Technische Universität München
Zentrum Mathematik
Lehrstuhl für Optimalsteuerung

Optimal Control of Time-Dependent Nonlinear Fluid-Structure Interaction

Lukas Failer

Vollständiger Abdruck der von Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende:		Univ.-Prof. Dr. Christina Kuttler
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Boris Vexler
	2.	Univ.-Prof. Dr. Thomas Richter (Universität Magdeburg)
	3.	Univ.-Prof. Dr. Winnifried Wollner (Technische Universität Darmstadt)

Die Dissertation wurde am 11.04.2017 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 23.06.2017 angenommen.

Abstract

For various applications in aeroelasticity and hemodynamics optimization problems constrained by fluid-structure interaction models have to be solved. In this thesis, we consider two settings — a linear and a nonlinear fluid-structure interaction model. We analyze well-posedness for optimal control of a linear fluid-structure interaction problem. Then we derive necessary optimality conditions and prove regularity results for the optimal control variable. Hereafter, we formally establish necessary optimality conditions for the optimization with a nonlinear fluid-structure interaction model. The resulting optimality system is discretized with a Petrov-Galerkin discretization method in time. For spatial discretization, we use a stabilized conforming finite element method. This enables the use of a dual-weighted residual error estimator to approximate the space, time, and control discretization error with respect to the cost functional. Furthermore, we can compute exact sensitivity information due to the Galerkin discretization approach. This is necessary to obtain efficient optimization algorithms. The resulting adaptive algorithm is tested numerically for several exemplary optimization problems.

Zusammenfassung

Für verschiedenste Anwendungen in den Bereichen Aeroelastizität und Hämodynamik muss ein durch ein Fluid-Struktur-Model beschränktes Optimierungsproblem gelöst werden. In dieser Arbeit betrachten wir zwei Konfigurationen, ein lineares und ein nichtlineares Fluid-Struktur Model. Wir analysieren die Wohlgestelltheit des Optimalsteuerungsproblems für ein lineares Fluid-Struktur Wechselwirkungsmodel. Danach leiten wir notwendige Optimaliätsbedingungen her und beweisen weiterführende Regularitätsergebnisse für die optimale Kontrolle. Des Weiteren entwickeln wir formal notwendige Optimalitätsbedingungen für die Optimierung mit einem nichtlinearen Fluid-Struktur-Model. Das dabei entstehende Optimalitätssystem wird mithilfe eines Petrow-Galerkin-Verfahrens in der Zeit diskretisiert. Für die Ortsdiskretisierung verwenden wir eine stabilisierte Finite-Elemente-Methode. Dies ermöglicht die Anwendung von residuenbasierten Fehlerschätzern zur Approximation des Zeit-, Orts- und Kontrolldiskretisierungsfehlers bezüglich des Kostenfunktionals. Dies ist eine notwendige Voraussetzung für effiziente Optimierungsalgorithmen. Der adaptive Algorithmus wird für verschieden Optimierungsprobleme numerisch getestet.

Contents

1. Introduction **1**

2. Definitions **7**

3. Optimal Control of Linear Fluid-Structure Interaction **9**
3.1. A Linear Fluid-Structure Interaction Problem 9
3.1.1. Linear Wave Equation 10
3.1.2. Stokes Equations 10
3.1.3. Coupling and Boundary Conditions 11
3.1.4. Monolithic Formulation of Fluid-Structure Interaction 11
3.2. Existence Theory for Linear Fluid-Structure Interaction 12
3.2.1. Known Results from the Literature 13
3.2.2. Novel Symmetric Weak Formulation 15
3.2.3. How are the Coupling Conditions Fulfilled? 19
3.3. Optimal Control Problem 20
3.3.1. Existence and Uniqueness of Optimal Solutions 21
3.3.2. Adjoint Equations 22
3.3.3. Necessary Optimality Conditions 25

4. Optimal Control of Nonlinear Fluid-Structure Interaction **33**
4.1. A Nonlinear Fluid-Structure Interaction Problem 33
4.1.1. Nonlinear Elastodynamics 34
4.1.2. Navier-Stokes Equations on a Moving Domain 36
4.1.3. Mesh Motion Equation 39
4.1.4. Coupling and Boundary Conditions 40
4.1.5. Monolithic ALE-Formulation for Fluid-Structure Interaction 42
4.2. Existence Theory for Nonlinear Fluid-Structure Interaction 43
4.3. Optimal Control Problem 44
4.3.1. Discussion of Optimal Solutions 47
4.3.2. Adjoint Equations 47
4.3.3. Necessary Optimality conditions 50

5. Discretization **53**
5.1. Time Discretization of the State Equation 53
5.1.1. Theta Time-Stepping Scheme 54
5.1.2. Divergence Condition 58
5.1.3. Fractional-Step Theta Time-Stepping Scheme 58
5.2. Time Discretization of the Adjoint Equation 58
5.2.1. Theta Time-Stepping Scheme 58

5.2.2. Adjoint Divergence Condition . . . 61
5.3. Space Discretization . . . 61
5.3.1. Finite Element Spaces . . . 61
5.3.2. Local Projection Stabilization . . . 62
5.3.3. Space-Discretized State and Adjoint Equations . . . 63
5.4. Discretization of the Control Variable . . . 65
5.5. Discrete Control Problem . . . 65

6. Solution Algorithms **67**
6.1. Newton Algorithm . . . 68
6.2. Linear Solver . . . 69
6.3. Limited memory BFGS Algorithm . . . 70

7. Dual-Weighted Residual Error Estimator **75**
7.1. A Posteriori Error Estimation for Optimal Control . . . 76
7.2. A Posteriori Error Estimation for Optimal Control of the Linear FSI Problem . 80
7.3. Evaluation of the Error Estimators . . . 83
7.3.1. Higher-Order Reconstruction . . . 83
7.3.2. Localization of the Error Estimator . . . 86
7.4. Adaptive Algorithm . . . 87

8. Numerical Examples **89**
8.1. Optimal Control of a Linear Fluid-Structure Interaction Problem . . . 89
8.2. A posteriori Error Estimation for a Nonlinear FSI Problem . . . 95
8.2.1. FSI-2 and FSI-3 Benchmark Configuration . . . 95
8.2.2. A Comparison of Various Time-Stepping Schemes . . . 97
8.2.3. FSI-3 Benchmark with new Inflow . . . 98
8.2.4. FSI-2 Benchmark with new Inflow . . . 100
8.2.5. Flapping Test . . . 102
8.3. Optimal Control of a Nonlinear Fluid-Structure Interaction Problem . . . 108
8.3.1. Optimal Control of Flow in an Elastic Channel . . . 108
8.3.2. Optimal Control of the FSI-2 Benchmark Example . . . 112
8.3.3. Optimal Control of the Mesh Motion Equation . . . 115

9. Conclusion and Outlook **125**

A. Transformation of Sobolev Spaces **127**

B. Linearization of the Nonlinear FSI Problem **129**

Acknowledgements **133**

Bibliography **135**

1. Introduction

Fluid-structure interaction is still the most prominent example of a multiphysics PDE system. Possible applications range from aeroelasticity, over mechanical engineering, to computational medicine and medical engineering. For references about these applications, we exemplary refer to the books [41, 66, 70, 40, 14]. In these books the isothermal, incompressible Navier-Stokes equations are coupled with nonlinear elasticity. The former system is of parabolic type and describes fluid flow, whereas the latter equation serves as a solid description and is of hyperbolic character.

More and more applications leading to optimal control, shape-optimization, and parameter estimation of fluid-structure interaction (FSI) are regarded recently. Especially in hemodynamic applications — in order to get a deeper understanding of the development of vascular diseases — patient specific properties have to be incorporated into the models. Hereby, advice for an appropriate treatment can be given. For example, in [25, 26, 27, 51, 103, 122, 117] patient specific boundary conditions and vessel material parameters are determined to simulate arterial blood flow. Similar approaches using gradient information have been proposed in [51, 24, 123] to estimate Young's modulus of an arteria. In [108], the authors apply reduced basis methods for a shape-optimization problem in this context.

Parameter estimation in the context of hemodynamics gain more and more importance as computer tomography (CT) and magnetic resonance imaging (MRI) evolve rapidly. Nowadays, already very accurate measurements of the movement of the vessel wall are possible and even averaged flow profiles in blood vessels can be provided, see [2, 25, 98]. To incorporate the data in the vascular models, it is necessary to improve the available parameter estimation and optimal control algorithms for fluid-structure interaction applications. Especially since only few approaches take the sensitivity information of the full time-dependent nonlinear system into account. For example in [49, 109], adjoint equations are derived for one-dimensional FSI configurations and in [130] for a stationary FSI problem. This motivates one of the main topics of this thesis, the systematic derivation of sensitivity information for the unsteady nonlinear FSI model. It allows to incorporate a large variety of measurements into the models. In addition, gradient-based approaches can handle high-dimensional control spaces, which enables the handling of time-dependent and distributed parameters.

Fluid-structure interaction problems have been extensively studied from theoretical and numerical point of view in the last decade. Various results on existence and regularity have been published for linear fluid-structure interaction. For example, in [53, 54, 3, 5, 8, 9] a detailed analysis of the regularity of the needed initial data and right-hand side has been realized. For the full nonlinear fluid-structure interaction problem only few results are available. For a smooth geometry and a damped solid model, existence results can be found in [47, 76, 86, 87, 77]. At the same time, efficient gradient-based algorithms for solving optimization problems governed by elliptic, parabolic, or hyperbolic equations are developed

and deeply analyzed in the literature. All these algorithms are based on optimality systems containing appropriate adjoint equations, and building necessary optimality conditions for considered problems. However, especially in the context of optimal control problems for unsteady FSI problems, optimality systems based on rigorous analysis cannot be found in the literature. In [119], the authors formally derive necessary optimality conditions for an optimal control problem of a nonlinear time-dependent FSI configuration by using shape derivatives. In [151] Gateaux-differentiability of the control to state mapping can be proven for a stationary FSI configuration. Further results on optimal feedback control of FSI can be found in [39, 101, 102, 104], where corresponding Riccati equations are derived.

We try to fill this gap by analyzing a model optimal control problem governed by an unsteady linear fluid-structure interaction problem. The Stokes equations are coupled with a linear wave equation on a domain with fixed interface. We establish necessary optimality conditions and analyze the regularity of the optimal solution. Although such a linear system neglects several problem-relevant phenomena we believe that our results provide an important step towards tackling optimal control problems for nonlinear FSI models. Based on the results on optimal control of the linear FSI model, we can at least formally derive necessary optimality conditions for optimal control of the nonlinear FSI configuration.

One of the main issues in this analysis, as well as in the numerical solution of FSI problems, is the incorporation of coupling conditions between the equations describing the behavior of the fluid and of the structure, respectively. A correct treatment of such conditions for the adjoint system is indispensable for a precise description of the information transport across the interface between the fluid and the structure, and as a consequence, for an accurate evaluation of the derivatives required in gradient-based optimization algorithms. A prominent FSI-coupling technique is based on an interface-tracking method; namely the nowadays standard Arbitrary Lagrangian-Eulerian (ALE) technique [52, 83, 85, 120, 64, 138, 139]. Here, the flow equations are re-written in such a way that their coordinate system matches the Lagrangian framework used to describe the structure mechanics. The resulting formulation using variational-monolithic coupling in the reference configuration is outlined in [84, 132, 56, 23].

The fact that the coupling conditions are directly incorporated in the variational formulation enables a straightforward derivation of sensitivity information. If we apply the Lagrange formalism, we will not have to take the coupling conditions into account. In addition, as the moving domain is hidden in the ALE-transformation, it is possible to linearize the problem directly, without having to compute shape derivatives. Meanwhile, the derivation of sensitivity information for partitioned approaches, whereby the fluid and solid equations are solved one after the other, is still an open question.

Furthermore such a monolithic formulation enables the natural usage of Galerkin finite element discretizations in space and time. This is particularly advantageous for optimal control problems, since the two approaches optimize-then-discretize, i.e., the discretization of the optimality system from continuous level, and discretize-then-optimize, i.e., discretization of the state equation and subsequent construction of the optimality system on the discrete level, lead to the same discretization scheme; see, e.g., [18, 32]. The recent results of Meidner and Richter in [111, 112] provide a Galerkin formulation in time of second-order equivalent to the strongly A-stable fractional-step theta time-stepping scheme. We extend these results to the nonlinear fluid-structure interaction problem and to its adjoint equation as for the coupled

hyperbolic-parabolic system, the fractional-step theta time-stepping scheme, see [37], turned out to provide excellent properties.

In addition the Galerkin discretization enables the use of a posteriori error estimators. In recent years, much effort has been spent on spatial discretization of fluid-structure interaction and in particular on adaptivity in space for higher accuracy of the overall solution or more importantly the accuracy of certain quantities of interest. Several studies on a posteriori error estimators using the dual-weighted residual method [19, 20] can be found in [55, 62, 144, 128, 150]. In [112, 111], time step control of parabolic problems and the Navier-Stokes equations have been developed for a Galerkin interpretation of fractional-step theta scheme. In addition, we refer, e.g., to [113, 114, 115, 116] for a posteriori error estimation and adaptivity for parabolic optimal control problems discretized by a Galerkin approach.

To solve optimization problems subject to a FSI model, we have to compute the solution of the fluid-structure interaction problem several times. To reduce the computational cost, we extend the dual-weighted residual method in [19, 20] to control the spatial and time-discretization error as well as the control discretization error for optimal control of a monolithic fluid-structure interaction problem. Thereby, we are able to implement an adaptive algorithm, which enables to solve the optimal control problems in several numerical examples very fast and accurately.

The rest of the thesis is organized as follows:

Definitions

We state the used notation, the Sobolev spaces and inner products needed throughout the thesis.

Optimal Control of Linear Fluid-Structure Interaction

In this chapter, we derive existence and regularity results for optimal control of a linear fluid-structure interaction model. We shortly motivate the underlying equations in Section 3.1. In Section 3.2, we summarize the existing regularity results for linear FSI published in the last decade and extend the results to a symmetric monolithic formulation. This formulation leads to an adjoint equation with the same structure as the considered linear FSI problem, which allows for a unified analytic treatment of state and adjoint equation. Section 3.3 is devoted to the analysis of the optimal control problem. We discuss the existence of the optimal solution for two model configurations. Finally, we derive regularity results for the adjoint system, which is indispensable to ultimately prove the optimality system.

Optimal Control of Nonlinear Fluid-Structure Interaction

In Chapter 4, we consider optimal control of the nonlinear fluid-structure interaction problem. In Section 4.1, we systematically derive the equations for solid and fluid mechanics from continuum mechanics. The moving fluid domain is transformed via an ALE-mapping on a reference domain. Thereafter, we can incorporate the coupling conditions in a variational way

and obtain the fully-coupled monolithic formulation. In Section 4.2 we shortly summarize existing regularity results for nonlinear fluid-structure interaction. Finally, in Section 4.3 we apply the Lagrange formalism to derive the formal optimality system.

Discretization

In Section 5.1, we discretize the nonlinear fluid-structure interaction problem in time using a Galerkin approach, as suggested in [111, 112], resulting in a fractional-step theta time-stepping scheme. The same technique is used, in Section 5.2, to derive a suitable time-stepping scheme for the adjoint equation. In Section 5.3, we introduce a finite element space to discretize the state and adjoint equation in space. As we are using equal order elements for velocity and pressure, we have to stabilize the system with a local projection stabilization (LPS). Finally, we state different possibilities for the control discretization, in Section 5.4, and present the discretized optimality system, in Section 5.5.

Solution Algorithms

In Chapter 6, we describe the structure of the applied optimization algorithm. We state the Newton algorithm used to solve the nonlinear fluid-structure interaction problem in every time step, in Section 6.1, and comment on possible linear solvers in Section 6.2. The sensitivity information derived from solving an adjoint equation backward in time is used to compute a control update with a limited memory BFGS algorithm presented in Section 6.3.

Dual-Weighted Residual Error Estimator

The use of a Galerkin discretization, presented in Chapter 5, makes the the use of a dual-weighted residual error estimator possible. The main theorem is stated in Section 7.1, whereby the theta time-stepping scheme and the LPS stabilization have to be taken into account. For optimal control of the linear fluid-structure interaction problem with tracking type functional, we are able to prove, in Section 7.2, that the remainder terms in the a posteriori error estimator can be neglected. To evaluate the a posteriori error estimators, in Section 7.3, we present a higher order reconstruction, which can be easily localized if a patch structure is available. The localized error indicators enable the use of an adaptive algorithm given in Section 7.4.

Numerical Examples

In the final Chapter 8, we present various numerical examples. In Section 8.1, we apply the a posteriori error estimator to optimal control of a linear fluid-structure interaction problem. We test the effectivity of the spatial, time, and control discretization error estimators. In Section 8.2, we apply the a posteriori estimator to the nonlinear fluid-structure interaction problem without control. We compute a slightly modified variant of the FSI-2 and FSI-3 benchmarks proposed in [142, 41, 40] and compare the convergence in a functional of interest using global and adaptive refinement in time. Then, we consider a flapping example and

equilibrate the spatial and time-discretization error. Motivated by hemodynamical applications, in Section 8.3.1, we control a Neumann boundary condition to enforce the energy of the pulsatile inflow to leave the channel again. By controlling the mean pressure at the boundary of the FSI-2 benchmark, we are able to reduce the amplitude of the oscillating flag, in Section 8.3.2. The ALE approach has significant difficulties if the transformation gets too large. For such cases we control the mesh motion equation in Section 8.3.3, to smoothen the ALE transformation. Here, the control variable is a distributed material parameter in the mesh motion equation.

2. Definitions

We consider a domain $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$, which is separated into two disjoint Lipschitz sub-domains Ω_s and Ω_f with $\overline{\Omega} = \overline{\Omega}_\mathrm{s} \cup \overline{\Omega}_\mathrm{f}$, as presented in Figure 2.1. Furthermore $\Gamma := \partial\Omega$ denotes the outer boundary, which is split into two parts, the solid outer boundary $\Gamma_\mathrm{s} := \Gamma \cap \overline{\Omega}_\mathrm{s}$ and the fluid outer boundary $\Gamma_\mathrm{f} := \Gamma \cap \overline{\Omega}_\mathrm{f}$, where we assume that Γ_s has positive measure in Γ. The common interface between the fluid domain Ω_f and the solid domain Ω_s is denoted by $\Gamma_\mathrm{i} := \overline{\Omega}_\mathrm{f} \cap \overline{\Omega}_\mathrm{s}$. Moreover, n_s is the unit outward normal vector on Γ_i with respect to the region Ω_s and $n_\mathrm{f} = -n_\mathrm{s}$ is the unit outward normal vector with respect to the region Ω_s. We introduce, in addition, the time interval $I := (0, T)$ with end time point T.

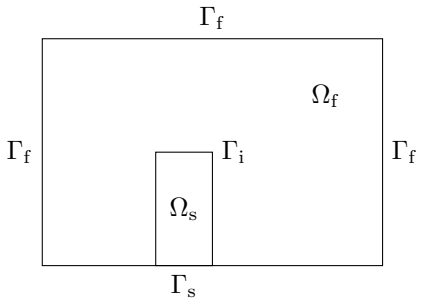

Figure 2.1.: An exemplary reference domain Ω

For several of the presented models, in this thesis, neither existence nor regularity results are available. We will nevertheless state very precise test and solution spaces in such a way that the weak formulations of the models are well defined. To be on the safe side, we assume rather more regularity then necessary. The presented models are going to be defined either in 2D or 3D. Thus, we introduce the vector valued Lebesgue spaces

$$H := L^2(\Omega)^d, \quad H_\mathrm{f} := L^2(\Omega_\mathrm{f})^d, \quad \text{and} \quad H_\mathrm{s} := L^2(\Omega_\mathrm{s})^d.$$

On the domain Ω_f we either use the Stokes equations or the Navier-Stokes equations to model fluid flow. The solution of such equations of parabolic type will be an element of the $W[0, T]$ space defined by

$$V_\mathrm{f} := \left\{ \varphi \in H^1(\Omega_\mathrm{f})^d \;\middle|\; \varphi = 0 \text{ on } \Gamma_\mathrm{f} \right\} \quad \text{and} \quad W_\mathrm{f}^v := \left\{ \varphi \;\middle|\; \varphi \in L^2(I; V_\mathrm{f}) \text{ and } \partial_t \varphi \in L^2(I; V_\mathrm{f}^*) \right\}$$

with trace zero on parts of the boundary and V_f^* the dual space of V_f.

For the pressure variable we introduce the following Lebesgue spaces

$$L_\mathrm{f} := \left\{ p \in L^2(\Omega_\mathrm{f}) \;\middle|\; (p, 1)_\mathrm{f} = 0 \right\}.$$

The elastic structure on the domain Ω_s can either be modeled by a linear Lamé system or by a nonlinear elastodynamics model. As these models are of hyperbolic type the solution will be an element of the Sobolev spaces:

$$V_\mathrm{s} := \left\{ \varphi \in H^1(\Omega_\mathrm{s})^d \;\middle|\; \varphi = 0 \text{ on } \Gamma_\mathrm{s} \right\}, \quad W_\mathrm{s}^u := \left\{ \varphi \;\middle|\; \varphi \in L^2(I; V_\mathrm{s}) \text{ and } \partial_t \varphi \in L^2(I; H_\mathrm{s}) \right\},$$
$$\text{and} \quad W_\mathrm{s}^v := \left\{ \varphi \;\middle|\; \varphi \in L^2(I; H_\mathrm{s}) \text{ and } \partial_t \varphi \in L^2(I; V_\mathrm{s}^*) \right\}.$$

The fluid domain is going to move with the elastic structure. Therefore we have to solve an auxiliary problem to define an extension operator. The solution of the so called mesh motion equation is an element of the Sobolev space

$$V_{\mathrm{f},0} := \left\{ \varphi \in H^1(\Omega_\mathrm{f})^d \;\middle|\; \varphi = 0 \text{ on } \Gamma_\mathrm{f} \cup \Gamma_\mathrm{i} \right\}$$

with trace zero on all boundaries of Ω_f.

As we follow a monolithic approach we embed the coupling conditions between the fluid model on Ω_f and the solid model on Ω_s in the Sobolev spaces. Thus we have to introduce on the whole domain Ω the spaces

$$V := \left\{ \varphi \in H^1(\Omega)^d \;\middle|\; \varphi = 0 \text{ on } \Gamma_\mathrm{f} \cup \Gamma_\mathrm{s} \right\}, \quad W^v := \left\{ \varphi \;\middle|\; \varphi \in L^2(I; V) \text{ and } \partial_t \varphi \in L^2(I; V^*) \right\},$$
$$\text{and} \quad W^u := \left\{ \varphi \;\middle|\; \varphi \in L^2(I; V) \text{ and } \partial_t \varphi \in L^2(I; H) \right\}.$$

To keep the notation as compact as possible here and in what follows, let

$$(u, v) := (u, v)_\Omega, \qquad \langle u, v \rangle_\mathrm{i} := \langle u, v \rangle_{\Gamma_\mathrm{i}},$$
$$(u, v)_\mathrm{f} := (u, v)_{\Omega_\mathrm{f}}, \qquad (u, v)_\mathrm{s} := (u, v)_{\Omega_\mathrm{s}},$$

be the L^2 inner product on Ω, its sub-domains Ω_f and Ω_s, and on Γ_i. Furthermore, we use the following notation for inner products on the space-time cylinder:

$$(\!(u, v)\!) = \int_0^T (u, v)_\Omega \, \mathrm{d}t, \qquad (\!\langle u, v \rangle\!)_\mathrm{i} = \int_0^T \langle u, v \rangle_\mathrm{i} \, \mathrm{d}t,$$
$$(\!(u, v)\!)_\mathrm{f} = \int_0^T (u, v)_{\Omega_\mathrm{f}} \, \mathrm{d}t, \qquad (\!(u, v)\!)_\mathrm{s} = \int_0^T (u, v)_{\Omega_\mathrm{s}} \, \mathrm{d}t.$$

3. Optimal Control of Linear Fluid-Structure Interaction

In this section, we will analyze optimal control of a tracking type functional subject to a linear fluid-structure interaction problem. On Ω_f, the fluid is described via the Stokes equations and the structure on Ω_s behaves according to the linear wave equation. We note that all the results presented in this chapter can be extended to the model problem, where the wave equation is replaced by the linear Lamé system. The control q is going to be either time-dependent or distributed in space and controlling the volume force.

Later, in Chapter 4 we are going to present a nonlinear FSI model, in which we couple the Navier-Stokes equations with a nonlinear hyperbolic equation. As the solid motion involves large stress-induced displacements, the fluid domain is not stationary. However, the energy between the fluid and structure system is mainly transported by forces and not due to the motion of the fluid domain. Therefore, the same difficulties occurring due to the coupling already appear in a linear FSI problem, where we couple a linear wave equation with a Stokes equation on a domain with fixed interface.

In Section 3.1, we first derive a monolithic formulation of the linear fluid-structure interaction problem. In Section 3.2, we take a closer look at existence and regularity results for the presented problem provided by literature and extend the results to a symmetric monolithic formulation. This formulation leads to an adjoint equation with the same structure as the considered linear FSI problem, which allows for a unified analytical and numerical treatment of the state and the adjoint systems. Then, in Section 3.3, we formulate a model optimal control problem governed by an unsteady linear FSI problem, establish necessary optimality conditions, and analyze the regularity of the optimal solutions.

Most of the results in this chapter have already appeared in similar form in [58].

3.1. A Linear Fluid-Structure Interaction Problem

In this chapter, the domain Ω is separated in two disjoint domains Ω_s and Ω_f as for example in Figure 2.1. The mechanics on the solid domain Ω_s are described by a wave equation and the fluid mechanics on Ω_f by the Stokes equations. We first state the two problems, define the coupling conditions between those two models, and then derive a monolithic formulation.

3.1.1. Linear Wave Equation

In Section 4.1.1, we will derive the nonlinear elastodynamics equations to model elastic materials. For small displacements u_s with $\|\nabla u_\mathrm{s}\| \ll 1$, the linearized elastodynamics equations simplify to the linear Lamé system. To make the following sections easier to read, we restrict to the special case of a linear wave equation. However, all results presented in this chapter can be extended to the model problem, where the wave equation is replaced by the linear Lamé system.

We have already rewritten the linear wave equation on Ω_s in a system of first order in time. The variables u_s denotes the structure displacement in Ω_s and v_s the velocity. If we define the bi-linear forms

$$a^S(u_\mathrm{s}, v_\mathrm{s})(\varphi) := (\!(\rho_\mathrm{s}\partial_t v_\mathrm{s}, \varphi)\!)_\mathrm{s} + (\!(\mu_\mathrm{s}\nabla u_\mathrm{s}, \nabla\varphi)\!)_\mathrm{s}, \tag{3.1}$$

$$a^V(u_\mathrm{s}, v_\mathrm{s})(\psi) := (\!(\partial_t u_\mathrm{s}, \psi)\!)_\mathrm{s} - (\!(v_\mathrm{s}, \psi)\!)_\mathrm{s}, \tag{3.2}$$

then the linear wave equation with the material parameters ρ_s and μ_s is given by

Problem 3.1 (Linear wave equation)**.** *For $u_{\mathrm{s},0} \in V_\mathrm{s}$, $v_{\mathrm{s},0} \in H_\mathrm{s}$, $g_\mathrm{s} \in L^2(I; H^{\frac{1}{2}}(\Gamma_\mathrm{i})^d)$ and the volume force $f_\mathrm{s} \in L^2(I; H_\mathrm{s})$ find a solution $(u_\mathrm{s}, v_\mathrm{s}) \in W_\mathrm{s}^u \times W_\mathrm{s}^v$ satisfying the initial conditions $(u_\mathrm{s}(0), v_\mathrm{s}(0)) = (u_{\mathrm{s},0}, v_{\mathrm{s},0})$ and*

$$\begin{aligned} a^S(u_\mathrm{s}, v_\mathrm{s})(\varphi) &= (\!(f_\mathrm{s}, \varphi)\!)_\mathrm{s} + \langle\!(g_\mathrm{s}, \varphi)\!\rangle_\mathrm{i} && \forall \varphi \in L^2(I; V_\mathrm{s}), \\ a^V(u_\mathrm{s}, v_\mathrm{s})(\psi) &= 0 && \forall \psi \in L^2(I; H_\mathrm{s}). \end{aligned} \tag{3.3}$$

Existence and uniqueness of such a linear wave are given in [106] and [100]. Further results on the wave equation with non homogenous boundary values can be found in [99].

3.1.2. Stokes Equations

A large class of fluids can be described by the Navier-Stokes equations, which we are going to derive later in Section 4.1.2. In some configurations, especially for low Reynolds numbers, it is possible to neglect the transport term in order to acquire the Stokes equations. In this section, we describe the fluid on Ω_f via such a linear model. The variables v_f and p_f denote the velocity and pressure of the fluid in Ω_f. We define the bi-linear forms

$$\begin{aligned} a^F(v_\mathrm{f}, p)(\varphi) &:= (\!(\rho_\mathrm{f}\partial_t v_\mathrm{f}, \varphi)\!)_\mathrm{f} + (\!(\nu_\mathrm{f}\nabla v_\mathrm{f}, \nabla\varphi)\!)_\mathrm{f} - (\!(p_\mathrm{f}, \operatorname{div}\varphi)\!)_\mathrm{f}, \\ a^D(v_\mathrm{f})(\xi) &:= (\!(\operatorname{div} v_\mathrm{f}, \xi)\!)_\mathrm{f}. \end{aligned} \tag{3.4}$$

Then, we obtain for the Stokes problem with material parameters ρ_f and ν_f:

Problem 3.2 (Stokes equations)**.** *For $v_{\mathrm{f},0} \in H_\mathrm{f}$, $g_\mathrm{f} \in L^2(I; L^2(\Gamma_\mathrm{i})^d)$ and $f_\mathrm{f} \in L^2(I; H_\mathrm{f})$ find a solution $(v_\mathrm{f}, p_\mathrm{f}) \in W_\mathrm{f}^v \times L^2(I; L_\mathrm{f})$ satisfying the initial conditions $v_\mathrm{f}(0) = v_{\mathrm{f},0}$ and*

$$\begin{aligned} a^F(v_\mathrm{f}, p_\mathrm{f})(\varphi) &= (\!(f_\mathrm{f}, \varphi)\!)_\mathrm{f} + \langle\!(g_\mathrm{f}, \varphi)\!\rangle_\mathrm{i} && \forall \varphi \in L^2(I; V_\mathrm{f}), \\ a^D(v_\mathrm{f})(\xi) &= 0 && \forall \xi \in L^2(I; L_\mathrm{f}). \end{aligned} \tag{3.5}$$

A very detailed analysis of the Stokes equations can be found in [137] and [69]. Therein, the authors proof existence of the Stokes equations for Dirichlet conditions. The results can be directly extended to inhomogeneous Neumann boundary conditions.

3.1.3. Coupling and Boundary Conditions

At the interface Γ_i, the momentum has to be conserved. Therefore, we demand the directional derivatives to coincide on $\Gamma_\mathrm{i} \times I$ (dynamic coupling condition):

$$\nu_\mathrm{f} \nabla v_\mathrm{f}\, n_\mathrm{f} - p_\mathrm{f} n_\mathrm{f} + \mu_\mathrm{s} \nabla u_\mathrm{s}\, n_\mathrm{s} = 0 \qquad \text{on } \Gamma_\mathrm{i} \times I. \tag{3.6}$$

Additionally, the fluid is not allowed to enter the structure domain (kinematic coupling condition). For slow fluid flow, we can assume a no slip condition, which implies that structure and fluid velocity have to be equal. Therefore, we demand

$$v_\mathrm{f} = v_\mathrm{s} \qquad \text{on } \Gamma_\mathrm{i} \times I \tag{3.7}$$

on the interface. At the outer boundaries Γ_f and Γ_s, we prescribe homogeneous Dirichlet boundary conditions.

$$\begin{aligned} v_\mathrm{f} &= 0 \qquad \text{on } \Gamma_\mathrm{f} \times I, \\ u_\mathrm{s} &= 0 \qquad \text{on } \Gamma_\mathrm{s} \times I. \end{aligned} \tag{3.8}$$

3.1.4. Monolithic Formulation of Fluid-Structure Interaction

Monolithic formulations are usually obtained by transforming a weak formulation of the FSI equations into a system of first order in time by introducing a structure velocity variable. Then, the kinematic coupling condition is enforced by choosing a smooth trial space for the common velocity variable defined on the whole domain. Furthermore, due to a test function defined in the same space, the dynamic coupling condition is automatically fulfilled. Thereby the coupling conditions are directly incorporated in the variational formulation. This simplifies the computation of sensitivities and allows for a natural usage of Galerkin finite element discretizations in space and time. Then the two approaches optimize-then-discretize and discretize-then-optimize lead to the same discretization scheme; see, e.g., [18]. As the velocity variable v now live on the whole domain, and as it is obvious that we mean with p the fluid pressure and with u the solid displacement, we neglect the indices f and s in the following.

For optimal control, the semi-linear form $a^V(\cdot)(\cdot)$ in (3.2) will result in an asymmetric formulation, see Section 3.3.2 for details. Because of this, we favor a slightly different formulation, which is motivated by an approach used by Johnson in [90] in the context of the wave equation. He suggests to introduce a velocity variable v, which fulfills

$$a^{\mathcal{V}}(u, v)(\psi) := \mu_\mathrm{s} (\!(\nabla v, \nabla \psi)\!)_\mathrm{s} - \mu_\mathrm{s} (\!(\nabla \partial_t u, \nabla \psi)\!)_\mathrm{s} = 0 \quad \forall \psi \in L^2(I; V_\mathrm{s}) \tag{3.9}$$

instead of (3.2). In the case of the here considered linear FSI problem, this leads to the self-adjoint symmetric monolithic formulation

Problem 3.3 (Monolithic linear fluid-structure interaction)**.** *Find a velocity $v \in W^v$, a pressure $p \in L^2(I; L_\mathrm{f})$, and a structure displacement $u \in W_\mathrm{s}^u$ fulfilling the initial conditions*

$$u(0) = u_0, \quad v(0)\big|_{\Omega_\mathrm{s}} = u_1, \quad \textit{and} \quad v(0)\big|_{\Omega_\mathrm{f}} = v_0 \tag{3.10}$$

and the weak formulation

$$a^F(v,p)(\varphi) + a^S(u,v)(\varphi) = (\!(f_\mathrm{f},\varphi)\!)_\mathrm{f} + (\!(f_\mathrm{s},\varphi)\!)_\mathrm{s} \qquad \forall \varphi \in L^2(I;V), \tag{3.11}$$

$$a^{\mathcal{V}}(u,v)(\psi) = 0 \qquad \forall \psi \in L^2(I;H_\mathrm{s}), \tag{3.12}$$

$$a^D(v)(\xi) = 0 \qquad \forall \xi \in L^2(I;L_\mathrm{f}). \tag{3.13}$$

The velocity v now describes the fluid velocity on Ω_f and the velocity of the structure on Ω_s. The wave equation would be well defined for a velocity $v \in L^2(I;H_\mathrm{s})$. However, to have the kinematic coupling condition fulfilled in a trace sense, we demand $v \in L^2(I;V)$. For smooth initial data and smooth right-hand side we prove existence and uniqueness of a solution for Problem 3.3 in Theorem 3.3.

We can write the symmetric weak form of the linear fluid-structure interaction problem in a very compact way by introducing the common solution variable $\mathbf{u} := (v,u,p)$ the test function $\boldsymbol{\varphi} := (\varphi,\psi,\xi)$ and the bi-linear form

$$a(\mathbf{u})(\boldsymbol{\varphi}) := a^F(v,p)(\varphi) + a^S(u,v)(\varphi) + a^{\mathcal{V}}(u,v)(\psi) + a^D(v)(\xi). \tag{3.14}$$

If we define in addition the trial space $X := W^v \times W^u_\mathrm{s} \times L^2(I;L_\mathrm{f})$ as well as the test space $Y := L^2(I;V) \times L^2(I;V_\mathrm{s}) \times L^2(I;L_\mathrm{f})$, then the monolithic formulation in Problem 3.3 reduces to :

Find $\mathbf{u} \in X$ such that $\mathbf{u}(0) = \mathbf{u}_0$ and

$$a(\mathbf{u})(\boldsymbol{\varphi}) = (\!(f_\mathrm{f},\varphi)\!)_\mathrm{f} + (\!(f_\mathrm{s},\varphi)\!)_\mathrm{s} \quad \forall \boldsymbol{\varphi} \in Y \tag{3.15}$$

Later, in Section 4 we will use the same notation for the semi-linear forms, test variables and test and trial spaces to state the nonlinear fluid-structure interaction problem. In doing so the similarity of the structure of linear and nonlinear problems immediately gets obvious. In addition we can describe the discretization of the nonlinear problem in Section 5. The abuse of notation makes it possible for the reader to immediately extend the results to the linear problem.

3.2. Existence Theory for Linear Fluid-Structure Interaction

Linear FSI configurations have been already analyzed in [53, 54, 3, 5, 8, 9], wherein the authors prove existence and regularity results. By introducing a damping term in the wave equation or in the coupling condition, uniform stability results are shown in [3, 9, 6, 7, 4, 87, 153] independently of the geometry. Therefore, we can build up on a vast number of already established results. That is not going to be the case for the nonlinear FSI problem as we see in Section 4.2. In the following we are going to briefly summarize the existing results in literature and then extend them to the symmetric monolithic formulation. Finally, we are going to analyze in which sense the coupling conditions are fulfilled.

Throughout this chapter, we assume the following properties of the initial data:

Assumption 1. The initial data u_0, u_1, and v_0 satisfy

$$u_0 \in V_\mathrm{s},\ \Delta u_0 \in H_\mathrm{s},\ u_1 \in V_\mathrm{s},\ v_0 \in V_\mathrm{f},\ \operatorname{div} v_0 = 0,\ \Delta v_0 \in H_\mathrm{f} \text{ and } v_0\big|_{\Gamma_\mathrm{i}} = u_1\big|_{\Gamma_\mathrm{i}}.$$

Furthermore, there exists a $p_0 \in H^1(\Omega_\mathrm{f})$ such that

$$(p_0 n_\mathrm{f} - \nu_\mathrm{f} \nabla v_0^T n_\mathrm{f})\big|_{\Gamma_\mathrm{i}} = (\mu_\mathrm{s} \nabla u_0^T n_\mathrm{s})\big|_{\Gamma_\mathrm{i}}.$$

For brevity, we introduce

$$\begin{aligned} A_0^2 := \|u_0\|^2_{H^1(\Omega_\mathrm{s})} + \|\Delta u_0\|^2_{L^2(\Omega_\mathrm{s})} + \|u_1\|^2_{H^1(\Omega_\mathrm{s})} + \|v_0\|^2_{H^1(\Omega_\mathrm{f})} \\ + \|\Delta v_0\|^2_{L^2(\Omega_\mathrm{f})} + \|p_0\|^2_{H^1(\Omega_\mathrm{f})}. \end{aligned} \tag{3.16}$$

The previously defined assumption on the initial data is not needed to guarantee existence of solutions for the linear fluid-structure interaction problem. Existence and regularity results for less regular initial data can be found, e.g., in [3, 5, 8, 9]. However, we assume more regular initial data, and the thereby implied higher regularity of solutions, to derive in Section 3.3.3 an optimality system for the considered optimal control problem.

3.2.1. Known Results from the Literature

The linear fluid-structure interaction problem was intensively studied in [53] and [54] by Du, Gunzburger, and coworkers. The following proposition, taken from there, ensures existence and uniqueness of a solution admitting optimal regularity with respect to the assumptions on the right-hand sides:

Proposition 3.1 (Theorems 3.2 and 3.4 in [53] and Theorem 2.3 in [54])**.** *Let u_0, u_1, and v_0 satisfy Assumption 1 and let $f_\mathrm{f} \in H^1(I; V_\mathrm{f}^*)$ and $f_\mathrm{s} \in H^1(I; H_\mathrm{s})$. Then, there exists a unique triplet (v, u, p) with*

$$v \in H^1(I; V_\mathrm{f}) \cap W^{1,\infty}(I; H_\mathrm{f}), \quad u \in W^{1,\infty}(I; V_\mathrm{s}) \cap W^{2,\infty}(I; H_\mathrm{s}), \quad p \in L^2(I; L_\mathrm{f})$$

satisfying the initial conditions $v(0) = v_0$ in H_f, $u(0) = u_0$ in V_s, and $\partial_t u(0) = u_1$ in H_s, as well the coupling condition

$$v\big|_{\Gamma_\mathrm{i}} = \partial_t u\big|_{\Gamma_\mathrm{i}} \text{ in } L^2(I; H^{\frac{1}{2}}(\Gamma_\mathrm{i})^d),$$

and almost everywhere in I

$$\begin{aligned} \rho_\mathrm{f}(\partial_t v, \varphi)_\mathrm{f} + \nu_\mathrm{f}(\nabla v, \nabla \varphi)_\mathrm{f} - (p, \operatorname{div} \varphi)_\mathrm{f} & \\ + \rho_\mathrm{s}(\partial_{tt} u, \varphi)_\mathrm{s} + \mu_\mathrm{s}(\nabla u, \nabla \varphi)_\mathrm{s} = (f_\mathrm{f}, \varphi)_\mathrm{f} + (f_\mathrm{s}, \varphi)_\mathrm{s} & \qquad \forall \varphi \in V, \\ (\xi, \operatorname{div} v)_\mathrm{f} = 0 & \qquad \forall \xi \in L_\mathrm{f}. \end{aligned}$$

Furthermore, the solution fulfills the a priori estimates

a)
$$\begin{aligned} \|v\|^2_{L^2(I;H^1(\Omega_\mathrm{f}))} + \|v\|^2_{L^\infty(I;L^2(\Omega_\mathrm{f}))} + \|u\|^2_{L^\infty(I;H^1(\Omega_\mathrm{s}))} + \|\partial_t u\|^2_{L^\infty(I;L^2(\Omega_\mathrm{s}))} \\ \leq C\big[\|f_\mathrm{f}\|^2_{L^2(I;V_\mathrm{f}^*)} + \|f_\mathrm{s}\|^2_{L^2(I;L^2(\Omega_\mathrm{s}))} + A_0^2\big], \end{aligned}$$

b)
$$\|\partial_t v\|^2_{L^2(I;H^1(\Omega_\mathrm{f}))} + \|\partial_t v\|^2_{L^\infty(I;L^2(\Omega_\mathrm{f}))} + \|\partial_t u\|^2_{L^\infty(I;H^1(\Omega_\mathrm{s}))} + \|\partial_{tt} u\|^2_{L^\infty(I;L^2(\Omega_\mathrm{s}))} \\ \leq C\big[\|f_\mathrm{f}\|^2_{H^1(I;V_\mathrm{f}^*)} + \|f_\mathrm{s}\|^2_{H^1(I;L^2(\Omega_\mathrm{s}))} + A_0^2\big],$$

c)
$$\|p\|^2_{L^2(I;L^2(\Omega_\mathrm{f}))} \leq C\big[\|f_\mathrm{f}\|^2_{H^1(I;V_\mathrm{f}^*)} + \|f_\mathrm{s}\|^2_{H^1(I;L^2(\Omega_\mathrm{s}))} + A_0^2\big],$$

where A_0^2 is defined in (3.16).

Remark 3.1. Under similar assumptions on the initial data, the stated result on existence, uniqueness, and regularity of solutions has also been proven in [3, 5, 8, 9] by Avalos and Triggiani for an elastic solid completely surrounded by a fluid.

Remark 3.2. Clearly, the solution given by Proposition 3.1 also fulfills the weak space-time formulation

$$\begin{aligned} \rho_\mathrm{f}(\!(\partial_t v, \varphi)\!)_\mathrm{f} + \nu_\mathrm{f}(\!(\nabla v, \nabla\varphi)\!)_\mathrm{f} - (\!(p, \operatorname{div}\varphi)\!)_\mathrm{f} & \\ +\rho_\mathrm{s}(\!(\partial_{tt} u, \varphi)\!)_\mathrm{s} + \mu_\mathrm{s}(\!(\nabla u, \nabla\varphi)\!)_\mathrm{s} = (\!(f_\mathrm{f}, \varphi)\!)_\mathrm{f} + (\!(f_\mathrm{s}, \varphi)\!)_\mathrm{s} & \qquad \forall \varphi \in L^2(I;V), \\ (\!(\xi, \operatorname{div} v)\!)_\mathrm{f} = 0 & \qquad \forall \xi \in L^2(I;L_\mathrm{f}). \end{aligned} \tag{3.17}$$

Later, we are going to control the system by right-hand sides, which are just in L^2 with respect to time. To do so, we will need the existence of solutions even if the right-hand side has less regularity than assumed in Proposition 3.1. The existence of such a solution operator is guaranteed by the results presented in [53, 54, 3, 5, 8, 9]. Therein the existence of a pressure variable can not be guaranteed. Thus, the fluid velocity variable is going to be an element of the divergence-free space

$$V_{\mathrm{f,div}} := \{\, v \in V_\mathrm{f} : \operatorname{div} v = 0 \text{ on } \Omega_\mathrm{f} \,\},$$

and we denote with

$$V_\mathrm{div} := \{\, v \in V : \operatorname{div} v = 0 \text{ on } \Omega_\mathrm{f} \,\}$$

the solution space on the whole domain Ω with divergence-free elements on the sub-domain Ω_f. Here, we will make use of the following result given in [53, 54]:

Proposition 3.2 (Theorem 2.5 in [53] and Theorem 2.2 in [54])**.** *Let u_0, u_1, and v_0 satisfy Assumption 1 and let $f_\mathrm{f} \in L^2(I;V_\mathrm{f}^*)$ and $f_\mathrm{s} \in L^2(I;H_\mathrm{s})$. Then, there exists a unique solution $\tilde{v} \in H^1(I;V_\mathrm{div}^*)$ with*

$$v = \tilde{v}\big|_{\Omega_\mathrm{f}} \in L^2(I;V_{\mathrm{f,div}}) \cap L^\infty(I;H_\mathrm{f})$$
$$\text{and} \quad u = \int_0^t \tilde{v}(s)\big|_{\Omega_\mathrm{s}}\, ds + u_0 \in L^\infty(I;V_\mathrm{s}) \cap W^{1,\infty}(I;H_\mathrm{s})$$

satisfying the initial conditions $v(0) = v_0$ in H_f, $u(0) = u_0$ in V_s, and $\partial_t u(0) = u_1$ in H_s, as well as the coupling condition

$$\int_0^t v(s)\big|_{\Gamma_\mathrm{i}}\, ds = u(t)\big|_{\Gamma_\mathrm{i}} - u_0\big|_{\Gamma_\mathrm{i}} \text{ in } L^2(I;H^{\frac{1}{2}}(\Gamma_\mathrm{i})^d)$$

and almost everywhere in I

$$\rho_\mathrm{f}(\partial_t\tilde{v},\varphi)_\mathrm{f} + \nu_\mathrm{f}(\nabla v,\nabla\varphi)_\mathrm{f} + \rho_\mathrm{s}(\partial_t\tilde{v},\varphi)_\mathrm{s} + \mu_\mathrm{s}(\nabla u,\nabla\varphi)_\mathrm{s} = (f_\mathrm{f},\varphi)_\mathrm{f} + (f_\mathrm{s},\varphi)_\mathrm{s} \quad \forall\varphi \in V_\mathrm{div}.$$

Furthermore, the solution (u, v) fulfills the a priori estimate a) in Proposition 3.1.

Remark 3.3. Clearly, the solution given by Proposition 3.2 fulfills also the space-time weak formulation

$$\rho_{\mathrm{f}}(\!(\partial_t \tilde{v}, \varphi)\!)_{\mathrm{f}} + \nu_{\mathrm{f}}(\!(\nabla v, \nabla \varphi)\!)_{\mathrm{f}} + \rho_{\mathrm{s}}(\!(\partial_t \tilde{v}, \varphi)\!)_{\mathrm{s}} + \mu_{\mathrm{s}}(\!(\nabla u, \nabla \varphi)\!)_{\mathrm{s}} \\ = (\!(f_{\mathrm{f}}, \varphi)\!)_{\mathrm{f}} + (\!(f_{\mathrm{s}}, \varphi)\!)_{\mathrm{s}} \quad \forall \varphi \in L^2(I; V_{\mathrm{div}}), \quad (3.18)$$

and for right-hand sides fulfilling the assumptions of Proposition 3.1, it coincides with the solution given there.

In [53, 54], the authors proved Proposition 3.2 under weaker assumptions on the initial data than in Assumption 1 at hand. For even weaker assumptions on the initial data, the results in [3, 5, 8, 9] show existence of a unique solution for linear FSI problems. The authors use semigroup theory and a new technique to derive a pressure-free formulation. Thereby, they were able to prove existence of a mild solution and optimal regularity results. As the goal of this thesis is to present an optimality system, which can be discretized by using a Galerkin finite element discretization in space and time, we require the solution to fulfill a weak formulation with explicit pressure and including the coupling conditions. Hence, we will from now on make use of initial data fulfilling Assumption 1, which ensures a regular solution as far as the right-hand sides fulfill the assumptions in Proposition 3.1.

Remark 3.4. In [53, 54], the authors demand $f_{\mathrm{f}} \in L^2(I; H_{\mathrm{f}})$ in Proposition 3.2 and furthermore $f_{\mathrm{f}} \in H^1(I; H_{\mathrm{f}})$, $u_0 \in H^2(\Omega_{\mathrm{s}})^d$, $v_0 \in H^2(\Omega_{\mathrm{f}})^d$ in Proposition 3.1. However, the proofs can directly be extended to $f_{\mathrm{f}} \in L^2(I; V_{\mathrm{f}}^*)$ and $f_{\mathrm{f}} \in H^1(I; V_{\mathrm{f}}^*)$ with initial data fulfilling Assumption 1 as stated above.

The results in [53, 54] and [8] are more general and also apply to Stokes flow coupled with linear elasticity equations. Therefore, all the results presented in the following are also extendable to formulations with stress tensors.

3.2.2. Novel Symmetric Weak Formulation

Now, we take a closer look to the symmetric weak formulation in Problem 3.3. The weak form consists of the bi-linear forms $a^S(\cdot)(\cdot)$ defined in (3.1), $a^{\mathcal{V}}(\cdot)(\cdot)$ defined in (3.9), as well as the bi-linear forms $a^F(\cdot)(\cdot)$ and $a^D(\cdot)(\cdot)$ defined in (3.4). The results in Section 3.2.1 enable us to prove existence and uniqueness of a solution of Problem 3.3. Note that the volume force g, appearing on the right-hand side of the equation introducing the structure velocity, has no physical interpretation but will occur later in the adjoint equation, see Section 3.3.

Theorem 3.3. *Let u_0, u_1, and v_0 satisfy Assumption 1 and let the right-hand sides fulfill $f_{\mathrm{f}} \in H^1(I; V_{\mathrm{f}}^*)$, $f_{\mathrm{s}} \in H^1(I; H_{\mathrm{s}})$, and $g \in L^2(I; H_{\mathrm{s}})$. Then, there exists a unique triplet (v, u, p) with*

$$v \in L^2(I; V) \cap W^{1,\infty}(I; H), \quad v\big|_{\Omega_{\mathrm{f}}} \in H^1(I; V_{\mathrm{f}}), \quad v\big|_{\Omega_{\mathrm{s}}} \in L^\infty(I; V_{\mathrm{s}}), \\ u \in L^\infty(I; V_{\mathrm{s}}) \cap H^1(I; V_{\mathrm{s}}), \quad p \in L^2(I; L_{\mathrm{f}}),$$

which satisfies the initial conditions

$$u(0) = u_0, \quad v(0)\big|_{\Omega_{\mathrm{s}}} = u_1, \quad \text{and} \quad v(0)\big|_{\Omega_{\mathrm{f}}} = v_0 \quad (3.19)$$

and the weak formulation

$$\begin{aligned} a^F(v,p)(\varphi) + a^S(u,v)(\varphi) &= (\!(f_\mathrm{f},\varphi)\!)_\mathrm{f} + (\!(f_\mathrm{s},\varphi)\!)_\mathrm{s} \quad \forall \varphi \in L^2(I;V), \\ a^{\mathcal{V}}(u,v)(\psi) &= (\!(g_\mathrm{s},\psi)\!)_\mathrm{s} \qquad\quad \forall \psi \in L^2(I;H_\mathrm{s}), \\ a^D(v)(\xi) &= 0 \qquad\qquad\quad \forall \xi \in L^2(I;L_\mathrm{f}). \end{aligned} \tag{3.20}$$

Furthermore, the solution fulfills the a priori estimates

a) $$\begin{aligned} &\|v\|^2_{L^\infty(I;L^2(\Omega))} + \|v\|^2_{L^2(I;H^1(\Omega_\mathrm{f}))} + \|u\|^2_{L^\infty(I;H^1(\Omega_\mathrm{s}))} \\ &\qquad \le C\big[\|f_\mathrm{f}\|^2_{L^2(I;V_\mathrm{f}^*)} + \|f_\mathrm{s}\|^2_{L^2(I;L^2(\Omega_\mathrm{s}))} + \|g\|^2_{L^2(I;L^2(\Omega_\mathrm{s}))} + A_0^2\big], \end{aligned}$$

b) $$\begin{aligned} &\|\partial_t v\|^2_{L^\infty(I;L^2(\Omega))} + \|\partial_t v\|^2_{L^2(I;H^1(\Omega_\mathrm{f}))} + \|v\|^2_{L^\infty(I;H^1(\Omega_\mathrm{s}))} + \|\partial_t u\|^2_{L^2(I;H^1(\Omega_\mathrm{s}))} \\ &\qquad \le C\big[\|f_\mathrm{f}\|^2_{H^1(I;V_\mathrm{f}^*)} + \|f_\mathrm{s}\|_{H^1(I;L^2(\Omega_\mathrm{s}))} + \|g\|^2_{L^2(I;L^2(\Omega_\mathrm{s}))} + A_0^2\big], \end{aligned}$$

c) $$\|p\|^2_{L^2(I;L^2(\Omega_\mathrm{f}))} \le C\big[\|f_\mathrm{f}\|^2_{H^1(I;V_\mathrm{f}^*)} + \|f_\mathrm{s}\|^2_{H^1(I;L^2(\Omega_\mathrm{s}))} + \|g\|^2_{L^2(I;L^2(\Omega_\mathrm{s}))} + A_0^2\big],$$

where A_0^2 *is defined in* (3.16).

Proof. Let $\hat{f}_\mathrm{f} := f_\mathrm{f}$ and $\hat{f}_\mathrm{s} := f_\mathrm{s} + \int_0^t g(s)\,\mathrm{d}s$. Due to the assumptions on the data, we have $\hat{f}_\mathrm{f} \in H^1(I;V_\mathrm{f}^*)$ and $\hat{f}_\mathrm{s} \in H^1(I;H_\mathrm{s})$. Therefore, Proposition 3.1 ensures for these right-hand sides and the given initial data existence of a unique triplet $(\hat{v}_\mathrm{f}, \hat{u}, \hat{p})$ solving (3.17). Next, we introduce a structure velocity $\hat{v}_\mathrm{s}$ by the setting $\hat{v}_\mathrm{s} = \partial_t \hat{u} \in L^\infty(I;V_\mathrm{s})$. Thus, $\hat{v}_\mathrm{s}$ fulfills

$$\mu_\mathrm{s}(\!(\nabla \hat{v}_\mathrm{s}, \nabla\psi)\!)_\mathrm{s} = \mu_\mathrm{s}(\!(\nabla \partial_t \hat{u}, \nabla\psi)\!)_\mathrm{s} \quad \forall \psi \in L^2(I;V_\mathrm{s}). \tag{3.21}$$

Now, we are prepared to introduce the global velocity $\hat{v}$ by setting $\hat{v}\big|_{\Omega_\mathrm{f}} := \hat{v}_\mathrm{f}$ and $\hat{v}\big|_{\Omega_\mathrm{s}} := \hat{v}_\mathrm{s}$. As $\hat{v}_\mathrm{f} \in L^2(I;V_\mathrm{f})$ and $\hat{v}_\mathrm{s} \in L^2(I;V_\mathrm{s})$, we get immediately $\hat{v} \in L^2(I;H)$. However, to obtain $\hat{v} \in L^2(I;V)$, we have to check that the weak partial derivatives $\hat{w}_i$ with $\hat{w}_i\big|_{\Omega_\mathrm{f}} := \partial_{x_i}\hat{v}_\mathrm{f}$ and $\hat{w}_i\big|_{\Omega_\mathrm{s}} := \partial_{x_i}\hat{v}_\mathrm{s}$ constitute the weak partial derivatives $\partial_{x_i}\hat{v}$ of $\hat{v}$ for $i = 1, 2, \ldots, d$. To this end, let $\varphi \in L^2(I;C_0^\infty(\Omega)^d)$. We obtain by the definition of the weak derivatives

$$\begin{aligned} (\!(\hat{v}, \partial_{x_i}\varphi)\!) &= (\!(\hat{v}_\mathrm{f}, \partial_{x_i}\varphi)\!)_\mathrm{f} + (\!(\hat{v}_\mathrm{s}, \partial_{x_i}\varphi)\!)_\mathrm{s} \\ &= -(\!(\partial_{x_i}\hat{v}_\mathrm{f}, \varphi)\!)_\mathrm{f} - (\!(\partial_{x_i}\hat{v}_\mathrm{s}, \varphi)\!)_\mathrm{s} + \langle\!\langle \hat{v}_\mathrm{f}, \varphi n_\mathrm{f}^T e_i \rangle\!\rangle_\mathrm{i} + \langle\!\langle \hat{v}_\mathrm{s}, \varphi n_\mathrm{s}^T e_i \rangle\!\rangle_\mathrm{i} \\ &= -(\!(\hat{w}_i, \varphi)\!) + \langle\!\langle \partial_t\hat{u} - \hat{v}_\mathrm{f}, \varphi n_\mathrm{s}^T e_i \rangle\!\rangle_\mathrm{i} = -(\!(\hat{w}_i, \varphi)\!), \end{aligned}$$

where the last step holds, since the kinematic coupling condition is valid due to Proposition 3.1. Therefore, it holds $\hat{v} \in L^2(I;V)$.

It remains to prove that $(\hat{v}, \hat{u}, \hat{p})$ solves the weak formulation (3.20). Due to the construction of $\hat{v}_\mathrm{s}$ by (3.21), we directly get

$$\rho_\mathrm{s}(\!(\partial_{tt}\hat{u}, \varphi)\!)_\mathrm{s} = \rho_\mathrm{s}(\!(\partial_t\hat{v}, \varphi)\!)_\mathrm{s} \quad \forall \varphi \in L^2(I;V_\mathrm{s}). \tag{3.22}$$

If we enter (3.22) in the weak formulation (3.17), we immediately obtain with (3.21) that the triplet $(\hat{v}, \hat{u}, \hat{p})$ solves the weak formulation (3.20) with the right-hand sides $\hat{f}_\mathrm{f}$, $\hat{f}_\mathrm{s}$, and $g = 0$.

In what follows, we construct a solution to (3.20) with the original right-hand sides f_f, f_s, and g. We define $\tilde{u}\colon I \to V_\mathrm{s}$ for almost all $t \in I$ by

$$\mu_\mathrm{s}(\nabla\tilde{u}(t), \nabla\psi)_\mathrm{s} := (-g(t), \psi)_\mathrm{s} \quad \forall\psi \in V_\mathrm{s}. \tag{3.23}$$

Standard elliptic theory guarantees the existence and uniqueness of $\tilde{u}(t)$ together with the estimate

$$\|\tilde{u}(t)\|_{H^1(\Omega_\mathrm{s})} \leq C\|g(t)\|_{L^2(\Omega_\mathrm{s})} \quad \text{for almost all } t \in I. \tag{3.24}$$

As $g \in L^2(I; H_\mathrm{s})$, the integration in time of the above inequality leads to $\tilde{u} \in L^2(I; V_\mathrm{s})$. Furthermore, integrating (3.23) in time twice implies

$$\mu_\mathrm{s}\left(\!\!\left(\nabla\int_0^t \tilde{u}(s)\,\mathrm{d}s, \nabla\psi\right)\!\!\right)_\mathrm{s} = -\left(\!\!\left(\int_0^t g(s)\,\mathrm{d}s, \psi\right)\!\!\right)_\mathrm{s} \quad \forall\psi \in L^2(I; V_\mathrm{s}).$$

Defining $u := \hat{u} + \int_0^t \tilde{u}\,\mathrm{d}s$, we directly obtain $u \in L^2(I; V_\mathrm{s})$. Since for $\varphi \in L^2(I; V)$ it holds $\psi = \varphi\big|_{\Omega_\mathrm{s}} \in L^2(I; V_\mathrm{s})$, we get for all $\varphi \in L^2(I; V)$ the identity

$$\begin{aligned}\mu_\mathrm{s}(\!(\nabla\hat{u}, \nabla\varphi)\!)_\mathrm{s} &= \mu_\mathrm{s}(\!(\nabla\hat{u}, \nabla\varphi)\!)_\mathrm{s} + \mu_\mathrm{s}\left(\!\!\left(\nabla\int_0^t \tilde{u}(s)\,\mathrm{d}s, \nabla\varphi\right)\!\!\right)_\mathrm{s} + \left(\!\!\left(\int_0^t g(s)\,\mathrm{d}s, \varphi\right)\!\!\right)_\mathrm{s} \\ &= \mu_\mathrm{s}(\!(\nabla u, \nabla\varphi)\!)_\mathrm{s} + \left(\!\!\left(\int_0^t g(s)\,\mathrm{d}s, \varphi\right)\!\!\right)_\mathrm{s}.\end{aligned}$$

Together with the definition of $\hat{f}_\mathrm{s}$, this implies that u, $v := \hat{v}$, and $p := \hat{p}$ solves the first equation of (3.20). Furthermore, since $\hat{u}$ and $\hat{v}$ solve the second equation of (3.20) with $g = 0$, we obtain for all $\psi \in L^2(I; V_\mathrm{s})$

$$\begin{aligned}\mu_\mathrm{s}(\!(\nabla v, \nabla\psi)\!)_\mathrm{s} - \mu_\mathrm{s}(\!(\nabla\partial_t u, \nabla\psi)\!)_\mathrm{s} &= \mu_\mathrm{s}(\!(\nabla\hat{v}, \nabla\psi)\!)_\mathrm{s} - \mu_\mathrm{s}(\!(\nabla\partial_t\hat{u}, \nabla\psi)\!)_\mathrm{s} - \mu_\mathrm{s}(\!(\nabla\tilde{u}, \nabla\psi)\!)_\mathrm{s} \\ &= -\mu_\mathrm{s}(\!(\nabla\tilde{u}, \nabla\psi)\!)_\mathrm{s} = (\!(g, \psi)\!)_\mathrm{s}.\end{aligned}$$

Therefore (v, u, p) solves the weak formulation (3.20) for the right-hand sides f_f, f_s, and g.

Proving the uniqueness remains. Let (v_1, u_1, p_1) and (v_2, u_2, p_2) be two solutions fulfilling the weak formulation (3.20) and the regularities assumed in Theorem 3.3. Define $\bar{v} := v_1 - v_2$, $\bar{u} := u_1 - u_2$ and $\bar{p} := p_1 - p_2$. It holds

$$\bar{v}\big|_{\Omega_\mathrm{f}}(0) = 0, \quad \bar{v}\big|_{\Omega_\mathrm{s}}(0) = 0, \quad \bar{u}(0) = 0,$$

and for almost all $t \in I$

$$\begin{aligned}\rho_\mathrm{f}(\partial_t\bar{v}(t), \varphi)_\mathrm{f} - (\bar{p}(t), \operatorname{div}\varphi)_\mathrm{f} + \nu_\mathrm{f}(\nabla\bar{v}(t), \nabla\varphi)_\mathrm{f} & \\ +\rho_\mathrm{s}(\partial_t\bar{v}(t), \varphi)_\mathrm{s} + \mu_\mathrm{s}(\nabla\bar{u}(t), \nabla\varphi)_\mathrm{s} = 0 & \qquad \forall\varphi \in V, \\ \mu_\mathrm{s}(\nabla\bar{v}(t), \nabla\psi)_\mathrm{s} - \mu_\mathrm{s}(\nabla\partial_t\bar{u}(t), \nabla\psi)_\mathrm{s} = 0 & \qquad \forall\psi \in V_\mathrm{s}, \\ (\xi, \operatorname{div}\bar{v}(t))_\mathrm{f} = 0 & \qquad \forall\xi \in L_\mathrm{f}.\end{aligned}$$

Choosing the test functions $\varphi = \bar{v}(t)$, $\psi = \bar{u}(t)$, and $\xi = \bar{p}(t)$, we get

$$\begin{aligned}\rho_\mathrm{f}(\partial_t\bar{v}(t), \bar{v}(t))_\mathrm{f} - (\bar{p}(t), \operatorname{div}\bar{v}(t))_\mathrm{f} + \nu_\mathrm{f}(\nabla\bar{v}(t), \nabla\bar{v}(t))_\mathrm{f} & \\ +\rho_\mathrm{s}(\partial_t\bar{v}(t), \bar{v}(t))_\mathrm{s} + \mu_\mathrm{s}(\nabla\bar{u}(t), \nabla\bar{v}(t))_\mathrm{s} = 0, & \\ \mu_\mathrm{s}(\nabla\bar{v}(t), \nabla\bar{u}(t))_\mathrm{s} - \mu_\mathrm{s}(\nabla\partial_t\bar{u}(t), \nabla\bar{u}(t))_\mathrm{s} = 0, & \\ (\bar{p}(t), \operatorname{div}\bar{v}(t))_\mathrm{f} = 0. &\end{aligned}$$

Because of the symmetry of the bilinear forms we obtain, for almost all $t \in I$,

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\rho_{\mathrm{f}}\|\bar{v}(t)\|^2_{L^2(\Omega_{\mathrm{f}})} + \nu_{\mathrm{f}}\|\nabla\bar{v}(t)\|^2_{L^2(\Omega_{\mathrm{f}})} + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\rho_{\mathrm{s}}\|\bar{v}(t)\|^2_{L^2(\Omega_{\mathrm{s}})} + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mu_{\mathrm{s}}\|\nabla\bar{u}(t)\|^2_{L^2(\Omega_{\mathrm{s}})} = 0.$$

Integrating this identity in time and noting the initial conditions, we are led to

$$\frac{1}{2}\rho_{\mathrm{f}}\|\bar{v}(t)\|^2_{L^2(\Omega_{\mathrm{f}})} + \nu_{\mathrm{f}}\int_0^t \|\nabla\bar{v}(s)\|^2_{L^2(\Omega_{\mathrm{f}})}\,\mathrm{d}s + \frac{1}{2}\rho_{\mathrm{s}}\|\bar{v}(t)\|^2_{L^2(\Omega_{\mathrm{s}})} + \frac{1}{2}\mu_{\mathrm{s}}\|\nabla\bar{u}(t)\|^2_{L^2(\Omega_{\mathrm{s}})} = 0$$

for almost all $t \in I$. This implies $\bar{v} = 0$ and, as $\bar{u}$ vanishes on $\Gamma_{\mathrm{s}} \subset \Gamma$ with $|\Gamma_{\mathrm{s}}| > 0$, also $\bar{u} = 0$. Thus we get, in particular for almost all $t \in I$

$$(\bar{p}, \operatorname{div}\varphi)_{\mathrm{f}} = 0 \quad \forall\varphi \in V,$$

and thus $\bar{p} = 0$ since $\bar{p} \in L_{\mathrm{f}}$. Therefore, the solution is unique.

According to Proposition 3.1, the solution $(\hat{v}_{\mathrm{f}}, \hat{u}, \hat{p})$ of (3.17) fulfills the estimates a)-c) given in Proposition 3.1 with right-hand side $\hat{f}_{\mathrm{f}} := f_{\mathrm{f}}$ and $\hat{f}_{\mathrm{s}} := f_{\mathrm{s}} + \int_0^t g(s)\,\mathrm{d}s$. As $\hat{v}_{\mathrm{s}}$ of the formulation (3.20) coincides to $\partial_t\hat{u}$, the estimates for $\partial_t\hat{u}$ from the Propositions 3.2 and 3.1 are valid for $\hat{v}\big|_{\Omega_{\mathrm{s}}} = \hat{v}_{\mathrm{s}}$, too. Hence, we have

$$\begin{aligned}
\|\hat{v}\|^2_{L^\infty(I;L^2(\Omega))} &+ \|\hat{v}\|^2_{L^2(I;H^1(\Omega_{\mathrm{f}}))} + \|\hat{u}\|^2_{L^\infty(I;H^1(\Omega_{\mathrm{s}}))} \\
&\le C\big[\|f_{\mathrm{f}}\|^2_{L^2(I;V_{\mathrm{f}}^*)} + \|f_{\mathrm{s}}\|^2_{L^2(I;L^2(\Omega_{\mathrm{s}}))} + \|g\|^2_{L^2(I;L^2(\Omega_{\mathrm{s}}))} + A_0^2\big], \\
\|\partial_t\hat{v}\|^2_{L^\infty(I;L^2(\Omega))} &+ \|\partial_t\hat{v}\|^2_{L^2(I;H^1(\Omega_{\mathrm{f}}))} + \|\hat{v}\|^2_{L^\infty(I;H^1(\Omega_{\mathrm{s}}))} \\
&\le C\big[\|f_{\mathrm{f}}\|^2_{H^1(I;V_{\mathrm{f}}^*)} + \|f_{\mathrm{s}}\|^2_{H^1(I;L^2(\Omega_{\mathrm{s}}))} + \|g\|^2_{L^2(I;L^2(\Omega_{\mathrm{s}}))} + A_0^2\big], \\
\|\hat{p}\|^2_{L^2(I;L^2(\Omega_{\mathrm{f}}))} &\le C\big[\|f_{\mathrm{f}}\|^2_{H^1(I;V_{\mathrm{f}}^*)} + \|f_{\mathrm{s}}\|^2_{H^1(I;L^2(\Omega_{\mathrm{s}}))} + \|g\|^2_{L^2(I;L^2(\Omega_{\mathrm{s}}))} + A_0^2\big].
\end{aligned}$$

Due to the setting $v = \hat{v}$ and $p = \hat{p}$, these estimates directly transfer to v and p. To estimate $u := \hat{u} + \int_0^t \tilde{u}(s)\,\mathrm{d}s$, $\tilde{u} \in V_{\mathrm{s}}$ given by (3.23) has to be bounded. By (3.24), we get for almost all $t \in I$ that

$$\begin{aligned}
\|u(t)\|^2_{H^1(\Omega_{\mathrm{s}})} &= \left\|\hat{u}(t) + \int_0^t \tilde{u}(s)\,\mathrm{d}s\right\|^2_{H^1(\Omega_{\mathrm{s}})} \\
&\le C\left[\|\hat{u}(t)\|^2_{H^1(\Omega_{\mathrm{s}})} + T\int_I \|\tilde{u}(s)\|^2_{H^1(\Omega_{\mathrm{s}})}\mathrm{d}s\right] \\
&\le C\big[\|\hat{u}(t)\|^2_{H^1(\Omega_{\mathrm{s}})} + \|g\|^2_{L^2(I;L^2(\Omega_{\mathrm{s}}))}\big],
\end{aligned}$$

which implies

$$\|u\|^2_{L^\infty(I;H^1(\Omega_{\mathrm{s}}))} \le C\big[\|\hat{u}\|^2_{L^\infty(I;H^1(\Omega_{\mathrm{s}}))} + \|g\|^2_{L^2(I;L^2(\Omega_{\mathrm{s}}))}\big].$$

Furthermore, we get with $\partial_t\hat{u} = \hat{v}\big|_{\Omega_{\mathrm{s}}}$ for almost all $t \in I$ that

$$\|\partial_t u(t)\|^2_{H^1(\Omega_{\mathrm{s}})} = \|\partial_t\hat{u}(t) + \tilde{u}(t)\|^2_{H^1(\Omega_{\mathrm{s}})} \le C\big[\|\hat{v}(t)\|^2_{H^1(\Omega_{\mathrm{s}})} + \|g(t)\|^2_{L^2(\Omega_{\mathrm{s}})}\big],$$

and consequently (limited through $g \in L^2(I; H_{\mathrm{s}})$)

$$\|\partial_t u\|^2_{L^2(I;H^1(\Omega_{\mathrm{s}}))} \le C\big[\|\hat{v}\|^2_{L^2(I;H^1(\Omega_{\mathrm{s}}))} + \|g\|^2_{L^2(I;L^2(\Omega_{\mathrm{s}}))}\big].$$

Together with the above estimates for $(\hat{v}, \hat{u}, \hat{p})$, we obtain the stated estimates for (v, p, u). □

Remark 3.5. If the right-hand side g lies in $L^\infty(I; H_{\mathrm{s}})$, then we also get an estimate for $\|\partial_t u\|_{L^\infty(I;H^1(\Omega_{\mathrm{s}}))}$, as in Proposition 3.1.

3.2.3. How are the Coupling Conditions Fulfilled?

In the following, we analyze in which sense the weak solution of (3.20) fulfills the original fluid-structure interaction problem, and especially in which sense the coupling conditions are fulfilled. To this end, we introduce the space $\widetilde{H}^{\frac{1}{2}}(\Gamma_\mathrm{i})$ in the spirit of [78, Definition 1.3.2.5] by

$$\widetilde{H}^{\frac{1}{2}}(\Gamma_\mathrm{i}) = \{\, v \in H^{\frac{1}{2}}(\Gamma_\mathrm{i}) : \tilde{v} \in H^{\frac{1}{2}}(\Gamma) \,\},$$

where $\tilde{v}$ denotes the continuation of v on Γ by zero.

Theorem 3.4. *Let the assumptions of Theorem 3.3 be fulfilled, and let (v, u, p) be the solution of (3.20) and in addition $f_\mathrm{f} \in L^2(I; H_\mathrm{f})$. Then, the kinematic coupling condition*

$$v\big|_{\Omega_\mathrm{f}} = v\big|_{\Omega_\mathrm{s}}$$

is valid in the sense of $L^2(I; H^{\frac{1}{2}}(\Gamma_\mathrm{i})^d) \cap H^{\frac{1}{2}}(I; L^2(\Gamma_\mathrm{i})^d)$. Furthermore, the dynamic coupling condition

$$\nu_\mathrm{f} \nabla v\, n_\mathrm{f} - p n_\mathrm{f} + \mu_\mathrm{s} \nabla u\, n_\mathrm{s} = 0$$

holds in $L^2(I; (\widetilde{H}^{\frac{1}{2}}(\Gamma_\mathrm{i})^d)^*)$.

Proof. By Theorem 3.3, we have that $v \in L^2(I; V)$ and $v \in W^{1,\infty}(I; H) \subset H^1(I; H)$. Hence, the trace results in [105, Theorem 2.1] imply that the kinematic coupling condition $v\big|_{\Omega_\mathrm{f}} = v\big|_{\Omega_\mathrm{s}}$ holds on Γ_i in the space $L^2(I; H^{\frac{1}{2}}(\Gamma_\mathrm{i})^d) \cap H^{\frac{1}{2}}(I; L^2(\Gamma_\mathrm{i})^d)$.

In the remaining part of this proof, we derive validity of the stated dynamic coupling condition. In (3.20), we choose test functions φ with $\varphi\big|_{\Omega_\mathrm{f}} \in L^2(I; C_0^\infty(\Omega_\mathrm{f})^d)$ and $\varphi\big|_{\Omega_\mathrm{s}} = 0$ to get

$$(\!(\rho_\mathrm{f} \partial_t v, \varphi)\!)_\mathrm{f} - (\!(\mathrm{div}(\nu_\mathrm{f} \nabla v + p\,\mathrm{Id}), \varphi)\!)_\mathrm{f} = (\!(f_\mathrm{f}, \varphi)\!)_\mathrm{f} \quad \forall \varphi \in L^2(I; C_0^\infty(\Omega_\mathrm{f})^d),$$

where $\mathrm{div}(\nu_\mathrm{f} \nabla v + p\,\mathrm{Id})$ is defined in the distributional sense. This is equivalent to

$$\mathrm{div}(\nu_\mathrm{f} \nabla v + p\,\mathrm{Id}) = \rho_\mathrm{f} \partial_t v - f_\mathrm{f} \quad \text{in } L^2(I; C_0^\infty(\Omega_\mathrm{f})^d)^*.$$

As Theorem 3.3 yields $v\big|_{\Omega_\mathrm{f}} \in W^{1,\infty}(I; H_\mathrm{f})$, we get by the assumption on $f_\mathrm{f} \in L^2(I; H_\mathrm{f})$ that $\rho_\mathrm{f} \partial_t v + f_\mathrm{f} \in L^2(I; H_\mathrm{f})$. We immediately obtain, that

$$\mathrm{div}(\nu_\mathrm{f} \nabla v + p\,\mathrm{Id}) = \rho_\mathrm{f} \partial_t v - f_\mathrm{f} \quad \text{in } L^2(I; H_\mathrm{f}) \tag{3.25}$$

and

$$\|\mathrm{div}(\nu_\mathrm{f} \nabla v + p\,\mathrm{Id})\|_{L^2(I; L^2(\Omega_\mathrm{f}))} \leq \|f_\mathrm{f}\|_{L^2(I; L^2(\Omega_\mathrm{f}))} + C \|\partial_t v\|_{L^2(I; L^2(\Omega_\mathrm{f}))},$$

where the right-hand side is bounded according to Theorem 3.3. The same approach, but instead choosing $\varphi\big|_{\Omega_\mathrm{f}} = 0$ and $\varphi\big|_{\Omega_\mathrm{s}} \in L^2(I; C_0^\infty(\Omega_\mathrm{s})^d)$, leads to

$$\mathrm{div}(\mu_\mathrm{s} \nabla u) = \rho_\mathrm{s} \partial_t v - f_\mathrm{s} \quad \text{in } L^2(I; H_\mathrm{s}) \tag{3.26}$$

and

$$\|\mathrm{div}(\mu_\mathrm{s} \nabla u)\|_{L^2(I; L^2(\Omega_\mathrm{s}))} \leq \|f_\mathrm{s}\|_{L^2(I; L^2(\Omega_\mathrm{s}))} + C \|\partial_t v\|_{L^2(I; L^2(\Omega_\mathrm{s}))},$$

where the right-hand side is again bounded according to Theorem 3.3. Thus, we obtain that

$$\begin{aligned}\nu_\mathrm{f}\nabla v + p\,\mathrm{Id} \in E(\Omega_\mathrm{f}) &:= \{\,\varphi \in L^2(I;H_\mathrm{f}) : \|\mathrm{div}\,\varphi\|_{L^2(I;L^2(\Omega_\mathrm{f}))} < \infty\,\},\\ \mu_\mathrm{s}\nabla u \in E(\Omega_\mathrm{s}) &:= \{\,\varphi \in L^2(I;H_\mathrm{s}) : \|\mathrm{div}\,\varphi\|_{L^2(I;L^2(\Omega_\mathrm{s}))} < \infty\,\}.\end{aligned}$$

According to [137, Ch. I §1 Theorem 1.1], the space $L^2(I;C_0^\infty(\Omega_\mathrm{f})^d)$ is dense in $E(\Omega_\mathrm{f})$ and $L^2(I;C_0^\infty(\Omega_\mathrm{s})^d)$ is dense in $E(\Omega_\mathrm{s})$. Therefore, following [69, p. 114] or [78, Theorems 1.5.3.10 and 1.5.3.11], we get

$$\begin{aligned}\|(\nu_\mathrm{f}\nabla v + p\,\mathrm{Id})^T n_\mathrm{f}\|_{L^2(I;(\widetilde{H}^{\frac{1}{2}}(\Gamma_\mathrm{i})^d)^*)} &\le \|\nu_\mathrm{f}\nabla v + p\,\mathrm{Id}\|_{L^2(I;L^2(\Omega_\mathrm{f}))}\\ &\quad + \|f_\mathrm{f}\|_{L^2(I;L^2(\Omega_\mathrm{f}))} + C\|\partial_t v\|_{L^2(I;L^2(\Omega_\mathrm{f}))}\\ \|(\mu_\mathrm{s}\nabla u)^T n_\mathrm{s}\|_{L^2(I;(\widetilde{H}^{\frac{1}{2}}(\Gamma_\mathrm{i})^d)^*)} &\le \|\mu_\mathrm{s}\nabla u\|_{L^2(I;L^2(\Omega_\mathrm{s}))}\\ &\quad + \|f_\mathrm{s}\|_{L^2(I;L^2(\Omega_\mathrm{s}))} + C\|\partial_t v\|_{L^2(I;L^2(\Omega_\mathrm{s}))}.\end{aligned}$$

According to [69, 78], this enables us to apply Gauss' theorem in (3.20) to obtain

$$\begin{aligned}\langle\!\langle \nu_\mathrm{f}\nabla v\, n_\mathrm{f} - p n_\mathrm{f}, \varphi\rangle\!\rangle_\mathrm{i} + \langle\!\langle \mu_\mathrm{s}\nabla u\, n_\mathrm{s}, \varphi\rangle\!\rangle_\mathrm{i} &= (\!(\mathrm{div}(\nu_\mathrm{f}\nabla v + p\,\mathrm{Id}) - \rho_\mathrm{f}\partial_t v + f_\mathrm{f}, \varphi)\!)_\mathrm{f},\\ &\quad + (\!(\mathrm{div}(\mu_\mathrm{s}\nabla u) - \rho_\mathrm{s}\partial_t v + f_\mathrm{s}, \varphi)\!)_\mathrm{s} \quad \forall \varphi \in L^2(I;V).\end{aligned}$$

This immediately implies by (3.26) and (3.25) that

$$\langle\!\langle \nu_\mathrm{f}\nabla v\, n_\mathrm{f} - p n_\mathrm{f} + \mu_\mathrm{s}\nabla u\, n_\mathrm{s}, \varphi\rangle\!\rangle_\mathrm{i} = 0 \quad \forall \varphi \in L^2(I;V)$$

and thus the dynamic coupling condition is fulfilled in $L^2(I;(\widetilde{H}^{\frac{1}{2}}(\Gamma_\mathrm{i})^d)^*)$. □

3.3. Optimal Control Problem

In the following, we consider the optimal control problem of a linearized FSI configuration given by

Problem 3.4 (Optimal Control Problem)**.**

$$\min_{q\in Q} \mathcal{J}(q,u,v) := \frac{\gamma_\mathrm{f}}{2}\int_I \|v - v_\mathrm{d}\|^2_{L^2(\Omega_\mathrm{f})}\,dt + \frac{\gamma_\mathrm{s}}{2}\int_I \|u - u_\mathrm{d}\|^2_{L^2(\Omega_\mathrm{s})}\,dt + \frac{\alpha}{2}\|q\|^2_Q, \tag{3.27}$$

subject to

$$\tilde{v} \in H^1(I;V^*_\mathrm{div}) \text{ with } v = \tilde{v}\big|_{\Omega_\mathrm{f}} \in L^2(I;V_{\mathrm{f,div}}) \text{ and } u = \int_0^t \tilde{v}(s)\big|_{\Omega_\mathrm{s}}\,ds + u_0 \in L^2(I;V_\mathrm{s})$$

satisfying the initial conditions $v(0) = v_0$, $u(0) = u_0$, *and* $\partial_t u(0) = u_1$ *and*

$$\begin{aligned}&\rho_\mathrm{f}(\!(\partial_t\tilde{v}, \varphi)\!)_\mathrm{f} + \nu_\mathrm{f}(\!(\nabla v, \nabla\varphi)\!)_\mathrm{f}\\ &\quad + \rho_\mathrm{s}(\!(\partial_t\tilde{v}, \varphi)\!)_\mathrm{s} + \mu_\mathrm{s}(\!(\nabla u, \nabla\varphi)\!)_\mathrm{s} = (\!(B_\mathrm{f}q, \varphi)\!)_\mathrm{f} + (\!(B_\mathrm{s}q, \varphi)\!)_\mathrm{s} \quad \forall \varphi \in L^2(I;V_\mathrm{div}),\\ &\qquad \int_0^t v(s)\big|_{\Gamma_\mathrm{i}}\,ds = u(t)\big|_{\Gamma_\mathrm{i}} - u_0\big|_{\Gamma_\mathrm{i}} \quad \text{in } L^2(I;H^{\frac{1}{2}}(\Gamma_\mathrm{i})^d),\\ &\qquad\qquad q_a \le q \le q_b.\end{aligned}$$

The control q is going to be either time-dependent or distributed in space and controlling the volume force through the linear operators B_f and B_s; see the two configurations in Section 3.3.1 for details. In addition, the control variable is subject to control constrains with the bounds $q_a, q_b \in \mathbb{R} \cup \{\pm\infty\}$ and $q_a < q_b$. The variables v_d and u_d are given desired states and $\alpha > 0$ is a given regularization parameter. To enable observation on both or just on one sub-domain, the parameters $\gamma_\mathrm{f}, \gamma_\mathrm{s} \geq 0$ can be chosen appropriately. The initial data are assumed to satisfy Assumption 1, and for the desired states we require $v_\mathrm{d} \in H^1(I; V_\mathrm{f}^*) \cap L^2(I; H_\mathrm{f})$ and $u_\mathrm{d} \in L^2(I; H_\mathrm{s})$. The fluid velocity v and solid displacement u only have to fulfill the pressure free weak formulation (3.18). This enables control variables in the space $L^2(I)$ as we will see in the next Section. For the optimal control we can prove higher regularity such that the optimal velocity and displacement variable are solutions of the symmetric weak formulation (3.20), as we see later.

3.3.1. Existence and Uniqueness of Optimal Solutions

We analyze two concrete configurations for the considered control problem:

Configuration C1 Let the control space given by $Q := (L^2(\Omega)^d)^N$ with $N \in \mathbb{N}$ and let $B_\mathrm{f} \colon Q \to H^1(I; H_\mathrm{f})$, as well as $B_\mathrm{s} \colon Q \to H^1(I; H_\mathrm{s})$ be linear and continuous operators, given for $q = (q^1, q^2, \dots, q^N) \in Q$ by

$$B_\mathrm{f} q = \sum_{i=1}^{N} g_\mathrm{f}^i q^i \big|_{\Omega_\mathrm{f}} \quad \text{and} \quad B_\mathrm{s} q = \sum_{i=1}^{N} g_\mathrm{s}^i q^i \big|_{\Omega_\mathrm{s}}.$$

Thereby, $g_\mathrm{s}^i, g_\mathrm{f}^i \in H^1(I)$, $i = 1, 2, \dots, N$, are given functions. The admissible set Q_ad is defined as

$$Q_\mathrm{ad} := \{\, q \in Q : q_a \leq q^i(x) \leq q_b, \text{ for almost all } x \in \Omega \text{ and } i = 1, 2, \dots, N \,\}.$$

Note that the inequality in the definition of Q_ad has to be understood componentwise for $q^i \in L^2(\Omega)^d$.

Configuration C2 Let the control space given by $Q := L^2(I)^N$ with $N \in \mathbb{N}$ and let $B_\mathrm{f} \colon Q \to L^2(I; H_\mathrm{f})$, as well as $B_\mathrm{s} \colon Q \to L^2(I; H_\mathrm{s})$ be linear continuous operators given for $q = (q^1, q^2, \dots, q^N) \in Q$ by

$$B_\mathrm{f} q = \sum_{i=1}^{N} q^i h^i \big|_{\Omega_\mathrm{f}} \quad \text{and} \quad B_\mathrm{s} q = \sum_{i=1}^{N} q^i h^i \big|_{\Omega_\mathrm{s}}.$$

Thereby, $h^i \in V_\mathrm{div}$, $i = 1, 2, \dots, N$, are given functions. The admissible set Q_ad is defined as

$$Q_\mathrm{ad} := \{\, q \in Q : q_a \leq q^i(t) \leq q_b, \text{ for almost all } t \in I \text{ and } i = 1, 2, \dots, N \,\}.$$

The assumption that h^i has to be divergence-free in Ω_f is taken for simplicity of the presentation. All results can be extended to $h^i \in V$ using a Helmholtz decomposition.

Since for both configurations, $B_\mathrm{f} q \in L^2(I; H_\mathrm{f})$ and $B_\mathrm{s} q \in L^2(I; H_\mathrm{s})$, Proposition 3.2 ensures the well-posedness of the so-defined control to state mapping $G \colon q \mapsto (v(q), u(q))$, with $(v(q), u(q))$ the solution of (3.18) for $f_\mathrm{f} := B_\mathrm{s} q$ and $f_\mathrm{s} := B_\mathrm{s} q$. The linearity of (3.18) and the estimate given in Proposition 3.2 imply the continuity of G.

Lemma 3.5. *The control to state mapping* $G\colon Q \to L^2(I;H_\mathrm{f}) \times L^2(I;H_\mathrm{s})$ *is an affine linear and continuous operator for both configurations* C1 *and* C2.

Proof. Let $(\hat{v},\hat{u})$ be the solution of (3.18) for $f_\mathrm{f} = f_\mathrm{s} = 0$. Furthermore, we denote by $G_0\colon Q \to L^2(I;H_\mathrm{f}) \times L^2(I;H_\mathrm{s})$ the linear part of G, defined by (3.18) with zero initial data for $f_\mathrm{f} := B_\mathrm{s}q$ and $f_\mathrm{s} := B_\mathrm{s}q$. Hence, the mapping $G\colon Q \to L^2(I;H_\mathrm{f}) \times L^2(I;H_\mathrm{s})$ can be expressed as

$$(v(q),u(q)) := Gq = (\hat{v},\hat{u}) + G_0 q.$$

Proposition 3.2 yields that $(\hat{v},\hat{u})$ is bounded in $L^\infty(I;H_\mathrm{f}) \times L^\infty(I;V_\mathrm{s})$ and G_0 is a bounded linear operator. Thus, the control to state mapping G is continuous in both considered configurations. □

By means of applying the control to state mapping G, the reduced functional $j\colon Q \to \mathbb{R}$ can be defined as

$$j(q) := \mathcal{J}(q,u(q),v(q)), \tag{3.28}$$

and the optimal control problem 3.4 can for both configurations be written in the compact form

$$\min_{q\in Q_\mathrm{ad}} j(q). \tag{3.29}$$

Theorem 3.6. *The considered optimal control problem* (3.29) *admits a unique solution for both configurations* C1 *and* C2.

Proof. Standard arguments, see, e.g. [141, Theorem 2.14], guarantee the existence of a unique optimal control $\bar{q} \in Q_\mathrm{ad}$. □

We emphasize that this existence result is also valid if the control is acting only on the domain Ω_s or Ω_f and if reference solutions are only given on sub-domains.

3.3.2. Adjoint Equations

The necessary optimality conditions to be derived in Section 3.3.3 will make use of an adjoint equation. However, we will not derive the adjoint equation for the optimal control problem 3.4 with the weak formulation (3.18). Instead we replace equation (3.18) by the symmetric weak formulation (3.20). For the resulting optimal control problem we derive formally the adjoint equation, then we prove existence and uniqueness of adjoint solutions for the symmetric adjoint equation. Later in Section 3.3.3 we are going to prove that the resulting optimal solution will fulfill the derived symmetric formulation.

Formal Lagrange Formalism

We replace equation (3.18) by the symmetric weak formulation (3.20) in the optimal control problem 3.4. As Theorem 3.3 does not guarantee existence of solutions for a right-hand side as in the two configurations presented in Section 3.3, we only derive formally the Lagrangian. As suggested in [141], we derive the adjoint equation without taking into account if the control to state mapping is well defined or differentiable.

Similar to the derivation of a Karush-Kuhn-Tucker system for discrete constraint optimization problems, we define Lagrange multipliers. Multiplying the strong linear fluid-structure interaction equation with a Lagrange multiplier $\mathbf{z} = (z^v, z^u, z^p)$ and integrating by parts immediately gets us the weak linear FSI formulation. The scalar products in equation (3.20) are well defined, if we choose the Lagrange multiplier $\mathbf{z} \in Y := L^2(I;V) \times L^2(I;V_\mathrm{s}) \times L^2(I;L_\mathrm{f})$ and the primal variable $\mathbf{u} = (v,u,p) \in X := W^v \times W^u_\mathrm{s} \times L^2(I;L_\mathrm{f})$. Then, we can define for the symmetric linear fluid-structure interaction problem the Lagrangian $\mathcal{L} : Q \times X \times Y \to \mathbb{R}$ by

$$\begin{aligned}\mathcal{L}(q,\mathbf{u},\mathbf{z}) :=& \mathcal{J}(q,u,v) - a(\mathbf{u})(\mathbf{z}) + (\!(B_\mathrm{f} q, z^v)\!)_\mathrm{f} + (\!(B_\mathrm{s} q, z^v)\!)_\mathrm{s} \\ &+ (u_0 - u(0), z^u(0))_\mathrm{s} + (v_0 - v(0), z^v(0))\end{aligned}$$

and $a(\mathbf{u})(\mathbf{z})$ is the symmetric bilinearform defined in (3.14).

If the triple $\mathbf{u} = (v,u,p)$ is the solution of the weak formulation (3.20) with the right hand side $f_\mathrm{s} = B_\mathrm{f} q$ and $f_\mathrm{s} = B_\mathrm{s} q$ and initial conditions (u_0, v_0), the useful identity

$$j(q) := \mathcal{J}(q, \mathbf{u}(q)) = \mathcal{L}(q, \mathbf{u}(q), \mathbf{z})$$

is true for arbitrary values $\mathbf{z} = (z^v, z^u, z^p) \in Y$. If we denote with $\delta\mathbf{u} = \mathbf{u}'_q(q)(\delta q)$ the derivative of the state variable with respect to the control, we obtain via the Lagrange functional the representation

$$j'(q)(\delta q) = \mathcal{L}'_q(q, \mathbf{u}(q), \mathbf{z})(\delta q) + \mathcal{L}'_\mathbf{u}(q, \mathbf{u}(q), \mathbf{z})(\delta \mathbf{u})$$

of the derivative of the reduced functional. If we choose $\mathbf{z} \in Y$ such that $\mathbf{z}$ solves the adjoint equation

$$\mathcal{L}'_\mathbf{u}(q, \mathbf{u}, \mathbf{z})(\boldsymbol{\varphi}) = 0 \quad \forall \boldsymbol{\varphi} \in X, \tag{3.30}$$

the derivative of the Lagrange functional with respect to $\mathbf{u}$ is zero in arbitrary directions $\boldsymbol{\varphi} = (\varphi, \psi, \xi) \in X$. Then we can evaluate the derivative of the reduced functional in an arbitrary direction $\delta q \in Q$ by evaluating

$$j'(q)(\delta q) = \mathcal{L}'_q(q, \mathbf{u}, \mathbf{z})(\delta q).$$

It is not necessary anymore to calculate the derivative of the solution $\mathbf{u}(q)$ with respect to the control.

In case of the linear fluid-structure interaction problem the adjoint variable $\mathbf{z} \in Y$ has to solve

$$a(\boldsymbol{\varphi})(\mathbf{z}) - (\psi(0), z^u(0))_\mathrm{s} - (\varphi(0), z^v(0)) = \mathcal{J}'_\mathbf{u}(q,u,v)(\boldsymbol{\varphi}) \quad \forall \boldsymbol{\varphi} \in X.$$

Without taking the regularity of the adjoint variable into account, we formally integrate by parts in time. Then we get, with $\mathcal{J}(q,u,v)$ the tracking type functional in (3.27), the adjoint equation

$$\begin{aligned} -\rho_{\mathrm{f}}(\!(\varphi,\partial_t z^v)\!)_{\mathrm{f}} + \nu_{\mathrm{f}}(\!(\nabla\varphi,\nabla z^v)\!)_{\mathrm{f}} + (\!(z^p,\operatorname{div}\varphi)\!)_{\mathrm{f}} & \\ -\rho_{\mathrm{s}}(\!(\varphi,\partial_t z^v)\!)_{\mathrm{s}} + \mu_{\mathrm{s}}(\!(\nabla\varphi,\nabla z^u)\!)_{\mathrm{s}} &= \gamma_{\mathrm{f}}(\!(v-v_{\mathrm{d}},\varphi)\!)_{\mathrm{f}} \quad \forall\varphi\in L^2(I;V), \\ \mu_{\mathrm{s}}(\!(\nabla\psi,\nabla z^v)\!)_{\mathrm{s}} + \mu_{\mathrm{s}}(\!(\nabla\psi,\nabla\partial_t z^u)\!)_{\mathrm{s}} &= \gamma_{\mathrm{s}}(\!(u-u_{\mathrm{d}},\psi)\!)_{\mathrm{s}} \quad \forall\psi\in L^2(I;V_{\mathrm{s}}), \\ -(\!(\xi,\operatorname{div} z^v)\!)_{\mathrm{f}} &= 0 \qquad\qquad \forall\xi\in L^2(I;L_{\mathrm{f}}) \end{aligned} \tag{3.31}$$

and the terminal conditions $z^v(T)=0$ and $z^u(T)=0$. Thus the derivative of the Lagrangian with respect to test variable results in a linear fluid-structure interaction problem only running back in time. We will prove existence and regularity results for the adjoint equation (3.31) in the following. Later, we will see that (3.31) is indeed the correct adjoint equation appearing in the optimality system.

Discussion of the Adjoint Equations for a Non-symmetric Formulation

Before analyzing the adjoint equations of the weak formulation (3.14), we investigate for a moment the optimal control problem subject to linear fluid-structure interaction, whereby we replace the bi-linear form $a^{\mathcal{V}}(\cdot)(\cdot)$ in the monolithic formulation (3.14) with $a^{V}(\cdot)(\cdot)$ given in 3.2. For this formulation, the formal Lagrange approach leads to the following adjoint equation:

$$\begin{aligned} -(\!(\varphi,z_t^v)\!)_{\mathrm{f}} + \nu_{\mathrm{f}}(\!(\nabla\varphi,\nabla z^v)\!)_{\mathrm{f}} + (\!(z^p,\operatorname{div}\varphi)\!)_{\mathrm{f}} & \\ -(\!(\varphi,z_t^v)\!)_{\mathrm{s}} - (\!(\varphi,z^u)\!)_{\mathrm{s}} &= \gamma_{\mathrm{f}}(\!(v-v_{\mathrm{d}},\varphi)\!)_{\mathrm{f}} \quad \forall\varphi\in L^2(I;V), \\ -(\!(\psi,z_t^u)\!)_{\mathrm{s}} + \mu_{\mathrm{s}}(\!(\nabla\psi,\nabla z^v)\!)_{\mathrm{s}} &= \gamma_{\mathrm{s}}(\!(u-u_{\mathrm{d}},\psi)\!)_{\mathrm{s}} \quad \forall\psi\in L^2(I;V_{\mathrm{s}}), \\ -(\!(\xi,\operatorname{div} z^v)\!)_{\mathrm{f}} &= 0 \qquad\qquad \forall\xi\in L^2(I;L_{\mathrm{f}}). \end{aligned}$$

Here, $z^v\big|_{\Omega_{\mathrm{f}}}$ describes the solution of an adjoint Stokes equation, and $z^v\big|_{\Omega_{\mathrm{s}}}$ the solution of an adjoint linear wave equation. However, as the bi-linear form formulation (3.14) with $a^{V}(\cdot)(\cdot)$ is not symmetric, the adjoint equation is a Stokes-wave system with new coupling conditions on Γ_{i}:

$$z^v\big|_{\Omega_{\mathrm{f}}} = z^v\big|_{\Omega_{\mathrm{s}}}, \quad \nu_{\mathrm{f}}\nabla z^v\, n_{\mathrm{f}} - z^p n_{\mathrm{f}} = 0, \quad \text{and} \quad \mu_{\mathrm{s}}\nabla z^v\, n_{\mathrm{s}} = 0.$$

In contrast to this, the advantage of the following optimality system lies in the fact that the adjoint equation is again a linear FSI problem and all numerical methods developed to solve the primal FSI problem can be utilized. Therefore, no additional difficulty occurs in the implementation and analysis.

Existence and Regularity for the Adjoint Equation

Due to the symmetry in (3.20), the adjoint equation is again a linear FSI problem. Therefore, we can use the already proved results, to get the following result on existence of a unique adjoint solution:

Theorem 3.7. *Let $v_\mathrm{d} \in H^1(I;V_\mathrm{f}^*) \cap L^2(I;H_\mathrm{f})$, $u_\mathrm{d} \in L^2(I;H_\mathrm{s})$, and let the initial data u_0, u_1, and v_0 satisfy Assumption 1. Furthermore, let $q \in Q$ be given as in configuration* C1 *or $q \in Q \cap H^1(I)^N$ be given for configuration* C2*, and let the triple (v,u,p) be the corresponding solution of* (3.20) *with $f_\mathrm{f} = B_\mathrm{f} q$, $f_\mathrm{s} = B_\mathrm{s} q$, and $g = 0$. Then, there exists a unique triple (z^v, z^u, z^p) with*

$$z^v \in L^2(I;V) \cap W^{1,\infty}(I;H), \quad z^v\big|_{\Omega_\mathrm{f}} \in H^1(I;V_\mathrm{f}), \quad z^v\big|_{\Omega_\mathrm{s}} \in L^\infty(I;V_\mathrm{s}),$$
$$z^u \in L^\infty(I;V_\mathrm{s}) \cap H^1(I;V_\mathrm{s}), \quad z^p \in L^2(I;L_\mathrm{f})$$

satisfying the terminal condition $z^v(T) = 0$, $z^u(T) = 0$ and the adjoint equation (3.31)*. Furthermore, the adjoint solution triple (z^v, z^u, z^p) fulfills the a priori estimates*

a)
$$\begin{aligned}\|z^v\|^2_{L^\infty(I;L^2(\Omega))} + \|z^v\|^2_{L^2(I;H^1(\Omega_\mathrm{f}))} + \|z^u\|^2_{L^\infty(I;H^1(\Omega_\mathrm{s}))} \\ \leq C\big[\|v_\mathrm{d}\|^2_{L^2(I;V_\mathrm{f}^*)} + \|u_\mathrm{d}\|^2_{L^2(I;L^2(\Omega_\mathrm{s}))}\big] \\ + C\big[\|B_\mathrm{f} q\|^2_{L^2(I;V_\mathrm{f}^*)} + \|B_\mathrm{s} q\|^2_{L^2(I;L^2(\Omega_\mathrm{s}))} + A_0^2\big],\end{aligned}$$

b)
$$\begin{aligned}\|\partial_t z^v\|^2_{L^\infty(I;L^2(\Omega))} + \|\partial_t z^v\|^2_{L^2(I;H^1(\Omega_\mathrm{f}))} + \|z^v\|^2_{L^\infty(I;H^1(\Omega_\mathrm{s}))} + \|\partial_t z^u\|^2_{L^2(I;H^1(\Omega_\mathrm{s}))} \\ \leq C\big[\|v_\mathrm{d}\|^2_{H^1(I;V_\mathrm{f}^*)} + \|u_\mathrm{d}\|_{L^2(I;L^2(\Omega_\mathrm{s}))}\big] \\ + C\big[\|B_\mathrm{f} q\|^2_{H^1(I;V_\mathrm{f}^*)} + \|B_\mathrm{s} q\|_{H^1(I;L^2(\Omega_\mathrm{s}))} + A_0^2\big],\end{aligned}$$

c)
$$\begin{aligned}\|z^p\|^2_{L^2(I;L^2(\Omega_\mathrm{f}))} \leq C\big[\|v_\mathrm{d}\|^2_{H^1(I;V_\mathrm{f}^*)} + \|u_\mathrm{d}\|^2_{L^2(I;L^2(\Omega_\mathrm{s}))})\big] \\ + C\big[\|B_\mathrm{f} q\|^2_{H^1(I;V_\mathrm{f}^*)} + \|B_\mathrm{s} q\|^2_{H^1(I;L^2(\Omega_\mathrm{s}))} + A_0^2\big],\end{aligned}$$

where A_0^2 is defined in (3.16)*.*

Proof. Based on the assumptions on the control q, with $f_\mathrm{f} = B_\mathrm{f} q$, $f_\mathrm{s} = B_\mathrm{s} q$, and $g = 0$ Theorem 3.3 ensures a solution $v \in H^1(I;H)$, $u \in L^2(I;V_\mathrm{s})$ of (3.20). Hence, the right-hand sides of the adjoint equation $f_\mathrm{f} := \gamma_\mathrm{f}(v - v_\mathrm{d})$ and $g := \gamma_\mathrm{s}(u - u_\mathrm{d})$ fulfill the required regularities $f_\mathrm{f} \in H^1(I;V_\mathrm{f}^*)$ and $g \in L^2(I;H_\mathrm{s})$ of Theorem 3.3. Furthermore, the initial conditions for the adjoint equation backwards in time $z_T^v = 0$ and $z_T^u = 0$ fulfill with $z_T^p = 0$ the assumptions on u_0, v_0, p_0, and u_1 of Theorem 3.3. As the considered adjoint equation (3.31) coincides with the state equation (3.20) after the transformation $t \mapsto -t$, there exists a unique adjoint solution (z^v, z^u, z^p) due to Theorem 3.3. The estimates follow immediately from Theorem 3.3, too. □

3.3.3. Necessary Optimality Conditions

Since the reduced functional j is convex due to the (affine) linear-quadratic structure of the considered control problem, the necessary and sufficient optimality condition for the optimal solution $\bar{q} \in Q_\mathrm{ad}$ of (3.29) reads as

$$j'(\bar{q})(\delta q - \bar{q}) \geq 0 \quad \forall \delta q \in Q_\mathrm{ad}. \tag{3.32}$$

Based on this, we derive in the sequel an optimality system separately for the configurations C1 and C2. Thereby, we make use of the self-adjoint formulation (3.20). Due to this symmetry,

the derivation of an optimality system for configuration C1 is straightforward. For configuration C2 however, this is not directly possible since for $q \in Q = L^2(I)^N$, the right-hand sides $B_\mathrm{f} q$ and $B_\mathrm{s} q$ do not fulfill the prerequisites of Theorem 3.3. Therefore, an additional approximation step will be necessary.

Control Distributed in Space (Configuration C1)

Here, the control $q \in Q = (L^2(\Omega)^d)^N$ acts as volume force through the linear operators B_f and B_s, as described in configuration C1. Since in this case it holds $B_\mathrm{f} q \in H^1(I; H_\mathrm{f})$ and $B_\mathrm{s} q \in H^1(I; H_\mathrm{s})$, the weak formulation (3.20) is applicable for $f_\mathrm{f} = B_\mathrm{f} q$ and $f_\mathrm{s} = B_\mathrm{s} q$ by Theorem 3.3. For the derivative of the reduced functional given by (3.28), we directly obtain the following representation:

Lemma 3.8. *Let the initial data u_0, u_1, and v_0 satisfy Assumption 1 and let the desired states fulfill $v_\mathrm{d} \in H^1(I; V_\mathrm{f}^*) \cap L^2(I; H_\mathrm{f})$ and $u_\mathrm{d} \in L^2(I; H_\mathrm{s})$. Let for given $q \in Q$ the triple (v, u, p) be the solution of (3.20) with $f_\mathrm{f} = B_\mathrm{f} q$, $f_\mathrm{s} = B_\mathrm{s} q$, and $g = 0$ guaranteed by Theorem 3.3. Furthermore, let (z^v, z^u, z^p) be the solution of the adjoint equation (3.31) guaranteed by Theorem 3.7. Then, the directional derivative of the reduced cost functional at q in direction $\delta q \in Q$ is given by*

$$j'(q)(\delta q) = \sum_{i=1}^{N} \left[(\!(g_\mathrm{f}^i \delta q^i, z^v)\!)_\mathrm{f} + (\!(g_\mathrm{s}^i \delta q^i, z^v)\!)_\mathrm{s} + \alpha(q^i, \delta q^i) \right].$$

Proof. By Theorem 3.3, the control to state map G can be understood as mapping from Q to $L^2(I; V) \times L^2(I; V_\mathrm{s}) \times L^2(I; L_\mathrm{f})$. Similar to the proof of Lemma 3.5, let $(\hat{v}, \hat{u}, \hat{p})$ be the solution of (3.20) for $f_\mathrm{f} = f_\mathrm{s} = g = 0$. Furthermore, we denote by $G_0 \colon Q \to L^2(I; H_\mathrm{f}) \times L^2(I; H_\mathrm{s}) \times L^2(I; L_\mathrm{f})$ the linear part G given by (3.20) for zero initial data and $f_\mathrm{f} = B_\mathrm{f} q$, $f_\mathrm{s} = B_\mathrm{s} q$, $g = 0$. Then, G can be written for $q \in Q$ as

$$(v(q), u(q), p(q)) = Gq = (\hat{v}, \hat{u}, \hat{p}) + G_0 q.$$

Hence, we get directly

$$j'(q)(\delta q) = \gamma_\mathrm{f} (\!(v - v_\mathrm{d}, \delta v)\!)_\mathrm{f} + \gamma_\mathrm{s} (\!(u - u_\mathrm{d}, \delta u)\!)_\mathrm{s} + \alpha \sum_{i=1}^{N} (q^i, \delta q^i) \tag{3.33}$$

for all $\delta q \in Q$, where $(\delta v, \delta u, \delta p) = G_0 \delta q$.

By construction, $(\delta v, \delta u, \delta p)$ solves for the right-hand sides $f_\mathrm{f} = B_\mathrm{f} \delta q$, $f_\mathrm{s} = B_\mathrm{s} \delta q$, $g = 0$, and for zero initial data the equation (3.20). Thus, we obtain by testing this equation with $(\varphi, \psi, \xi) = (z^v, z^u, z^p) \in L^2(I; V) \times L^2(I; V_\mathrm{s}) \times L^2(I; L_\mathrm{f})$ the identity

$$\begin{aligned}
\rho_\mathrm{f} (\!(\delta v_t, z^v)\!)_\mathrm{f} - (\!(\delta p, \operatorname{div} z^v)\!)_\mathrm{f} + \nu_\mathrm{f} (\!(\nabla \delta v, \nabla z^v)\!)_\mathrm{f} & \\
+ \rho_\mathrm{s} (\!(\delta v_t, z^v)\!)_\mathrm{s} + \mu_\mathrm{s} (\!(\nabla \delta u, \nabla z^v)\!)_\mathrm{s} &= \sum_{i=1}^{N} \left[(\!(g_\mathrm{f}^i \delta q^i, z^v)\!)_\mathrm{f} + (\!(g_\mathrm{s}^i \delta q^i, z^v)\!)_\mathrm{s} \right], \\
\mu_\mathrm{s} (\!(\nabla \delta v, \nabla z^u)\!)_\mathrm{s} - \mu_\mathrm{s} (\!(\nabla \delta u_t, \nabla z^u)\!)_\mathrm{s} &= 0, \\
(\!(z^p, \operatorname{div} \delta v)\!)_\mathrm{f} &= 0.
\end{aligned}$$

Testing with $(\varphi, \psi, \xi) = (\delta v, \delta u, \delta p) \in L^2(I;V) \times L^2(I;V_\mathrm{s}) \times L^2(I;L_\mathrm{f})$ in the adjoint equation (3.31) yields

$$\begin{aligned} -\rho_\mathrm{f}(\!(\delta v, \partial_t z^v)\!)_\mathrm{f} + \nu_\mathrm{f}(\!(\nabla \delta v, \nabla z^v)\!)_\mathrm{f} + (\!(z^p, \operatorname{div} \delta v)\!)_\mathrm{f} & \\ -\rho_\mathrm{s}(\!(\delta v, \partial_t z^v)\!)_\mathrm{s} + \mu_\mathrm{s}(\!(\nabla \delta v, \nabla z^u)\!)_\mathrm{s} &= \gamma_\mathrm{f}(\!(v - v_\mathrm{d}, \delta v)\!)_\mathrm{f}, \\ -(\!(\delta p, \operatorname{div} z^v)\!)_\mathrm{f} &= 0, \\ \mu_\mathrm{s}(\!(\nabla \delta u, \nabla z^v)\!)_\mathrm{s} + \mu_\mathrm{s}(\!(\nabla \delta u, \nabla \partial_t z^u)\!)_\mathrm{s} &= \gamma_\mathrm{s}(\!(u - u_\mathrm{d}, \delta u)\!)_\mathrm{s}. \end{aligned}$$

As the adjoint solution (z^v, z^u, z^p) has zero initial conditions at $t = T$, and as $(\delta v, \delta u, \delta p)$ has zero initial conditions at $t = 0$, the boundary terms vanish when using integration by parts in time. If we insert the equations into each other, then we obtain for any $\delta q \in Q$

$$\gamma_\mathrm{f}(\!(v - v_\mathrm{d}, \delta v)\!)_\mathrm{f} + \gamma_\mathrm{s}(\!(u - u_\mathrm{d}, \delta u)\!)_\mathrm{s} = \sum_{i=1}^{N} \left[(\!(g_\mathrm{f}^i \delta q^i, z^v)\!)_\mathrm{f} + (\!(g_\mathrm{s}^i \delta q^i, z^v)\!)_\mathrm{s}\right].$$

Together with (3.33), this implies the assertion. □

Combining the condition (3.32) and Lemma 3.8 implies the following representation and regularity for the optimal control $\bar{q}$ in terms of the pointwise projection $\mathrm{P}_{\mathrm{Q_{ad}}}$ on the admissible set Q_ad, given by

$$\mathrm{P}_{\mathrm{Q_{ad}}} \colon L^2(\Omega)^d \to L^2(\Omega)^d, \quad \mathrm{P}_{\mathrm{Q_{ad}}}(r)(x) := \max(q_a, \min(q_b, r(x)))$$

for almost all $x \in \Omega$, where the projection has to be applied componentwise for $r \in L^2(\Omega)^d$.

Lemma 3.9. *Let the assumptions of Lemma 3.8 be fulfilled. Then, for configuration* C1 *the optimal solution $\bar{q} \in Q_{ad}$ of the considered optimal control problem* (3.29) *fulfills for $i = 1, 2, \ldots, N$:*

$$\bar{q}^i\big|_{\Omega_\mathrm{f}} = \mathrm{P}_{\mathrm{Q_{ad}}}\left(-\frac{1}{\alpha}\int_I g_\mathrm{f}^i(t) z^v(t,\cdot)\ dt\right), \quad \bar{q}^i\big|_{\Omega_\mathrm{s}} = \mathrm{P}_{\mathrm{Q_{ad}}}\left(-\frac{1}{\alpha}\int_I g_\mathrm{s}^i(t) z^v(t,\cdot)\ dt\right).$$

Thus, for the optimal control holds $\bar{q}\big|_{\Omega_\mathrm{f}} \in (H^1(\Omega_\mathrm{f})^d)^N$ and $\bar{q}\big|_{\Omega_\mathrm{s}} \in (H^1(\Omega_\mathrm{s})^d)^N$.

Proof. The necessary optimality condition (3.32) can be written as

$$\left(\int_I g_\mathrm{f}^i z^v\ \mathrm{d}t, \delta q^i - \bar{q}^i\right)_\mathrm{f} + \left(\int_I g_\mathrm{s}^i z^v\ \mathrm{d}t, \delta q^i - \bar{q}^i\right)_\mathrm{s} + \alpha(\bar{q}^i, \delta q^i - \bar{q}^i) \geq 0 \quad \forall \delta q^i \in Q_\mathrm{ad}.$$

Using the projection $\mathrm{P}_{\mathrm{Q_{ad}}}$, this can be expressed for $i = 1, 2, \ldots, N$ as

$$\bar{q}^i\big|_{\Omega_\mathrm{f}} = \mathrm{P}_{\mathrm{Q_{ad}}}\left(-\frac{1}{\alpha}\int_I g_\mathrm{f}^i(t) z^v(t,\cdot)\ \mathrm{d}t\right), \quad \bar{q}^i\big|_{\Omega_\mathrm{s}} = \mathrm{P}_{\mathrm{Q_{ad}}}\left(-\frac{1}{\alpha}\int_I g_\mathrm{s}^i(t) z^v(t,\cdot)\ \mathrm{d}t\right).$$

Theorem 3.7 ensures $z^v \in L^2(I;V)$, and together with

$$\|\mathrm{P}_{\mathrm{Q_{ad}}}(r)\|_{H^1(\Omega)} \leq \|r\|_{H^1(\Omega)}$$

we conclude that $\bar{q}^i\big|_{\Omega_\mathrm{f}} \in H^1(\Omega_\mathrm{f})^d$ and $\bar{q}^i\big|_{\Omega_\mathrm{s}} \in H^1(\Omega_\mathrm{s})^d$. □

The optimal solution $\bar{q} \in Q_{\mathrm{ad}}$ solves the optimality system presented in the following theorem:

Theorem 3.10. *Let the initial data u_0, u_1, and v_0 satisfy Assumption 1. Furthermore, let $v_{\mathrm{d}} \in H^1(I;V_{\mathrm{f}}^*) \cap L^2(I;H_{\mathrm{f}})$ and $u_{\mathrm{d}} \in L^2(I;H_{\mathrm{s}})$. Then, the optimal solution $\bar{q} \in Q_{ad}$ of optimal control problem* (3.29) *for configuration* C1 *fulfills $\bar{q}\big|_{\Omega_{\mathrm{f}}} \in (H^1(\Omega_{\mathrm{f}})^d)^N$ and $\bar{q}\big|_{\Omega_{\mathrm{s}}} \in (H^1(\Omega_{\mathrm{s}})^d)^N$ as well as the following necessary optimality condition:*

1. *The optimal state $(\bar{v},\bar{u},\bar{p}) = (v(\bar{q}),u(\bar{q}),p(\bar{q}))$ solves*

$$\begin{aligned}
\rho_{\mathrm{f}}(\!(\partial_t \bar{v},\varphi)\!)_{\mathrm{f}} - (\!(\bar{p},\operatorname{div}\varphi)\!)_{\mathrm{f}} + \nu_{\mathrm{f}}(\!(\nabla\bar{v},\nabla\varphi)\!)_{\mathrm{f}} & \\
+\rho_{\mathrm{s}}(\!(\partial_t \bar{v},\varphi)\!)_{\mathrm{s}} + \mu_{\mathrm{s}}(\!(\nabla\bar{u},\nabla\varphi)\!)_{\mathrm{f}} = (\!(B_{\mathrm{f}}\bar{q},\varphi)\!)_{\mathrm{f}} + (\!(B_{\mathrm{s}}\bar{q},\varphi)\!)_{\mathrm{s}} \quad & \forall\varphi \in L^2(I;V),\\
\mu_{\mathrm{s}}(\!(\nabla\bar{v},\nabla\psi)\!)_{\mathrm{s}} - \mu_{\mathrm{s}}(\!(\nabla\partial_t\bar{u},\nabla\psi)\!)_{\mathrm{s}} = 0 \quad & \forall\psi \in L^2(I;V_{\mathrm{s}}),\\
(\!(\xi,\operatorname{div}\bar{v})\!)_{\mathrm{f}} = 0 \quad & \forall\xi \in L^2(I;L_{\mathrm{f}}).
\end{aligned}$$

2. *The optimal adjoint $(\bar{z}^v,\bar{z}^u,\bar{z}^p) = (z^v(\bar{q}),z^u(\bar{q}),z^p(\bar{q}))$ solves*

$$\begin{aligned}
-\rho_{\mathrm{f}}(\!(\varphi,\partial_t\bar{z}^v)\!)_{\mathrm{f}} + \nu_{\mathrm{f}}(\!(\nabla\varphi,\nabla\bar{z}^v)\!)_{\mathrm{f}} + (\!(\bar{z}^p,\operatorname{div}\varphi)\!)_{\mathrm{f}} & \\
-\rho_{\mathrm{s}}(\!(\varphi,\partial_t\bar{z}^v)\!)_{\mathrm{s}} + \mu_{\mathrm{s}}(\!(\nabla\varphi,\nabla\bar{z}^u)\!)_{\mathrm{s}} = \gamma_{\mathrm{f}}(\!(\bar{v}-v_{\mathrm{d}},\varphi)\!)_{\mathrm{f}} \quad & \forall\varphi \in L^2(I;V),\\
\mu_{\mathrm{s}}(\!(\nabla\psi,\nabla\bar{z}^v)\!)_{\mathrm{s}} + \mu_{\mathrm{s}}(\!(\nabla\psi,\nabla\partial_t\bar{z}^u)\!)_{\mathrm{s}} = \gamma_{\mathrm{s}}(\!(\bar{u}-u_{\mathrm{d}},\psi)\!)_{\mathrm{s}} \quad & \forall\psi \in L^2(I;V_{\mathrm{s}}),\\
-(\!(\xi,\operatorname{div}\bar{z}^v)\!)_{\mathrm{f}} = 0 \quad & \forall\xi \in L^2(I;L_{\mathrm{f}}).
\end{aligned}$$

3. *It holds for $i = 1,2,\ldots,N$ that*

$$\bar{q}^i\big|_{\Omega_{\mathrm{f}}} = \mathrm{P}_{\mathrm{Q_{ad}}}\left(-\frac{1}{\alpha}\int_I g_{\mathrm{f}}^i(t)\bar{z}^v(t,\cdot)\;dt\right), \quad \bar{q}^i\big|_{\Omega_{\mathrm{s}}} = \mathrm{P}_{\mathrm{Q_{ad}}}\left(-\frac{1}{\alpha}\int_I g_{\mathrm{s}}^i(t)\bar{z}^v(t,\cdot)\;dt\right).$$

Remark 3.6. Thereby, the optimal state variable $(\bar{v},\bar{u},\bar{p}) = (v(\bar{q}),u(\bar{q}),p(\bar{q}))$ and adjoint variable $(\bar{z}^v,\bar{z}^u,\bar{z}^p) = (z^v(\bar{q}),z^u(\bar{q}),z^p(\bar{q}))$ fulfill the a priori estimates in Theorem 3.3 and Theorem 3.7.

Time-dependent Control (Configuration C2)

In the following, the control $q \in Q = (L^2(I))^N$ is controlling the volume force through the linear operators B_{f} and B_{s} described in configuration C2. As Theorem 3.3 does not guarantee existence of a unique solution of (3.20) for a right-hand side $f_{\mathrm{f}} = B_{\mathrm{f}}q \in L^2(I;H_{\mathrm{f}})$ and $f_{\mathrm{s}} = B_{\mathrm{s}}q \in L^2(I;H_{\mathrm{s}})$, we cannot directly proceed as for Configuration 1. Therefore, we will make use of a smooth sequence in $Q \cap (H^1(I))^N$ converging against the optimal solution. For smooth controls, the symmetric formulation (3.20) can be utilized, and a priori estimates for the adjoint then lead to higher regularity also for the limit. Then, we are able to derive the optimality system similar to the configuration C1.

Lemma 3.11. *Let the initial data u_0, u_1, and v_0 satisfy Assumption 1. Furthermore, let $v_{\mathrm{d}} \in H^1(I;V_{\mathrm{f}}^*) \cap L^2(I;H_{\mathrm{f}})$ and $u_{\mathrm{d}} \in L^2(I;H_{\mathrm{s}})$. Additionally, for a given control $q \in Q \cap (H^1(I))^N$, let the triple (v,u,p) be the solution of* (3.20) *with $f_{\mathrm{f}} = B_{\mathrm{f}}q$, $f_{\mathrm{s}} = B_{\mathrm{s}}q$, and $g = 0$ guaranteed by Theorem 3.3. Furthermore, let (z^v,z^u,z^p) be the solution of the adjoint equation* (3.31)

guaranteed by Theorem 3.7. Then, the directional derivative of the reduced cost functional j at q in direction $\delta q \in Q$ is given by

$$j'(q)(\delta q) = \sum_{i=1}^{N} \left[(\!(h^i \delta q^i, z^v)\!) + \alpha \int_I q^i \delta q^i \ dt \right].$$

Proof. Since for $q \in Q \cap (H^1(I))^N$ it holds $B_{\mathrm{f}} q \in H^1(I; H_{\mathrm{f}})$ and $B_{\mathrm{s}} q \in H^1(I; H_{\mathrm{s}})$, we proceed as in Lemma 3.8 to obtain

$$j'(q)(\delta q) = \sum_{i=1}^{N} \left[(\!(h^i \delta q^i, z^v)\!) + \alpha \int_I q^i \delta q^i \ \mathrm{d}t \right]$$

for all $\delta q \in Q \cap (H^1(I))^N$. By the density of $(H^1(I))^N$ in $(L^2(I))^N$ with respect to the $(L^2(I))^N$ topology, we obtain the assertion. □

In the next lemma, we prove that the optimal control $\bar{q}$ lies in $Q \cap (H^1(I))^N$ such that the representation derived in Lemma 3.11 is also valid for $\bar{q}$. Therefore, we will introduce also for configuration C2 the pointwise projection $\mathrm{P}_{\mathrm{Q_{ad}}}$ on the admissible set Q_{ad} given here by

$$\mathrm{P}_{\mathrm{Q_{ad}}} \colon L^2(I) \to L^2(I), \quad \mathrm{P}_{\mathrm{Q_{ad}}}(r)(t) := \max(q_a, \min(q_b, r(t))) \quad \text{for almost all } t \in I.$$

Lemma 3.12. *Let the assumptions of Lemma 3.11 be fulfilled. Then, the optimal solution $\bar{q} \in Q_{ad}$ of the considered optimal control problem* (3.29) *for configuration* C2 *lies in $(H^1(I))^N$.*

Proof. Let $\bar{q} \in Q_{ad}$ be the optimal solution. We consider a smooth sequence (q_n) with $q_n \in Q \cap (H^1(I))^N$ and $q_n \to \bar{q}$ in Q. As in the proof of Lemma 3.8, according to Theorem 3.3, we have that $(v_n, u_n, p_n) = G q_n$ solves (3.20) with right-hand sides $f_{\mathrm{f}} = B_{\mathrm{f}} q_n$, $f_{\mathrm{s}} = B_{\mathrm{s}} q_n$, and $g = 0$. The velocity and displacement have at least the regularities $v_n \in H^1(I; H_{\mathrm{f}})$ and $u_n \in L^2(I; V_{\mathrm{s}})$. Therefore, Theorem 3.7 guarantees the existence of a unique adjoint solution (z_n^v, z_n^u, z_n^p) of (3.31) with $v_n - v_{\mathrm{d}}$ and $u_n - u_{\mathrm{d}}$ in the right-hand side. By Lemma 3.11, we have

$$j'(q_n)(\delta q) = \sum_{i=1}^{N} \left[(\!(h_{\mathrm{f}}^i \delta q^i, z_n^v)\!) + \alpha \int_I q_n^i \delta q^i \ \mathrm{d}t \right]$$

for all $\delta q \in Q$. Due to estimate a) in Theorem 3.3, the linearity of (3.20), and the boundedness of h^i in V_{div}, we get for $m, n \in \mathbb{N}$ the estimate

$$\begin{aligned}
&\|v_n - v_m\|^2_{L^2(I;H^1(\Omega_{\mathrm{f}}))} + \|u_n - u_m\|^2_{L^2(I;H^1(\Omega_{\mathrm{s}}))} \\
&\quad \le C \left[\sum_{i=1}^{N} \|h_{\mathrm{f}}^i (q_n^i - q_m^i)\|^2_{L^2(I;V_{\mathrm{f}}^*)} + \sum_{i=1}^{N} \|h_{\mathrm{s}}^i (q_n^i - q_m^i)\|^2_{L^2(I;L^2(\Omega_{\mathrm{s}}))} \right] \\
&\quad \le C \sum_{i=1}^{N} \|q_n^i - q_m^i\|^2_{L^2(I)}.
\end{aligned}$$

Furthermore, due to estimate a) in Theorem 3.7, the adjoint variables fulfill the estimate

$$\begin{aligned}
&\|z_n^v - z_m^v\|^2_{L^\infty(I;L^2(\Omega))} + \|z_n^v - z_m^v\|^2_{L^2(I;H^1(\Omega_\mathrm{f}))} + \|z_n^u - z_m^u\|^2_{L^2(I;H^1(\Omega_\mathrm{s}))} \\
&\quad \le C\left[\sum_{i=1}^N \|h_\mathrm{f}^i(q_n^i - q_m^i)\|^2_{L^2(I;V_\mathrm{f}^*)} + \sum_{i=1}^N \|h_\mathrm{s}^i(q_n^i - q_m^i)\|^2_{L^2(I;L^2(\Omega_\mathrm{s}))}\right] \\
&\quad \le C\sum_{i=1}^N \|q_n^i - q_m^i\|^2_{L^2(I)}.
\end{aligned}$$

If we consider in (3.31) test functions $\varphi \in L^2(I;V_\mathrm{div})$ that are divergence-free in the fluid domain Ω_f, we get the estimate

$$\begin{aligned}
\|\partial_t z_n^v - \partial_t z_m^v\|^2_{L^2(I;V_\mathrm{div}^*)} \le C\big[&\|z_n^v - z_m^v\|^2_{L^2(I;H^1(\Omega_\mathrm{f}))} + \|z_n^u - z_m^u\|^2_{L^2(I;H^1(\Omega_\mathrm{s}))} \\
&+ \|v_n - v_m\|_{L^2(I;L^2(\Omega_\mathrm{f}))}\big].
\end{aligned}$$

By combining the above estimates, we derive for the adjoint $z_n^v - z_m^v$ the bound

$$\|\partial_t z_n^v - \partial_t z_m^v\|^2_{L^2(I;V_\mathrm{div}^*)} + \|z_n^v - z_m^v\|^2_{L^\infty(I;L^2(\Omega))} \le C\sum_{i=1}^N \|q_n^i - q_m^i\|^2_{L^2(I)}.$$

As $q_n \to \bar{q}$ in $(L^2(I))^N$, it holds $\|q_n^i - q_m^i\|^2_{L^2(I)} \to 0$ for $m,n \to \infty$. Thus, z_n^v is a Cauchy sequence in $H^1(I;V_\mathrm{div}^*) \cap L^\infty(I;H)$, and therefore there exists the limit $\tilde{z}^v \in H^1(I;V_\mathrm{div}^*) \cap L^\infty(I;H)$ such that

$$z_n^v \to \tilde{z}^v \quad \text{in} \quad H^1(I;V_\mathrm{div}^*) \cap L^\infty(I;H).$$

Since we assumed $h^i \in V_\mathrm{div}$ and $\delta q^i \in L^2(I)$, the product fulfills $h^i\delta q^i \in L^2(I;V_\mathrm{div})$. This implies, due to $L^2(I;V_\mathrm{div}) \hookrightarrow L^2(I;H)$ and due to the convergence of z_n^v in $L^\infty(I;H)$, that

$$j'(q_n)(\delta q) = \sum_{i=1}^N \left[\langle\!\langle h^i\delta q^i, z_n^v\rangle\!\rangle + \alpha\int_I q_n^i\delta q^i \,\mathrm{d}t\right] \to \sum_{i=1}^N \left[\langle\!\langle h^i\delta q^i, \tilde{z}^v\rangle\!\rangle + \alpha\int_I \bar{q}^i\delta q^i \,\mathrm{d}t\right].$$

In addition, the directional derivative of the reduced cost functional $j'(\cdot)(\delta q)$ is continuous as the control to state mapping $G\colon Q \to L^2(I;H_\mathrm{f}) \times L^2(I;H_\mathrm{s})$ is affine-linear and continuous accordingly to Lemma 3.5. Therefore, the convergence $q_n \to \bar{q}$ in Q implies in addition $j'(q_n)(\delta q) \to j'(\bar{q})(\delta q)$ and we obtain the identity

$$j'(\bar{q})(\delta q) = \sum_{i=1}^N \left[\langle\!\langle h^i\delta q^i, \tilde{z}^v\rangle\!\rangle + \alpha\int_I \bar{q}^i\delta q^i \,\mathrm{d}t\right].$$

As the optimal solution $\bar{q}$ fulfills the necessary optimality condition (3.32), we get the optimality condition

$$\sum_{i=1}^N \left[\langle\!\langle h^i(\delta q^i - \bar{q}^i), \tilde{z}^v\rangle\!\rangle + \alpha\int_I \bar{q}^i(\delta q^i - \bar{q}^i) \,\mathrm{d}t\right] \ge 0 \quad \forall \delta q \in Q_{ad}.$$

Using the projection $\mathrm{P}_{\mathrm{Q_{ad}}}$ on the admissible set Q_ad, this can be expressed as

$$\bar{q}^i = \mathrm{P}_{\mathrm{Q_{ad}}}(r^i) \quad \text{with} \quad r^i = -\frac{1}{\alpha}\int_\Omega h^i(x)\tilde{z}^v(\cdot,x) \,\mathrm{d}x, \quad i = 1,2,\ldots,N.$$

The time regularity of the limit $\tilde{z}^v \in H^1(I; V^*_{\mathrm{div}})$ and the assumed regularity of $h \in V_{\mathrm{div}}$ imply

$$\partial_t r^i(t) = -\frac{1}{\alpha}\langle h^i, \partial_t \tilde{z}^v(t,\cdot)\rangle_{V_{\mathrm{div}}\times V^*_{\mathrm{div}}}, \quad i = 1, 2, \ldots, N$$

for almost all $t \in I$, and consequently that $r^i \in H^1(I)$. Using the stability of the projection

$$\|\mathrm{P}_{\mathrm{Q_{ad}}}(r)\|_{H^1(I)} \leq \|r\|_{H^1(I)},$$

we obtain the asserted regularity $\bar{q} \in (H^1(I))^N$. □

Then, the optimal solution $\bar{q} \in Q_{ad}$ of the considered optimal control problem in configuration C2 fulfills the following theorem:

Theorem 3.13. *Let the initial data u_0, u_1, and v_0 satisfy Assumption 1. Furthermore, let the desired states fulfill $v_{\mathrm{d}} \in H^1(I; V_{\mathrm{f}}^*) \cap L^2(I; H_{\mathrm{f}})$ and $u_{\mathrm{d}} \in L^2(I; H_{\mathrm{s}})$. Then, the optimal solution $\bar{q} \in Q_{ad}$ of the considered optimal control problem* (3.29) *for configuration* C2 *fulfills $\bar{q} \in (H^1(I))^N$ and the following necessary optimality condition:*

1. The optimal state $(\bar{v}, \bar{u}, \bar{p}) = (v(\bar{q}), u(\bar{q}), p(\bar{q}))$ solves

$$\begin{aligned}
\rho_{\mathrm{f}}(\!(\partial_t \bar{v}, \varphi)\!)_{\mathrm{f}} - (\!(\bar{p}, \operatorname{div}\varphi)\!)_{\mathrm{f}} + \nu_{\mathrm{f}}(\!(\nabla\bar{v}, \nabla\varphi)\!)_{\mathrm{f}} & \\
+\rho_{\mathrm{s}}(\!(\partial_t \bar{v}, \varphi)\!)_{\mathrm{s}} + \mu_{\mathrm{s}}(\!(\nabla\bar{u}, \nabla\varphi)\!)_{\mathrm{f}} = (\!(B_{\mathrm{f}}\bar{q}, \varphi)\!)_{\mathrm{f}} + (\!(B_{\mathrm{s}}\bar{q}, \varphi)\!)_{\mathrm{s}} & \quad \forall \varphi \in L^2(I; V), \\
\mu_{\mathrm{s}}(\!(\nabla\bar{v}, \nabla\psi)\!)_{\mathrm{s}} - \mu_{\mathrm{s}}(\!(\nabla\partial_t\bar{u}, \nabla\psi)\!)_{\mathrm{s}} = 0 & \quad \forall \psi \in L^2(I; V_{\mathrm{s}}), \\
(\!(\xi, \operatorname{div}\bar{v})\!)_{\mathrm{f}} = 0 & \quad \forall \xi \in L^2(I; L_{\mathrm{f}}).
\end{aligned}$$

2. The optimal adjoint $(\bar{z}^v, \bar{z}^u, \bar{z}^p) = (z^v(\bar{q}), z^u(\bar{q}), z^p(\bar{q}))$ solves

$$\begin{aligned}
-\rho_{\mathrm{f}}(\!(\varphi, \partial_t \bar{z}^v)\!)_{\mathrm{f}} + \nu_{\mathrm{f}}(\!(\nabla\varphi, \nabla\bar{z}^v)\!)_{\mathrm{f}} + (\!(\bar{z}^p, \operatorname{div}\varphi)\!)_{\mathrm{f}} & \\
-\rho_{\mathrm{s}}(\!(\varphi, \partial_t \bar{z}^v)\!)_{\mathrm{s}} + \mu_{\mathrm{s}}(\!(\nabla\varphi, \nabla\bar{z}^u)\!)_{\mathrm{s}} = \gamma_{\mathrm{f}}(\!(\bar{v} - v_{\mathrm{d}}, \varphi)\!)_{\mathrm{f}} & \quad \forall \varphi \in L^2(I; V), \\
\mu_{\mathrm{s}}(\!(\nabla\psi, \nabla\bar{z}^v)\!)_{\mathrm{s}} + \mu_{\mathrm{s}}(\!(\nabla\psi, \nabla\partial_t\bar{z}^u)\!)_{\mathrm{s}} = \gamma_{\mathrm{s}}(\!(\bar{u} - u_{\mathrm{d}}, \psi)\!)_{\mathrm{s}} & \quad \forall \psi \in L^2(I; V_{\mathrm{s}}), \\
-(\!(\xi, \operatorname{div}\bar{z}^v)\!)_{\mathrm{f}} = 0 & \quad \forall \xi \in L^2(I; L_{\mathrm{f}}).
\end{aligned}$$

3. It holds for $i = 1, 2, \ldots, N$ that

$$\bar{q}^i = \mathrm{P}_{\mathrm{Q_{ad}}}\left(-\frac{1}{\alpha}\int_\Omega h^i(x)\bar{z}^v(\cdot, x)\ dx\right).$$

Proof. As $\bar{q} \in Q \cap (H^1(I))^N$, one can choose in the proof of Lemma 3.12 the sequence $q_n = \bar{q}$. This immediately implies that $\tilde{z}^v = \bar{z}^v$.

□

Remark 3.7. Thereby, the optimal state variable $(\bar{v}, \bar{u}, \bar{p}) = (v(\bar{q}), u(\bar{q}), p(\bar{q}))$ and adjoint variable $(\bar{z}^v, \bar{z}^u, \bar{z}^p) = (z^v(\bar{q}), z^u(\bar{q}), z^p(\bar{q}))$ fulfill the a priori estimates in Theorem 3.3 and Theorem 3.7.

4. Optimal Control of Nonlinear Fluid-Structure Interaction

For several applications in hemodynamics and aerodynamics, the fluid flows in or around an elastic object. Due to the interaction between solid and fluid, the models cannot be regarded separately. In addition, due to the large forces, solid deformations can get quite large. Then, the solid mechanics can not be modeled by a linear Lamé system anymore and we have to take the movement of the fluid domain into account. As the system is highly dynamical, it is very difficult to predict how the system will react on changes in the boundary conditions and of material parameters. This motivates to consider an optimal control problem to steer the solution against a desired state or to choose such a parameter that the simulation correlates with measurements.

In this chapter, we are going to derive a monolithic formulation of the nonlinear fluid-structure interaction problem, in Section 4.1, using a nonlinear elastodynamics model for the solid and the Arbitrary Lagrangian-Eulerian formulation to transform the Navier-Stokes equations on a reference domain. After a short discussion about theoretical results for the nonlinear FSI model, in Section 4.2, we state the considered optimal control problem, in Section 4.3, and derive formally an optimality system.

4.1. A Nonlinear Fluid-Structure Interaction Problem

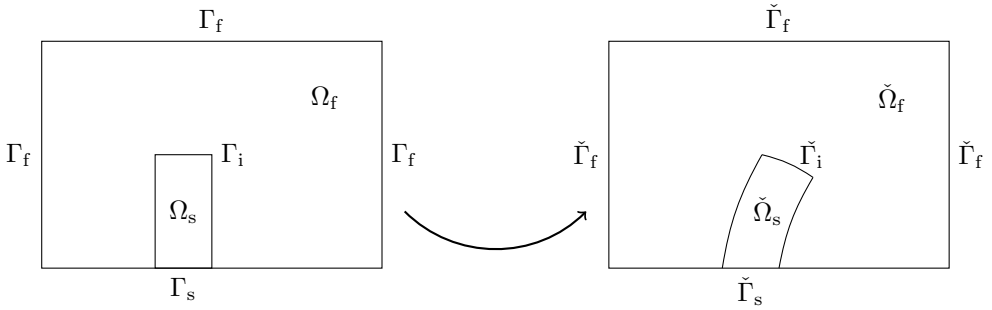

Figure 4.1.: An exemplary reference domain Ω and the transformed current domain $\check{\Omega}$

The domain $\Omega \subset \mathbb{R}^d$ with $d = 2, 3$ is separated as in Chapter 3 in two disjoint Lipschitz sub-domains Ω_s and Ω_f. For the boundaries we use again the same notation (see Section 2). However, in comparison to the linear fluid-structure interaction problem, the interface between fluid and solid is not fixed in this chapter. Therefore, we have to introduce the transformed current domain $\check{\Omega}(t)$. In this section, we follow mainly the notation and ideas presented

in [126, 40, 70, 149]. Therein, all variables defined in the reference domain in Lagrangian coordinates are denoted with a "hat" sign. As we are going to follow the ALE-approach, which transforms all variables on a reference domain, we mark all variables on the current domain instead to make the following chapters easier to read. We denote with the "inverse hat" symbol all variables defined on the current domain and all other variables are now defined on the reference domain.

In structure mechanics, it is natural to define all variables on the reference domain. This is called Lagrangian point of view. In fluid mechanics we are not interested in the movement of a single fluid particle but in the behavior of the fluid at a certain fixed point in the current domain. This is called Eulerian point of view. To overcome this mismatch we are going to use an Arbitrary Lagrangian-Eulerian technique, see, e.g., [52, 83, 85, 120, 64, 138, 139], to transform the Navier-Stokes equations to a reference domain. For optimal control this formulation is of advantage. As all variables are defined on the reference domain the linearization of the semi-linear form, needed to derive sensitivity information, can be computed straightforward. However, the ALE method can only tackle minor movements of the elastic structure. To model for example a ball falling in a liquid, alternative approaches such as Eulerian formulations presented in [68, 67, 126] have to be used. Therein, the solid equations are transformed on the moving physical domain.

Remark 4.1. We use the same notation and names for the semi-linear forms as in the Chapter 3.2 for linear fluid-structure interaction. However, due to the nonlinear models and the moving fluid domain the semi-linear forms differ. We nevertheless decided for this abuse of notation such that we only have to regard the nonlinear problem in the following chapters, but immediately it gets clear what we would have to do for the linear problem.

4.1.1. Nonlinear Elastodynamics

Basic principles in continuum mechanics enable to describe the large deformations of the elastic structure in Ω_s. We are going to introduce the Green-Lagrange strain tensor first, then state the equation resulting from conservation of momentum and the Saint-Venant Kirchhof material law as exemplary example for the stress strain relation.

Kinematic

To describe the relation between the solid at initial time and the current deformed solid, we introduce the transformation

$$T_\mathrm{s} : \Omega_\mathrm{s} \times I \to \check{\Omega}_\mathrm{s}(t) \quad \text{with} \quad T_\mathrm{s} : (x,t) \longmapsto \check{x}(x,t)$$

mapping every particle $x \in \Omega_\mathrm{s}$ to its location in the current domain $\check{\Omega}_\mathrm{s}(t)$ at time t. In general, we choose the initial configuration as reference domain Ω_s, but it is also possible to take an arbitrary domain as reference domain. For every particle, we define the displacement $u_\mathrm{s}(x,t) := \check{x}(x,t) - x$ and the material velocity $v_\mathrm{s}(x,t) := \partial_t u_\mathrm{s}(x,t)$. The description of the movement of every particle is called Lagrangian point of view. Deformations of an arbitrary volume in Ω_s can be characterized via the deformation gradient

$$F_\mathrm{s}(x,t) = \mathrm{Id} + \nabla u_\mathrm{s}(x,t).$$

The determinant of the deformation gradient $J_\mathrm{s} = \det(F_\mathrm{s})$ describes the local change of volume and the local length change of a line segment is given by the Green-Lagrange strain tensor

$$E_\mathrm{s} = \frac{1}{2}(F_\mathrm{s}^T F_\mathrm{s} - \mathrm{Id}).$$

Conservation Principle

The momentum has to be conserved over time if no forces are active. Only if boundary forces g_s or volume forces f_s are active, the momentum changes. Via Cauchy's stress theorem we postulate the existence of a stress tensor Σ_s. The second Piola Kirchhoff stress tensor $\Sigma_\mathrm{s}(x,t)n_\mathrm{s}$ denotes the surface tensions on the surface of an imaginary infinitesimally small cube around the point $x \in \Omega_\mathrm{s}$ in the reference configuration. The equation of conservation of momentum then results in

$$\begin{aligned} \rho_\mathrm{s}\partial_{tt}u_\mathrm{s} - \mathrm{div}(F_\mathrm{s}\Sigma_\mathrm{s}) &= \rho_\mathrm{s}f_\mathrm{s} \quad && \text{in } \Omega_\mathrm{s} \times I, \\ F_\mathrm{s}\Sigma_\mathrm{s}n_\mathrm{s} &= g_\mathrm{s} && \text{on } \Gamma_\mathrm{i} \times I. \end{aligned} \tag{4.1}$$

As most elastic materials are compressible, the density in the material can change over time. Thus, we define with ρ_s the density at initial time $t = 0$. Conservation of angular momentum demands the stress tensor to be symmetric.

Material Law

Strain of a material usually results in stress. The exact correlation cannot be derived from physical laws but has to be measured in experiments. A large variety of homogenous and isotropic materials behaves according to the Saint-Venant Kirchhof material law

$$\Sigma_\mathrm{s} = \lambda_\mathrm{s}\,\mathrm{tr}(E_\mathrm{s})\,\mathrm{Id} + 2\mu_\mathrm{s}E_\mathrm{s}.$$

Due to the linear strain stress relation, the material law is very simple and we are going to use it as an exemplary material law in this thesis. The Lamé parameters λ_s and μ_s are given by

$$\lambda_\mathrm{s} := \frac{\mathcal{E}_\mathrm{s}}{2(1+\nu_\mathrm{s})}, \quad \mu_\mathrm{s} := \frac{\mathcal{E}_\mathrm{s}\nu_s}{(1+\nu_\mathrm{s})(1-\nu_\mathrm{s})},$$

whereby the Young's modulus $\mathcal{E}_\mathrm{s}$ describes the stiffness and ν_s the compressibility of a material. For more sophisticated material laws to model blood vessels we refer to the book [66].

Nonlinear Elastodynamics Equation

The nonlinear elastodynamics equation (4.1) is of second order in time. By introducing a material velocity, we can rewrite the equations as system of first order. This will make the coupling of fluid and structure equations easier. In Section 3.1.4, we mainly introduced the velocity variable in an unconventional way to be able to prove existence and uniqueness of the adjoint equation. For the optimal control problem of a nonlinear FSI problem, we are

only going to derive the optimal control problem in a formal way. Therefore, we demand as commonly done in literature:

$$(\!(\partial_t u_\mathrm{s} - v_\mathrm{s}, \psi)\!)_\mathrm{s} = 0 \quad \forall \psi \in L^2(I; H_\mathrm{s}).$$

A further discussion can be found in Section 4.3.2.

By integrating by parts, we obtain the weak nonlinear elastodynamics equation, whereby we have again written the problem on the space-time cylinder as in Chapter 3.

Problem 4.1 (Nonlinear elastodynamics equation)**.** *Given the volume force* $f_\mathrm{s} \in L^2(I; H_\mathrm{s})$ *and the boundary stress* $g_\mathrm{s} \in L^2(I; H^{\frac{1}{2}}(\Gamma_\mathrm{i})^d)$ *at the interface* Γ_i*, find the displacement field* $u_\mathrm{s} \in W_\mathrm{s}^u$ *and solid velocity* $v_\mathrm{s} \in W_\mathrm{s}^v$ *such that* $u_\mathrm{s}(0) = u_{\mathrm{s},0}$*,* $v_\mathrm{s}(0) = v_{\mathrm{s},0}$ *and*

$$\begin{aligned} a^S(u_\mathrm{s}, v_\mathrm{s})(\varphi) &= (\!(\rho_\mathrm{s} f, \varphi)\!)_\mathrm{s} + \langle\!\langle g_\mathrm{s}, \varphi \rangle\!\rangle_\mathrm{i} && \forall \varphi \in L^2(I; V_\mathrm{s}), \\ a^V(u_\mathrm{s}, v_\mathrm{s})(\psi) &= 0 && \forall \psi \in L^2(I; H_\mathrm{s}), \end{aligned}$$

whereby

$$\begin{aligned} a^S(u_\mathrm{s}, v_\mathrm{s})(\varphi) &:= (\!(\rho_\mathrm{s} \partial_t v_\mathrm{s}, \varphi)\!)_\mathrm{s} + (\!(F_\mathrm{s} \Sigma_\mathrm{s}, \nabla \varphi)\!)_\mathrm{s}, \\ a^V(u_\mathrm{s}, v_\mathrm{s})(\psi) &:= (\!(\partial_t u_\mathrm{s} - v_\mathrm{s}, \psi)\!)_\mathrm{s}. \end{aligned} \tag{4.2}$$

Its analysis is difficult and most results only exist for small deformations. We would like to refer to [45] for existence of the stationary solution. For small deformations u_s and very small changes in deformation $\|\nabla u_\mathrm{s}\| \ll 1$, a linear strain tensor can be used. This results in the well analyzed Navier-Lamé system.

4.1.2. Navier-Stokes Equations on a Moving Domain

In fluid mechanics, we are interested in the velocity and pressure of the fluid at a certain points and not in the movement of a single fluid particle. Thereby, we are able to calculate, lift and drag values at the boundary. Therefore, the Eulerian point of view is the method of choice.

Navier-Stokes Equations in Eulerian Framework

In comparison to solid mechanics, not the strain but the temporal variation of strain will be the key quantity to model the internal forces. Therefore we introduce the strain rate

$$\dot{\varepsilon}_\mathrm{f} := \frac{1}{2}(\check{\nabla} \check{v}_\mathrm{f} + \check{\nabla} \check{v}_\mathrm{f}^T) \tag{4.3}$$

defining the rate of length change. Just like in solid mechanics, we postulate the existence of a stress tensor. The Cauchy stress tensor $\check{\sigma}_\mathrm{f} \check{n}$ denotes the internal force on an imaginary surface of an infinitesimally small cube around the point $\check{x} \in \check{\Omega}_\mathrm{f}(t)$. However, in comparison to the definition of the Green-Lagrange strain tensor, the Cauchy stress tensor is defined in the current configuration. Let $\check{f}_\mathrm{f}$ be a volume force on the current domain and $\check{g}_\mathrm{f}$ a boundary

force on parts of the boundary of the current domain $\check{\Omega}_\mathrm{f}(t)$, then the equations of conservation of momentum result in

$$\begin{aligned} \check{\rho}_\mathrm{f}\partial_t\check{v}_\mathrm{f} + \check{\rho}_\mathrm{f}(\check{v}_\mathrm{f}\cdot\check{\nabla})\check{v}_\mathrm{f} - \widetilde{\mathrm{div}}(\check{\sigma}_\mathrm{f}) &= \check{\rho}_\mathrm{f}\check{f}_\mathrm{f} \quad \text{in } \check{\Omega}_\mathrm{f}(t)\times I \\ \check{\sigma}_\mathrm{f}\check{n} &= \check{g}_\mathrm{f} \quad \text{on } \check{\Gamma}_\mathrm{i}(t)\times I. \end{aligned} \tag{4.4}$$

As already for solid mechanics, conservation of angular moment needs the stress tensor $\check{\sigma}$ to be symmetric. In this thesis, we are going to consider only incompressible fluids, which are homogeneous at initial time. Therefore, the equation of mass conservation simplifies to

$$\widetilde{\mathrm{div}}(\check{v}_\mathrm{f}) = 0 \quad \text{in } \check{\Omega}_\mathrm{f}(t)\times I. \tag{4.5}$$

If the velocity field $\check{v}$ is divergence free, mass is going to be conserved.

Material Law

In this thesis, we consider only Navier-Stokes fluids. For these fluids, the stress tensor depends linearly on the strain rate. Due to the incompressibility assumption we have $\mathrm{tr}(\dot{\varepsilon}_\mathrm{f}) = 0$ and the stress tensor simplifies to

$$\check{\sigma}_\mathrm{f} = -\check{p}_\mathrm{f}\,\mathrm{Id} + \check{\rho}_\mathrm{f}\check{\nu}_\mathrm{f}(\check{\nabla}\check{v}_\mathrm{f} + \check{\nabla}\check{v}_\mathrm{f}^T).$$

Thereby, $\check{p}$ denotes the undetermined pressure and $\check{\nu}_\mathrm{f}$ the kinematic viscosity of the fluid.

Arbitrary Lagrangian-Eulerian (ALE) Transformation

We consider configurations where either a fluid flows around an elastic object or flows in an elastic tube. Due to the elastic behavior of the solid, the fluid domain changes over time. In this thesis, we introduce the Arbitrary Lagrangian-Eulerian (ALE) transformation $T_\mathrm{f} : \Omega_\mathrm{f}\times I \to \check{\Omega}_\mathrm{f}(t)$ mapping a reference domain Ω_f on the Eulerian domain $\check{\Omega}_\mathrm{f}(t)$. The ALE-mapping can be expressed by $T_\mathrm{f}(x,t) := x + u_\mathrm{f}(x,t)$, using an artificial displacement field u_f, similar as already for the mapping T_s to describe the kinematics of the elastic domain Ω_s. If T_f is a diffeomorphism on the space-time cylinder in the space $C(I;C^1(\Omega_\mathrm{f}))\cap C^1(I;C(\Omega_\mathrm{f}))$, we can transform the domain integrals in the weak Navier-Stokes equation to the fixed reference domain Ω_f. Thereby, the deformation gradient of the ALE -mapping with $F_\mathrm{f} := \mathrm{Id} + \nabla u_\mathrm{f}$ and its determinant $J_\mathrm{f} := \det(F_\mathrm{f})$ are going to occur. We refer to [126, 40, 70, 149] for more details of the transformation of the space and time derivatives. On the reference domain we define a velocity $v_\mathrm{f}(x,t) := \check{v}_\mathrm{f}(T_\mathrm{f}(x,t),t)$ and a pressure variable $p_\mathrm{f}(x,t) := \check{p}_\mathrm{f}(T_\mathrm{f}(x,t),t)$ for all $x\in\Omega_\mathrm{f}$. Then, we get for example the transformed Cauchy stress tensor

$$\sigma_\mathrm{f} = \rho_\mathrm{f}\nu_\mathrm{f}(\nabla v_\mathrm{f}F_\mathrm{f}^{-1} + F_\mathrm{f}^{-T}\nabla v_\mathrm{f}^T) - p_\mathrm{f}\,\mathrm{Id}\,.$$

As for homogenous fluids, the density $\check{\rho}_\mathrm{f}$ and the kinematic viscosity $\check{\nu}_\mathrm{f}$ are constant, the values $\rho_\mathrm{f} = \check{\rho}_\mathrm{f}$ and $\nu_\mathrm{f} = \check{\nu}_\mathrm{f}$ are equivalent in the current and reference domain.

Problem 4.2 (Navier-Stokes equations in ALE coordinates)**.** *Let $f_\mathrm{f} \in L^2(I;H_\mathrm{f})$ be a volume force and $g_\mathrm{f} \in L^2(I;L^2(\Gamma_\mathrm{i})^d)$. We assume the transformation $u_\mathrm{f} \in C(I;C^1(\Omega_\mathrm{f}))\cap C^1(I;C(\Omega_\mathrm{f}))$ to be given. Then find the velocity $v_\mathrm{f} \in W_\mathrm{f}^v$ and pressure $p_\mathrm{f} \in L^2(I;L_\mathrm{f})$ such that $v_\mathrm{f}(0) = v_{\mathrm{f},0}$ and*

$$\begin{aligned} a^F(v_\mathrm{f},u_\mathrm{f},p_\mathrm{f})(\varphi) + a^\Gamma(u_\mathrm{f},v_\mathrm{f})(\varphi) &= (\!(\rho_\mathrm{f} f_\mathrm{f},\varphi)\!)_\mathrm{f} + \langle\!\langle g_\mathrm{f},\varphi\rangle\!\rangle_\mathrm{i} && \forall \varphi \in L^2(I;V_\mathrm{f}) \\ a^D(v_\mathrm{f},u_\mathrm{f})(\xi) &= 0 && \forall \xi \in L^2(I;L_\mathrm{f}) \end{aligned}$$

where the transformed momentum equation $a^F(\cdot)(\cdot)$ and the incompressibility constraint $a^D(\cdot)(\cdot)$ are defined by:

$$\begin{aligned} a^F(v_\mathrm{f},u_\mathrm{f},p_\mathrm{f})(\varphi) :=& (\!(J_\mathrm{f}\rho_\mathrm{f}\partial_t v_\mathrm{f},\varphi)\!)_\mathrm{f} + (\!(J_\mathrm{f}\rho_\mathrm{f}(F_\mathrm{f}^{-1}(v_\mathrm{f}-\partial_t u_\mathrm{f})\cdot\nabla)v_\mathrm{f},\varphi)\!)_\mathrm{f} \\ &+ (\!(J_\mathrm{f}\sigma_\mathrm{f}F_\mathrm{f}^{-T},\nabla\varphi)\!)_\mathrm{f}, \\ a^D(v_\mathrm{f},u_\mathrm{f})(\xi) :=& (\!(\mathrm{div}(J_\mathrm{f}F_\mathrm{f}^{-1}v_\mathrm{f}),\xi)\!)_\mathrm{f}. \end{aligned} \tag{4.6}$$

Remark 4.2. The volume force f_f is defined on the reference domain here. If we only have the data on the moving domain $\check{\Omega}_\mathrm{f}$, we have to transform the values on the reference domain. Then the right-hand side depends on the transformation u_f.

Only if the transformation $T_\mathrm{f} \in C(I;C^1(\Omega_\mathrm{f})) \cap C^1(I;C(\Omega_\mathrm{f}))$ is a diffeomorphism, then the Navier-Stokes equations in ALE coordinates is equivalent to the Navier-Stokes equation on the moving domain. We only prove this result assuming existence of a smooth solution. We define on the moving domain $\check{\Omega}_\mathrm{f}(t)$ the Sobolev spaces

$$\check{V}_\mathrm{f}(t) := \left\{ \check{\varphi} \in H^1(\check{\Omega}_\mathrm{f}(t))^d \;\middle|\; \check{\varphi} = 0 \text{ on } \check{\Gamma}_\mathrm{f} \right\}, \quad \check{H}_\mathrm{f}(t) := L^2(\check{\Omega}_\mathrm{f}(t))^d$$
$$\text{and} \quad \check{L}_\mathrm{f} := \left\{ \xi \in L^2(\check{\Omega}_\mathrm{f}(t)) \;\middle|\; \int_{\check{\Omega}_\mathrm{f}(t)} \xi \,\mathrm{d}x = 0 \right\}$$

Lemma 4.1. *Let $\Omega_\mathrm{f} \subset \mathbb{R}^d$ with $d = 2,3$ be a smooth domain and $T_\mathrm{f} : \Omega_\mathrm{f} \times I \to \check{\Omega}_\mathrm{f}(t)$ be a $C(I;C^1(\Omega_\mathrm{f})) \cap C^1(I;C(\Omega_\mathrm{f}))$-diffeomorphism. Then for every solution*

$$v \in \left\{ \varphi \;\middle|\; \varphi \in L^2(I;V_\mathrm{f}) \text{ and } \partial_t\varphi \in L^2(I;H_\mathrm{f}) \right\}, \quad p \in L^2(I;L_\mathrm{f})$$

of Problem 4.2, there exist a solution

$$\check{v} \in \left\{ \check{\varphi} \;\middle|\; \check{\varphi} \in L^2(I;\check{V}_\mathrm{f}(t)) \text{ and } \partial_t\check{\varphi} \in L^2(I;\check{H}_\mathrm{f}(t)) \right\}, \quad \check{p} \in L^2(I;\check{L}_\mathrm{f})$$

of the Navier-Stokes equations 4.4 and 4.5 on the moving domain in weak form with $v(x,t) = \check{v}(T_\mathrm{f}(x,t),t))$ and $p(x,t) = \check{p}(T_\mathrm{f}(x,t),t))$ almost everywhere.

Proof. This result directly follows by mapping the two different formulations of the Navier-Stokes equations and by using the equivalence of the solution spaces in Appendix A. □

4.1.3. Mesh Motion Equation

Let the movement of the interface $\check{\Gamma}_\mathrm{i}$ be given by the displacement u_Γ, then we can derive the ALE transformation u_f on the fluid domain Ω_f by computation of an extension of the interface movement u_Γ. As we have seen in Lemma 4.1, the ALE mapping of the reference domain Ω_f on the moving fluid domain $\check{\Omega}_\mathrm{f}$ has to be very smooth. Several strategies have been suggested in literature to calculate the ALE transformation from boundary information. The methods differ with respect to computational cost and regularity of the transformation. In general, the more regularity we would like to achieve the higher the computational cost. We just state four standard methods given for example in [23, 126]. All methods have in common that we have to solve a partial differential equation, but lack the C^1 regularity needed in Lemma 4.1.

Laplace Equation

For a given deformation of the interface $u_\Gamma \in L^2(I; H^{\frac{1}{2}}(\Gamma_\mathrm{i}))$ we calculate the ALE transformation $T_\mathrm{f}(x,t) = x + u_\mathrm{f}(x,t)$ by solving a Laplace problem on the space-time cylinder with $u_\mathrm{f} \in \left\{ \varphi \in L^2(I; V_\mathrm{f}) \;\middle|\; \varphi = u_\Gamma \text{ on } \Gamma_\mathrm{i} \times I \right\}$ and

$$\left(\!\left(\nabla u_\mathrm{f}, \nabla \varphi\right)\!\right)_\mathrm{f} = 0 \quad \forall \varphi \in L^2(I; V_{\mathrm{f},0}).$$

It is well known that even for convex domains we only have $u_\mathrm{f} \in L^2(I; H^2(\Omega_\mathrm{f})^d)$. If we have reentering corners, we must even expect corner singularities.

Elasticity Equation

Instead of the Laplace problem, the authors in [140] suggest to solve an artificial linear elasticity equation with linear strain tensor $\varepsilon_\mathrm{m} := \frac{1}{2}(\nabla u_\mathrm{f} + \nabla u_\mathrm{f}^T)$ and Saint-Venant Kirchhoff material law. Then, the Cauchy stress tensor σ_m is defined by

$$\sigma_\mathrm{m} := \lambda_\mathrm{m} \operatorname{tr}(\varepsilon_\mathrm{m}) \operatorname{Id} + 2\mu_\mathrm{m}\varepsilon_\mathrm{m}. \tag{4.7}$$

The ALE transformation $T_\mathrm{f}(x,t) = x + u_\mathrm{f}(x,t)$ then is given by

Problem 4.3 (Mesh motion equation)**.** *Find* $u_\mathrm{f} \in \left\{ \varphi \in L^2(I; V_\mathrm{f}) \;\middle|\; \varphi = u_\Gamma \text{ on } \Gamma_\mathrm{i} \times I \right\}$ *such that*

$$a^M(u_\mathrm{f})(\psi) := \left(\!\left(\sigma_\mathrm{m}, \nabla \psi\right)\!\right)_\mathrm{f} = 0 \quad \forall \psi \in L^2(I; V_{\mathrm{f},0}). \tag{4.8}$$

The distributed artificial Lamé parameters $\lambda_\mathrm{m} \in L^\infty(\Omega_\mathrm{f})$ and $\mu_\mathrm{m} \in L^\infty(\Omega_\mathrm{f})$ are given by the

$$\lambda_\mathrm{m} := \frac{\mathcal{E}_\mathrm{m}}{2(1+\nu_\mathrm{m})}, \quad \mu_\mathrm{m} := \frac{\mathcal{E}_\mathrm{m}\nu_\mathrm{m}}{(1+\nu_\mathrm{m})(1-\nu_\mathrm{m})}.$$

As the equation has a physical interpretation, we know how the modification of the artificial Lamé parameters can improve the mesh motion. For example, in areas of reentering corners it can be reasonable to choose the Young's modulus $\mathcal{E}_\mathrm{m}$ higher so that these areas are only translated but not sheared or compressed due to the transformation. In addition, numerical experiments have shown that it can be beneficial to choose a negative poisson ratio ν_m.

Nonlinear Elasticity Equation

If the determinant of the ALE transformation $J_\mathrm{f} = \det(F_\mathrm{f})$ converges to zero, the finite element solution algorithm will break down. For a detailed discussion see Section 8.3.3. The authors in [136] suggest to multiply Young's modulus in the stress tensor σ_m in (4.7) with J_f^{-1}, which results in the Lame parameters

$$\lambda_\mathrm{m}(u_\mathrm{f}) := \frac{\mathcal{E}_\mathrm{m}}{2J_\mathrm{f}(1+\nu_\mathrm{m})}, \quad \mu_\mathrm{m}(u_\mathrm{f}) := \frac{\mathcal{E}_\mathrm{m}\nu_\mathrm{m}}{J_\mathrm{f}(1+\nu_\mathrm{m})(1-\nu_\mathrm{m})}.$$

If we use this method, the stiffness automatically will increase in the areas with extreme volume changes and the determinant of the ALE transformation can be bounded away from zero. In comparison to the previous approach, the equation is not linear any more. However, for small deformations, the problem behaves very similar to the linear elasticity equation as extension operator.

Bi-Laplace Equation

Nevertheless, for large displacements the presented extension methods will fail. If we solve instead a bi-Laplace problem, we have more freedom in the choice of boundary conditions. Numerical experiments in [23] have shown that thereby much larger displacements can be tackled, even with constant material parameters. In this approach, we calculate the ALE mapping $u_\mathrm{f} \in \left\{ \varphi \in L^2(I; H_0^1(\Omega_\mathrm{f})^d) \;\middle|\; \varphi = u_\Gamma \text{ on } \Gamma_\mathrm{i} \times I \right\}$ and $w_\mathrm{f} \in L^2(I; H^1(\Omega_\mathrm{f})^d)$ such that

$$\begin{aligned} (\!(\mu_\mathrm{m} \nabla u_\mathrm{f}, \nabla\varphi)\!)_\mathrm{f} &= (\!(w_\mathrm{f}, \varphi)\!)_\mathrm{f} && \forall \varphi \in L^2(I; H_0^1(\Omega_\mathrm{f})^d), \\ (\!(\mu_\mathrm{m} \nabla w_\mathrm{f}, \nabla\psi)\!)_\mathrm{f} &= \langle\!\langle g_\mathrm{m}, \psi \rangle\!\rangle_\mathrm{i} && \forall \psi \in L^2(I; H^1(\Omega_\mathrm{f})^d). \end{aligned}$$

Thereby, we are able to enforce the slope of the extension g_m at the boundary in addition. In comparison to the previous methods no pre-knowledge to choose the optimal Lamé parameters is necessary. However, the computational cost rises immensely, as we have to introduce the additional variable w_f.

4.1.4. Coupling and Boundary Conditions

Coupling Conditions

At the interface Γ_i, we expect that no gaps are able to evolve between solid and fluid. Therefore, the movement of the fluid and solid boundary has to coincide at the interface Γ_i. The displacement of the solid domain u_s has to match the ALE mapping u_f:

$$u_\mathrm{f} = u_\mathrm{s} \text{ on } \Gamma_\mathrm{i} \times I \qquad \text{(geometric coupling condition)}.$$

In addition, we assume a no-slip condition such that the viscid fluid follows the motion of its neighboring solid. Thus the solid and fluid velocities have to coincide at the interface Γ_i, such that

$$v_\mathrm{f} = v_\mathrm{s} \text{ on } \Gamma_\mathrm{i} \times I \qquad \text{(kinematic coupling condition)}.$$

Furthermore, the surface tensions on the fluid boundary and solid boundary have to be equal at the interface. As we have transformed the fluid equation on the reference domain, fluid and solid stresses are given in the same coordinate system and we demand

$$F_{\mathrm{s}}\Sigma_{\mathrm{s}}n = J_{\mathrm{f}}\sigma_{\mathrm{f}}F_{\mathrm{f}}^{-T}n \text{ on } \Gamma_{\mathrm{i}} \times I \qquad \text{(dynamic coupling condition)}.$$

Inflow and Do-Nothing Outflow Condition

If we describe the flow through a channel, we have to enforce boundary conditions at the outflow boundary $\Gamma_{\mathrm{out}} \subset \Gamma_{\mathrm{f}}$. A common practice is to describe no condition at all. For such a "natural" boundary condition, the velocity vectors get deflected and sheer out of line at the outflow boundary. To allow at least Poiseuille flow to leave the domain without deflection, we enforce the do-nothing condition

$$\check{\rho}_{\mathrm{f}}\check{\nu}_{\mathrm{f}}\check{\nabla}\check{v} - \check{p}\,\mathrm{Id} = 0 \quad \text{on } \check{\Gamma}_{\mathrm{out}}.$$

The do-nothing outflow condition works excellent in several configurations. For example, vorticities can leave the domain. In addition, on every straight outflow boundary-line segment Γ_{out}, that is enclosed by no-slip Dirichlet boundaries, it holds $\int_{\Gamma_{\mathrm{out}}} p \, \mathrm{d}\Gamma = 0$. Thereby the do-nothing outflow condition normalizes the pressure. In the ALE configuration this results in the additional boundary term

$$a^{\Gamma}(u_{\mathrm{f}}, v_{\mathrm{f}})(\varphi) := -\langle\!\langle \rho_{\mathrm{f}}\nu_{\mathrm{f}}F_{\mathrm{f}}^{-T}\nabla v_{\mathrm{f}}^{T}, \varphi \rangle\!\rangle_{\Gamma_{\mathrm{out}}}.$$

Zero mean pressure at the outflow condition is not physiologically accurate in bio-medical applications. For the simulation of an arteria the whole remaining arterial network has to be taken into account. We refer to [65] for an overview on alternative boundary conditions used in hemodynamics to model the pulsating pressure at the outflow boundary.

For the flow through a channel as regarded later in Chapter 8 we have to define inflow conditions. We are going to use a parabolic inflow profile enforced by Dirichlet conditions v^D on the inflow boundary $\Gamma_{\mathrm{in}} \subset \Gamma_{\mathrm{f}}$.

If we apply a Dirichlet inflow condition as well as a do-nothing outflow condition, the velocity variable v_f will be an element of the space

$$\tilde{V}_{\mathrm{f}} := \left\{ \varphi \in H^1(\Omega_{\mathrm{f}})^d \;\middle|\; \varphi = 0 \text{ on } \Gamma_{\mathrm{f}} \setminus (\Gamma_{\mathrm{in}} \cup \Gamma_{\mathrm{out}}) \text{ and } \varphi = v^D \text{ on } \Gamma_{\mathrm{in}} \right\}$$
$$\text{and} \quad \widetilde{W_{\mathrm{f}}^{v}} := \left\{ \varphi \;\middle|\; \varphi \in L^2(I; \tilde{V}_{\mathrm{f}}) \text{ and } \partial_t\varphi \in L^2(I; \tilde{V}_{\mathrm{f}}^{*}) \right\}.$$

For brevity we are going to enforce zero Dirichlet conditions on Γ_{f} in the following Sections and Chapters. Everything can be extended to configurations with inflow and outflow condition.

4.1.5. Monolithic ALE-Formulation for Fluid-Structure Interaction

Monolithic formulations are well known for nonlinear FSI problems and they are used for example in [84, 132, 56, 23, 149] to obtain robust numerical algorithms keeping errors occurring from the coupling conditions small.

Usually the solid velocity has only limited regularity $v_\mathrm{s} \in W_\mathrm{s}^v$, however to enforce the kinematic coupling condition in a $H^{\frac{1}{2}}$ trace sense we will need $v_\mathrm{s}\big|_{\Gamma_\mathrm{i}} \in L^2(I; H^{\frac{1}{2}}(\Gamma_\mathrm{i}))$. We will not be able to close this gap, but assume that every solution of the solid problem will fulfill the additional smoothness $v_\mathrm{s} \in L^2(I; V_\mathrm{s})$. At least for the linear fluid-structure interaction problem, we were able to prove in Theorem 3.3 this additional regularity for smooth initial data and right-hand side. This enables to enforce the kinematic coupling condition by choosing a smooth trial space for the common velocity variable $v \in W^v$ defined on the whole domain. The same approach is used to enforce the geometric coupling condition. We demand for the displacement $u \in W^u$, whereby u on Ω_s is the physical solid displacement and on Ω_f the variable u is the artificial ALE mapping. Furthermore, due to a test function defined in the same velocity space V, the dynamic coupling condition is automatically fulfilled. As the variable u and v now live on the whole domain, we neglect the indices s and f as already in the linear case. In the case of the here considered nonlinear FSI problem this leads to

Problem 4.4 (Fluid-structure interaction problem)**.** *Find a velocity $v \in W^v$, a displacement $u \in W^u$ and a pressure $p \in L^2(I; L_\mathrm{f})$ fulfilling the weak formulation:*

$$\begin{aligned} a^F(v,u,p)(\varphi) + a^S(u,v)(\varphi) &= (\!(\rho_\mathrm{f} f_\mathrm{f}, \varphi)\!)_\mathrm{f} + (\!(\rho_\mathrm{s} f_\mathrm{s}, \varphi)\!)_\mathrm{s} && \forall \varphi \in L^2(I; V), \\ a^M(u)(\psi) + a^V(u,v)(\psi) &= 0 && \forall \psi \in L^2(I; V_{\mathrm{f},0} \oplus H_\mathrm{s}), \\ a^D(v,u)(\xi) &= 0 && \forall \xi \in L^2(I; L_\mathrm{f}) \end{aligned}$$

with initial conditions $v(0) = v_0$ and $u(0) = u_0$.

We denote by $V_{\mathrm{f},0} \oplus H_\mathrm{s}$ the sum of the space $V_{\mathrm{f},0}$ defined on the fluid domain Ω_f and the space H_s defined on the solid domain Ω_s. Thereby, we have

$$V_{\mathrm{f},0} \oplus H_\mathrm{s} := \left\{ \varphi \in H \;\middle|\; \varphi\big|_{\Omega_\mathrm{f}} \in V_{\mathrm{f},0} \text{ and } \varphi\big|_{\Omega_\mathrm{s}} \in H_\mathrm{s} \right\}.$$

Remark 4.3. For Problem 4.4 we cannot prove existence of a unique solution. The spaces are chosen in such a way that the semi-linear forms are well defined. The coupling conditions are embedded in the same way into the variational formulation as for linear fluid-structure interaction configuration in Problem 3.3. If the solution holds more regularity in time, we can adapt the proof of Theorem 3.4 and show that the coupling conditions are fulfilled in the trace sense.

To write the FSI problem more compact, we introduce for the volume force the linear operator

$$f(\varphi) := (\!(\rho_\mathrm{f} f_\mathrm{f}, \varphi)\!)_\mathrm{f} + (\!(\rho_\mathrm{s} f_\mathrm{s}, \varphi)\!)_\mathrm{s}.$$

Remark 4.4. Here the volume force f_f is defined on the reference domain Ω_f. If the volume force is given on the physical domain $\check{\Omega}_\mathrm{f}$, we have to transform the data on the reference domain Ω_f. Then the right-hand side will depend on the solution of the ALE transformation u.

If we define in addition the solution space

$$X := W^v \times W^u \times L^2(I; L_{\mathrm{f}}),$$

and the combined test space

$$Y := L^2(I; V) \times L^2(I; V_{\mathrm{f},0} \oplus H_{\mathrm{s}}) \times L^2(I; L_{\mathrm{f}}),$$

we can write the continuous fluid-structure interaction problem in a very compressed form as

Problem 4.5 (Fluid-structure interaction problem)**.** *Find* $\mathbf{u} = (v, u, p) \in X$ *with initial conditions* $v(0) = v_0$ *and* $u(0) = u_0$ *such that*

$$a(\mathbf{u})(\boldsymbol{\varphi}) = f(\varphi) \quad \forall \boldsymbol{\varphi} \in Y \tag{4.9}$$

with $\boldsymbol{\varphi} = (\varphi, \psi, \xi)$ *and*

$$a(\mathbf{u})(\boldsymbol{\varphi}) := a^F(\mathbf{u})(\varphi) + a^S(\mathbf{u})(\varphi) + a^V(\mathbf{u})(\varphi) + a^M(\mathbf{u})(\psi) + a^D(\mathbf{u})(\xi). \tag{4.10}$$

We apply here the same notation as already used for the linear fluid-structure interaction formulation in Section 3.1.4, not only for the semi-linear form, but in addition the label X for the solution space and Y for the test space. In the next chapters, we will only tackle the nonlinear fluid-structure interaction problem. Thereby, we will denote with X and Y always the here defined spaces. Nevertheless, the presented techniques can be applied to the linear model, too. The abuse of notations enables the reader to extend the results on his own.

4.2. Existence Theory for Nonlinear Fluid-Structure Interaction

We will assume in the following that there exists a unique smooth solution for Problem 4.5. Nevertheless the existence and regularity results in literature do not provide such a result. Nonlinear fluid-structure interaction problems have been analyzed in great detail in the last decade. The first results have been published by Coutand and Shkoller in [47, 48]. We state here their main theorem taken from [48], wherein the authors consider a similar nonlinear fluid-structure interaction configuration in ALE coordinates with Saint-Venant Kirchhof material law as in Problem 4.4.

Theorem 4.2 (Therorem 1 in [48])**.** *Let* $\Omega \subset \mathbb{R}^3$ *be a bounded domain of class* H^4*, and let* Ω_{s} *be an open set of class* H^4 *such that* $\overline{\Omega}_{\mathrm{s}} \subset \Omega$ *and* $\Omega = \overline{\Omega}_{\mathrm{s}} \cup \Omega_{\mathrm{f}}$*. Let* $f \in H^n(I; H^{3-n}(\Omega))$ *for* $n = 0, ..., 3$*,* $f(0) \in H^4(\Omega)$ *and* $\partial_t f(0) \in H^4(\Omega)$*. Assume the initial data satisfies*

$$u_0 = 0, v_0 \in H^1(\Omega), v_0\big|_{\Omega_{\mathrm{f}}} \in H^6(\Omega_{\mathrm{f}}), v_0\big|_{\Omega_{\mathrm{s}}} \in H^6(\Omega_{\mathrm{s}})$$

with $div(v_0) = 0$ *on* Ω_{f} *and a set of compatibility conditions for the initial data on the interface (see [48] for details).*

Then there exists $\bar{T} \in I$ *depending on* v_0, f, *and* Ω_f *such that there exists a unique velocity* $v \in L^2(\bar{I}; H_0^1(\Omega))$ *and pressure* $p \in H^n(\bar{I}; H^{3-n}(\Omega_\mathrm{f})$ *with* $n = 0, ..., 2$ *solving the nonlinear fluid structure interaction problem in ALE coordinates on* $\bar{I} = (0, \bar{T})$. *Furthermore*

$$\begin{aligned} v\big|_{\Omega_\mathrm{f}} &\in H^n(\bar{I}; H^{k-n}(\Omega_\mathrm{f})) \text{ with } n = 0, .., 4, \\ \int_0^t v\big|_{\Omega_\mathrm{s}}\, dt &\in H^n(\bar{I}; H^{k-n}(\Omega_\mathrm{s})) \text{ with } n = 0, .., 4, \\ u \in C^0(\bar{I}; H^1(\Omega)), u\big|_{\Omega_\mathrm{f}} &\in C^0(\bar{I}; H^4(\Omega_\mathrm{f})), u\big|_{\Omega_\mathrm{s}} \in C^0(\bar{I}; H^4(\Omega_\mathrm{s})). \end{aligned}$$

The existence result has been extended in [86, 93] by Kukavica, Tuffaha and coworkers for lesser regularity assumptions on the initial data. By introducing a damping parameter in solid and coupling condition, the authors in [87] where able to prove global existence of a solution for fluid-structure interaction. We would like to emphasize that existence results for more general domains, as considered in the numerical calculations later, only have been proven for the linear FSI problem as for example in [53, 54, 3, 5], yet. The authors in [48] need the outer boundary as well as the initial solid boundary to be extreme smooth to prove Theorem 4.2. Furthermore, the solid has to be fully immersed by the fluid. Hence, configurations as given in Figure 2.1 are not included in this result.

Further results can be found in [76, 77] by Grandmont and Hillairet, who use a plate model to describe the solid motion. A further class of nonlinear FSI models with fixed interface for fluid flow around a smooth elastic object is analyzed in [12, 13, 95], where in [13, 95] the authors even prove the existence of strong solutions for smooth initial conditions. A similar result can be found in [3, 94].

4.3. Optimal Control Problem

Several motivations for optimal control of fluid-structure interaction come from hemodynamics. The arterial system of every patient varies widely. The geometry of the aorta can be visualized using MRI images. However, not only the geometry but also the surrounding system and the material parameters of the vascular system are patient-specific. To use the simulation as a diagnostic tool, we have to take all these unknown variables into account. As in vivo measurements are usually not at hand, we have to solve a parameter estimation problem.

To simulate diseased arteries, it is important to reproduce the correct pressure values. This is only possible if we have appropriate patient specific boundary conditions at hand modeling the whole arterial network. For example in [63] and [65] Formaggia and coworkers couple 3D fluid-structure interaction models with 1D/0D models for the remaining network. As a reduced model, very often a windkessel model is used as presented in [148]. In [135] Spilker and Taylor first suggest an optimal control problem to calibrate the windkessel parameters. Therein, the whole arterial system is modeled by a reduced model. Gradient information is obtained by finite difference approximations. In [122], a sequential unscented Kalman filter is used to determine the windkessel parameters. Wall and coworkers suggest in [88] an adjoint based approach having only the systolic and diastolic pressure available for the calibration. However, the sensitivity information is only determined for the windkessel model. The windkessel model

is very simple to implement, but fails to capture wave traveling phenomena. To be able to resolve additional physical aspects in the models, it could be reasonable to allow time-dependent parameters. Alternatively, a 3D fluid-structure interaction model can be coupled with a reduced 1D fluid-structure interaction model. In [110] and [50], parameter estimation approaches to calibrate the parameters in such 1D models are presented. In both articles the authors use a quasi Newton algorithm and adjoint sensitivity information. An extension taking sensitivity information into account of the full 3D model coupled with a 1D/0D model has not been applied yet.

Furthermore, the effects of the surrounding tissue and organs has to be taken into account. In [118], Moireau, Bertoglio, and coworkers divide the vessel wall in regions with distinct tissue support and adjust the boundary condition in each region. In [117], they apply a Kalman filter to automatically adapt the boundary condition such that the simulation matches the time-resolved wall motion determined by a computed tomography (CT) angiographic scan. A natural extension of their approach would be to allow the parameters in the boundary condition to be distributed along the boundary.

The deformability of soft tissue can be an important index for the detection of anomalies or diseases. In [27], Bertoglio, Moirea, and Gerbeau use a sequential reduced Kalman filter to determine the Young's modulus in the region of an aneurism. They divide the aneurism in 5 to 10 regions and assume the Young's modulus to be constant in every region. In [25], they extend the model applying a more realistic material law and use an image registration algorithm to determine the vessel displacement at arbitrary time points as reference data. Peregio, Veneziani, and Vergara compute in [123] sensitivity information to estimate the wall stiffness. To reduce the computational time, they solve in every time-point an optimal control problem. As the mesh motion is discretized via an explicit time-stepping scheme, no sensitivity information of the mesh motion equation has to be computed. Similar to the articles [27, 25, 117], the estimated parameters are updated in every time-step and the forward simulation only runs once. To be able to incorporate better measurements, which are only available at specific time-points, it would be reasonable to compute the sensitivities on the space-time cylinder. In addition, as a homogenous distribution of the stiffness can only be assumed for healthy patients, it is necessary to extend these results to distributed stiffness parameters.

Naturally, the arterial system is never at rest. Hence, as initial condition we have to choose the fluid flow and the displacement at one time point in the cardiac cycle. However, measurements usually do not provide us with these distributed information. One possibility is to start the system at rest and run several cardiac cycles until a periodic solution appears. An alternative would be to control the initial conditions such that the simulation matches given measurements at certain time points.

However, not only in hemodynamics are a large variety of applications for optimization problems constrained by fluid-structure interaction models available. Shape optimization has a long tradition in aerodynamics. To take the elastic reaction of the controlled structure into account more and more aeroelasticity models are beeing used during the optimization process.

We are going to take the following chapters as general as possible such that the methods developed in the following can be applied to most of these applications. Nevertheless, to keep

everything readable, we will restrict us to the case with a control variable $q \in L^2(I; L^2(\Omega_\mathrm{f}))$ or $q \in L^2(I; L^2(\Omega_\mathrm{s}))$ acting as volume force in the right-hand side. We are going to comment, if we have to proceed differently in the case of a Neumann boundary in the following chapters. The extension to a control variable only distributed in space or to time-dependent parameters is straightforward throughout the next chapters. In Section 8.3.3, we will control the artificial material parameters in the mesh motion equation to obtain a smoother ALE transformation. Therein, we are going to highlight shortly the additional difficulties appearing if we want to control a distributed material parameter.

In fluid-structure interaction applications, we are often interested in the stresses in the solid domain or in the forces acting on the coupling interface. For example hemodynamic forces, including wall stresses, play a critical role in the development of stenosis according to [92]. In the case of an aneurism, high stresses will result in a rupture. Hence, the authors in [142] suggest to compute the drag value

$$\mathcal{J}(u) = \int_I \int_{\Gamma_\mathrm{i}} -\Sigma(u) n e_1 \, \mathrm{d}x \, \mathrm{d}t \tag{4.11}$$

for the benchmark examples along a moving flag. To reduce the boundary forces, the goal would be to minimize such a time averaged drag functional. Furthermore, over the last years great effort has been taken to use the simulation as a diagnostic tool. To obtain patient specific simulations, measurements have to correlate with the simulation. For example in [88], the systolic and diastolic pressure at certain points is used to determine patient specific boundary conditions. In the last decade computed tomography (CT) and magnetic resonance imaging (MRI) have advanced to the state, where detailed state and motion of the vessel wall can be provided as in [25, 2]. In [98], even a data assimilation approach using flow field data provided by a 4D phase contrast MRI has been presented. If we want the simulation to coincide with measurements, we have to minimize a tracking type functional as for example

$$\mathcal{J}(u,v) := \frac{\gamma_\mathrm{f}}{2} \int_{\tilde{I}} \|v - v_\mathrm{d}\|^2_{L^2(\tilde{\Omega}_\mathrm{f})} \, \mathrm{d}t + \frac{\gamma_\mathrm{s}}{2} \int_{\tilde{I}} \|u - u_\mathrm{d}\|^2_{L^2(\tilde{\Omega}_\mathrm{s})} \, \mathrm{d}t. \tag{4.12}$$

Therein, v_d is a given desired fluid velocity field and u_d a given desired displacement field of the elastic structure. These measurements can be given on certain time intervals $\tilde{I} \subset I$ and on parts of the domain $\tilde{\Omega}_\mathrm{s} \subset \Omega_\mathrm{s}$ or $\tilde{\Omega}_\mathrm{f} \subset \Omega_\mathrm{f}$. If measurements are only available at the end-timepoint, we have the tracking type functional

$$\mathcal{J}(u,v) := \frac{\gamma_\mathrm{f}}{2} \|v(T) - v_\mathrm{d}\|^2_{L^2(\tilde{\Omega}_\mathrm{f})} + \frac{\gamma_\mathrm{s}}{2} \|u(T) - u_\mathrm{d}\|^2_{L^2(\tilde{\Omega}_\mathrm{s})}. \tag{4.13}$$

Very often, only measurements at certain points $x_{obs} \in \Omega$ are available, then we have a point functional, as for example

$$\mathcal{J}(u,v) := \frac{\gamma_\mathrm{f}}{2} \int_{\tilde{I}} (v(x_{obs}) - v_\mathrm{d}(x_{obs}))^2 \, \mathrm{d}t + \frac{\gamma_\mathrm{s}}{2} \int_{\tilde{I}} (u(x_{obs}) - u_\mathrm{d}(x_{obs}))^2 \, \mathrm{d}t. \tag{4.14}$$

We will present further examples for the cost functional in Section 8.

All the presented examples have in common that they are solved via minimization of the cost functional constrained by a fluid-structure interaction problem. In a very abstract form, we obtain the optimal control problem:

Problem 4.6 (Optimal Control Problem)**.**

$$\min_{q\in Q} \mathcal{J}(q,\mathbf{u}) = \mathcal{J}(\mathbf{u}) + \frac{\alpha}{2}\|q\|_Q^2$$

subject to $\mathbf{u} \in X$ *with* $v(0) = v_0$, $u(0) = u_0$ *and*

$$\begin{aligned} a(\mathbf{u})(\boldsymbol{\varphi}) &= f(q)(\varphi) \quad \forall \boldsymbol{\varphi} \in Y, \\ q_a \leq & q \leq q_b. \end{aligned}$$

Thereby, $a(\mathbf{u})(\boldsymbol{\varphi})$ is the semi-linear form defined in (4.10) of the monolithic nonlinear fluid-structure interaction problem. The semi-linear form consist of the fluid problem $a^F(\cdot)(\cdot)$ and $a^D(\cdot)(\cdot)$ defined in (4.6), the solid problem $a^S(\cdot)(\cdot)$ and $a^V(\cdot)(\cdot)$ defined in (4.2) and the mesh motion equation $a^M(\cdot)(\cdot)$ defined in (4.8). The control $q \in Q$ enters the right-hand side as volume force. For a control $q \in L^2(I; L^2(\Omega_\mathrm{f}))$ on the fluid domain we have $f(q)(\varphi) := (\!(q, \varphi)\!)_\mathrm{f}$ and for a control $q \in L^2(I; L^2(\Omega_\mathrm{s}))$ on the solid domain we have $f(q)(\varphi) := (\!(q, \varphi)\!)_\mathrm{f}$. In addition the control variable is subject to control constraints with the bounds $q_a, q_b \in \mathbb{R} \cup \{\pm\infty\}$ and $q_a < q_b$. We will denote with $\mathcal{J}(\mathbf{u}, q)$ the regularized cost functional with Tikhonov regularization and regularization parameter $\alpha > 0$ in the following.

4.3.1. Discussion of Optimal Solutions

Prooving existence of an optimal solution of the nonlinear optimal control problem 4.6 is not as straightforward as in the linear case. Due to the nonlinearity of the control-to-state mapping, we cannot guarantee lower semicontinuity of the reduced functional $j(q) = \mathcal{J}(q, \mathbf{u}(q))$, even if the functional $\mathcal{J}(q, \mathbf{u})$ is convex and continuous. We want to refer to [141] for existence results for optimal control of semilinear parabolic problems and to [79] as well as [1] for existence results for optimal control of the Navier-Stokes equations. Therein, the authors need, in addition to the existence of the control-to-state mapping, its boundedness. Then, they can prove weak convergence of a mimimization sequence $(q_i, \mathbf{u}_i) \rightharpoonup (\bar{q}, \bar{\mathbf{u}})$ in appropriate spaces. To prove that the limit $\bar{\mathbf{u}}$ solves the state equation with control $\bar{q}$, the authors have to take the limit of the resulting sequence of weak formulations. Thereby, the limit in the nonlinear terms has to be analyzed.

Hence, even if existence and uniqueness results were available for the nonlinear time dependent fluid-structure interaction problem, existence results for the here considered optimal control problem are far from trivial.

We want to highlight at this point uniqueness of the optimal solution cannot be guaranteed as we have a nonlinear optimal control problem.

4.3.2. Adjoint Equations

For optimal control of a stationary fluid-structure interaction problem in [130], the necessary optimality system is derived from using the Lagrange formalism. We follow the same approach, as already in Section 3.3.2 for the linear optimal control problem, now for the nonlinear optimal control problem presented in Problem 4.6.

Formal Lagrange Formalism

For all the cited configurations at the beginning of this section, we can use the Lagrange formalism to compute formally the gradient of the reduced cost functional. As in Section 3.3.2, we define the Lagrangian $\mathcal{L} : Q \times X \times Y \to \mathbb{R}$ by

$$\begin{aligned}\mathcal{L}(q, \mathbf{u}, \mathbf{z}) :=& \mathcal{J}(q, \mathbf{u}) - a(\mathbf{u})(\mathbf{z}) + f(q)(\varphi) \\ &+ (u_0 - u(0), z^u(0)) + (v_0 - v(0), z^v(0)),\end{aligned}$$

with $\mathbf{u} = (v, u, p)$ the solution of the fluid-structure interaction problem and the Lagrange multipliers $\mathbf{z} = (z^v, z^u, z^p)$. Here, $a(\mathbf{u})(\mathbf{z})$ is the semi-linear form defined in (4.10) with the control q acting as volume force in the right-hand side $f(q)(\varphi)$. If the triple $\mathbf{u} = (v, u, p)$ is the solution of the nonlinear fluid-structure interaction configuration in Problem 4.5, we can express again the reduced functional $j(q) := \mathcal{J}(q, \mathbf{u}(q))$ via the Lagrangian $j(q) = \mathcal{L}(q, \mathbf{u}(q), \mathbf{z})$ for arbitrary values $\mathbf{z} = (z^v, z^u, z^p) \in Y$. Hence, we get, as for the linear FSI problem, the representation of the derivative of the reduced functional

$$j'(q)(\delta q) = \mathcal{L}'_q(q, \mathbf{u}(q), \mathbf{z})(\delta q) + \mathcal{L}'_{\mathbf{u}}(q, \mathbf{u}(q), \mathbf{z})(\delta \mathbf{u}),$$

whereby $\delta \mathbf{u} = \mathbf{u}'_q(q)(\delta q)$ is the derivative of the state solution with respect to the control variable. If we choose $\mathbf{z} \in Y$ such that $\mathbf{z}$ solves the adjoint equation

$$\mathcal{L}'_{\mathbf{u}}(q, \mathbf{u}, \mathbf{z})(\boldsymbol{\varphi}) = 0 \quad \forall \boldsymbol{\varphi} \in X, \tag{4.15}$$

the derivative of the Lagrange functional with respect to $\mathbf{u}$ is zero independent of $\delta \mathbf{u}$. Then, we can evaluate the derivative of the reduced functional in an arbitrary direction $\delta q \in Q$ by evaluating

$$j'(q)(\delta q) = \mathcal{L}'_q(q, \mathbf{u}, \mathbf{z})(\delta q).$$

In the case of the nonlinear fluid-structure interaction problem, we obtain the adjoint equation

Problem 4.7 (Adjoint FSI Problem)**.** *Find $\mathbf{z} \in Y$ solving the adjoint equation*

$$a'_{\mathbf{u}}(\mathbf{u})(\boldsymbol{\varphi}, \mathbf{z}) - (\psi(0), z^u(0)) - (\varphi(0), z^v(0)) = \mathcal{J}'_{\mathbf{u}}(\mathbf{u})(\boldsymbol{\varphi}) \quad \forall \boldsymbol{\varphi} \in X.$$

Hereby, we denote with $a'_{\mathbf{u}}(\mathbf{u})(\boldsymbol{\varphi}, \mathbf{z})$ the linearization of $a(\mathbf{u})(\mathbf{z})$ with respect to $\mathbf{u}$ in direction $\boldsymbol{\varphi}$. As we only consider the case of control acting as volume force in the right-hand side in this section, the adjoint equation does not depend on the control variable. If the control variable enters for example as parameter into the equations, then the adjoint equation will also depend on the control variable.

Remark 4.5. In comparison to Section 3.3.2, we are not able to prove strictly an optimality system. Hence, we cannot be sure if the optimal state and adjoint solution will be an element of the chosen spaces in the Lagrange functional. Nevertheless, for a large class of examples the formally derived adjoint equation is indeed the correct adjoint equation, as for example in Section 3.3 for optimal control of linear fluid-structure interaction.

Remark 4.6. If we have nonhomogenous Dirichlet conditions on parts of the outer boundary, we can embed the Dirichlet conditions into the space X. For example to enforce an inflow profile we can proceed as suggested in Section 4.1.4. However, the test functions in the weak state equation still have trace zero on the Dirichlet boundary. Hence the adjoint solution has zero Dirichlet conditions on the outer boundary.

Discussion of the Adjoint Equations

To derive the adjoint equation, we have to compute the derivative of the semi-linear form $a(\mathbf{u})(\mathbf{z})$ with respect to the velocity variable v, the displacement u, and the pressure p. In the monolithic formulation (4.10) the fluid equations are mapped on a fixed reference domain by the ALE transformation. All differentiations can now be carried out on the reference domain. If the fluid problem is formulated in Eulerian-coordinates, the moving fluid domain must be carefully included. Then the theory of shape calculus has to be used to compute sensitivity information as in [119, 31, 30]. In the following these shape derivatives are hidden in the derivatives with respect to the ALE transformation.

We have summarized in Appendix B how to compute the linearization of the individual terms, whereby we follow [126, 149]. Then the adjoint equation can be written in more detail as

Problem 4.8. *Find* $(z^v, z^u, z^p) \in L^2(I;V) \times L^2(I;V_{\mathrm{f},0} \oplus H_{\mathrm{s}}) \times L^2(I;L_{\mathrm{f}})$ *solving*

$$
\begin{aligned}
a_v'^F(\mathbf{u})(\varphi, z^v) + a_v'^D(\mathbf{u})(\varphi, z^p) - (\varphi(0), z^v(0)) & \\
+a_v'^S(\mathbf{u})(\varphi, z^v) + a_v'^V(\mathbf{u})(\varphi, z^u) = \mathcal{J}_v'(\mathbf{u})(\varphi) & \qquad \forall \varphi \in W^v, \\
a_u'^F(\mathbf{u})(\psi, z^v) + a_u'^D(\mathbf{u})(\psi, z^p) + a_u'^M(\mathbf{u})(\psi, z^u) & \\
+a_u'^S(\mathbf{u})(\psi, z^v) + a_u'^V(\mathbf{u})(\psi, z^u) - (\psi(0), z^u(0)) = \mathcal{J}_u'(\mathbf{u})(\psi) & \qquad \forall \psi \in W^u, \\
a_p'^F(\mathbf{u})(\xi, z^v) = \mathcal{J}_p'(\mathbf{u})(\xi) & \qquad \forall \xi \in L^2(I;L_{\mathrm{f}}).
\end{aligned}
$$

If we take a closer look, we see, that the derivative of the fluid semi-linear form $a^F(\cdot)(\cdot)$ and $a^D(\cdot)(\cdot)$, defined in (4.6), with respect to velocity and pressure results in a linearized adjoint Navier-Stokes problem on the moving domain given by the displacement u. The differentiation of the nonlinear elastodynamics equations, consisting of the semi-linear form $a^S(\cdot)(\cdot)$ and $a^V(\cdot)(\cdot)$, defined in (4.2), results in an adjoint wave equation. Together the derivative of the fluid equation $a^F(\cdot)(\cdot)$ and the derivative of the mesh motion equation $a^M(\cdot)(\cdot)$ with respect to the displacement u, describe the sensitivity information of the moving fluid domain.

In the state equation (4.10), the fluid and solid models are coupled due to the dynamic and kinematic coupling condition. The mesh motion equation is coupled by the geometric coupling condition. We enforce the coupling condition in a variational way in the monolithic formulation in Problem 4.5. Thereby, we choose the velocity $v \in W^v$ and displacement $u \in W^u$ in an adequate trial spaces to enforce the kinematic and geometric coupling conditions. In addition, the test variable $\varphi \in L^2(I;V)$ guarantees the dynamic coupling condition. Hence, for the state equation, we have two Dirichlet coupling conditions and one Neumann coupling condition. Now, we only have the variable $z^v \in L^2(I;V)$ defined on the whole domain in the adjoint equation and instead two test functions $\varphi \in W^v$ and $\psi \in W^u$ defined on the whole domain. Therefore, we now have enforced in a variational manner one Dirichlet and two Neumann coupling conditions in the adjoint equation.

For a tracking type functional (4.12), the sensitivity information of the functional enters the adjoint equation as volume force. In the case of a boundary functional as (4.11), we get a Neumann boundary condition. If we only have end observation, the functional (4.13) changes the terminal conditions of the adjoint equation. In the case of a point functional, we have a dirac function in the right-hand side in the adjoint equation.

Symmetric Formulation for the Nonlinear FSI Problem

In Section 3.1.4 and Section 3.3.2, we highlight that we introduce in the linear fluid-structure interaction model the velocity variable in the wave equation in an unconventional way. This enables us to use the same existence and uniqueness results for the linear state and linear adjoint equations.

In the last section, we derived the adjoint equation by calculating the linearization of the FSI Problem around the time dependent solution (u, v) and then taking the adjoint of the linearized operator. We could see that the resulting linear adjoint equation is not just a linearized fluid-structure interaction problem running back in time. If we would just adapt the approach in Section 3.1.4 by introducing the velocity variable in a new way this would not change. We still would have a change in the coupling conditions as the sensitivity of the mesh motion enters the adjoint structure equations via Neumann coupling conditions.

Furthermore, numerical experiments in Section 8.1 have not shown a faster or more stable convergence of the optimal solution, despite using the modified symmetric formulation. As we are neither going to prove existence nor regularity results for the adjoint equation and as we could not see any numerical benefit, we decided to use the standard formulation to get a first order system in the elastodynamics equation.

4.3.3. Necessary Optimality conditions

The optimal control $\bar{q}$ of the considered control problem has to be an element of the admissible set

$$Q_{ad} = \{\, q \in Q \mid q_a \leq q \leq q_b \,\}$$

and has to fulfill the necessary optimality condition

$$j'(\bar{q})(\delta q - \bar{q}) \geq 0 \quad \forall \delta q \in Q_{ad}.$$

The derivatives of the reduced functional $j(q) := J(q, \mathbf{u}(q))$ in direction of the admissible set have to vanish. Using the Lagrange formalism, we were able to derive a representation of the derivative of the reduced functional. Using this approach, we see that the optimal state, adjoint, and control variable have to constitute a stationary point of the Lagrangian $\mathcal{L}$.

1. The optimal state $\bar{\mathbf{u}} = (v, u, p) \in X$ fulfills the initial conditions $u(0) = u_0$ and $v(0) = v_0$ and solves the fluid structure interaction Problem 4.5. Therefore the state variable $\bar{\mathbf{u}} = \mathbf{u}(\bar{q}) \in X$ has to solve
$$\mathcal{L}'_{\mathbf{z}}(\bar{q}, \bar{\mathbf{u}}, \bar{\mathbf{z}})(\boldsymbol{\varphi}) = 0 \quad \forall \boldsymbol{\varphi} \in Y.$$

2. The optimal adjoint $\bar{\mathbf{z}} = \mathbf{z}(\bar{q}) \in Y$ has to solve
$$\mathcal{L}'_{\mathbf{u}}(\bar{q}, \bar{\mathbf{u}}, \bar{\mathbf{z}})(\boldsymbol{\varphi}) = 0 \quad \forall \boldsymbol{\varphi} \in X.$$
Thus $\mathbf{z} \in Y$ solves the solution of the adjoint fluid-structure interaction Problem 4.8.

3. It holds in addition the necessary optimality condition
$$j'(\bar{q})(\delta q - \bar{q}) = \mathcal{L}'_q(\bar{q}, \bar{\mathbf{u}}, \bar{\mathbf{z}})(\delta q - \bar{q}) \geq 0 \quad \forall \delta q \in Q_{ad}.$$

In the case of a control q in the control space $Q = L^2(I; L^2(\Omega))$ acting as volume force in the right-hand side this results in

$$\frac{\alpha}{2}(q, \delta q - \bar{q})_Q + (\!(\delta q - \bar{q}, \bar{\mathbf{z}})\!) \geq 0 \quad \forall \delta q \in Q_{ad}.$$

If we define the pointwise projection $\mathrm{P}_{\mathrm{Q_{ad}}}$ on the admissible set Q_{ad} by

$$\mathrm{P}_{\mathrm{Q_{ad}}} \colon L^2(I; L^2(\Omega)) \to L^2(I; L^2(\Omega)), \quad \mathrm{P}_{\mathrm{Q_{ad}}}(r)(t,x) := \max(q_a, \min(q_b, r(t,x))),$$

we can rewrite the necessary optimality condition to

$$\bar{q} = -\frac{1}{\alpha}\,\mathrm{P}_{\mathrm{Q_{ad}}}\left(\bar{z}^v\right).$$

5. Discretization

Usually, the finite element method is used to discretize a time-dependent PDE in space and a finite difference approach is applied to discretize in time. The mixture of variational discretization and finite differences has several drawbacks. For example "optimize" and "discretize" do not have to commute. If we first discretize the state equation and derive the optimality system, using the discrete Lagrangian, the system can differ to the system we get if we discretize the continuous optimality system. If we use a Petrov-Galerkin discretization on the space-time cylinder instead, we can guarantee that "optimize" and "discretize" commute up to quadrature errors. This is important, as we need the exact discrete gradient to get fast convergence of the optimization algorithm presented in Section 6.3. In addition, the variational discretization in time enables a posteriori error estimation, as we are going to see in Section 7.

We restrict ourselves in this chapter to a control q in the control space $Q = L^2(I; H)$ acting as volume force $f(q)(\varphi) := (\!(q, \varphi)\!)_{\mathrm{f}} + (\!(q, \varphi)\!)_{\mathrm{s}}$ on the solid and fluid domain. All the results derived in this chapter can be directly extended to the case with a control only time dependent or only acting on parts of the domain. For a Neumann boundary control we could proceed similarly.

We first introduce a temporal discretization of the nonlinear fluid-structure interaction problem in detail in Section 5.1. We apply the Petrov-Galerkin discretization suggested by Meidner and Richter in [111] and [112] to systematically derive a fractional-step theta time-stepping scheme for the fluid-structure interaction model in Problem 4.5. Then, we derive an adjoint time-stepping scheme in Section 5.2. In Section 5.3, we briefly comment on the difficulties in spatial discretization and state the finite element spaces used. Finally, we discretize the control variable and state the resulting fully discrete necessary optimality condition in Section 5.5. We could use the same methods to obtain a fully discrete optimality system for the linear fluid-structure interaction problem.

5.1. Time Discretization of the State Equation

The nonlinear elastodynamics equations are of hyperbolic type and energy-conserving over time. Therefore, either explicit time-stepping schemes with small time steps or implicit time-stepping schemes with small damping should be used. The fractional-step theta time-stepping scheme presented in [37] has shown to have excellent properties to discretize the Navier-Stokes equations. As the second order time-stepping scheme is A-stable and has only little numerical dissipation, numerical experiments in [131] have shown that it works excellent for monolithic fluid-structure interaction, too.

5.1.1. Theta Time-Stepping Scheme

The Galerkin discretization approach in [111, 112] enables to derive arbitrary theta time-stepping schemes. For $\theta = 0.5$ the method coincides with a Crank-Nicolson time-stepping scheme and for $\theta = 1$ with a backward Euler method. In addition, we can combine three steps to the fractional-step theta time-stepping scheme. In the following, we partition the time Interval $I = (0, T)$ in M subintervals $I_m = (t_m, t_{m-1})$ with $t_0 < t_1 < ... < t_M = T$. On every time interval I_m, we define a parameter $\theta_m \in [0, 1]$ and the interval length $k_m := t_m - t_{m-1}$.

A very detailed introduction and motivation for continuous Galerkin discretization in time can be found in the textbook [57] by Eriksson, Estep, Hansbo, and Johnson. Those authors show that the Crank-Nicolson time-stepping scheme is closely connected to piecewise linear trial and piecewise constant test functions. Motivated by these results, we are going to discretize the velocity and displacement field using a piecewise linear trial space $W_k^v \subset W^v$ and $W_k^u \subset W^u$ with

$$W_k^u = W_k^v := \left\{ v \in C(\bar{I}; V) \;\middle|\; v\big|_{I_m} \in P_1(I_m, V), m = 1, 2, ..., M \right\}.$$

To take the pressure fully implicit we use the piecewise constant and discontinuous test space $L_k^p \subset L^2(I; L_\mathrm{f})$ with

$$L_k^p = \left\{ p \in L^2(I; L_\mathrm{f}) \;\middle|\; p\big|_{I_m} \in P_0(I_m, L_\mathrm{f}), m = 1, 2, ..., M \right\}.$$

For the incompressibility constraint the test space $L^2(I; L_\mathrm{f})$ is replaced by the piecewise constant space L_k^p, too. Thereby the semidiscrete pressure can later be interpreted as Lagrange multiplier of the divergence condition in each time interval I_m. To obtain an arbitrary theta time-stepping scheme, we choose for the momentum equation and the extension a discontinuous test space $L_{k,\theta}^v \subset L^2(I; V)$ and $L_{k,\theta}^u \subset L^2(I; V_{\mathrm{f},0} \oplus H_\mathrm{s})$ with

$$L_{k,\theta}^v = \left\{ \varphi \in L^2(I; V) \;\middle|\; \varphi\big|_{I_m} \in P_0^\theta(I_m, V), m = 1, 2, ..., M, \varphi(0) \in V \right\},$$
$$L_{k,\theta}^u = \left\{ \varphi \in L^2(I; V_{\mathrm{f},0} \oplus H_\mathrm{s}) \;\middle|\; \varphi\big|_{I_m} \in P_0^\theta(I_m, V_{\mathrm{f},0} \oplus H_\mathrm{s}), m = 1, 2, ..., M, \varphi(0) \in V_{\mathrm{f},0} \oplus H_\mathrm{s} \right\}.$$

Thereby the slope, as suggested in [111], depends on the parameter θ_m on each time interval I_m, as we choose

$$P_0^\theta(I_m, V) := \{ \omega_{\theta,m}\varphi \mid \varphi \in V \} \text{ and } P_0^\theta(I_m, V_{\mathrm{f},0} \oplus H_\mathrm{s}) := \{ \omega_{\theta,m}\psi \mid \psi \in V_{\mathrm{f},0} \oplus H_\mathrm{s} \}$$

with

$$\omega_{\theta,m}(t) = 1 + 6\Big(\theta_m - \frac{1}{2}\Big)\frac{2t - t_{m-1} - t_m}{\Delta t_m}. \tag{5.1}$$

Therefore, if we look for the semidiscrete solution $\mathbf{u}_k := (v_k, u_k, p_k) \in X_k := W_k^v \times W_k^u \times L_k^p$ and if we replace the test space Y in Problem 4.5 with the just defined semidiscrete space $Y_{k,\theta} := L_{k,\theta}^v \times L_{k,\theta}^u \times L_k^p$, we automatically obtain the Petrov Galerkin semidiscretization of the state equation:

Problem 5.1. *Find for $q \in Q$ a solution $\mathbf{u}_k \in X_k$ with initial conditions $v_k(0) = v_0$ and $u_k(0) = u_0$ such that*

$$\begin{aligned} a^F(\mathbf{u}_k)(\varphi) + a^S(\mathbf{u}_k)(\varphi) &= f(q)(\varphi) && \forall \varphi \in L^v_{k,\theta}, \\ a^M(\mathbf{u}_k)(\psi) + a^V(\mathbf{u}_k)(\psi) &= 0 && \forall \psi \in L^u_{k,\theta}, \\ a^D(\mathbf{u}_k)(\xi) &= 0 && \forall \xi \in L^p_k. \end{aligned}$$

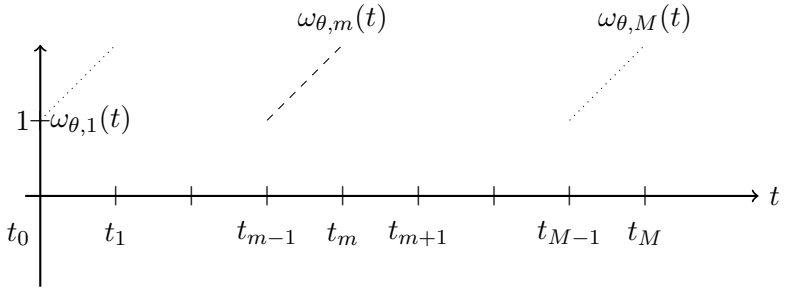

Figure 5.1.: Plot of the basis functions $\omega_{\theta,m}$

Now we exploit the fact, that every function $(\varphi, \psi, \xi) \in L^v_{k,\theta} \times L^u_{k,\theta} \times L^p_k$ can be represented as linear combination of the basis functions $\omega_{\theta,m}$ given in Figure 5.1 and characteristic function $\mathcal{X}_{I_m}$ such that

$$\varphi = \sum_{m=0}^{M} \varphi_m \omega_{\theta,m}(t), \quad \psi = \sum_{m=0}^{M} \psi_m \omega_{\theta,m}(t) \quad \text{and} \quad \xi = \sum_{m=1}^{M} \xi_m \mathcal{X}_{I_m}(t)$$

with $(\varphi_m, \psi_m, \xi_m) \in V \times V_{\mathrm{f},0} \oplus H_{\mathrm{s}} \times L_{\mathrm{f}}$.

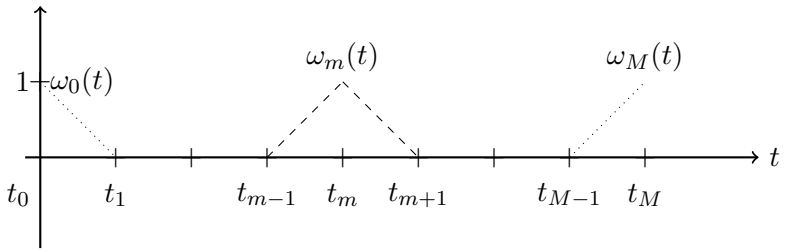

Figure 5.2.: Plot of the basisfunctions ω_m

As $(v_k, u_k) \in W^v_k \times W^u_k$ are piecewise linear functions in time, we can use the representation

$$v_k = \sum_{m=0}^{M} v_{k,m} \omega_m(t) \quad \text{and} \quad u_k = \sum_{m=0}^{M} u_{k,m} \omega_m(t)$$

with $v_{k,m} := v_k(t_m)$ and $u_{k,m} := u_k(t_m)$, whereby $\omega_m(t)$ is a standard hat function as in Figure 5.2 on the time grid. In addition, we can use for the semidiscretized pressure the representation

$$p_k = \sum_{m=1}^{M} p_{k,m} \mathcal{X}_{I_m}(t) \quad \text{with} \quad p_{k,m} := \frac{1}{k_m} \int_{I_m} p_k \mathrm{d}t.$$

For nonlinear problems, as the considered fluid-structure interaction problem at hand, we have to evaluate the time-integrals with an appropriate quadrature rule. Instead of using a quadrature rule of high order, we are going to evaluate the inner product in time using the following theta-dependent trapezoidal rule

$$\int_{I_m} h(t)\omega_{\theta,m}(t)\ \mathrm{d}x \approx k_m\{\theta_m h(t_m^-) + (1-\theta_m)h(t_{m-1}^+)\} \tag{5.2}$$

suggested in [111] with

$$h(t_m^+) := \lim_{\delta t \to 0^+} h(t_m + \delta t) \text{ and } \quad h(t_m^-) := \lim_{\delta t \to 0^-} h(t_m + \delta t).$$

Thereby, the resulting time-stepping scheme coincides with a standard theta time-stepping scheme and inherits its well known properties. For smooth functions multiplied with the theta-dependent function $\omega_{\theta,m}$, the quadrature rule is of second order convergence.

Lemma 5.1. *Let $h \in W^{2,1}(I_m, \mathbb{R})$ and $\omega_{\theta,m}$ defined by (5.1) for $m = 1, ..., M$. Then it holds*

$$\int_{I_m} h(t)\omega_{\theta,m}\ dt = k_m\{\theta_m h(t_m^-) + (1-\theta_m)h(t_{m-1}^+)\} + R_m(h).$$

The remainder $R_m(h)$ satisfies

$$|R_m(h)| \leq k_m^2 \|\partial_t^2 v\|_{L^1(I_m,\mathbb{R})}.$$

The proof of Lemma 5.1 can be found in [111]. As we have chosen for the incompressibility condition the test function $\xi \in L_k^p$ the θ-quadrature rule would only be of first order here. So we use a standard trapezoidal rule instead to evaluate $a^D(\cdot)(\cdot)$.

Remark 5.1. For the linear fluid-structure interaction model the theta-dependent quadrature rule evaluates exactly the time integrals in the bi-linear $a(\cdot)(\cdot)$, with exception of the divergence condition.

If we apply the stated quadrature rule and take into account that we have for the derivative in time of $(v_k, u_k) \in W_k^v \times W_k^u$ on every time interval

$$\partial_t v_k\big|_{I_m} = \frac{1}{k_m}(v_{k,m} - v_{k,m-1}) \quad \text{and} \quad \partial_t u_k\big|_{I_m} = \frac{1}{k_m}(u_{k,m} - u_{k,m-1}),$$

then we obtain

Problem 5.2. *Find for $q \in Q$ a solution $\mathbf{u}_k \in X_k$ with initial conditions $v_k(0) = v_0$ and $u_k(0) = u_0$ such that*

$$a_k(\mathbf{u}_k)(\boldsymbol{\varphi}) = f_k(q)(\varphi) \quad \forall \boldsymbol{\varphi} \in Y_{k,\theta} \tag{5.3}$$

with

$$\begin{aligned} a_k(\mathbf{u}_k)(\boldsymbol{\varphi}) = \sum_{m=1}^{M} \big[& a_{k,m}^F(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\varphi_m) + a_{k,m}^S(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\varphi_m) + a_{k,m}^M(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\psi_m) \\ & + a_{k,m}^V(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\psi_m) + a_{k,m}^D(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\xi_m) \big], \\ f_k(q)(\varphi) = \sum_{m=1}^{M} & f_{k,m}(q)(\varphi_m). \end{aligned}$$

The particular semi-linear forms $a^F_{k,m}(\cdot,\cdot)(\cdot)$, $a^S_{k,m}(\cdot,\cdot)(\cdot)$, $a^M_{k,m}(\cdot,\cdot)(\cdot)$, $a^V_{k,m}(\cdot,\cdot)(\cdot)$, as well as $a^D_{k,m}(\cdot,\cdot)(\cdot)$ and the expression $f_{k,m}(\cdot)(\cdot)$ are explained in detail in the following.

In detail, we get for $m = 1, ..., M$ with the test functions $(\varphi_m, \psi_m, \xi_m) \in V \times V_{\mathrm{f},0} \oplus H_\mathrm{s} \times L_\mathrm{f}$ on every time-interval I_m the following time discretized: semi-linear forms.

$$
\begin{aligned}
a^F_{k,m}(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\varphi_m) :=& \left(\rho_\mathrm{f}\left[\theta_m J_{k,m} + (1-\theta_m) J_{k,m-1}\right](v_{k,m} - v_{k,m-1}), \varphi_m\right)_\mathrm{f} \\
&- \Big[\theta_m \left(\rho_\mathrm{f} J_{k,m} F^{-1}_{k,m}(u_{k,m} - u_{k,m-1}) \cdot \nabla v_{k,m}, \varphi_m\right)_\mathrm{f} \\
&\qquad + (1-\theta_m)\left(\rho_\mathrm{f} J_{k,m-1} F^{-1}_{k,m-1}(u_{k,m} - u_{k,m-1}) \cdot \nabla v_{k,m-1}, \varphi_m\right)_\mathrm{f}\Big] \\
&+ k_m\Big[\theta_m \left(\rho_\mathrm{f} J_{k,m}(F^{-1}_{k,m} v_{k,m} \cdot \nabla) v_{k,m}), \varphi_m\right)_\mathrm{f} \\
&\qquad + (1-\theta_m)\left(\rho_\mathrm{f} J_{k,m-1}(F^{-1}_{k,m-1} v_{k,m-1} \cdot \nabla) v_{k,m-1}), \varphi_m\right)_\mathrm{f}\Big] \\
&+ k_m\Big[\theta_m \left(J_{k,m}(\sigma_\mathrm{f})_{k,m} F^{-T}_{k,m}, \nabla\varphi_m\right)_\mathrm{f} \\
&\qquad + (1-\theta_m)\left(J_{k,m-1}(\sigma_\mathrm{f})_{k,m} F^{-T}_{k,m-1}, \nabla\varphi_m\right)_\mathrm{f}\Big], \\
a^S_{k,m}(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\varphi_m) :=& \left(\rho_\mathrm{s}(v_{k,m} - v_{k,m-1}), \varphi_m\right)_\mathrm{s} \\
&+ k_m\Big[\theta_m \left(F_m(\Sigma_\mathrm{s})_{k,m}, \nabla\varphi_m\right)_\mathrm{s} + (1-\theta_m)\left(F_{k,m-1}(\Sigma_\mathrm{s})_{k,m}, \nabla\varphi_m\right)_\mathrm{s}\Big], \\
a^M_{k,m}(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\psi) :=& k_m\left[\theta_m \left((\sigma_\mathrm{m})_{k,m}, \nabla\psi_m\right)_\mathrm{f} + (1-\theta_m)\left((\sigma_\mathrm{m})_{k,m}, \nabla\psi_m\right)_\mathrm{f}\right], \\
a^V_{k,m}(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\psi) :=& \left(u_{k,m} - u_{k,m-1}, \psi_m\right)_\mathrm{s} - k_m\left[\left(\theta_m v_{k,m} + (1-\theta_m) v_{k,m-1}, \psi_m\right)_\mathrm{s}\right], \\
a^D_{k,m}(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\xi) :=& \frac{k_m}{2}\left(\mathrm{div}(J_{k,m} F^{-1}_{k,m} v_{k,m}) + \mathrm{div}(J_{k,m-1} F^{-1}_{k,m-1} v_{k,m-1}), \xi_m\right)_\mathrm{f},
\end{aligned}
$$

and after applying the theta quadrature rule (5.2) on the right-hand side we obtain on the time interval I_m

$$
f_{k,m}(q)(\varphi_m) := k_m\Big[\theta_m \left(q(t^-_m), \varphi_m\right)_\mathrm{s} + (1-\theta_m)\left(q(t^+_{m-1}), \varphi_m\right)_\mathrm{s}\Big].
$$

Thereby, we denote with $J_{k,m}$, $F_{k,m}$, $(\sigma_\mathrm{f})_{k,m}$, $(\sigma_\mathrm{m})_{k,m}$ and $(\Sigma_\mathrm{s})_{k,m}$ the evaluation of the respective terms at the time point t^-_m and for the values with index $m-1$ the evaluation at t^+_{m-1}.

Due to the linearity of the semilinear form $a_k(\mathbf{u}_k)(\boldsymbol{\varphi})$ with respect to the test function $\boldsymbol{\varphi}$, we can rewrite the semidiscretized nonlinear fluid-structure interaction as time-stepping scheme.

Problem 5.3. *Let $(\theta_m)_{m=1}^M \in [0,1]^M$ and $(v_{k,0}, u_{k,0}) = (v_0, u_0)$. Find for $q \in Q$ a solution $(\mathbf{u}_{k,m})_{m=1}^M \in V^M \times V^M \times L_\mathrm{f}^M$ by iterating for $m = 1, ..., M$:*

$$
\begin{aligned}
a^F_{k,m}(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\varphi) + a^S_{k,m}(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\varphi) &= f_{k,m}(q)(\varphi) && \forall \varphi \in V, \\
a^M_{k,m}(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\psi) + a^V_{k,m}(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\psi) &= 0 && \forall \psi \in V_{\mathrm{f},0} \oplus H_\mathrm{s}, \\
a^D_{k,m}(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\xi) &= 0 && \forall \xi \in L_\mathrm{f}.
\end{aligned}
$$

5.1.2. Divergence Condition

The divergence condition in Problem 5.3 is only fulfilled in an average sense in every time interval I_m as we demand in the time-stepping scheme in every step $m = 1, ..., M$

$$\frac{k_m}{2}\left(\operatorname{div}(J_{k,m}F_{k,m}^{-1}v_{k,m}) + \operatorname{div}(J_{k,m-1}F_{k,m-1}^{-1}v_{k,m-1}), \xi\right)_{\mathrm{f}} = 0 \quad \forall \xi \in L_{\mathrm{f}}.$$

Thereby, errors in the divergence condition are conserved and can accumulate over time. Thus, we are going to enforce the divergence condition to be fulfilled in every time point, by using instead

$$a_{k,m}^{D}(\mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\xi) = k_m \left(\operatorname{div}(J_{k,m}F_{k,m}^{-1}v_{k,m}), \xi\right)_{\mathrm{f}}.$$

If the initial condition is divergence-free, both formulations are equivalent. A detailed discussion of the divergence condition for the Navier-Stokes equations can be found in [112].

5.1.3. Fractional-Step Theta Time-Stepping Scheme

For $\theta_m = 1$ the resulting time-stepping scheme corresponds to an implicit backward Euler time-stepping scheme. This is known to be A-stable and of first order. The severe damping prevents oscillations in the elastic structure. Hence, the Euler scheme is inapplicable for fluid-structure interaction. For $\theta = 0.5$ we obtain the second order Crank-Nicolson time-stepping scheme. The Crank-Nicolson method conserves energy and thereby errors can accumulate during the simulation of diffusion equations. Thus, very small time steps are necessary to have a stable algorithm for the fluid-structure interaction model. To derive a fractional-step theta scheme, we will combine three steps to one macro time step with step size k_n and $n = 1, \ldots, M/3$ and for $m = 3 \cdot n$

$$\begin{aligned}(k_m, k_{m+1}, k_{m+2}) &:= (\alpha k_n, (1-2\alpha)k_n, \alpha k_n),\\ (\theta_m, \theta_{m+1}, \theta_{m+2}) &:= (\theta, (1-\theta), \theta).\end{aligned}$$

If we choose

$$\theta = \frac{1-2\alpha}{1-\alpha} \text{ and } \alpha := 1 - \sqrt{\frac{1}{2}}$$

the resulting fractional-step theta time-stepping scheme is known to be A-stable and of second order accuracy [37]. Due to its little numerical dissipation it is especially well suited for the coupled problem of wave equation and Navier-Stokes equations considered in this thesis. In addition, the resulting fractional-step theta time-stepping scheme in Problem 5.3 coincides with the time discrete Petrov-Galerkin formulation of Problem 5.1 up to numerical quadrature for the given choice of parameters.

5.2. Time Discretization of the Adjoint Equation

5.2.1. Theta Time-Stepping Scheme

As already in the previous Section 5.1, we replace the continuous spaces by the suggested semi-discrete spaces X_k and $Y_{k,\theta}$. Therefore, the Petrov-Galerkin semidiscretized adjoint problem writes as

Problem 5.4. *Find* $(z_k^v, z_k^u, z_k^p) \in Y_{k,\theta} = L_{k,\theta}^v \times L_{k,\theta}^u \times L_k^p$ *such that*

$$a'_{\mathbf{u}}(\mathbf{u}_k)(\boldsymbol{\varphi}, \mathbf{z}_k) - (\psi(0), z_k^u(0)) - (\varphi(0), z_k^v(0)) = \mathcal{J}'_{\mathbf{u}}(\mathbf{u}_k)(\boldsymbol{\varphi}) \quad \forall \boldsymbol{\varphi} \in X_k. \tag{5.4}$$

As test and trial spaces have been switched in the adjoint equation, the adjoint solutions z_k^v and z_k^u are now elements of the θ dependent space $L_{k,\theta}^v$ and $L_{k,\theta}^u$. Every test function $(\varphi, \psi, \xi) \in L_k^v \times L_k^u \times L_k^p$ can be presented as

$$\varphi = \sum_{m=0}^{M} \varphi_m \omega_m(t), \quad \psi = \sum_{m=0}^{M} \psi_m \omega_m(t), \quad \xi = \sum_{m=1}^{M} \xi_m \mathcal{X}_{I_m}(t)$$

with $(\varphi_m, \psi_m, \xi_m) \in V \times V \times L_{\mathrm{f}}$, too. Thereby, ω_m is again a hat function on the given time grid as in Figure 5.2. To evaluate the time integrals in the adjoint equation, we use the possibility to write functions in $X_{k,\theta}^v \times X_{k,\theta}^u \times X_k^p$ as linear combination of $\omega_{\theta,m}$ and $\mathcal{X}_{I_m}$. We get

$$z_k^v = \sum_{m=0}^{M} z_{k,m}^v \omega_{\theta,m}(t), \quad z_k^u = \sum_{m=0}^{M} z_{k,m}^u \omega_{\theta,m}(t), \quad \text{and} \quad z_k^p = \sum_{m=1}^{M} z_{k,m}^p \mathcal{X}_{I_m}(t)$$

$$\text{with} \quad z_{k,m}^v := \frac{1}{k_m} \int_{I_m} z_k^v \mathrm{dt}, \quad z_{k,m}^u := \frac{1}{k_m} \int_{I_m} z_k^u \mathrm{dt}, \quad \text{and} \quad z_{k,m}^p := \frac{1}{k_m} \int_{I_m} z_k^p \mathrm{dt}.$$

Applying again the suggested quadrature rule (5.2) and exploiting the linearity of the equation (5.4) with respect to the test functions results again in a time-stepping scheme. Thereby, the derivative of the cost functional is either evaluated exactly or with a quadrature rule of higher order such that the discretization error can be neglected. It is only important to use the same quadrature rule as used for the evaluation of the functional to compute the exact discrete gradient.

To illustrate how to derive the time-stepping scheme, we evaluate exemplary the integral terms occurring if we test the adjoint transport term with $\varphi = \varphi_m \omega_m$. We apply the quadrature rule (5.2) on the time interval I_m and get

$$\begin{aligned}
&\int_{I_m} (\rho_{\mathrm{f}} J_k (F_k^{-1}(v_k - \partial_t u_k) \cdot \nabla) \omega_m(t) \varphi_m), z_{k,m}^v \omega_{\theta,m})_{\mathrm{f}} \ \mathrm{d}t = \\
&k_m \theta_m ((\rho_{\mathrm{f}} J_k(t_m^-)(F_k^{-1}(t_m^-)(v_k(t_m^-) - \partial_t u_k(t_m^-)) \cdot \nabla) \omega_m(t_m^-) \varphi_m, z_{k,m}^v)_{\mathrm{f}} \\
+k_m (1-\theta_m) ((&\rho_{\mathrm{f}} J_k(t_{m-1}^+)(F_k^{-1}(t_{m-1}^+)(v_k(t_{m-1}^+) - \partial_t u_k(t_{m-1}^+)) \cdot \nabla) \omega_m(t_{m-1}^+) \varphi_m, z_{k,m}^v)_{\mathrm{f}} \\
&= k_m \theta_m ((\rho_{\mathrm{f}} J_{k,m}(F_{k,m}^{-1}(v_{k,m} - \frac{u_{k,m} - u_{k,m-1}}{k_m}) \cdot \nabla) \varphi_m, z_{k,m}^v)_{\mathrm{f}}.
\end{aligned}$$

Similarly we obtain on the time interval I_{m+1}

$$\begin{aligned}
&\int_{I_{m+1}} (\rho_{\mathrm{f}} J_k (F_k^{-1}(v_k - \partial_t u_k) \cdot \nabla) \omega_m(t) \varphi_m), z_{k,m+1}^v \omega_{\theta,m+1})_{\mathrm{f}} \ \mathrm{d}t = \\
k_{m+1} \theta_{m+1} ((&\rho_{\mathrm{f}} J_k(t_{m+1}^-)(F_k^{-1}(t_{m+1}^-)(v_k(t_{m+1}^-) - \partial_t u_k(t_{m+1}^-)) \cdot \nabla) \omega_{m+1}(t_{m+1}^-) \varphi_m, z_{k,m+1}^v)_{\mathrm{f}} \\
+k_{m+1} (1-\theta_{m+1}) ((&\rho_{\mathrm{f}} J_k(t_m^+)(F_k^{-1}(t_m^+)(v_k(t_m^+) - \partial_t u_k(t_m^+)) \cdot \nabla) \omega_m(t_m^+) \varphi_m, z_{k,m+1}^v)_{\mathrm{f}} \\
&= k_{m+1} (1-\theta_{m+1}) ((\rho_{\mathrm{f}} J_{k,m}(F_{k,m}^{-1}(v_{k,m} - \frac{u_{k,m+1} - u_{k,m}}{k_{m+1}}) \cdot \nabla) \varphi_m, z_{k,m+1}^v)_{\mathrm{f}}.
\end{aligned}$$

As $\varphi = \varphi_m \omega_m$ has compact support only on the time intervals I_m and I_{m+1} we get

$$\int_I (\rho_{\mathrm{f}} J_k (F_k^{-1}(v_k - \partial_t u_k) \cdot \nabla)\omega_m(t)\varphi_m), z_k^v)_{\Omega_{\mathrm{f}}} \,\mathrm{d}t =$$
$$k_m \theta_m ((\rho_{\mathrm{f}} J_{k,m} (F_{k,m}^{-1}(v_{k,m} - \frac{u_{k,m} - u_{k,m-1}}{k_m}) \cdot \nabla)\varphi_m, z_{k,m}^v)_{\mathrm{f}}$$
$$+ k_{m+1}(1-\theta_{m+1})((\rho_{\mathrm{f}} J_{k,m} (F_{k,m}^{-1}(v_{k,m} - \frac{u_{k,m+1} - u_{k,m}}{k_{m+1}}) \cdot \nabla)\varphi_m, z_{k,m+1}^v)_{\mathrm{f}}.$$

Due to the implicit structure of the adjoint equation, we do not integrate the adjoint equation first in time before replacing the continuous spaces with its discrete equivalent. Thus, we have to evaluate terms with the time derivative of the test function $\partial_t(\varphi_m \omega_m)$. We just exploit the fact that $\partial_t \omega_m\big|_{I_m} = \frac{1}{k_m}$, $\partial_t \omega_m\big|_{I_{m+1}} = \frac{1}{k_{m+1}}$ and zero elsewhere. Using the same techniques as before, we get for example

$$\int_I -(\rho_{\mathrm{f}} J_k (F_k^{-1} \partial_t \omega_m(t)\psi_m \cdot \nabla) v_k), z_k^v)_{\Omega_{\mathrm{f}}} \,\mathrm{d}x =$$
$$-((1-\theta_m)\rho_{\mathrm{f}} J_{k,m-1}(F_{k,m-1}^{-1}\psi_m \cdot \nabla)v_{k,m-1} + \theta_m \rho_{\mathrm{f}} J_{k,m}(F_{k,m}^{-1}\psi_m \cdot \nabla)v_{k,m}, z_{k,m}^v)_{\mathrm{f}}$$
$$+((1-\theta_{m+1})\rho_{\mathrm{f}} J_{k,m}(F_{k,m}^{-1}\psi_m \cdot \nabla)v_{k,m} + \theta_{m+1}\rho_{\mathrm{f}} J_{k,m+1}(F_{k,m+1}^{-1}\psi_m \cdot \nabla)v_{k,m+1}, z_{k,m+1}^v)_{\mathrm{f}}.$$

Remark 5.2. Note that the first and last step of the adjoint time-stepping scheme will look slightly differently. The basis function $\omega_M(t)$ and $\omega_0(t)$ only have support on the time interval I_M and I_1. To make clear the structure of the resulting time-stepping scheme we sketch the structure in the following. The semi-linear forms in the adjoint equation can be defined in the same spirit as the state equation after systematically applying the theta quadrature rule (5.2).

1. For $m = M$, find $\mathbf{z}_{k,M} \in V \times V_{\mathrm{f},0} \oplus H_{\mathrm{s}} \times L_{\mathrm{f}}$ such that $\forall \boldsymbol{\varphi} \in V \times V \times L_{\mathrm{f}}$:

$$(a'_{\mathbf{u}})_{k,M}(\mathbf{u}_{k,M}, \mathbf{u}_{k,M-1})(\boldsymbol{\varphi}, \mathbf{z}_{k,M}) = \mathcal{J}'_{\mathbf{u},M}(\mathbf{u}_{k,M}, \mathbf{u}_{k,M-1})(\boldsymbol{\varphi}).$$

2. For $m = M-1, ..., 1$ find $\mathbf{z}_{k,m} \in V \times V_{\mathrm{f},0} \oplus H_{\mathrm{s}} \times L_{\mathrm{f}}$ such that $\forall \boldsymbol{\varphi} \in V \times V \times L_{\mathrm{f}}$:

$$(a'_{\mathbf{u}})_{k,m}(\mathbf{u}_{k,m+1}, \mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\boldsymbol{\varphi}, \mathbf{z}_{k,m}, \mathbf{z}_{k,m+1}) = \mathcal{J}'_{\mathbf{u},m}(\mathbf{u}_{k,m+1}, \mathbf{u}_{k,m}, \mathbf{u}_{k,m-1})(\boldsymbol{\varphi}).$$

3. For $m = 0$ find $\mathbf{z}_{k,0} \in V \times V_{\mathrm{f},0} \oplus H_{\mathrm{s}} \times L_{\mathrm{f}}$ such that $\forall \boldsymbol{\varphi} \in V \times V \times L_{\mathrm{f}}$:

$$(a'_{\mathbf{u}})_{k,0}(\mathbf{u}_{k,1}, \mathbf{u}_{k,0})(\boldsymbol{\varphi}, \mathbf{z}_{k,1}) - (\psi, z_{k,0}^u) - (\varphi, z_{k,0}^v) = \mathcal{J}'_{\mathbf{u},0}(\mathbf{u}_{k,0}, \mathbf{u}_{k,1})(\boldsymbol{\varphi}).$$

As we can see, to compute the adjoint solution $\mathbf{z}_m$, we need the adjoint solution $\mathbf{z}_{m+1}$ from the previous time step. Hence, the adjoint equation runs backward in time. In contrast to the forward problem the adjoint time-stepping scheme depends on θ_m and θ_{m+1} and on the solution $\mathbf{u}_{m-1}$, $\mathbf{u}_m$, and $\mathbf{u}_{m+1}$. Therefore, the adjoint time-stepping scheme has not the same structure as a fractional-step theta scheme. The adjoint value $\mathbf{z}_{k,0}$ gets important if we control the initial condition.

5.2.2. Adjoint Divergence Condition

The averaged divergence condition causes an "average adjoint pressure term". As we changed the divergence condition in Section 5.1.2, we have to adapt the adjoint time-stepping scheme to get a consistent formulation. Therefore, we replace in every time step $m = M, .., 1$ the adjoint pressure dependent terms

$$\begin{aligned}
(a_v'^D)_{k,m}(\mathbf{u}_{k,m})(\varphi, z^p_{k,m}) =& k_m(\mathrm{div}(J_{k,m}F^{-1}_{k,m}\varphi), z^p_{k,m}))_{\Omega_\mathrm{f}} \\
&+ k_{m+1}(\mathrm{div}(J_{k,m}F^{-1}_{k,m}\varphi), z^p_{k,m+1}))_{\Omega_\mathrm{f}}, \\
(a_u'^D)_{k,m}(\mathbf{u}_{k,m})(\psi, z^p_{k,m}) =& k_m(\mathrm{div}(J_{k,m}\,\mathrm{tr}(F^{-1}_{k,m}\nabla\psi)F^{-1}_{k,m}v_{k,m}), z^p_{k,m}))_{\Omega_\mathrm{f}} \\
&+ k_{m+1}(\mathrm{div}(J_{k,m}\,\mathrm{tr}(F^{-1}_{k,m}\nabla\psi)F^{-1}_{k,m}v_{k,m}), z^p_{k,m+1}))_{\Omega_\mathrm{f}} \\
&- (\mathrm{div}(J_{k,m}F^{-1}_{k,m}\nabla\psi F^{-1}_{k,m}v_{k,m}), z^p_{k,m})_{\Omega_\mathrm{f}} \\
&- k_{m+1}(\mathrm{div}(J_{k,m}F^{-1}_{k,m}\nabla\psi F^{-1}_{k,m}v_{k,m}), z^p_{m+1})_{\Omega_\mathrm{f}},
\end{aligned}$$

by the terms

$$\begin{aligned}
(a_v'^D)_{k,m}(\mathbf{u}_{k,m})(\varphi, z^p_{k,m}) =& k_m(\mathrm{div}(J_{k,m}F^{-1}_{k,m}\varphi), z^p_{k,m}))_{\Omega_\mathrm{f}}, \\
(a_u'^D)_{k,m}(\mathbf{u}_{k,m})(\psi, z^p_{k,m}) =& k_m(\mathrm{div}(J_{k,m}\,\mathrm{tr}(F^{-1}_{k,m}\nabla\psi)F^{-1}_{k,m}v_{k,m}), z^p_{k,m}))_{\Omega_\mathrm{f}} \\
&- (\mathrm{div}(J_{k,m}F^{-1}_{k,m}\nabla\psi F^{-1}_{k,m}v_{k,m}), z^p_{k,m})_{\Omega_\mathrm{f}}.
\end{aligned}$$

Thereby the adjoint pressure is taken fully implicitly again. Now the semi-discretized adjoint and semidiscretized-state equation match again.

5.3. Space Discretization

The time discretized formulation in Problem 5.3 is the starting point for a conforming Galerkin discretization in space. To this end, to find an approximate solution to the continuous problem, we construct finite dimensional subspaces $V_h^{(s)} \times V_h^{(s)} \times L_{\mathrm{f},h}^{(s)} \subset V \times V \times L_\mathrm{f}$ of order s .

5.3.1. Finite Element Spaces

For simplicity, we assume the boundary to be polygonal. We use two dimensional shape-regular meshes (see [35, 36, 44]) relaxed by introducing hanging nodes to allow for local mesh refinement. Cells are allowed to have nodes, which lie on midpoints of faces of neighboring cells. A mesh consists of quadrilateral or hexahedral cells K. They perform a non-overlapping cover of the computation domain Ω. As the fluid reference domain is fixed we can assume in addition, that the interface Γ_i does not cross any cell. Thereby the mesh always resolves the interface. The resulting triangulations is denoted by $\mathcal{T}_h = \{K\}$. Every sequence of meshes is supposed to have a patch structure: every element $K \in \mathcal{T}_h$ is part of 2^d elements arising from the same element by uniform refinement. The patch structure is needed to construct a local projection stabilization filter and later to evaluate the a posteriori error estimators in

Section 7.3. The discretization parameter in the reference configuration is denoted by h and is the maximal cell-wise constant that is given by the diameter h_K of the cell K.

On the mesh $\mathcal{T}_h$ we define as in [36, 44] the usual space of isoparametric finite element functions $\mathcal{V}_h^{(s)} \subset H^1(\Omega)$ with degree s and on the subset $\mathcal{T}_{\mathrm{f},h} \subset \mathcal{T}_h$ on the fluid domain the space

$$L_{\mathrm{f},h}^{(s)} := \left\{ \varphi|_{\Omega_\mathrm{f}} \;\middle|\; \varphi \in \mathcal{V}_h^{(s)} \text{ and } (\varphi, 1)_\mathrm{f} = 0 \right\}.$$

For the velocity and displacement variable we use the vector valued space

$$V_h^{(s)} = \left\{ \varphi \in (\mathcal{V}_h^{(s)})^d \;\middle|\; \varphi = 0 \text{ on } \Gamma_\mathrm{s} \cup \Gamma_\mathrm{f} \right\}$$

and for the pressure the space $L_{\mathrm{f},h}^{(s)}$. Thus we choose for fluid velocity and pressure the same space of equal order. In the numerical examples in Chapter 8, we always apply a do-nothing outflow condition. Thereby, the pressure is automatically normalized.

Concerning the test functions, it is unclear how to discretize the test space $V_{\mathrm{f},0} \oplus H_\mathrm{s}$. The first possibility is to split the space again and define finite element spaces on the fluid and solid subdomain. Unfortunately, the definition of finite element spaces on subdomains or the use of different spaces for test- and trial functions is in many software packages not possible. An alternative is the use of the finite element space $V_h^{(s)} \subset V_\mathrm{f} \oplus H_\mathrm{s}$ as test space. Now the test function does not vanish anymore on the interface Γ_i. Then, we have to subtract the natural boundary term $\langle\!\langle \sigma_\mathrm{m}(u)^T n, \psi \rangle\!\rangle_\mathrm{i}$ in the weak mesh motion equation. We define the space discretized mesh motion equation

$$\begin{aligned} a_{kh,m}^M(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\psi) :=& a_{k,m}^M(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\psi) \\ & - \langle \theta_m(\sigma_\mathrm{m})_{kh,m} n + (1-\theta_m)(\sigma_\mathrm{m})_{kh,m-1} n, \psi \rangle_\mathrm{i}. \end{aligned}$$

If we evaluate the additional boundary term via a residual method, it becomes clear that we obtain the same stiffness matrix as with separate finite element spaces on fluid and solid domain. Consequently, if we evaluate the boundary integral accurately enough the finite element functions on the interface will only contribute to the solid. Thus, there is no direct back coupling of the mesh motion into the solid. As the software package RoDoBo [133] used for numerical experiments later supports neither finite element spaces on subdomains, nor the evaluation of integrals along the interface, we use the test space V_h and neglect the additional boundary integral. Instead we apply a small elasticity parameter in the mesh motion such that the back coupling is small.

5.3.2. Local Projection Stabilization

As equal order elements are not inf-sup stable we have to stabilize the pressure. For instance, the streamline upwind Petrov-Galerkin (SUPG) method, introduced by Brooks and Hughes [38], can be used for fluid-structure interaction problems as in [146]. However, conventional residual-based stabilized finite elements for flow problems have the disadvantage, that "discretize" and "optimize" does not necessarily commute. If we first stabilize the state equation and then calculate the adjoint equation for the SUPG stabilization method, then the adjoint equation differs from the adjoint equation we get if we first calculate the adjoint equation

and then stabilize state and adjoint equation. This is not the case for the symmetric local projection stabilization (LPS) suggested by Becker and Braak in [16, 17]. A very detailed discussion on optimal control of the Navier-Stokes equations with stabilized finite elements can be found in [32].

For the considered examples, we are not in a convection dominant regime. Therefore, we focus on the pressure stabilization. We introduce the spatial interpolation operator

$$i_{2h/s}^{(1)} : \mathcal{V}_h^{(s)} \to \begin{cases} \mathcal{V}_{2h}^{(1)}, & s = 1 \\ \mathcal{V}_h^{(1)}, & s = 2. \end{cases}$$

For bi/tri-linear finite elements, the interpolation in the space $\mathcal{V}_{2h}^{(1)}$ on the coarse grid can be easily reconstructed, by making use of the assumed patch structure. Now we can define the filtering operator $\pi : \mathcal{V}_h^{(s)} \to \mathcal{V}_h^{(s)}$ by $\pi = \mathrm{Id} - i_{2h/s}^{(1)}$. Due to the moving fluid domain we obtain the stabilization term

$$\int_{I_m} \sum_{K \in \mathcal{T}_h} \alpha_K \big(J_{kh} F_{kh}^{-1} \nabla \pi(p_{kh}), F_{kh}^{-1} \nabla \pi(\xi) \big)_K.$$

Thereby, the stabilization parameter α_K defined on each cell K in the mth time step is given by

$$\alpha_K = \alpha_0 \frac{\check{h}_K^2}{6\nu + \check{h}_K \|v_{kh}\|_K}.$$

Due to the transformation we receive an additional stabilization term in the adjoint equation. As a result, "discretize" and "optimize" do not commute anymore. In addition the stabilization term now is highly nonlinear. Therefore, we decide to neglect the transformation and to use instead

$$a_{kh,m}^{\mathrm{LPS}}(p_{kh,m})(\xi) := k_m \sum_{K \in \mathcal{T}_h} \alpha_K \big(\nabla \pi(p_{kh,m}), \nabla \pi(\xi) \big)_K$$

with the approximate $\check{h}_K \approx h_K$. We still got very accurate solutions for this stabilization technique at least for the considered benchmark examples calculated within this thesis.

5.3.3. Space-Discretized State and Adjoint Equations

Altogether, we obtain the temporally and spatially discretized formulation :

Problem 5.5. *Let* $(\theta_m)_{m=1}^M \in [0,1]^M$ *and* $(v_{kh,0}, u_{kh,0}) = (v_{0,h}, u_{0,h})$. *Find* $(\mathbf{u}_{kh,m})_{m=1}^M \in (V_h^{(s)})^M \times (V_h^{(s)})^M \times (L_{\mathrm{f},h}^{(s)})^M$ *by iterating for* $m = 1, ..., M$:

$$\begin{aligned} a_{k,m}^F(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\varphi) + a_{k,m}^S(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\varphi) &= f_{k,m}(q)(\varphi) && \forall \varphi \in V_h^{(s)}, \\ a_{kh,m}^M(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\psi) + a_{k,m}^V(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\psi) &= 0 && \forall \psi \in V_h^{(s)}, \\ a_{k,m}^D(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\xi) + a_{kh,m}^{LPS}(p_{kh,m})(\xi) &= 0 && \forall \xi \in L_{\mathrm{f},h}^{(s)}. \end{aligned}$$

Remark 5.3. On every cell the inner products now consist of sums and products of polynomial functions. We can either evaluate the integrals over the domains Ω_s and Ω_f exactly or at least use a Gauß quadrature rule of high order, such that the quadrature error can be neglected.

Taken all together the discrete solution $\mathbf{u}_{kh}$ defined on the space-time cylinder is an element of the space-time discretized space $X_{kh} \subset X$ with $X_{kh} := W_{kh}^u \times W_{kh}^v \times L_{kh}^p$ and

$$W_{kh}^u := W_{kh}^v := \left\{ v \in C(\bar{I}; V_h^{(s)}) \;\middle|\; v\big|_{I_m} \in P_1(I_m, V_h^{(s)}), m = 1, 2, ..., M \right\},$$
$$L_{kh}^p := \left\{ p \in L^2(I; L_{\mathrm{f},h}^{(s)}) \;\middle|\; p\big|_{I_m} \in P_0(I_m, L_{\mathrm{f},h}^{(s)}), m = 1, 2, ..., M \right\}.$$

If we define in addition the space and time discretized test space $Y_{kh,\theta} := L_{kh,\theta}^v \times L_{kh,\theta}^u \times L_{kh}^p$ with

$$L_{kh,\theta}^v = L_{kh,\theta}^u := \left\{ \varphi \in L^2(I; V_h^{(s)}) \;\middle|\; \varphi\big|_{I_m} \in P_0^\theta(I_m, V_h^{(s)}), m = 1, 2, ..., M, \varphi(0) \in V_h^{(s)} \right\},$$

we can now write the discretized fluid-structure interaction equation in Problem 5.5 in a more compact form. The discrete solution of the fluid-structure interaction problem has to fulfill

Problem 5.6. *Find for $q \in Q$ a solution $\mathbf{u}_{kh} \in X_{kh}$ with initial conditions $v_{kh}(0) = v_0$ and $u_{kh}(0) = u_0$ such that*

$$a_{kh}(\mathbf{u}_{kh})(\boldsymbol{\varphi}) = f_k(q)(\boldsymbol{\varphi}) \quad \forall \varphi \in Y_{kh,\theta} \tag{5.5}$$

with

$$\begin{aligned} a_{kh}(\mathbf{u}_{kh})(\boldsymbol{\varphi}) := \sum_{m=1}^{M} \big[& a_{k,m}^F(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\varphi_m) + a_{k,m}^S(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\varphi_m) \\ & + a_{kh,m}^M(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\psi_m) + a_{k,m}^V(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\psi_m) \\ & + a_{k,m}^D(\mathbf{u}_{kh,m}, \mathbf{u}_{kh,m-1})(\xi_m) + a_{kh,m}^{LPS}(p_{kh,m})(\xi_m) \big], \\ f_k(q)(\varphi) := \sum_{m=1}^{M} & f_{k,m}(q)(\varphi_m). \end{aligned}$$

For the adjoint equation this method then results in

Problem 5.7. *Find $\mathbf{z}_{kh} \in Y_{kh,\theta}$*

$$a'_{kh,\mathbf{u}}(\mathbf{u}_{kh})(\mathbf{z}_{kh}, \boldsymbol{\varphi}) - (\psi(0), z_{kh}^u(0)) - (\varphi(0), z_{kh}^v(0)) = J'_{\mathbf{u}}(\mathbf{u}_{kh})(\boldsymbol{\varphi}) \quad \forall \varphi \in X_{kh}. \tag{5.6}$$

Thereby the semi-linear form $a'_{kh,\mathbf{u}}(\cdot)(\cdot,\cdot)$ is defined in the same spirit as the semi-linear forms in the state equation. We would like to highlight that the adjoint mesh motion equation in every adjoint time step is given by

$$\begin{aligned} ({a'_u}^M)_{kh,m}(\mathbf{u}_{kh,m})(\psi, z_{kh,m}^u, z_{kh,m+1}^u) = {} & ({a'_u}^M)_{k,m}(\mathbf{u}_{kh,m})(\psi, z_{kh,m}^u, z_{kh,m+1}^u) \\ & - [k_m \theta_m \langle \sigma'_{\mathrm{m},u}(\psi) n, z_{kh,m}^u \rangle_{\mathrm{i}} + k_{m+1} \theta_{m+1} \langle \sigma'_{\mathrm{m},u}(\psi) n, z_{kh,m+1}^u \rangle_{\mathrm{i}}]. \end{aligned}$$

In the adjoint equation the directional derivative is applied on the test variable now in the additional boundary term. Therefore, the boundary integral on the interface Γ_{i} enforces the original zero Dirichlet boundary condition in some very weak sense. As the chosen modified LPS stabilization term is symmetric, we would have gotten the same formulation if we had applied the Lagrange formalism on the stabilized Lagrangian.

5.4. Discretization of the Control Variable

Up to this point the control space has not been discretized. It is possible to solve optimal control problems without explicitly discretizing the control space, see [81]. For a control $q \in L^2(I;H)$ acting as volume force in the right-hand side, this would result in $q_{kh} \in L^v_{kh,\theta}$. Every time step in the theta time-stepping scheme consist of three substeps. Therefore, we are going to refine the time-grid by bisection of every macro time interval and then introducing the intermediate steps again later in the adaptive algorithm (Algorithm 7.1). The resulting time-grid is not going to be a subgrid of the previous time-grid. To start the optimization algorithm with the control calculated on the coarse grid then is not straightforward. The approach presented in [81] exhibits the difficulty that if the control enters in a nonlinear fashion the implementation of the derived "natural" discretization is usually not straightforward. In addition, it is quite often not necessary to use a high dimensional control space. Hence, we discretize the control space independent of the state and dual variable. Similar to the primal variable, we discretize first in time and later in space.

We partition the time interval $I = (0,T)$ in N subintervals $I_n = (t_n, t_{n-1})$ with $t_0 < t_1 < ... < t_n = T$. We demand the control time grid to be a subgrid of the macro time grid introduced in Section 5.1.3. The set of time points used for the control then has to be a subset of the time points used for the discretization of the state equation. For instance, we can either discretize the control using a piecewise linear trial space $Q_d \subset Q$ with

$$Q_d = \{q_k \in C(\bar{I};H) | q_k\big|_{I_n} \in P_1(I_n,H), n = 1,2,...,N\}$$

or we can use the piecewise constant and discontinuous test space

$$Q_d = \{q_k \in L^2(I;H) | q_k\big|_{I_n} \in P_0(I_m,H), n = 1,2,...,N\}.$$

Both spaces are easy to handle and due to the Galerkin structure, we are able to estimate the discretization error of the control variable as for state and dual variable in Chapter 7.

If the control variable is not only time-dependent, but distributed in space, we have to choose an appropriate Galerkin space here, too. For example, it can be natural to have jumps in the desired parameter. In that case we should use the space of cellwise constant functions. Alternatively, we can use the finite element space $Q_d = \mathcal{V}_h \subset H$, which already has to be available in the used software package to discretize the state and adjoint variables.

For a Neumann boundary control, the space discretization is done by traces of functions $v_h \in V_h$.

5.5. Discrete Control Problem

We denote in the following with $Q_d \subset Q$ the fully discretized control space and let $(q_\sigma, \mathbf{u}_\sigma) \in Q_d \times X_{kh}$ be the solution of the fully discrete optimal control problem:

Problem 5.8 (Fully discrete optimal control problem)**.**

$$\min_{q_\sigma \in Q_d} \mathcal{J}(q_\sigma, \mathbf{u}_\sigma) = \mathcal{J}(\mathbf{u}_\sigma) + \|q_\sigma\|^2_Q$$

subject to $\mathbf{u}_\sigma \in X_{kh}$ *with* $v_\sigma(0) = v_0$, $u_\sigma(0) = u_0$ *and*

$$a_{kh}(\mathbf{u}_\sigma)(\boldsymbol{\varphi}) = f_k(q_\sigma)(\boldsymbol{\varphi}) \quad \forall \boldsymbol{\varphi} \in Y_{kh,\theta},$$
$$q_a \leq q_\sigma \leq q_b.$$

If we define the fully discrete Lagragian $\mathcal{L}_{kh} : X_{kh} \times Y_{kh,\theta} \times Q_d \to \mathbb{R}$ with

$$\begin{aligned}\mathcal{L}_{kh}(q_\sigma, \mathbf{u}_\sigma, \mathbf{z}_\sigma) :=& \mathcal{J}(q_\sigma, \mathbf{u}_\sigma) - a_{kh}(\mathbf{u}_\sigma)(\mathbf{z}_\sigma) + f_k(q_\sigma)(z_\sigma^v) \\ &+ (u_0 - u_\sigma(0), z_\sigma^u(0)) + (v_0 - v_\sigma(0), z_\sigma^v(0))\end{aligned}$$

and apply the Lagrange formalism on the discrete Lagrangian $\mathcal{L}_{kh}$, the optimal solution $(q_\sigma, \mathbf{u}_\sigma)$ of the discrete optimal control problem has to solve the Karush-Kuhn-Tucker condition:

1. The optimal state $\bar{\mathbf{u}}_\sigma = \mathbf{u}_\sigma(\bar{q}_\sigma) \in X_{kh}$ has to solve

   $$\mathcal{L}'_{kh,\mathbf{z}}(\bar{q}_\sigma, \bar{\mathbf{u}}_\sigma, \bar{\mathbf{z}}_\sigma)(\boldsymbol{\varphi}) = 0 \quad \forall \boldsymbol{\varphi} \in Y_{kh,\theta}.$$

   If we compare the resulting equations with (5.5), we can observe that $\bar{\mathbf{u}}_\sigma = (\bar{v}_\sigma, \bar{u}_\sigma, \bar{p}_\sigma) \in X_{kh}$ fulfills the initial conditions $u_\sigma(0) = u_0$ and $v_\sigma(0) = v_0$ and solves the space-time discretized fluid structure interaction formulation in Problem 5.6 with control $q_\sigma \in Q_d \subset Q$.

2. The optimal adjoint variable $\bar{\mathbf{z}}_\sigma = \mathbf{z}_\sigma(\bar{q}) \in Y_{kh,\theta}$ has to solve

   $$\mathcal{L}'_{kh,\mathbf{u}}(\bar{q}_\sigma, \bar{\mathbf{u}}_\sigma, \bar{\mathbf{z}}_\sigma)(\boldsymbol{\varphi}) = 0 \quad \forall \boldsymbol{\varphi} \in X_{kh}.$$

   If we compare the resulting fully discrete adjoint equation with the adjoint equation (5.6), we notice that $\bar{\mathbf{z}}_\sigma \in Y_{kh,\theta}$ is a solution of the adjoint fluid-structure interaction equation in Problem 5.7.

3. For the optimal control $\bar{q}_\sigma \in Q_{ad,d} := \{\, q \in Q_d \mid q_a \leq q \leq q_b \,\}$ it has to hold the necessary optimality condition

   $$\mathcal{L}'_{kh,q}(\bar{q}_\sigma, \bar{\mathbf{u}}_\sigma, \bar{\mathbf{z}}_\sigma)(\delta q - \bar{q}_\sigma) \geq 0 \quad \forall \delta q \in Q_{ad,d}.$$

We can observe that the same conditions occur if we discretize the necessary optimality condition in Section 4.3.3 systematically. Therefore, “optimize” and “discretize” commute here.

6. Solution Algorithms

To compute the optimal solution $(q_\sigma, \mathbf{u}_\sigma) \in Q_d \times X_{kh}$ of the considered fully discrete optimization configuration in Problem 5.8, we can use standard optimization algorithms as gradient as well as Newton algorithm. Because of the Galerkin discretization approach "optimize" and "discretize" commute. We can define the optimization algorithm either for the fully discrete optimization problem or on the continuous level and discretize afterward. We are going to follow the second approach, as we thereby obtain convergence of the optimization algorithm in the natural norms.

To compute the gradient in an arbitrary direction, we use the Lagrange formalism presented in Section 3.3.2 and 4.3.2. If the variable $\mathbf{u} \in X$ is the solution of the fluid-structure interaction problem and $\mathbf{z} \in Y$ solves the adjoint equation, then we have shown that we get

$$j'(q)(\delta q) = L'_q(q, \mathbf{u}, \mathbf{z})(\delta q).$$

For one gradient evaluation, we have to solve two PDEs. First we have to solve the nonlinear state equation over the whole time frame. Second, we have to compute the adjoint equation back in time. As the fluid-structure interaction problem is still a nonlinear problem after discretization in space and time, we apply a Newton algorithm presented in Section 6.1 in every time step. In Section 6.2, we comment on the linear solvers which can be used to solve the linear systems occurring in every Newton step. Then, using the sensitivity information of the adjoint solution we can apply a Quasi-Newton algorithm. Here we use the Limited Memory BFGS algorithm (LBFGS) presented in Section 6.3. The reduced storage consumption of the LBFGS algorithm enables to solve optimal control problems with high dimensional control spaces. An overview how the different algorithms interact is given in the abstract Algorithm 6.1.

We want to highlight that the structure of the presented optimization algorithm differs from the approaches used in literature for optimal control of fluid-structure interaction. At least in the literature cited in the introduction to Section 4.3, the control variable is updated in every time step during the solution process. In Algorithm 6.1 the state and adjoint equation is solved over the whole time-interval and the control is not updated during the forward and backward simulation. Such an approach is most commonly used in the optimal control community to have a consistent optimization algorithm. With refinement of the control space as well as refinement of state and adjoint variable we can expect convergence of the discrete optimal solution against the solution of the continuous optimal control problem. We refer to [114, 116] for a detailed proof for optimal control of a parabolic PDE. Furthermore, the additional sensitivity information with respect to time enables the use of cost functionals with observation only at certain time-points, as for example the end-time point.

Algorithm 6.1: Abstract optimization algorithm

Set $q_\sigma \in Q_d$, $\mathbf{u}_\sigma(0) = \mathbf{u}_0$
while $\|\nabla j(q_\sigma)\| > tol$ **do**
  Compute the solution $\mathbf{u}_\sigma$ of the fluid-structure interaction equation with control q_σ
  for $m = 1 : M$ **do**
    In every time step apply the Newton algorithm
    while $\rho_m > tol$ **do**
      Construct Jacobian A_m
      Solve Newton system $A_m \mathbf{w} = b_m$
      Evaluate residual of the state equation ρ_m and update $\mathbf{u}_{\sigma,m} = \mathbf{u}_{\sigma,m} + \mathbf{w}$
    end while
  end for
  Compute the solution $\mathbf{z}_\sigma$ of the adjoint fluid-structure interaction problem
  for $m = M : 1$ **do**
    Construct adjoint Jacobian A_m^T
    Solve linear system $A_m^T \mathbf{z}_{\sigma,m} = b_m$
  end for
  Evaluate gradient $\nabla j(q_\sigma)$ and compute BFGS update $H d_\sigma = -\nabla j(q_\sigma)$
  Update BFGS Matrix H
  Set $q_\sigma = q_\sigma + d_\sigma$
end while

6.1. Newton Algorithm

In every LBFGS step we have to solve the fully discrete state (see Problem 5.5) and the adjoint fluid-structure interaction problem (see Problem 5.7). The adjoint equation is known to be a linear problem. On the contrary the fluid-structure interaction problem is highly nonlinear. To compute a solution numerically, we apply a Newton algorithm in every time step on the fully monolithic formulation. The resulting algorithm is very robust and fast converging.

To write the system of equations in Problem 5.5 in a compact way we introduce the LPS stabilized fully-discrete semi-linear form $a_{kh,m}(\mathbf{u}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\boldsymbol{\varphi})$:

$$\begin{aligned} a_{kh,m}(\mathbf{u}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\boldsymbol{\varphi}) :=& a^F_{k,m}(\mathbf{u}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\varphi) + a^S_{k,m}(\mathbf{u}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\varphi) \\ &+ a^M_{kh,m}(\mathbf{u}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\psi) + a^V_{k,m}(\mathbf{u}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\psi) \\ &+ a^D_{k,m}(\mathbf{u}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\xi) + a^{\mathrm{LPS}}_{kh,m}(p_{k,m})(\xi). \end{aligned}$$

For detailed definition of those particular semi-linear forms we refer to Problem 5.5 and Problem 5.2. To compute a solution of the nonlinear fluid-structure interaction equation we have to determine the solution $\mathbf{u}_{\sigma,m} \in V_h^{(s)} \times V_h^{(s)} \times L^{(s)}_{\mathrm{f},h}$ in every time step such that

$$a_{kh,m}(\mathbf{u}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\boldsymbol{\varphi}) = f_{k,m}(q_\sigma)(\boldsymbol{\varphi}) \quad \forall \boldsymbol{\varphi} \in V_h^{(s)} \times V_h^{(s)} \times L^{(s)}_{\mathrm{f},h} \tag{6.1}$$

having the solution of the previous time step $\mathbf{u}_{\sigma,m-1} \in V_h^{(s)} \times V_h^{(s)} \times L^{(s)}_{\mathrm{f},h}$ available.

To compute the solution $\mathbf{u}_{\sigma,m} \in V_h^{(s)} \times V_h^{(s)} \times L_{\mathrm{f},h}^{(s)}$ we apply a Newton algorithm on (6.1). We start the Newton algorithm with the solution of the previous time step $\mathbf{u}_{\sigma,m-1}$. If the initial guess $\mathbf{u}_{\sigma,m-1}$ is not close enough, a simple damping strategy can often help to provide convergence. For such complex systems as the fluid-structure interaction problem often times finite difference approaches are used to approximate the Jacobian. The difficulty thereby lies in the choice of parameters such that the accuracy of the Jacobian is high. Hence, we build up the exact Jacobian:

$$a'_{kh,m}(\mathbf{u}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\mathbf{w}_\sigma, \boldsymbol{\varphi}) = \lim_{s\to 0} \frac{d}{ds} a_{kh,m}(\mathbf{u}_{\sigma,m} + s\mathbf{w}_\sigma, \mathbf{u}_{\sigma,m-1})(\boldsymbol{\varphi})\Big|_{s=0}.$$

To compute the linearization we can apply the same technique as we used in Section 4.3.2 for deriving the adjoint equation. With the slight difference of now applying the linearization on the time space discretized semi-linear form. In Appendix B, we summarize in more detail how to compute the linearization of the fluid-structure interaction problem. The linearization of the space-time discretized formulation works in the same way. The main difficulty lies again in the domain motion hidden in the ALE mapping. The applied Newton algorithm is summarized in Algorithm 6.2.

Algorithm 6.2: Newton algorithm to solve nonlinear equations

Set $\mathbf{u}_{\sigma,m} = \mathbf{u}_{\sigma,m-1}$
while $\rho_m > tol$ **do**
  Compute Newton update $\mathbf{w}_\sigma$ such that

$$a'_{kh,m}(\mathbf{u}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\mathbf{w}_\sigma, \boldsymbol{\varphi}) = f_{k,m}(q_\sigma)(\boldsymbol{\varphi}) - a_{kh,m}(\mathbf{u}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\boldsymbol{\varphi})$$

  Set $\tilde{\mathbf{u}}_{\sigma,m} := \mathbf{u}_{\sigma,m} + \mathbf{w}_\sigma$
  Compute residual $\rho_m := \|f_{k,m}(q_\sigma)(\cdot) - a_{kh,m}(\tilde{\mathbf{u}}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\cdot)\|$
  while $\rho_m > \rho_{m-1}$ **do**
    $i = i + 1$
    $\tilde{\mathbf{u}}_{\sigma,m} := \mathbf{u}_{\sigma,m} + \gamma^i \mathbf{w}_\sigma$
    Compute residual $\rho_m = \|f_{k,m}(q_\sigma)(\cdot) - a_{kh,m}(\tilde{\mathbf{u}}_{\sigma,m}, \mathbf{u}_{\sigma,m-1})(\cdot)\|$
  end while
  Set $\mathbf{u}_{\sigma,m} := \tilde{\mathbf{u}}_{\sigma,m}$
end while

6.2. Linear Solver

By applying the Newton algorithm on the space-time discretized formulation (6.1), we can rewrite the resulting linear system to solve for a Newton step in compact form as

$$A_m w_\sigma = b_m.$$

The Jacobian matrix A_m and the vector b_m depend on the control q_σ, on the solution $\mathbf{u}_{\sigma,m}$ and $\mathbf{u}_{\sigma,m-1}$. Hence, we actually have to build up the Jacobian in every Newton step. We

would like to refer to [126] for an analysis of the structure of the system matrix A_m and a convergence test of the Newton algorithm if the Jacobian is not updated in every Newton step. For the adjoint equation the discretization in space and time results in the adjoint system

$$A_m^T \mathbf{z}_{\sigma,m} = b_m.$$

Although the adjoint problem is linear, we have to build up the matrix A_m^T in every time step as the matrix depends on the solution of the state equation $\mathbf{u}_{\sigma,m+1}$,$\mathbf{u}_{\sigma,m}$ and $\mathbf{u}_{\sigma,m-1}$.

For the presented numerical examples in Chapter 8, we use the direct solver UMFPACK, which uses an incomplete LU decomposition to compute a solution of the linear system. For more complex configurations, it is necessary to use iterative solvers as the computational time and memory use of a direct solver gets too large. As the derived linear system is huge, very ill conditioned and without structures like symmetry or positivity, the construction of preconditioners is not straightforward.

To circumvent this difficulty, fluid-structure interaction problems are often solved by using partitioned approaches. The problem is split into the Navier-Stokes system, the mesh motion equation and the elastic structure system. Then, one after the other the problems are solved separately and coupled by boundary conditions. A good overview of partitioned approaches and semi-implicit approaches is given in [60, 97, 61] and in the books [40, 41]. Now, well tuned preconditioners for the subproblems can be used to solve the resulting linear systems, which is why these algorithms are widely used at the moment. Such weakly coupled partitioned approaches show instabilities if the coupling is stiff ("added mass effect") as in most hemodynamic applications, see [43]. Motivated by partitioned algorithms several preconditioners for the monolithic fluid-structure interaction formulation have been developed in the last decade as for instance [11, 71, 10, 129].

6.3. Limited memory BFGS Algorithm

As the evaluation of the gradient $j'(q)(\delta q)$ includes the solution of the fluid-structure interaction problem, it is computationally quite costly. So fast converging algorithms needing few gradient evaluations should be preferred. Gradient methods converge slowly and need a step size control, which involves additional evaluations of the functional. Therefore we follow the suggestion in [82] and use an optimization algorithm of second order.

If we apply Newton's method on the necessary optimality condition

$$j'(\bar{q})(\delta q) = 0 \quad \forall \delta q \in Q,$$

we have to solve

$$(\nabla^2 j(q)\delta q, \delta\tau)_Q = -(\nabla j(q), \delta\tau)_Q \quad \forall \delta q, \delta\tau \in Q,$$

whereby the gradient $\nabla j(q) \in Q$ and the Hessian $\nabla^2 j(q) : Q \to Q$ are defined by Riesz's representation

$$(\nabla j(q), \delta q)_Q := j'(q)(\delta q) \quad \forall \delta q \in Q,$$
$$(\nabla^2 j(q)\delta q, \delta\tau)_Q := j''(q)(\delta q, \delta\tau) \quad \forall \delta q, \delta\tau \in Q.$$

For a wide range of optimal control problems with tracking type functional the Newton algorithm is known to be convergent with second order (see for example [82]). To evaluate the action of the Hessian on a vector, we have to solve two additional PDEs, the so called tangential and additional adjoint equation. To avoid having to evaluate these PDEs, we are going to use a quasi-Newton method, which uses an approximation of the Hessian. As quasi-Newton methods still have locally superlinear convergence properties they provide a good alternative to Newton algorithms. In the following, we describe a BFGS algorithm. We will denote with $H_k : Q \to Q$ the approximation of the Hessian which is defined via an updating formula.

Algorithm 6.3: BFGS algorithm for continuous optimization problem

---

Set $q_0 \in Q$, $H_0 : Q \to Q$, $\varepsilon > 0$ and $l = 0$
while $\|\nabla j(q_l)\|_Q > \varepsilon$ **do**
  Solve

$$(H_l d_l, \delta q)_Q = -(\nabla j(q_l), \delta q)_Q \quad \forall \delta q \in Q$$

  Set $q_{l+1} = q_l + d_l$ and $y_l = \nabla j(q_{l+1}) - \nabla j(q_l)$
  Set

$$H_{l+1} := H_l + \frac{(y_l, \cdot)_Q}{(d_l, y_l)_Q} y_l - \frac{(H_l d_l, \cdot)_Q}{(d_l, H_l d_l)_Q} H_l d_l$$

  Set $l = l + 1$
end while

In every step the algorithm only needs one forward solution of the fluid-structure interaction problem and one backward solve of its adjoint problem. A detailed convergence analysis of Algorithm 6.3 in the context of PDE-constrained optimization can be found in [82] and [91].

After discretization of the control space, the control $q_\sigma \in Q_d$ can be expressed by its basis functions $\varphi^i \in Q_d$ such that $q_\sigma = \sum_{i=1}^n \bar{q}^i \varphi^i$ with $\bar{q} = (\bar{q}^1, ..., \bar{q}^n) \in \mathbb{R}^n$ and $n = \dim(Q_d)$. Then Riesz's representation of the gradient and the approximation of the Hessian H can be written with respect to the basis functions, too. We get as representation of the gradient

$$\nabla j(q_\sigma) = \sum_{i=1}^n \bar{g}^i \varphi^i$$

and if we apply H on $\delta q = \sum_{i=1}^n \overline{\delta q}^i \varphi^i$ with $\overline{\delta q} \in \mathbb{R}^d$ we obtain the representation

$$H\delta q = \sum_{j=1}^n \overline{\delta q}^j H\varphi^j = \sum_{i=1}^n \Big(\sum_{j=1}^n \overline{H}^{i,j} \overline{\delta q}^i \Big) \varphi^i$$

with the vector $\bar{g} \in \mathbb{R}^n$ and the symmetric matrix $\overline{H} \in \mathbb{R}^{n \times n}$. In the following by writing $\bar{g}$ we will always denote the vector representation of $\nabla j(q_\sigma)$, which we obtain by solving the state and adjoint equation. The inner products in Q then simplify for $\delta q, \delta\tau \in Q_d$ to

$$(\delta q, \delta\tau)_Q = \overline{\delta q}^T \overline{M} \overline{\delta\tau}$$

with the control mass Matrix $\overline{M} \in \mathbb{R}^{n \times n}$ fulfilling $\overline{M}^{i,j} = (\varphi^i, \varphi^j)_Q$ for $i, j = 1, ...n$. Then, Algorithm 6.3 reduces to Algorithm 6.4.

Algorithm 6.4: BFGS algorithm for discrete optimization problem

Set $\bar{q}_0 \in \mathbb{R}^n$, $\overline{H}_0 : \mathbb{R}^{n\times n}$ symmetric, positive definite , $\varepsilon > 0$ and $l = 0$
while $\bar{g}_l^T \overline{M} \bar{g}_l > \varepsilon$ **do**
Solve $\overline{M}\overline{H}_l \bar{d}_l = -\overline{M}\bar{g}_l$
Set $\bar{q}_{l+1} = \bar{q}_l + \bar{d}_l$ and $\bar{y}_l = \bar{g}_{l+1} - \bar{g}_l$
Set

$$\overline{M}\overline{H}_{l+1} := \overline{M}\overline{H}_l + \frac{\overline{M}\bar{y}_l(\overline{M}\bar{y}_l)^T}{\bar{y}_l^T \overline{M}\bar{d}_l} - \frac{\overline{M}\overline{H}_l \bar{d}_l \bar{d}_l^T \overline{M}\overline{H}_l}{\bar{d}_l^T \overline{M}\overline{H}_l \bar{d}_l}$$

Set $l = l + 1$
end while

In order to have not only local convergence, the authors in [72] suggest to use a Powell-Wolfe step size control instead of an Armijo rule, which is usually used for Newton-type algorithms, to guarantee that the update $\overline{H}_{l+1}$ is still symmetric and positive definite. This is essential to guarantee well posedeness of Algorithm 6.4. The convergence results in [82, 91] can immediately be extended to Algorithm 6.4, if we use a Galerkin discretization approach. Thus we can expect at least linear convergence for a sequence of controls $\bar{q}_l \in \mathbb{R}^d$. For locally superlinear convergence, the authors have to demand as starting matrix $H_0 = J_{qq}(\bar{q})$, which corresponds for standard L^2-Tikhonov regularization to $\overline{H}_0 = \alpha\,\mathrm{Id}$. An additional advantage of the BFGS algorithm is that we do not have to compute the inverse of $\overline{H}_l$, but can instead use the inverse updating formula

$$\overline{B}_{l+1} := \overline{B}_l + \frac{(\bar{d}_l - \overline{B}_l\overline{M}\bar{y}_l)\bar{d}_l^T + \bar{d}_l(\bar{d}_l - \overline{B}_l\overline{M}\bar{y}_l)^T}{(\overline{M}\bar{y}_l)^T\bar{d}_l} - \frac{(\bar{d}_l - \overline{B}_l\overline{M}\bar{y}_l)^T\overline{M}\bar{y}_l\bar{d}_l\bar{d}_l^T}{((\overline{M}\bar{y}_l)^T\bar{d}_l)^2}.$$

Using elementary transformations, one can show that $\overline{B}_{l+1}\overline{M}\overline{H}_{l+1} = \mathrm{Id}$. For more details, we refer the reader to [72].

If we store the update $\bar{d}_l$ and $\bar{y}_l$ in every step, the Matrix $\overline{B}_{l+1}$ can be constructed in every step alternatively via the formula

$$\begin{aligned}\overline{B}_{l+1} =& \overline{V}_l^T\overline{V}_{l-1}^T\cdots\overline{V}_0^T\overline{B}_0\overline{V}_0\cdots\overline{V}_{l-1}\overline{V}_l \\ &+ \rho_0\overline{V}_l^T\overline{V}_{l-1}^T\cdots\overline{V}_1^T\bar{d}_0\bar{d}_0^T\overline{V}_0\cdots\overline{V}_{l-1}\overline{V}_l \\ &\vdots \\ &+ \rho_l\bar{d}_l\bar{d}_l^T,\end{aligned}$$

whereby the variables ρ_l and $\overline{V}_l$ are defined as $\rho_l = \frac{1}{(\overline{M}\bar{y}_l)^T\bar{d}_l}$ and $\overline{V}_l = \mathrm{Id} - \rho_l\overline{M}\bar{y}_l\bar{d}_l^T$. Especially for highly nonlinear problems, the gradient steps at the beginning are probably not of interest for the BFGS algorithm after some time. One possibility would be to restart the algorithm. Alternatively the construction formula for $\overline{B}_{l+1}$ enables to use only the gradient

information $\bar{y}_l$ and update $\bar{d}_l$ of the last m steps. We get

$$\begin{aligned}\overline{\mathcal{B}}_{l+1} =& \overline{V}_l^T \overline{V}_{l-1}^T \cdots \overline{V}_{l-m_l+1}^T \overline{B}_0 \overline{V}_{l-m_l+1} \cdots \overline{V}_{l-1} V_l \\ &+ \rho_{l-m_l+1} \overline{V}_l^T \overline{V}_{l-1}^T \cdots \overline{V}_{l-m_l+2}^T \bar{d}_{l-m_l+1} \bar{d}_{l-m_l+1}^T \overline{V}_{l-m_l+2} \cdots \overline{V}_{l-1} \overline{V}_l \\ &\vdots \\ &+ \rho_l \bar{d}_l \bar{d}_l^T .\end{aligned}$$

The advantage of using only the last m steps is especially interesting if the dimension of the control space gets large and storing of the full matrix $\overline{B}_{l+1}$ gets costly. As we only have to store the vectors $\bar{d}_l$ and $\bar{y}_l$ the storage place needed reduces from $O(n^2)$ to $O(mn)$. Typically $m \in 3, 4..., 9$ can be chosen very small such that the reduction of storage place is large. In addition, as we are going to see in Algorithm 6.5, the action of the matrix $\overline{\mathcal{B}}_{l+1}$ on a vector can be computed with $O(mn)$ evaluations without constructing the matrices $\overline{V}_l$. However, we do not have locally superlinear convergence for the LBFGS algorithm anymore. Nevertheless still very fast convergence can be seen in several numerically examples.

Algorithm 6.5: Limited Memory BFGS algorithm

Set $\bar{q}_0 \in \mathbb{R}^n$, $\overline{B}_0 : \mathbb{R}^{n \times n}$ symmetric, positive definite, $\varepsilon > 0$ and $l = 0$
Solve $\bar{d}_0 = -\overline{B}_0 \bar{g}_0$
Set $\bar{q}_1 = \bar{q}_0 + \bar{d}_0$ and $l = 0$
while $\bar{g}_{l+1}^T \overline{M} \bar{g}_{l+1} > \varepsilon$ **do**
  Set $\bar{y}_l = \bar{g}_{l+1} - \bar{g}_l$ and $m_l = \min\{l, m\}$
  for $j = l - m_l + 1, ..., l$ **do**
    $\rho_j = \frac{1}{\bar{d}_j^T \overline{M} \bar{y}_j}$
  end for
  Set $\bar{a}_{l+1} = \overline{M} \bar{g}_{l+1}$
  for $i = l, l-1, ..., l - m_l + 1$ **do**
    $\alpha_i = \rho_i \bar{d}_i^T \bar{a}_{i+1}$
    $\bar{a}_i = \bar{a}_{i+1} - \alpha_i \overline{M} \bar{y}_i$
  end for
  $\bar{b}_{l-m_l+1} = \overline{B}_0 \bar{a}_{l-m_l+1}$
  for $i = l - m_l + 1, l - m_l + 2, ..., l$ **do**
    $\beta_i = \rho_i (\overline{M} \bar{y}_i)^T \bar{b}_i$
    $\bar{b}_{i+1} = \bar{b}_i + (\alpha_i - \beta_i) \bar{d}_i$
  end for
  Set $\bar{d}_{l+1} = \bar{b}_{l+1}$, $\bar{q}_{l+2} = \bar{q}_{l+1} + \bar{d}_{l+1}$ and $l = l + 1$
end while

A very detailed proof of the presented Limited Memory BFGS Algorithm 6.5 can be found in [72]. For the first m steps, the algorithm coincides with the standard BFGS Algorithm 6.4. Numerical experiments have shown that it is advantageous to start the LBFGS algorithm in every step with the matrix $B_0 = \gamma_l \alpha^{-1} M^{-1}$ with $\gamma_l = \frac{d_l^T M y_l}{\|y_l\|^2}$. For more detail we refer to [74], [107] or [154].

7. Dual-Weighted Residual Error Estimator

In this chapter, we are going to derive a dual-weighted residual error estimator for optimal control of a fluid-structure interaction problem. The error estimator is based on the Dual-Weighted Residual method (DWR) developed by Becker and Rannacher [19, 20]. Such DWR estimators have already been applied successfully for estimating the discretization error of partial differential equations for example in [20, 80, 125, 33, 34, 28] and have also been extended to optimal control problems in [21, 22, 18, 113]. Especially for the latter, it is particularly advantageous to use the DWR estimator. The error in the cost functional can be estimated and the computational cost is quite small as we do not need to calculate the solution of an additional auxiliary problem.

For simplicity, we assume that we have an optimal control problem with no control constraints. We estimate the error between the solution $(q, \mathbf{u}) \in Q \times X$ of the continuous optimal control problem (Problem 4.6) and the solution $(q_\sigma, u_\sigma) \in Q_d \times X_{kh}$ of the discretized problem (Problem 5.8) with respect to the cost functional $\mathcal{J}$:

$$\mathcal{J}(q, \mathbf{u}) - \mathcal{J}(q_\sigma, \mathbf{u}_\sigma) \approx \eta_k + \eta_h + \eta_d.$$

Thereby the errors caused due to discretization of the state variable in time and space are approximated by η_k and η_h. The error due to discretization of the control variable is given by η_d. Splitting up into η_k, η_h and η_d allows for balancing the different error contributions within an adaptive refinement algorithm. The optimal control computed on a coarse grid is usually an excellent initial guess to start the optimization algorithm on the fine grid. Thus we can profit additionally from the solution process on different grids.

Concerning the error due to time discretization we follow the approach suggested by Meidner and Richter in [111] and [112], which can be used to compute the error with respect to the functional of interest for arbitrary theta time-stepping schemes. In [75] a similar approach for the damped Crank-Nicolson time-stepping scheme can be found. A DWR-error estimator to estimate the spatial discretization error for the Navier-Stokes equations has been successfully applied in [28]. In [55, 128, 62, 144] and [150] a posteriori error estimation for a stationary fluid-structure interaction problem has been analyzed. We are going to extend these results to error estimation for the time-dependent fluid-structure interaction problem.

We first state in Section 7.1 the residuals needed to evaluate the DWR estimator, then we prove in Section 7.2 that the remainder terms are of higher-order for an optimal control problem constrained by the linear fluid-structure interaction problem. In Section 7.3, we derive a higher-order reconstruction to evaluate the weights. The evaluation of the error via a residual functional enables a localization of the error estimator and thereby adaptive refinement. In Section 7.3, we present one localization strategy to refine the space, time, and control grid. The local error indicators enable adaptive refinement as introduced in Section 7.4.

7.1. A Posteriori Error Estimation for Optimal Control

In order to separate the influences of the different discretization steps on the error in the cost functional we split the error into

$$\begin{aligned}\mathcal{J}(q,\mathbf{u})-\mathcal{J}(q_\sigma,\mathbf{u}_\sigma) =&\big(\mathcal{J}(q,\mathbf{u})-\mathcal{J}(q_k,\mathbf{u}_k)\big)+\big(\mathcal{J}(q_k,\mathbf{u}_k)-\mathcal{J}(q_{kh},\mathbf{u}_{kh})\big)\\ &+\big(\mathcal{J}(q_{kh},\mathbf{u}_{kh})-\mathcal{J}(q_\sigma,\mathbf{u}_\sigma)\big),\end{aligned}$$

whereby we denote by $(q_k,\mathbf{u}_k)\in Q\times X_k$ the time-discretized solution and by $(q_{kh},\mathbf{u}_{kh})\in Q\times X_{kh}$ the optimal solution after space and time discretization with a control in the continuous space. As already stated in the introduction of this Chapter the solution of the continuous optimal control problem (Problem 4.6) is given by $(q,\mathbf{u})\in Q\times X$ and the solution of the fully discretized problem (Problem 5.8) is constituted by $(q_\sigma,u_\sigma)\in Q_d\times X_{kh}$. We respectively apply the error estimator derived in [19, 20] on those three pairs.

To solve the nonlinear FSI problem numerically as suggested in Section 6.1 and 6.2, we have to choose on the one side appropriate discrete Galerkin spaces and apply on the other side a quadrature rule to evaluate the integrals in time and space. To obtain an energy conserving, stable time-stepping scheme, which transports information only in one direction, we evaluate the time-integrals using the theta-trapezoidal rule (5.2). Therefore, the semi-discretized optimal solution $(q_k,\mathbf{u}_k)\in Q\times X_k$ is a stationary point of the time discretized Lagrange functional

$$\begin{aligned}\mathcal{L}_k(q_k,\mathbf{u}_k,\mathbf{z}_k) =&\mathcal{J}(q_k,\mathbf{u}_k)-a_k(\mathbf{u}_k)(\mathbf{z}_k)+f_k(q_k)(z_k^v)\\ &+(u_0-u_k(0),z_k^u(0))+(v_0-v_k(0),z_k^v(0)).\end{aligned}$$

The detailed structure of the time-discretized semi-linear form $a_k(\cdot)(\cdot)$ is laid out in Problem 5.2.

By using equal order elements to discretize in space, which are not inf-sub stable, we make use of a local projection stabilization. Thus the space-time discretized optimal solution $(q_{kh},\mathbf{u}_{kh})\in Q\times X_{kh}$ is a stationary point of the stabilized Lagrangian

$$\begin{aligned}\mathcal{L}_{kh}(q_{kh},\mathbf{u}_{kh},\mathbf{z}_{kh}) =&\mathcal{J}(q_{kh},\mathbf{u}_{kh})-a_{kh}(\mathbf{u}_{kh})(\mathbf{z}_{kh})+f_k(q_{kh})(z_{kh}^v)\\ &+(u_0-u_{kh}(0),z_{kh}^u(0))+(v_0-v_{kh}(0),z_{kh}^v(0)).\end{aligned}$$

We refer to Problem 5.6 for the exact definition of the space-time discretized semi-linear form $a_{kh}(\cdot)(\cdot)$. In this section, we assume that we use a quadrature rule of higher-order for the evaluation of the cost functional such that the cost functional is either evaluated exactly or the quadrature error is of such high-order that it can be neglected. Then, following the ideas in [111, 112] and [28] we obtain:

Theorem 7.1. *Let $(q,\mathbf{u},\mathbf{z})$, $(q_k,\mathbf{u}_k,\mathbf{z}_k)$, $(q_{kh},\mathbf{u}_{kh},\mathbf{z}_{kh})$, be stationary points of $\mathcal{L}$, $\mathcal{L}_k$ and $\mathcal{L}_{kh}$ on different levels of discretization and let $(q_\sigma,\mathbf{u}_\sigma,\mathbf{z}_\sigma)$ be a stationary point of $\mathcal{L}_{kh}$ on the fully discrete space $Q_d\times X_{kh}\times Y_{kh,\theta}$ that is*

$$\mathcal{L}'(q,\mathbf{u},\mathbf{z})(\delta q,\delta\mathbf{u},\delta\mathbf{z})=0 \qquad \forall(\delta q,\delta\mathbf{u},\delta\mathbf{z})\in Q\times X\times Y, \tag{7.1}$$

$$\mathcal{L}_k'(q_k,\mathbf{u}_k,\mathbf{z}_k)(\delta q_k,\delta\mathbf{u}_k,\delta\mathbf{z}_k)=0 \qquad \forall(\delta q_k,\delta\mathbf{u}_k,\delta\mathbf{z}_k)\in Q\times X_k\times Y_{k,\theta}, \tag{7.2}$$

$$\mathcal{L}_{kh}'(q_{kh},\mathbf{u}_{kh},\mathbf{z}_{kh})(\delta q_{kh},\delta\mathbf{u}_{kh},\delta\mathbf{z}_{kh})=0 \qquad \forall(\delta q_{kh},\delta\mathbf{u}_{kh},\delta\mathbf{z}_{kh})\in Q\times X_{kh}\times Y_{kh,\theta}, \tag{7.3}$$

$$\mathcal{L}_{kh}'(q_\sigma,\mathbf{u}_\sigma,\mathbf{z}_\sigma)(\delta q_\sigma,\delta\mathbf{u}_\sigma,\delta\mathbf{z}_\sigma)=0 \qquad \forall(\delta q_\sigma,\delta\mathbf{u}_\sigma,\delta\mathbf{z}_\sigma)\in Q_d\times X_{kh}\times Y_{kh,\theta}. \tag{7.4}$$

Then, there holds for the error with respect to the cost functional $\mathcal{J}$ due to discretization in space and time and of the control

$$\begin{aligned}
\mathcal{J}(q,\mathbf{u}) - \mathcal{J}(q_k,\mathbf{u}_k) =& \frac{1}{2}\big[\mathcal{L}'_{\mathbf{u}}(q_k,\mathbf{u}_k,\mathbf{z}_k)(\mathbf{u}-\hat{\mathbf{u}}_k) + \mathcal{L}'_{\mathbf{z}}(q_k,\mathbf{u}_k,\mathbf{z}_k)(\mathbf{z}-\hat{\mathbf{z}}_k)\big] \\
&+ [f(q_k)(z_k^v) - a(\mathbf{u}_k)(\mathbf{z}_k)] + R_{k,1} + R_{k,2}, \\
\mathcal{J}(q_k,\mathbf{u}_k) - \mathcal{J}(q_{kh},\mathbf{u}_{kh}) =& \frac{1}{2}\big[\mathcal{L}'_{k,\mathbf{u}}(q_{kh},\mathbf{u}_{kh},\mathbf{z}_{kh})(\mathbf{u}_k-\hat{\mathbf{u}}_{kh}) + \mathcal{L}'_{k,\mathbf{z}}(q_{kh},\mathbf{u}_{kh},\mathbf{z}_{kh})(\mathbf{z}_k-\hat{\mathbf{z}}_{kh})\big], \\
&+ a_{kh}^{LPS}(\mathbf{u}_{kh})(\mathbf{z}_{kh}) + R_{kh,1} + R_{kh,2} \\
\mathcal{J}(q_{kh},\mathbf{u}_{kh}) - \mathcal{J}(q_\sigma,\mathbf{u}_\sigma) =& \frac{1}{2}\mathcal{L}'_{kh,q}(q_\sigma,\mathbf{u}_\sigma,\mathbf{z}_\sigma)(q_{kh}-\hat{q}_\sigma) + R_\sigma.
\end{aligned}$$

Here $(\hat{\mathbf{u}}_k,\hat{\mathbf{z}}_k) \in X_k \times Y_{k,\theta}$ $(\hat{\mathbf{u}}_{kh},\hat{\mathbf{z}}_{kh}) \in X_{kh} \times Y_{kh,\theta}$ and $\hat{q}_\sigma \in Q_d$ can be chosen arbitrarily.

Proof. For the error due to discretization in time and the application of the theta-quadrature rule (5.2) we obtain

$$\begin{aligned}
\mathcal{J}(q,\mathbf{u}) - \mathcal{J}(q_k,\mathbf{u}_k) &= \mathcal{L}(q,\mathbf{u},\mathbf{z}) - \mathcal{L}_k(q_k,\mathbf{u}_k,\mathbf{z}_k) \\
&= \underbrace{\big[\mathcal{L}(q,\mathbf{u},\mathbf{z}) - \mathcal{L}(q_k,\mathbf{u}_k,\mathbf{z}_k)\big]}_{I_1} + \underbrace{\big[\mathcal{L}(q_k,\mathbf{u}_k,\mathbf{z}_k) - \mathcal{L}_k(q_k,\mathbf{u}_k,\mathbf{z}_k)\big]}_{I_2}
\end{aligned}$$

as $(q,\mathbf{u})$ and $(q_k,\mathbf{u}_k)$ fulfill due to (7.1) and (7.2) the continuous and time-discretized state equation. For the first term I_1 we follow the idea in [19, 20] and receive with $e = (q-q_k, \mathbf{u}-\mathbf{u}_k, \mathbf{z}-\mathbf{z}_k)$ by using the main theorem of calculus

$$I_1 = \int_0^1 \mathcal{L}'((q_k,\mathbf{u}_k,\mathbf{z}_k) + se)(e)\ \mathrm{d}s.$$

If we evaluate this integral with the trapezoidal rule and take (7.1) into account, then we obtain

$$\begin{aligned}
I_1 &= \frac{1}{2}\mathcal{L}'(q_k,\mathbf{u}_k,\mathbf{z}_k)(e) + \frac{1}{2}\mathcal{L}'(q,\mathbf{u},\mathbf{z})(e) + R_{k,1} \\
&= \frac{1}{2}\mathcal{L}'(q_k,\mathbf{u}_k,\mathbf{z}_k)(e) + R_{k,1}
\end{aligned}$$

with the remainder term

$$R_{k,1} = \frac{1}{2}\int_0^1 \mathcal{L}'''((q_k,\mathbf{u}_k,\mathbf{z}_k) + se)(e,e,e)\cdot s\cdot(s-1)\ \mathrm{d}s.$$

We split the the weight $e = (q-q_k, \mathbf{u}-\mathbf{u}_k, \mathbf{z}-\mathbf{z}_k)$ into two parts

$$e = \big(q-\hat{q}_k, \mathbf{u}-\hat{\mathbf{u}}_k, \mathbf{z}-\hat{\mathbf{z}}_k\big) + \big(\hat{q}_k-q_k, \hat{\mathbf{u}}_k-\mathbf{u}_k, \hat{\mathbf{z}}_k-\mathbf{z}_k\big)$$

with arbitrary values $(\hat{q}_k,\hat{\mathbf{z}}_k,\hat{\mathbf{u}}_k) \in Q \times X_k \times Y_{k,\theta}$ and get

$$I_1 = \frac{1}{2}\mathcal{L}'(q_k,\mathbf{u}_k,\mathbf{z}_k)(q-\hat{q}_k, \mathbf{u}-\hat{\mathbf{u}}_k, \mathbf{z}-\hat{\mathbf{z}}_k) + \underbrace{\frac{1}{2}\mathcal{L}'(q_k,\mathbf{u}_k,\mathbf{z}_k)(\hat{q}_k-q_k, \hat{\mathbf{u}}_k-\mathbf{u}_k, \hat{\mathbf{z}}_k-\mathbf{z}_k)}_{R_{k,2}} + R_{k,1}.$$

The last term I_2 estimates the error between the Lagrangian $\mathcal{L}$ and the semi-discrete Lagrangian $\mathcal{L}_k$ and reduces to

$$I_2 = f(q_k)(z_k^v) - a(\mathbf{u}_k, \mathbf{z}_k)$$

as $(q_k, \mathbf{u}_k)$ fulfills (7.2).

We can apply exactly the same approach for the spatial error estimator and get

$$\begin{aligned}\mathcal{J}(q_k, \mathbf{u}_k) - \mathcal{J}(q_{kh}, \mathbf{u}_{kh}) &= \mathcal{L}_k(q_k, \mathbf{u}_k, \mathbf{z}_k) - \mathcal{L}_{kh}(q_{kh}, \mathbf{u}_{kh}, \mathbf{z}_{kh}) = \\ &= \underbrace{\left[\mathcal{L}_k(q_k, \mathbf{u}_k, \mathbf{z}_k) - \mathcal{L}_k(q_{kh}, \mathbf{u}_{kh}, \mathbf{z}_{kh})\right]}_{I_3} + \underbrace{\left[\mathcal{L}_k(q_{kh}, \mathbf{u}_{kh}, \mathbf{z}_{kh}) - \mathcal{L}_{kh}(q_{kh}, \mathbf{u}_{kh}, \mathbf{z}_{kh})\right]}_{I_4}.\end{aligned}$$

As the time discretized Lagrangian $\mathcal{L}_k$ and the space discretized Lagrangian $\mathcal{L}_{kh}$ only differ in the stabilization term we get

$$I_4 = a_{kh}^{\mathrm{LPS}}(\mathbf{u}_{kh})(\mathbf{z}_{kh}).$$

For the first term I_3 we follow the same approach as before and get for $e = (q_k - q_{kh}, \mathbf{u}_k - \mathbf{u}_{kh}, \mathbf{z}_k - \mathbf{z}_{kh})$ using again the theorem of calculus

$$I_3 = \frac{1}{2}\mathcal{L}_k'(q_{kh}, \mathbf{u}_{kh}, \mathbf{z}_{kh})(q_k - \hat{q}_{kh}, \mathbf{u}_k - \hat{\mathbf{u}}_{kh}, \mathbf{z}_k - \hat{\mathbf{z}}_{kh}) + R_{kh,1} + R_{kh,2}.$$

The remainder term $R_{kh,2}$ occurs by enforcing the weight to be an interpolation error and the remainder term $R_{kh,1}$ has again the structure

$$R_{kh,1} = \frac{1}{2}\int_0^1 \mathcal{L}_k'''((q_{kh}, \mathbf{u}_{kh}, \mathbf{z}_{kh}) + se)(e, e, e) \cdot s \cdot (s-1)\ \mathrm{d}s.$$

Finally, we have to take a closer look at the control discretization error. We apply again the main theorem of calculus and evaluate the integral with the trapezoidal rule. Then we obtain

$$\begin{aligned}\mathcal{J}(q_{kh}, \mathbf{u}_{kh}) - \mathcal{J}(q_\sigma, \mathbf{u}_\sigma) &= \mathcal{L}_{kh}(q_{kh}, \mathbf{u}_{kh}, \mathbf{z}_{kh}) - \mathcal{L}_{kh}(q_\sigma, \mathbf{u}_\sigma, \mathbf{z}_\sigma) = \\ &= \frac{1}{2}\mathcal{L}_{kh}'(q_\sigma, \mathbf{u}_\sigma, \mathbf{z}_\sigma)(q_{kh} - \hat{q}_\sigma, \mathbf{u}_{kh} - \hat{\mathbf{u}}_\sigma, \mathbf{z}_{kh} - \hat{\mathbf{z}}_\sigma) + R_{\sigma,1}\end{aligned}$$

With $e = (q_{kh} - q_\sigma, \mathbf{u}_{kh} - \mathbf{u}_\sigma, \mathbf{z}_{kh} - \mathbf{z}_\sigma)$ the remainder $R_{\sigma,1}$ has again the structure

$$R_{\sigma,1} = \frac{1}{2}\int_0^1 \mathcal{L}_k'''((q_\sigma, \mathbf{u}_\sigma, \mathbf{z}_\sigma) + se)(e, e, e) \cdot s \cdot (s-1)\ \mathrm{d}s.$$

As we have Galerkin-orthogonality here, we do not receive an additional remainder term as in the previous cases. The values $\hat{q}_k \in Q$, $\hat{q}_{kh} \in Q$, $\hat{\mathbf{u}}_\sigma \in X_{kh}$, and $\hat{\mathbf{z}}_\sigma \in Y_{kh,\theta}$ can be chosen arbitrarily. We can choose, for example, $\hat{q}_k = q \in Q$ and $\hat{q}_{kh} = q_k \in Q$ such that the derivatives with respect to the control $\mathcal{L}_q'(\cdot)(\cdot)$ and $\mathcal{L}_{k,q}'(\cdot)(\cdot)$ vanish in the error estimators of time and space discretization. If we choose in addition $\hat{\mathbf{u}}_\sigma = \mathbf{u}_{kh} \in X_{kh}$ and $\hat{\mathbf{z}}_\sigma = \mathbf{z}_{kh} \in Y_{kh,\theta}$ the derivatives of the Lagrange functional $\mathcal{L}_{kh}$ with respect to primal and state variable will vanish in the error estimator of the control. □

Due to the discrepancy between the time discretized semi-linear form $a_k(\cdot)(\cdot)$ in Problem 5.2 and the semi-linear form $a(\cdot)(\cdot)$ in the Petrov-Galerkin formulation in Problem 5.1, the time discretization error consists of both the Galerkin discretization error and the numerical quadrature error. The quadrature error cannot be neglected as the applied quadrature rule is only of second order accuracy and therefore of the same order as the approximation error. The spatial error estimator consists of the Galerkin discretization error and an estimator measuring the mismatch between discrete and continuous semi-linear form due to the additional LPS stabilization.

Remark 7.1. If we choose for the arbitrary values $\hat{\mathbf{u}}_k, \hat{\mathbf{z}}_k, \hat{\mathbf{u}}_{kh}, \hat{\mathbf{z}}_{kh}$ and $\hat{q}_\sigma$ the interpolation of $\mathbf{u}, \mathbf{z}, \mathbf{u}_k, \mathbf{z}_k$ and q_{kh} in the semi-discrete and discrete spaces, the weights will mainly represent an interpolation error. Therefore, for evaluating the residuals later, we have to find a good approximation of the interpolation error.

Bear in mind that it is possible to extend the a posteriori error estimator to estimate the discretization error with respect to an arbitrary functional of interest not coinciding with the cost functional. For more information on this topic, we would like to refer to [113].

The presented DWR error estimator cannot only be used for optimal control problems, but also to estimate the error in a functional of interest for simulations of arbitrary PDEs. In order to compute characteristic values of our fluid-structure interaction simulation as accurate as possible we can define the Lagrangian

$$\begin{aligned}\mathcal{L}(\mathbf{u},\mathbf{z}) =& \mathcal{J}(\mathbf{u}) - a(\mathbf{u})(\mathbf{z}) + f(\mathbf{z}) \\ &+ (u_0 - u(0), z^u(0)) + (v_0 - v(0), z^v(0)).\end{aligned}$$

Thereby, $\mathcal{J}(\mathbf{u})$ can be an arbitrary functional of interest and $a(\cdot)(\cdot)$ is the semi-linear form given in (4.10). As state and adjoint solutions are again stationary points of the Lagrangian we can directly adapt the approach of Theorem 7.1. The derivative of the Lagrange functional vanishes with respect to the control variable and we get here

Theorem 7.2. *Let $(\mathbf{u},\mathbf{z})$, $(\mathbf{u}_k,\mathbf{z}_k)$, $(\mathbf{u}_{kh},\mathbf{z}_{kh})$ be stationary points of $\mathcal{L}$, $\mathcal{L}_k$ and $\mathcal{L}_{kh}$ on different levels of discretization, that is*

$$\begin{aligned}\mathcal{L}'(\mathbf{u},\mathbf{z})(\delta\mathbf{u},\delta\mathbf{z}) &= 0 && \forall(\delta\mathbf{u},\delta\mathbf{z}) \in X \times Y, \\ \mathcal{L}_k'(\mathbf{u}_k,\mathbf{z}_k)(\delta\mathbf{u}_k,\delta\mathbf{z}_k) &= 0 && \forall(\delta\mathbf{u}_k,\delta\mathbf{z}_k) \in X_k \times Y_{k,\theta}, \\ \mathcal{L}_{kh}'(\mathbf{u}_{kh},\mathbf{z}_{kh})(\delta\mathbf{u}_{kh},\delta\mathbf{z}_{kh}) &= 0 && \forall(\delta\mathbf{u}_{kh},\delta\mathbf{z}_{kh}) \in X_{kh} \times Y_{kh,\theta}.\end{aligned}$$

Then, there holds for the error with respect to the functional of interest $\mathcal{J}$ due to discretization in space and time

$$\begin{aligned}\mathcal{J}(\mathbf{u}) - \mathcal{J}(\mathbf{u}_k) =& \frac{1}{2}\left[\mathcal{L}_{\mathbf{u}}'(\mathbf{u}_{kh},\mathbf{z}_{kh})(\mathbf{u}-\hat{\mathbf{u}}_k) + \mathcal{L}_{\mathbf{z}}'(\mathbf{u}_{kh},\mathbf{z}_{kh})(\mathbf{z}-\hat{\mathbf{z}}_k)\right] \\ &+ [f(\mathbf{z}_k) - a(\mathbf{u}_k)(\mathbf{z}_k)] + R_{k,1} + R_{k,2}, \\ \mathcal{J}(\mathbf{u}_k) - \mathcal{J}(\mathbf{u}_{kh}) =& \frac{1}{2}\left[\mathcal{L}_{k,\mathbf{u}}'(\mathbf{u}_{kh},\mathbf{z}_{kh})(\mathbf{u}_k-\hat{\mathbf{u}}_{kh}) + \mathcal{L}_{k,\mathbf{z}}'(\mathbf{u}_{kh},\mathbf{z}_{kh})(\mathbf{z}_k-\hat{\mathbf{z}}_{kh})\right], \\ &+ a_{kh}^{LPS}(\mathbf{u}_{kh})(\mathbf{z}_{kh}) + R_{kh,1} + R_{kh,2}.\end{aligned}$$

Here $(\hat{\mathbf{u}}_k,\hat{\mathbf{z}}_k) \in X_k \times Y_{k,\theta}$ $(\hat{\mathbf{u}}_{kh},\hat{\mathbf{z}}_{kh}) \in X_{kh} \times Y_{kh,\theta}$ can be chosen arbitrarily and the remainder terms $R_{k,1}, R_{k,2}, R_{kh,1}$ and $R_{kh,2}$ have the same structure as in Theorem 7.1.

To calculate the discretization error in terms of the functional of interest $\mathcal{J}(\mathbf{u})$ we have to solve in addition an adjoint equation to obtain sensitivity information. As the adjoint equation is linear, this additional cost does not carry weight for highly nonlinear problems.

Remark 7.2. For the fractional-step theta time-stepping scheme, second order convergence only occurs in the macro time-steps. The intermediate steps of the fractional-step theta method are not necessarily precise approximations to the solution. Hence, these steps should be omitted in the functional evaluation. Therefore we introduce the piecewise linear interpolation $i_{3k}^{(1)} : X_k \to X_k^{\text{macro}}$, such that $i_{3k}^{(1)}\mathbf{u}_k$ and $\mathbf{u}_k$ coincide on the macro time points t_{3n} and X_k^{macro} is defined in on the macro time grid $t_0 < t_3 < ... < t_{3n} < ... < t_M$ just like X_k. By splitting the error in the cost functional in

$$\mathcal{J}(\mathbf{u}) - \mathcal{J}(i_{3k}\mathbf{u}_k) = [\mathcal{J}(\mathbf{u}) - \mathcal{J}(\mathbf{u}_k)] + [\mathcal{J}(\mathbf{u}_k) - \mathcal{J}(i_{3k}\mathbf{u}_k)]$$

we then can apply again the derived a posteriori error estimator for the first term, and the last term can be evaluated exactly.

7.2. A Posteriori Error Estimation for Optimal Control of the Linear FSI Problem

In this Section, we are going to take a closer look at the time-discretization error estimator in Theorem 7.1 for optimal control of a linear fluid-structure interaction problem. We minimize the tracking type functional (3.27) constrained by the symmetric linear fluid-structure interaction model. The time dependent control $q \in L^2(I)$ enters the equation as volume force as in Configuration 2. In Section 3.3.3, we have proven the existence of a unique solution for this configuration. The focus in this section will be on proving that the remainder terms in Theorem 7.1 are of higher-order and therefore can be neglected in the evaluation of the error estimators. As defined in Section 3.3.2, we have the Lagrange functional $\mathcal{L} : Q \times X \times Y \times \to \mathbb{R}$

$$\begin{aligned}\mathcal{L}(q, \mathbf{u}, \mathbf{z}) :=& \mathcal{J}(q, \mathbf{u}) - a(\mathbf{u})(\mathbf{z}) + (\!(B_{\mathrm{f}} q, z^v)\!)_{\mathrm{f}} + (\!(B_{\mathrm{s}} q, z^v)\!)_{\mathrm{s}} \\ &+ (u_0 - u_k(0), z_k^u(0)) + (v_0 - v_k(0), z_k^v(0))\end{aligned}$$

and $a(\cdot)(\cdot)$ the bi-linear form given in (3.14) modeling the linear FSI configuration. We denote with $B_{\mathrm{f}} : Q \to L^2(I; H_{\mathrm{f}})$ and $B_{\mathrm{s}} : Q \to L^2(I; H_{\mathrm{s}})$ the linear operators mapping the control in the right-hand side, defined in Configuration 2 in Section 3.3.1.

According to Theorem 7.1, we can estimate the error due to time-discretization by

$$\begin{aligned}\mathcal{J}(q, \mathbf{u}) - \mathcal{J}(q_k, \mathbf{u}_k) =& \frac{1}{2}\big[\mathcal{L}'_{\mathbf{u}}(q_k, \mathbf{u}_k, \mathbf{z}_k)(\mathbf{u} - \hat{\mathbf{u}}_k) + \mathcal{L}'_{\mathbf{z}}(q_k, \mathbf{u}_k, \mathbf{z}_k)(\mathbf{z} - \hat{\mathbf{z}}_k)\big] \\ &+ [(\!(B_{\mathrm{f}} q_k, z_k^v)\!)_{\mathrm{f}} + (\!(B_{\mathrm{s}} q_k, z_k^v)\!)_{\mathrm{s}} - a(q_k, \mathbf{u}_k)(\mathbf{z}_k)] + R_{k,1} + R_{k,2}.\end{aligned}$$

To be able to evaluate the estimator we would like to neglect for arbitrary values $\hat{q}_k, \hat{\mathbf{u}}_k, \hat{\mathbf{z}}_k \in X_k$ the remainder terms

$$R_{k,1} = \frac{1}{2}\int_0^1 \mathcal{L}'''((q_k, \mathbf{u}_k, \mathbf{z}_k) + se)(e, e, e) \cdot s \cdot (s-1)\ \mathrm{d}s$$

with $e = (q - q_k, \mathbf{u} - \mathbf{u}_k, \mathbf{z} - \mathbf{z}_k)$ and the remainder

$$R_{k,2} = \frac{1}{2}\mathcal{L}'(q_k, \mathbf{u}_k, \mathbf{z}_k)(\hat{q}_k - q_k, \hat{\mathbf{u}}_k - \mathbf{u}_k, \hat{\mathbf{z}}_k - \mathbf{z}_k).$$

At least for the linear-fluid-structure interaction problem, we can prove in the following Lemma that the remainder terms $R_{k,1}$ and $R_{k,2}$ will even vanish.

Lemma 7.3. *For optimal control of the linear FSI problem with tracking-type functional and control $q \in L^2(I)$ as in Configuration 2 (see Section 3.3.1) the remainder terms $R_{k,1}$ and $R_{k,2}$ in Theorem 7.1 vanish.*

Proof. For the quadratic tracking type functional $\mathcal{J}(q, \mathbf{u})$ in (3.27) the remainder $R_{k,1}$ vanishes as the third derivative of the Lagrange functional is zero in all directions due to the linearity of the fluid-structure interaction problem. Therefore, we only have to take a closer look at the remainder $R_{k,2}$:

$$\begin{aligned} R_{k,2} &= \frac{1}{2}\mathcal{L}'(q_k, \mathbf{u}_k, \mathbf{z}_k)(\hat{q}_k - q_k, \hat{\mathbf{u}}_k - \mathbf{u}_k, \hat{\mathbf{z}}_k - \mathbf{z}_k) \\ &= \frac{1}{2}\mathcal{L}'_{\mathbf{z}}(q_k, \mathbf{u}_k, \mathbf{z}_k)(\hat{\mathbf{z}}_k - \mathbf{z}_k) + \frac{1}{2}\mathcal{L}'_{\mathbf{u}}(q_k, \mathbf{u}_k, \mathbf{z}_k)(\hat{\mathbf{u}}_k - \mathbf{u}_k) + \frac{1}{2}\mathcal{L}'_q(q_k, \mathbf{u}_k, \mathbf{z}_k)(\hat{q}_k - q_k) \\ &=: \frac{1}{2}R_{2,1} + \frac{1}{2}R_{2,2} + \frac{1}{2}R_{2,3}. \end{aligned}$$

As $\mathbf{z}_k \in Y_{k,\theta}$, we can rewrite the adjoint variables z_k^v and z_k^u on every time interval as $z_k^v|_{I_m} = z_{k,m}^v \omega_{\theta,m}$ and $z_k^u|_{I_m} = z_{k,m}^u \omega_{\theta,m}$, like we did in Section 5.1. As the semi-linear form $a(\cdot)(\cdot)$ is bi-linear and $\mathbf{u}_k$ is linear on every time interval I_m, all time integrals in the bi-linear form, with the exception of the divergence condition, have the structure

$$\sum_{m=1}^{M} \int_{I_m} h_m \omega_{\theta,m}$$

with h_m linear in time. According to Lemma 5.1, the theta-quadrature rule (5.2) evaluates the time integral exactly. For the divergence condition we applied the trapezoidal rule which, in this case, is exact, too. Therefore, we obtain

$$a(\mathbf{u}_k)(\hat{\mathbf{z}}_k - \mathbf{z}_k) = a_k(\mathbf{u}_k)(\hat{\mathbf{z}}_k - \mathbf{z}_k).$$

According to Theorem 3.13 the optimal solution q fulfills

$$q = -\frac{1}{\alpha}\int_\Omega h(x) z_k^v(\cdot, x)\, \mathrm{d}x.$$

We remind the reader that the control is mapped in the right-hand side through the linear operators B_f and B_s. Thereby we multiply the time dependent control with the function $h \in V_\mathrm{div}$ as defined in Configuration 2. Hence, the optimal control variable q_k inherits the time discretization of the adjoint velocity variable z_k, which is an element of the time-discrete theta dependent space $z_k \in L_{k,\theta}^v$. Thus we can assume without loss of generality

$$q_k \in Q_{k,\theta} := \{\varphi_k \in L^2(I) | \varphi_k|_{I_m} \in P_0^\theta(I_m), m = 1, 2, ..., M \text{ and } \varphi_k(0) \in \mathbb{R}\}.$$

Therefore the quadrature rule (5.2) evaluates the control dependent terms exactly and we get

$$\begin{aligned} R_{2,1} &= \frac{1}{2}[a(\mathbf{u}_k)(\hat{\mathbf{z}}_k - \mathbf{z}_k) + (\!(B_\mathrm{f} q_k, \hat{\mathbf{z}}_k - \mathbf{z}_k)\!)_\mathrm{f} + (\!(B_\mathrm{s} q_k, \hat{\mathbf{z}}_k - \mathbf{z}_k)\!)_\mathrm{s}] \\ &= \frac{1}{2}\mathcal{L}'_{k,\mathbf{z}}(q_k, \mathbf{u}_k, \mathbf{z}_k)(\hat{\mathbf{z}}_k - \mathbf{z}_k) = 0. \end{aligned}$$

As state and adjoint equation have the same structure, we can apply the same argumentation for the adjoint. Here, we only have to take a closer look at the discretization of the right-hand side. For the given quadratic tracking type functional we would choose a two-point Gauß quadrature rule to evaluate the cost functional exactly. Then we get

$$\begin{aligned} R_{2,2} &= a(\mathbf{u})(\hat{\mathbf{u}}_k - \mathbf{u}_k) - \mathcal{J}'_{\mathbf{u}}(\hat{\mathbf{u}}_k - \mathbf{u}_k) = a_k(\mathbf{u})(\hat{\mathbf{u}}_k - \mathbf{u}_k) - \mathcal{J}'_{\mathbf{u}}(\hat{\mathbf{u}}_k - \mathbf{u}_k) \\ &= \frac{1}{2}\mathcal{L}'_{k,\mathbf{u}}(q_k, \mathbf{u}_k, \mathbf{z}_k)(\hat{\mathbf{u}}_k - \mathbf{u}_k) = 0. \end{aligned}$$

Finally we consider the derivative with respect to the control

$$R_{2,3} = (\!(B_\mathrm{f}(\hat{q}_k - q_k), \mathbf{z}_k)\!)_\mathrm{f} + (\!(B_\mathrm{s}(\hat{q}_k - q_k), \mathbf{z}_k)\!)_\mathrm{s} - \alpha(\!(q_k, \hat{q}_k - q_k)\!).$$

As already mentioned, without restricting the control space we have $q_k \in Q_{k,\theta}$, therefore we follow the same argumentation as above. As the operators B_f and B_s are linear and not time-dependent, the values $B_\mathrm{f}(\hat{q}_k - q_k)$ and $B_\mathrm{s}(\hat{q}_k - q_k)$ are piecewise linear on every time interval. Therefore, we can exactly evaluate with the quadrature rule (5.2) all integrals in time and we get

$$R_{2,3} = \frac{1}{2}\mathcal{L}'_{k,q}(q_k, \mathbf{u}_k, \mathbf{z}_k)(\hat{q}_k - q_k) = 0.$$

As $(q_k, \mathbf{u}_k, \mathbf{z}_k)$ are stationary points and fulfill (7.2), the remainder term $R_{k,2}$ vanishes.

□

In [111] the authors apply the DWR error estimator to nonlinear parabolic equations and in [112] to the Navier-Stokes equations. Thereby, they prove for both examples that the remainder terms $R_{k,1}$ and $R_{k,2}$ are of third order and can be neglected. These results immediately enable extending of Lemma 7.3 to optimal control problems constrained by nonlinear parabolic equations or the Navier-Stokes equations. Even in these cases the remainder terms $R_{k,1}$ and $R_{k,2}$ are of higher-order. In [128], the author claims that for dual-weighted residual error estimators in space for stationary fluid-structure interaction the remainder terms can be assumed to be of third order, as long as the ALE transformation is smooth enough. Hence when Theorem 7.1 is applied later to optimal control of the nonlinear fluid-structure problem, it is reasonable to assume that the remainder terms are small and can be neglected.

7.3. Evaluation of the Error Estimators

7.3.1. Higher-Order Reconstruction

One possibility to deal with the arbitrary values $\hat{\mathbf{u}}_k, \hat{\mathbf{z}}_k, \hat{\mathbf{u}}_{kh}, \hat{\mathbf{z}}_{kh}$ and $\hat{q}_\sigma$ is to choose the interpolation of the exact solution in the discretized and semi-discretized spaces. Then the so called weights, as for example the term $\mathbf{u} - \hat{\mathbf{u}}_k$, are mainly interpolation errors. Nevertheless, the weights in the residual error estimator in Theorem 7.1 still contain either the exact solution $(\mathbf{u}, \mathbf{z}, p)$ or the semi discretized solutions. However, only the fully discrete solution $(\mathbf{u}_\sigma, \mathbf{z}_\sigma, q_\sigma)$ can be computed. Motivated by super convergence results, as in [134, 19], we use as approximation of $(\mathbf{u}, \mathbf{z}, q)$, a higher-order reconstruction. This approach relies on "super closeness" of the higher-order interpolants to the exact solution, see [19]. Such results can be found for the finite element discretization of the Poisson equation and the Stokes equations in [29]. Such super convergence results, have not been proven for the considered fluid-structure interaction problem, yet. An alternative would be to solve the fluid-structure interaction problem on a finer grid. This would result in a very high computational effort. More information about alternative methods to higher-order interpolation can be found in [19]. We are going to apply a higher-order reconstruction in the following, which performs well for the numerical examples in Chapter 8.

We construct linear operators which enable to approximate the interpolation errors in the time, spatial and control discretized spaces. The applied interpolation operator thereby only builds on the computed solution. We first derive an approximation of the weights for estimating the time-discretization error. Thereby we follow [28] and [111]. Then we extend the method to the weights in the spatial and control discretization error estimator.

Higher-Order Reconstruction in Time

The discretized velocity variable $v_\sigma \in W_{kh}^v$ and the displacement $u_\sigma \in W_{kh}^v$ are piecewise linear in time on every time-interval. We define the higher order approximation $i_{2k}^{(2)}(u_\sigma)$ and $i_{2k}^{(2)}(v_\sigma)$, which map in the space of piecewise quadratic polynomials defined on unions of two adjacent subintervals. The action of the operator $i_{2k}^{(2)}$ is depicted in Figure 7.1.

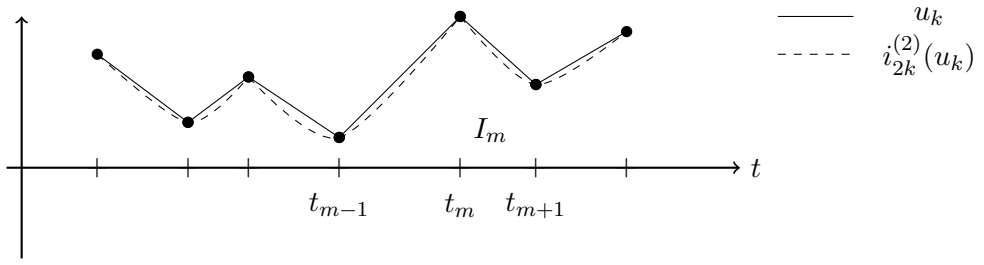

Figure 7.1.: Continous piecewise quadratic interpolation $i_{2k}^{(2)}$ of a continuous piecewise linear function u_k.

As the intermediate steps of the fractional-step theta time-stepping scheme are not necessarily precise approximations, we should use the piecewise quadratic reconstruction of the solution

on the macro steps suggested in [111, 112]. Therefore we use for the fraction step theta scheme the interpolation $i_{2k}^{(2)}$ depicted in Figure 7.2, whereby we keep the notation for the interpolation operator. In the following chapters if we use the fractional-step theta time-stepping scheme the reconstruction will always be defined on the macro time grid.

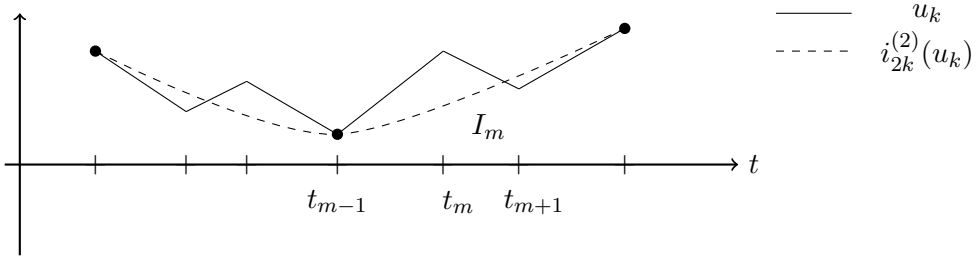

Figure 7.2.: Continous piecewise quadratic interpolation $i_{2k}^{(2)}$ of a continuous piecewise linear function u_k on macro time steps.

The semi-discretized pressure $p_\sigma \in L^2(I; L_\mathrm{f})_{kh}$ is piecewise constant, on every time interval. Here several possibilities are reasonable for reconstructing an approximation in the piecewise linear finite element space. We decided to interpolate the pressure values at the nodes in the time grid with a linear function. The action of the operator $i_k^{(1)}$ is depicted in Figure 7.3. For the fractional-step theta method we use the reconstruction of the pressure on the macro time-grid.

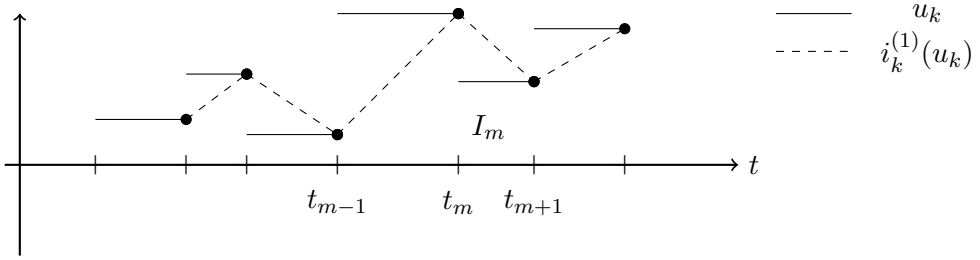

Figure 7.3.: Continous piecewise linear interpolation $i_k^{(1)}$ of a discontinuous piecewise constant function u_k.

It is unclear how the reconstruction for the adjoint variable has to look like. We do not have any super-convergence results available. However, numerical experiments have shown that as long as the interpolation operator is well defined, a lot of different approaches work well. As the slope of the test function is needed to artificially introduce the theta value, we should use for the interpolation the values at the middle of each time-interval. Then, we get the operator $i_{2k,\theta}^{(2)}$ as in Figure 7.4. Such a piecewise quadratic reconstruction has kinks in the middle of the time intervals. By shifting the interpolation to the time points of the time-grid the implementation is much easier. As numerical experiments have not shown any disadvantages, we are going to use the shifted version in our numerical calculations later.

For the adjoint pressure, we use again the interpolation operator $i_k^{(1)}$. Therefore we approximate the weights $\mathbf{u} - \hat{\mathbf{u}}_k$ and $\mathbf{z} - \hat{\mathbf{z}}_k$ using the interpolation of the discrete solution $\mathbf{u}_\sigma$ and $\mathbf{z}_\sigma$

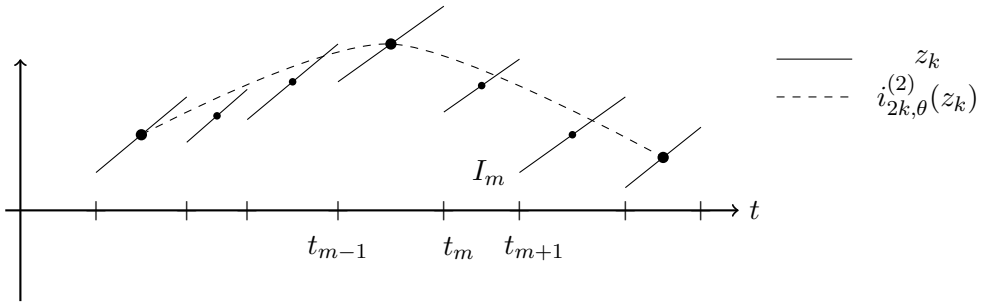

Figure 7.4.: Continous piecewise quadratic interpolation $i^{(2)}_{2k,\theta}$ of a discontinuous piecewise linear function z_k.

in a higher finite element space and we get

$$\begin{aligned}
\mathbf{u} - \hat{\mathbf{u}}_k &\approx (i^{(2)}_{2k}(v_\sigma) - v_\sigma, i^{(2)}_{2k}(u_\sigma) - u_\sigma, i^{(1)}_{2k}(p_\sigma) - p_\sigma) \\
&=: i_k(\mathbf{u}_\sigma) - \mathbf{u}_\sigma, \\
\mathbf{z} - \hat{\mathbf{z}}_k &\approx (i^{(2)}_{3k,\theta}(z^v_\sigma) - z^v_\sigma, i^{(2)}_{3k,\theta}(z^u_\sigma) - z^u_\sigma, i^{(1)}_{3k}(z^p_\sigma) - z^p_\sigma) \\
&=: i_{k,\theta}(\mathbf{z}_\sigma) - \mathbf{z}_\sigma.
\end{aligned}$$

In the case of the fractional-step theta time stepping scheme, we denote with i_k and $i_{k,\theta}$ the reconstruction on the macro time grid.

Higher-Order Reconstruction in Space

The same techniques can be used to define a spacial interpolation operator $i^{(2s)}_{2h}$ interpolating in the space of finite elements with order $2s$ on macrocells. We combine in 2D four adjacent cells to a macro cell. The interpolation operator thereby can be implemented very easily if we have a patch structure.

$$\begin{aligned}
\mathbf{u}_k - \hat{\mathbf{u}}_{kh} &\approx (i^{(2s)}_{2h}(v_\sigma) - v_\sigma, i^{(2s)}_{2h}(u_\sigma) - u_\sigma, i^{(2s)}_{2h}(p_\sigma) - p_\sigma) \\
&=: i_h(\mathbf{u}_\sigma) - \mathbf{u}_\sigma, \\
\mathbf{z}_k - \hat{\mathbf{z}}_{kh} &\approx (i^{(2s)}_{2h}(z^v_\sigma) - z^v_\sigma, i^{(2s)}_{2h}(z^u_\sigma) - z^u_\sigma, i^{(2s)}_{2h}(z^v_\sigma) - z^v_\sigma) \\
&=: i_h(\mathbf{z}_\sigma) - \mathbf{z}_\sigma.
\end{aligned}$$

Higher-Order Reconstruction of the Control Variable

For the control variable we can use the same approaches and we get

$$q_{kh} - q_\sigma \approx i_d(q_\sigma) - q_\sigma.$$

If the control variable is chosen piecewise constant in time, we have $i_d(q_\sigma) = i^{(1)}_k(q_\sigma)$ and in the case of a piecewise linear control space we can use $i_d(q_\sigma) = i^{(2)}_{2k}(q_\sigma)$. For a control distributed in space, we have $i_d(q_\sigma) = i^{(2)}_{2h}(q_\sigma)$.

Approximation of the Weights

Thus, by replacing all unknowns in the residuals and weights by fully discrete solutions or its interpolation in a higher finite element space we get using fractional-step theta time-stepping

$$\begin{aligned}
\eta_k \approx & \frac{1}{2}\big[\mathcal{L}'_{\mathbf{u}}(q_\sigma, \mathbf{u}_\sigma, \mathbf{z}_\sigma)(i_k(\mathbf{u}_\sigma) - \mathbf{u}_\sigma) + \mathcal{L}'_{\mathbf{z}}(q_\sigma, \mathbf{u}_\sigma, \mathbf{z}_\sigma)(i_{k,\theta}(\mathbf{z}_\sigma) - \mathbf{z}_k)\big] \\
& + [f(q_k)(z_k^v) - a(\mathbf{u}_k)(\mathbf{z}_k)], \\
\eta_h \approx & \frac{1}{2}\big[\mathcal{L}'_{k,\mathbf{u}}(q_\sigma, \mathbf{u}_\sigma, \mathbf{z}_\sigma)(i_h(\mathbf{u}_\sigma) - \mathbf{u}_\sigma) + \mathcal{L}'_{k,\mathbf{z}}(q_\sigma, \mathbf{u}_\sigma, \mathbf{z}_\sigma)(i_h(\mathbf{z}_\sigma) - \mathbf{z}_\sigma)\big] \\
& + a_{kh}^{\mathrm{LPS}}(\mathbf{u}_{kh})(\mathbf{z}_{kh}), \\
\eta_d \approx & \frac{1}{2}\mathcal{L}'_{kh,q}(q_\sigma, \mathbf{u}_\sigma, \mathbf{z}_\sigma)(i_d(q_\sigma) - q_\sigma).
\end{aligned}$$

It would be possible to use the higher reconstruction also for the terms in the residual, although several examples in literature show that this does not seem to be necessary. We refer for instance to the sources cited at the beginning of this chapter.

7.3.2. Localization of the Error Estimator

A posteriori error estimators can be used to quantify the discretization error. This enables to define a stopping criteria such that the solution is calculated up to a certain accuracy. In addition, by evaluating the error estimator of space, time, and control discretization, we can compare the estimators. Then we can refine either the control, space, or time grid respectively to their contribution to the overall error. After having localized the a posteriori error estimators we can refine in addition locally the space, time, and control grid.

For the time-discretization error, we can calculate the absolute value of the error estimator on every time interval I_m. When splitting the space discretization error estimator in cell-wise contributions, large overestimation of the actual error due to oscillatory behavior of the residuals can be observed (we would like to refer to [42]). One commonly used way to overcome this effect is to cell-wise integrate by parts. Then, the local error indicators consist of the strong residual of the equations, as well as of jump terms over the faces of cells. However, in the context of fluid-structure interaction this is too costly. Instead, we follow Braak and Ern [33], who suggested an alternative filtering method without having to integrate by parts.

In addition to the interpolation operator $i_{2h}^{(2s)}$ in the space of finite elements with order $2s$ on a coarse grid, we define the projection $i_{2h}^{(s)}$ mapping in the space of finite elements with order s, this time on the coarse grid. Furthermore, we define the projection $i_h^{(s)}$ in the space of finite elements with order s on the fine grid. We can observe that for arbitrary functions $\mathbf{u} \in X$:

$$i_{2h}^{(s)}(\mathbf{u} - i_h^{(s)}(\mathbf{u})) = 0.$$

Therefore, we can add this additional term and get as an approximation for the weight with $\hat{\mathbf{u}} = i_h(\mathbf{u})$

$$\mathbf{u} - \hat{\mathbf{u}} = \mathbf{u} - i_h^{(s)}(\mathbf{u}) + i_{2h}^{(s)}(\mathbf{u} - i_h^{(s)}(\mathbf{u})) = \mathbf{u} - i_{2h}^{(s)}(\mathbf{u}) - i_h^{(s)}(\mathbf{u} - i_{2h}^{(s)}(\mathbf{u})).$$

If we now replace the continuous solution $\mathbf{u}$ with its approximation $i_{2h}^{(2)}(\mathbf{u}_\sigma)$, we get the weight

$$\mathbf{u}_k - \hat{\mathbf{u}}_{kh} \approx i_{2h}^{(2s)}(\mathbf{u}_\sigma) - i_{2h}^{(s)}(\mathbf{u}_\sigma) - i_h^{(s)}(i_{2h}^{2s}(\mathbf{u}_\sigma) - i_{2h}^{(s)}(\mathbf{u}_\sigma)). \tag{7.5}$$

Evaluating now the error estimator with the filtered weight (7.5) cell-wise it can be shown that the local error indicator is effective for several examples and does not overestimate the error anymore.

Hence, we need the patch structure to first construct an aproximation of the exact solution in a higher finite-element space and to define then a filtering operator to localize the error estimator. Alternative localization methods have been discussed in [132] without needing a patch structure.

7.4. Adaptive Algorithm

We present in this section an adaptive refinement algorithm for optimal control, which uses the derived localized a posteriori error estimators:

$$\Sigma_k := \{\eta_{k,m} | m = 1, ..., M\}, \quad \Sigma_h := \{\eta_{h,K} | K \in \mathcal{T}_h\} \quad \text{and} \quad \Sigma_q := \{\eta_{q,i} | i = 1, ..., n\}.$$

To get a consistent optimization algorithm, we solve in every adaption step first the optimization problem up to a certain accuracy. Only if the optimization algorithm has converged, the discrete solution will fulfill the necessary optimality condition plus will be a stationary point of the discrete Lagrangian. Numerical experiments show, that if the calculated solution is not close to the optimal solution the DWR estimator is not accurate. After refinement in space and time, the optimization problem is solved once again. Quasi-Newton and Newton algorithm are known to have fast convergence close to the optimal solution. Therefore, we expect that we only need a few steps on the refined grid if we start the LBFGS algorithm with the optimal control calculated on the coarse grid. We have summarized the different steps needed for an adaptive algorithm for optimal control in Algorithm 7.1.

There are different possibilities to choose the cells and time intervals to be refined. The easiest method is to calculate an averaged error per cell, whereby all cells with larger error indicators have to be refined. We will follow a slightly different strategy which is presented for example in [127]. If we have a set of localized error estimators $\Sigma = \{\eta_1, ..., \eta_N\}$, we compute first a permutation $(i_1, ..., i_N)$ of the indices $1, ..., N$ such that

$$|\eta_{i_1}| \geq |\eta_{i_2}| \geq ... \geq |\eta_{i_N}|.$$

The subset of indicators $\Sigma_R \subset \Sigma$ to be refined is chosen as coherent queue $\Sigma_R = \{\eta_{i_1}, ..., \eta_{i_r}\}$ with r being the solution of the minimization problem

$$r := \operatorname*{argmin}_{1 \leq r \leq N} \mathcal{E}(r) N(r)^\beta.$$

Thereby,

$$\mathcal{E}(r) = \sum_{i=1}^{N} |\eta_i| - \sum_{i=1}^{r} (1 - 2^{-\alpha}) |\eta_i|$$

Algorithm 7.1: Abstract adaptive algorithm for optimal control

Choose initial temporal and spatial discretization
while $\eta > \text{tol}_1$ **do**
  while $\|\nabla j(q_\sigma)\| > \text{tol}_2$ **do**
    for $m = 1, ..., M$ **do**
      For given control q_σ find $\mathbf{u}_{\sigma,m}$ solving the discretized state equation
    end for
    for $m = M, ..., 0$ **do**
      Find $\mathbf{z}_{\sigma,m}$ solving the discretized adjoint equation
    end for
    Calculate LBFGS update Δq_σ and define $q_\sigma = \Delta q_\sigma + q_\sigma$
  end while
  Evaluate the localized a posteriori estimators $\eta_{k,m}$, $\eta_{h,K}$, $\eta_{d,i}$
  Determine time intervals/cells with large error contribution and set $\eta = \eta_k + \eta_h + \eta_d$
  Refine the selected time intervals/cells
end while

is a prediction of the discretization error after refinement of the cell. The parameter α denotes the expected order of convergence and $N(r)$ corresponds to the number of degrees of freedom of the refined discretization. The quotient of the degree of polynomials used for discretization and the dimension of the domain is given by β . In the case of a fractional-step theta time-stepping scheme we have $\beta = 1$ and $\alpha = 2$. The optimal value r is determined by successively testing. Then, we refine all cells with indicators in the set Σ_R.

Adaptive algorithms are very efficient to resolve singularities occurring for example around corners. As we are going to see in Chapter 8, the solution in time is often very smooth and the time discretization error is uniformly distributed. In that case, there is no advantage in adaptive refinement of the time-grid compared to global refinement. However, the DWR estimator enables in addition to compare the spatial, time, and control discretization error. Thereby, we only refine in space and time, as long as the errors are similar in size. As we are going to see in Section 8.2, the possibility to equilibrate the errors is one of the main advantages of the DWR error estimates and can reduce the computational cost enormously.

In [89], a good overview on heuristic error estimators for different time-discretization schemes for the Navier-Stokes equations can be found. Therein, an error indicator is calculated in every time step and the time step size for the following step can be immediately adjusted. Such an approach seems to be efficient, as the PDE only has to be solved once. On the downside, such error estimators do not allow error estimation with respect to a goal functional and do not take the discretization error in the adjoint equation into account. Furthermore, to get a consistent optimization algorithm, we cannot adapt the time-grid each time we solve the state equation. Therefore, the advantage of only having to solve the PDE once is not given here.

8. Numerical Examples

In the following section, we test the limited memory BFGS algorithm presented in Section 6 for optimal control of the linear and nonlinear fluid-structure interaction problems in various numerical examples. Furthermore, we estimate the time, spatial, and control discretization error using the dual-weighted residual a posteriori estimator presented in Section 7. In all following examples, we provide computations using the software package RoDoBo [133]. We use the fractional-step theta time-stepping scheme derived by Galerkin discretization in Section 5 and bi-linear or bi-quadratic finite elements in space. We denote with M the number of time steps and with N the degrees of freedom of the finite element space in one component. All computations are repeated on different levels of discretization to analyze the behavior of the discretization error with global or adaptive refinement. In Section 8.1, we first evaluate the accuracy of the presented a posteriori error estimator for optimal control of a linear fluid-structure interaction problem. In Section 8.2, we apply the error estimator on the nonlinear fluid-structure interaction problem. There, we estimate the error with respect to a functional of interest and solve no optimal control problem. Finally, we solve, in Section 8.3, several optimal control problems constrained by the nonlinear model.

8.1. Optimal Control of a Linear Fluid-Structure Interaction Problem

To test the a posteriori error estimator presented in Chapter 7, we consider a simple optimal control problem governed by a linear fluid-structure interaction problem.

Configuration

The fluid flows through an elastic tube as presented in Figure 8.1. The channel has the length

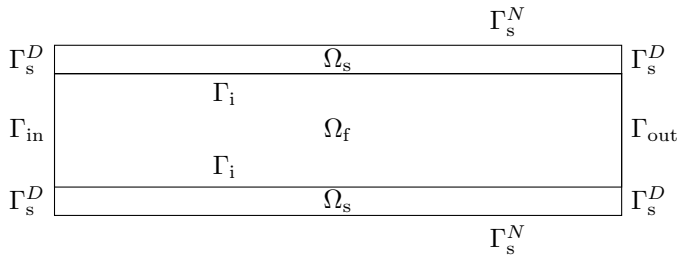

Figure 8.1.: Geometry for fluid flow through an elastic channel.

of 5 cm and the height of 1.2 cm. The surrounding elastic layers have a height of 0.1 cm each. At the inflow boundary Γ_{in}, we enforce a parabolic inflow profile

$$v(0, y) := 4(y + 0.5)(0.5 - y)v_{\text{in}}(t) \tag{8.1}$$

with the time varying inflow $v_{\text{in}}(t)$ plotted in Figure 8.2.

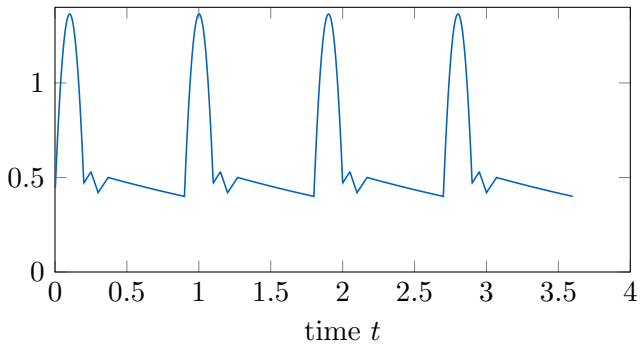

Figure 8.2.: Inflow velocity $v_{\text{in}}(t)$ plotted over time t.

At the outflow boundary Γ_{out} we use a do-nothing outflow condition as presented in Section 4.1.4. The elastic beam is clamped at the left and right boundary Γ_{s}^{D} and free at the top and bottom boundary Γ_{s}^{N}. Motivated by applications in hemodynamics, we choose the parameters given in Table 8.1, which have already been used for a similar configuration in [149]. The beam is assumed to be at rest in the initial configuration and we assume a constant parabolic flow in the channel. Hence, we have the initial conditions $u_0 = 0$, $v_0\big|_{\Omega_{\text{s}}} = 0$ and $v_0\big|_{\Omega_{\text{f}}}(x, y) = 4(y + 0.5)(0.5 - y)v_{\text{in}}(0)$.

Even if the geometry seems simple, the configuration has its difficulties. The varying fluid inflow excites the beam to oscillate in its eigenmode. As the resulting beam velocity at the interface can be quite large, the beam movement sincerely influences the fluid flow.

Table 8.1.: Material parameters (flow of a fluid through elastic channel)

ν_{f}	λ_{s}	μ_{s}	ρ_{s}	ρ_{f}	T	γ_{s}
$3.5 \cdot 10^{-2} \frac{\text{cm}^2}{\text{s}}$	$2.0 \cdot 10^{6} \frac{\text{g}}{\text{cm s}^2}$	$0.5 \cdot 10^{6} \frac{\text{g}}{\text{cm s}^2}$	$1.0 \frac{\text{g}}{\text{cm}^3}$	$0.5 \frac{\text{g}}{\text{cm}^3}$	3.6	10^5

Optimal Control Problem

The goal of this numerical example is to validate the effectivity of the a posteriori error estimator. The focus thereby will be on evaluating how accurate the estimator is able to approximate the time discretization error, for a fractional-step theta time-stepping scheme, and the control discretization error. As cost functional, we use a tracking type functional (3.27) with observation on the solid domain and Tikhonov regularization

$$J(q, u) := \frac{\gamma_{\text{s}}}{2} \int_I \|u - u_{\text{d}}\|^2_{L^2(\Omega_{\text{s}})} \, \mathrm{d}t + \frac{\alpha}{2} \|q\|^2_Q.$$

We decided to control a volume force in the solid in x- and y-direction

$$f_{\mathrm{s}} = q \quad \text{on } \Omega_s \times I.$$

Due to the symmetry of the configurations, we control the volume force on the upper and lower elastic layer at once. There is no control on the fluid domain Ω_{f}. The control q is assumed to be constant on the solid domain and time-varying. Thus we have the control space $Q = L^2(I) \times L^2(I)$.

If not stated differently, we use in this section bi-quadratic finite elements (the space $V_h^{(2)}$) to discretize in space and a fractional-step theta time-stepping scheme as presented in Section 5.1.3. The time-dependent control variable is either approximated by piecewise linear or piecewise constant functions. We stop the LBFGS Algorithm 6.1 if the gradient is reduced by a factor 10^{-3}. For the regularization parameter we chose the value $\alpha = 10^{-1}$.

We have shown in Section 7.2 for the linear fluid-structure interaction problem that the remainder terms are zero. Hence, the discrepancy between the a posteriori error estimator and the actual discretization error only comes from using an approximation of the exact solution in a higher finite element space. The quotient between error estimator and discretization error is called effectivity index I_{eff}. Values close to one indicate that both values are equivalent. If the effectivity index is close to zero or very large the error estimator either under- or overestimates the discretization error. As we are not able to acquire the exact discretization error, we compute a reference solution $\mathcal{J}_{\mathrm{ref}}$ of the cost functional on a fine grid to evaluate the effectivity indices.

Separability of Discretization Errors

Table 8.2.: Time and control discretization error estimator on different discretization levels

N	M	$\dim(Q_d)$	η_k	η_q
1617	288	194		$-3.57 \cdot 10^{-4}$
1617	576	194		$-3.50 \cdot 10^{-4}$
1617	1152	194		$-3.36 \cdot 10^{-4}$
1617	2304	194		$-3.31 \cdot 10^{-4}$
1617	4608	194		$-3.30 \cdot 10^{-4}$
1617	2304	98	$1.43 \cdot 10^{-2}$	
1617	2304	194	$1.41 \cdot 10^{-2}$	
1617	2304	386	$1.41 \cdot 10^{-2}$	
1617	2304	770	$1.41 \cdot 10^{-2}$	
1617	2304	1538	$1.41 \cdot 10^{-2}$	
425	576	386	$3.22 \cdot 10^{-1}$	$-1.13 \cdot 10^{-4}$
1617	576	386	$2.05 \cdot 10^{-1}$	$-3.48 \cdot 10^{-5}$
6305	576	386	$1.62 \cdot 10^{-1}$	$-2.07 \cdot 10^{-5}$
24897	576	386	$1.49 \cdot 10^{-1}$	$-1.75 \cdot 10^{-5}$

The equilibration of the space, time, and control discretization errors is only possible, if the errors are independent of each other. To test the separability of the a posteriori estimator,

we refine in space or time or the control variable and keep the other discretization spaces constant. In Table 8.2, we present the resulting error estimators. We denote with N the degrees of freedom of the spatial discretization and with M the number of time steps. As we can see under refinement of the spatial discretization, the control and time discretization error estimator stay constant. The same is valid for refining in time or space.

Control Discretization Error Estimator

Now we would like to evaluate how accurate the control estimator approximates the exact discretization errors. We give, in Table 8.3, the effectivity indices for the control estimator, when discretizing the control by piecewise linear functions in time. In Table 8.4, we present the effectivity indices using piecewise constant functions in time as control space. By using the fractional-step theta time-stepping scheme, we evaluate the functional on the macro time grid as suggested in Remark 7.2. As we are just interested in the control discretization error, the reference solution $\mathcal{J}_{\text{ref}}$ is computed on the same space and time grid as the results in Table 8.3 and Table 8.4, except for the control, for which we use a piecewise linear control space with $\dim(Q_d) = 1538$.

As we can see in Table 8.4 and Table 8.3 on fine control grids we get excellent effectivity indices. The underestimation of the error on the coarse grid is a sign that the higher order reconstruction of the exact solution only works well if the computed control already approximates the exact control very accurately. Comparing Table 8.3 and Table 8.4, we can see a much faster convergence of the control discretization error for the linear control space than for the piecewise constant control space. This can be expected as the variationally discretized solution q_{kh} can be approximated more accurate with control functions in the linear space than with piecewise constant functions.

Table 8.3.: Effectivitiy indices of the control discretization error estimator η_d using a linear control space

N	M	$\dim(Q_d)$	$\mathcal{J}(i_{3k}^{(1)}(\mathbf{u}_\sigma)))$	η_q	$\mathcal{J}_{\text{ref}} - \mathcal{J}(i_{3k}^{(1)}(\mathbf{u}_\sigma)))$	I_{eff}
1617	2304	98	6.23	$-5.80 \cdot 10^{-3}$	$-1.03 \cdot 10^{-2}$	0.56
1617	2304	194	6.22	$-3.31 \cdot 10^{-4}$	$-3.49 \cdot 10^{-4}$	0.94
1617	2304	386	6.22	$-1.92 \cdot 10^{-5}$	$-1.84 \cdot 10^{-5}$	1.03
1617	2304	770	6.22	$-1.42 \cdot 10^{-6}$	$-1.03 \cdot 10^{-6}$	1.37

Table 8.4.: Effectivitiy indices of the control discretization error estimator η_d using a piecewise constant control space

N	M	$\dim(Q_d)$	$\mathcal{J}(i_{3k}^{(1)}(\mathbf{u}_\sigma)))$	η_q	$\mathcal{J}_{\text{ref}} - \mathcal{J}(i_{3k}^{(1)}(\mathbf{u}_\sigma)))$	I_{eff}
1617	2304	96	6.28	$-2.85 \cdot 10^{-2}$	$-6.21 \cdot 10^{-2}$	0.45
1617	2304	192	6.24	$-1.39 \cdot 10^{-2}$	$-1.67 \cdot 10^{-2}$	0.83
1617	2304	384	6.22	$-4.06 \cdot 10^{-3}$	$-4.20 \cdot 10^{-3}$	0.96
1617	2304	768	6.23	$-1.03 \cdot 10^{-3}$	$-1.00 \cdot 10^{-3}$	1.02

Time Discretization Error Estimator

Now we keep the control discretization constant and refine in time. For the fractional-step theta time-stepping scheme, we get optimal results in Table 8.5. The error in the cost functional reduces constantly with refinement of the timegrid and the error estimator neither overestimates nor underestimates the error extremely. We computed here again the reference value $\mathcal{J}_{\text{ref}}$ on a fine time grid with $M = 4608$ time steps and kept the spatial and control discretization as in Table 8.5.

Table 8.5.: Effectivitiy indices of the time discretization error estimator η_k

N	M	$\dim(Q_d)$	$\mathcal{J}(i_{3k}^{(1)}(\mathbf{u}_\sigma)))$	η_k	$\mathcal{J}_{\text{ref}} - \mathcal{J}(i_{3k}^{(1)}(\mathbf{u}_\sigma)))$	I_{eff}
1617	288	194	5.44	$5.84 \cdot 10^{-1}$	$7.93 \cdot 10^{-1}$	0.73
1617	576	194	6.01	$2.05 \cdot 10^{-1}$	$2.25 \cdot 10^{-1}$	0.91
1617	1152	194	6.18	$5.52 \cdot 10^{-2}$	$5.49 \cdot 10^{-2}$	1.01
1617	2304	194	6.22	$1.41 \cdot 10^{-2}$	$1.11 \cdot 10^{-2}$	1.27

Error Estimator Spatial Discretization

To test the spatial a posteriori error estimator, we used in comparison to the previous computations bi-linear finite elements. The reference value $\mathcal{J}_{\text{ref}}$ was computed on a fine grid with $N = 24897$ degrees of freedom and second-order finite elements. The effectivity indices in Table 8.6 for the spatial discretization estimator are surprisingly not perfect. On the coarse grid the error is underestimated and on finer grids slightly overestimated. A closer look at the convergence rate of the error in the cost functional, shows, that the grids we computed on are still too coarse to be in the convergence range. This can be an explanation for the suboptimal higher order reconstruction in the error estimator.

Table 8.6.: Effectivitiy indices of the spatial discretization error estimator η_h (bi-linear finite elements)

N	M	$\dim(Q_d)$	$\mathcal{J}(i_{3k}^{(1)}(\mathbf{u}_\sigma)))$	η_h	$\mathcal{J}_{\text{ref}} - \mathcal{J}(i_{3k}^{(1)}(\mathbf{u}_\sigma)))$	I_{eff}
117	576	386	3.50	$6.99 \cdot 10^{0}$	$2.06 \cdot 10^{0}$	3.39
425	576	386	9.44	$-1.10 \cdot 10^{1}$	$-3.87 \cdot 10^{0}$	2.86
1617	576	386	5.10	$2.91 \cdot 10^{-1}$	$4.55 \cdot 10^{-1}$	0.63
6305	576	386	7.16	$-1.00 \cdot 10^{0}$	$-1.60 \cdot 10^{0}$	0.62
24897	576	386	5.91	$-4.16 \cdot 10^{-1}$	$-3.54 \cdot 10^{-1}$	1.17

If we use bi-quadratic finite elements instead of bi-linear finite elements, we get a fast and uniform reduction of the error in the cost functional, as we can see in Table 8.7. As the error estimator for bi-quadratic finite elements is not fully supported by the software library RoDoBo [133] yet, we cannot present effectivity indices here. One explanation for the reduced convergence rate of the error in the cost functional for bi-linear finite elements could be the use of very anisotropic cells as well as locking effects due to the thin beam structure.

Table 8.7.: Error in the cost functional under spatial discretization (quadratic finite elements)

N	M	$\dim(Q_d)$	$\mathcal{J}(i_{3k}^{(1)}(\mathbf{u}_\sigma)))$	$\mathcal{J}_{\text{ref}} - \mathcal{J}(i_{3k}^{(1)}(\mathbf{u}_\sigma)))$
117	576	386	5.13	$4.30 \cdot 10^{-1}$
425	576	386	7.53	$-1.97 \cdot 10^{0}$
1617	576	386	6.01	$-4.52 \cdot 10^{-1}$
6305	576	386	5.65	$-9.88 \cdot 10^{-2}$

Test of the LBFGS Algorithm

In Section 6.3, we claimed that only the gradient information of very few steps is necessary to construct an appropriate approximation of the Hessian. We use in the following the gradient of the last $m = 3$, $m = 6$ and the last $m = 12$ steps. We compute the solution using bi-quadratic finite elements on a mesh with $N = 1617$ degrees of freedom and a time grid with $M = 2304$ time nodes. For the control, we use a piecewise linear discretization. The optimization algorithm terminates in all cases if the gradient is reduced by a factor of 10^{-3}.

If we use the gradient information of the last $m = 12$ steps, the LBFGS algorithm already converges after 9 steps independent of the dimension of the control space. As no gradient information is omitted, the LBFGS algorithm coincides with the BFGS algorithm. This explains the extreme fast super-linear convergence we could see in the last steps. If we use the gradient information of the last $m = 3$ or $m = 6$ steps, the convergence rate is still fast, nevertheless, we loose the super-linear convergence. Hence, we have to choose the termination criteria very reasonably as reducing the gradient by additional 10% can mean a serious number of extra optimization loops. At least for the optimal control problem at hand, the number of optimization loops in Table 8.8 is in all cases independent of the control dimension.

Table 8.8.: Number of optimization loops in the LBFGS algorithm

$\dim(Q_d)$	LBFGS with $m = 3$	LBFGS with $m = 6$	LBFGS with $m = 12$
50	11	9	9
98	9	12	9
194	11	12	9
386	11	12	9
770	9	12	9
1538	11	12	9

Implementation of the Solid Velocity

For proving existence and regularity results for state and adjoint equation at the same time, we introduced, in Section 3.1.4, the solid velocity in an unconventional way. The question arises if this formulation has any numerical advantages. Hence, we compute the optimal control problem demanding

$$(\!(\nabla \partial_t u, \nabla \psi)\!)_{\text{s}} - (\!(\nabla v, \nabla \psi)\!)_{\text{s}} = 0 \quad \forall \psi \in L^2(I; V_{\text{s}}),$$

as in (3.9). We denote the resulting cost functional value with $\mathcal{J}_1(\mathbf{u})$. In addition we solve the optimal control problem of the linear fluid-structure interaction problem by using the standard formulation to introduce the solid velocity variable as in (3.3):

$$(\!(u_t, \psi)\!)_{\mathrm{s}} - (\!(v, \psi)\!)_{\mathrm{s}} = 0 \quad \forall \psi \in L^2(I; H_{\mathrm{s}}).$$

We denote the resulting value of the cost functional with $\mathcal{J}_2(\mathbf{u})$. The example is computed on a fixed fine grid with $M = 1617$ degrees of freedom, $N = 2304$ time steps and a control space with $\dim(Q_d) = 386$. Similar to previous examples we compute a reference solution for the cost functional $\mathcal{J}_{\mathrm{ref}}$ on a fine grid with $M = 6305$, $N = 2304$ time steps and a control space with $\dim(Q_d) = 770$.

Comparing the functional values we can see that the choice of implementation rather has no influence. The difference in the cost functionals given by $\mathcal{J}_1(\mathbf{u}_\sigma) - \mathcal{J}_2(\mathbf{u}_\sigma) = 3.49 \cdot 10^{-7}$ is much smaller then the actual discretization error $\mathcal{J}_1(\mathbf{u}_\sigma) - \mathcal{J}_{\mathrm{ref}} = 3.943660 \cdot 10^{-1}$ and $\mathcal{J}_2(\mathbf{u}_\sigma) - \mathcal{J}_{\mathrm{ref}} = 3.943656 \cdot 10^{-1}$. Moreover, the algorithm terminates in both cases after 11 steps. Hence, we cannot see any numerical benefit in using either the formulation (3.9) or (3.3) here. Therefore, we do not expect any disadvantages, if we use in the case of the nonlinear fluid-structure interaction problem the standard formulation (3.3).

8.2. A posteriori Error Estimation for a Nonlinear FSI Problem

In this section, we discuss several numerical examples in order to demonstrate our algorithmic developments. We apply the a posteriori error estimator in Theorem 7.2 on a nonlinear fluid-structure interaction configuration. In this section, we only estimate the discretization error with respect to a functional of interest and do not solve an optimization problem. The first two examples are the original FSI-2 and FSI-3 benchmark tests proposed in [142, 41, 40]. We compute them with the originally proposed inflow profile as well as with a modified inflow profile. In the latter, we decrease the inflow (after full oscillations have been developed) in order to study how the error estimator can deal with time step refinement and coarsening. In the final example, we consider a flapping test proposed in [73] that is inspired by hemodynamics applications. The results have not been published yet. A comparison with heuristic error estimators will be the topic of the article [59] currently in preparation.

8.2.1. FSI-2 and FSI-3 Benchmark Configuration

We define first the configuration and material parameters for both the FSI-2 and FSI-3 benchmark. In the FSI-2 and FSI-3 benchmark a beam of length 0.35m and height 0.02m is attached to a cylinder with radius $r = 0.05$m and circle-center $C = (0.2, 0.2)$. The cylinder is located in a channel of length 2.5m and height 0.41m filled by a fluid as in Figure 8.3. On the cylinder and on the outer boundary Γ_{f} as well as on Γ_{s} we enforce zero Dirichlet boundary conditions. At the outflow boundary Γ_{out} we apply a do-nothing outflow condition. The parabolic inflow profile on Γ_{in} is given by:

$$v(0, y) := 1.5y(0.41 - y)\frac{4}{0.41^2} v_{\mathrm{in}}(t). \tag{8.2}$$

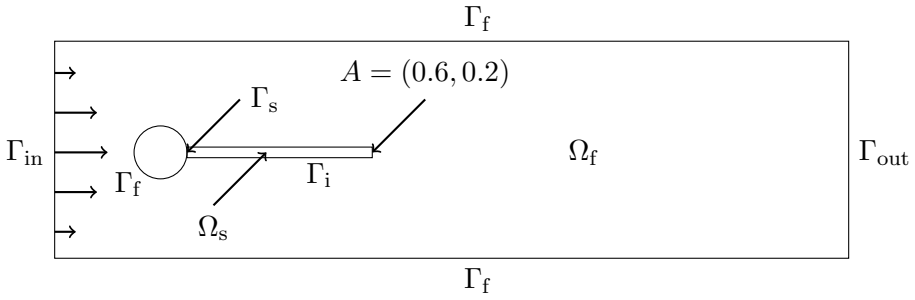

Figure 8.3.: Geometry for flow around cylinder with elastic beam.

In the original benchmark the inflow velocity $v_{\text{in}}(t)$ is slowly increased and kept constant after some time. We will use in addition a second inflow profile which decreases again after some time. We consider the FSI-2 configuration on the time interval $I = (0, 25)$ and analyze the FSI-3 configuration on the time interval $I = (0, 12)$. The material parameters are chosen as proposed in [142, 41, 40] and listed in Table 8.9.

Table 8.9.: Material parameters for the FSI-2, FSI-3 and Flapping Test.

	FSI-2	FSI-3	Flapping
ν_f	$10^{-3}\,\frac{m^2}{s}$	$10^{-3}\,\frac{m^2}{s}$	$10^{-1}\,\frac{cm^2}{s}$
λ_s	$0.5 \cdot 10^6\,\frac{kg}{ms^2}$	$2.0 \cdot 10^6\,\frac{kg}{ms^2}$	$2.0 \cdot 10^7\,\frac{g}{cms^2}$
μ_s	$2.0 \cdot 10^6\,\frac{kg}{ms^2}$	$8.0 \cdot 10^6\,\frac{kg}{ms^2}$	$8.0 \cdot 10^7\,\frac{g}{cms^2}$
ν_m	-0.1	-0.1	-0.1
ρ_s	$10^4\,\frac{kg}{m^3}$	$10^3\,\frac{kg}{m^3}$	$10^2\,\frac{g}{cm^3}$
ρ_f	$10^3\,\frac{kg}{m^3}$	$10^3\,\frac{kg}{m^3}$	$10^2\,\frac{g}{cm^3}$

To compare the computed solutions $\mathbf{u}_{kh} = (v_{kh}, u_{kh}, p_{kh})$ of the time-space discretized FSI model in Problem 5.5 on different temporal meshes, we evaluate several time-dependent functionals. As in [142, 41, 40], we calculate the displacement u in x- and y-direction at the point $A = (0.6, 0.2)$ and integrate the values over time to have a functional value:

$$\mathcal{J}_1(\mathbf{u}) := \int_I u_1^2(A,t)\ \mathrm{d}t, \quad \mathcal{J}_2(\mathbf{u}) := \int_I u_2^2(A,t)\ \mathrm{d}t. \tag{8.3}$$

In addition, we calculate the drag at the combined outer boundary of cylinder and solid flag which depends on the solid displacement u and the fluid velocity v and pressure p:

$$\mathcal{J}_3(\mathbf{u}) := \int_I \int_{\Gamma_{f,\text{circle}}} -\sigma_f(v) n e_1\ \mathrm{d}x\ \mathrm{d}t + \int_I \int_{\Gamma_i} -\Sigma(u) n e_1\ \mathrm{d}x\ \mathrm{d}t. \tag{8.4}$$

A further characteristic value of the fluid velocity field v is the vorticity value. Due to the boundary layers, we use the Okubo-Weiss criterion which was first introduced in [121, 147]. The vorticity is given by the positive values of $\det(\nabla v)$. As the vorticity value is defined on the moving domain we have to transform the criteria to the reference domain. To be able to

calculate sensitivities of the functional we use the regularized version

$$\mathcal{J}_4(\mathbf{u}) := \int_I \int_{\Omega_\mathrm{f}} Jg(\det(\nabla v F^{-1})\ \mathrm{d}t\ \mathrm{d}x \text{ with } \quad g(x) = \begin{cases} 0 & x < 0 \\ \frac{t^3}{1+t^2} & x > 0. \end{cases} \tag{8.5}$$

suggested in [96]. In addition to the functional evaluation on the whole fluid domain, we calculate the vorticity value in the area $\tilde{\Omega}_\mathrm{f} := \{v \in \Omega_\mathrm{f} | x > 0.9\}$ behind the beam:

$$\mathcal{J}_5(\mathbf{u}) := \int_{\tilde{\Omega}_\mathrm{f}} \int_I Jg(\det(\nabla v F^{-1})\ \mathrm{d}t. \tag{8.6}$$

8.2.2. A Comparison of Various Time-Stepping Schemes

Before we start with detailed analysis of the adaptive algorithm, we briefly want to motivate our choice to work with the fractional-step theta time-stepping scheme. To do so, we perform a computational analysis in which we compare the behavior of the 2nd order unconditionally stable Crank-Nicolson scheme ($\theta_m = 0.5$), the 2nd order strictly A-stable shifted Crank-Nicolson scheme ($\theta_m = 0.5 + \varepsilon_m$), and the fractional-step theta scheme.

As shown in [149], for FSI-2, both the shifted Crank-Nicolson and fractional-step theta scheme work well. However, one drawback of the shifted Crank-Nicolson scheme is the determination of the shift value. On the one hand stability can only be observed for the shift value $\varepsilon_m \approx k_m$. On the other hand for the shift value must hold $\varepsilon_m < 0.5$, since otherwise $\theta_m \in [0, 1]$ is violated. Especially on an adaptively refined time grid it is unclear how the shift value has to be chosen.

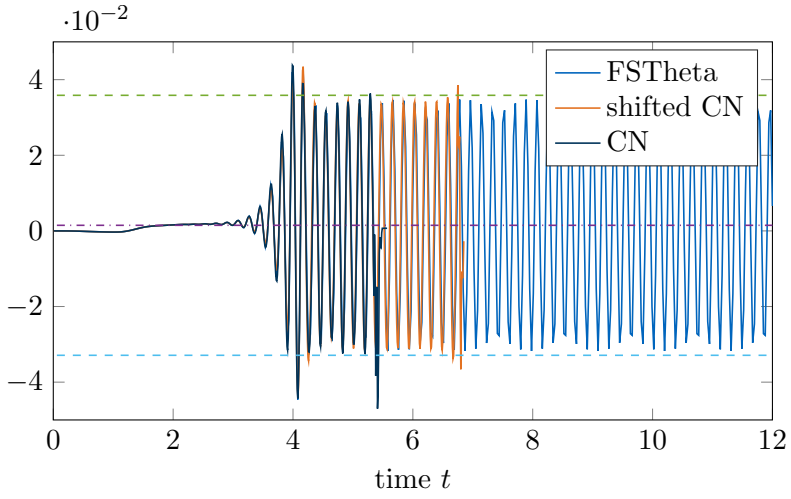

Figure 8.4.: Comparison of various time-stepping schemes showing blow-up for the two Crank-Nicolson schemes in the u_2 displacements at $A = (0.6, 0.2)$ with time step $k_m = 10^{-2}$ for the FSI-3 benchmark

For FSI-3 such comparison has not yet been made. We use the smooth increasing inflow

profile

$$v_{\text{in}}(t) = \begin{cases} 1 - 1\cos(\frac{\pi t}{2}) & t \leq 2 \\ 2 & t > 2 \end{cases}$$

and compute the solution on a rather coarse time-grid with the time step size $k_m = 10^{-2}$. We plot the resulting displacement u at the tip of the flag A in y-direction in Figure 8.4. We observe that both Crank-Nicolson and the shifted version will fail after some time. Since adaptive algorithms only pay off, if we are able to solve the partial differential equations already on rather course grids, we conclude that the fractional-step theta time-stepping scheme seems to be the optimal choice as time-stepping scheme in combination with adaptivity.

8.2.3. FSI-3 Benchmark with new Inflow

To avoid artificial oscillations and large numerical errors, the authors in [142] suggest to increase the inflow velocity slowly and smoothly. In the following, we choose an inflow profile, which lacks this regularity to test if the adaptive algorithm is able to tackle the kinks in the inflow $v_{\text{in}}(t)$. The inflow velocity $v_{\text{in}}(t)$ is given by

$$v_{\text{in}}(t) = \begin{cases} t & t \leq 2 \\ 2 & 2 < t \leq 6 \\ 8 - t & 6 < t \leq 7 \\ 1 & \text{else,} \end{cases} \tag{8.7}$$

and decreases again after full oscillations have been developed, to enforce a varying behavior over time.

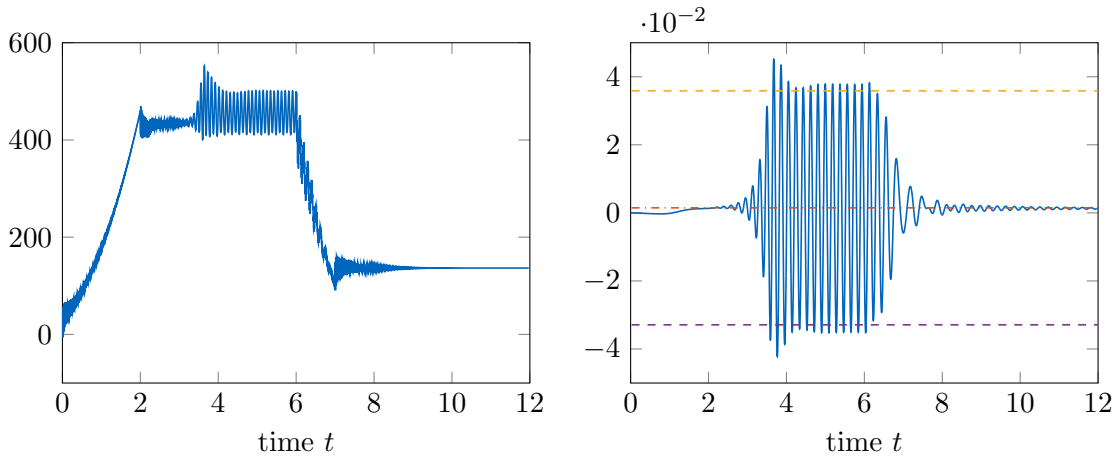

Figure 8.5.: Solution of the FSI-3 benchmark using the new inflow profile (8.7). Drag value at the flag and cylinder (left) and displacement u in y-direction (right) plotted over time t.

In Figure 8.5, we plotted the displacement at the tip of the flag A and the drag value on the cylinder and interface over time (see definition of $\mathcal{J}_3(\mathbf{u})$ in (8.4)). The solution is computed

with bi-quadratic finite elements on a mesh with $N = 4128$ degrees of freedom and $M = 144384$ time steps for the fractional-step theta time-stepping scheme. As in the standard FSI-3 benchmark the beam starts to oscillate after some time, whereby frequency and amplitude are equivalent to the standard benchmark for $t < 6$. The moment the inflow decreases, the amplitude of the oscillations reduces and the configuration converges to a steady state. This behavior can also be seen in the drag value.

Functional of Interest: Vorticity

To compare the accuracy of the computed solution on different time grids, we choose the vorticity functional $\mathcal{J}_5(\mathbf{u})$ defined on $\tilde{\Omega}_\mathrm{f} := \{v \in \Omega_\mathrm{f} | x > 0.9\}$. The error in the functional of interest due to time discretization is plotted in Figure 8.6 over the number of time steps. The space discretization is kept constant with $N = 4128$. As reference value for the functional $\mathcal{J}_\mathrm{ref}$ we use the solution on a very fine equidistant time grid with $M = 144384$ time steps and the same spatial discretization. The error in the functional of interest thereby converges slightly faster using the adaptive algorithm with the DWR error estimator as refinement indicator, instead of global refinement in time. In addition, the estimates η_k of the time discretization error in Figure 8.6 are very close to the exact values of the discretization error.

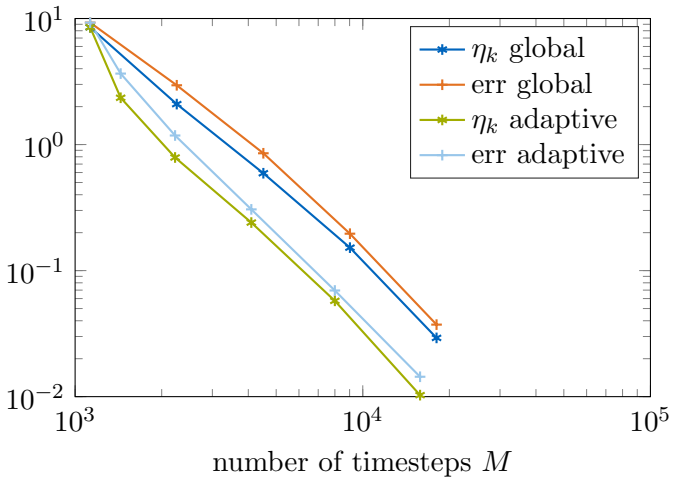

Figure 8.6.: Error in the vorticity functional $\mathcal{J}_5(\mathbf{u})$ and time discretization error estimator η_k for the FSI-3 benchmark plotted over number of time steps M.

Functional of Interest: Drag

Furthermore, we evaluate the functional $\mathcal{J}_3(\mathbf{u})$, measuring the drag around the cylinder and the flag, as well as the time discretization error estimator on different time grids with M time steps. As already done in Section 8.1, we evaluate the effectivity indices I_eff refining either globally or adaptively in time. The effectivity indice I_eff is defined as quotient of time discretization error estimator η_k and exact discretization error $\mathcal{J}_\mathrm{ref} - \mathcal{J}(\mathbf{u}_{kh})$. As reference value for the functional $\mathcal{J}_\mathrm{ref}$ we use again the solution on a very fine equidistant time grid with $M = 144384$ time steps.

Table 8.10.: Effectivity indices I_{eff} for the time discretization error estimator η_k with respect to $\mathcal{J}_3(\mathbf{u})$ on globally refined time grids

M	1128	2256	4512	9024	18048	36096		
I_{eff}	0.86	0.84	0.94	1.35	0.69	1.18		

Table 8.11.: Effectivity indices I_{eff} for the time discretization error estimator η_k with respect to $\mathcal{J}_3(\mathbf{u})$ on adaptively refined time grids

M	1128	1458	2310	4122	4890	7440	9420	9420
I_{eff}	0.86	0.84	0.94	1.98	2.85	0.89	0.86	0.88

The effectivity indices in Tables 8.10 and 8.11 are very close to one. Hence, the error in the functional of interest is neither over- nor underestimated. Thus, the error estimator can be used additionally as a reliable stopping criteria for the simulation. Furthermore, if error estimates for the space discretization are available, we can decide if it is necessary to refine in space and time or just in one of the two. We refer to the example in Section 8.2.5, where we equilibrate space and time discretization error.

To get a fast converging adaptive algorithm, having an accurate localization of the error is also important to know where to refine the time grid. In Figure 8.7, we plotted the time step size k_m over time t after 1, 3 and 6 refinement steps. The DWR algorithm refines the areas where the beam oscillates with faster frequency as we would expect. In addition, the adaption strategy does not refine the time grid for $t > 8$, when the solution converges against a stationary solution. Surprisingly the algorithm refines quite heavily close to the initial condition. However, taking a closer look at the drag values in Figure 8.5, we can see that the fast increasing inflow causes oscillations in the drag value.

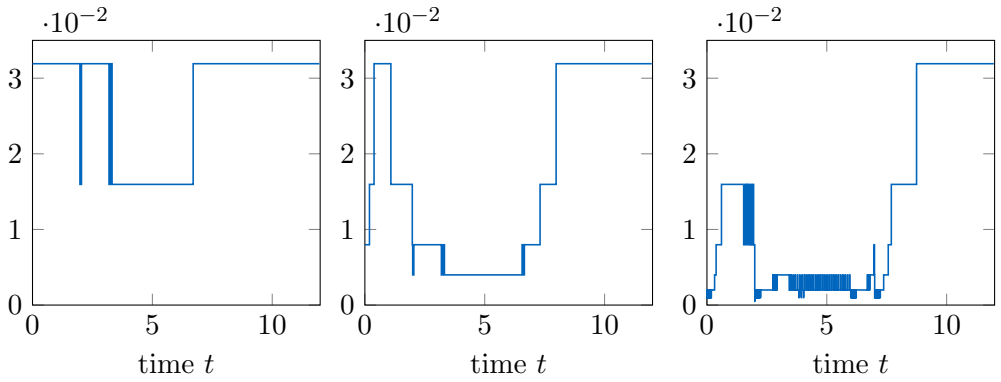

Figure 8.7.: Time step size k_m plotted over time t after 1 (left), 3 (middle) and 6 (right) adaptive refinements for the FSI-3 benchmark.

8.2.4. FSI-2 Benchmark with new Inflow

The geometry for the FSI-2 benchmark is identical to the FSI-3 benchmark as given in Figure 8.3. However, as the material parameters and maximum inflow vary, the beam starts to

oscillate with lower frequency and with higher amplitude. As in the previous simulation, we choose an inflow profile

$$v_{\text{in}}(t) = \begin{cases} 0.5t, & t \leq 2 \\ 1, & 2 < t \leq 14 \\ 8 - 0.5t, & 14 < t \leq 15 \\ 0.5, & \text{else} \end{cases} \tag{8.8}$$

rising linearly and after some constant period, the inflow profile decreases again.

In Figure 8.8, we plotted the displacement u in x- and y-direction at the tip of the flag at the point $A = (0.6, 0.2)$ over time. Here we used again bi-quadratic finite elements with $N = 4128$ and $M = 149760$ times steps in the fractional-step theta time-stepping scheme. We will use the resulting solution as reference solution in the following. In comparison to the FSI-3 configuration (see Figure 8.5), the damping of the beam due to the fluid flow is much slower, especially as the simulation runs here until $T = 25$ instead of $T = 12$. In the time frame $t \in [10, 14]$, the frequency and amplitude coincide again with the reference values of the standard FSI-2 benchmark.

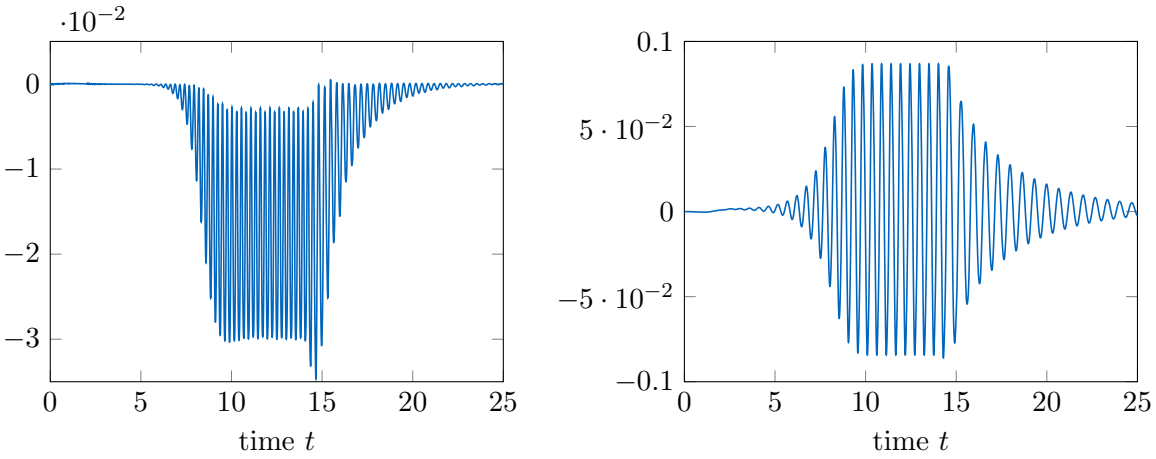

Figure 8.8.: Solution of the FSI-2 benchmark using the new inflow profile (8.8). Displacement u in x-direction (left) and y-direction(right) at the tip of the flag A plotted over time.

To analyze the convergence rate of the fractional-step theta time-stepping scheme, we plotted in Figure 8.9 the error in the functional $\mathcal{J}_2(\mathbf{u})$. The functional contains the displacement values u at the tip of the flag A in y-direction integrated over time. Due to the lower inflow velocity, the solution is compared to the FSI-3 benchmark smoother in time. In addition, even at the end time point T the solution has not reached a steady state. That is why global refinement is reasonable here. The adaptive algorithm does not provide any advantages with respect to convergence in time.

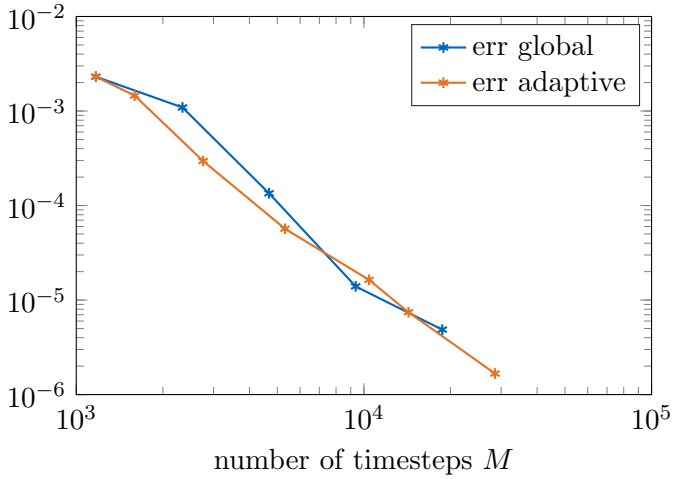

Figure 8.9.: Error in the functional $\mathcal{J}_2(\mathbf{u})$ for the FSI-2 benchmark plotted over number of time steps M.

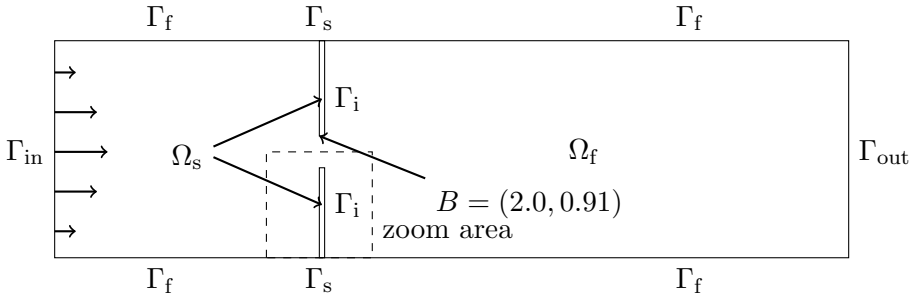

Figure 8.10.: Flow through two elastic beams.

8.2.5. Flapping Test

In [73] an alternative test configuration was suggested. The fluid in a channel of length 8cm and height 1.61cm flows through the gap between two vertical elastic beams of width 0.0121cm and height 0.7cm (see Figure 8.10), whereby the fluid flow induces a flapping of the two beams. At the inflow boundary Γ_{in} we enforce a parabolic inflow profile

$$v(0, y) := y(1.61 - y)\frac{4}{1.61^2}v_{in}(t).$$

Motivated by applications in hemodynamics, we choose for the mean inflow velocity $v_{in}(t)$ the periodic profile given in Figure 8.11. This configuration was developed in [73] to test alternative methods to the ALE mapping used in this thesis. For large inflow velocities, the algorithm fails as the fluid-structure interaction equation cannot be solved anymore due to the large displacement of the thin beam. Hence, the maximal value of parabolic inflow profile is chosen in such a way, that the mesh motion can still be handled by the ALE method.

As in the FSI benchmark, we enforce zero Dirichlet boundary conditions at the outer boundary Γ_f and Γ_s and a do-nothing outflow condition at the outflow boundary Γ_{out}. The computations

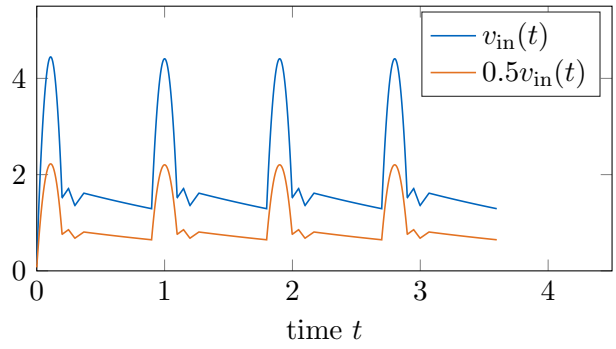

Figure 8.11.: Inflow velocities $v_{\text{in}}(t)$ and $0.5v_{\text{in}}(t)$ plotted over time t

are carried out on the time interval $I = (0, 3.6)$ and the material parameter values are given in Table 8.9.

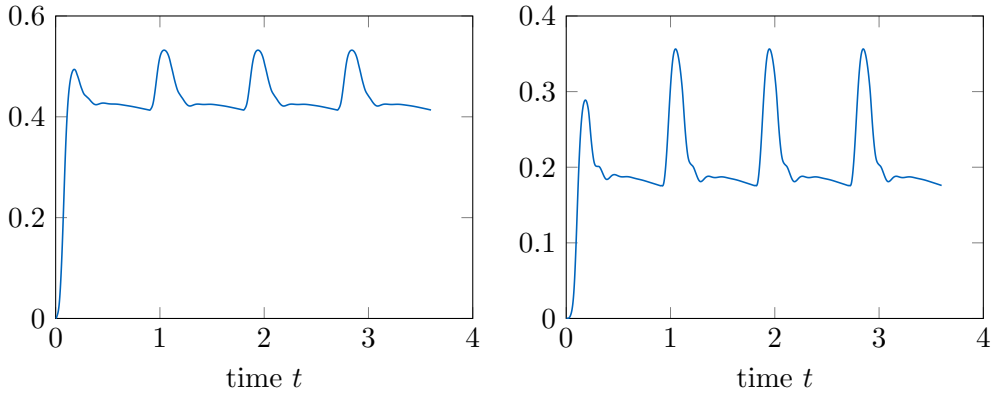

Figure 8.12.: Solution of the flapping test. Displacement u at the point B in x-direction (left) and displacement u at the point B in y-direction (right) plotted over time t.

To visualize the behavior of the space-time discretized FSI model in Problem 5.5 for the given configuration, we plotted the displacement of the flag at the point $B = (2.0, 0.91)$, the drag around the two beams, and the vorticity over time in Figure 8.12 and 8.13. For the simulation, we use bi-quadratic finite elements with $N = 5537$ degrees of freedom and the fractional-step theta time-stepping scheme with $M = 184320$ time steps. Due to the varying inflow, the beams are at rest most of the time. The moment the inflow profile increases the gap between the beams increases as well. In addition the change in the fluid velocity induces large changes in the drag and vorticity values.

As in the previous configurations, we evaluate a functional of interest to compare the time and spatial discretization error on different time and spatial grids. Here, we choose the vorticity value on the domain Ω_{f} given by the functional $\mathcal{J}_4(\mathbf{u})$ defined in (8.5).

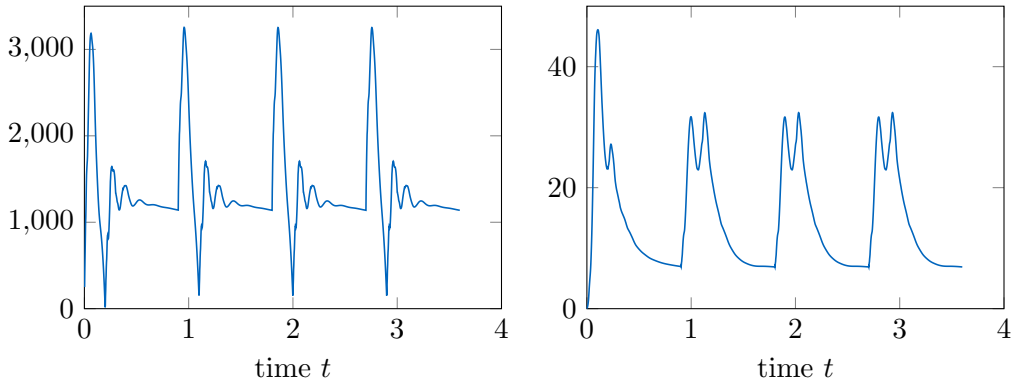

Figure 8.13.: Solution of the flapping test. Drag value (left) and vorticity value (right) plotted over time t.

Adaptive Refinement in Time

First, we solve the configuration on different globally refined time grids. Then, we use the DWR error estimator presented in Theorem 7.2 to refine the time grid adaptively. We plotted again the error in vorticity functional over the number of time steps M. The discretization error is computed using as reference value $\mathcal{J}_{\text{ref}}$ the vorticity functional computed on a fine grid. The reference solution is solved on the same spatial grid with $N = 5537$ degrees of freedom and $M = 184320$ time steps. Furthermore we use bi-quadratic finite elements. As we can see in Figure 8.14, the error in the functional of interest reduces much faster if we use the adaptive algorithm in comparison to the error by global refinement of the time grid.

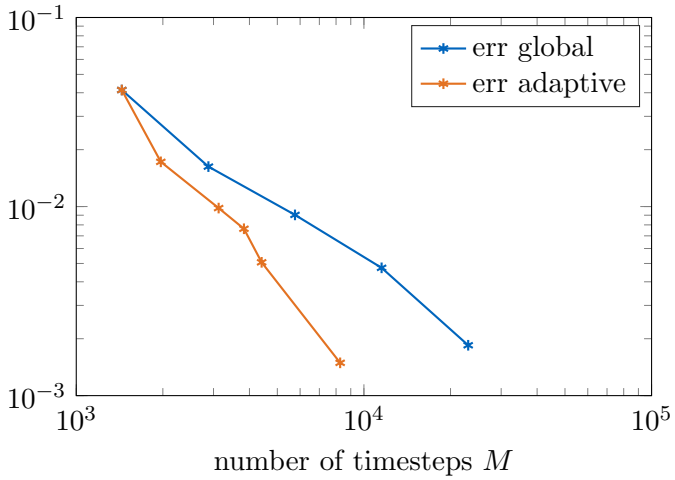

Figure 8.14.: Error in the vorticity functional $\mathcal{J}_4(\mathbf{u})$ for the flapping test plotted over number of time steps M.

If we take a closer look on the resulting time step size in Figure 8.15 we see that the adaptive algorithm chooses locally very small time steps. The adaptive algorithm refines especially in the areas of large inflow velocity, but also sees the kinks in the inflow profile and refines there.

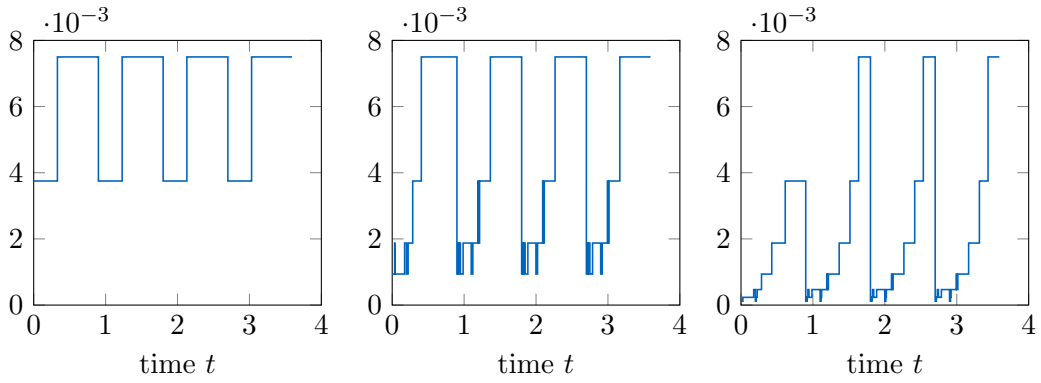

Figure 8.15.: Time step size k_m plotted over time t after 1 (left), 3 (middle) and 6 (right) adaptive refinements for the flapping test.

Space-Time Adaptive Refinement with Equilibration

One of the major advantages of dual-weighted residual error estimators is the possibility to equilibrate the spatial and time discretization error. To demonstrate the efficiency of such an approach, we refine globally in space and time and compare the error in the vorticity functional with the discretization error we obtain by refining adaptively. The adaptive algorithms only refines in space and time, if the error indicators η_k and η_h have the same order of magnitude.

In comparison to the previous sections, we use bi-linear finite elements here. The reference solution is computed on a fine grid with bi-quadratic finite elements with $N = 20273$ degrees of freedom and $M = 23040$ time steps. In addition, we reduced the inflow velocity to $0.5v_{\text{in}}(t)$ as given in Figure 8.11.

Table 8.12.: Effectivity indices of spatial and time discretization error estimator using constant Young's modulus $\mathcal{E}_m$ (left) and modified distributed Young's modulus $\mathcal{E}_m$ in the mesh motion equation for the flapping test

$\mathcal{E}_m$ const		$\mathcal{E}_m$ distributed	
$N \cdot M$	I_{eff}	$N \cdot M$	I_{eff}
$5.05 \cdot 10^5$	$4.70 \cdot 10^{-1}$	$5.05 \cdot 10^5$	$4.82 \cdot 10^{-1}$
$1.06 \cdot 10^6$	$7.21 \cdot 10^{-1}$	$1.06 \cdot 10^6$	$6.85 \cdot 10^{-1}$
$2.35 \cdot 10^6$	$-3.09 \cdot 10^0$	$2.35 \cdot 10^6$	$8.98 \cdot 10^{-1}$
$5.11 \cdot 10^6$	$-9.76 \cdot 10^0$	$5.44 \cdot 10^6$	$1.09 \cdot 10^0$
$1.06 \cdot 10^7$	$-4.43 \cdot 10^1$	$1.28 \cdot 10^7$	$6.73 \cdot 10^{-1}$
$1.19 \cdot 10^7$	$1.56 \cdot 10^3$	$3.41 \cdot 10^7$	$1.01 \cdot 10^2$

To be able to handle the large deformations we weight the Young's modulus $\mathcal{E}_m$ in the mesh motion equation with the inverse determinant of the deformation gradient J^{-1}. Nevertheless the regularity of the ALE transformation is disturbing. The volume of the cells in the transformed mesh reach a critical value and the angles in the cells around the tip of the flag are either extremely small or very large. If we set the mean inflow slightly larger, the Newton algorithm will not converge anymore.

The lack of regularity of the mesh motion around the tip causes the dual weighted residual error estimator to rise on the critical cells. However, as we can see in Table 8.12 the error estimator overestimates the discretization error heavily. We give in Table 8.12 the effectivity indices I_{eff} of the DWR error estimator of the combined error of time and spatial discretization on the adaptively refined grid. As the mesh deteriorates in the area around the tip of the flag, the determinant of the deformation gradient $J = \det(F)$ converges to zero. The neglected remainder terms in the a posteriori error estimator, including the derivatives of F^{-1}, cannot be assumed small anymore. This effect was already predicted in [128].

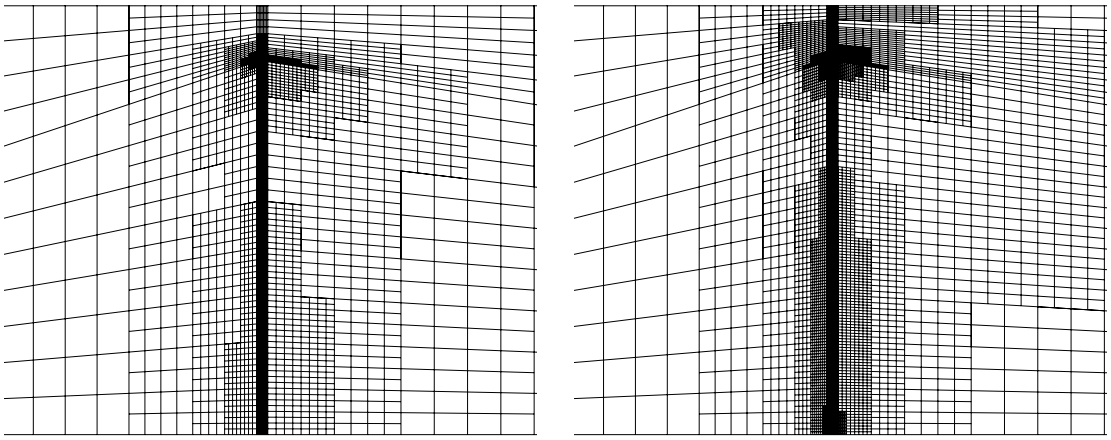

Figure 8.16.: Mesh after 6 refinement steps with constant Young's modulus (left) and after 6 refinement steps with modified distributed Young's modulus (right). Zoom into the area of the beam.

As the error estimator is especially located in the area around the tip of the two flags the adaptive algorithm only refines very locally in this region. The mesh after 6 refinement steps is given in Figure 8.16. However, the refinement of the mesh around the flag does not increase the regularity of the ALE transformation. That is why the quality of the a posteriori error estimator does not improve with local refinement.

In Figure 8.17, we plotted the discretization error in the vorticity functional over the degrees of freedom of the space-time discretized fluid-structure interaction problem ($N \cdot M$). The error in $\mathcal{J}_4(\mathbf{u})$ converges using the adaptive algorithm much faster then by global refinement in space and time. However, the overestimation of the error results in local refinement around the tip of the flag, which does not contribute to a further decrease in the discretization error, as we can see in Figure 8.17.

To improve the regularity of the ALE transformation, we use now a distributed value for the Young's modulus $\mathcal{E}_m$ in the mesh motion equation. The chosen distribution thereby is motivated by the results of the optimal control problem in Section 8.3.3, see Figure 8.32. Thereby, the mesh properties of the cells around the tip can be conserved.

The effectivity indices in Table 8.12 are now very close to the desired value one. Only on the finest level the cell distortion influences the error estimator. In comparison to the mesh generated by solving a linear elasticity equation with constant Young's modulus, the adaptive algorithm refines now in a much wider region and not only at the tip of the flag as we can see in Figure 8.16.

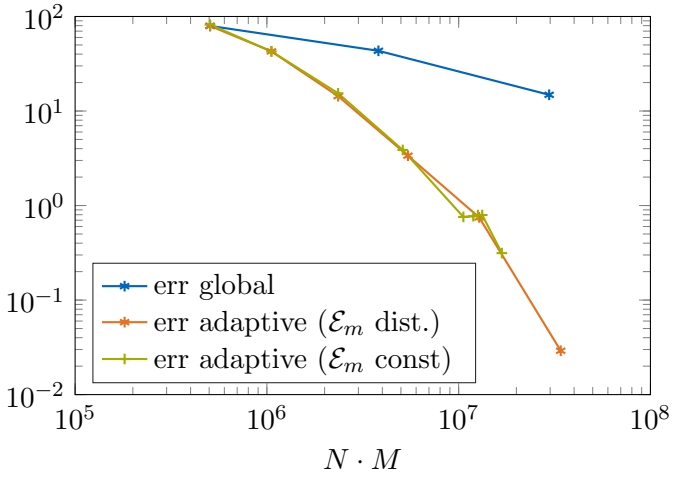

Figure 8.17.: Error in the vorticity functional for the flapping test plotted over the degrees of freedom on the space-time cylinder ($N \cdot M$) using global and adaptive refinement strategies.

Altogether we can see in Figure 8.17 an extreme fast reduction of the discretization error using the adaptive algorithm in comparison to global refinement in space and time. For the identical number of degrees of freedom of spatial and time discretization we can compute the vorticity functional 10^3-times more accurate by applying the equilibrated adaptive approach. As the error estimates of the time discretization error, given in Table 8.13, are much smaller than the spatial discretization error the algorithm only refines in space. Hence, this example shows an additional advantage of error estimators in space and time, because they indicate which errors dominate and we can avoid unnecessary refinement.

Table 8.13.: Error estimators in space η_h and time η_k for the adaptive refinement algorithm with equilibration

N	M	η_h	η_k	η
351	1440	$-3.84 \cdot 10^{1}$	$4.43 \cdot 10^{-2}$	$-3.84 \cdot 10^{1}$
733	1440	$-2.93 \cdot 10^{1}$	$3.10 \cdot 10^{-2}$	$-2.93 \cdot 10^{1}$
1633	1440	$-1.29 \cdot 10^{1}$	$1.81 \cdot 10^{-2}$	$-1.29 \cdot 10^{1}$
3777	1440	$-3.70 \cdot 10^{0}$	$9.90 \cdot 10^{-3}$	$-3.69 \cdot 10^{0}$
8885	1440	$-5.23 \cdot 10^{-1}$	$9.19 \cdot 10^{-3}$	$-5.13 \cdot 10^{-1}$

Nevertheless, this example also reveals the limits of dual-weighted residual error estimators in space for FSI. For small solid displacements the method works excellent, but if the ALE transformation gets too irregular the error estimator tend to overestimate the error and local mesh refinement does not contribute to a smoothing of the ALE transformation.

8.3. Optimal Control of a Nonlinear Fluid-Structure Interaction Problem

We present in the following two numerical examples with Neumann boundary control on either the fluid (in Section 8.3.2) or the solid boundary (in Section 8.3.1) for a nonlinear fluid-structure interaction model. Furthermore, we state the results of an optimal control problem with distributed control parameter in Section 8.3.3. The diversity of the different numerical examples make clear the large variety of applications for optimal control constrained by fluid-structure interaction.

8.3.1. Optimal Control of Flow in an Elastic Channel

We consider once again the configuration with fluid flow through a channel with elastic walls. In comparison to the example in Section 8.1, we now use the nonlinear fluid-structure interaction model. The geometry in Figure 8.18 is identical to the geometry in Section 8.1. We only renamed the right solid boundary to Γ_{q}.

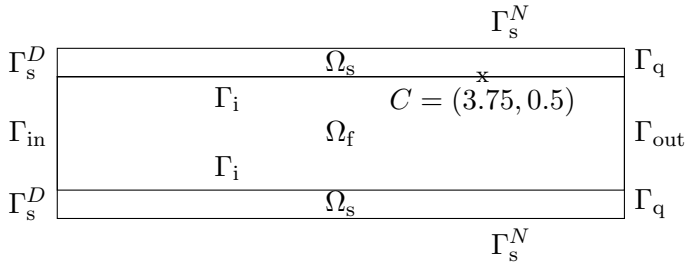

Figure 8.18.: Geometry for fluid flow through an elastic channel

As already stated in Section 4.3, to model arterial flow it is necessary to incorporate the whole arterial network. Very often reduced 0D and 1D models are used. We refer to [124] for an overview on reduced models and the different techniques available to couple 3D models with the reduced models at the outflow boundary. The coupling of fluid and elasticity equations in FSI enables the propagation of pressure waves in a channel. Therefore, the outflow condition has to be chosen carefully to let the pressure wave leave the system, see [145, 63]. As the outflow boundary is generated by artificially truncation of the arterial system, the outflow conditions cannot be physically deduced. This motivates to control the forces on the solid outflow boundary Γ_{q} and the mean pressure on Γ_{out}, instead of adapting parameters in a reduced model.

In this section, we only control the time dependent tangential force

$$(\Sigma_{\mathrm{s}} n)^T e_2 = q \quad \text{ on } I \times \Gamma_{\mathrm{q}}$$

on the solid boundary Γ_{q} with $q \in L^2(I)$. For the fluid outflow we use again the do-nothing condition. The elastic wall is fixed in x-direction and free to move in y-direction at the control boundary Γ_{q}. The control should be chosen in such a way, that the energy induced in the

structure due to the pulsating inflow can leave the domain. Therefore, we minimize the kinetic and elastic energy in the vessel wall

$$\mathcal{J}_E(\mathbf{u}) = \frac{1}{2}\int_I (\Sigma_\mathrm{s}, E_\mathrm{s})_\mathrm{s}\ \mathrm{d}t + \frac{1}{2}\int_I \|\rho_\mathrm{s} v\|^2_{L^2(\Omega_\mathrm{s})}\ \mathrm{d}t.$$

To have a well-posed optimal control problem we add a Tikhonov regularization and minimize the functional

$$\mathcal{J}(q, \mathbf{u}) = \mathcal{J}_E(\mathbf{u}) + \frac{\alpha}{2}\|q\|^2_{L^2(I)}$$

subject to the fluid-structure interaction model presented in Section 4.4.

We want to emphasize that the example is only motivated by hemodynamics. The configuration is not at all close to a real arteria. It is mainly thought to demonstrate various possibilities in this area. In addition, we use the configuration to test the dual-weighted residual error estimator for optimal control of the nonlinear fluid-structure interaction problem.

The inflow profile is identical to (8.1) in Section 8.1. We have again a parabolic inflow profile with $v_\mathrm{in}(t)$ given in Figure 8.2 and use a constant flow as initial condition, as in Section 8.1. Due to the symmetry of the configuration we assume the control on the upper and lower control boundary to be identical with opposing sign. Furthermore, we use again similar parameters as in Section 8.1, see Table 8.1. Only for the fluid and solid density we use now the values $\rho_\mathrm{s} := 1.2\frac{\mathrm{g}}{\mathrm{cm}^3}$ and $\rho_\mathrm{f} := 1.0\frac{\mathrm{g}}{\mathrm{cm}^3}$.

We apply the fractional-step theta time-stepping scheme and use bi-quadratic finite elements as well as a piecewise linear control space as suggested in Chapter 5.

Discussion of the Optimal Solution

The optimal control problem is solved on different space, time, and control grids by the LBFGS algorithm 6.1 presented in Chapter 6. As we can see in Table 8.14, the error in the cost functional reduces uniformly after refinement. As reference value $\mathcal{J}_\mathrm{ref}$ we thereby used the solution on the finest grid with $N = 6305$ degrees of freedom, $M = 2880$ time steps and $\dim(Q_d) = 2883$.

Table 8.14.: Value and error of the optimized cost functional $\mathcal{J}(q_\sigma, u_\sigma)$ for different discretization levels for flow in an elastic channel

N	M	$\dim(Q_d)$	$\mathcal{J}(u_\sigma)$	$\mathcal{J}_\mathrm{ref} - \mathcal{J}(u_\sigma)$
117	360	363	$1.082 \cdot 10^{-1}$	$1.125 \cdot 10^{-2}$
425	720	723	$1.178 \cdot 10^{-1}$	$1.654 \cdot 10^{-3}$
1617	1440	1443	$1.190 \cdot 10^{-1}$	$4.403 \cdot 10^{-4}$
6305	2880	2883	$1.194 \cdot 10^{-1}$	-

Comparing the energy on the finest grid in the solid with and without control, we can see that we were able to reduce the elastic and kinetic energy from $\mathcal{J}_E(\mathbf{u}) = 3.8 \cdot 10^{-1}$ to $\mathcal{J}_E(\mathbf{u}) = 8.87 \cdot 10^{-2}$. If we take a closer look at the energy values plotted over time in Figure 8.19, only the elastic energy at the beginning could not be reduced.

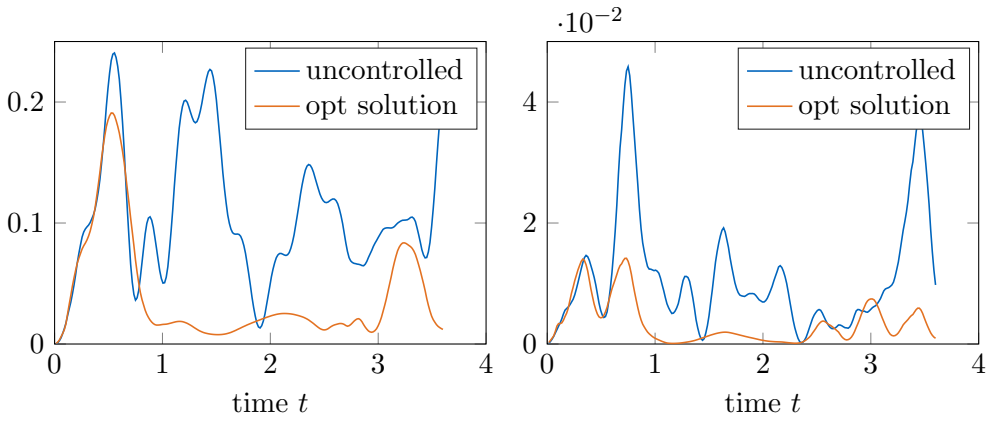

Figure 8.19.: Elastic energy (left) and kinematic energy (right) in the solid plotted over time in the uncontrolled and controlled configuration for flow through an elastic channel.

If we regard the displacement at the point C plotted over time in Figure 8.20, we can see that the oscillations in the solid could be significantly reduced by the optimization algorithm. The maximal values stay constant over time and do not increase anymore after every inflow pulse.

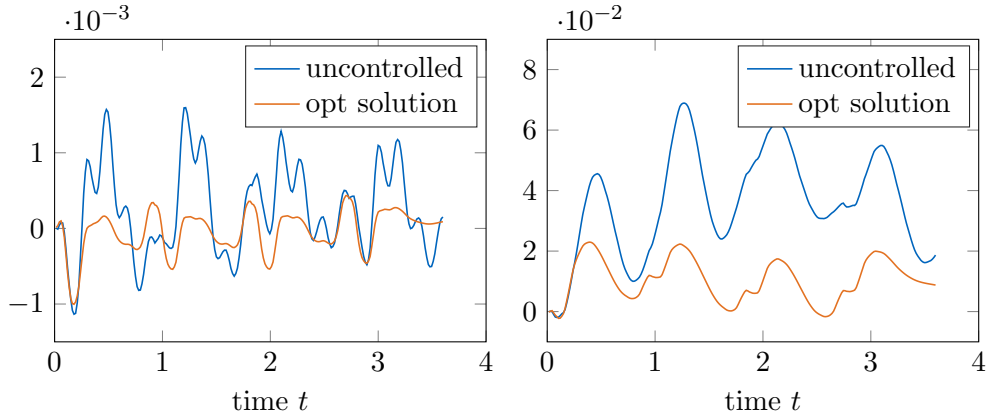

Figure 8.20.: Displacement at the point C in x-direction (left) and y-direction (right) plotted over time in the uncontrolled and controlled configuration for flow through an elastic channel.

The resulting optimal control on the finest grid level plotted in Figure 8.21 is highly distributed in time. This justifies the use of the high-dimensional control space here. If we take a closer look at the computed control, we see a significant decent in the control variable a few seconds before the inflow profile rises. This decent causes the detour in the displacement profile of the point C in Figure 8.20. The control induces a movement of the elastic layer, which counteracts the deformation due to the inflow.

To get a more realistic model, unphysical high values of the boundary forces can be avoided by the enforcement of control constraints. For vascular models it would be necessary to control the fluid outflow condition, too. The do-nothing outflow condition sets the outflow pressure to be zero. Thereby, the pressure wave is always reflected.

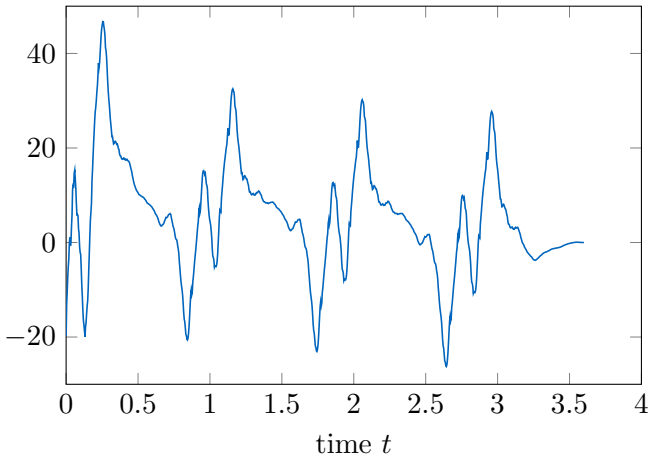

Figure 8.21.: Optimal control q plotted over time for optimal control of flow through an elastic channel.

Adaptive Refinement Using A Posteriori Estimator

Now we would like to test if the dual-weighted residual error estimator works well for optimal control of the nonlinear fluid-structure interaction problem. We only consider the time and control discretization error estimator and compute in the following on a spatial grid with $N = 1617$ degrees of freedom. We refine either globally in time or use the a posteriori error estimator to refine locally the time and control time grid. The errors are plotted over the degrees of freedom of time and control discretization in Figure 8.22. Thereby, the discretization error of the cost functional reduces much faster using the adaptive algorithm.

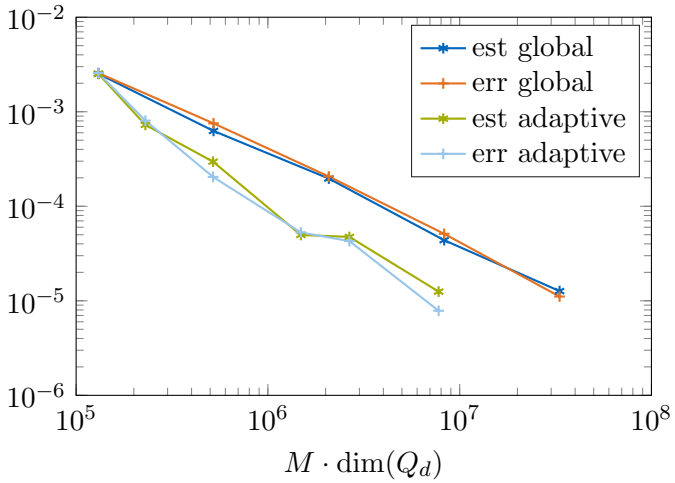

Figure 8.22.: Error in cost functional for optimal control of flow through an elastic channel plotted over degrees of freedom of the time and control discretization spaces $(M \cdot \dim(Q_d))$

We would like to emphasize that the number of degrees of freedom is here not directly con-

nected to the computational cost. As we have seen in Section 8.1, the number of optimization loops needed in the LBFGS algorithm is at least for the presented linear example independent of the control dimension.

8.3.2. Optimal Control of the FSI-2 Benchmark Example

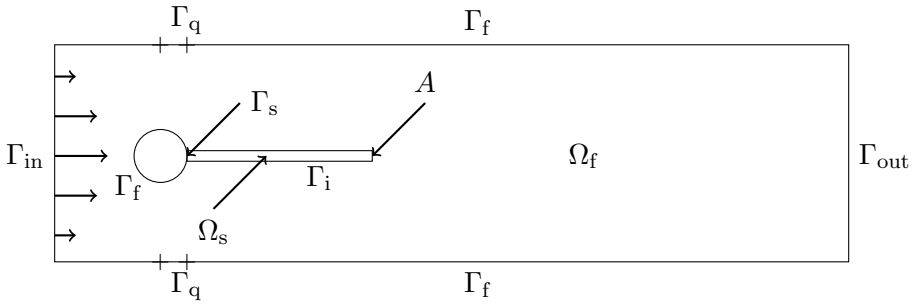

Figure 8.23.: Geometry for flow around cylinder with elastic beam and additional control boundary Γ_q.

We consider again the FSI-2 benchmark configuration presented in Section 8.2.1. We adapt a Neumann boundary control problem suggested by Becker in [15] for the Navier-Stokes benchmark, on the FSI-2 benchmark. We add an additional boundary Γ_q on the upper and lower channel wall and control the mean pressure in the outflow condition on Γ_q, see Figure 8.23. Thereby, we are able to control indirectly the in- and outflow at the boundary Γ_q. We use again the parameters from the FSI-benchmark given in Table 8.9 and the smoothly increasing parabolic inflow profile (8.2) with

$$v_{in}(t) = \begin{cases} 0.5 - 0.5\cos(\frac{\pi t}{2}) & t \leq 2 \\ 1 & t > 2. \end{cases}$$

The simulation is computed on the time interval $I = (0; 15)$.

Optimal Control Configuration

We describe on the boundary Γ_q the outflow condition

$$(\nabla v + p\,\mathrm{Id})n = qn \quad \text{on } \Gamma_q \times I,$$

whereby we control the mean pressure q here. For $q = 0$, the standard do-nothing outflow condition is enforced at Γ_q. The FSI-2 system is known to be highly dynamical. The beam starts to oscillate and sincerely influences the flow behavior. To be able to influence the system, we choose a time depend control variable

$$q \in Q := L^2((5; T)) \times L^2((5; T)).$$

We only control the system after the time-point $t > 5$ as the system is very sensitive to numerical errors in the beginning. Thereby, we can increase the inflow profile smoothly and after a stable flow developed, we start to control the system.

The control is chosen in such a way that the oscillations of the beam are reduced after the time point $t > 8$. Hence we minimize the tracking type functional

$$\min_{q \in Q} \mathcal{J}(q, \mathbf{u}) := \frac{1}{2} \int_8^T \|u\|^2_{L^2(\Omega_s)} \, \mathrm{d}t + \frac{\alpha}{2} \|q\|^2_Q \tag{8.9}$$

with L^2-Tikhonov regularization of the control variable.

Numerical Example

The FSI-2 benchmark example is highly dynamical, therefore we need to compute with very small time steps even for the fractional-step theta time-stepping scheme. In addition to resolve the forces at the interface Γ_i, a very fine mesh has to be used. To improve the accuracy of the simulation, we apply the DWR estimator presented in Theorem 7.1 on the cost functional $\mathcal{J}(q, \mathbf{u})$. In addition, we equilibrate the control error estimator η_q, the time discretization error estimator η_k and the spatial discretization error η_h. We have seen in Section 8.2.5 that thereby, we can avoid ridiculous refinement and we can reduce the computational cost severely.

The control space is discretized by the space of continuous and piecewise linear functions. We use the fractional-step theta time-stepping scheme to discretize the optimal control problem in time and bi-linear finite elements in space. For the Tikhonov regularization the parameter $\alpha = 2 \cdot 10^{-11}$ is chosen. The LBFGS algorithm terminates if the gradient could be reduced by a factor 10^{-3}.

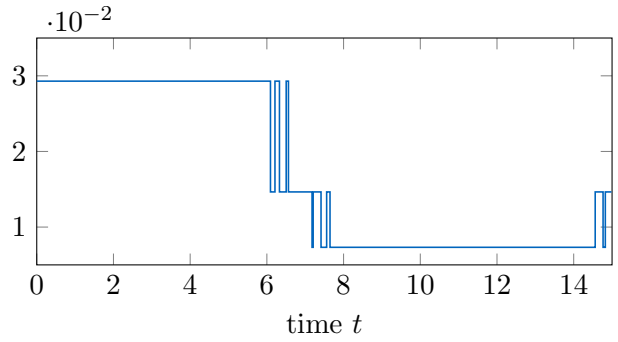

Figure 8.24.: Time step size k_m plotted over time t after 2 adaptive refinement cycles for optimal control of the FSI-2 benchmark.

After two refinement cycles the control space is discretized with $\dim(Q_d) = 258$ degrees of freedom. Furthermore we have $M = 8500$ time steps and a spatial discretization with $N = 3895$ degrees of freedom. We plotted in Figure 8.24 the resulting time step size and in Figure 8.25 the refined mesh. As we only have observation of the cost functional in the time interval $(8, 15)$, the adaptive algorithm mainly refines in this interval of interest. Furthermore the moment the algorithm starts to choose smaller time step sizes, correlates with the moment the oscillations in the flag set in.

Moreover the algorithm refines very locally around the flag and the circle. Only thereby the pressure field at the circle boundary and interface can be computed accurately. As the

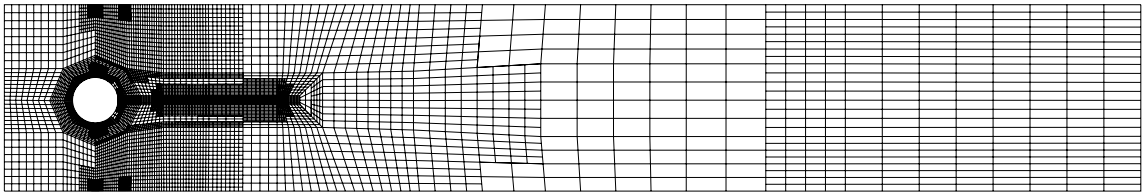

Figure 8.25.: Adaptively refined mesh for optimal control of the FSI-2 benchmark.

boundary forces at the interface Γ_i severely influences the behavior of the solid motion, it is reasonable that the adaptive algorithm refines here. In addition, the algorithm refines very locally around the control boundary Γ_q. As we enforce a Neumann boundary condition on Γ_q and zero Dirichlet boundary conditions at Γ_f, we have a "corner" of angle 180°, between a Neumann and Dirichlet boundary condition. Standard elliptic PDE theory predicts already for angles larger then 90° corner singularities. These singularities in the fluid velocity field and pressure can only can be resolved by local refinement, hence the adaptive algorithm refines very locally here.

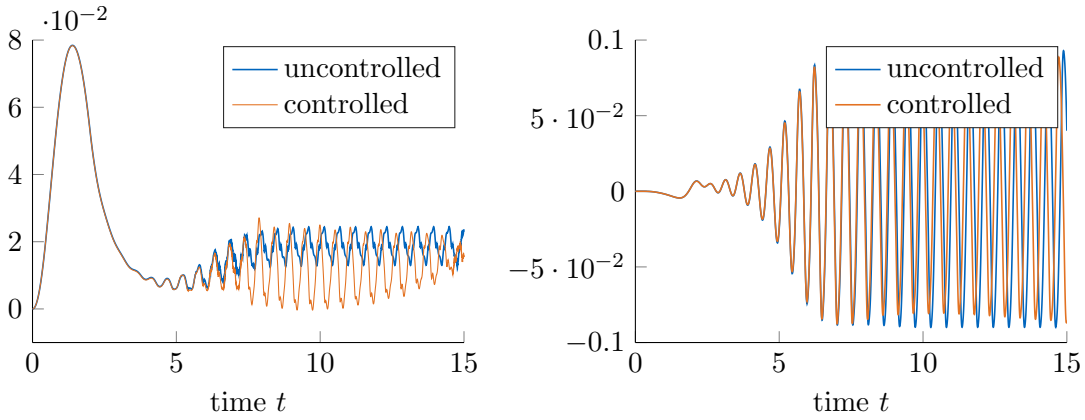

Figure 8.26.: Flux at Γ_q (left) and displacement in y-direction at the tip of the flag A(right) plotted over time for optimal control of the FSI-2 benchmark. Plotted is the controlled vs uncontrolled solution.

We start the optimization algorithm with control $q(t) = 0$. Thereby, we have in the uncontrolled configurations a do-nothing outflow. The fluid can leave and enter the domain through the control boundary Γ_q. Despite the slight modification of the additional outflow boundary condition the flag starts to oscillate as for the standard FSI-2 benchmark. We plotted the displacement u at the tip of the flag A in y-direction in Figure 8.26 for the uncontrolled configuration on the two times refined mesh. If we compare the solution with the displacement in Figure 8.8, we can see that the amplitude and frequency correlate with the standard benchmark.

We plotted in Figure 8.26 in addition the flux $\int_{\Gamma_q} v_{kh} n \, dx$ leaving the lower boundary Γ_q. While the inflow velocity v_{in} rises, the outflow is very strong, but after some time the flux reduces and the fluid flows across the boundary Γ_q. The displacements of the flag are very

large in the FSI-2 benchmark configuration. The moment the tip of the flag reaches its minimal value, the gap between wall and elastic beam gets very thin. This enforces the fluid to leave the domain through the lower control boundary Γ_{q} and explains the oscillations in the flux profile.

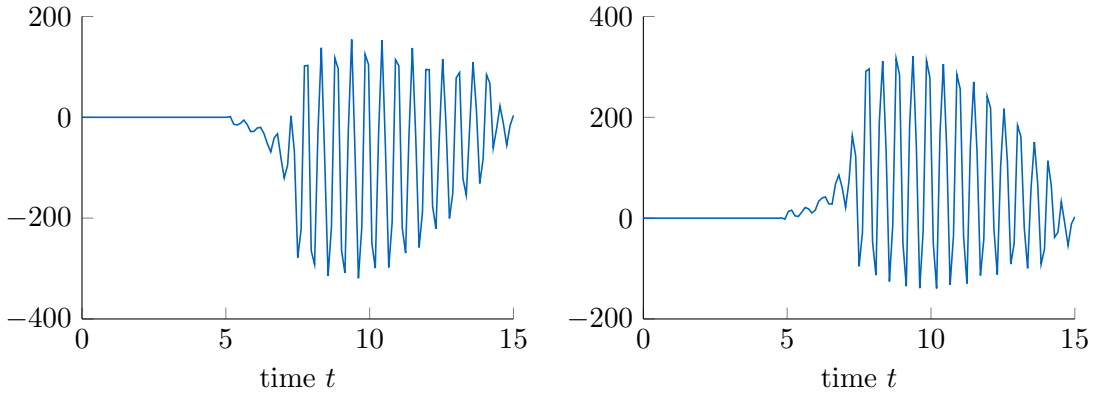

Figure 8.27.: Optimal control q at the upper (left) and lower (right) control boundary Γ_{q} plotted over time for optimal control of the FSI-2 benchmark.

In Figure 8.27 the optimal control is plotted over time at the upper and lower control boundary. The control variable is still ragged as the control error estimator is much smaller, then the time and spatial discretization error. Hence, the adaptive algorithm only refines in space and time. Due to the almost symmetrical configuration the control on upper and lower control boundary Γ_{q} have the same values with opponent sign.

The optimal control sincerely influences the flow behavior at the boundary Γ_{q}, as we can see in Figure 8.26. The amplitude in the flux oscillations increases and the flux reduces at certain time-points so heavily such that we almost have inflow. In addition, we are able to reduce the amplitude of the oscillating flag slightly. The value of the cost functional can be reduced from $\mathcal{J}(q_\sigma, u_\sigma) = 2.63 \cdot 10^{-5}$ to $\mathcal{J}(\bar{q}_\sigma, \bar{u}_\sigma) = 2.48 \cdot 10^{-5}$. To get an impression of the flow behavior we plotted in Figure 8.28 the velocity field in x-direction at the time points $t = 14.74$, $t = 14.90$ and $t = 15.00$. Here we can see again the large deformation of the solid beam, which influences the flow behavior sincerely. The outflow on the control boundary Γ_{q} is relatively small with respect to the outflow at the outflow boundary Γ_{out}. Hence the flow behavior is very similar to the solution of the FSI-2 benchmark.

8.3.3. Optimal Control of the Mesh Motion Equation

In the FSI model in Problem 4.5 the solution of the velocity and pressure variable computed on a reference domain, coincides according to Lemma 4.1 with the solution on the moving fluid domain $\check{\Omega}_{\mathrm{f}}$, if the ALE mapping is smooth enough. The moment where the solid deformation becomes too large all mesh motion equations suggested in literature lack the necessary regularity. In numerical experiments the loss of regularity leads to a reduced convergence of the Newton solver until the solution algorithm finally breaks down. In these cases re-meshing is necessary or the Eulerian approach as in [68, 67, 126] has to be used. But already for

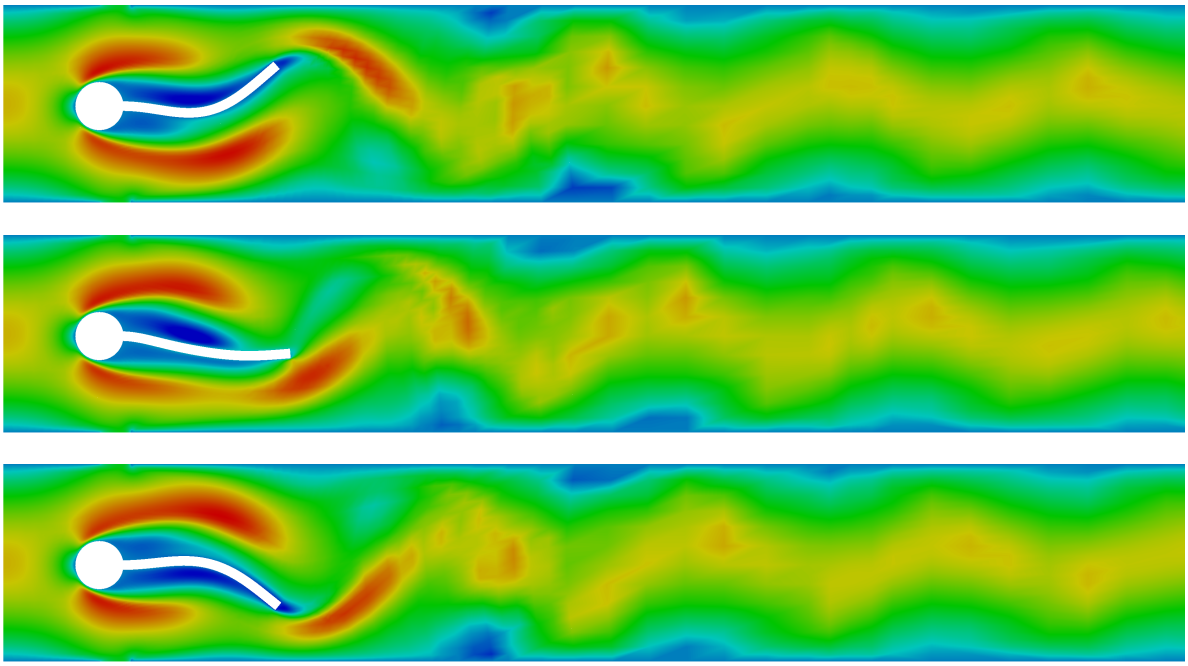

Figure 8.28.: Fluid velocity field v in x direction on moving domain for optimal control of the FSI-2 benchmark at the time points $t = 14.74$, $t = 14.90$ and $t = 15.00$.

moderate solid deformations, a non regular transformation can reduce the convergence rate of the discretization error.

Instead of solving the Navier-Stokes equations on the moving domain $\check{\Omega}_\mathrm{f}$, we transform the fluid equations via ALE mapping $T_\mathrm{f} : I \times \Omega_\mathrm{f} \to \check{\Omega}_\mathrm{f}$ on a reference domain Ω_f. There we choose in Section 5.3 a triangulation and define the finite element spaces, whereby the ALE transformation is chosen in the same space as the solution variable. This approach is very similar to isoparametric elements, where we would like to approximate curved boundaries more accurately. To do so the basic finite element space V_h is defined on the base polyhedral domain Ω. In addition we define a one-to-one continuous mapping $T_h : \Omega \to \mathbb{R}^n$ with $T_h \in V_h$ and $\check{\Omega} := T_h(\Omega)$. Now the discrete solution approximating the physical quantity we are interested in, is an element of the isoparametric equivalent finite element space

$$\check{V}_h := \left\{ w(T_h^{-1}(x)) \;\middle|\; x \in T_h(\Omega), w \in V_h \right\}. \tag{8.10}$$

In the following, we want to examine more closely what condition on the transformation T_h are necessary such that the space $\check{V}_h$ enables good approximations of functions $\check{w} \in W^{2,p}(\check{\Omega})$. If the discretized ALE mapping fulfills similar assumptions, then we can expect the velocity $\check{v}$ and pressure $\check{p}$ living on the moving domain $\check{\Omega}_\mathrm{f}$ to be approximated accurately, too. In addition, we motivate, why these assumptions are connected to the regularity of the transformed mesh. Then we define an optimal control problem to automatically improve the smoothness of the discrete ALE transformation.

Approximation Theory for Isoparametric-Equivalent Spaces

A necessary step to prove approximation results for a function $\check{w} \in W^{2,p}(\check{\Omega})$ in the isoparametric equivalent finite element space $\check{V}_h$ is the availability of interpolation results. We state in the following such an interpolation result taken from [36] (Theorem 4.4.20) for standard finite elements. Then we extend the result to the interpolation operator in the isoparametric space, whereby we follow the proof of Theorem 1 in [46].

Theorem 8.1 (Theorem 4.4.20 in [36])**.** *Let $\{\mathcal{T}_h\}$, $0 < h < 1$ be a non-degenerate family of subdivisions of a polyhedral domain $\Omega \subset \mathbb{R}^d$ with B_K the circumradius and*

$$\operatorname{diam}(B_K) \geq \rho \operatorname{diam}(K) \quad \forall K \in \mathcal{T}_h.$$

Let V_h be a finite element space on $\mathcal{T}_h$ with degree 1. Then there exists a positive constant C depending on ρ and on the reference element such that for $0 \leq l \leq 1$

$$\Big(\sum_{K \in \mathcal{T}_h} \|w - I^h w\|_{W^{l,p}(K)} \Big)^{1/p} \leq C h^{2-l} |w|_{W^{2,p}(\Omega)}. \tag{8.11}$$

for all $w \in W^{2,p}(\Omega)$ and I^h the interpolation operator into V_h.

For a regular family of meshes Theorem 8.1 guarantees, that with refinement of the mesh the interpolation error reduces by the factor h^{2-l}, $0 \leq l \leq 1$. In the following we define assumptions on the transformation $T_h \in V_h$, which guarantee the same order of convergence for the interpolation operator in the isoparametric finite element space $\check{V}_h$.

Let Ω be a polyhedral reference domain, where we define the finite element space V_h. Furthermore let $T_h : \Omega \to \mathbb{R}^d$ be a piecewise polynomial mapping in V_h and we define the physical domain $\check{\Omega} := T_h(\Omega)$. If the transformation fulfills the following bounds independent of h

$$\max_{K \in \mathcal{T}_h} \|J_h\|_{L^\infty(K)} \leq C \quad \text{and} \quad \max_{K \in \mathcal{T}_h} \|(\nabla T_h)^{-1}\|_{L^\infty(K)} \leq C, \tag{8.12}$$

$$\max_{K \in \mathcal{T}_h} \|J_h^{-1}\|_{L^\infty(K)} \leq C \quad \text{and} \quad \max_{K \in \mathcal{T}_h} \|\nabla^2 T_h\|_{L^\infty(K)} \leq C, \tag{8.13}$$

then the interpolation $\check{I}^h$ of a function $\check{w} \in W^{2,p}(\check{\Omega})$ in the isoparametric equivalent finite element space $\check{V}_h$ satisfies:

Theorem 8.2. *Let $\{\mathcal{T}_h\}$, $0 < h < 1$ be a non-degenerate family of subdivisions of a polyhedral domain $\Omega \subset \mathbb{R}^d$ with*

$$\operatorname{diam}(B_K) \geq \rho \operatorname{diam}(K) \quad \forall K \in \mathcal{T}_h.$$

Let V_h be a finite element space on $\mathcal{T}_h$ with degree 1. Suppose the piecewise polynomial mapping $T_h \in V_h$ exists which satisfies properties (8.13) and (8.12) above. Then there exists a positive constant C depending on ρ and on the reference element such that for $0 \leq l \leq 1$

$$\Big(\sum_{K \in \mathcal{T}_h} \|\check{w} - \check{I}^h \check{w}\|_{W^{l,p}(T_h(K))} \Big)^{1/p} \leq C h^{2-l} |w|_{W^{2,p}(\check{\Omega})} \tag{8.14}$$

for all $\check{w} \in W^{2,p}(\check{\Omega})$. Here $\check{I}^h \check{w}$ denotes the isoparametric interpolant defined by $\check{I}^h \check{w}(T_h(x)) = I^h w(x)$ for all $x \in \Omega$, where $w(x) := \check{w}(T_h(x))$ for all $x \in \Omega$ and I^h is the global interpolant for the base finite element space V_h.

Proof. We follow the proof of Theorem 1 in [46]. Any function $\check{w} \in W^{2,p}(T_h(K))$ can be written as $\check{w} = w \circ T_h$ with $w \in W^{2,p}(K)$. Thus by Lemma 3 in [46] there exists a constant C such that:

$$\|\check{w} - \check{I}^h\check{w}\|^p_{L^p(T_h(K))} \leq C\|J_h\|_{L^\infty(K)}\|w - I^h w\|^p_{L^p(K)},$$
$$|\check{w} - \check{I}^h\check{w}|^p_{W^{1,p}(T_h(K))} \leq C\|J_h\|_{L^\infty(K)}|w - I^h w|^p_{W^{1,p}(K)}\|(\nabla T_h)^{-1}\|^p_{L^\infty(K)}.$$

If we apply the given bounds in (8.12) for the transformation T_h and sum over all cells $K \in \mathcal{T}_h$ we obtain

$$\Big(\sum_{K\in\mathcal{T}_h}\|\check{w} - \check{I}^h\check{w}\|_{L^p(T_h(K))}\Big)^{1/p} \leq C\Big(\sum_{K\in\mathcal{T}_h}\|w - I^h w\|_{L^p(K)}\Big)^{1/p}, \tag{8.15}$$
$$\Big(\sum_{K\in\mathcal{T}_h}|\check{w} - \check{I}^h\check{w}|_{W^{1,p}(T_h(K))}\Big)^{1/p} \leq C\Big(\sum_{K\in\mathcal{T}_h}|w - I^h w|_{W^{1,p}(K)}\Big)^{1/p}. \tag{8.16}$$

The estimates given for the interpolation operator I in Theorem 8.1 lead to

$$\Big(\sum_{K\in\mathcal{T}_h}\|w - I^h w\|_{L^p(T_h(K))}\Big)^{1/p} \leq Ch^2|w|_{W^{2,p}(\Omega)},$$
$$\Big(\sum_{K\in\mathcal{T}_h}\|w - I^h w\|_{W^{1,p}(T_h(K))}\Big)^{1/p} \leq Ch|w|_{W^{2,p}(\Omega)}.$$

We apply now Lemma 3 in [46] once again on $|w|_{W^{2,p}(K)}$ and take the bounds in (8.13) on the transformation T_h into account, then we obtain

$$\begin{aligned}|w|^p_{W^{2,p}(K)} &\leq C\|J_h^{-1}\|_{L^\infty(K)}\Big(|\check{w}|^p_{W^{1,p}(T_h(K))}|\nabla^2 T_h|^p_{L^\infty(K)} \quad + |\check{w}|^p_{W^{2,p}(T_h(K))}\|\nabla T_h\|^{2p}_{L^\infty(K)}\Big)\\ &\leq C|\check{w}|^p_{W^{2,p}(T_h(K))}.\end{aligned}$$

If we combine the estimate on $|w|_{W^{2,p}(K)}$ with (8.15) and (8.16) , we obtain the estimate (8.14) in Theorem 8.2.

□

Optimal Control Problem

These approximation results in Theorem 8.2 are valid in the same way for our fluid sub-problem in the space-time discretized fluid-structue interaction model in Problem 5.5. As we define the finite element spaces for the fluid and pressure variable on the reference domain, the physically relevant discrete pressure and velocity on the moving domain are now elements of such an isoparametric equivalent finite element space. Hence, we can expect optimal convergence results for the velocity $\check{v}_{kh} := v_{kh}(T^{-1}_{\mathrm{f},kh}(\check{x},t))$ and the pressure $\check{p}_{kh} := p_{kh}(T^{-1}_{\mathrm{f},kh}(\check{x},t))$ with $\check{x} \in \check{\Omega}_\mathrm{f}(t)$, if the discrete ALE transformation $T_{\mathrm{f},kh}$ fulfills in every time step the estimates (8.13) and (8.12). This implies especially that the determinant of the deformation gradient $\|J_{kh}\|_{L^\infty(\Omega_\mathrm{f})} < C$ and $\|J^{-1}_{kh}\|_{L^\infty(\Omega_\mathrm{f})} < C$ is bounded independent of h. Furthermore for every family of triangulations $\{\mathcal{T}_h\}$ we need for piecewise bi-linear finite elements for all $K \in \mathcal{T}_h$

$$\|\nabla^2 T_{\mathrm{f},kh}\|_{W^\infty(K)} \leq C \quad \text{and} \quad \|\nabla T^{-1}_{\mathrm{f},kh}\|_{L^\infty(K)} \leq C.$$

If we compute, as suggested in Section 4.1.3, the ALE transformation by solving a Laplace or elasticy problem on a nonconvex domain, the ALE mapping will in general not fulfill this regularity for a sequence of globally refined meshes. Hence, the constant C is going to rise after refinement. This motivates to choose the elasticity module in the mesh motion equation problem-dependent, such that the constant C keeps small even on fine grids. To choose the parameter automatically, we can solve on every mesh an optimal control problem.

We are going to construct a functional measuring the smoothness of the ALE transformation, which can be easily computed and used for an optimal control approach. The functional will consist of geometric mesh properties as the volume change of the cells due to the transformation and the change of the angle of the elements. Both values are elementary properties of the Jacobian matrix $\nabla T_{\mathrm{f},kh}$ we would like to bound. Moreover, the improvement of the mesh regularity of the transformed mesh is quite natural. The convergence rate of velocity $v_{kh}(T_{\mathrm{f},kh}^{-1}(\check{x},t))$ and the pressure $p_{kh}(T_{\mathrm{f},kh}^{-1}(\check{x},t))$ in the isoparametric equivalent finite element spaces will be connected to the convergence rate, which we would obtain, if we defined finite elements on the transformed grid on $\check{\Omega}$. Standard convergence results demand a regular mesh with angles $0^\circ \ll \beta \ll 180^\circ$. To avoid the cells to deteriorate, the determinant of the deformation gradient $J_{kh} = \det(\nabla T_{\mathrm{f},kh})$ has in addition to be bounded.

Functional Measuring Change of Angle The transformation gradient $\nabla T_{\mathrm{f},kh}$ transforms the unit vectors in x- and y-direction e_1 and e_2 to $\tilde{e}_1 = \nabla T_{\mathrm{f},kh} e_1$ and $\tilde{e}_2 = \nabla T_{\mathrm{f},kh} e_2$. Therefore, the angle in the transformed cells change by a value of $\beta = \arccos(\frac{\tilde{e}_1 \cdot \tilde{e}_2}{\|\tilde{e}_x\|\|\tilde{e}_2\|}) - 90^\circ$. Assuming the reference grid to be regular with angles around 90°, the change of the angle β should be close to zero. This motivates to enforce $-1 \ll \frac{\tilde{e}_1 \cdot \tilde{e}_2}{\|\tilde{e}_1\|\|\tilde{e}_2\|} \ll 1$ by minimization of the functional

$$\mathcal{J}_\beta(u_{kh}) = \int_I \int_{\Omega_\mathrm{f}} \Big(1 - \frac{\tilde{e}_1 \cdot \tilde{e}_2}{\|\tilde{e}_1\|\|\tilde{e}_2\|}\Big)^{-1} + \Big(\frac{\tilde{e}_1 \cdot \tilde{e}_2}{\|\tilde{e}_1\|\|\tilde{e}_2\|} + 1\Big)^{-1} \,\mathrm{d}t\,\mathrm{d}x. \tag{8.17}$$

Functional Measuring Change of Volume To avoid cells to degenerate, the volume change of the cells should be kept as small as possible. The local change of volume is given by the determinant $J_{kh} = \det(\nabla T_{\mathrm{f},kh})$. The approximation theory for isoparametric elements demands in addition $\|J_{kh}\|_{L^\infty(\Omega_\mathrm{f})} \leq C$ and $\|J_{kh}^{-1}\|_{L^\infty(\Omega_\mathrm{f})} \leq C$. Hence, a good idea would be to minimize

$$\mathcal{J}_V(u_{kh}) = \int_I \int_{\Omega_\mathrm{f}} \frac{1}{J_{kh}^p} + J_{kh}^p \,\mathrm{d}t\,\mathrm{d}x. \tag{8.18}$$

The functional $\mathcal{J}_V(u_{kh})$ rises if cells get too large which corresponds to large values of J_{kh}. In addition, if cells get too small J_{kh}^{-1} increases. The parameter p can control how severe extreme values of J_{kh} are penalized. In the following, we will choose $p = 2$.

Both functionals depend on the derivative of the ALE transformation $\nabla T_{\mathrm{f},kh}$ and will rise, if the regularity of the transformation diminishes. In the following we minimize a combined functional by choosing a distributed Young's modulus in the mesh motion equation (4.8). Minimization of the chosen functionals will not enforce the very strict state constraints demanded by the approximation theory. Hence the constant C in Theorem 8.2 will rise with refinement.

Nevertheless the approach will help to keep the h-dependent constant in the estimates small. We consider the optimal control problem:

$$\min_{q_\sigma \in Q_d} \mathcal{J}_{\mathrm{ALE}}(q_\sigma, \mathbf{u}_\sigma) := \gamma_1 \mathcal{J}_V(u_\sigma) + \gamma_2 \mathcal{J}_\beta(u_\sigma) + \frac{\alpha}{2}\|q_\sigma\|_Q \tag{8.19}$$

subject to $\mathbf{u}_\sigma \in X_{kh}$ with $v_\sigma(0) = v_0$, $u_\sigma(0) = u_0$ and

$$a_{kh}(q_\sigma, \mathbf{u}_\sigma)(\boldsymbol{\varphi}) = 0 \quad \forall \boldsymbol{\varphi} \in Y_{kh,\theta}, \tag{8.20}$$

$$q_a \leq q_\sigma \leq q_b. \tag{8.21}$$

We denote with $a_{kh}(q_\sigma, \mathbf{u}_\sigma)(\boldsymbol{\varphi})$ the space-time discretized semi-linear form of the fluid-structure interaction problem given in (5.5), whereby the control parameter q_σ enters the mesh motion equation 4.8 via the first and second Lamé parameter. We have now:

$$\lambda_{\mathrm{m}}(q_\sigma, u_\sigma) := \frac{q_\sigma}{2J_\sigma(1+\nu_{\mathrm{m}})} \quad \text{and} \quad \mu_{\mathrm{m}}(q_\sigma, u_\sigma) := \frac{q_\sigma \nu_{\mathrm{m}}}{J_\sigma(1+\nu_{\mathrm{m}})(1-\nu_{\mathrm{m}})}.$$

Standard regularity results demand Young's Modulus $\mathcal{E}_{\mathrm{m}} \in L^\infty(\Omega_{\mathrm{f}})$. As we now control Young's Modulus, the control bound's have to be finite. To receive a positive definite extension operator, we demand q_a to be a positive value. The parameters γ_1 and γ_2 weight the two functionals.

The control constraints can for example be enforced due to logarithmic penalization terms and a Tikhonov regularization $\|q_\sigma\|_{L^2(\Omega_{\mathrm{f}})}$ as in [143]. During the optimization process the penalization parameter can be decreased such that the control constraints get sharper and sharper. As we are only interested in finding a distribution of Young's modulus which significantly improves the mesh regularity, we keep the penalization parameter constant during the following computations.

This approach taken for itself is very costly, but if we are already solving an optimal control problem with cost functional $\tilde{\mathcal{J}}(\tilde{q}, \mathbf{u})$ and control $\tilde{q} \in \tilde{Q}$, the optimization of the mesh motion can be done simultaneously. We minimize on the discrete level the combined functional $\alpha\tilde{\mathcal{J}}(\tilde{q}_d, \mathbf{u}_{kh}) + \beta\mathcal{J}_{\mathrm{ALE}}(q_d, \mathbf{u}_{kh})$ over the combined control space $\tilde{Q}_d \times Q_d$. Then, we can compute sensitivity information of both optimization problems at once.

Numerical Example

We have seen previously that the ALE mapping reaches its limits in the flapping example in Section 8.2.5. This example is very challenging as the displacement of the beams is quite large and generating a mesh on the underlying geometry is not simple. We have reentering corners at the tip of the beam and, in addition, the beam is very thin. However, we have to demand a very fine grid to resolve the structure displacement. Hence, the generated grid already consist of very anisotropic cells in the region between the beam.

We optimize in the following the mesh regularity of the transformed mesh by minimizing $\mathcal{J}_{\mathrm{ALE}}(u_{kh})$. Numerical experiments have shown that only the minimization of the functional $\mathcal{J}_\beta(u_{kh})$ can cause some cells to collapse and the algorithm breaks down. In contrast if we only minimize the functional $\mathcal{J}_V(u_{kh})$, the angles in the transformed mesh barely improve.

The algorithm is initiated on a coarse grid. Then the cells around the tip of the flag are relatively large and even for large deformations the algorithm does not break down. After refinement, we project the optimal control computed on the coarse grid to the finer grid and can restart the optimization algorithm having a smoothed ALE-transformation at hand. Thereby, we can compute the solution on fine grids, whereas the standard algorithm already would break down.

The configuration is identical to Section 8.2.5. We use again the parabolic pulsating inflow profile with the slower mean inflow profile $0.5v_{\text{in}}(t)$ given in Figure 8.11. The configuration is given in Figure 8.10 and the used material parameters can be found in Table 8.9. For the optimization algorithm we use a regularization parameter $\alpha = 10^{-4}$. The control is chosen in the space of piecewise constant functions, whereby we use the same mesh for the control variable as for the state and adjoint variable. We discretize the FSI problem with a fractional-step theta time-stepping scheme in time and use bi-linear finite elements in space. The presented solutions are computed on a mesh with $N = 5145$ degrees of freedom, $M = 5761$ time steps and $\dim(Q_d) = 4992$. Furthermore the weights $\gamma_1 = 1.0$ and $\gamma_2 = 0.1$ have shown to work well for the considered configuration.

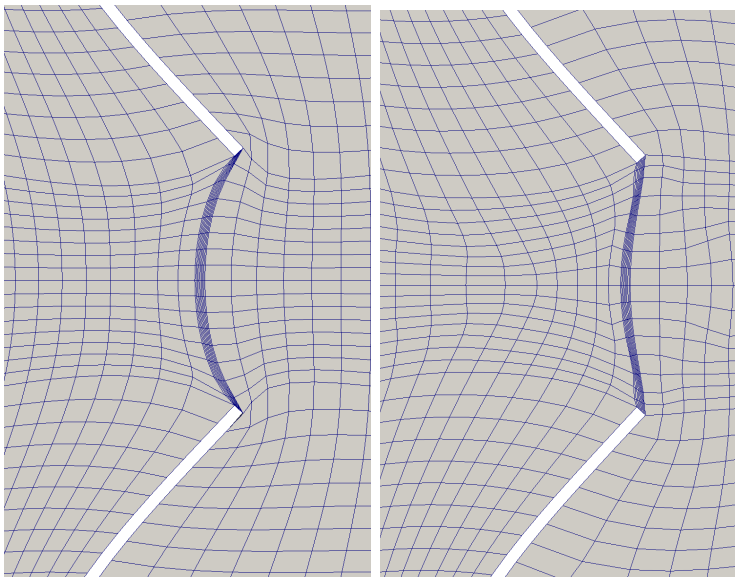

Figure 8.29.: Transformed mesh at time $t = 1$ for the flapping test with constant Young's modulus (left) and distributed modified Young's modulus (right). Zoom into the region around the two flapping beams.

The resulting transformed mesh is plotted in Figure 8.29 in the region of the flapping beam. On the left side, we can see the transformed grid using a constant Young‘s modulus. Especially the cells at the tip of the flag are extremely sheared. In comparison, on the right side, the grid generated using the optimal distributed Young's modulus is very smooth and regular.

In Figure 8.30 the minimal angles of cells on the transformed grid are given. To have a better visualization we plotted the values on the reference domain. As we can see in Figure 8.30, the number of cells with an extreme minimal angle in the transformed grid could be reduced.

There are no cells around the tip of the flag anymore with values in the dark area of angles between 0° and 20°. The minimum angle in the mesh rises from 7.67° to 35.7°.

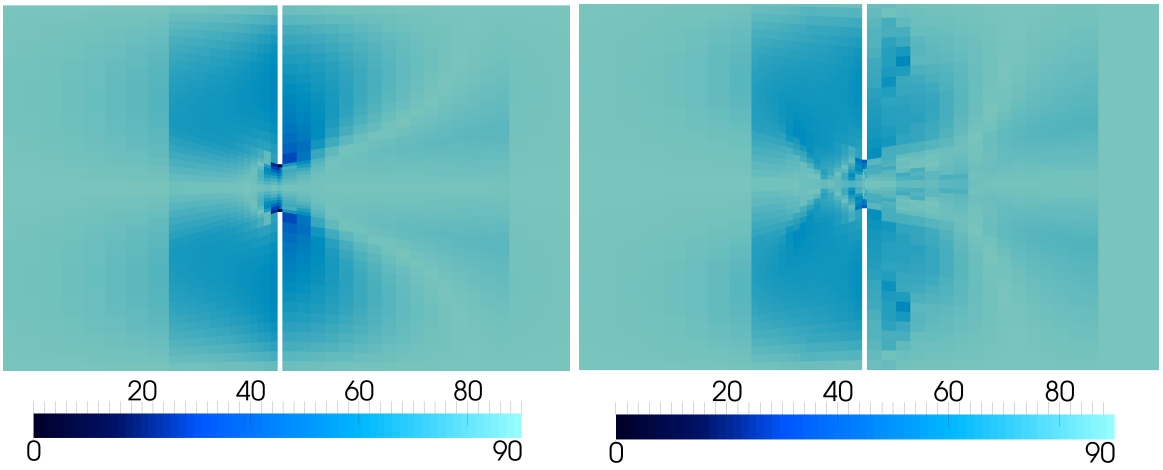

Figure 8.30.: Minimal angle in the cells of the transformed mesh for the flapping test at time $t = 1$ with constant Young's modulus (left) and distributed modified Young's modulus (right) plotted on the reference domain. Zoom into the region around the two flapping beams.

In Figure 8.31 we plotted the ratio between the area of each cell on the transformed and on the initial mesh defined on the reference domain. After optimization the volume change is not localized anymore to the areas around the tip of the flag. As we can see in Figure 8.31, the number of cells with extreme volume change are reduced whereas the number of cells with medium volume change rises. As the overall area has to be conserved the large transformation of the beams has to result in a transformation of the cells. Hence, the optimization algorithm just distributes the deformation to more cells.

The algorithm achieves this optimal mesh motion due to an increase of the stiffness of the cells around the flag (red and green area) as we can see in Figure 8.32. At the same time the algorithm also decreases the stiffness in large areas of the fluid domain (dark blue area).

In Section 8.2.5, we used a dual-weighted residual error estimator to refine locally. Thereby, we could observe that the effectivity indices I_{eff} are not good if the mesh motion is not smooth. With an approximation of the optimal distributed Young's modulus plotted in Figure 8.32, we were able to improve the effectivity indices sincerly. Hence, it could be reasonable to combine the dual-weighted residual error estimator with the mesh smoothing optimization algorithm. But estimation of the discretization error with respect to the mesh regularity functional $\mathcal{J}_{\text{ALE}}(q_\sigma, u_\sigma)$, which is only defined on the discrete level, is not reasonable. Hence we have to solve an additional adjoint equation to estimate the error with respect to a new functional of interest. We refer to [113] for a detailed description of such a posteriori error estimators, where one estimates the discretization error with respect to a functional of interest not coinciding with the cost functional.

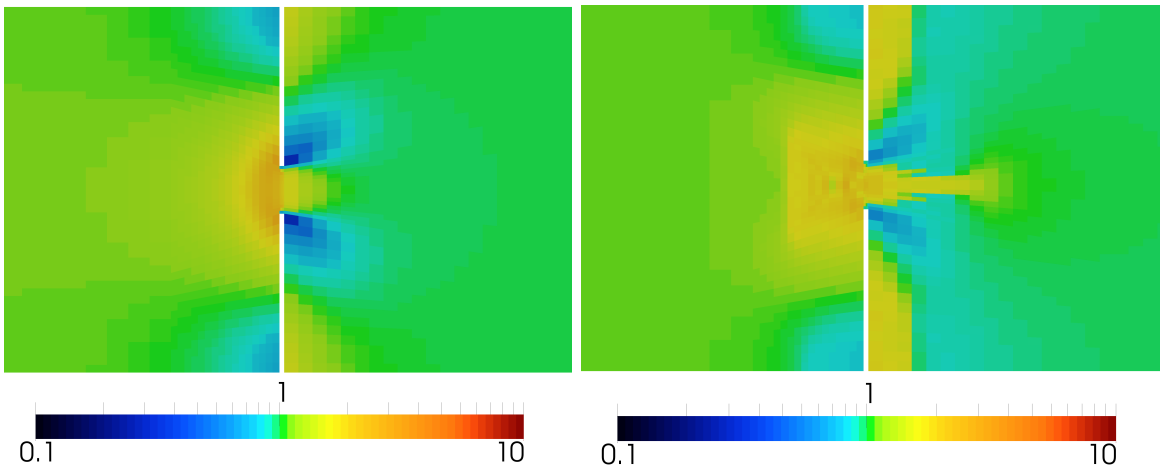

Figure 8.31.: Change of area of the transformed cells for the flapping test at time $t = 1$ with constant Young's modulus (left) and distributed modified Young's modulus(right) plotted on the reference domain. Zoom into the region around the two flapping beams.

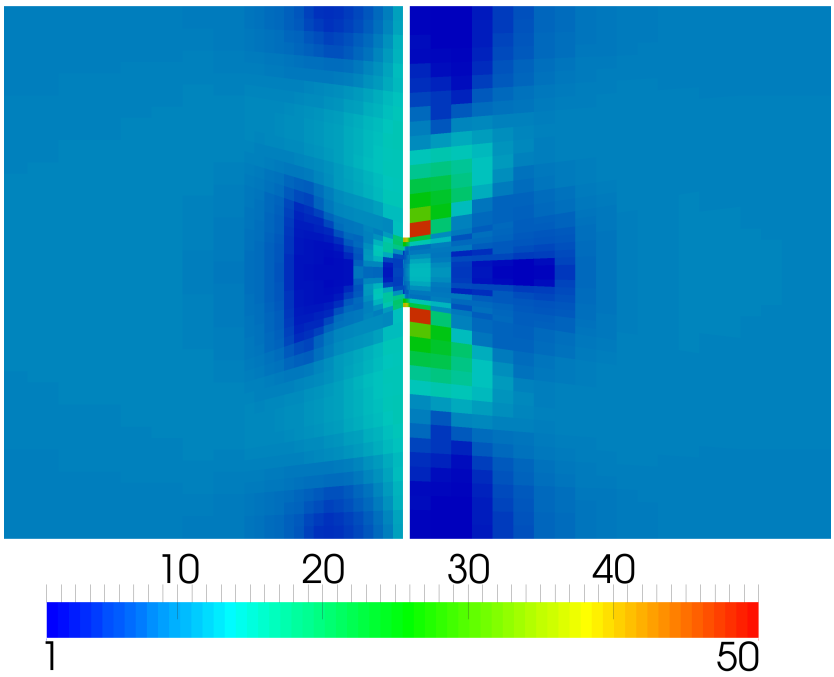

Figure 8.32.: Optimal distributed Young's Modulus for the flapping test at time $t = 1$ plotted on the reference domain. Zoom into the region around the two flapping beams.

9. Conclusion and Outlook

We were able to prove well-posedness for a set of optimal control problems subject to linear fluid-structure interaction. The additional regularity of the optimal solutions enables the use of the symmetric weak state and adjoint formulation. Hence, we can use a Galerkin approach to discretize the optimality system in space as well as in time. First numerical experiments have shown convergence of the error in the cost functional with refinement of the spatial, time, and control grid. To have convergence result for a more general set of configurations a priori error bounds should be proven for the Galerkin discretization error in the future.

The DWR error estimator presented in Section 7 shows excellent effectivity indices even for complex configurations as the FSI benchmark and the Flapping test. The time discretization error especially pays off, if we equilibrate the spatial and time discretization error. For the chosen geometries with thin beams, bi-quadratic finite element have shown to converge much more stable. Hence, the computations should be repeated with adaptive refinement in space and the use of bi-quadratic finite elements for the presented examples.

The presented optimal control problem of the Young‘s modulus helps to smooth the mesh motion. However, the LBFGS algorithm needs several optimization loops to compute the optimal solution in the high dimensional control space. In combination with dual-weighted residual error estimation or in combination with a second optimization problem the additional computational cost would rather be negligible. Beyond that, the combination enables optimal control with the ALE technique for a wider class of problems.

For optimal control of the mean pressure in the FSI-2 configuration the transition between Neumann and Dirichlet boundary conditions makes the solution difficult. Therefore, Dirichlet control of the velocity at the control boundary Γ_{q} could perhaps help to enable a further reduction of the oscillations.

In this thesis, we applied the optimization algorithm on test configurations as the FSI-2 and FSI-3 benchmark. As a next step more sophisticated examples from hemodynamics should be considered. Especially the choice of the outflow conditions in vascular models and the calibration of the usually used reduced models is still not fully understood. The presented results should be transferable to such configurations. Thereby, sensitivity information of the whole time dependent model are available and can help to determine the parameters in the reduced models more accurately. Furthermore, the use of time dependent and distributed parameters leaves the door open for the extension of the vascular models.

To sum up, optimal control works well for a large class of test examples. A posteriori estimation helps to reduce the computational cost and enables fast solution of the optimization problem. Nevertheless, already for the rather simple test configuration very robust algorithms are needed. The complexity of the systems requires a very fine mesh resolution and a small time step size. Therefore, the computational cost is still high for optimal control despite

the adaptive approach. Hence, further investigations are necessary until the algorithms can provide reliable results, enabling patient specific diagnosis in hemodynamics.

A. Transformation of Sobolev Spaces

Lemma A.1. *Let Ω and $\check{\Omega}$ be two domains in $\mathbb{R}^d$ and let $T : \Omega \times I \to \check{\Omega}(t)$ with $T(\Omega) = \check{\Omega}(t)$ be a $C(I; C^1(\Omega)) \cap C^1(I; C(\Omega))$-diffeomorphism. Then the composition operators*

$$\begin{aligned}
&\check{\varphi} = \varphi \circ T^{-1} \quad \forall \varphi \in \left\{ \varphi \;\middle|\; \varphi \in L^2(I; H^1(\Omega)) \text{ and } \partial_t \check{\varphi} \in L^2(I; L^2(\Omega)) \right\} \\
&\varphi = \check{\varphi} \circ T \quad \forall \check{\varphi} \in \left\{ \check{\varphi} \;\middle|\; \check{\varphi} \in L^2(I; H^1(\check{\Omega}(t))) \text{ and } \partial_t \check{\varphi} \in L^2(I; L^2(\check{\Omega}(t))) \right\}
\end{aligned}$$

are continuous. Hence the stated Sobolev spaces are equivalent.

Proof. We follow here the proof by Wloka of Lemma 4.2 in [152]. Let

$$\varphi \in \left\{ \varphi \;\middle|\; \varphi \in L^2(I; H^1(\Omega)) \text{ and } \partial_t \varphi \in L^2(I; L^2(\Omega)) \right\}.$$

We first consider the norm of the space $H^1(\check{\Omega})$ and show with $F = \nabla T$ and $J = \det(F)$

$$\begin{aligned}
\|\check{\varphi}\|^2_{L^2(I;H^1(\check{\Omega}(t)))} &= \int_I \int_{\check{\Omega}(t)} |\check{\varphi}(\check{x},t)|^2 \,\mathrm{d}\check{x}\,\mathrm{d}t + \int_I \int_{\check{\Omega}(t)} |\check{\nabla}\check{\varphi}(\check{x},t)|^2 \,\mathrm{d}\check{x}\,\mathrm{d}t \\
&= \int_I \int_{\Omega} J|\varphi(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t + \int_I \int_{\Omega} J|\nabla\varphi(x,t)F^{-1}|^2 \,\mathrm{d}x\,\mathrm{d}t \\
&\le C\|\varphi\|^2_{L^2(I;H^1(\Omega))}.
\end{aligned}$$

Now we try to find an estimate for the time derivative $\partial_t\check{\varphi}$

$$\begin{aligned}
\|\partial_t\check{\varphi}\|_{L^2(I;L^2(\check{\Omega}(t)))} &= \int_I \int_{\check{\Omega}(t)} |\partial_t\check{\varphi}(\check{x},t)|^2 \,\mathrm{d}\check{x}\,\mathrm{d}t \\
&= \int_I \int_{\Omega} |J\partial_t\varphi(x,t) - JF^{-1}\partial_t T(x,t) \circ \nabla\varphi|^2 \,\mathrm{d}x\,\mathrm{d}t \\
&\le C\|\partial_t\varphi\|^2_{L^2(I;L^2(\Omega))} + C\|\varphi\|^2_{L^2(I;H^1(\Omega))}.
\end{aligned}$$

If we combine the two estimates we get

$$\|\check{\varphi}\|_{L^2(I;H^1(\check{\Omega}(t)))} + \|\partial_t\check{\varphi}\|_{L^2(I;L^2(\check{\Omega}(t)))} \le C\big[\|\varphi\|_{L^2(I;H^1(\Omega))} + \|\partial_t\varphi\|_{L^2(I;L^2(\Omega))}\big].$$

The estimate for $\varphi = \check{\varphi} \circ T$ with given $\check{\varphi}$ is equivalent. □

B. Linearization of the Nonlinear FSI Problem

In order to compute sensitivities of the functional of interest we need to calculate in Section 4.3.2 the derivative of the Lagrange functional

$$\begin{aligned}\mathcal{L}(q,\mathbf{u},\mathbf{z}) :=& \mathcal{J}(q,\mathbf{u}) - a(\mathbf{u})(\mathbf{z}) + f(q)(\varphi) \\ &+ (u_0 - u(0), z^u(0)) + (v_0 - v(0), z^v(0))\end{aligned}$$

with respect to the state solution $\mathbf{u}$. This includes the linearization of the semi-linear form $a(\mathbf{u})(\mathbf{z})$ defined in (4.10) with respect to the veloctiy variable v, the displacement u and the pressure p. This has already be done in [56, 126, 149]. We summarize once more the resulting derivatives of the single sub-problems in the following. To take the derivative of the ALE-transformation we need to compute the derivative of the deformation gradient F and of its inverses F^{-1}. Therefore, we have to apply the following Lemma, which is taken from [126], several times.

Lemma B.1. *Let $F := \mathrm{Id} + \nabla u$ and $J = \det F$ its determinant. It holds :*

$$\begin{aligned}
(i) \qquad & \frac{dF}{du}(\psi) = \nabla\psi \\
(ii) \qquad & \frac{dF^T}{du}(\psi) = \nabla\psi^T \\
(iii) \qquad & \frac{dF^{-1}}{du}(\psi) = -F^{-1}\nabla\psi F^{-1} \\
(iv) \qquad & \frac{dJ}{du}(\psi) = J\,\mathrm{tr}(F^{-1}\nabla\psi)
\end{aligned}$$

Proof. The relations (i) and (ii) are directly available. If we take the derivative on the relation $\mathrm{Id} = F^{-1}F$ and apply the chain rule we get

$$0 = \frac{d}{du}F^{-1}F = \frac{dF^{-1}}{du}F + F^{-1}\frac{dF}{du} \Rightarrow \frac{dF^{-1}}{du} = -F^{-1}\frac{dF}{du}F^{-1}$$

The relation (iv) can be shown by component wise calculations.

□

If we apply Lemma B.1 systematically on the semilinear $a(\mathbf{u})(\mathbf{z})$ form in the Lagrange func-

tional, we get as linearization for the semi-linear $a^F(\mathbf{u})(\mathbf{z})$ form, modeling the fluid:

$$\begin{aligned}
a_p'^F(\mathbf{u})(\xi, z^v) =& (\!(J\frac{d\sigma_\mathrm{f}}{dp}(\xi)F^{-T}, \nabla z^v)\!)_\mathrm{f} \\
a_v'^F(\mathbf{u})(\varphi, z^v) =& (\!(\rho_\mathrm{f} J\partial_t\varphi, z^v)\!)_\mathrm{f} + (\!(J\frac{d\sigma_\mathrm{f}}{dv}(\varphi)F^{-T}, \nabla z^v)\!)_\mathrm{f} \\
&+ (\!(\rho_\mathrm{f} J(F^{-1}\varphi\cdot\nabla)v), z^v)\!)_\mathrm{f} + (\!(\rho_\mathrm{f} J(F^{-1}(v-\partial_t u)\cdot\nabla)\varphi), z^v)\!)_\mathrm{f} \\
a_u'^F(\mathbf{u})(\psi, z^v) =& (\!(\rho_\mathrm{f} J\,\mathrm{tr}(F^{-1}\nabla\psi)\partial_t v, z^v)\!)_\mathrm{f} - (\!(\rho_\mathrm{f} J(F^{-1}\partial_t\psi\cdot\nabla)v), z^v)\!)_\mathrm{f} \\
&+ (\!(\rho_\mathrm{f} J\,\mathrm{tr}(F^{-1}\nabla\psi)(F^{-1}(v-\partial_t u)\cdot\nabla)v), z^v)\!)_\mathrm{f} \\
&- (\!(\rho_\mathrm{f} J(F^{-1}\nabla\psi F^{-1}(v-\partial_t u)\cdot\nabla)v), z^v)\!)_\mathrm{f} \\
&+ (\!(J\,\mathrm{tr}(F^{-1}\nabla\psi)\sigma_\mathrm{f}F^{-T}, \nabla z^v)\!)_\mathrm{f} - (\!(J\sigma_\mathrm{f}F^{-T}\nabla\psi^T F^{-T}, \nabla z^v)\!)_\mathrm{f} \\
&+ (\!(J\frac{d\sigma_\mathrm{f}}{du}(\psi)F^{-T}, \nabla z^v)\!)_\mathrm{f}.
\end{aligned}$$

Thereby the derivative of the fluid stress tensor is given by

$$\begin{aligned}
\frac{d\sigma_\mathrm{f}}{dp}(\xi) &= -\xi\,\mathrm{Id} \\
\frac{d\sigma_\mathrm{f}}{dv}(\varphi) &= \rho_\mathrm{f}\nu_\mathrm{f}(\nabla\varphi F^{-1} + F^{-T}\nabla\varphi^T) \\
\frac{d\sigma_\mathrm{f}}{du}(\psi) &= \rho_\mathrm{f}\nu_\mathrm{f}(-\nabla v_\mathrm{f}F^{-1}\nabla\psi F^{-1} - F^{-T}\nabla\psi F^{-T}\nabla v_\mathrm{f}^T).
\end{aligned}$$

The linearization of the divergence condition $a^D(\mathbf{u})(\mathbf{z})$ results in

$$\begin{aligned}
a_v'^D(\mathbf{u})(\varphi, z^p) =& (\!(\mathrm{div}(JF^{-1}\varphi), z^p)\!)_\mathrm{f} \\
a_u'^D(\mathbf{u})(\psi, z^p) =& (\!(\mathrm{div}(J\,\mathrm{tr}(F^{-1}\nabla\psi)F^{-1}v), z^p)\!)_\mathrm{f} - (\!(\mathrm{div}(JF^{-1}\nabla\psi F^{-1}v), z^p)\!)_\mathrm{f}.
\end{aligned}$$

Now we have to take in addition semi-linear form $a^S(\mathbf{u})(\mathbf{z})$ and $a^V(\mathbf{u})(\mathbf{z})$ modeling the solid into account. If we calculate the derivates of the nonlinear elastodynamic equations we obtain:

$$\begin{aligned}
a_v'^S(\mathbf{u})(\varphi, z^v) =& (\!(\rho_\mathrm{s}\partial_t\varphi, z^v)\!)_\mathrm{s} \\
a_u'^S(\mathbf{u})(\psi, z^v) =& (\!(\nabla\psi\Sigma_\mathrm{s}, \nabla z^v)\!)_\mathrm{s} + (\!(F\frac{d\Sigma_\mathrm{s}}{du}(\psi), \nabla z^v)\!)_\mathrm{s} \\
a_v'^V(\mathbf{u})(\varphi, z^u) =& (\!(-\varphi, z^u)\!)_\mathrm{s} \\
a_u'^V(\mathbf{u})(\psi, z^u) =& (\!(\psi_t, z^u)\!)_\mathrm{s}.
\end{aligned}$$

Thereby the derivative of the solid stress tensor are given by

$$\begin{aligned}
\frac{d\Sigma_\mathrm{s}}{du}(\psi) &= \lambda_\mathrm{s}\,\mathrm{tr}(\frac{dE_\mathrm{s}}{du}(\psi))\,\mathrm{Id} + 2\mu_\mathrm{s}\frac{dE_\mathrm{s}}{du}(\psi) \\
\frac{dE_\mathrm{s}}{du}(\psi) &= \frac{1}{2}(\nabla\psi^T F + F^T\nabla\psi).
\end{aligned}$$

Finally we have to compute the derivative of the mesh motion $a^M(\mathbf{u})(z^u)$. This leads for the harmonic extension to

$$\begin{aligned} a_u'^M(\mathbf{u})(\psi, z^u) &= (\!(\frac{d\sigma_m}{du}(\psi), \nabla z^u)\!)_{\mathrm{f}} \\ &= (\!(\lambda_{\mathrm{m}} \operatorname{tr}(\frac{1}{2}(\nabla\psi + \nabla\psi^T))\operatorname{Id} + 2\mu_{\mathrm{m}}\frac{1}{2}(\nabla\psi + \nabla\psi^T), \nabla z^u)\!)_{\mathrm{f}}. \end{aligned}$$

At first sight, we would have to demand, that velocity and deformations are two times-differentiable to calculate the semi-linearform $a^D(\mathbf{u})(\mathbf{z})$ and its derivatives. But according to [149, 126] the divergence terms can be restated into:

$$\begin{aligned} &(\!(\operatorname{div}(JF^{-1}\varphi), z^p)\!)_{\mathrm{f}} = (\!(J\operatorname{tr}(\nabla\varphi F^{-1}), z^p)\!)_{\mathrm{f}} \\ &(\!(\operatorname{div}(J\operatorname{tr}(F^{-1}\nabla\psi)F^{-1}v), z^p)\!)_{\mathrm{f}} = (\!(J\operatorname{tr}(F^{-1}\nabla\psi)\operatorname{tr}(\nabla v F^{-1}), z^p)\!)_{\mathrm{f}} \\ &(\!(\operatorname{div}(JF^{-1}\nabla\psi F^{-1}v), z^p)\!)_{\mathrm{f}} = (\!(J\operatorname{tr}(\nabla v F^{-1}\nabla\psi F^{-1}), z^p)\!)_{\mathrm{f}}. \end{aligned}$$

Therefore, it is sufficient to use in our numerical calculations subspaces of H^1 Sobolev-spaces.

Acknowledgements

First of all I sincerely thank my supervisor Prof. Dr. Boris Vexler for suggesting this interesting dissertation topic. I am deeply thankful for his guidance, help, advice as well as his patience. I am very grateful for his encouragement to participate at international conferences and for making my associate membership in the IGDK possible.

My sincere gratitude goes to Prof. Dr. Thomas Richter for providing new impulses for the development of this thesis and sharing his knowledge on FSI. I would like to thank Prof. Dr. Winnifried Wollner for reviewing this thesis and the interesting discussions we had on conferences.

I gratefully acknowledge the funding received from the the federal ministry of education and research (BMBF) within the project ExtremSimOpt: Modellierung, Simulation und Optimierung von Strömungsvorgängen unter Extrembedingungen. Furthermore, I would like to thank the IGSSE graduate school for the financial support on the two month stay at Linz. I deeply thank Thomas Wick for inviting me for a research visit to RICAM. Thank you for two intensive months of collaboration and discussions.

Special thanks to my mentor Dominik for his guidance and the numerous scientific discussions. Thank you for the untiring effort to answer my countless questions on RoDoBo and C++, as well as helping to smooth my latex style.

Furthermore, I would like to thank all my colleagues, especially I want to mention Daniel A., Lucas, Andreas, Korbinian, Dominik M., Jonny, Christian, Ira, Dominik H., Thomas W., Daniel W. and Bernhard for the enjoyable time at the chair and their helpfulness. Thanks Lucas for helping me through the world of semigroup theory. Special thanks go to all members of the IGDK, among others for three incredible summerworkshops.

Finally I am grateful to my family, friends and girlfriend for their support in every kind of way.

Bibliography

[1] F. Abergel and R. Temam, *On some control problems in fluid mechanics*, Theoret. Comput. Fluid Dynamic., 1 (1990), pp. 303–325.

[2] L. Asner, M. Hadjicharalambous, R. Chabiniok, D. Peresutti, E. Sammut, J. Wong, G. Carr-White, P. Chowienczyk, J. Lee, A. King, et al., *Estimation of passive and active properties in the human heart using 3D tagged MRI*, Biomech. Model. Mechanobiol., 15 (2016), pp. 1121–1139.

[3] G. Avalos, I. Lasiecka, and R. Triggiani, *Higher regularity of a coupled parabolic-hyperbolic fluid-structure interactive system*, Georgian Math. J., 15 (2008), pp. 403–437.

[4] ——, *Beyond lack of compactness and lack of stability of a coupled parabolic-hyperbolic fluid-structure system*, in Optimal control of coupled systems of partial differential equations, vol. 158 of Internat. Ser. Numer. Math., Birkhäuser Verlag, Basel, 2009, pp. 1–33.

[5] G. Avalos and R. Triggiani, *The coupled PDE system arising in fluid/structure interaction. I. Explicit semigroup generator and its spectral properties*, in Fluids and waves, vol. 440 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2007, pp. 15–54.

[6] ——, *Uniform stabilization of a coupled PDE system arising in fluid-structure interaction with boundary dissipation at the interface*, Discrete Contin. Dyn. Syst., 22 (2008), pp. 817–833.

[7] ——, *Boundary feedback stabilization of a coupled parabolic-hyperbolic Stokes-Lamé PDE system*, J. Evol. Equ., 9 (2009), pp. 341–370.

[8] ——, *Semigroup well-posedness in the energy space of a parabolic-hyperbolic coupled Stokes-Lamé PDE system of fluid-structure interaction*, Discrete Contin. Dyn. Syst. Ser. S, 2 (2009), pp. 417–447.

[9] ——, *Fluid-structure interaction with and without internal dissipation of the structure: a contrast study in stability*, Evol. Equ. Control Theory, 2 (2013), pp. 563–598.

[10] S. Badia, F. Nobile, and C. Vergara, *Robin-Robin preconditioned Krylov methods for fluid-structure interaction problems*, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 2768–2784.

[11] S. Badia, A. Quaini, and A. Quarteroni, *Splitting methods based on algebraic factorization for fluid-structure interaction*, SIAM J. Sci. Comput., 30 (2008), pp. 1778–1805.

[12] V. Barbu, Z. Grujić, I. Lasiecka, and A. Tuffaha, *Existence of the energy-level weak solutions for a nonlinear fluid-structure interaction model*, in Fluids and waves, vol. 440 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2007, pp. 55–82.

[13] ——, *Smoothness of weak solutions to a nonlinear fluid-structure interaction model*, Indiana Univ. Math. J., 57 (2008), pp. 1173–1207.

[14] Y. Bazilevs, K. Takizawa, and T. Tezduyar, *Computational Fluid-Structure Interaction: Methods and Applications*, Wiley, 2013.

[15] R. Becker, *Adaptive Finite Elements for Optimal Control Problems*, habilitation, University of Heidelberg, 2001.

[16] R. Becker and M. Braack, *A finite element pressure gradient stabilization for the Stokes equations based on local projections*, Calcolo, 38 (2001), pp. 173–199.

[17] R. Becker and M. Braack, *A two-level stabilization scheme for the Navier-Stokes equations*, in Numerical mathematics and advanced applications, Springer, Berlin, 2004, pp. 123–130.

[18] R. Becker, D. Meidner, and B. Vexler, *Efficient numerical solution of parabolic optimization problems by finite element methods*, Optim. Methods Softw., 22 (2007), pp. 813–833.

[19] R. Becker and R. Rannacher, *A feed-back approach to error control in finite element methods: basic analysis and examples*, East-West J. Numer. Math., 4 (1996), pp. 237–264.

[20] R. Becker and R. Rannacher, *An optimal control approach to a posteriori error estimation in finite element methods*, Acta Numer., 10 (2001), pp. 1–102.

[21] R. Becker and B. Vexler, *A posteriori error estimates for parameter identification*, in Numerical mathematics and advanced applications, Springer, Berlin, 2004, pp. 131–140.

[22] ——, *Mesh refinement and numerical sensitivity analysis for parameter calibration of partial differential equations*, J. Comput. Phys., 206 (2005), pp. 95–110.

[23] A. Bermúdez de Castro, D. Gómez, P. Quintela, and P. Salgado, eds., *Numerical mathematics and advanced applications*, Springer-Verlag, Berlin, 2006.

[24] L. Bertagna, M. D'Elia, M. Perego, and A. Veneziani, *Data assimilation in cardiovascular fluid-structure interaction problems: an introduction*, in Fluid-structure interaction and biomedical applications, Adv. Math. Fluid Mech., Birkhäuser/Springer, Basel, 2014, pp. 395–481.

[25] C. Bertoglio, D. Barber, N. Gaddum, I. Valverde, M. Rutten, P. Beerbaum, P. Moireau, R. Hose, and J. Gerbeau, *Identification of artery wall stiffness: In vitro validation and in vivo results of data assimilation procedure applied to a 3D fluid-structure interaction model*, J. Biomech., 47 (2014), pp. 1027–34.

[26] C. Bertoglio, D. Chapelle, M. A. Fernández, J.-F. Gerbeau, and P. Moireau, *State observers of a vascular fluid-structure interaction model through measurements in the solid*, Comput. Methods Appl. Mech. Engrg., 256 (2013), pp. 149–168.

[27] C. Bertoglio, P. Moireau, and J.-F. Gerbeau, *Sequential parameter estimation for fluid–structure problems: application to hemodynamics*, Int. J. Numer. Methods Biomed. Eng., 28 (2012), pp. 434–455.

[28] M. Besier and R. Rannacher, *Goal-oriented space-time adaptivity in the finite element Galerkin method for the computation of nonstationary incompressible flow*, Internat. J. Numer. Methods Fluids, 70 (2012), pp. 1139–1166.

[29] H. Blum, Q. Lin, and R. Rannacher, *Asymptotic error expansion and Richardson extrapolation for linear finite elements*, Numer. Math., 49 (1986), pp. 11–37.

[30] L. Bociu and J.-P. Zolésio, *Linearization of a coupled system of nonlinear elasticity and viscous fluid*, in Modern aspects of the theory of partial differential equations, vol. 216 of Oper. Theory Adv. Appl., Birkhäuser/Springer Basel AG, Basel, 2011, pp. 93–120.

[31] ——, *Sensitivity analysis for a free boundary fluid-elasticity interaction*, Evol. Equ. Control Theory, 2 (2013), pp. 55–79.

[32] M. Braack, *Optimal control in fluid mechanics by finite elements with symmetric stabilization*, SIAM J. Control Optim., 48 (2009), pp. 672–687.

[33] M. Braack and A. Ern, *Adaptive computation of reactive flows with local mesh refinement and model adaptation*, in Numerical mathematics and advanced applications, Springer, Berlin, 2004, pp. 159–168.

[34] M. Braack and T. Richter, *Solving multidimensional reactive flow problems with adaptive finite elements*, in Reactive flows, diffusion and transport, Springer, Berlin, 2007, pp. 93–112.

[35] D. Braess, *Finite elements*, Cambridge University Press, Cambridge, second ed., 2001.

[36] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, vol. 15 of Texts in Applied Mathematics, Springer-Verlag, New York, 1994.

[37] M. O. Bristeau, R. Glowinski, and J. Périaux, *Numerical methods for the Navier-Stokes equations. Applications to the simulation of compressible and incompressible viscous flows*, in Finite elements in physics (Lausanne, 1986), North-Holland, Amsterdam, 1987, pp. 73–187.

[38] A. N. Brooks and T. J. R. Hughes, *Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations*, Comput. Methods Appl. Mech. Engrg., 32 (1982), pp. 199–259. FENOMECH '81, Part I (Stuttgart, 1981).

[39] F. Bucci and I. Lasiecka, *Optimal boundary control with critical penalization for a PDE model of fluid-solid interactions*, Calc. Var. Partial Differential Equations, 37 (2010), pp. 217–235.

[40] H.-J. Bungartz, M. Mehl, and M. Schäfer, eds., *Fluid structure interaction. II*, vol. 73 of Lecture Notes in Computational Science and Engineering, Springer, Heidelberg, 2010. Modelling, simulation, optimization.

[41] H.-J. Bungartz and M. Schäfer, eds., *Fluid-structure interaction*, vol. 53 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin, 2006. Modelling, simulation, optimisation.

[42] C. Carstensen and R. Verfürth, *Edge residuals dominate a posteriori error estimates for low order finite element methods*, SIAM J. Numer. Anal., 36 (1999), pp. 1571–1587.

[43] P. Causin, J. F. Gerbeau, and F. Nobile, *Added-mass effect in the design of partitioned algorithms for fluid-structure problems*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 4506–4527.

[44] P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978. Studies in Mathematics and its Applications, Vol. 4.

[45] ——, *Mathematical elasticity. Vol. I*, vol. 20 of Studies in Mathematics and its Applications, North-Holland Publishing Co., Amsterdam, 1988. Three-dimensional elasticity.

[46] P. G. Ciarlet and P.-A. Raviart, *Interpolation theory over curved elements, with applications to finite element methods*, Comput. Methods Appl. Mech. Engrg., 1 (1972), pp. 217–249.

[47] D. Coutand and S. Shkoller, *Motion of an elastic solid inside an incompressible viscous fluid*, Arch. Ration. Mech. Anal., 176 (2005), pp. 25–102.

[48] D. Coutand and S. Shkoller, *The interaction between quasilinear elastodynamics and the Navier-Stokes equations*, Arch. Rational Mech. Anal., 179 (2006), pp. 303–352.

[49] J. Degroote, M. Hojjat, E. Stavropoulou, R. Wüchner, and K.-U. Bletzinger, *Partitioned solution of an unsteady adjoint for strongly coupled fluid-structure interactions and application to parameter identification of a one-dimensional problem*, Struct. Multidiscip. Optim., 47 (2013), pp. 77–94.

[50] ——, *Partitioned solution of an unsteady adjoint for strongly coupled fluid-structure interactions and application to parameter identification of a one-dimensional problem*, Struct. Multidiscip. Optim., 47 (2013), pp. 77–94.

[51] M. D'Elia, L. Mirabella, T. Passerini, M. Perego, M. Piccinelli, C. Vergara, and A. Veneziani, *Modeling of Physiological Flows*, vol. 5 of MS&A - Modeling, Simulation and Applications, Springer Milan, 2012, ch. Applications of variational data assimilation in computational hemodynamics, pp. 363–394.

[52] J. Donéa, P. Fasoli-Stella, and S. Giuliani, *Lagrangian and Eulerian finite element techniques for transient fluid-structure interaction problems*, in Trans. 4th Int. Conf. on Structural Mechanics in Reactor Technology, 1977, p. Paper B1/2.

[53] Q. Du, M. D. Gunzburger, L. S. Hou, and J. Lee, *Analysis of a linear fluid-structure interaction problem*, Discrete Contin. Dyn. Syst., 9 (2003), pp. 633–650.

[54] ——, *Semidiscrete finite element approximations of a linear fluid-structure interaction problem*, SIAM J. Numer. Anal., 42 (2004), pp. 1–29.

[55] T. Dunne, *An Eulerian approach to fluid-structure interaction and goal-oriented mesh adaption*, Int. J. Numer. Methods in Fluids, 51 (2006), pp. 1017–1039.

[56] T. Dunne, R. Rannacher, and T. Richter, *Numerical simulation of fluid-structure interaction based on monolithic variational formulations*, in Fundamental trends in fluid-structure interaction, vol. 1 of Contemp. Chall. Math. Fluid Dyn. Appl., World Sci. Publ., Hackensack, NJ, 2010, pp. 1–75.

[57] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Computational differential equations*, Cambridge University Press, Cambridge, 1996.

[58] L. Failer, D. Meidner, and B. Vexler, *Optimal control of a linear unsteady fluid-structure interaction problem*, J. Optim. Theory Appl., 170 (2016), pp. 1–27.

[59] L. Failer and T. Wick, *Adaptive time-step control for nonlinear fluid-structure interaction.* in preparation (2017).

[60] M. A. Fernández, *Coupling schemes for incompressible fluid-structure interaction: implicit, semi-implicit and explicit*, SeMA Journal, 55 (2011), pp. 59–108.

[61] M. A. Fernández, J.-F. Gerbeau, and C. Grandmont, *A projection semi-implicit scheme for the coupling of an elastic structure with an incompressible fluid*, Internat. J. Numer. Methods Engrg., 69 (2007), pp. 794–821.

[62] P. W. Fick, E. H. van Brummelen, and K. G. van der Zee, *On the adjoint-consistent formulation of interface conditions in goal-oriented error estimation and adaptivity for fluid-structure interaction*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 3369–3385.

[63] L. Formaggia, J. F. Gerbeau, F. Nobile, and A. Quarteroni, *On the coupling of 3D and 1D Navier-Stokes equations for flow problems in compliant vessels*, Comput. Methods Appl. Mech. Engrg., 191 (2001), pp. 561–582.

[64] L. Formaggia and F. Nobile, *A stability analysis for the arbitrary Lagrangian Eulerian formulation with finite elements*, East-West J. Numer. Math., 7 (1999), pp. 105–131.

[65] L. Formaggia, F. Nobile, A. Quarteroni, and A. Veneziani, *Multiscale modelling of the circulatory system: a preliminary analysis*, Comput. Visual. Sci., 2 (1999), pp. 75–83.

[66] L. Formaggia, A. Quarteroni, and A. Veneziani, eds., *Cardiovascular mathematics*, vol. 1 of MS&A. Modeling, Simulation and Applications, Springer-Verlag Italia, Milan, 2009. Modeling and simulation of the circulatory system.

[67] S. Frei and T. Richter, *A second-order time-stepping scheme for parabolic interface problems with moving interfaces.* accepted for publication at ESAIM: M2AN (2016).

[68] ——, *A locally modified parametric finite element method for interface problems*, SIAM J. Numer. Anal., 52 (2014), pp. 2315–2334.

[69] G. P. Galdi, *An introduction to the mathematical theory of the Navier-Stokes equations. Vol. I*, vol. 38 of Springer Tracts in Natural Philosophy, Springer-Verlag, New York, 1994. Linearized steady problems.

[70] G. P. Galdi and R. Rannacher, eds., *Fundamental trends in fluid-structure interaction*, vol. 1 of Contemporary Challenges in Mathematical Fluid Dynamics and Its Applications, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2010.

[71] M. W. Gee, U. Küttler, and W. A. Wall, *Truly monolithic algebraic multigrid for fluid-structure interaction*, Internat. J. Numer. Methods Engrg., 85 (2011), pp. 987–1016.

[72] C. Geiger and C. Kanzow, *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben*, Springer-Verlag, 2013.

[73] A. J. Gil, A. Arranz Carreño, J. Bonet, and O. Hassan, *The immersed structural potential method for haemodynamic applications*, J. Comput. Phys., 229 (2010), pp. 8613–8641.

[74] J. C. Gilbert and C. Lemaréchal, *Some numerical experiments with variable-storage quasi-Newton algorithms*, Math. Programming, 45 (1989), pp. 407–435.

[75] C. Goll, R. Rannacher, and W. Wollner, *The damped Crank-Nicolson time-marching scheme for the adaptive solution of the Black-Scholes equation*, J. Comput. Finance, 18 (2015), pp. 1–37.

[76] C. Grandmont, *Existence of weak solutions for the unsteady interaction of a viscous fluid with an elastic plate*, SIAM J. Math. Anal., 40 (2008), pp. 716–737.

[77] C. Grandmont and M. Hillairet, *Existence of global strong solutions to a beam-fluid interaction system*, Arch. Ration. Mech. Anal., 220 (2016), pp. 1283–1333.

[78] P. Grisvard, *Elliptic problems in nonsmooth domains*, vol. 24 of Monographs and Studies in Mathematics, Pitman (Advanced Publishing Program), Boston, MA, 1985.

[79] M. D. Gunzburger and S. Manservisi, *The velocity tracking problem for Navier-Stokes flows with bounded distributed controls*, SIAM J. Control Optim., 37 (1999), pp. 1913–1945.

[80] R. Hartmann, *Multitarget error estimation and adaptivity in aerodynamic flow simulations*, SIAM J. Sci. Comput., 31 (2008), pp. 708–731.

[81] M. Hinze, *A variational discretization concept in control constrained optimization: the linear-quadratic case*, Comput. Optim. Appl., 30 (2005), pp. 45–61.

[82] M. Hinze and K. Kunisch, *Second order methods for optimal control of time-dependent fluid flow*, SIAM J. Control Optim., 40 (2001), pp. 925–946.

[83] C. Hirt, A. Amsden, and J. Cook, *An arbitrary Lagrangian-Eulerian computing method for all flow speeds*, J. Comput. Phys., 14 (1974), pp. 227–253.

[84] J. Hron and S. Turek, *A monolithic FEM/multigrid solver for an ALE formulation of fluid-structure interaction with applications in biomechanics*, in Fluid-structure interaction, vol. 53 of Lect. Notes Comput. Sci. Eng., Springer, Berlin, 2006, pp. 146–170.

[85] T. Hughes, W. Liu, and T. Zimmermann, *Lagrangian-Eulerian finite element formulation for incompressible viscous flows*, Comput. Methods Appl. Mech. Engrg., 29 (1981), pp. 329–349.

[86] M. Ignatova, I. Kukavica, I. Lasiecka, and A. Tuffaha, *On well-posedness for a free boundary fluid-structure model*, J. Math. Phys., 53 (2012), pp. 115624, 13.

[87] ——, *On well-posedness and small data global existence for an interface damped free boundary fluid-structure model*, Nonlinearity, 27 (2014), pp. 467–499.

[88] M. Ismail, W. A. Wall, and M. W. Gee, *Adjoint-based inverse analysis of windkessel parameters for patient-specific vascular models*, J. Comput. Phys., 244 (2013), pp. 113–130.

[89] V. John and J. Rang, *Adaptive time step control for the incompressible Navier-Stokes equations*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 514–524.

[90] C. Johnson, *Discontinuous Galerkin finite element methods for second order hyperbolic problems*, Comput. Methods Appl. Mech. Engrg., 107 (1993), pp. 117–129.

[91] C. T. Kelley and E. W. Sachs, *Quasi-Newton methods and unconstrained optimal control problems*, SIAM J. Control Optim., 25 (1987), pp. 1503–1516.

[92] M. K. Krishnamoorthy, R. K. Banerjee, Y. Wang, J. Zhang, A. S. Roy, S. F. Khoury, L. J. Arend, S. Rudich, and P. Roy-Chaudhury, *Hemodynamic wall shear stress profiles influence the magnitude and pattern of stenosis in a pig av fistula*, Kidney Int., 74 (2008), pp. 1410–1419.

[93] I. Kukavica and A. Tuffaha, *Solutions to a fluid-structure interaction free boundary problem*, Discrete Contin. Dyn. Syst., 32 (2012), pp. 1355–1389.

[94] I. Kukavica, A. Tuffaha, and M. Ziane, *Strong solutions to a nonlinear fluid structure interaction system*, J. Differential Equations, 247 (2009), pp. 1452–1478.

[95] ——, *Strong solutions to a Navier-Stokes-Lamé system on a domain with a non-flat boundary*, Nonlinearity, 24 (2011), pp. 159–176.

[96] K. Kunisch and B. Vexler, *Optimal vortex reduction for instationary flows based on translation invariant cost functionals*, SIAM J. Control Optim., 46 (2007), pp. 1368–1397.

[97] U. Küttler, M. Gee, C. Förster, A. Comerford, and W. A. Wall, *Coupling strategies for biomedical fluid-structure interaction problems*, Int. J. Numer. Methods Biomed. Eng., 26 (2010), pp. 305–321.

[98] P. Lamata, A. Pitcher, S. Krittian, D. Nordsletten, M. M. Bissell, T. Cassar, A. J. Barker, M. Markl, S. Neubauer, and N. P. Smith, *Aortic relative pressure components derived from four-dimensional flow cardiovascular magnetic resonance*, Magn. Reson. Med., 72 (2014), pp. 1162–1169.

[99] I. Lasiecka, J.-L. Lions, and R. Triggiani, *Nonhomogeneous boundary value problems for second order hyperbolic operators*, J. Math. Pures Appl. (9), 65 (1986), pp. 149–192.

[100] I. Lasiecka and R. Triggiani, *Recent Advances in Regularity of Second-order Hyperbolic Mixed Problems, and Applications*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1994, pp. 104–162.

[101] I. Lasiecka and A. Tuffaha, *Optimal feedback synthesis for Bolza control problem arising in linearized fluid structure interaction*, in Optimal control of coupled systems of partial differential equations, vol. 158 of Internat. Ser. Numer. Math., Birkhäuser Verlag, Basel, 2009, pp. 171–190.

[102] ——, *Riccati theory and singular estimates for a Bolza control problem arising in linearized fluid-structure interaction*, Systems Control Lett., 58 (2009), pp. 499–509.

[103] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza, *A reduced computational and geometrical framework for inverse problems in hemodynamics*, Int. J. Numer. Methods Biomed. Eng., 29 (2013), pp. 741–776.

[104] J. Lequeurre, *Quelques résultats d'existence de contrôlabilité et de stabilisation pour des systémes couplés fluid-structure*, PhD thesis, L'Université de Toulouse, 2011.

[105] J.-L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications. Vol. I*, Springer-Verlag, New York-Heidelberg, 1972.

[106] ——, *Non-homogeneous boundary value problems and applications. Vol. II*, Springer-Verlag, New York-Heidelberg, 1972.

[107] D. C. Liu and J. Nocedal, *On the limited memory BFGS method for large scale optimization*, Math. Programming, 45 (1989), pp. 503–528.

[108] A. Manzoni, A. Quarteroni, and G. Rozza, *Shape optimization for viscous flows by reduced basis methods and free-form deformation*, Internat. J. Numer. Methods Fluids, 70 (2012), pp. 646–670.

[109] V. Martin, F. Clément, A. Decoene, and J.-F. Gerbeau, *Parameter identification for a one-dimensional blood flow model*, in CEMRACS 2004—mathematics and applications to biology and medicine, vol. 14 of ESAIM Proc., EDP Sci., Les Ulis, 2005, pp. 174–200.

[110] ——, *Parameter identification for a one-dimensional blood flow model*, in CEMRACS 2004—mathematics and applications to biology and medicine, vol. 14 of ESAIM Proc., EDP Sci., Les Ulis, 2005, pp. 174–200.

[111] D. Meidner and T. Richter, *Goal-oriented error estimation for the fractional step theta scheme*, Comput. Methods Appl. Math., 14 (2014), pp. 203–230.

[112] ——, *A posteriori error estimation for the fractional step theta discretization of the incompressible Navier-Stokes equations*, Comput. Methods Appl. Mech. Engrg., 288 (2015), pp. 45–59.

[113] D. Meidner and B. Vexler, *Adaptive space-time finite element methods for parabolic optimization problems*, SIAM J. Control Optim., 46 (2007), pp. 116–142 (electronic).

[114] ——, *A priori error estimates for space-time finite element discretization of parabolic optimal control problems. I. Problems without control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1150–1177.

[115] ——, *A priori error estimates for space-time finite element discretization of parabolic optimal control problems. II. Problems with control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1301–1329.

[116] ——, *A priori error analysis of the Petrov-Galerkin Crank-Nicolson scheme for parabolic optimal control problems*, SIAM J. Control Optim., 49 (2011), pp. 2183–2211.

[117] P. Moireau, C. Bertoglio, N. Xiao, C. A. Figueroa, C. A. Taylor, D. Chapelle, and J.-F. Gerbeau, *Sequential identification of boundary support parameters in a fluid-structure vascular model using patient image data*, Biomech. Model. Mechanobiol., 12 (2013), pp. 475–496.

[118] P. Moireau, N. Xiao, M. Astorino, C. A. Figueroa, D. Chapelle, C. A. Taylor, and J.-F. Gerbeau, *External tissue support and fluid–structure simulation in blood flows*, Biomech. Model. Mechanobiol., 11 (2012), pp. 1–18.

[119] M. Moubachir and J.-P. Zolésio, *Moving shape analysis and control*, vol. 277 of Pure and Applied Mathematics (Boca Raton), Chapman & Hall/CRC, Boca Raton, FL, 2006. Applications to fluid structure interactions.

[120] W. Noh, *A time-dependent two-space-dimensional coupled Eulerian-Lagrangian code*, vol. 3 of Methods Comput. Phys., Academic Press, New York, 1964, pp. 117–179.

[121] A. Okubo, *Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences*, Deep Sea Research and Oceanographic Abstracts, 17 (1970), pp. 445 – 454.

[122] S. Pant, B. Fabrges, J.-F. Gerbeau, and I. E. Vignon-Clementel, *A methodological paradigm for patient-specific multi-scale cfd simulations: from clinical measurements to parameter estimates for individual analysis*, Int. J. Numer. Method. Biomed. Eng., 30 (2014), pp. 1614–1648.

[123] M. Perego, A. Veneziani, and C. Vergara, *A variational approach for estimating the compliance of the cardiovascular tissue: an inverse fluid-structure interaction problem*, SIAM J. Sci. Comput., 33 (2011), pp. 1181–1211.

[124] A. Quarteroni, A. Veneziani, and C. Vergara, *Geometric multiscale modeling of the cardiovascular system, between theory and practice*, Comput. Methods Appl. Mech. Engrg., 302 (2016), pp. 193–252.

[125] R. Rannacher and F.-T. Suttmeier, *A feed-back approach to error control in finite element methods: application to linear elasticity*, Comput. Mech., 19 (1997), pp. 434–446.

[126] T. Richter, *Finite elements for fluid-structure interactions.* in preparation (2017).

[127] ——, *Parallel Multigrid Method for Adaptive Finite Elements with Application to 3D Flow Problems*, PhD thesis, University of Heidelberg, 2005.

[128] T. Richter, *Goal-oriented error estimation for fluid-structure interaction problems*, Comput. Methods Appl. Mech. Engrg., 223/224 (2012), pp. 28–42.

[129] T. Richter, *A monolithic geometric multigrid solver for fluid-structure interactions in ALE formulation*, Internat. J. Numer. Methods Engrg., 104 (2015), pp. 372–390.

[130] T. Richter and T. Wick, *Optimal control and parameter estimation for stationary fluid-structure interaction problems*, SIAM J. Sci. Comput., 35 (2013), pp. B1085–B1104.

[131] T. Richter and T. Wick, *On time discretizations of fluid-structure interactions*, in Multiple Shooting and Time Domain Decomposition Methods, T. Carraro, M. Geiger, S. Körkel, and R. Rannacher, eds., Contributions in Mathematical and Computational Science, 2015, pp. 377–400.

[132] T. Richter and T. Wick, *Variational localizations of the dual weighted residual estimator*, J. Comput. Appl. Math., 279 (2015), pp. 192–208.

[133] RoDoBo, *A C++ library for optimization with stationary and nonstationary PDEs*. `http://www.rodobo.org/`.

[134] M. Schmich and B. Vexler, *Adaptivity with dynamic meshes for space-time finite element discretizations of parabolic equations*, SIAM J. Sci. Comput., 30 (2007/08), pp. 369–393.

[135] R. L. Spilker and C. A. Taylor, *Tuning multidomain hemodynamic simulations to match physiological measurements*, Ann. Biomed. Eng., 38 (2010), pp. 2635–2648.

[136] K. Stein, T. Tezduyar, and R. Benney, *Mesh moving techniques for fluid-structure interactions with large displacements*, Journal of Applied Mechanics, 70 (2003), pp. 58–63.

[137] R. Temam, *Navier-Stokes equations*, vol. 2 of Studies in Mathematics and its Applications, North-Holland Publishing Co., Amsterdam, third ed., 1984. Theory and numerical analysis, With an appendix by F. Thomasset.

[138] T. Tezduyar, M. Behr, and J. Liou, *A new strategy for finite element computations involving moving boundaries and interfaces - the deforming-spatial-domain/space-time procedure: I. the concept and the preliminary numerical tests*, Comp. Methods Appl. Mech. Engrg., 94 (1992), pp. 339–351.

[139] T. Tezduyar, M. Behr, S. Mittal, and J. Liou, *A new strategy for finite element computations involving moving boundaries and interfaces - the deforming-spatial-domain/space-time procedure: II. Computation of free-surface flows, two-liquid flows, and flows with drifting cylinders*, Comp. Methods Appl. Mech. Engrg., 94 (1992), pp. 353–371.

[140] T. E. Tezduyar, M. Behr, S. Mittal, and A. Johnson, *Computation of unsteady incompressible flows with the finite element methods space-time formulations, iterative strategies and massively parallel implementations*, New Methods in Transient Analysis, ASME: New York, PVP-vol. 246/AMD-vol. 143 (1992), pp. 7–24.

[141] F. Tröltzsch, *Optimale Steuerung partieller Differentialgleichungen*, Vieweg+Teubner Verlag, 2005.

[142] S. Turek and J. Hron, *Proposal for numerical benchmarking of fluid-structure interaction between an elastic object and laminar incompressible flow*, in Fluid-structure interaction, vol. 53 of Lect. Notes Comput. Sci. Eng., Springer, Berlin, 2006, pp. 371–385.

[143] M. Ulbrich and S. Ulbrich, *Primal-dual interior-point methods for PDE-constrained optimization*, Math. Program., 117 (2009), pp. 435–485.

[144] K. G. van der Zee, E. H. van Brummelen, I. Akkerman, and R. de Borst, *Goal-oriented error estimation and adaptivity for fluid-structure interaction using exact linearized adjoints*, Comput. Methods Appl. Mech. Engrg., 200 (2011), pp. 2738–2757.

[145] I. E. Vignon-Clementel, C. A. Figueroa, K. E. Jansen, and C. A. Taylor, *Outflow boundary conditions for three-dimensional finite element modeling of blood flow and pressure in arteries*, Comput. Methods Appl. Mech. Engrg., 195 (2006), pp. 3776–3796.

[146] W. A. Wall, *Fluid-Struktur-Interaktion mit stabilisierten Finiten Elementen*, PhD thesis, University of Stuttgart, 1999.

[147] J. Weiss, *The dynamics of enstrophy transfer in two-dimensional hydrodynamics*, Phys. D, 48 (1991), pp. 273–294.

[148] N. Westerhof, J.-W. Lankhaar, and B. E. Westerhof, *The arterial windkessel*, Med. Biol. Eng. Comput., 47 (2009), pp. 131–141.

[149] T. Wick, *Adaptive Finite Element Simulation of Fluid-Structure Interaction with Application to Heart-Valve Dynamics*, PhD thesis, University of Heidelberg, 2011.

[150] ——, *Goal-oriented mesh adaptivity for fluid-structure interaction with application to heart-valve settings*, Arch. Mech. Engrg., 59 (2012), pp. 73–99.

[151] T. Wick and W. Wollner, *On the differentiability of fluid-structure interaction problems with respect to the problem data*, RICAM report, 16 (2014).

[152] J. Wloka, *Partial differential equations*, Cambridge University Press, Cambridge, 1987.

[153] X. Zhang and E. Zuazua, *Long-time behavior of a coupled heat-wave system arising in fluid-structure interaction*, Arch. Ration. Mech. Anal., 184 (2007), pp. 49–120.

[154] M. Zhu, J. L. Nazareth, and H. Wolkowicz, *The quasi-Cauchy relation and diagonal updating*, SIAM J. Optim., 9 (1999), pp. 1192–1204. Dedicated to John E. Dennis, Jr., on his 60th birthday.