is depending on $u \in H^1(\Omega)$. Since we assumed the domain Ω to be a convex polygon, we can conclude that $p \in H^2(\Omega) \cap H_0^1(\Omega)$ and that the equation

$$-\Delta p = u, \tag{5.4}$$

is valid in the sense of $L^2(\Omega)$, see, e.g., [33]. Consequently, we obtain by Green's first formula and the first equation of (5.3) the relation

$$\langle \partial_n p, v \rangle_\Gamma = \langle \nabla p, \nabla v \rangle_{L^2(\Omega)} - \langle u, v \rangle_{L^2(\Omega)} = 0,$$

for all $v \in H^1(\Omega)$, from which we conclude $\partial_n p = 0$ on Γ. Thus we have $p \in H_0^2(\Omega)$.

Now we consider the second equation of (5.3) for all test functions $q \in H_0^2(\Omega)$, since we have the inclusion $H_0^2(\Omega) \subset H_0^1(\Omega)$. By applying integration by parts we obtain

$$\langle -u, \Delta q \rangle_{L^2(\Omega)} = \langle f, q \rangle_\Omega,$$

for all $q \in H_0^2(\Omega)$. Now we can insert (5.4) and obtain

$$\langle \Delta p, \Delta q \rangle_{L^2(\Omega)} = \langle f, q \rangle_\Omega,$$

for all $q \in H_0^2(\Omega)$. This problem has now a unique solution $p \in H_0^2(\Omega)$. In particular we have $p \in H_0^2(\Omega) \cap H^3(\Omega)$, see Remark 5.1. As a consequence, we also have a unique auxiliary variable by the relation $u = -\Delta p \in H^1(\Omega)$. This concludes the proof. □

As we have seen, the main disadvantage of the formulation (5.3) is the lack of the ellipticity of the bilinear form $a(\cdot,\cdot)$. In order to apply the theory of saddle point problems a new formulation was recently proposed in [86], where the solution space for the auxiliary variable u is enriched. This means, it is based on a formulation in the space

$$H_{\Delta}^{-1}(\Omega) = \left\{ v \in L^2(\Omega) : \Delta v \in H^{-1}(\Omega) \right\},$$

equipped with the corresponding norm

$$\|v\|_{H_{\Delta}^{-1}(\Omega)} = \left(\|v\|_{L^2(\Omega)}^2 + \|\Delta v\|_{H^{-1}(\Omega)}^2 \right)^{1/2}.$$

The corresponding variational formulation for the problem (5.1) is then given as follows. Find $(u,p) \in H_{\Delta}^{-1}(\Omega) \times H_0^1(\Omega)$ such that

$$\begin{aligned} \langle u, v \rangle_{L^2(\Omega)} + \langle p, \Delta v \rangle_{\Omega} &= 0, \\ \langle \Delta u, q \rangle_{\Omega} &= -\langle f, q \rangle_{\Omega}, \end{aligned}$$

for all $(v,q) \in H_{\Delta}^{-1}(\Omega) \times H_0^1(\Omega)$. Note that we have for the Sobolev space $H_{\Delta}^{-1}(\Omega)$ the inclusions

$$H^1(\Omega) \subset H_{\Delta}^{-1}(\Omega) \subset L^2(\Omega).$$

For the details we refer to [86].

5.2 Discretization and error estimates

Within this section we introduce a mixed finite element discretization for the biharmonic equation of first kind (5.1). The main advantage of the variational formulation (5.3) is that we can use for instance standard piecewise linear and globally continuous finite elements. In the following we introduce a discrete variational formulation, show the existence and uniqueness of its solution and comment on corresponding error estimates. Afterwards we present some numerical examples.

Let us consider an admissible, shape-regular and globally quasi-uniform triangulation $\mathcal{T}_h$ of the domain Ω into triangles or tetrahedra, denoted by T. We introduce the finite dimensional subspaces

$$\mathcal{V}_h = \operatorname{span}\{\varphi_i^1\}_{i=1}^{n_I+n_C} \subset H^1(\Omega), \quad \mathcal{Q}_h = \operatorname{span}\{\varphi_i^1\}_{i=1}^{n_I} \subset H_0^1(\Omega), \tag{5.5}$$

where φ_i^1 denotes the piecewise linear and globally continuous shape functions, for all $i = 1, \ldots, n_I + n_C$. Note that $n_I = \dim \mathcal{Q}_h$ denotes the number of interior degrees of freedom, while n_C denotes the number of degrees of freedom on the boundary, which implies $\dim \mathcal{V}_h = n_I + n_C$.

The discrete variational formulation, corresponding to (5.3), is then given as follows. Find $(u_h, p_h) \in \mathcal{V}_h \times \mathcal{Q}_h$ such that

$$\begin{aligned} a(u_h, v_h) - b(v_h, p_h) &= 0, \\ b(u_h, q_h) &= \langle f, q_h \rangle_\Omega, \end{aligned} \tag{5.6}$$

for all $(v_h, q_h) \in \mathcal{V}_h \times \mathcal{Q}_h$.

In the following we shall investigate the existence and uniqueness of a discrete solution. In particular we discuss the choice of the finite element spaces $\mathcal{V}_h$ and $\mathcal{Q}_h$. Let us recall that the standard theory for saddle point problems is not applicable since the bilinear form $a(\cdot,\cdot)$ is neither elliptic in $H^1(\Omega)$ nor in Ker B. Nevertheless we can show an inf–sup condition for the bilinear form $b(\cdot,\cdot)$ under a certain assumption on the finite element spaces.

Lemma 5.1. *Let us assume for the finite element spaces $\mathcal{V}_h$ and $\mathcal{Q}_h$ the inclusion $\mathcal{Q}_h \subset \mathcal{V}_h$, then the following discrete inf–sup condition is valid,*

$$\sup_{0 \neq v_h \in \mathcal{V}_h} \frac{b(v_h, q_h)}{\|v_h\|_{H^1(\Omega)}} \geq \widetilde{c}_S \|q_h\|_{H^1(\Omega)},$$

for all $q_h \in \mathcal{Q}_h$.

Proof. Due to the inclusion $\mathcal{Q}_h \subset \mathcal{V}_h$ and Friedrichs inequality we have

$$\sup_{0 \neq v_h \in \mathcal{V}_h} \frac{b(v_h, q_h)}{\|v_h\|_{H^1(\Omega)}} \geq \frac{|q_h|^2_{H^1(\Omega)}}{\|q_h\|_{H^1(\Omega)}} \geq (1 + c_F^{-2})^{-1} \|q_h\|_{H^1(\Omega)} = \widetilde{c}_S \|q_h\|_{H^1(\Omega)},$$

for all $q_h \in \mathcal{Q}_h$. This concludes the proof. □

In particular we observe for the finite element spaces in (5.5) the inclusion

$$\mathcal{Q}_h \subset \mathcal{V}_h,$$

and thus we can conclude the discrete inf–sup condition of Lemma 5.1.

For the equivalent linear system of the discrete variational formulation (5.6), let us introduce the standard mass, stiffness matrices and right-hand side vector by

$$M_h[j,i] = \langle \varphi_i^1, \varphi_j^1 \rangle_{L^2(\Omega)}, \quad A_h[\ell,i] = \langle \nabla \varphi_i^1, \nabla \varphi_\ell^1 \rangle_{L^2(\Omega)}, \quad f_h[\ell] = \langle f, \varphi_\ell^1 \rangle_\Omega,$$

for all $i, j = 1, \ldots, n_I + n_C$ and $\ell = 1, \ldots, n_I$, respectively. Note that $M_h \in \mathbb{R}^{(n_I+n_C)\times(n_I+n_C)}$ and $A_h \in \mathbb{R}^{(n_I+n_C)\times n_I}$. With the isomorphisms $u_h \leftrightarrow \underline{u} \in \mathbb{R}^{n_I+n_C}$ and $p_h \leftrightarrow \underline{p} \in \mathbb{R}^{n_I}$, the corresponding equivalent linear system for the discrete variational formulation (5.6) is given by

$$\begin{pmatrix} M_h & -A_h^\top \\ A_h & \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{0} \\ \underline{f}_h \end{pmatrix}. \tag{5.7}$$

For the existence and uniqueness of a discrete solution, the following result is valid.

Theorem 5.2. *Let us consider the finite element spaces $\mathcal{V}_h$ and $\mathcal{Q}_h$, defined in* (5.5)*. Then for the discrete variational formulation* (5.6) *exists a unique solution* $(u_h, p_h) \in \mathcal{V}_h \times \mathcal{Q}_h$*.*

Proof. Let us consider the linear system (5.7), which is equivalent to the discrete variational formulation (5.6). The mass matrix M_h is obviously invertible. Further, the discrete inf–sup condition of Lemma 5.1 implies that A_h has full rank, in particular rank $A_h = n_I$. Note that this can be also seen from the fact that we have

$$A_h = (A_{II}, A_{IC}),$$

by the separation of interior and boundary degrees of freedom. The matrix $A_{II} \in \mathbb{R}^{n_I \times n_I}$, associated to the interior degrees of freedom, is of full rank since it corresponds to a Dirichlet boundary value problem. As a consequence we have that rank $M_h >$ rank A_h, since rank $M_h = n_I + n_C$. Thus the linear system has a unique solution $(\underline{u}, \underline{p})^\top$. Since we have the isomorphisms $\underline{u} \leftrightarrow u_h \in \mathcal{V}_h$ and $\underline{p} \leftrightarrow p_h \in \mathcal{Q}_h$, we have a unique solution $(u_h, p_h) \in \mathcal{V}_h \times \mathcal{Q}_h$. This concludes the proof. □

In the following we comment on error estimates for the discrete variational formulation (5.6). Note that the standard procedure for the derivation of error estimates for saddle point problems is not applicable, due to the lack of the ellipticity of the bilinear form $a(\cdot,\cdot)$. Thus other technique have to be taken into consideration. The following result is obtained by [69, Theorem 1].

Theorem 5.3. *Let Ω be a convex polygon and let the exact solution of problem* (5.3) *be regular enough, i.e. $p \in H_0^1(\Omega) \cap H^4(\Omega)$. Then there holds the error estimate*

$$\|p - p_h\|_{L^2(\Omega)} + h^{1/2} |\ln h| \|u - u_h\|_{L^2(\Omega)} \leq ch |\ln h|^2 \|p\|_{H^4(\Omega)}. \tag{5.8}$$

We shall note that the error estimate (5.8) is far away from optimal, which would mean of second order when using linear basis functions.

5.3 Numerical results

In this section we present some numerical examples for the biharmonic equation of first kind, in particular we are interested in the order of convergence of the mixed finite element discretization. Therefore we consider the discrete variational formulation (5.6) with piecewise linear and globally continuous finite elements, being equivalent to the linear system (5.7). In the following we present numerical result for the two and three dimensional model problems, where we consider as a computational domains $\Omega = (0, \frac{1}{2})^n$ and $\Omega = B_{1/2}(0)$, both

for $n = 2,3$. The corresponding exact solutions, for both computational domains, are given by

$$p = \begin{cases} 2^{-n} \prod_{i=1}^{n} (\cos(4\pi x_i) - 1) & \text{for } \Omega = (0, \frac{1}{2})^n, \\ 2^n \exp\left(\sum_{i=1}^{n} x_i^2 - \frac{1}{4} \right)^{-1} & \text{for } \Omega = B_{1/2}(0). \end{cases} \tag{5.9}$$

We shall note that the exact solution is regular.

Example 1

In the first numerical example we consider the two dimensional model problem and for both domains a uniform triangulation $\mathcal{T}_h$ with $N = 4^{L+1}$ elements for all refinement levels L. The initial triangulation into 4 elements is done via the diagonals and its refinement is obtained by congruent triangles via the edge midpoints. In the following we present discretization errors and estimated order of convergences for the different problems.

L	DoFs	$\|u - u_h\|_{L^2(\Omega)}$	eoc	$\|p - p_h\|_{L^2(\Omega)}$	eoc
0	6	7.70425 e–01	–	1.19188 e–02	–
1	18	1.24647 e+00	-0.69	2.05581 e–02	-0.79
2	66	4.46978 e–01	1.48	3.41882 e–03	2.59
3	258	1.22460 e–01	1.87	1.20219 e–03	1.51
4	1 026	4.11720 e–02	1.57	3.19848 e–04	1.91
5	4 098	1.06405 e–02	1.95	8.12152 e–05	1.98
6	16 386	2.68684 e–03	1.99	2.04068 e–05	1.99
7	65 538	6.73385 e–04	2.00	5.10939 e–06	2.00
8	262 146	1.68452 e–04	2.00	1.27774 e–06	2.00
9	1 048 578	4.21188 e–05	2.00	3.19525 e–07	2.00
10	4 194 306	1.05299 e–05	2.00	8.00297 e–08	2.00
observed			2.00		2.00

Table 5.1: Errors and eoc for $\Omega = B_{1/2}(0)$, $n = 2$.

In Table 5.1 and Table 5.2 we present the obtained numerical results. The computations were performed until level $L = 10$, where the total number of degrees of freedom (DoFs) is $2n_I + n_C$ on each refinement level L. For both considered problems we observe second order of convergence with respect to the $L^2(\Omega)$ norm for both variables.

Example 2

In order to illustrate the convergence also for the three dimensional model problem, we consider as a second example $\Omega = B_{1/2}(0)$ and $\Omega = (0, \frac{1}{2})^3$ for $n = 3$ as a computational

L	DoFs	$\|u-u_h\|_{L^2(\Omega)}$	eoc	$\|p-p_h\|_{L^2(\Omega)}$	eoc
0	6	1.37310 e+01	–	8.89720 e–02	–
1	18	9.66346 e+00	0.51	8.45632 e–02	0.07
2	66	4.08556 e+00	1.24	3.43389 e–02	1.30
3	258	1.10158 e+00	1.89	9.83510 e–03	1.80
4	1 026	2.84739 e–01	1.95	2.57842 e–03	1.93
5	4 098	7.20626 e–02	1.98	6.54275 e–04	1.98
6	16 386	1.80891 e–02	1.99	1.64289 e–04	1.99
7	65 538	4.52866 e–03	2.00	4.11010 e–05	2.00
8	262 146	1.13251 e–03	2.00	1.03032 e–05	2.00
9	1 048 578	2.83306 e–04	2.00	2.55362 e–06	2.01
10	4 194 306	7.08294 e–05	2.00	6.39816 e–07	2.00
observed			2.00		2.00

Table 5.2: Errors and eoc for $\Omega = (0, \frac{1}{2})^2$, $n = 2$.

domain. As before we consider a uniform triangulation $\mathcal{T}_h$ with $N = 32 \cdot 8^L$ for the ball and $N = 12 \cdot 8^L$ elements for all refinement level L, performed until $L = 6$.

L	DoFs	$\|u-u_h\|_{L^2(\Omega)}$	eoc	$\|p-p_h\|_{L^2(\Omega)}$	eoc
0	20	1.71952 e+00	–	1.48296 e–02	–
1	104	1.50511 e+00	0.19	1.37640 e–02	0.11
2	720	7.38096 e–01	1.03	3.97069 e–03	1.79
3	5 536	2.12025 e–01	1.80	1.22816 e–03	1.69
4	43 840	6.62657 e–02	1.68	3.44279 e–04	1.83
5	349 824	1.72605 e–02	1.94	8.84893 e–05	1.96
6	2 796 800	4.37609 e–03	1.98	2.22934 e–05	1.99
observed			2.00		2.00

Table 5.3: Errors and eoc for $\Omega = B_{1/2}(0)$, $n = 3$.

From the results in Table 5.3 for the ball $B_{1/2}(0)$ we observe, as in the previous example, second order of convergence with respect to the $L^2(\Omega)$ norm, for both variables. This, on the other hand, is not that case for the results, presented in Table 5.4 for the cube $(0, \frac{1}{2})^3$. Here we observe a reduced convergence rate for the auxiliary variable u. This means, even tough the error estimate (5.8) might not be optimal, optimal convergence rates, meaning second order, are in general not obtained. For a better understanding we have additionally plotted the error $\|u-u_h\|_{L^2(\Omega)}$ for this case in Figure 5.1. As we can see the order of convergence of the error is close to $h^{3/2}|\ln h|$. It might be possible to improve the error estimate (5.8), which can be seen as a challenging future work.

L	DoFs	$\|u-u_h\|_{L^2(\Omega)}$	eoc	$\|p-p_h\|_{L^2(\Omega)}$	eoc
0	10	1.10988 e+01	–	5.54297 e–02	–
1	44	8.33870 e+00	0.41	5.24166 e–02	0.08
2	280	5.39055 e+00	0.63	3.29991 e–02	0.67
3	2 096	2.16529 e+00	1.32	1.28502 e–02	1.36
4	16 480	7.31154 e–01	1.57	3.93101 e–03	1.71
5	131 264	2.64592 e–01	1.47	1.06426 e–03	1.89
6	1 048 960	1.13068 e–01	1.23	2.72788 e–04	1.96
observed			1.20		2.00

Table 5.4: Errors and eoc for $\Omega = (0, \frac{1}{2})^3$, $n = 3$.

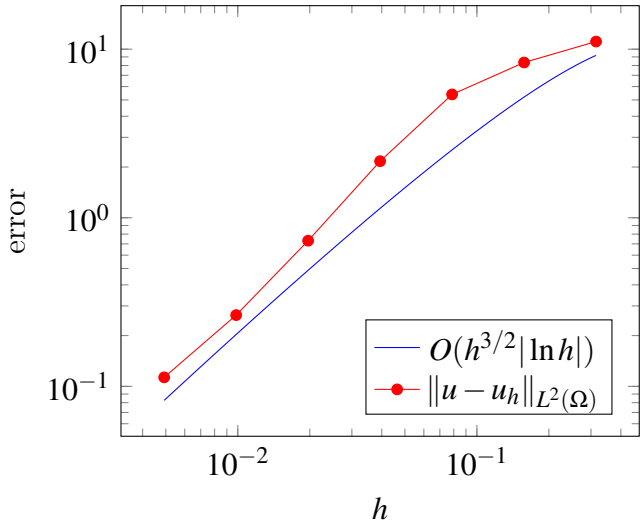


Figure 5.1: Order of convergence of the error $\|u-u_h\|_{L^2(\Omega)}$ for $\Omega = (0, \frac{1}{2})^3$, $n = 3$, Table 5.4.

5.4 Concluding remarks

In this chapter we have analyzed a mixed finite element method for the biharmonic equation of first kind. In particular we have discussed the existence and uniqueness of a solution and commented on corresponding error estimates. Numerical examples have been presented, with focus on the order of convergence.

In the upcoming chapter we discuss the construction of robust preconditioners for the mixed finite element method. We shall note that this might not be straight forward, which is due to the lack of the ellipticity of the bilinear form $a(\cdot,\cdot)$. This results shall be afterwards applied to the optimal boundary control problems which were discussed in Chapter 3.

6 PRECONDITIONING STRATEGIES FOR THE BIHARMONIC EQUATION

In this chapter, we derive a robust preconditioner for the mixed finite element formulation of the biharmonic equation of first kind (5.1). Therefore, we consider the variational formulation (5.6) where all unknown variables are discretized by piecewise linear and globally continuous finite elements. This formulation is then equivalent to the linear system (5.7), which is the starting point for the construction of the preconditioner.

The construction of preconditioners for the biharmonic problem started in the 1990's. For a review on existing preconditioners for the biharmonic equation we refer to [4, 8, 48, 61, 62]. First results on iterative methods for the mixed finite element discretization for the biharmonic equation were given for Uzawa–type methods in [48], and for a multilevel algorithm in [61], where convergence was shown assuming $H^3(\Omega)$–regularity of the solution. In [8], the authors considered a preconditioned conjugate gradient method for solving the Schur complement system with respect to the primal variable. Afterwards, in [62] it was shown that the Schur complement matrix is spectrally equivalent to a mesh depending norm where the related preconditioner is realized by a special factorization. A variable V–cycled multigrid approach was considered for piecewise quadratic or higher order shape functions in [35]. Further, a W–cycled multigrid method was analyzed with a sufficiently high number of smoothing steps. More recently in [56], an arbitrary black box multigrid approach for the biharmonic equation was studied. A different approach for the iterative solution of the mixed finite element formulation is based on the elimination of all interior degrees of freedom. This requires the solution of two Dirichlet boundary value problems for the Poisson equation, and results in a Schur complement system to find the Dirichlet datum of the dual variable, see, e.g., [17, 31].

We consider the construction of a block diagonal preconditioner, see for instance [22], where the main difficulty lies in the construction of a preconditioner for the Schur complement. It turns out, as already shown in [62], that the Schur complement with respect to the boundary is related to a harmonic extension of the boundary data. This implies an equivalent norm in the piecewise defined fractional Sobolev space $\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)$. First, we consider a representation of this particular Sobolev norm by using locally defined single layer boundary integral operators. In fact, this approach corresponds to an additive Schwarz method, see therefore [36,57]. Although this method results in a constant bound of the spectral condition number, its realization requires the inversion of a block diagonal matrix including a coarse problem. Instead, one may use multilevel representations of fractional Sobolev norms [10, 26, 58, 60], i.e. of $H^{-1/2}(\Gamma)$, which results in an almost optimal preconditioner where the spectral condition number is constant up to a logarithmic term, see also [41].

The chapter is organized as follows: In the first part we introduce fractional Sobolev spaces on the boundary which are needed for the construction of the preconditioner for the Schur complement system. Afterwards, several general spectral equivalence results are stated. It turns out that the Schur complement equation induces an equivalent norm in the fractional Sobolev space $\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)$ for which we construct two different types of preconditioners. The first is derived from boundary element methods via the single layer boundary integral operator. For the second preconditioner we consider a multilevel approach of BPX type. Afterwards we propose two different preconditioner for the global system. Several numerical examples illustrate the obtained theoretical results.

Let us recall from Chapter 5, the linear system (5.7), which is the starting point for the construction of a preconditioner. According to interior and boundary degrees of freedom we can decompose the vector $\underline{u} = (\underline{u}_I, \underline{u}_C)^\top \in \mathbb{R}^{n_I+n_C}$, where we denote by n_I and n_C the number of interior and boundary degrees of freedom, respectively. Consequently we can rewrite the linear system (5.7), by using $A_{IC} = A_{CI}^\top \in \mathbb{R}^{n_I \times n_C}$, as

$$\begin{pmatrix} M_{II} & M_{IC} & -A_{II} \\ M_{CI} & M_{CC} & -A_{CI} \\ A_{II} & A_{IC} & \end{pmatrix} \begin{pmatrix} \underline{u}_I \\ \underline{u}_C \\ \underline{p} \end{pmatrix} = \begin{pmatrix} \underline{0} \\ \underline{0} \\ \underline{f}_h \end{pmatrix},$$

or, by a simple reordering of the variables and introducing $\widetilde{\underline{p}}_I = -\underline{p}$, $\underline{f}_I = \underline{f}_h$, as

$$\begin{pmatrix} M_{II} & A_{II} & M_{IC} \\ A_{II} & & A_{IC} \\ M_{CI} & A_{CI} & M_{CC} \end{pmatrix} \begin{pmatrix} \underline{u}_I \\ \widetilde{\underline{p}}_I \\ \underline{u}_C \end{pmatrix} = \begin{pmatrix} \underline{0} \\ \underline{f}_I \\ \underline{0} \end{pmatrix}. \tag{6.1}$$

Note that the matrix of the linear system (6.1) is symmetric and indefinite. As mentioned above, the aim is the construction of a block diagonal preconditioner, which is of the form

$$\begin{pmatrix} C_{A_h} & \\ & C_{T_h} \end{pmatrix}, \tag{6.2}$$

where $C_{A_h} \in \mathbb{R}^{2n_I \times 2n_I}$ is a preconditioner for the upper left 2×2 block and $C_{T_h} \in \mathbb{R}^{n_C \times n_C}$ is a preconditioner for the Schur complement, of the matrix in (6.1). The latter one, as in most cases, is the more difficult one to construct and treated first within this section. From the linear system (6.1) we obtain

$$\underline{u}_I = A_{II}^{-1}\big[\underline{f}_I - A_{IC}\underline{u}_C\big],$$

and

$$\widetilde{\underline{p}}_I = -A_{II}^{-1}\big[M_{II}\underline{u}_I + M_{IC}\underline{u}_C\big] = -A_{II}^{-1}M_{II}A_{II}^{-1}\underline{f}_I + A_{II}^{-1}M_{II}A_{II}^{-1}A_{IC}\underline{u}_C - A_{II}^{-1}M_{IC}\underline{u}_C,$$

which results in the Schur complement system

$$\begin{aligned} \big[M_{CC} - M_{CI}A_{II}^{-1}A_{IC} - A_{CI}A_{II}^{-1}M_{IC} + A_{CI}A_{II}^{-1}M_{II}A_{II}^{-1}A_{IC}\big]\underline{u}_C \\ = \big[A_{CI}A_{II}^{-1}M_{II} - M_{CI}\big]A_{II}^{-1}\underline{f}_I. \end{aligned} \tag{6.3}$$

In order to use for example a preconditioned conjugate gradient scheme for an iterative solution of the linear system (6.3) we need to have a preconditioner C_{T_h} for the Schur complement matrix

$$T_h = M_{CC} - M_{CI}A_{II}^{-1}A_{IC} - A_{CI}A_{II}^{-1}M_{IC} + A_{CI}A_{II}^{-1}M_{II}A_{II}^{-1}A_{IC}. \tag{6.4}$$

For all $\underline{v}_C \in \mathbb{R}^{n_C}$ we rewrite the induced bilinear form as

$$\begin{aligned}(T_h\underline{v}_C, \underline{v}_C) &= (M_{CC}\underline{v}_C, \underline{v}_C) - 2(M_{CI}A_{II}^{-1}A_{IC}\underline{v}_C, \underline{v}_C) + (M_{II}A_{II}^{-1}A_{IC}\underline{v}_C, A_{II}^{-1}A_{IC}\underline{v}_C)\\ &= (M_{CC}\underline{v}_C, \underline{v}_C) + 2(M_{CI}\underline{v}_I, \underline{v}_C) + (M_{II}\underline{v}_I, \underline{v}_I),\end{aligned}$$

in which we define

$$\underline{v}_I := -A_{II}^{-1}A_{IC}\underline{v}_C \in \mathbb{R}^{n_I}.$$

By using the isomorphism $\underline{v} = (\underline{v}_I, \underline{v}_C)^\top \in \mathbb{R}^{n_I+n_C} \leftrightarrow v_h \in \mathcal{V}_h$ we finally obtain

$$(T_h\underline{v}_C, \underline{v}_C) = (M_h\underline{v}, \underline{v}) = \langle v_h, v_h\rangle_{L^2(\Omega)} = \|v_h\|^2_{L^2(\Omega)}. \tag{6.5}$$

Note, that $v_h \in \mathcal{V}_h$ is the discrete harmonic extension of $v_{h|\Gamma} \leftrightarrow \underline{v}_C \in \mathbb{R}^{n_C}$ which is the unique solution of the variational problem

$$\langle \nabla v_h, \nabla q_h\rangle_{L^2(\Omega)} = 0,$$

for all $q_h \in \mathcal{Q}_h$, where the finite element space $\mathcal{Q}_h$ is defined in (5.5). In order to construct a robust preconditioner we need to have equivalence estimates for the harmonic extension v_h in the $L^2(\Omega)$ norm.

6.1 Sobolev spaces and trace theorems

In this subsection we introduce the Sobolev spaces and trace theorems which are needed for the construction and the analysis of the preconditioners for the biharmonic equation (5.1). We shall not give a review on all details, just state the main ideas which are needed further on. For the basics we refer to [1, 30, 33, 53, 73].

Since the boundary $\Gamma = \partial\Omega$ was assumed to be piecewise smooth, and thus is decomposable into $J \in \mathbb{N}$ smooth parts, we have

$$\Gamma = \bigcup_{i=1}^{J} \overline{\Gamma}_i, \quad \Gamma_i \cap \Gamma_j = \emptyset \quad \text{for all } i \neq j,\ i,j \in \{1,\ldots,J\}.$$

Let $s \in [0, 3/2]$ and we define the Sobolev space of piecewise smooth functions

$$H^s_{\mathrm{pw}}(\Gamma) = \Big\{ v \in L^2(\Gamma) : v|_{\Gamma_i} \in H^s(\Gamma_i),\ i = 1,\ldots,J \Big\}, \tag{6.6}$$

with the corresponding norm

$$\|v\|_{H^s_{\mathrm{pw}}(\Gamma)} = \left(\sum_{i=1}^{J} \|v_{|\Gamma_i}\|^2_{H^s(\Gamma_i)} \right)^{1/2}.$$

Note that with Sobolev–Slobodeckii norm, see, e.g., [73, p. 36], we can easily show

$$\|v\|_{H^s_{\mathrm{pw}}(\Gamma)} \leq \|v\|_{H^s(\Gamma)},$$

for all $v \in H^s(\Gamma)$ and consequently have the inclusion $H^s(\Gamma) \subset H^s_{\mathrm{pw}}(\Gamma)$. We recall that for any smooth and open part $\Gamma_i \subset \Gamma$, $i = 1, \ldots, J$ we have

$$\widetilde{H}^s(\Gamma_i) = \left\{ v = \widetilde{v}_{|\Gamma_i} : \widetilde{v} \in H^s(\Gamma),\ \mathrm{supp}\ \widetilde{v} \subset \overline{\Gamma}_i \right\},$$

with the norm

$$\|v\|_{\widetilde{H}^s(\Gamma_i)} = \inf_{\widetilde{v} \in H^s(\Gamma) : \widetilde{v}_{|\Gamma_i} = v} \|\widetilde{v}\|_{H^s(\Gamma)}.$$

These definitions motivate the following space

$$\widetilde{H}^s_{\mathrm{pw}}(\Gamma) = \left\{ v \in L^2(\Gamma) : v|_{\Gamma_i} \in \widetilde{H}^s(\Gamma_i),\ i = 1, \ldots, J \right\}, \tag{6.7}$$

with the corresponding norm

$$\|v\|_{\widetilde{H}^s_{\mathrm{pw}}(\Gamma)} = \left(\sum_{i=1}^{J} \|v_{|\Gamma_i}\|^2_{\widetilde{H}^s(\Gamma_i)} \right)^{1/2}.$$

For the dual spaces we have

$$H^{-s}(\Gamma_i) = [\widetilde{H}^s(\Gamma_i)]^*, \quad \widetilde{H}^{-s}(\Gamma_i) = [H^s(\Gamma_i)]^*,$$

for all $i = 1, \ldots, J$ and $s \in [0, 3/2]$, respectively. Moreover, the dual spaces of (6.6) and (6.7) are then given by

$$\widetilde{H}^{-s}_{\mathrm{pw}}(\Gamma) = \prod_{i=1}^{J} \widetilde{H}^{-s}(\Gamma_i), \quad H^{-s}_{\mathrm{pw}}(\Gamma) = \prod_{i=1}^{J} H^{-s}(\Gamma_i),$$

with the norms

$$\|\widetilde{\psi}\|_{\widetilde{H}^{-s}_{\mathrm{pw}}(\Gamma)} = \left(\sum_{i=1}^{J} \|\widetilde{\psi}_{|\Gamma_i}\|^2_{\widetilde{H}^{-s}(\Gamma_i)} \right)^{1/2}, \quad \|\psi\|_{H^{-s}_{\mathrm{pw}}(\Gamma)} = \left(\sum_{i=1}^{J} \|\psi_{|\Gamma_i}\|^2_{H^{-s}(\Gamma_i)} \right)^{1/2},$$

for all $\widetilde{\psi} \in \widetilde{H}^{-s}_{\mathrm{pw}}(\Gamma)$, $\psi \in H^{-s}_{\mathrm{pw}}(\Gamma)$ and all $s \in [0, 3/2]$. The relation of the dual spaces is pointed out in the following lemma.

Lemma 6.1. *For all $s \in [0,3/2]$ we have $\widetilde{H}^{-s}_{\mathrm{pw}}(\Gamma) \subset H^{-s}(\Gamma)$, i.e. for all $\psi \in \widetilde{H}^{-s}_{\mathrm{pw}}(\Gamma)$ there holds the inequality*

$$\|\psi\|_{H^{-s}(\Gamma)} \leq \|\psi\|_{\widetilde{H}^{-s}_{\mathrm{pw}}(\Gamma)}. \tag{6.8}$$

Proof. Let $\psi \in \widetilde{H}^{-s}_{\mathrm{pw}}(\Gamma)$, then there holds

$$\|\psi\|_{H^{-s}(\Gamma)} = \sup_{0\neq v\in H^s(\Gamma)} \frac{\langle \psi, v\rangle_\Gamma}{\|v\|_{H^s(\Gamma)}} \leq \sup_{0\neq v\in H^s_{\mathrm{pw}}(\Gamma)} \frac{\sum_{i=1}^{J} \langle \psi_{|\Gamma_i}, v_{|\Gamma_i}\rangle_{\Gamma_i}}{\|v\|_{H^s_{\mathrm{pw}}(\Gamma)}} \leq \|\psi\|_{\widetilde{H}^{-s}_{\mathrm{pw}}(\Gamma)}.$$

□

For a further discussion of the above Sobolev spaces we refer to [30,53,73], and the references therein.

For the trace theorems in $H^2(\Omega)$ the following definitions are needed. We denote by

$$I_{i,j} := \overline{\Gamma}_i \cap \overline{\Gamma}_j \quad \text{for all } i \neq j,\ i,j \in \{1,\dots,J\},$$

the interface between two smooth boundary parts, and, by

$$v_i = v_{|\Gamma_i},$$

the restriction of $v \in H^s(\Gamma)$, $s \in [0,3/2]$ to Γ_i for all $i = 1,\dots,J$. Note that there holds $I_{i,j} = I_{j,i}$ for all $i \neq j$, $i,j = 1,\dots,J$. Then we define the compatibility condition

$$v_{i|I_{i,j}} = v_{j|I_{i,j}} \quad \text{for all } i \neq j,\ i,j \in \{1,\dots,J\}. \tag{6.9}$$

This means that we have continuity of the function across the interfaces.

The following statements are the Dirichlet and Neumann trace theorems and inverse trace theorems for the Sobolev space $H^2(\Omega)$ where Ω is a piecewise $C^{1,1}$ domain, i.e. the individual parts Γ_i are of class $C^{1,1}$ for all $i = 1,\dots,J$. The corresponding proofs can be found for instance in [33, Chapter 1.5.2], [30, Theorem I.1.6 and Remark 1.1, p. 9] and [38, Theorem 4.2.1, p. 178].

Theorem 6.1 (Dirichlet trace theorem)**.** *Let $\Omega \subset \mathbb{R}^n$, $n = 2,3$, be a bounded Lipschitz domain with piecewise smooth boundary Γ. For all $v \in H^2(\Omega)$ we have $v_{|\Gamma} \in H^{3/2}_{\mathrm{pw}}(\Gamma)$, satisfying the compatibility condition* (6.9)*, and*

$$\|v_{|\Gamma}\|_{H^{3/2}_{\mathrm{pw}}(\Gamma)} \leq c_{T,D}\|v\|_{H^2(\Omega)}.$$

Vice versa, for each $\lambda \in H^{3/2}_{\mathrm{pw}}(\Gamma)$ which is satisfying the compatibility condition (6.9) *there exists a $v \in H^2(\Omega)$ with $v_{|\Gamma} = \lambda$ on Γ and*

$$\|v\|_{H^2(\Omega)} \leq c_{IT,D}\|\lambda\|_{H^{3/2}_{\mathrm{pw}}(\Gamma)}.$$

Theorem 6.2 (Neumann trace theorem)**.** *Let $\Omega \subset \mathbb{R}^n$, $n = 2,3$, be a bounded Lipschitz domain with piecewise smooth boundary Γ. For all $v \in H^2(\Omega)$ we have $\partial_n v \in H^{1/2}_{\mathrm{pw}}(\Gamma)$ and*

$$\|\partial_n v\|_{H^{1/2}_{\mathrm{pw}}(\Gamma)} \leq c_{T,N} \|v\|_{H^2(\Omega)}.$$

Vice versa, for each $\lambda \in H^{1/2}_{\mathrm{pw}}(\Gamma)$ there exists a $v \in H^2(\Omega)$ with $\partial_n v = \lambda$ on Γ and

$$\|v\|_{H^2(\Omega)} \leq c_{IT,N} \|\lambda\|_{H^{1/2}_{\mathrm{pw}}(\Gamma)}.$$

An extended version, see [30, Theorem I.1.6, p. 9 and p. 184], of the second statement of Theorem 6.2 is given as follows, where we enforce zero Dirichlet boundary conditions.

Lemma 6.2. *Let $\Omega \subset \mathbb{R}^n$, $n = 2,3$, be a bounded Lipschitz domain with piecewise smooth boundary Γ. For each $\lambda \in H^{1/2}_{\mathrm{pw}}(\Gamma)$ there exists a $v \in H^2(\Omega) \cap H^1_0(\Omega)$ with $\partial_n v = \lambda$ on Γ and*

$$\|v\|_{H^2(\Omega)} \leq \widetilde{c}_{IT,N} \|\lambda\|_{H^{1/2}_{\mathrm{pw}}(\Gamma)}.$$

Proof. Let $\lambda \in H^{1/2}_{\mathrm{pw}}(\Gamma)$ be given. From Theorem 6.2 follows that there is $\widetilde{v} \in H^2(\Omega)$ with $\partial_n \widetilde{v} = \lambda$ on Γ and

$$\|\widetilde{v}\|_{H^2(\Omega)} \leq c_{IT,N} \|\lambda\|_{H^{1/2}_{\mathrm{pw}}(\Gamma)}.$$

Now, from Theorem 6.1 follows that $\widetilde{v}_{|\Gamma} \in H^{3/2}_{\mathrm{pw}}(\Gamma)$ and satisfies the compatibility condition (6.9). Let us consider now the problem

$$\begin{aligned} \Delta^2 w &= 0 && \text{in } \Omega, \\ w &= \widetilde{v}_{|\Gamma} && \text{on } \Gamma, \\ \partial_n w &= 0 && \text{on } \Gamma. \end{aligned}$$

Multiplication with a test function $\varphi \in H^2_0(\Omega)$ and applying integration by parts twice lead to the following variational formulation. Find $w \in H^2(\Omega)$ with $w = \widetilde{v}_{|\Gamma}$ and $\partial_n w = 0$ on Γ such that

$$\langle \Delta w, \Delta \varphi \rangle_{L^2(\Omega)} = 0,$$

for all test functions $\varphi \in H^2_0(\Omega)$. We decompose $w = w_0 + \mathcal{E}\widetilde{v}_{|\Gamma}$ with $w_0 \in H^2_0(\Omega)$ and a continuous extension $\mathcal{E} : H^{3/2}_{\mathrm{pw}}(\Gamma) \to H^2(\Omega)$ due to Theorem 6.1, c.f. also [30, p. 16–17]. Thus we obtain

$$\|\Delta w_0\|^2_{L^2(\Omega)} = \langle \Delta w_0, \Delta w_0 \rangle_{L^2(\Omega)} = -\langle \Delta \mathcal{E}\widetilde{v}_{|\Gamma}, \Delta w_0 \rangle_{L^2(\Omega)} \leq \|\Delta \mathcal{E}\widetilde{v}_{|\Gamma}\|_{L^2(\Omega)} \|\Delta w_0\|_{L^2(\Omega)},$$

and consequently, due to the norm equivalence in $H_0^2(\Omega)$,

$$\|w_0\|_{H^2(\Omega)} \simeq \|\Delta w_0\|_{L^2(\Omega)} \leq \|\Delta \mathcal{E}\widetilde{v}_{|\Gamma}\|_{L^2(\Omega)} \leq \|\mathcal{E}\widetilde{v}_{|\Gamma}\|_{H^2(\Omega)}.$$

This means we get by Theorem 6.1 the estimate

$$\|w\|_{H^2(\Omega)} \leq c\|\mathcal{E}\widetilde{v}_{|\Gamma}\|_{H^2(\Omega)} \leq cc_{IT,D}\|\widetilde{v}_{|\Gamma}\|_{H^{3/2}_{\mathrm{pw}}(\Gamma)} \leq cc_{IT,D}c_{T,D}\|\widetilde{v}\|_{H^2(\Omega)} = c_2\|\widetilde{v}\|_{H^2(\Omega)}.$$

Choosing now $v = \widetilde{v} - w \in H^2(\Omega)$ satisfies $\partial_n v = \lambda$ and $v_{|\Gamma} = 0$ on Γ. Consequently we obtain

$$\|v\|_{H^2(\Omega)} \leq \|\widetilde{v}\|_{H^2(\Omega)} + \|w\|_{H^2(\Omega)} \leq (1+c_2)\|\widetilde{v}\|_{H^2(\Omega)} \leq (1+c_2)c_{IT,N}\|\lambda\|_{H^{1/2}_{\mathrm{pw}}(\Gamma)},$$

which concludes the proof. □

6.2 Spectral equivalence estimates

Let us consider for some given $z \in H^{1/2}(\Gamma)$ the following homogeneous Dirichlet problem

$$\begin{aligned} -\Delta u_z &= 0 \qquad \text{in } \Omega, \\ u_z &= z \qquad \text{on } \Gamma. \end{aligned} \tag{6.10}$$

Then the corresponding unique solution $u_z \in H^1(\Omega)$ is called harmonic extension. The results of the previous subsection are required to prove the following theorem.

Theorem 6.3. *Let $\Omega \subset \mathbb{R}^n$, $n = 2,3$, be a convex and bounded Lipschitz domain with piecewise smooth boundary Γ. For $z \in H^{1/2}(\Gamma)$ let $u_z \in H^1(\Omega)$ be the harmonic extension of problem* (6.10)*. Then there hold the spectral equivalence inequalities*

$$\|z\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} \simeq \|u_z\|_{L^2(\Omega)},$$

for all $z \in H^{1/2}(\Gamma)$.

Proof. We will first prove the upper estimate $\|u_z\|_{L^2(\Omega)} \leq c_2\|z\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)}$. For a test function $v \in H^2(\Omega) \cap H_0^1(\Omega)$ we have, since $u_z \in H^1(\Omega)$ is the harmonic extension of $z \in H^{1/2}(\Gamma)$, by applying duality and the Hölder inequality,

$$\begin{aligned} \langle -\Delta v, u_z\rangle_{L^2(\Omega)} &= \langle \nabla v, \nabla u_z\rangle_{L^2(\Omega)} - \langle \partial_n v, u_z\rangle_\Gamma \\ &= -\langle \partial_n v, z\rangle_\Gamma = -\sum_{i=1}^J \langle \partial_n v_{|\Gamma_i}, z_{|\Gamma_i}\rangle_{\Gamma_i} \\ &\leq \sum_{i=1}^J \|\partial_n v_{|\Gamma_i}\|_{H^{1/2}(\Gamma_i)}\|z_{|\Gamma_i}\|_{\widetilde{H}^{-1/2}(\Gamma_i)} \leq \|\partial_n v\|_{H^{1/2}_{\mathrm{pw}}(\Gamma)}\|z\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)}. \end{aligned}$$

Let $w_z \in H_0^1(\Omega)$ be the unique solution of the Dirichlet problem

$$\begin{aligned} -\Delta w_z &= u_z \qquad \text{in } \Omega, \\ w_z &= 0 \qquad \text{on } \Gamma. \end{aligned}$$

Since Ω is assumed to be convex, we have $w_z \in H^2(\Omega) \cap H_0^1(\Omega)$, which satisfies

$$\|w_z\|_{H^2(\Omega)} \leq c \|u_z\|_{L^2(\Omega)}.$$

Thus, with Lemma 6.2 we conclude

$$\begin{aligned} \|u_z\|_{L^2(\Omega)}^2 &= \langle -\Delta w_z, u_z \rangle_{L^2(\Omega)} = -\langle \partial_n w_z, z \rangle_\Gamma \\ &\leq \|\partial_n w_z\|_{H_{\mathrm{pw}}^{1/2}(\Gamma)} \|z\|_{\widetilde{H}_{\mathrm{pw}}^{-1/2}(\Gamma)} \leq c \|w_z\|_{H^2(\Omega)} \|z\|_{\widetilde{H}_{\mathrm{pw}}^{-1/2}(\Gamma)} \\ &\leq c_2 \|u_z\|_{L^2(\Omega)} \|z\|_{\widetilde{H}_{\mathrm{pw}}^{-1/2}(\Gamma)}. \end{aligned}$$

This proves the upper estimate and it remains to show the inequality

$$c_1 \|z\|_{\widetilde{H}_{\mathrm{pw}}^{-1/2}(\Gamma)} \leq \|u_z\|_{L^2(\Omega)},$$

for all $z \in H^{1/2}(\Gamma)$. From Lemma 6.2 we conclude that for any $\lambda \in H_{\mathrm{pw}}^{1/2}(\Gamma)$ there exists a $v \in H^2(\Omega) \cap H_0^1(\Omega)$ such that $\partial_n v = \lambda$ on Γ is satisfied and

$$\|v\|_{H^2(\Omega)} \leq \widetilde{c}_{\mathrm{IT,N}} \|\lambda\|_{H_{\mathrm{pw}}^{1/2}(\Gamma)}.$$

Hence we find

$$\begin{aligned} \|z\|_{\widetilde{H}_{\mathrm{pw}}^{-1/2}(\Gamma)} &= \sup_{0 \neq \lambda \in H_{\mathrm{pw}}^{1/2}(\Gamma)} \frac{\langle z, \lambda \rangle_\Gamma}{\|\lambda\|_{H_{\mathrm{pw}}^{1/2}(\Gamma)}} = \sup_{0 \neq \lambda \in H_{\mathrm{pw}}^{1/2}(\Gamma)} \frac{\langle \Delta v, u_z \rangle_{L^2(\Omega)}}{\|\lambda\|_{H_{\mathrm{pw}}^{1/2}(\Gamma)}} \\ &\leq \sup_{0 \neq \lambda \in H_{\mathrm{pw}}^{1/2}(\Gamma)} \frac{\|\Delta v\|_{L^2(\Omega)} \|u_z\|_{L^2(\Omega)}}{\|\lambda\|_{H_{\mathrm{pw}}^{1/2}(\Gamma)}} \leq \widetilde{c}_{\mathrm{IT,N}} \|u_z\|_{L^2(\Omega)}, \end{aligned}$$

which completes the proof. □

Now we are in a position to state the required spectral equivalence inequalities for the Schur complement T_h, defined in (6.4). Let us consider piecewise linear and globally continuous shape functions on the boundary by restriction to the boundary, i.e. $\phi_k := \varphi_{n_I+k|\Gamma}$ for all $k = 1, \ldots, n_C$. Then we can define the finite element trace space

$$Z_h := \operatorname{span}\{\phi_k\}_{k=1}^{n_C} = \mathcal{V}_{h|\Gamma} = \operatorname{span}\{\varphi_{n_I+k|\Gamma}\}_{k=1}^{n_C} \subset H^{1/2}(\Gamma). \tag{6.11}$$

Theorem 6.4. *For all $\underline{z}_C \in \mathbb{R}^{n_C} \leftrightarrow z_h \in Z_h$ there hold the spectral equivalence inequalities*

$$(T_h \underline{z}_C, \underline{z}_C) \simeq \|z_h\|_{\widetilde{H}_{\mathrm{pw}}^{-1/2}(\Gamma)}^2.$$

Proof. For $z_h \in Z_h \leftrightarrow \underline{z}_C \in \mathbb{R}^{n_C}$ let $u_{z_h} \in H^1(\Omega)$ be the harmonic extension for which we have, by Theorem 6.3,

$$c_1 \|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} \leq \|u_{z_h}\|_{L^2(\Omega)} \leq c_2 \|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)}. \tag{6.12}$$

On the other hand, by defining $\underline{u}_I = -A_{II}^{-1} A_{IC} \underline{z}_C$ and by setting $\underline{u} = (\underline{u}_I, \underline{z}_C)^\top \leftrightarrow u_{z_h,h} \in \mathcal{V}_h$, which is the discrete harmonic extension of z_h, we obtain by using (6.5),

$$(T_h \underline{z}_C, \underline{z}_C) = \|u_{z_h,h}\|^2_{L^2(\Omega)}.$$

Since $u_{z_h,h} \in \mathcal{V}_h$ is the standard finite element approximation of $u_{z_h} \in H^1(\Omega)$, we have, by applying the spectral equivalence (6.12), the standard finite element error estimate in $L^2(\Omega)$, the continuity of the Dirichlet trace of the harmonic extension $u_{z_h} \in H^1(\Omega)$, and an inverse inequality, the estimate

$$\begin{aligned}
\|u_{z_h,h}\|_{L^2(\Omega)} &\leq \|u_{z_h}\|_{L^2(\Omega)} + \|u_{z_h,h} - u_{z_h}\|_{L^2(\Omega)} \\
&\leq c_2 \|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} + c_3 h \|u_{z_h}\|_{H^1(\Omega)} \\
&\leq c_2 \|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} + c_4 h \|z_h\|_{H^{1/2}(\Gamma)} \\
&\leq c_2 \|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} + c_5 \|z_h\|_{H^{-1/2}(\Gamma)}.
\end{aligned}$$

Now the upper estimate follows by using (6.8). To prove the reverse estimate we first have, by using an inverse inequality, and the bound of the Dirichlet trace of the discrete harmonic extension $u_{z_h,h} \in \mathcal{V}_h \subset H^1(\Omega)$,

$$\|u_{z_h,h}\|_{L^2(\Omega)} \geq c_6 h \|u_{z_h,h}\|_{H^1(\Omega)} \geq c_7 h \|z_h\|_{H^{1/2}(\Gamma)} = c_7 \underbrace{\frac{h\|z_h\|_{H^{1/2}(\Gamma)}}{\|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)}}}_{=:\alpha} \|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)}.$$

On the other hand we conclude, as above,

$$\begin{aligned}
\|u_{z_h,h}\|_{L^2(\Omega)} &\geq \|u_{z_h}\|_{L^2(\Omega)} - \|u_{z_h,h} - u_{z_h}\|_{L^2(\Omega)} \\
&\geq c_1 \|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} - c_4 h \|z_h\|_{H^{1/2}(\Gamma)} = (c_1 - c_4 \alpha) \|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)}.
\end{aligned}$$

In particular we have

$$\|u_{z_h,h}\|_{L^2(\Omega)} \geq \max\{c_7\alpha, c_1 - \alpha c_4\} \|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)},$$

and by using

$$\min_{\alpha>0} \max\{c_7\alpha, c_1 - \alpha c_4\} = \frac{c_1 c_7}{c_7 + c_4} > 0,$$

the proof is concluded. □

6.3 Schur complement preconditioners

In this section we discuss two different approaches for the construction of a preconditioner for the Schur complement matrix T_h as defined in (6.4). By this we mean a discrete operator, which reflects the spectral equivalence inequalities in $\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)$, see Theorem 6.4. The first approach is based on the local single layer boundary integral operator. Further, we also present a multilevel approach of BPX type which sometimes can be more useful. Therefore we need additional spectral equivalence inequalities which relates the Schur complement T_h to a Sobolev norm in $H^{-1/2}(\Gamma)$, which is, as we will see, in general only optimal up to a logarithmic factor. Corresponding numerical examples illustrate the obtained theoretical results.

6.3.1 The SLP preconditioner

For the Schur complement matrix T_h, defined in (6.4) we know form Theorem 6.4 that it satisfies the spectral equivalence inequalities

$$(T_h \underline{z}_C, \underline{z}_C) \simeq \|z_h\|^2_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} = \sum_{i=1}^{J} \|z_{h|\Gamma_i}\|^2_{\widetilde{H}^{-1/2}(\Gamma_i)}, \tag{6.13}$$

for all $\underline{z}_C \in \mathbb{R}^{n_C} \leftrightarrow z_h \in Z_h$, where the finite element trace space Z_h is defined in (6.11). For the construction of a preconditioning matrix C_{T_h} it is therefore sufficient to find a computable representation of the local Sobolev norms $\|\cdot\|^2_{\widetilde{H}^{-1/2}(\Gamma_i)}$, for all $i = 1,\dots,J$, which can be done by using local boundary integral operators, see, e.g., [38, 73].

For a given $\psi_i \in \widetilde{H}^{-1/2}(\Gamma_i)$, $i = 1,\dots,J$, we define the local single layer boundary integral operator as

$$(V_i \psi_i)(x) = \int_{\Gamma_i} U^*(x,y)\,\psi_i(y)\,ds_y, \tag{6.14}$$

for $x \in \Gamma_i$, where

$$U^*(x,y) = \begin{cases} -\dfrac{1}{2\pi}\log|x-y| & \text{for } n = 2, \\ \dfrac{1}{4\pi|x-y|} & \text{for } n = 3, \end{cases}$$

is the fundamental solution of the Laplace operator. It turns out that the local single layer boundary integral operator $V_i : \widetilde{H}^{-1/2}(\Gamma_i) \to H^{1/2}(\Gamma_i)$ is bounded and elliptic in $\widetilde{H}^{-1/2}(\Gamma_i)$, see, e.g., [54, Theorem 2.4], and hence the duality product

$$\|\psi_i\|^2_{V_i} := \langle V_i \psi_i, \psi_i \rangle_{\Gamma_i} \simeq \|\psi_i\|^2_{\widetilde{H}^{-1/2}(\Gamma_i)}, \tag{6.15}$$

defines an equivalent norm in $\widetilde{H}^{-1/2}(\Gamma_i)$, for all $i = 1,\ldots,J$. Note that for the two-dimensional case we assume that the length of all Γ_i are less than 4, see [54, Theorem 2.4]. By combining the spectral equivalence inequalities (6.13) and (6.15) we therefore conclude for the Schur complement the spectral equivalence inequalities

$$(T_h \underline{z}_C, \underline{z}_C) \simeq \sum_{i=1}^{J} \langle V_i z_{h|\Gamma_i}, z_{h|\Gamma_i} \rangle_{\Gamma_i} = \sum_{i=1}^{J} (A_i^\top V_{i,h} A_i \underline{z}_C, \underline{z}_C), \tag{6.16}$$

for all $\underline{z}_C \in \mathbb{R}^{n_C} \leftrightarrow z_h \in Z_h$. In (6.16), $V_{i,h} \in \mathbb{R}^{n_{C,i} \times n_{C,i}}$ with $n_{C,i} = \dim Z_{h|\Gamma_i}$ is the Galerkin boundary element matrix of the local single layer boundary integral operator given by

$$V_{i,h}[\ell,k] = \langle V_i \phi_{k,i}, \phi_{\ell,i} \rangle_{\Gamma_i},$$

for all $k,\ell = 1,\ldots,n_{C,i}$, and

$$Z_{h|\Gamma_i} = \operatorname{span}\{\phi_{k,i}\}_{k=1}^{n_{C,i}},$$

is the localized finite element space, for all $i = 1,\ldots,J$. The relation between the global and local degrees of freedom is described by connectivity matrices $A_i \in \mathbb{R}^{n_{C,i} \times n_C}$. The spectral equivalence inequalities (6.16) imply the definition of the preconditioning matrix

$$C_{\mathrm{SLP}} := \sum_{i=1}^{J} A_i^\top V_{i,h} A_i, \tag{6.17}$$

where the spectral condition number of the preconditioned system,

$$\kappa(C_{\mathrm{SLP}}^{-1} T_h) \leq c, \tag{6.18}$$

is bounded by a constant which is independent of the discretization.

The application of the preconditioning matrix C_{SLP}^{-1} requires the solution of a linear system, $\underline{v} = C_{\mathrm{SLP}}^{-1} \underline{r}$. Since the preconditioner (6.17) corresponds to an additive Schwarz method, see e.g. [11], for the discrete single layer boundary integral operator [36, 57], it can be realized by solving local subproblems which correspond to all the interior degrees of freedom within Γ_i, and by solving a coarse Schur complement system which corresponds to all degrees of freedom along the interfaces. In the particular case of a two-dimensional polygonal bounded domain the dimension of the coarse system coincides with the number of corner points, which is in general rather small. The situation can be quite different when considering more general three-dimensional polyhedral domains. This motivates the use of global preconditioning strategies such as a multilevel approach wich implies a spectral equivalent preconditioner in $H^{-1/2}(\Gamma)$.

6.3.2 The BPX preconditioner

For the definition of a global preconditioning matrix C_{T_h} in $H^{-1/2}(\Gamma)$ we need to have, in addition to (6.13), the spectral equivalence inequalities

$$\|z_h\|^2_{\widetilde{H}^{-1/2}_{\rm pw}(\Gamma)} \simeq \|z_h\|^2_{H^{-1/2}(\Gamma)},$$

for all $\underline{z}_C \in \mathbb{R}^{n_C} \leftrightarrow z_h \in Z_h$. As we will see this estimate is depending on a logarithmic term of the mesh size h. First, we recall from (6.8) the estimate

$$\|z_h\|_{H^{-1/2}(\Gamma)} \le \|z_h\|_{\widetilde{H}^{-1/2}_{\rm pw}(\Gamma)}, \tag{6.19}$$

for all $\underline{z}_C \in \mathbb{R}^{n_C} \leftrightarrow z_h \in Z_h$, where Z_h is the finite element trace space defined in (6.11). The proof of the reverse inequality is more involved.

Theorem 6.5. *Let the boundary Γ be piecewise smooth and let $z_h \in Z_h$ where the mesh size h is assumed to be sufficiently small. Then there holds the estimate*

$$\|z_h\|_{\widetilde{H}^{-1/2}_{\rm pw}(\Gamma)} \le \overline{c}_2 J[1-\log h]\|z_h\|_{H^{-1/2}(\Gamma)}. \tag{6.20}$$

Proof. The proof of (6.20) follows the ideas as used for the analysis of the additive Schwarz method for the single layer boundary integral operator, see, e.g., [36,57]. For $s \in (0,\frac{1}{2})$ and $i = 1,\dots,J$ we first have, see, e.g., [54, Lemma 2.3],

$$\|\phi\|_{\widetilde{H}^s(\Gamma_i)} \le \frac{c}{1/2-s}\|\phi\|_{H^s(\Gamma_i)},$$

for all $\phi \in H^s(\Gamma_i)$. By using a duality argument and the inverse inequality we therefore conclude, for $\varepsilon \in (0,\frac{1}{2})$,

$$\begin{aligned}\|z_h\|_{\widetilde{H}^{-1/2}(\Gamma_i)} \le \|z_h\|_{\widetilde{H}^{-1/2+\varepsilon}(\Gamma_i)} &= \sup_{0\neq\phi\in H^{1/2-\varepsilon}(\Gamma_i)} \frac{\langle z_h,\phi\rangle_{\Gamma_i}}{\|\phi\|_{H^{1/2-\varepsilon}(\Gamma_i)}} \\ &\le \frac{c}{\varepsilon}\sup_{0\neq\phi\in H^{1/2-\varepsilon}(\Gamma_i)} \frac{\langle z_h,\phi\rangle_{\Gamma_i}}{\|\phi\|_{\widetilde{H}^{1/2-\varepsilon}(\Gamma_i)}} \le \frac{c}{\varepsilon}\|z_h\|_{H^{-1/2+\varepsilon}(\Gamma_i)} \le \frac{\widetilde{c}}{\varepsilon}h^{-\varepsilon}\|z_h\|_{H^{-1/2}(\Gamma_i)}.\end{aligned}$$

Finally, by choosing $\varepsilon = -1/\log h \in (0,\frac{1}{2})$, which is satisfied for sufficient small h, we obtain

$$\|z_h\|_{\widetilde{H}^{-1/2}(\Gamma_i)} \le c[1-\log h]\|z_h\|_{H^{-1/2}(\Gamma_i)}.$$

Now the assertion follows by summing up, and by using again a duality argument,

$$\begin{aligned}
\|z_h\|^2_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} &= \sum_{i=1}^{J} \|z_h\|^2_{\widetilde{H}^{-1/2}(\Gamma_i)} \leq c^2[1-\log h]^2 \sum_{i=1}^{J} \|z_h\|^2_{H^{-1/2}(\Gamma_i)} \\
&\leq c^2[1-\log h]^2 \left(\sum_{i=1}^{J} \|z_h\|_{H^{-1/2}(\Gamma_i)} \right)^2 \\
&= c^2[1-\log h]^2 \left(\sum_{i=1}^{J} \sup_{0\neq\phi_i\in\widetilde{H}^{1/2}(\Gamma_i)} \frac{\langle z_h,\phi_i\rangle_{\Gamma_i}}{\|\phi_i\|_{\widetilde{H}^{1/2}(\Gamma_i)}} \right)^2 \\
&= c^2[1-\log h]^2 \Big(\sup_{\phi=\sum_{i=1}^{J} \frac{\phi_i}{\|\phi_i\|_{\widetilde{H}^{1/2}(\Gamma_i)}},\ 0\neq\phi_i\in\widetilde{H}^{1/2}(\Gamma_i)} \langle z_h,\phi\rangle_\Gamma \Big)^2 \\
&\leq c^2[1-\log h]^2 \Big(\sup_{\phi\in H^{1/2}(\Gamma),\ \|\phi\|_{H^{1/2}(\Gamma)}\leq J} \langle z_h,\phi\rangle_\Gamma \Big)^2 \\
&\leq c^2 J^2[1-\log h]^2 \|z_h\|^2_{H^{-1/2}(\Gamma)},
\end{aligned}$$

which completes the proof. □

By combining the spectral equivalence inequalities (6.13) with (6.19) and (6.20) we conclude the spectral equivalence inequalities

$$\widetilde{c}_1\|z_h\|^2_{H^{-1/2}(\Gamma)} \leq (T_h \underline{z}_C, \underline{z}_C) \leq \widetilde{c}_2 J^2[1-\log h]^2 \|z_h\|^2_{H^{-1/2}(\Gamma)}, \tag{6.21}$$

for all $\underline{z}_C \in \mathbb{R}^{n_C} \leftrightarrow z_h \in Z_h$. It remains to find a preconditioner which is spectrally equivalent to the discrete norm in $H^{-1/2}(\Gamma)$. One possibility is the use again a boundary integral operator, namely the stabilized discrete hypersingular boundary integral operator as a preconditioner of opposite order [75]. However, in the following we consider a geometric multilevel operator, see [10,60], for piecewise linear and globally continuous basis functions on the boundary to represent the norm in $H^{-1/2}(\Gamma)$. Other choices involve algebraic or artificial multilevel operators as considered in, e.g., [26,58,72].

For the construction of the multilevel preconditioner we consider a sequence of admissible globally quasi–uniform nested finite element meshes $\{\mathcal{T}_{h_i}\}_{i\in\mathbb{N}_0}$ of mesh size $h_i \simeq 2^{-i}$. Let $\{\mathcal{V}_{h_i}\}_{i\in\mathbb{N}_0} \subset H^1(\Omega)$ denote the related sequence of finite element spaces with piecewise linear and globally continuous basis functions. Then we consider the restrictions on the boundary,

$$Z_i := \mathrm{span}\{\phi^i_k\}_{k=1}^{n^i_C} = \mathcal{V}_{h_i|\Gamma} = \mathrm{span}\{\varphi^i_{n_{I,i}+k|\Gamma}\}_{k=1}^{n^i_C} \subset H^{1/2}(\Gamma),$$

with $i \in \mathbb{N}_0$. This results in a sequence of nested spaces of the form

$$Z_0 \subset Z_1 \subset \ldots \subset Z_L = Z_{h_L} \subset Z_{L+1} \subset \ldots \subset H^{1/2}(\Gamma),$$

where L denotes the current level of interest. With respect to the boundary element spaces Z_i, $i \in \mathbb{N}_0$, of piecewise linear globally continuous shape functions ϕ_k^i we introduce, for a given $z \in L^2(\Gamma)$, the L^2–projection $Q_i : L^2(\Gamma) \to Z_i$ as the unique solution $Q_i z \in Z_i$ of the variational problem

$$\langle Q_i z, v_{h_i} \rangle_{L^2(\Gamma)} = \langle z, v_{h_i} \rangle_{L^2(\Gamma)},$$

for all $v_{h_i} \in Z_i$. In addition we set $Q_{-1} := 0$. It turns out, see, e.g., [10, 60, 73], that the multilevel operator

$$B_{-1/2} := \sum_{i=0}^{\infty} h_i (Q_i - Q_{i-1}) : H^{-1/2}(\Gamma) \to H^{1/2}(\Gamma), \tag{6.22}$$

defines an equivalent norm in $H^{-1/2}(\Gamma)$, and that its inverse operator is given by

$$B_{-1/2}^{-1} = B_{1/2} = \sum_{i=0}^{\infty} h_i^{-1} (Q_i - Q_{i-1}) : H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma).$$

As in [73, Corollary 13.7] we finally conclude that the preconditioner C_{T_h} of the Schur complement T_h is given by

$$C_{\mathrm{BPX}} := M_{h_L} B_{h_L}^{-1} M_{h_L}, \tag{6.23}$$

where

$$M_{h_L}[\ell, k] = \langle \phi_k^L, \phi_\ell^L \rangle_{L^2(\Gamma)}, \quad B_{h_L}[\ell, k] = \langle B_{1/2} \phi_k^L, \phi_\ell^L \rangle_\Gamma,$$

for all $k, \ell = 1, \ldots, n_C^L$ denote the standard mass matrix and the Galerkin matrix of the multilevel operator $B_{1/2}$. Moreover, by using the spectral equivalence inequalities (6.21) we conclude the following bound for the spectral condition number of the preconditioned system,

$$\kappa(C_{\mathrm{BPX}}^{-1} T_h) \leq c J^2 [1 - \log h]^2. \tag{6.24}$$

Note that J depends on the geometry of Ω, but not on the discretization.

For the application $\underline{v} = C_{\mathrm{BPX}}^{-1} \underline{r}$, e.g., within a conjugate gradient scheme, we obtain, as for the standard BPX multilevel approach [10, 73], the representation

$$\underline{v} = \sum_{i=0}^{L} \alpha_i R_i M_{h_i}^{-1} R_i^\top \underline{r},$$

with coefficients

$$\alpha_i = \begin{cases} \dfrac{1}{h_L} & \text{for } i = L, \\ \dfrac{1}{h_i} - \dfrac{1}{h_{i+1}} & \text{for } i = 0, \ldots, L-1, \end{cases}$$

where $R_i : \mathbb{R}^{n_{C,i}} \to \mathbb{R}^{n_{C,L}}$ is the prolongation operator which is related to the nested sequence of piecewise linear finite element spaces on the boundary. While for the application of multilevel preconditioners for (pseudo)differential operators of positive order one can replace the inverse mass matrices $M_{h_i}^{-1}$ by its diagonals, this is not possible in the case of the Schur complement T_h which is the Galerkin discretization of a (pseudo)differential operator of order -1, in particular we have

$$\alpha_i < 0,$$

for all $i = 0, \ldots, L-1$. As a consequence we need to use the inverse mass matrices $M_{h_i}^{-1}$, or its approximation, which can be realized at low cost by a few conjugate gradient iterations.

6.3.3 Numerical results

For the numerical experiments we consider the biharmonic Dirichlet boundary value problem (5.1) in the domains $\Omega = B_{1/2}(0)$ and $\Omega = (0, \frac{1}{2})^n$, both for $n = 2, 3$. The linear system (6.3) is solved by a conjugate gradient scheme without (CG) and with (PCG) preconditioning up to a relative error reduction of $\varepsilon = 10^{-8}$. In all following tables we present, for a sequence of different levels L, the number of required PCG iterations, and the related numbers $n_{I,L}$ and $n_{C,L}$ of degrees of freedom in the interior and on the boundary, respectively.

Example 1

In the first numerical example we choose the right-hand side f as in the previous section of this chapter, namely (5.9).

For the two-dimensional test problems we first consider the discrete single layer boundary integral operator preconditioner (6.17), see Table 6.1. As expected from the estimate (6.18) we observe a constant number of PCG iterations for both computational domains. Next we consider the multilevel preconditioner (6.23), the results are given in Table 6.1 too. In the case of the circular domain $\Omega = B_{1/2}(0)$ with a smooth boundary $\Gamma = \partial\Omega$ we observe a constant number of PCG iterations since the Sobolev spaces $H^{-1/2}(\Gamma) = \widetilde{H}_{\mathrm{pw}}^{-1/2}(\Gamma)$ coincide. In contrast to the circular domain, for the polygonal bounded domain $\Omega = (0, \frac{1}{2})^2$ we observe a slightly increasing number of PCG iterations, which corresponds to the logarithmic behavior of the spectral condition number bound (6.24).

Example 2

In this example the right–hand side f is chosen by an arbitrary vector with values in $[-1, 1]$, generated by `rand()`.

			$\Omega = B_{1/2}(0)$		$\Omega = (0,\frac{1}{2})^2$	
L	$n_{I,L}$	$n_{C,L}$	C_{SLP}	C_{BPX}	C_{SLP}	C_{BPX}
0	1	4	1	1	1	1
1	5	8	5	5	5	5
2	25	16	10	10	7	6
3	113	32	13	12	5	5
4	481	64	12	11	9	9
5	1 985	128	13	13	11	12
6	8 065	256	14	14	11	14
7	32 513	512	13	15	11	14
8	130 561	1 024	13	15	11	15
9	523 265	2 048	12	15	11	15
10	2 095 105	4 096	12	15	11	16

Table 6.1: PCG iterations for C_{SLP} and C_{BPX} preconditioner, $n = 2$.

			$\Omega = B_{1/2}(0)$		$\Omega = (0,\frac{1}{2})^2$	
L	$n_{I,L}$	$n_{C,L}$	CG iter	PCG iter	CG iter	PCG iter
0	1	4	1	1	1	1
1	5	8	3	5	3	5
2	25	16	9	11	9	11
3	113	32	19	15	22	16
4	481	64	25	16	30	18
5	1 985	128	33	17	38	19
6	8 065	256	43	17	49	20
7	32 513	512	53	16	63	21
8	130 561	1 024	70	16	80	22
9	523 265	2 048	88	16	101	23
10	2 095 105	4 096	114	16	128	24

Table 6.2: Iterations for the multilevel preconditioner C_{BPX}, $n = 2$.

In Table 6.2 we present iteration numbers for the multilevel preconditioner (6.23) for the two-dimensional test problems. As in Example 1, we observe a constant number of PCG iterations for $\Omega = B_{1/2}(0)$, while for the square $\Omega = (0,\frac{1}{2})^2$ we obtain a logarithmic factor, corresponding to the spectral condition number bound (6.24). This logarithmic behavior is more obvious when considering the three-dimensional test problem with $\Omega = (0,\frac{1}{2})^3$, see Table 6.3.

	$\Omega = B_{1/2}(0)$				$\Omega = (0,\frac{1}{2})^2$			
L	$n_{I,L}$	$n_{C,L}$	CG iter	PCG iter	$n_{I,L}$	$n_{C,L}$	CG iter	PCG iter
0	1	18	12	12	1	8	2	2
1	19	66	23	16	9	26	18	15
2	231	258	31	25	91	98	29	24
3	2 255	1 026	43	29	855	386	41	27
4	19 871	4 098	54	30	7 471	1 538	55	30
5	166 719	16 386	71	31	62 559	6 146	72	34
6	1 365 631	65 538	91	31	512 191	24 578	94	37

Table 6.3: Iterations for the multilevel preconditioner C_{BPX}, $n = 3$.

Example 3 (non-convex domains)

So far we presented numerical results for convex domains. Even tough we assumed for the spectral equivalence estimates in this section the domain to be convex, we would like to demonstrate the effectiveness of the preconditioner for non-convex domains too.

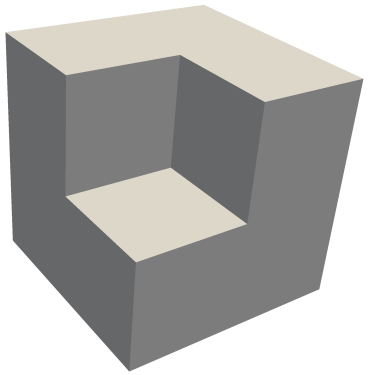


Figure 6.1: Fichera cube, $n = 3$.

In particular we consider the following two cases, the two-dimensional L-shaped domain $\Omega = (-\frac{1}{2},\frac{1}{2})^2 \backslash (-\frac{1}{2},0)^2$ with $J = 6$ and the Fichera cube $\Omega = (0,1)^3 \backslash (0,\frac{1}{2})^3$ with $J = 9$ smooth boundary parts. Note, the number of iterations will depend on the number of smooth parts J, see (6.24). In both cases the right-hand side f is an arbitrary vector with values in $[-1,1]$, as in the previous example.

In Table 6.4 we present corresponding iteration numbers. We observe a logarithmic behavior of the iteration numbers for the preconditioned system, influenced by the number of smooth parts J.

	L–shape, $n=2$				Fichera cube, $n=3$			
L	$n_{I,L}$	$n_{C,L}$	CG iter	PCG iter	$n_{I,L}$	$n_{C,L}$	CG iter	PCG iter
0	3	8	8	8	2	26	22	23
1	17	16	16	14	34	98	39	36
2	81	32	25	19	397	386	49	48
3	353	64	33	20	3 803	1 538	65	59
4	1 473	128	43	21	33 207	6 146	84	71
5	6 017	256	54	23	277 359	24 578	108	79
6	24 321	512	69	25				
7	97 793	1 024	91	26				
8	392 193	2 048	116	29				
9	1 570 817	4 096	154	30				
10	6 287 361	8 192	197	34				

Table 6.4: Iterations for the multilevel preconditioner C_{BPX} for non-convex domains.

6.4 Global preconditioners

In order to solve the global linear system (6.1) using an iterative method efficiently, a global preconditioner (6.2) is needed. In the previous section we have seen the construction of preconditioners C_{T_h} for the Schur complement equation. Within this section we propose, without any numerical analysis, two preconditioners C_{A_h} for the upper left block

$$\begin{pmatrix} M_{II} & A_{II} \\ A_{II} & \end{pmatrix}. \tag{6.25}$$

Note that this 2×2 block is symmetric and indefinite. Let us denote by $C_{\mathrm{MG}} \in \mathbb{R}^{n_I \times n_I}$ the preconditioning matrix, which is obtained when we apply a V–cycled geometrical multigrid $V(2,2)$, i.e. with 2 pre-smoothing and 2 post-smoothing steps, to the stiffness matrix $A_{II} \in \mathbb{R}^{n_I \times n_I}$. For a further discussion on multigrid methods we refer for instance to [34].

From the theory of saddle point problems, see, e.g., [12, 22, 73], it is easy to prove that the upper left 2×2 block (6.25) is spectrally equivalent to

$$\begin{pmatrix} A_{II} & \\ & A_{II} \end{pmatrix}.$$

Nevertheless, numerical results show that this preconditioner, where the individual blocks are realized by C_{MG} in combination with C_{BPX} for a preconditioned GMRES method, fails. The iteration numbers do not grow as fast as without preconditioning, but no robustness with respect to the discretization can be observed.

Alternatively, we propose for C_{A_h} the following two preconditioners, where A_{II} is already

realized by C_{MG},

$$C_{A_h,1} = \begin{pmatrix} & C_{\mathrm{MG}} \\ C_{\mathrm{MG}} & \end{pmatrix}, \quad C_{A_h,2} = \begin{pmatrix} M_{II} & C_{\mathrm{MG}} \\ C_{\mathrm{MG}} & \end{pmatrix}.$$

The corresponding inverse of the preconditioners are then given by

$$C_{A_h,1}^{-1} = \begin{pmatrix} & C_{\mathrm{MG}}^{-1} \\ C_{\mathrm{MG}}^{-1} & \end{pmatrix}, \quad C_{A_h,2}^{-1} = \begin{pmatrix} C_{\mathrm{MG}}^{-1} & \\ C_{\mathrm{MG}}^{-1} & -C_{\mathrm{MG}}^{-1} M_{II} C_{\mathrm{MG}}^{-1} \end{pmatrix}.$$

Note that the preconditioner $C_{A_h,2}$ is the upper left block (6.25) itself, where the stiffness matrices are replaced by its multigrid realization C_{MG}. Further, it is important to mention, that if C_{MG} is a good preconditioner for A_{II}, it is in general not true that $C_{\mathrm{MG}} M_{II}^{-1} C_{\mathrm{MG}}$ is a good one for $A_{II} M_{II}^{-1} A_{II}$. In the following numerical results we consider for the Schur complement preconditioner C_{T_h} always the BPX preconditioner C_{BPX} (6.23).

6.4.1 Numerical results

We present numerical results for the preconditioning of the linear system (6.1). The global preconditioner is given in (6.2), where C_{A_h} is either $C_{A_h,1}$ or $C_{A_h,2}$ and the Schur complement preconditioner C_{T_h} is in all considered cases C_{BPX} defined in (6.23). As computational domains we consider $\Omega = B_{1/2}(0)$ and $\Omega = (0, \frac{1}{2})^n$, both for $n = 2, 3$. For the iterative solution of the linear system (6.1) we use generalized minimal residual method without (I) and with ($C_{A_h,1}$ or $C_{A_h,2}$) preconditioning up to a relative error reduction of $\varepsilon = 10^{-8}$. In all following tables we present, for a sequence of different levels L, the number of required PGMRES iterations, and the related numbers DoFs $= 2n_{I,L} + n_{C,L}$ denoting the global number of degrees of freedom.

Example 1

In this example we consider the right-hand side f generated by the exact solutions (5.9), with the results presented in Table 6.5. As expected we get in the case of $\Omega = B_{1/2}(0)$ nearly constant iteration numbers and in the case of $\Omega = (0, \frac{1}{2})^2$ we obtain a logarithmic factor, as we have seen it for the Schur complement preconditioner in the previous section. Further, we observe that the iteration numbers of the preconditioner $C_{A_h,2}$ are smaller in comparison to $C_{A_h,1}$. This is due to the fact that we consider for $C_{A_h,2}$ the inverse of the upper left 2×2 block where we replace A_{II} by its approximation C_{MG}.

		$\Omega = B_{1/2}(0)$			$\Omega = (0,\frac{1}{2})^2$		
L	DoFs	I	$C_{A_h,1}$	$C_{A_h,2}$	I	$C_{A_h,1}$	$C_{A_h,2}$
0	6	3	3	3	3	3	3
1	18	10	6	5	6	6	4
2	66	37	9	7	27	9	7
3	258	157	13	9	163	16	12
4	1 026	>500	15	11	>500	23	18
5	4 098		21	17		28	21
6	16 386		25	18		38	31
7	65 538		27	19		51	36
8	262 146		28	20		59	40
9	1 048 578		30	20		64	43
10	4 194 306		31	21		69	46

Table 6.5: GMRES iterations with and without preconditioning, $n = 2$.

		$\Omega = B_{1/2}(0)$			$\Omega = (0,\frac{1}{2})^2$		
L	DoFs	I	$C_{A_h,1}$	$C_{A_h,2}$	I	$C_{A_h,1}$	$C_{A_h,2}$
0	6	3	3	3	3	3	3
1	18	15	13	11	15	13	9
2	66	59	21	16	59	27	18
3	258	236	29	22	234	36	26
4	1 026	>500	35	26	>500	47	33
5	4 098		40	28		53	37
6	16 386		43	29		59	40
7	65 538		47	31		64	43
8	262 146		49	33		70	46
9	1 048 578		51	33		76	50
10	4 194 306		53	35		81	54

Table 6.6: GMRES iterations with and without preconditioning, $n = 2$.

Example 2

In order to check the reliability for the preconditioner we choose again for the right-hand side f an arbitrary vector with values in $[-1,1]$, generated by `rand()`. The corresponding numerical results are presented in Table 6.6 for $n = 2$ and in Table 6.7 for $n = 3$. For the ball $\Omega = B_{1/2}(0)$ we observe again almost constant iteration numbers, while for the $\Omega = (0,\frac{1}{2})^n$, with $n = 2,3$ we observe the known logarithmic behavior in the iteration numbers. As in the previous examples the preconditioner $C_{A_h,2}$ performs better.

	$\Omega = B_{1/2}(0)$				$\Omega = (0,\frac{1}{2})^3$			
L	DoFs	I	$C_{A_h,1}$	$C_{A_h,2}$	DoFs	I	$C_{A_h,1}$	$C_{A_h,2}$
0	20	11	11	11	10	4	4	4
1	104	53	24	20	44	30	24	19
2	720	451	35	30	280	196	46	36
3	5 536	>500	41	34	2 096	>500	58	42
4	43 840		45	37	16 480		66	49
5	349 824		48	39	131 264		76	54
6	2 796 800		52	40	1 048 960		87	60

Table 6.7: GMRES iterations with and without preconditioning, $n = 3$.

Example 3 (non-convex domains)

As in the numerical results for the preconditioning of Schur complement equation in Section 6.3, we consider numerical results for the case of non-convex domains too. As previously we consider the two-dimensional L-shaped domain $\Omega = (-\frac{1}{2},\frac{1}{2})^2 \backslash (-\frac{1}{2},0)^2$ with $J = 6$ and the Fichera cube $\Omega = (0,1)^3 \backslash (0,\frac{1}{2})^3$ with $J = 9$ smooth boundary parts.

	L–shape, $n = 2$				Fichera cube, $n = 3$			
L	DoFs	I	$C_{A_h,1}$	$C_{A_h,2}$	DoFs	I	$C_{A_h,1}$	$C_{A_h,2}$
0	14	14	14	12	30	22	21	22
1	50	46	29	22	166	92	47	40
2	194	174	39	29	1 180	>500	70	57
3	770	>500	50	36	9 144		88	71
4	3 074		59	40	72 560		111	87
5	12 290		66	45	579 296		133	103
6	49 154		72	48	4 632 000		159	122
7	196 610		77	52				
8	786 434		81	55				
9	3 145 730		89	59				
10	12 582 914		95	69				

Table 6.8: GMRES iterations with and without preconditioning.

In Table 6.8 we present the corresponding numerical results. We observe for both domains a logarithmic behavior of the iteration numbers and a significant influence of the number of smooth boundary parts J.

6.5 Concluding remarks

In this chapter we have considered the construction of preconditioners for the biharmonic equation of first kind (5.1). At first, we have developed a preconditioner for the Schur complement system with respect to the boundary degrees of freedom. The analysis has shown a constant condition number for the local single layer potential preconditioner C_{SLP} (6.17) and a logarithmic dependency for the BPX preconditioner C_{BPX} (6.23). For both cases, the presented numerical examples illustrate these properties. In the last part of the chapter we have presented two global, block diagonal, preconditioners. Numerical examples have affirmed the expected behavior.

As an outlook we would like to mention the extension of the spectral equivalence estimates for the Schur complement preconditioner to the case of non-convex domains. Another interesting and important question is a possible modification of the BPX preconditioner in order to eliminate the logarithmic dependency in the condition number. Also, we would like to name a possible application to adaptive methods, in h and p, as well as the analysis of the global preconditioner.

Furthermore, we would like to mention the application of these ideas for the construction of preconditioners to boundary control problems in the energy space, which will be the topic of the next chapter.

7 PRECONDITIONING STRATEGIES FOR OPTIMAL BOUNDARY CONTROL PROBLEMS

The aim of this chapter is to present a unified analysis and construction of preconditioners for boundary control problems. While the main focus of the research is aimed to the construction and numerical analysis of preconditioners for distributed control problems, see, e.g., [21, 68, 85], less work seems to be done in the case of preconditioners for boundary control problems. In a recent article, [32], multilevel preconditioners for the solution of unconstrained Neumann boundary control problems were considered. The proposed preconditioner turns out to be robust with respect to all regularization parameters. However, since the discretization of the related optimality system is a mixed finite element scheme, a discrete stability condition for the discrete Neumann control and the discrete Dirichlet trace of the state has to be ensured. While this stability condition excludes the use of standard piecewise linear finite elements for the state and associated piecewise constants for the Neumann control, piecewise linear basis functions are used for the approximation of the control in [32]. Although one may consider basis functions which are discontinuous across corners or edges, such an approach seems not to be very practicable from an application point of view.

As we have seen in (3.28) and (3.57), the Schur complement system, solved for both the Dirichlet and Neumann boundary control problem, is a linear combination of the Schur complement matrix which is related to a mixed finite element approximation of the biharmonic equation of first kind, and of the finite element approximation of the Steklov–Poincaré operator which realizes the Dirichlet to Neumann map related to the homogeneous partial differential equation. Moreover, the latter is scaled by the cost or regularization parameter. Consequently the same ideas for preconditioning of Dirichlet and Neumann boundary control problems can be applied. Since the system matrix is an additive composition of discrete pseudo differential operators of different order, i.e., of orders ± 1, an appropriate preconditioner has to take care of this behavior. In particular, multilevel operators, see, e.g., [10, 60], are known to be applicable in this situation. Within this section we discuss the construction of preconditioners for the optimal Dirichlet boundary control problem (3.1)–(3.2). In particular, we focus on the robustness of the preconditioner with respect to the mesh size h and the cost coefficient ϱ, i.e., iteration numbers are independent of these quantities.

This chapter is organized as follows: In the beginning we recall the Schur complement equation with respect to the control and present corresponding spectral equivalence estimates. Afterwards, we present two different type of preconditioners for the model problem. First, we consider a preconditioner as a combination of the local single layer potential operator

and the hypersingular operator. For this approach the proof of a constant spectral condition number is presented. Since it might be more applicable to use a multilevel method in practice, we also present the construction of corresponding multilevel BPX type preconditioner. In particular, it is an extended version for the biharmonic equation, discussed in Chapter 6, where we obtain a logarithmic dependency in the spectral condition number for domains with a piecewise smooth boundary. Numerical examples illustrate the obtained theoretical results.

7.1 Schur complement preconditioners

Starting point is, as in the construction of the preconditioner for the biharmonic equation, the linear system (3.25). Due to the separation of interior and boundary degrees of freedom corresponding, i.e. $\underline{u} = (\underline{u}_I, \underline{u}_C)^\top$ with $\underline{u}_I \in \mathbb{R}^{n_I}$ and $\underline{u}_C \in \mathbb{R}^{n_C}$, we obtain the linear system (3.26)

$$\begin{pmatrix} M_{II} + \varrho A_{II} & A_{II} & M_{IC} + \varrho A_{IC} \\ A_{II} & & A_{IC} \\ M_{CI} + \varrho A_{CI} & A_{CI} & M_{CC} + \varrho A_{CC} \end{pmatrix} \begin{pmatrix} \underline{u}_I \\ \widetilde{\underline{p}}_I \\ \underline{u}_C \end{pmatrix} = \begin{pmatrix} \widetilde{\underline{f}}_I \\ \widetilde{\underline{f}}_I \\ \underline{f}_C \end{pmatrix}.$$

Our aim is the construction of a block diagonal preconditioner

$$\begin{pmatrix} C_{A_h} & \\ & C_{T_h + \varrho S_h} \end{pmatrix}, \tag{7.1}$$

where $C_{A_h} \in \mathbb{R}^{2n_I \times 2n_I}$ denotes a preconditioner for the upper left 2×2 block, while $C_{T_h+\varrho S_h} \in \mathbb{R}^{n_C \times n_C}$ denotes the one corresponding to the Schur complement system. From the linear system (3.26) we obtain the Schur complement system (3.27),

$$\begin{aligned} \Big[\big[M_{CC} - M_{CI} A_{II}^{-1} A_{IC} - A_{CI} A_{II}^{-1} M_{IC} + A_{CI} A_{II}^{-1} M_{II} A_{II}^{-1} A_{IC} \big] \\ + \varrho \big[A_{CC} - A_{CI} A_{II}^{-1} A_{IC} \big] \Big] \underline{u}_C \\ = \big[A_{CI} A_{II}^{-1} M_{II} - M_{CI} \big] A_{II}^{-1} \underline{f}_I + \widetilde{\underline{f}}_C - A_{CI} A_{II}^{-1} \widetilde{\underline{f}}_I, \end{aligned}$$

where the corresponding Schur complement matrix (3.28) is given by

$$\begin{aligned} T_h + \varrho S_h = M_{CC} - M_{CI} A_{II}^{-1} A_{IC} - A_{CI} A_{II}^{-1} A_{IC} + A_{CI} A_{II}^{-1} M_{II} A_{II}^{-1} A_{IC} \\ + \varrho \big[A_{CC} - A_{CI} A_{II}^{-1} A_{IC} \big]. \end{aligned}$$

As it was mentioned in the end of Section 3.1, it consists of a biharmonic part T_h, see (6.4), and the cost coefficient ϱ weighted part $S_h = A_{CC} - A_{CI} A_{II}^{-1} A_{IC}$, which is the discrete Galerkin matrix of the Steklov–Poincaré operator, see [39, 59]. Note that the Schur complement

system with the unknown $\underline{u}_C$ is an equation for the discrete control z_h, since we have the isomorphism $z_h \leftrightarrow \underline{u}_C \in \mathbb{R}^{n_C}$.

For all $\underline{v}_C \in \mathbb{R}^{n_C}$ we introduce

$$\underline{v}_I := -A_{II}^{-1} A_{IC} \underline{v}_C \in \mathbb{R}^{n_I},$$

and thus rewrite the induced bilinear form of the Schur complement (3.28) as

$$\begin{aligned}((T_h + \varrho S_h)\underline{v}_C, \underline{v}_C) &= (M_{CC}\underline{v}_C, \underline{v}_C) - 2(M_{CI}A_{II}^{-1}A_{IC}\underline{v}_C, \underline{v}_C) + (M_{II}A_{II}^{-1}A_{IC}\underline{v}_C, A_{II}^{-1}A_{IC}\underline{v}_C) \\ &\quad + \varrho\left[(A_{CC}\underline{v}_C, \underline{v}_C) - (A_{CI}A_{II}^{-1}A_{IC}\underline{v}_C, \underline{v}_C)\right] \\ &= (M_{CC}\underline{v}_C, \underline{v}_C) + 2(M_{CI}\underline{v}_I, \underline{v}_C) + (M_{II}\underline{v}_I, \underline{v}_I) \\ &\quad + \varrho\left[(A_{CC}\underline{v}_C, \underline{v}_C) - 2(A_{CI}A_{II}^{-1}A_{IC}\underline{v}_C, \underline{v}_C) + (A_{CI}A_{II}^{-1}A_{IC}\underline{v}_C, \underline{v}_C)\right] \\ &= (M_{CC}\underline{v}_C, \underline{v}_C) + 2(M_{CI}\underline{v}_I, \underline{v}_C) + (M_{II}\underline{v}_I, \underline{v}_I) \\ &\quad + \varrho\left[(A_{CC}\underline{v}_C, \underline{v}_C) - 2(A_{CI}\underline{v}_I, \underline{v}_C) + (A_{II}\underline{v}_I, \underline{v}_I)\right].\end{aligned}$$

By using the isomorphism $\underline{v} = (\underline{v}_I, \underline{v}_C)^\top \in \mathbb{R}^{n_I+n_C} \leftrightarrow v_h \in \mathcal{V}_h$ we finally obtain

$$\begin{aligned}((T_h + \varrho S_h)\underline{v}_C, \underline{v}_C) &= (M_h\underline{v}, \underline{v}) + \varrho(A_h\underline{v}, \underline{v}) = \langle v_h, v_h\rangle_{L^2(\Omega)} + \varrho\langle\nabla v_h, \nabla v_h\rangle_{L^2(\Omega)} \\ &= \|v_h\|^2_{L^2(\Omega)} + \varrho|v_h|^2_{H^1(\Omega)}.\end{aligned} \tag{7.2}$$

As in Chapter 6, we would like to point out that $v_h \in \mathcal{V}_h$ is the discrete harmonic extension of $v_{h|\Gamma} \leftrightarrow \underline{v}_C \in \mathbb{R}^{n_C}$ which is the unique solution of the variational problem

$$\langle\nabla v_h, \nabla q_h\rangle_{L^2(\Omega)} = 0,$$

for all $q_h \in \mathcal{Q}_h$.

In order to construct the preconditioner for the Schur complement $T_h + \varrho S_h$ we need spectral equivalence estimates with respect to the boundary and correspondingly efficient invertible operators. In Chapter 6 we already treated the term $\|\cdot\|_{L^2(\Omega)}$ of (7.2) for a harmonic function. Here, we have to consider the additional term $|\cdot|_{H^1(\Omega)}$. Since $v_h \in \mathcal{V}_h$, constructed as above, is harmonic, it easily follows that

$$c_1|v_{h|\Gamma}|^2_{H^{1/2}(\Gamma)} \le |v_h|^2_{H^1(\Omega)} \le c_2|v_{h|\Gamma}|^2_{H^{1/2}(\Gamma)}. \tag{7.3}$$

Further, let us consider piecewise linear and globally continuous shape functions on the boundary by restriction to the boundary, i.e. $\phi_k := \varphi_{n_I+k|\Gamma}$ for all $k = 1, \ldots, n_C$. Then we can define, as in Chapter 6, the finite element trace space

$$Z_h := \operatorname{span}\{\phi_k\}_{k=1}^{n_C} = \mathcal{V}_{h|\Gamma} = \operatorname{span}\{\varphi_{n_I+k|\Gamma}\}_{k=1}^{n_C} \subset H^{1/2}(\Gamma). \tag{7.4}$$

According to Theorem 6.4 and (7.3) the following statement holds true.

Theorem 7.1. *Let $\Omega \subset \mathbb{R}^n$, $n = 2,3$, be a convex and bounded Lipschitz domain with a piecewise smooth boundary Γ. Then, for all $\underline{z}_C \in \mathbb{R}^{n_C} \leftrightarrow z_h \in Z_h$ there hold the spectral equivalence inequalities*

$$((T_h + \varrho S_h)\underline{z}_C, \underline{z}_C) \simeq \|z_h\|^2_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} + \varrho |z_h|^2_{H^{1/2}(\Gamma)}. \tag{7.5}$$

The following Lemma might be of interest for small cost coefficients and be advantageous in implementation of some preconditioner, where the semi-norm part can be replaced by the full norm.

Lemma 7.1. *Let the assumptions of Theorem 7.1 be valid and let us further assume a cost coefficient $\varrho < 1$. Then, for all $\underline{z}_C \in \mathbb{R}^{n_C} \leftrightarrow z_h \in Z_h$ there hold the spectral equivalence inequalities*

$$((T_h + \varrho S_h)\underline{z}_C, \underline{z}_C) \simeq \|z_h\|^2_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} + \varrho \|z_h\|^2_{H^{1/2}(\Gamma)}.$$

Proof. We only need to show the lower estimate, since the upper one is trivially satisfied by Theorem 7.1. By the definition of the norm in $\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)$ we obtain

$$\|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} = \sup_{0 \neq \phi \in H^{1/2}_{\mathrm{pw}}(\Gamma)} \frac{\langle z_h, \phi \rangle_\Gamma}{\|\phi\|_{H^{1/2}_{\mathrm{pw}}(\Gamma)}} \geq \frac{1}{|\Gamma|} \langle z_h, 1 \rangle_\Gamma,$$

and consequently, because of the equivalence of norms

$$\begin{aligned}
\|z_h\|^2_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} + \varrho |z_h|^2_{H^{1/2}(\Gamma)} &\geq c \left(\|z_h\|^2_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} + \langle z_h, 1 \rangle^2_\Gamma \right) + \varrho |z_h|^2_{H^{1/2}(\Gamma)} \\
&\geq c \left(\|z_h\|^2_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} + \varrho \langle z_h, 1 \rangle^2_\Gamma + \varrho |z_h|^2_{H^{1/2}(\Gamma)} \right) \\
&\geq c \left(\|z_h\|^2_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} + \varrho \|z_h\|^2_{H^{1/2}(\Gamma)} \right),
\end{aligned}$$

which concludes the proof. □

Note that the spectral equivalence estimates (7.5) have the advantage that they are satisfied for all cost coefficients $\varrho > 0$.

For the iterative solution of the Schur complement equation (3.27) with the symmetric, positive definite and parameter–dependent system matrix $T_h + \varrho S_h$ we need to have a preconditioning matrix $C_{T_h + \varrho S_h}$ which is robust with respect to the discretization parameter h, and with respect to the cost coefficient ϱ.

In what follows we present two preconditioners for the Schur complement equation (3.27). First we consider a preconditioner, derived from boundary element methods, which is a

combination of the local single layer potential for the $\widetilde{H}_{\mathrm{pw}}^{-1/2}(\Gamma)$ part, see Chapter 6, and the hypersingular operator for the $H^{1/2}(\Gamma)$ part. This combination results in a constant spectral condition number of the preconditioned system. The second preconditioner is based on the multilevel idea, where a BPX type preconditioner is derived. As in the case of the biharmonic equation we obtain a spectral condition which is constant up to a logarithmic term. Numerical examples illustrate the obtained theoretical results.

7.1.1 The SLP–HYP preconditioner

As it was shown in Chapter 6, the local single layer potential induces an spectral equivalent norm to the norm of $\widetilde{H}_{\mathrm{pw}}^{-1/2}(\Gamma)$. It remains to construct an operator which reflects the semi-norm in $H^{1/2}(\Gamma)$. Consequently, the weighted sum of both would then satisfy the spectral equivalence estimate (7.5).

For a given $\psi \in H^{1/2}(\Gamma)$, we define the double layer potential

$$(W\psi)(x) = \int_{\Gamma} \left[\partial_{n,y} U^*(x,y)\right] \psi(y)\, ds_y,$$

for $x \in \Omega$, where $U^*(x,y)$ denotes the fundamental solution of the Laplace equation. From this we can deduce the so-called hypersingular boundary integral operator $D : H^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$ defined by the negative Neumann trace of double layer potential, i.e.

$$(D\psi)(x) = -\partial_n (W\psi)(x),$$

for $x \in \Gamma$. The duality product, induced by the hypersingular operator, represents an equivalent semi-norm in $H^{1/2}(\Gamma)$, i.e.

$$\langle D\psi, \psi \rangle_{\Gamma} \simeq |\psi|^2_{H^{1/2}(\Gamma)},$$

for all $\psi \in H^{1/2}(\Gamma)$, see, e.g., [73, Theorem 6.24].

Let us recall from Chapter 5 that a piecewise smooth boundary Γ, can be decomposed into smooth parts, i.e.

$$\Gamma = \bigcup_{i=1}^{J} \Gamma_i.$$

Therefore we consider the preconditioner (6.17) for the $H_{\mathrm{pw}}^{-1/2}\Gamma$ norm, with the local single layer boundary integral operators $V_i : \widetilde{H}^{-1/2}(\Gamma_i) \to H^{1/2}(\Gamma_i)$, for all $i = 1, \ldots, J$. In combination with the hypersingular operator leads this to the following Schur complement preconditioner

$$C_{\mathrm{SLP\text{–}HYP}} := \sum_{i=1}^{J} A_i^{\top} V_{i,h} A_i + \varrho D_h. \tag{7.6}$$

Further, we obtain for the preconditioned system, due to the spectral equivalence of both operators, the spectral condition number

$$\kappa(C_{\mathrm{SLP-HYP}}^{-1}(T_h + \varrho S_h)) \leq c,$$

bounded by a constant which independent of the discretization, i.e. independent of the mesh size h and the cost coefficient ϱ.

7.1.2 The BPX preconditioner

For the derivation of a multilevel preconditioner for the Schur complement matrix T_h of the biharmonic equation, see Chapter 6, we obtained in Theorem 6.5 the spectral equivalence estimates

$$\|z_h\|_{H^{-1/2}(\Gamma)} \leq \|z_h\|_{\widetilde{H}^{-1/2}_{\mathrm{pw}}(\Gamma)} \leq \overline{c}_2 J[1-\log h]\|z_h\|_{H^{-1/2}(\Gamma)},$$

which are satisfied for all $z_h \in Z_h$. Note that the logarithmic factor does not appear when Γ is smooth. Hence, by combining the above estimate with (7.5) we conclude the spectral equivalence inequalities

$$\begin{aligned}\widetilde{c}_1 \left[\|z_h\|^2_{H^{-1/2}(\Gamma)} + \varrho|z_h|^2_{H^{1/2}(\Gamma)}\right] &\leq ((T_h + \varrho S_h)\underline{z}_C, \underline{z}_C) \\ &\leq \widetilde{c}_2 \max\left\{1, J^2[1-\log h]^2\right\}\left[\|z_h\|^2_{H^{-1/2}(\Gamma)} + \varrho|z_h|^2_{H^{1/2}(\Gamma)}\right],\end{aligned} \tag{7.7}$$

for all $\underline{z}_C \in \mathbb{R}^{n_C} \leftrightarrow z_h \in Z_h$. Since the spectral equivalence inequalities (7.7) involve fractional Sobolev norms of different order, an appropriate preconditioner has to take this behavior into consideration. A possible choice are multilevel preconditioners, see, e.g., [10, 60].

As in the construction of a multilevel preconditioner for the biharmonic equation in Chapter 6, we consider sequence of admissible globally quasi–uniform nested finite element meshes $\{\mathcal{T}_{h_i}\}_{i\in\mathbb{N}_0}$ of mesh size $h_i \simeq 2^{-i}$. Let $\{V_{h_i}\}_{i\in\mathbb{N}_0} \subset H^1(\Omega)$ denote the related sequence of finite element spaces with piecewise linear and globally continuous shape function. Let $Z_i = V_{h_i|\Gamma}$ denote the restrictions on the boundary. This results in a sequence of nested spaces of the form

$$Z_0 \subset Z_1 \subset \ldots Z_L = Z_h \subset Z_{L+1} \subset \cdots \subset H^{1/2}(\Gamma),$$

where L denotes the current level of interest. With respect to the boundary element spaces Z_i of piecewise linear and globally continuous shape functions, we introduce the L^2 projection $Q_i : L^2(\Gamma) \to Z_i$, $i \in \mathbb{N}_0$, and we set $Q_{-1} := 0$. It turns out, see, e.g., [10, 60, 73], that the multilevel operator

$$B_s := \sum_{i=0}^{\infty} h_i^{-2s}(Q_i - Q_{i-1}),$$

defines an equivalent norm in $H^s(\Gamma)$, for all $|s| < \frac{3}{2}$. For $s = -\frac{1}{2}$ the operator $B_{-1/2}$ defines an equivalent norm in $H^{-1/2}(\Gamma)$, and consequently a preconditioner for the complement matrix T_h of the biharmonic equation. It remains, a preconditioner for the semi–norm in $H^{1/2}(\Gamma)$. Therefore we consider

$$|v|^2_{H^{1/2}(\Gamma)} = \|v\|^2_{H^{1/2}(\Gamma)} - \|v\|^2_{L^2(\Gamma)} \simeq \langle B_{1/2}v, v\rangle_\Gamma - \|v\|^2_{L^2(\Gamma)} = \langle (B_{1/2} - B_0)v, v\rangle_\Gamma,$$

where we used the representation

$$\|v\|^2_{L^2(\Gamma)} = \langle B_0 v, v\rangle_{L^2(\Gamma)},$$

for all $v \in H^{1/2}(\Gamma)$. Thus, we finally define the preconditioning operator as

$$A := \varrho(B_{1/2} - B_0) + B_{-1/2},$$

i.e. we conclude the multilevel representation

$$A := \sum_{i=0}^{\infty} \left[\varrho(h_i^{-1} - 1) + h_i\right](Q_i - Q_{i-1}).$$

Note that we obtain for inverse multilevel operator the representation

$$A^{-1} = \sum_{i=0}^{\infty} \left[\varrho(h_i^{-1} - 1) + h_i\right]^{-1}(Q_i - Q_{i-1}).$$

As in [73, Corollary 13.7] we finally conclude that the preconditioner $C_{T_h + \varrho S_h}$ of the Schur complement matrix $T_h + \varrho S_h$ (3.28) is given by

$$C_{\mathrm{BPX}} = M_{h_L} B_{h_L}^{-1} M_{h_L}, \tag{7.8}$$

where

$$M_{h_L}[\ell, k] = \langle \phi_k^L, \phi_\ell^L\rangle_{L^2(\Gamma)}, \quad B_{h_L}[\ell, k] = \langle A^{-1}\phi_k^L, \phi_\ell^L\rangle_\Gamma$$

for all $k, \ell = 1, \ldots, n_{C,L}$ denote the standard mass matrix and the Galerkin matrix of the multilevel operator A^{-1}. For the application $\underline{v} = C_{\mathrm{BPX}}^{-1}\underline{r}$, e.g., within a conjugate gradient scheme, we obtain, as for the standard multilevel approach [10, 73], the representation

$$\underline{v} = \sum_{i=0}^{L} \alpha_i R_i M_{h_i}^{-1} R_i^\top \underline{r},$$

with coefficients

$$\alpha_i = \begin{cases} \dfrac{h_L}{\varrho(1-h_L) + h_L^2} & \text{for } i = L, \\ \dfrac{h_i}{\varrho(1-h_i) + h_i^2} - \dfrac{h_{i+1}}{\varrho(1-h_{i+1}) + h_{i+1}^2} & \text{for } i = 0, \ldots, L-1, \end{cases}$$

where $R_i : \mathbb{R}^{n_{C,i}} \to \mathbb{R}^{n_{C,L}}$ is the prolongation operator which is related to the nested sequence of piecewise linear finite element spaces on the boundary. While in the application of multilevel preconditioners for (pseudo)differential operators of positive order one can replace the inverse mass matrices $M_{h_i}^{-1}$ by their diagonals, this is in general not possible in the current situation, since the coefficients α_i might be negative. Note that this depends on h and ϱ. Consequently we need to use the inverse mass matrices $M_{h_i}^{-1}$, or its approximation, which can be realized by a few conjugate gradient iterations.

To summarize, due to the spectral equivalence inequalities (7.7) the spectral condition number of the preconditioned system is bounded as

$$\kappa(C_{\mathrm{BPX}}^{-1}(T_h + \varrho S_h)) \leq cJ^2[1 - \log h]^2, \tag{7.9}$$

where the constant c neither depends on the discretization parameter h, nor on the cost coefficient ϱ. Note that the logarithmic factor does not appear when Γ is smooth and that J depends only on the geometry Ω and not on the discretization.

7.1.3 Numerical results

In the following we present numerical examples which illustrate the obtained theoretical results from the previous part. As computational domains we consider $\Omega = B_{1/2}(0)$ and $\Omega = (0, \frac{1}{2})^n$, both for $n = 2, 3$, and the given data (3.29), i.e.

$$f = 0, \qquad \overline{u} = \left(\sum_{i=1}^{n} (x_i(x_i - 1/2) + 1)^2 \right)^{1/2},$$

for different cost coefficients $\varrho > 0$. The linear system (3.27) is solved by a conjugate gradient scheme with preconditioning up to a relative error reduction of $\varepsilon = 10^{-8}$, where the preconditioner is either the combination of local single layer potential and hypersingular boundary integral operator $C_{\mathrm{SLP\text{–}HYP}}$ (7.6) or the multilevel BPX preconditioner C_{BPX} (7.8). In the following tables we present, for a sequence of different refinement levels L and cost coefficients $\varrho = 10^{-i}$, the number of required PCG iterations. For the related numbers $n_{I,L}$ and $n_{C,L}$, denoting the degrees of freedom in the interior and on the boundary, respectively, we refer to the numerical results in Chapter 6.

Example 1

In the first numerical example we consider the preconditioner arising from boundary element methods, $C_{\mathrm{SLP\text{–}HYP}}$. The corresponding results for the tow dimensional model problem are presented in Table 7.1 and Table 7.2. We observe in both cases, i.e. for $\Omega = B_{1/2}(0)$ and $\Omega = (0, \frac{1}{2})^n$, for all cost coefficient $\varrho = 10^{-i}$, $i = 0, \ldots, 10$, constant iteration numbers with respect to the mesh size h. Nevertheless, the iterations slightly vary with respect to

the cost coefficients, i.e. we obtain increasing iterations until $\varrho = 10^{-5}$ and afterwards they decrease again. This phenomena was also observed in [9].

$L\backslash i$	0	1	2	3	4	5	6	7	8	9	10
0	3	3	3	3	3	3	3	3	3	3	3
1	5	5	5	5	5	5	5	5	5	5	5
2	7	9	9	9	9	9	9	9	9	9	9
3	9	12	14	14	13	13	13	13	13	13	13
4	10	12	16	17	16	15	14	14	14	14	14
5	9	12	16	18	18	16	15	15	15	15	15
6	9	11	15	18	18	17	15	15	15	15	15
7	8	11	15	17	18	17	15	14	14	14	14
8	8	10	14	16	17	17	16	14	14	14	14
9	7	9	13	15	16	17	17	15	14	13	13
10	6	9	13	14	15	16	16	16	14	13	13

Table 7.1: $C_{\text{SLP–HYP}}$ iterations for $\varrho = 10^{-i}$, $\Omega = B_{1/2}(0)$.

$L\backslash i$	0	1	2	3	4	5	6	7	8	9	10
0	2	2	2	2	2	2	2	2	2	2	2
1	2	2	2	2	2	2	2	2	2	2	2
2	3	3	3	3	3	3	3	3	3	3	3
3	5	5	5	5	5	5	5	5	5	5	5
4	9	9	9	9	9	9	9	9	9	9	9
5	9	10	11	13	15	12	11	10	10	10	10
6	9	10	11	14	16	15	12	11	11	11	11
7	9	10	11	13	16	16	14	11	11	11	11
8	9	10	11	13	16	17	16	13	11	11	11
9	9	9	10	12	15	16	16	15	12	11	11
10	9	9	10	12	15	16	16	16	13	11	11

Table 7.2: $C_{\text{SLP–HYP}}$ iterations for $\varrho = 10^{-i}$, $\Omega = (0, \frac{1}{2})^2$.

Example 2

As a second example we consider the BPX preconditioner C_{BPX}. First we consider $\Omega = B_{1/2}(0)$ for $n = 2$, where we observe constant iteration numbers with respect to the mesh size h, see Table 7.3. On the other hand, we observe in Table 7.4 and Table 7.5 a logarithmic factor of h in the iteration numbers for the cube for $n = 2, 3$, which corresponds to the theoretical results in the previous subsection. Furthermore we observe, as for the boundary element preconditioner, an increase and decrease with respect to the cost coefficient, see [9].

$L\backslash i$	0	1	2	3	4	5	6	7	8	9	10
0	2	2	2	2	2	2	2	2	2	2	2
1	5	5	5	5	5	5	5	5	5	5	5
2	9	9	9	9	8	8	8	8	8	8	8
3	15	14	16	14	12	11	11	11	11	11	11
4	17	17	20	22	15	12	12	12	12	12	12
5	18	17	22	28	20	14	13	13	13	13	13
6	18	17	23	29	27	18	13	13	13	13	13
7	18	17	21	29	30	24	15	13	13	13	13
8	18	17	21	29	30	30	19	13	13	13	13
9	18	17	21	29	30	31	26	16	13	13	13
10	18	17	21	29	30	30	31	21	14	12	12

Table 7.3: C_{BPX} iterations for $\varrho = 10^{-i}$, $\Omega = B_{1/2}(0)$.

$L\backslash i$	0	1	2	3	4	5	6	7	8	9	10
0	2	2	2	2	2	2	2	2	2	2	2
1	2	2	2	2	2	2	2	2	2	2	2
2	3	3	3	3	3	3	3	3	3	3	3
3	5	5	5	5	5	5	5	5	5	5	5
4	9	9	9	9	9	9	9	9	9	9	9
5	15	15	16	17	17	13	13	12	12	12	12
6	18	17	20	26	31	23	14	14	14	14	14
7	19	18	20	29	41	37	19	15	15	15	15
8	18	18	20	29	42	49	29	16	15	15	15
9	18	18	20	29	43	54	45	22	16	15	15
10	19	18	21	29	44	55	55	36	18	16	16

Table 7.4: C_{BPX} iterations for $\varrho = 10^{-i}$, $\Omega = (0, \frac{1}{2})^2$.

Remark 7.1. *In all considered numerical examples we have seen constant or constant up to logarithmic term iteration numbers of the preconditioned linear system. Even though, we can prove the robustness of both preconditioners with respect to the cost coefficient ϱ, the numerical results show a slight increase and decrease. As mentioned before, this behavior was observed previously in a different context in* [9]*, where the authors gave no explanation for this phenomena.*

We would like to mention that in the BPX preconditioner C_{BPX} the coefficients α_i can be negative and positive, depending on h_i and ϱ. The refinement level on which the α_i turn from negative coefficients to positive ones, differ for different ϱ. This might be an indicator for the behavior of the iteration numbers. The question if a modified preconditioner, in particular of multilevel BPX type, can capture this behavior is open.

$L\backslash i$	0	1	2	3	4	5	6	7	8	9	10
0	4	2	2	2	2	2	2	2	2	2	2
1	14	14	14	13	13	13	13	13	13	13	13
2	30	34	40	27	23	22	22	22	22	22	22
3	36	42	63	56	32	27	27	27	27	27	27
4	38	46	73	87	50	30	29	29	29	39	29
5	38	47	76	108	86	44	33	33	33	33	33
6	38	45	75	116	121	74	40	37	37	37	37

Table 7.5: C_{BPX} iterations for $\varrho = 10^{-i}$, $\Omega = (0, \frac{1}{2})^3$.

7.2 Concluding remarks

In this chapter we have developed a preconditioner for the Schur complement equation of the optimal Dirichlet boundary control problem. Therefore, the ideas of the biharmonic equation of first kind form Chapter 6, were applicable. In particular, two preconditioners were introduced, where the first one was motivated from boundary element methods. It was possible to prove the robustness of this approach with respect to the mesh size h and the cost coefficient ϱ. For practical reasons we have also presented a mulitlevel method of BPX type. For this approach we have proven robustness up to a logarithmic term of the mesh size h. Numerical examples illustrated these results.

Since the Schur complement matrices for the Dirichlet and Neumann boundary control have the same structure, see (3.28) and (3.57), and in particular coincide in the case of the Laplace equation, we can use the same ideas for preconditioning of the Neumann boundary control problem.

As an outlook we would like to mention the construction of a global robust preconditioner. Moreover, it would be interesting to apply these ideas for the construction of a robust preconditioner for the Dirichlet and Neumann boundary control of the Stokes equations. And finally, to the Navier–Stokes equations and their time depended version, which could be used for many applications, see also Chapter 4.

Outlook and open problems

In the following, we summarize the open problems and possible extensions of this work. Even tough, some of them were already mention in the concluding remarks of the individual chapters, we would like to recapitulate them in a more compact form here.

In Chapter 6 and 7, we presented spectral equivalence estimates for the Schur complement preconditioner of the biharmonic equation and the optimal boundary control problem, where we made the crucial assumption that the domain has to be convex or the boundary is smooth. In the numerical results we presented some examples for non-convex domains, such as the L-shaped domain and the Fichera cube, too. For both we obtained the expected behavior of the preconditioner, even though the derived theory is not applicable. Consequently, it would be interesting to extend the spectral equivalence estimates to results for the non-convex case, which seems to be a challenging problem. Another interesting point is the study of the preconditioner for the corresponding adaptive problem, where local mesh refinement has to be taken into account. Further, we have already mentioned that the analysis of the global preconditioners should be included in a possible future work.

In the case of the optimal boundary control problems in Chapter 7, we observed for the preconditioner a slight variety in the iteration numbers with respect to the cost coefficient. This might be due to the fact that the coefficient in the multilevel method changes their sign, which is dependent on the mesh size and the cost coefficient. The question is, if it is possible to construct a modified multilevel preconditioner which is able to capture this problem. This can be a challenging and interesting problem, which could be also applied for several other problems. Moreover, we would like to extend this work to robust preconditioner for the boundary control of the Stokes equations. Here, we would have an additional parameter, namely the viscosity, which has to be taken into account in order to obtain robustness of the preconditioner.

In Chapter 3 we discussed error estimates for the optimal boundary control problem. An important an interesting question are the optimal error estimates on the boundary for the control, where we gain a factor of up to $1/2$ in the order of convergence, assuming the solution is regular enough. In order to prove these error estimates we would like to apply the ideas of [3, 55].

Within this thesis, we did not discuss any box constraints. One possibility to treat such additional constraints and their connection to Signorini type boundary conditions is discussed in [59]. Consequently, we may ask what happens in such a case with the preconditioner and which modifications we would have to make. As a last point we would like to mention the discretization of the control, for the Neumann control. We have eliminated the control

in the variational formulation for the Neumann control problem, which has the advantage that an inf–sup condition is not needed and the control can be calculated in a post processing step. In the case of box constraint the situation might be different, which is an interesting question to investigate.

For the application part of this thesis we would like to mention the optimization of wall shear stresses as it was discussed at the end of Chapter 4. Further, the study of the hemodynamic indicators for the fluid-structure-interaction problem and the construction of corresponding preconditioners would be challenging.

BIBLIOGRAPHY

[1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces.* Academic Press, New York, London, 2003.

[2] M. Anand and K. Rajagopal. A shear thinning viscoelastic fluid model for describing the flow of blood. *Int. J. of card. Med. Sci.*, 4(2):55–68, 2004.

[3] T. Apel, J. Pfefferer, and A. Rösch. Finite element error estimates on the boundary with application to optimal control. *DFG Priority Program 1253, Erlangen*, Preprint SPP1253-136, 2012.

[4] D. Bahlmann and U. Langer. A fast solver for the first biharmonic boundary value problem. *Numer. Math.*, 63(3):297–313, 1992.

[5] M. Beneš. Mixed initial-boundary value problem for the three-dimensional Navier-Stokes equations in polyhedral domains. *Discrete Contin. Dyn. Syst.*, (Dynamical systems, differential equations and applications. 8th AIMS Conference. Suppl. Vol. I):135–144, 2011.

[6] P. B. Bochev and C. R. Dohrmann. A stabilized finite element method for the Stokes problem based on polynomial pressure projections. *Internat. J. Numer. Methods Fluids*, 46(2):183–201, 2004.

[7] M. Bollhöfer, R. A. Römer, and O. Schenk. On large-scale diagonalization techniques for the Anderson model of localization. *SIAM Rev.*, 50(1):91–112, 2008.

[8] D. Braess and P. Peisker. On the numerical solution of the biharmonic equation and the role of squaring matrices for preconditioning. *IMA J. Numer. Anal.*, 6(4):393–404, 1986.

[9] J. H. Bramble, J. E. Pasciak, and P. S. Vassilevski. Computational scales of Sobolev norms with application to preconditioning. *Math. Comp.*, 69(230):463–480, 2000.

[10] J. H. Bramble, J. E. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 55(191):1–22, 1990.

[11] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods.* Springer, New York, 2008.

[12] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods.* Springer, New York, 1991.

[13] F. Brezzi and P. A. Raviart. Mixed finite element methods for 4th order elliptic equations. In *Topics in numerical analysis, III (Proc. Roy. Irish Acad. Conf., Trinity Coll., Dublin, 1976)*, pages 33–56, London, 1977. Academic Press.

[14] S. Chien, S. Usami, R. J. Dellenback, M. I. Gregersen, L. B. Nanninga, and M. M. Guest. Blood Viscosity: Influence of Erythrocyte Aggregation. *Science*, 157(3790):829–831, 1967.

[15] S. Chien, S. Usami, R. J. Dellenback, M. I. Gregersen, L. B. Nanninga, and M. M. Guest. Blood Viscosity: Influence of Erythrocyte Deformation. *Science*, 157(3790):827–829, 1967.

[16] P. G. Ciarlet. *The finite element method for elliptic problems.* North–Holland, Amsterdam, 1978.

[17] P. G. Ciarlet and P. A. Raviart. A mixed finite element method for the biharmonic equation. In *Mathematical aspects of finite elements in partial differential equations (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1974)*, number 33, pages 125–145, New York, 1974.

[18] K. S. Cunningham and A. I. Gotlieb. The role of shear stress in the pathogenesis of atherosclerosis. *Laboratory Investigation*, 85:9–23, 2005.

[19] J. C. de los Reyes and K. Kunisch. A semi-smooth Newton method for control constrained boundary optimal control of the Navier-Stokes equations. *Nonlinear Anal.*, 62(7):1289–1316, 2005.

[20] P. Deuflhard. *Newton methods for nonlinear problems. Affine invariance and adaptive algorithms.* Springer, Berlin, 2004.

[21] S. H. Dollar, T. Rees, and A. J. Wathen. Optimal solvers for PDE-constrained optimization. *SIAM J. Sci. Comput.*, 32(1):271–298, 2010.

[22] H. C. Elman, D. J. Silvester, and A. J. Wathen. *Finite elements and fast iterative solvers: with applications in incompressible fluid dynamics.* Oxford University Press, New York, 2005.

[23] L. Formaggia, A. Quarteroni, and A. Veneziani, editors. *Cardiovascular mathematics.* MS&A. Modeling, Simulation and Applications, 1. Springer Italia, Milan, 2009.

[24] L. P. Franca and S. L. Frey. Stabilized finite element methods. II. The incompressible Navier-Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, 99(2–3):209–233, 1992.

[25] L. P. Franca, T. J. R. Hughes, and R. Stenberg. Stabilized Finite Element Methods for the Stokes Problem. *Incomp. Comp. Fluid Dynamics*, (4):87–107, 1993.

[26] S. A. Funken and E. P. Stephan. The BPX preconditioner for the single layer potential operator. *Appl. Anal.*, 67(3–4):327–340, 1997.

[27] G. P. Galdi. An introduction to the Navier-Stokes initial-boundary value problem. In *Fundamental directions in mathematical fluid mechanics*, Adv. Math. Fluid Mech., pages 1–70. Birkhäuser, Basel, 2000.

[28] G. P. Galdi, R. Rannacher, A. M. Robertson, and S. Turek. *Hemodynamical flows.* Oberwolfach Seminars, Birkhäuser, Basel, 2009.

[29] A. M. Gambaruto, J. Janela, A. Moura, and A. Sequeira. Sensitivity of hemodynamics in a patient specific cerebral aneurysm to vascular geometry and blood rheology. *Math. Biosci. Eng.*, 8(2):409–423, 2011.

[30] V. Girault and P. A. Raviart. *Finite Element Methods for Navier-Stokes Equations.* Springer, New York, 1986.

[31] R. Glowinski and O. Pironneau. Numerical methods for the first biharmonic equation and the two-dimensional Stokes problem. *SIAM Rev.*, 21(2):167–212, 1979.

[32] E. Goncalves and M. Sarkis. Analysis of Robust Parameter-Free Multilevel Methods for Neumann Boundary Control Problems. *Comput. Methods Appl. Math.*, 13(2):207–235, 2013.

[33] P. Grisvard. *Elliptic problems in nonsmooth domains.* Pitman, Boston, 1985.

[34] W. Hackbusch. *Multigrid methods and applications.* Springer, Berlin, 1985.

[35] M. R. Hanisch. Multigrid preconditioning for the biharmonic Dirichlet problem. *SIAM J. Numer. Anal.*, 30(1):184–214, 1993.

[36] N. Heuer. Additive Schwarz method for the p-version of the boundary element method for the single layer potential operator on a plane screen. *Numer. Math.*, 88(3):485–511, 2001.

[37] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints.* Mathematical Modelling: Theory and Applications, 23. Springer, New York, 2009.

[38] G. C. Hsiao and W. L. Wendland. *Boundary integral equations.* Springer, Berlin, 2008.

[39] L. John. Stabilized finite element methods for Dirichlet boundary control problems in fluid mechanics. Master's thesis, Institute of Computational Mathematics, Graz University of Technology, 2011.

[40] L. John, P. Pustějovská, and O. Steinbach. On the influence of the wall shear stress vector form on hemodynamic indicators. *Reports of the Institute of Computational Mathematics*, Report 2013/6, Graz University of Technology, 2013.

[41] L. John and O. Steinbach. Schur complement preconditioners for the biharmonic Dirichlet boundary value problem. *Reports of the Institute of Computational Mathematics*, Report 2013/4, Graz University of Technology, 2013.

[42] L. John and O. Steinbach. Schur complement preconditioners for boundary control problems. *Reports of the Institute of Computational Mathematics*, Report 2014/4, Graz University of Technology, 2014.

[43] K. N. Kayembe, M. Sasahara, and F. Hazama. Cerebral aneurysms and variations in the circle of Willis. *Stroke*, 15(5):846–850, 1984.

[44] H. Koch and V. A. Solonnikov. L_p-estimates for a solution to the nonstationary Stokes equations. *J. Math. Sci.*, 106(3):3042–3072, 2001.

[45] D. N. Ku. Blood Flow in Arteries. *Annual Review of Fluid Mechanics*, 29:399–434, 1997.

[46] D. N. Ku, D. P. Giddens, C. K. Zarins, and S. Glagov. Pulsatile flow and atherosclerosis in the human carotid bifurcation. Positive correlation between plaque location and low oscillating shear stress. *Arteriosclerosis*, 5(3):293–302, 1985.

[47] K. Kunisch and B. Vexler. Optimal vortex reduction for instationary flows based on translation invariant cost functionals. *SIAM J. Control Optim.*, 46(4):1368–1397, 2007.

[48] U. Langer. Zur numerischen Lösung des ersten biharmonischen Randwertproblems. *Numer. Math.*, 50(3):291–310, 1987.

[49] J. L. Lions. *Optimal control of systems governed by partial differential equations.* Springer, New York, 1971.

[50] A. M. Malek, S. L. Alper, and S. Izumo. Hemodynamic shear stress and its role in atherosclerosis. *JAMA*, 282(21):2035–2042, 1999.

[51] J. Málek, J. Nečas, and M. Růžička. On the non-Newtonian incompressible fluids. *Math. Models Methods Appl. Sci.*, 3(1):35–63, 1993.

[52] J. Málek and K. R. Rajagopal. Mathematical issues concerning the Navier-Stokes equations and some of its generalizations. In *Evolutionary equations. Vol. II*, Handb. Differ. Equ., pages 371–459. Elsevier/North-Holland, Amsterdam, 2005.

[53] W. McLean. *Strongly elliptic systems and boundary integral equations.* Cambridge University Press, Cambridge, 2000.

[54] W. McLean and O. Steinbach. Boundary element preconditioners for a hypersingular integral equation on an interval. *Adv. Comput. Math.*, 11(4):271–286, 1999.

[55] J. M. Melenk and B. Wohlmuth. Quasi-optimal approximation of surface based Lagrange multipliers in finite element methods. *SIAM J. Numer. Anal.*, 50(4):2064–2087, 2012.

[56] M. D. Mihajlović and D. Silvester. A black-box multigrid preconditioner for the biharmonic equation. *BIT*, 44(1):151–163, 2004.

[57] P. Mund, E. P. Stephan, and J. Weiße. Two-level methods for the single layer potential in $\mathbb{R}^3$. *Computing*, 66(3):243–266, 1998.

[58] G. Of. An efficient algebraic multigrid preconditioner for a fast multipole boundary element method. *Computing*, 82(2–3):139–155, 2008.

[59] G. Of, T. X. Phan, and O. Steinbach. An energy space finite element approach for elliptic Dirichlet boundary control problems. *Numer. Math.*, 2014. published online.

[60] P. Oswald. *Multilevel finite element approximation. Theory and applications.* Teubner, Stuttgart, 1994.

[61] P. Peisker. A multilevel algorithm for the biharmonic problem. *Numer. Math.*, 46(4):623–634, 1985.

[62] P. Peisker. On the numerical solution of the first biharmonic equation. *RAIRO Modél. Math. Anal. Numér.*, 22(4):655–676, 1988.

[63] T. X. Phan. *Boundary Element Methods for Boundary Control Problems.* Monographic Series TU Graz, Computation in Engineering and Science, vol. 9, 2011.

[64] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations.* Springer, Berlin, 1997.

[65] A. Quarteroni and A. Valli. *Domain decomposition methods for partial differential equations.* Oxford University Press, New York, 1999.

[66] S. Ramalho, A. Moura, A. M. Gambaruto, and A. Sequeira. Sensitivity to outflow boundary conditions and level of geometry description for a cerebral aneurysm. *Int. J. Numer. Methods Biomed. Eng.*, 28(6–7):697–713, 2012.

[67] H. Samady, P. Eshtehardi, M. C. McDaniel, J. Suo, S. S. Dhawan, C. Maynard, L. H. Timmins, A. A. Quyyumi, and D. P. Giddens. Coronary artery wall shear stress is associated with progression and transformation of atherosclerotic plaque and arterial remodeling in patients with coronary artery disease. *Circulation*, 124(7):779–788, 2011.

[68] J. Schöberl and W. Zulehner. Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems. *SIAM J. Matrix Anal. Appl.*, 29(3):752–773, 2007.

[69] R. Scholz. A mixed method for 4th order problems using linear finite elements. *RAIRO Anal. Numér.*, 12(1):85–90, iii, 1978.

[70] D. M. Sforza, C. M. Putman, and J. R. Cebral. Hemodynamics of Cerebral Aneurysms. *Annual Review of Fluid Mechanics*, 41:91–107, 2009.

[71] J. Simon. On the existence of the pressure for solutions of the variational Navier-Stokes equations. *J. Math. Fluid Mech.*, 1(3):225–234, 1999.

[72] O. Steinbach. *Stability estimates for hybrid coupled domain decomposition methods.* Springer, Berlin, 2003.

[73] O. Steinbach. *Numerical Approximation Methods for Elliptic Boundary Value Problems. Finite and Boundary Elements.* Springer, New York, 2008.

[74] O. Steinbach. Boundary element methods for variational inequalities. *Numer. Math.*, 126(1):173–197, 2014.

[75] O. Steinbach and W. L. Wendland. The construction of some efficient preconditioners in the boundary element method. *Adv. Comput. Math.*, 9(1–2):191–216, 1998.

[76] R. Temam. *Navier-Stokes Equations: Theory and Numerical Analysis.* American Mathematical Society, Providence, 2001.

[77] T. E. Tezduyar and S. Sathe. Stabilization parameters in SUPG and PSPG formulations. *J. Comput. Appl. Mech.*, 4(1):71–88, 2003.

[78] M. Thiriet. *Biology and Mechanics of Blood Flows.* Springer, New York, 2008.

[79] G. B. Thurston. Viscoelasticity of Human Blood. *Biophys. J.*, 12(9):1205–1217, 1972.

[80] G. B. Thurston. Rheological parameters for the viscosity viscoelasticity and thixotropy of blood. *Biorheology*, 16(3):149–162, 1979.

[81] F. Tröltzsch. *Optimal control of partial differential equations. Theory, methods and applications.* American Mathematical Society, Providence, 2010.

[82] S. Turek, L. Rivkind, J. Hron, and R. Glowinski. Numerical study of a modified time-stepping θ-scheme for incompressible flow simulations. *J. Sci. Comput.*, 28(2–3):533–547, 2006.

[83] S. M. Wasserman and J. N. Topper. Adaptation of the endothelium to fluid flow: in vitro analyses of gene expression and in vivo implications. *Vasc. Med.*, 9(1):35–45, 2004.

[84] K. K. Yeleswarapu, M. V. Kameneva, K. R. Rajagopal, and J. F. Antaki. The flow of blood in tubes: theory and experiment. *Mech. Res. Commun.*, 25(3):257–262, 1998.

[85] W. Zulehner. Efficient solvers for saddle point problems with applications to PDE-constrained optimization. In T. Apel and O. Steinbach, editors, *Advanced finite element methods and applications*, volume 66 of *Lect. Notes Appl. Comput. Mech.*, pages 197–216. Springer, Heidelberg, 2013.

[86] W. Zulehner. The Ciarlet-Raviart Method for Biharmonic Problems on General Polygonal Domains: Mapping Properties and Preconditioning. *Reports of the Institute of Computational Mathematics*, Report 2013/7, University of Linz, 2013.

Monographic Series TU Graz

Computation in Engineering and Science

Vol. 1 Steffen Alvermann

Effective Viscoelastic Behaviour of Cellular Auxetic Materials
2008
ISBN 978-3-902465-92-4

Vol. 2 Sendy Fransiscus Tantono

The Mechanical Behaviour of a Soilbag under Vertical Compression
2008
ISBN 978-3-902465-97-9

Vol. 3 Thomas Rüberg

Non-conforming FEM/BEM Coupling in Time Domain
2008
ISBN 978-3-902465-98-6

Vol. 4 Dimitrios E. Kiousis

Biomechanical and Computational Modeling of Atherosclerotic Arteries
2008
ISBN 978-3-85125-023-7

Vol. 5 Lars Kielhorn

A Time-Domain Symmetric Galerkin BEM for Viscoelastodynamics
2009
ISBN 978-3-85125-042-8

Vol. 6 Gerhard Unger

Analysis of Boundary Element Methods for Laplacian Eigenvalue Problems
2009
ISBN 978-3-85125-081-7

Monographic Series TU Graz

Computation in Engineering and Science

Vol. 7 Gerhard Sommer

Mechanical Properties of Healthy and Diseased Human Arteries
2010
ISBN 978-3-85125-111-1

Vol. 8 Mathias Nenning

Infinite Elements for Elasto- and Poroelastodynamics
2010
ISBN 978-3-85125-130-2

Vol. 9 Thanh Xuan Phan

Boundary Element Methods for Boundary Control Problems
2011
ISBN 978-3-85125-149-4

Vol. 10 Loris Nagler

Simulation of Sound Transmission through Poroelastic Plate-like Structures
2011
ISBN 978-3-85125-153-1

Vol. 11 Markus Windisch

Boundary Element Tearing and Interconnecting Methods for Acoustic and Electromagnetic Scattering
2011
ISBN: 978-3-85125-152-4

Monographic Series TU Graz

Computation in Engineering and Science

Vol. 12 Christian Walchshofer

Analysis of the Dynamics at the Base of a Lifted Strongly Buoyant Jet Flame Using Direct Numerical Simulation
2011
ISBN 978-3-85125-185-2

Vol. 13 Matthias Messner

Fast Boundary Element Methods in Acoustics
2012
ISBN 978-3-85125-202-6

Vol. 14 Peter Urthaler

Analysis of Boundary Element Methods for Wave Propagation in Porous Media
2012
ISBN 978-3-85125-216-3

Vol. 15 Peng Li

Boundary Element Method for Wave Propagation in Partially Saturated Poroelastic Continua
2012
ISBN 978-3-85125-236-1

Vol. 16 Andreas J. Schriefl

Quantification of Collagen Fiber Morphologies in Human Arterial Walls
2012
ISBN 978-3-85125-238-5

Vol. 17 Thomas S. E. Eriksson

Cardiovascular Mechanics
2013
ISBN 978-3-85125-277-4

Monographic Series TU Graz

Computation in Engineering and Science

Vol. 18 Jianhua Tong

Biomechanics of Abdominal Aortic Aneurysms
2013
ISBN 978-3-85125-279-8

Vol. 19 Jonathan Rohleder

Titchmarsh–Weyl Theory and Inverse Problems for Elliptic Differential Operators
2013
ISBN 978-3-85125-283-5

Vol. 20 Martin Neumüller

Space-Time Methods
2013
ISBN 978-3-85125-290-3

Vol. 21 Michael J. Unterberger

Microstructurally-Motivated Constitutive Modeling of Cross-Linked Filamentous Actin Networks
2013
ISBN 978-3-85125-303-0

Vol. 22 Vladimir Lotoreichik

Singular Values and Trace Formulae for Resolvent Power Differences of Self-Adjoint Elliptic Operators
2013
ISBN 978-3-85125-304-7

Vol. 23 Michael Meßner

A Fast Multipole Galerkin Boundary Element Method for the Transient Heat Equation
2014
ISBN 978-3-85125-350-4

Monographic Series TU Graz

Computation in Engineering and Science

Vol. 24 Lorenz Johannes John

Optimal Boundary Control in Energy Spaces
2014
ISBN 978-3-85125-373-3