Optimal Dirichlet boundary control problems of high-lift configurations with control and integral state constraints

vorgelegt von

Diplom-Wirtschaftsmathematiker
Dipl.-Math.oec. Christian John

aus Berlin

Von der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. John M. Sullivan
Gutachter: Prof. Dr. Fredi Tröltzsch
Gutachter: Prof. Dr. Arnd Rösch

Tag der wissenschaftlichen Aussprache: 13.07.2011

Berlin 2011

D 83

Abstract

This thesis investigates optimal control problems related to Navier-Stokes equations. We investigate two control problems related to the aerodynamic optimization of flows around airfoils in high-lift configurations.

The first issue is the steady state maximization of lift subject to restrictions on the drag. This leads to a Dirichlet boundary control problem for the stationary Navier-Stokes equations with constrained control functions belonging to $\mathbf{L}^2$ under an integral state constraint. The control space $\mathbf{L}^2$ makes it necessary to deal with very weak solutions of the Navier-Stokes equations and because of the low regularity of control and state, we reformulate the cost functional and the integral state constraint. We derive first-order necessary and second-order sufficient optimality conditions and treat the problem numerically by direct solution of the associated nonsmooth optimality system and additionally by an SQP-method, which convergence we proved.

The second part is based on a k-ω-WILCOX98 turbulence model, describing the nonstationary behavior of the fluid closer to the reality. To deal with the curse of dimension, we discuss a reduced-order model (ROM) by adapting a small system of ODEs to solutions computed with the full model. Based on this ROM, we investigate an optimal control problem theoretically and numerically.

Acknowledgment

I think it is impossible to write a doctoral dissertation without some kind of support. I am grateful to the DFG (SFB 557 'Control of complex turbulent shear flows') supporting me in the first three years financially and to the FAZIT-Stiftung for their financial support in the last 9 months.

My first and biggest gratitude goes to my supervisor Prof. Dr. Fredi Tröltzsch for the interesting topic, his helpful advice and his personality in general. I thank Dr. Daniel Wachsmuth, who worked with me together for about 2 years when i started my project at the TU Berlin, for introducing me into the topic of optimal control of Navier-Stokes equations, his interest in my work and many helpful inspirations. Many thanks go to my colleagues in the research group 'Optimization on PDEs' at the TU Berlin, especially Kristof Altmann for introducing me into COMSOL Multiphysics, and to Prof. Dr. Arnd Rösch for his willingness to review this thesis.

Furthermore, my gratitude to B.R. Noack and M. Schlegel for the good and successful cooperation. They introduced me into the topic of proper orthogonal decomposition and reduced-order modeling and supported me with both words and deeds. I also want to thank M. Luchtenburg for explaining me his reduced-order model, M. Nestler for his helpful assistance and the group of Prof. Thiele, especially B. Günther and A. Carnarius, for providing me with simulations of the URANS system.

Finally, i thank my friends and my family.

Contents

Chapter 1

Introduction

In this thesis, we study optimal control problems related to Navier-Stokes equations, describing the motion of fluid. We investigate minimizations of functionals subject to state equations. The objective functionals depend on the velocity field u, the pressure p and the control function g.

Our main concern is maximization the lift of an airplane, while drag remains beyond a given threshold. Therefore, we consider an objective functional $J(u, p, g)$ characterizing the lift, the Navier-Stokes equations as state equations and a constraint on the drag.

In given literature, there are two different approaches to get influence on the flow around a body. The first one is the possibility of passive control.

There are several possibilities of passive control, e.g. passive blowing, roughness and shaping. Passive noise control devices include shields of rigid and compliant walls, mufflers, silencers, resonators and absorbent materials, see [43] for more details. The idea behind most of them is to reduce vortices and make the airstream around the wing smoother.

The second ansatz is active flow control, which was investigated in particular by the SFB 557 'Control of complex turbulent shear flows'. Here, little slits are installed on a part of the wing, where suction and blowing of air is possible to reduce vortices.

Generally, flow control is a research field gaining a lot of interest in both academic research and industry. It is researched by engineers (experimental and computation fluid dynamics), mathematicians (control theory and optimization) and physicists.

In this work, the following optimal flow control problem is considered: active control of the flow of a fluid around an aircraft by means of suction and blowing on the wing to influence the resulting lift and drag. The associated background of applications in fluid mechanics, active separation control, was the subject of various papers written from an engineering point of view and has been proven to be effective in experiments as well as simulations. We only mention [17, 19, 87, 88, 89, 112], whose considerations are close to our setting, see Chapter 7 to 9.

The first part of this thesis deals with the steady-state problem. Here, we assume a low Reynolds number so that we avoid the discussion of turbulence. Furthermore, we consider a simplified control model, which is composed of the cost functional, the steady-state Navier-Stokes equations, and constraints on the control function as well as the state, for a mathematical investigation. First, the steady-state Navier-Stokes equations, describing the motion of the fluid around the wing, are investigated and we clarify the following questions.

1. What is the best (suitable) definition of a solution of the state equation for the formulated problem?
2. What are the requirements such that a (unique) solution for the state equations exists?
3. What preliminary results can be found in the existing literature?
4. What regularity assumptions are needed?
5. What are the requirements such that a solution for the stated optimization problem exists?
6. How are the necessary and sufficient optimality conditions formulated?
7. What is an appropriate numerical optimization method and does it converge?

We will characterize optimality of control strategies for our setting by necessary and sufficient optimality conditions.

Let us now describe the setting of our optimization problem in detail. Here, Ω is an open bounded domain of $\mathbb{R}^n$, $n = 2, 3$ with boundary Γ, which is assumed to be sufficiently smooth, more details later. The velocity field of the fluid is denoted by u and the pressure by p. The control is a boundary velocity field denoted by g and the viscosity parameter $\nu = 1/Re$ is a positive number. Let us denote the surface measure by $\mathrm{d}s(x)$ or short $\mathrm{d}s$. The term ∇ denotes the gradient and Δ the Laplace operator, which is applied componentwise. The resulting force of the fluid on the wing embedded in the fluid in direction $\vec{e}$ is given as the boundary integral

$$F_{\vec{e}} = \int_{\Gamma_w} \left(\nu \frac{\partial y}{\partial \mathfrak{n}_w} - p\mathfrak{n}_w \right) \cdot \vec{e} \, \mathrm{d}s, \tag{1.0.1}$$

where Γ_w is the boundary of the wing with its outer normal $\mathfrak{n}_w$. Since $\mathfrak{n}_w$ points into the fluid, the normal $\mathfrak{n}_w$ is the negative of the outer normal of the fluid domain Ω, $\mathfrak{n}_w = -\mathfrak{n}$. Let the vectors $\vec{e_l}$ and $\vec{e_d}$ indicate the directions of lift and drag. Then, we are able to calculate the lift and the drag with the boundary integral (1.0.1) where $\vec{e_l}$ or $\vec{e_d}$ have to be inserted instead of $\vec{e}$. Here, $\vec{e_d}$ is the normalized vector directed opposite to the gravity, and $\vec{e_l}$ is the normalized vector in the direction opposing the main flow field.

The optimization problem is then formulated as follows: Find a control u in $L^2(\Gamma)^n$ that maximizes the lift

$$-\int_{\Gamma_w} \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n}\right) \cdot \vec{e_l} \; \mathrm{d}s \tag{1.0.2}$$

subject to the steady state Navier-Stokes equations describing the motion of the fluid

$$\begin{aligned} -\nu\Delta u + (u \cdot \nabla)u + \nabla p &= 0 \quad \text{in } \Omega \\ \operatorname{div} u &= 0 \quad \text{in } \Omega \\ u &= g \quad \text{on } \Gamma_c, \\ u &= 0 \quad \text{on } \Gamma \setminus \Gamma_c, \end{aligned} \tag{1.0.3}$$

the convex control constraints

$$g(x) \in G \text{ a.e. on } \Gamma_c, \tag{1.0.4}$$

and the maximal drag constraint

$$-\int_{\Gamma_w} \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n}\right) \cdot \vec{e_d} \; \mathrm{d}s \leq D_0. \tag{1.0.5}$$

The boundary Γ_w is a curve satisfying

$$\int_{\Gamma_w} \mathfrak{n} \; \mathrm{d}s = 0. \tag{1.0.6}$$

As shown later, the pressure is only unique up to a constant. The constraint (1.0.6) avoids that this constant changes the objective functional arbitrarily. The control acts on a part of the boundary of the body $\Gamma_c \subset \Gamma_w$ and homogeneous Dirichlet boundary conditions are prescribed on the boundary $\Gamma \setminus \Gamma_c$.

The set of admissible controls G is a bounded, convex, closed, and nonempty subset of $\mathbb{R}^n$. Furthermore, we assume $0 \in G$, which gives us the option to turn off the control admissible in the optimization problem. For a more detailed discussion of such convex control constraints, we refer to [110].

Let us shortly review available literature on analysis of optimal control problems for the Navier-Stokes equations. Starting with Abergel and Temam [2] there is an ever growing list of contributions. Let us only mention the work by Gunzburger, Hou, and Svobodny [51], Gunzburger and Manservisi [52], Hinze and Kunisch [54, 55], Kunisch and de los Reyes [28], de los Reyes and Yousept [30], de los Reyes and Tröltzsch [29], Abergel and Casas [1], Casas [24, 23], Tröltzsch and Roubiček [84] and Wachsmuth [103]. Finite-element error estimates can be found in the work of Casas, Mateos and Raymond [20]. Optimal flow control problems with state constraints were studied by Griesse and Reyes [50], Reyes and Kunisch [80].

The novelty of the first part of this thesis is that it combines the use of very low regular boundary controls, i.e. in $L^2(\Gamma)$, and integral state constraints. There are only a few contributions to optimal control theory using Dirichlet controls in L^2, see for instance Kunisch and Vexler [61] and Casas, Mateos and Raymond [20]. In the context of steady-state Navier-Stokes equations this is a new and promising approach, since the use of L^2-controls yields localizable optimality conditions, whereas the use of, for instance, $H^{1/2}(\Gamma)$-controls yields optimality conditions containing non-local boundary operators.

In view of the low L^2-regularity of the controls, the boundary integrals (1.0.2) and (1.0.5) are no longer well-defined, since the velocity field u is not regular enough to admit traces on the boundary. Therefore, we transform the boundary integrals into volume integrals leading in case of the drag constraint to a non-standard mixed control-state constraint, see Section 3.2 below.

As it is well-known, the steady-state Navier-Stokes equations are solvable in suitable spaces. If the data and/or the Reynolds number $1/\nu$ are small enough then the solution will be unique. To judge whether this condition is fulfilled in a concrete application is a delicate issue in particular in the case of inhomogeneous boundary conditions, see the discussion in the monograph of Galdi [44]. Hence, instead of assuming smallness of the data, we assume non-singularity at the optimal control, which is equivalent to unique solvability of a certain linearized equation, see Section 2.4.2.

By assuming the existence of a linearized Slater point to the state constraint (1.0.5), we are able to prove first-order necessary optimality conditions, see Section 4.1. For the special case of smooth controls, the resulting optimality system simplifies considerably, see Section 4.3. Furthermore, we state a second-order sufficient optimality condition for the problem under consideration. The first part of this thesis is complemented by numerical experiments on a high-lift configuration. One numerical approach was to solve the associated nonsmooth first-order optimality system of two coupled Navier-Stokes equations. At the end of this topic, we also implemented an SQP method with a penalty term in the cost functional to handle the integral state constraint and proved its convergence.

Afterwards, the second part of the paper deals with a nonstationary problem, considering a model closer to a real setting, that accounts also for turbulence. Here, the flow is computed on the basis of a $k - \omega$ WILCOX98 turbulence model, see (7.0.2), including the nonstationary Navier-Stokes equations and high Reynolds numbers. There are many approaches of turbulence modeling. Let us just mention Reynolds-averaged Navier-Stokes (RANS) based models, one equation models and also two equation models, for instance the $k - \epsilon$ model and the $k - \omega$ models. For every group, one can find many variations, see for instance [114].

The curse of dimension and the inherent nonlinearity leads to very large computing times so that a mathematical optimization analogous to the stationary case is fairly unrealistic. In [19], a generic high-lift configuration was investigated and one forward solution of the turbulence model took about 48 hours. In the case of the SCCH configuration, which we consider here, the computation time was nearly twice that number. We think that model reduction is a method of choice to avoid such extremely long computation times. To this aim, many authors have considered proper orthogonal decomposition (POD), see e.g. [4, 62, 63, 111]. The problem is that, in this case, we have to insert the POD basis as a Galerkin basis in the full turbulence model, which is a time consuming task. Thus, we decided to follow an alternative approach by Luchtenburg, Noack et al. [66, 74] of building a low-order dynamical system based on uniform oscillation by parameter identification.

The original full optimization problem consists of the cost functional, calculating the lift, the nonstationary Navier-Stokes equations, the associated $k - \omega$ WILCOX98 turbulence model and of constraints on the control. In the nonstationary part, we discard for simplicity constraints on the drag. The control function is considered as a periodic function $t \mapsto g$,

$$g = B \cos(\omega t),$$

where B is the actuation amplitude and ω is the actuation frequency. In our problem, we consider only the actuation amplitude B as the optimization parameter. For further work, the actuation amplitude is one possible additional optimization parameter.

In Chapter 7, we consider our optimization problem in detail with the turbulence model and introduce the standard POD method.

Unfortunately, the POD method does not identify clear frequencies and amplitudes. Thus, in Chapter 8 we filter the POD mode coefficients in a way that we neglect fluctuations of frequencies and amplitudes to get clear structures and to calculate the associating modes. With the help of these coefficients, we build a reduced-order model (ROM), similar to Luchtenburg et al. [66], consisting only of four ODEs describing uniform oscillations. We identify the inherent parameters. Based on the results of the ROM, a formula for the lift calculation is created and so we are able to establish an optimization problem based on this new lift formula and the low-order dynamical system.

In Chapter 9, we establish an optimality system, consisting of an objective cost functional, calculating the lift of the aircraft based on the POD coefficients, the dynamical system of 4 ODEs as state equations and constraints on the control, to investigate the optimization problem numerically . We need only a fraction of time for one solution of the state equations compared to the full system while the optimal actuation amplitude is almost the same.

The main results of the stationary part, except the SQP-method and its convergence, have been published partially word for word in a joint work with D. Wachsmuth [58], while Chapter 8 is published partially word for word in a joint work with B.R. Noack, M. Schlegel, F. Tröltzsch and D. Wachsmuth [57].

Chapter 2

The steady-state Navier-Stokes equation

The Navier-Stokes equations are a mathematical model to describe the motion of fluid flow. Claude Louis Marie Henri Navier ($\star$10. February 1785 in Dijon; †21. August 1836 in Paris) was a French mathematician and physicist and Sir George Gabriel Stokes ($\star$13. August 1819 in Skreen, County Sligo; †1. February 1903 in Cambridge) was an Irish mathematician and physicist. They were the first trying to derive equations of motions for fluid.

The nonlinear incompressible steady state Navier-Stokes system with inhomogeneous Dirichlet boundary condition is given in its dimensionless form as follows

$$\begin{aligned} -\nu\Delta u + (u\cdot\nabla)u + \nabla p &= f \quad \text{in } \Omega \\ \operatorname{div} u &= 0 \quad \text{in } \Omega \\ u &= g \quad \text{on } \Gamma_c, \\ u &= 0 \quad \text{on } \Gamma \setminus \Gamma_c. \end{aligned} \tag{2.0.1}$$

Here, the velocity field is denoted by u and the pressure by p. In this section, we investigate the steady-state Navier-Stokes equations mathematically, this includes the spaces for the solutions. After that, we have to clarify in what sense the solutions are defined.

2.1 Function spaces

Let us first define the spaces of p-integrable functions and summarize some of their basic properties. The functions are defined on a domain $\Omega \subset \mathbb{R}^n$ with boundary Γ, which is assumed to be sufficiently smooth, I'll explain later.

Definition 2.1. *Let Ω be an open subset of $\mathbb{R}^n$ and $1 \leq p < \infty$. The set of* *p-integrable functions is defined as*

$$L^p(\Omega) = \{u : \Omega \to \mathbb{R}; u \textit{ is measurable and } \int\limits_\Omega |u|^p \mathrm{d}x < \infty\},$$

endowed with the norm

$$\|u\|_{L^p} = \{\int\limits_\Omega |u(x)|^p \mathrm{d}x\}^{1/p}.$$

For $p = \infty$*, we define*

$$L^\infty(\Omega) = \{u : \Omega \to \mathbb{R};\ u \textit{ is measurable and } |u(x)| \leq C \textit{ a.e. in } \Omega,\ C > 0\}$$

and introduce the norm

$$\|u\|_{L^\infty} = \inf\{C : |u(x)| \leq C \textit{ a.e. in } \Omega\}.$$

We will provide some basic inequalities to deal with Lebesgue integrable functions. At first, we have the well-known theorem:

Theorem 2.2 (Hölder)**.** *Let* $u \in L^p(\Omega)$ *and* $v \in L^q(\Omega)$ *with* $1 < p, q < \infty$ *and*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Then $uv \in L^1(\Omega)$ *and*

$$\|uv\|_{L^1(\Omega)} \leq \|u\|_{L^p(\Omega)} \|v\|_{L^q(\Omega)}.$$

The spaces $L^p(\Omega)$ are Banach spaces for $1 \leq p \leq \infty$ and reflexive for $1 < p < \infty$. In $L^2(\Omega)$, a scalar product can be defined by

$$(u, v)_{L^2(\Omega)} = \int\limits_\Omega uv \, \mathrm{d}x$$

and a Hilbert space structure is also obtained. The space of infinitely differentiable functions with compact support is denoted by $\mathcal{D}(\Omega)$ and its dual, the distributions space, by $\mathcal{D}'(\Omega)$.

The Sobolev space $W^{m,p}(\Omega)$ is the space of $L^p(\Omega)$ functions whose weak derivative up to order m is also in $L^p(\Omega)$: For these spaces a norm is introduced in the following way:

$$|u|_{W^{m,p}} := \left(\sum_{|j|<m} \|D^j u\|_{L^p}^p \right)^{1/p}.$$

In the case $p = 2$, the space $H^m(\Omega) := W^{m,2}(\Omega)$ is a Hilbert space with scalar product

$$(u, v)_{H^m} = \sum_{|j| \leq m} (Dju, Djv)_{L^2}.$$

The closure of $\mathcal{D}(\Omega)$ in the $W^{m,p}(\Omega)$ norm is denoted by $W_0^{m,p}(\Omega)$. For more details of Sobolev spaces and a proof for the two following Sobolev imbedding Theorems, we refer to [3].

Theorem 2.3 (Sobolev imbedding Theorem). *Let Ω be a domain in $\mathbb{R}^n$ with the cone property. Then the imbeddings*

$$\begin{aligned}
&W^{m,p}(\Omega) \hookrightarrow L^q(\Omega), \; \textit{for } mp < n \textit{ and } 1 \leq q < \frac{np}{n - mp}, \\
&W^{m,p}(\Omega) \hookrightarrow L^q(\Omega), \; \textit{for } mp = n \textit{ and } 1 \leq q < \infty, \\
&W^{m,p}(\Omega) \hookrightarrow C(\bar{\Omega}) \;, \; \textit{for } mp > n
\end{aligned}$$

are continuous.

In the two-dimensional case the imbeddings $H^1(\Omega) \hookrightarrow L^q(\Omega)$ for $1 \leq q < \infty$ and $W^{1,p}(\Omega) \hookrightarrow C(\bar{\Omega})$ for $p > 2$ are continuous.

Theorem 2.4. *Let Ω be a domain in $\mathbb{R}^n$ with $\mathcal{C}^1$-boundary.*

1. *Suppose $mp < n$ and $n - mp < n$. This leads to*

$$W^{j+m,p}(\Omega) \hookrightarrow W^{m,q}(\Omega)$$

for $p \leq q \leq np/(n - mp)$.

2. *Suppose $mp = n$. Then, we obtain*

$$W^{j+m,p}(\Omega) \hookrightarrow W^{j,q}(\Omega)$$

for $p \leq q < \infty$.

Inhomogeneous boundary values will be defined by the trace of $W^{k,p}(\Omega)$ on the boundary.

Theorem 2.5. *Let Ω be a bounded Lipschitz-domain and $1 \leq p \leq \infty$. Then there exists a linear and continuous mapping $\tau : W^{1,p}(\Omega) \to L^p(\Gamma)$ with*

$$(\tau u)(x) = u(x) \;\; \textit{a.e. on } \Gamma$$

for all $u \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$.

For a proof of this theorem and the next two, we refer to [3].

Definition 2.6. *The element τu is defined by the trace of u on Γ and the mapping is called the trace operator.*

Theorem 2.7. *Let $m \geq 1$, $m \in \mathbb{Z}$ and Γ of class $\mathcal{C}^{m-1,1}$. Then, we obtain that the trace operator τ is for $mp < n$ continuous from $W^{m,p}(\Omega)$ to $L^r(\Gamma)$, if $1 \leq r \leq \dfrac{(n-1)p}{n-mp}$. For $mp = n$, we get that τ is continuous for all $1 \leq r < \infty$.*

Another statement of [3] is:

Theorem 2.8. *Let Ω be of class $\mathcal{C}^m$, $m \geq 1$, $k \in \mathbb{Z}$, and $1 < p < \infty$. Then the trace operator τ is continuous from $W^{m,p}(\Omega)$ to $W^{m-1/p,p}(\Gamma)$.*

Because we have to deal with functions satisfying $\operatorname{div} u = 0$, we introduce the space

$$\mathcal{V} := \{u \in \mathcal{D}(\Omega)^n : \operatorname{div} u = 0\}$$

The closure of $\mathcal{V}$ in the $\mathbf{H}_0^1$-norm is denoted by V and if Ω is an open bounded Lipschitz set, it can be characterized as

$$V = \{u \in \mathbf{H}_0^1(\Omega) : \operatorname{div} u = 0\}.$$

We assume that the boundary of Ω is in C^2. The outer unit normal on Γ is denoted by $\mathfrak{n}$. The boundary Γ is the union of m connected components, $\Gamma = \bigcup_{j=1}^m \Gamma_j$.

Furthermore, we define the following spaces on Ω and Γ:

$$\begin{aligned}
\mathbf{H}^s(\Omega) &= \{v \in H^s(\Omega)^n : \operatorname{div} v = 0 \text{ on } \Omega,\ \langle u \cdot \mathfrak{n}, 1\rangle_{H^{-1/2}(\Gamma_j),H^{1/2}(\Gamma_j)} = 0 \\
&\qquad\qquad \forall j \in \{1,\dots,m\}\},\ s \geq 0,\\
\mathbf{H}_0^s(\Omega) &= \{v \in H^s(\Omega)^n : \operatorname{div} v = 0 \text{ on } \Omega,\ u = 0 \text{ on } \Gamma\},\ s \geq 1/2,\\
\mathbf{H}^s(\Gamma) &= \{v \in H^s(\Gamma) : \textstyle\int_{\Gamma_j} u \cdot \mathfrak{n} = 0\ \forall\ j = 1 \dots m\},\ s \geq 0,\\
\mathbf{L}^p(\Omega) &= \{v \in L^p(\Omega)^n : \operatorname{div} v = 0\},\ p \geq 1,\\
\mathbf{L}^p(\Gamma) &= L^p(\Gamma)^n,\ p \geq 1,\\
\mathbf{H}^{-s}(\Gamma) &= (\mathbf{H}^s(\Gamma))',\ s \geq 0,\\
\mathbf{H}^{-s}(\Omega) &= (\mathbf{H}^s(\Omega) \cap \mathbf{H}_0^1(\Omega))', s \geq 1\\
\mathbf{W}^{m,p}(\Omega) &= W^{m,p}(\Omega)^n,\\
\mathbf{W}^{m,p}(\Gamma) &= W^{m,p}(\Gamma)^n,
\end{aligned}$$

and the differential operators for vector-valued functions u and scalar-valued functions p:

$$\begin{aligned}
\Delta u \in \mathbb{R}^n &: (\Delta u)_i = \Delta u_i, && i = 1, \ldots, n, \\
\nabla p \in \mathbb{R}^n &: (\nabla p)_i = \frac{\partial}{\partial x_i} p, && i = 1, \ldots, n, \\
\frac{\partial}{\partial \mathfrak{n}} u = \mathfrak{n} \nabla u \in \mathbb{R}^n &: (\frac{\partial}{\partial \mathfrak{n}} u)_i = \frac{\partial}{\partial \mathfrak{n}} u_i, && i = 1, \ldots, n, \\
(u \cdot \nabla) u \in \mathbb{R}^n &: ((u \cdot \nabla) u)_i = \sum_{j=1}^{n} u_j \frac{\partial u_i}{\partial x_j}, && i = 1, \ldots, n.
\end{aligned}$$

Remark 2.9. *As mentioned later, the terms* $\langle u \cdot \mathfrak{n}, 1 \rangle_{H^{-1/2}(\Gamma_j), H^{1/2}(\Gamma_j)} = 0 \; \forall j \in \{1, \ldots, m\}$ *are important to obtain the existence of very weak solutions for arbitrary large data. The trace operator is only defined for* $s \geq 1/2$.

Before considering the solvability of the Navier-Stokes equations, we want to have a look on the Stokes equations.

2.2 The Stokes equations

The Stokes equations are similar to the Navier-Stokes ones, but without the nonlinear term. They are given by

$$\begin{aligned}
-\nu \Delta u + \nabla p &= f \quad \text{in } \Omega \\
\operatorname{div} u &= 0 \quad \text{in } \Omega \\
u &= g \quad \text{on } \Gamma_c, \\
u &= 0 \quad \text{on } \Gamma \setminus \Gamma_c.
\end{aligned} \tag{2.2.1}$$

Let us define the bilinear form $a(u, v) := \nu(\nabla u, \nabla v)$. Then we call u a weak solution of (2.2.1), if

$$a(u, v) = (f, v) \quad \forall v \in V$$

is satisfied. The next theorem guarantees existence and uniqueness of solutions for (2.2.1) and we refer to [27, 47, 97] for the theory.

Theorem 2.10. *For every* $f \in \mathbf{H}^{-1}(\Omega)$ *and* $g \in \mathbf{H}^{1/2}(\Gamma)$ *with*

$$\int_{\Gamma_c} g \cdot \mathfrak{n} \, \mathrm{d}s = 0$$

the Stokes equation (2.2.1) *has a unique solution* $(u, p) \in \mathbf{H}^1(\Omega) \times L^2(\Omega)$, *where* p *is unique up to a constant.*

The next theorem,see [44, Vol. 1, Thm. IV.6.1], shows that the solution (u, p) is more regular if the data (f, g) is more regular.

Theorem 2.11. *Let u be a solution of the Stokes problem* (2.2.1) *on a bounded domain $\Omega \subset \mathbb{R}^n$, $n \geq 2$ of class $\mathcal{C}^{m+2}$, $m \geq 0$ corresponding to*

$$f \in \mathbf{W}^{m,q}(\Omega), \quad g \in \mathbf{W}^{m+2-1/q,q}(\Gamma).$$

Then, we obtain

$$u \in \mathbf{W}^{m+2,q}(\Omega), \quad p \in W^{m+1,q}(\Omega).$$

Moreover, the following estimate holds

$$\|u\|_{W^{m+2,q}(\Omega)} + \|p\|_{W^{m+1,q}(\Omega)} \leq c(\|f\|_{W^{m,q}(\Omega)} + \|g\|_{W^{m+2-1/q,q}(\Gamma)}) \tag{2.2.2}$$

with a constant $c = c(m, n, q, \Omega)$.

Corollary 2.12. *The special case $q = 2$ and $m = 0$: For $f \in L^2(\Omega)$ and $g \in H^{3/2}(\Gamma)$, we obtain $u \in \mathbf{H}^2(\Omega)$ and $p \in H^1(\Omega)$ as solutions for the Stokes equations* (2.2.1)*. Let us define the associated control-to-state operators*

$$G : \mathbf{H}^{3/2}(\Gamma) \to \mathbf{H}^2(\Omega), \; g \mapsto u$$

with $f = 0$ and

$$S : \mathbf{L}^2(\Omega) \to \mathbf{H}^2(\Omega), \; f \mapsto u$$

with $g = 0$.

We even get the following theorem, see [26, Vol. 6, Thm. 1.11].

Theorem 2.13. *Let Ω be an open bounded set of $\mathbb{R}^n$, $n = 2, 3$, of class $\mathcal{C}^r$, $r = \max\{m + 2, 2\}$, $m \geq -1$, $m \in \mathbb{Z}$. Let*

$$f \in \mathbf{W}^{m,q}(\Omega) \quad g \in \mathbf{W}^{m+1,q}(\Gamma)$$

satisfy the compatibility condition

$$\int_{\Gamma_c} g \cdot \mathfrak{n} \, \mathrm{d}s = 0.$$

Then there exists a unique solution

$$(u, p) \in \mathbf{W}^{m+2,q}(\Omega) \times W^{m+1,q}(\Omega)$$

of the Stokes problem (2.2.1) *(p is unique up to a additive constant), satisfying* (2.2.2)*.*

Let us specify, in what sense we want to solve the Navier-Stokes equations. The first and standard way is to consider weak solutions.

2.3 Weak formulation

See [44] for more details on weak solutions of (1.0.3). For simplicity, we define some functions $b : \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) \to \mathbb{R}$

$$b(u, v, w) := ((u \cdot \nabla)v, w)_{\mathbf{L}^2(\Omega)} \tag{2.3.1}$$

and $B : \mathbf{H}^1(\Omega) \mapsto (\mathbf{H}^1(\Omega))'$ for $u, v \in \mathbf{H}^1(\Omega)$

$$\langle B(u), v\rangle_{(\mathbf{H}^1(\Omega))', \mathbf{H}^1(\Omega)} := b(u, u, v). \tag{2.3.2}$$

Due to to the quadratic character of B, its differentiability is easily to see. The first Fréchet derivative $B'(\bar{u})u$ of B with respect to $\bar{u}$ is a functional of $(\mathbf{H}^1(\Omega))'$ and has the following form

$$\langle B'(\bar{u})u, v\rangle_{(\mathbf{H}^1(\Omega))', \mathbf{H}^1(\Omega)} = b(\bar{u}, u, v) + b(u, \bar{u}, v)$$

with $v \in \mathbf{H}^1(\Omega)$. The second Fréchet derivative $B''(\bar{u})[\tilde{u}, \hat{u}]$ is given by

$$\langle B''(\bar{u})[\tilde{u}, \hat{u}], v\rangle_{(\mathbf{H}^1(\Omega))', \mathbf{H}^1(\Omega)} = b(\tilde{u}, \hat{u}, v) + b(\hat{u}, \tilde{u}, v).$$

Additionally, we get the following properties for the nonlinearity b and B, see [55],

Lemma 2.14. *The nonlinear term b and B fulfills*

1. $|b(u, v, w)| \leq c\|u\|^{1/2}_{\mathbf{L}^2(\Omega)}\|u\|^{1/2}_{\mathbf{H}^2(\Omega)}\|v\|_{\mathbf{H}^1(\Omega)}\|w\|_{\mathbf{L}^2(\Omega)}, \ \forall u \in \mathbf{H}^2(\Omega), v \in \mathbf{H}^1(\Omega), w \in \mathbf{L}^2(\Omega),$
2. $|b(u, v, w)| \leq c\|u\|_{\mathbf{H}^1(\Omega)}\|v\|_{\mathbf{H}^1(\Omega)}\|w\|_{\mathbf{H}^1(\Omega)}, \ \forall u, v, w \in \mathbf{H}^1(\Omega),$
3. $\langle B''(\bar{u})[\tilde{u}, \hat{u}], v\rangle_{(\mathbf{H}^1(\Omega))', \mathbf{H}^1(\Omega)} = \langle B'(\tilde{u})\hat{u}, v\rangle_{(\mathbf{H}^1(\Omega))', \mathbf{H}^1(\Omega)}$
4. $\frac{1}{2}\langle B''(\bar{u})[u, u], v\rangle_{(\mathbf{H}^1(\Omega))', \mathbf{H}^1(\Omega)} = \langle B(u), v\rangle_{(\mathbf{H}^1(\Omega))', \mathbf{H}^1(\Omega)}$ *for* $u = \tilde{u} = \hat{u}$

with a constant $c \in \mathbb{R}$.

Multiplying (1.0.3) by test functions $(v, q) \in V \times \mathbf{L}^2(\Omega)$, we obtain by partial integration the following weak formulation

$$a(u, v) + b(u, u, v) = \langle f, v\rangle_{V', V} \qquad \text{in } \Omega \tag{2.3.3}$$

$$\tau u = g \qquad \text{on } \Gamma_c \tag{2.3.4}$$

where $\tau : \mathbf{H}^1 :\to \mathbf{H}^{\frac{1}{2}}(\Gamma_c)$ is the trace operator.

Theorem 2.15. *Let $\Omega \in \mathbb{R}^n$ be a bounded locally Lipschitz domain of $\mathbb{R}^n$, $n = 2, 3$ with $\Gamma := \partial\Omega$ constituted by $m+1$ connected components $\Gamma_1, \ldots, \Gamma_{m+1}$, $m \geq 0$,*

$$f \in \mathbf{H}^{-1}(\Omega) \text{ and } g \in \mathbf{H}^{1/2}(\Gamma_c)$$

with

$$\int_{\Gamma_c} g \cdot \mathfrak{n} \, \mathrm{d}s = 0.$$

Then, there exists under additional assumptions, see [44, Vol. 2, Thm. VIII.4.1, Thm. VIII.4.2], at least one solution $(u, p) \in \boldsymbol{H}^1(\Omega) \times L^2(\Omega)$ for (2.0.1) satisfying the estimate

$$\|p\|_{L^2(\Omega)} \leq c(\|f\|_{\mathbf{H}^{-1}(\Omega)} + \|u\|^2_{\mathbf{H}^1(\Omega)} + \nu\|u\|_{\mathbf{H}^1(\Omega)}).$$

Furthermore, there exists $c_1 = c_1(n, \Omega)$ such that if

$$\|g\|_{1/2,2(\Gamma)} \leq c_1\nu/2,$$

u verifies

$$\|u\|_{\mathbf{H}^1(\Omega)} \leq c_2(\|f\|_{\mathbf{H}^{-1}(\Omega)} + \|g\|^2_{\mathbf{H}^{\frac{1}{2}}(\Gamma)} + (1+\nu)\|g\|_{\mathbf{H}^{\frac{1}{2}}(\Gamma)}).$$

with $c_2 = c_2(n, \Omega)$. If

$$\|f\|_{\mathbf{H}^{-1}(\Omega)} + \|g\|^2_{\mathbf{H}^{\frac{1}{2}}(\Gamma)} + (1+\nu)\|g\|_{\mathbf{H}^{\frac{1}{2}}(\Gamma)} < c_3\nu^2$$

is additionally satisfied with $c_3 = \min\{c_1, 1/c_2k\}$, then u is unique and p is unique up to a constant.

Proof. For the proof and the additional assumptions, we refer to [44, Vol. 2, Thm. VIII.4.1, Thm. VIII.4.2]. □

We can obtain some extra regularity of the solution (u, p), if the right hand side is smooth enough, see [44, Vol. 2, Thm. VIII.5.2]:

Theorem 2.16. *Let Ω be a bounded domain of $\mathbb{R}^n$, $n \geq 2$, of class $\mathcal{C}^2$. Let*

$$u \in \mathbf{W}^{1,2}(\Omega) \cap \mathbf{L}^n(\Omega)$$

be divergence-free, satisfying (2.3.3) for some $p \in L^2(\Omega)$ and for all $v \in C_0^\infty(\Omega)$. Then, if

$$f \in \mathbf{L}^q(\Omega), \quad g \in \mathbf{W}^{2-1/q,q}(\Gamma_c)$$

with

$$q \in (1, \infty), \quad \text{if } n = 2,$$

while

$$q \in (2n/(n+2), \infty), \quad \textit{if } n > 2,$$

it follows that

$$(u,p) \in \mathbf{W}^{2,q}(\Omega) \times W^{1,q}(\Omega).$$

Moreover, if Ω *is of class* $\mathcal{C}^{m+2}$ *and*

$$f \in \mathbf{W}^{m,q}(\Omega), \quad g \in \mathbf{W}^{m+2-1/q,q}(\Gamma_c)$$

with $m \geq 1$ *and*

$$q \in (1, \infty), \quad \textit{if } n = 2,$$

while

$$q \in (n/2, \infty), \quad \textit{if } n > 2,$$

then

$$(u,p) \in \mathbf{W}^{m+2,q}(\Omega) \times W^{m+1,q}(\Omega).$$

We obtain by Theorem 2.16 with $q = 2$:

Corollary 2.17. *Let* $f \in \mathbf{L}^2(\Omega)$ *and* $g \in \mathbf{H}^{3/2}(\Gamma_c)$ *. Then, it follows that* $(u,p) \in \mathbf{H}^2(\Omega) \times H^1(\Omega)$*.*

2.4 Very weak formulation

Here, we have to resort to the notation of very weak solutions, since in general weak solutions in $u \in \mathbf{H}^1(\Omega)$ do not exist due to the desired low regularity $\mathbf{L}^2(\Gamma_c)$ of boundary data. The theory in this section is based on [69, 40, 39].

Definition 2.18. *Let* $g \in \mathbf{H}^0(\Gamma)$ *be given. Then we call* $u \in \mathbf{L}^{2n/(n-1)}(\Omega)$ *a* very weak solution *of the state equation* (2.0.1) *if for all test functions* $v \in \mathbf{H}_0^2(\Omega)$, $\pi \in H^1(\Omega)$ *it holds*

$$\int_\Omega (u \cdot (-\nu \Delta v) - (u \cdot \nabla) v u) \, \mathrm{d}x + \int_\Gamma g \cdot \nu \frac{\partial v}{\partial \mathfrak{n}} \, \mathrm{d}s = \langle f, v \rangle_{H^{-1}(\Omega), H^1(\Omega)} \tag{2.4.1a}$$

and

$$\int_\Omega \nabla \pi \cdot u \, \mathrm{d}x - \int_\Gamma (g \cdot \mathfrak{n}) \pi \, \mathrm{d}s = 0. \tag{2.4.1b}$$

Here, the first equation is obtained by partially integrating the Navier-Stokes equations twice and using the equation:

$$\begin{aligned}\int_\Omega (u\cdot\nabla)u\cdot v\;\mathrm{d}x &= \sum_{i,j=1}^n \int_\Omega u_i \frac{\partial u_j}{\partial x_i} v_j\;\mathrm{d}x\\ &= \sum_{i,j=1}^n -\int_\Omega u_i \frac{\partial v_j}{\partial x_i} u_j + \frac{\partial u_i}{\partial x_i} v_j u_j\;\mathrm{d}x + \int_\Gamma u_i u_j \mathfrak{n}_i v_j\;\mathrm{d}s\\ &= -\int_\Omega (u\cdot\nabla)v\cdot u + u\cdot v \operatorname{div} u\;\mathrm{d}x + \int_\Gamma (u\cdot\mathfrak{n})u\cdot v\;\mathrm{d}s\end{aligned}$$

Then, we can reformulate the first equation of (2.0.1) to (2.4.1a) by

$$\begin{aligned}&\int_\Omega (-\nu\Delta u + (u\cdot\nabla)u + \nabla p)v\;\mathrm{d}x - \langle f, v\rangle_{H^{-1}(\Omega),H^1(\Omega)}\\ =&\int_\Omega \nu\nabla u\cdot\nabla v - (u\cdot\nabla)v\cdot u + u\cdot v\operatorname{div} u + p\cdot\operatorname{div} v\;\mathrm{d}x\\ &+\int_\Gamma (u\cdot\mathfrak{n})u\cdot v - \nu\frac{\partial u}{\partial\mathfrak{n}} v\;\mathrm{d}s - \langle f, v\rangle_{H^{-1}(\Omega),H^1(\Omega)}\\ =&\int_\Omega u\cdot(-\nu\Delta v) - (u\cdot\nabla)v\cdot u\;\mathrm{d}x + \int_\Gamma g\cdot\nu\frac{\partial v}{\partial\mathfrak{n}}\;\mathrm{d}s - \langle f, v\rangle_{H^{-1}(\Omega),H^1(\Omega)}.\end{aligned}$$

The second equation (2.4.1b) is the weak formulation of $\operatorname{div} u = 0$:

$$\begin{aligned}\int_\Omega \operatorname{div} u\pi\mathrm{d}x &= -\int_\Omega u\cdot\nabla\pi\mathrm{d}x + \int_\Gamma (u\cdot\mathfrak{n})\pi\;\mathrm{d}s\\ &= -\int_\Omega u\cdot\nabla\pi\mathrm{d}x + \int_{\Gamma_c} (g\cdot\mathfrak{n})\pi\;\mathrm{d}s.\end{aligned}$$

Moreover as discussed in [40], it incorporates the Dirichlet boundary condition for the normal component of u, since the term $\int_\Gamma g\cdot\nu\frac{\partial v}{\partial\mathfrak{n}}\;\mathrm{d}s$ acts only on tangential components. For example, let $n=3$ and and v be a smooth function. Then it holds $\frac{\partial v}{\partial\mathfrak{n}}(curlv)\times\mathfrak{n}$, which implies that $\frac{\partial v}{\partial\mathfrak{n}}$ is orthogonal to the outer normal, and thus the product $\int_\Gamma u\cdot\nu\frac{\partial v}{\partial\mathfrak{n}}\;\mathrm{d}s$ in (2.4.1a) acts only on the tangential component of u.

The existence of very weak solutions with inhomogeneous Dirichlet boundary conditions is discussed in [39, 40, 45, 69]. We remark that it is essential to have $\int_{\Gamma_i} g\cdot\mathfrak{n}\;\mathrm{d}s = 0$ for all connected components of Γ to obtain existence of solutions for arbitrary large data. For boundary data in $\mathbf{H}^0(\Gamma)$ it holds the following.

Remark 2.19. *If the data are regular and the problem has a variational solution* $(u,p)\in\mathbf{H}^1(\Omega)\times L^2(\Omega)$*, then it is easy to see that the variational solution is also a very weak solution.*

In the next theorem, we investigate the linearized state equation, see [69, Theorem 3].

Theorem 2.20. *Let $g \in \mathbf{L}^2(\Gamma_c)$ and $z \in \mathbf{L}^{2n/(n-1)}(\Omega)^2$. Then the linearized problem*

$$\begin{aligned} -\nu\Delta u + (z \cdot \nabla)u + \nabla p = f & \quad \text{in } \Omega \\ \operatorname{div} u = 0 & \quad \text{in } \Omega \\ u = g & \quad \text{on } \Gamma_c \end{aligned}$$

has a unique very weak solution $u \in \mathbf{L}^{2n/(n-1)}(\Omega)$ and additionally, we obtain

$$\|u\|_{\mathbf{L}^{2n/(n-1)}(\Omega)} < c(1 + \|z\|_{\mathbf{L}^{2n/(n-1)}(\Omega)})\|g\|_{\mathbf{L}^2(\Gamma_c)}.$$

Let us now formulate the main result of this section.

Theorem 2.21. *For every $f \in \mathbf{H}^{-1}(\Omega)$ and $g \in \mathbf{H}^0(\Gamma)$, there exists a very weak solution $u \in \mathbf{L}^{2n/(n-1)}(\Omega)$ of* (2.0.1). *In the two-dimensional case, this solution belongs to $\mathbf{H}^{1/2}(\Omega)$.*

The existence proof and a discussion of the smallness assumption of the data and/or the Reynolds number $Re = 1/\nu$ can be found in [69]. The $\mathbf{H}^{1/2}$-regularity for the 2d-case result can be proven following the lines of [69]. Unique solvability with respect to less regular data, i.e. in $W^{-1/q,q}(\Gamma)$ is investigated in the articles by Farwig, Galdi, Sohr [39, 40, 45]. Once existence of a solution is proven, the pressure field p can be reconstructed by means of De Rham's Lemma, see for instance [97].

One can find in [69] that there exists a distribution $p \in W^{-1,2n/(n-1)}$ such that

$$-\Delta u - (u \cdot \nabla)u + \nabla p = 0$$

holds in the sense of distributions.

In view of the existence result, let us define for abbreviation the state space

$$\mathbf{U} := \mathbf{L}^{2n/(n-1)}(\Omega).$$

2.4.1 More regular solutions

Let us briefly show that more regular boundary data in $\mathbf{L}^p(\Gamma)$ yields more regular solutions. In the following considerations, we will split the state u in two parts, $u = u_0 + u_1$. The function u_0 is defined as the unique very weak solution to the Stokes equation with inhomogeneous Dirichlet boundary conditions

$$\begin{aligned} -\nu\Delta u_0 + \nabla p_0 = 0 & \quad \text{in } \Omega \\ \operatorname{div} u_0 = 0 & \quad \text{in } \Omega \\ u_0 = g & \quad \text{on } \Gamma. \end{aligned} \tag{2.4.2}$$

Then $u_1 = u - u_0$ solves the following equation subject to homogeneous Dirichlet boundary conditions

$$\begin{aligned} -\nu\Delta u_1 + (u\cdot\nabla)u_1 + \nabla p_1 &= -(u\cdot\nabla)u_0 && \text{in } \Omega\\ \operatorname{div} u_1 &= 0 && \text{in } \Omega\\ u_1 &= 0 && \text{on } \Gamma. \end{aligned} \tag{2.4.3}$$

As one can easily see, both systems are uniquely solvable. At first, let us prove higher regularity of u_0.

Lemma 2.22. *Let $g \in \mathbf{L}^p(\Gamma) \cap \mathbf{H}^0(\Gamma)$, $p \ge 2$, be given. Then the solution of* (2.4.2) *satisfies $u_0 \in \mathbf{L}^q(\Omega)$, where q is given by*

$$q = \begin{cases} \frac{np}{n-1} & \textit{if } 2 \le p < \infty,\\ +\infty & \textit{if } p = \infty, n = 3. \end{cases} \tag{2.4.4}$$

Proof. The mapping $g \mapsto u_0$ is linear and continuous from $\mathbf{W}^{-1/p,p}(\Gamma)$ to $\mathbf{L}^p(\Omega)$ and from $\mathbf{W}^{1-1/p,p}(\Gamma)$ to $\mathbf{W}^{1,p}(\Omega)$, see e.g. [12, 25]. By interpolation arguments, we have continuity of this solution mapping from $\mathbf{L}^p(\Gamma)$ to $\mathbf{W}^{1/p,p}(\Omega)$. The claim follows by the imbedding argument $\mathbf{W}^{1/p,p}(\Omega) \hookrightarrow \mathbf{L}^q(\Omega)$, see Theorem 2.3. The result for $p = \infty$, $n = 3$ can be found in [91]. □

Applying this result, we can prove higher regularity of the function u_1 and in consequence of the solution u of the nonlinear system.

Lemma 2.23. *If the boundary data g is in $\mathbf{L}^p(\Gamma) \cap \mathbf{H}^0(\Gamma)$, $p \ge 2$ $(n = 2)$ or $p \ge 4$ $(n = 3)$, then u belongs to $\mathbf{L}^q(\Omega)$ with q given by* (2.4.4).

Proof. Let us first consider the 2d-case, $n = 2$. Then we have by Theorem 2.21 $u \in \mathbf{L}^4(\Omega)$. This implies $(u\cdot\nabla)u \in \mathbf{H}^{-1}(\Omega)$, hence $u_1 = u - u_0$ solves

$$\begin{aligned} -\nu\Delta u_1 + \nabla p_1 &= -(u\cdot\nabla)u && \text{in } \Omega\\ \operatorname{div} u_1 &= 0 && \text{in } \Omega\\ u_1 &= 0 && \text{on } \Gamma, \end{aligned} \tag{2.4.5}$$

and we have $u_1 \in \mathbf{H}^1_0(\Omega) \hookrightarrow \mathbf{L}^t(\Omega)$ for all $0 < t < \infty$.

In the 3d-case, $n = 3$, Theorem 2.21 gives $u \in \mathbf{L}^3(\Omega)$. Since $p \ge 4$ by assumption, Lemma 2.22 yields $u_0 \in L^{3p/2}(\Omega) \subset L^6(\Omega)$. Taking $v \in \mathbf{H}^1_0(\Omega)$, we find after integration by parts

$$|\int_\Omega (u\cdot\nabla)u_0 v \,\mathrm{d}x| = |-\int_\Omega (u\cdot\nabla)v u_0 \,\mathrm{d}x| \le c\|\nabla v\|_{\mathbf{L}^2(\Omega)}\|u_0\|_{\mathbf{L}^6(\Omega)}\|u\|_{\mathbf{L}^3(\Omega)}.$$

Then $(u\cdot\nabla)u_0$ belongs to $\mathbf{H}^{-1}(\Omega)$, and the solution $u_1 = u - u_0$ of (2.4.3) belongs to $\mathbf{H}^1_0(\Omega) \hookrightarrow \mathbf{L}^6(\Omega)$. Hence $u = u_0 + u_1$ is in $\mathbf{L}^6(\Omega)$ as well. This in turn gives $(u\cdot\nabla)u \in W^{-1,r}(\Omega)$, $r = \frac{6p}{p+4} \geq 4$ with the following argument:

$$|\int_\Omega (u\cdot\nabla)u_0 v \,\mathrm{d}x| \leq c\|\nabla v\|_{\mathbf{L}^{r'}(\Omega)}\|u_0\|_{\mathbf{L}^{3p/2}(\Omega)}\|u\|_{\mathbf{L}^6(\Omega)}.$$

By [25], the function u_1 as solution of (2.4.5) belongs then to $\mathbf{W}^{1,4}(\Omega) \hookrightarrow \mathbf{L}^\infty(\Omega)$.

In both cases, ($n = 2, 3$), the function u_1 is as regular as u_0, hence the solution u belongs to the space $\mathbf{L}^q(\Omega)$ with q as in (2.4.4). □

2.4.2 Regularity assumption

It is well known that the stationary Navier-Stokes equations are uniquely solvable if the data/the control function g is small, see for example [69, Theorem 4].

Hence, if we want to have a unique response u to each control g we would have to impose restrictions on the control to enforce uniqueness of solutions. This technique is widely employed in optimal control of the stationary Navier-Stokes equations, see e.g. [50, 80, 81, 84, 103]. We will however proceed without a smallness assumption and therefore with non-uniqueness of the solutions. Since we allow multiple solutions of the state equation, we have to clarify the meaning of optimality.

Definition 2.24. *A pair $(\bar u, \bar g)$ is called* locally optimal*, if there exist $\rho_u, \rho_g > 0$ such that $J(\bar u, \bar g) \leq J(u, g)$ holds for all admissible pairs (u, g) with $\|u - \bar u\|_{\mathbf{U}} < \rho_u$ and $\|g - \bar g\|_{\mathbf{L}^2(\Gamma_c)} < \rho_g$.*

Here, a pair (u, g) is admissible *if it satisfies the constraints* (1.0.3)–(1.0.5).

Instead of enforcing uniqueness of solutions for all controls, we will impose the following regularity condition on an optimal state. A similar assumption is used to derive error estimates for distributed control in [20].

Definition 2.25. *A pair $(\bar u, \bar g) \in \mathbf{U} \times \mathbf{H}^0(\Gamma)$ is called* non-singular*, if the linearized Navier-Stokes equation*

$$\begin{aligned} -\nu\Delta u + B'(\bar u)u + \nabla p &= f \quad \text{in } \Omega \\ \operatorname{div} u &= 0 \quad \text{in } \Omega \\ u &= g \quad \text{on } \Gamma, \end{aligned} \tag{2.4.6}$$

admits a unique very weak solution $u \in \mathbf{U}$ *for all* $g \in \mathbf{H}^0(\Gamma)$ *and* $f \in \mathbf{H}^{-1}(\Omega)$*. Moreover, we assume that the solution mapping* $g \mapsto u$ *for* $f = 0$ *is linear and continuous from* $\mathbf{H}^0(\Gamma)$ *to* $\mathbf{H}^{1/2}(\Omega)$*, and the mapping* $f \mapsto u$ *for* $g = 0$ *is linear and continuous from* $\mathbf{H}^{-1}(\Omega)$ *to* $\mathbf{H}^1_0(\Omega)$*.*

This condition is fulfilled, if the state $\bar{u}$ is small enough [79, Lemma B.1]. The assumption of non-singularity implies that the state equation can be solved uniquely in a neighborhood of the reference/optimal control and state, confer [20, Theorem 2.5] for a proof using an implicit function theorem.

Theorem 2.26. *Let* $(\bar{u}, \bar{g}) \in \mathbf{U} \times \mathbf{H}^0(\Gamma)$ *be a non-singular solution of* (1.0.3)*. Then there exist an open neighborhood* $\mathcal{O}(\bar{g})$ *of* $\bar{g}$ *in* $\mathbf{H}^0(\Gamma)$*, an open neighborhood* $\mathcal{O}(\bar{u})$ *of* $\bar{u}$ *in* $\mathbf{U}$*, and a mapping* S *from* $\mathcal{O}(\bar{g}) \subset \mathbf{H}^0(\Gamma)$ *to* $\mathcal{O}(\bar{u}) \subset \mathbf{U} = \mathbf{L}^{2n/(n-1)}(\Omega)$ *of class* C^2 *such that, for all* $g \in \mathcal{O}(\bar{g})$*,* $S(g) = u$ *is the unique very weak solution in* $\mathcal{O}(\bar{u})$ *of* (1.0.4)*.*

Furthermore, the action of the first Fréchet derivative $u = S'(\bar{g})g$ *is given by the unique very weak solution of the linearized equation* (2.4.6)*.*

Chapter 3

The optimal control problem

Since we will work with very weak solutions $u \in \mathbf{U} = \mathbf{L}^{2n/(n-1)}(\Omega)$, we have to clarify the meaning of the boundary integrals in the objective functional (1.0.2) and the state constraint (1.0.3). These would be well-defined if the regularity $\frac{\partial u}{\partial \mathfrak{n}} \in L^1(\Gamma)$ could be guaranteed. This is not fulfilled for very weak solutions from $\mathbf{H}^{1/2}(\Omega)$ or $\mathbf{U}$. We will thus extend the linear functionals in (1.0.2) and (1.0.3) to the larger space $\mathbf{U}$. The idea of reformulating the functionals in this way arises by reading the paper of Braack and Richter [18].

In the first section of this chapter, we introduce reformulated boundary integrals for (1.0.2) and (1.0.3) with once partial integration. They are well-defined for $u \in \mathbf{H}^1(\Omega)$ and we need them in Section 4.3. Here, we consider $g \in G$ and the associated state $u \in \mathbf{U}$. Therefore, we have to reformulate (1.0.2) and (1.0.3) again with partially integrating twice in the second section.

3.1 Reformulation of the boundary integrals

Let us assume for a while that $u, p \in C^1(\Omega)^n$ are classical solutions of (1.0.4) to the control g. Multiplying the state equations (1.0.4) with a function $\varphi_i \in \mathbf{H}^1(\Omega)$, we obtain by partial integration

$$\begin{aligned} 0 &= (-\nu\Delta u + (u\cdot\nabla)u + \nabla p, \varphi_i) \\ &= \int_\Omega (\nu\nabla u \cdot \nabla\varphi_i + (u\cdot\nabla)u\varphi_i)\,\mathrm{d}x - \int_\Gamma \left(\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n}\right)\varphi_i\,\mathrm{d}s. \end{aligned}$$

In order to represent the functionals in (1.0.2), (1.0.3), let us introduce functions φ_i, $i \in \{d, l\}$, that take suitable chosen boundary values. Let φ_i denote the weak solutions of the Stokes equations

$$\begin{aligned} -\Delta\varphi_i + \nabla\pi_i &= 0 \quad \text{in } \Omega \\ \operatorname{div}\varphi_i &= 0 \quad \text{in } \Omega \end{aligned} \tag{3.1.1}$$

with boundary values

$$\varphi_l = \begin{cases} \vec{e}_l & \text{on } \Gamma_w \\ 0 & \text{on } \Gamma\backslash\Gamma_w \end{cases}, \quad \varphi_d = \begin{cases} \vec{e}_d & \text{on } \Gamma_w \\ 0 & \text{on } \Gamma\backslash\Gamma_w \end{cases}.$$

Using the transformation above, we can write for $i \in \{d, l\}$.

$$\begin{aligned} -\int_{\Gamma_w} \left(\nu\frac{\partial u}{\partial\mathfrak{n}} - p\mathfrak{n}\right)\vec{e}_i \,\mathrm{d}s &= -\int_{\Gamma} \left(\nu\frac{\partial u}{\partial\mathfrak{n}} - p\mathfrak{n}\right)\varphi_i \,\mathrm{d}s \\ &= \int_{\Omega} (\nu\nabla u \cdot \nabla\varphi_i + (u\cdot\nabla)u\varphi_i)\,\mathrm{d}x. \end{aligned}$$

Taking the right-hand side of this expression, we define the functional

$$\hat{F}_{\vec{e}_i}(u) = \int_{\Omega} (\nu\nabla u \cdot \nabla\varphi_i + (u\cdot\nabla)u\varphi_i)\,\mathrm{d}x \tag{3.1.2}$$

for $i \in \{d, l\}$. This function is well-defined for $u \in \mathbf{H}^1(\Omega)$, but for $u \in \mathbf{U}$ we have to reformulate the boundary integrals once again.

3.2 Further reformulation of the boundary integrals

Let us now assume that $u \in C^2(\Omega)^n, p \in C^1(\Omega)^n$ are a classical solution of (1.0.4) to g. Multiplying the state equation (1.0.4) with $\varphi_i \in \mathbf{H}^2(\Omega)$, we obtain by partially integrating twice

$$\begin{aligned} 0 &= (-\nu\Delta u + (u\cdot\nabla)u + \nabla p, \varphi_i) \\ &= \int_{\Omega} (\nu\nabla u \cdot \nabla\varphi_i + (u\cdot\nabla)u\varphi_i)\,\mathrm{d}x - \int_{\Gamma} \left(\nu\frac{\partial u}{\partial\mathfrak{n}} - p\mathfrak{n}\right)\varphi_i \mathrm{d}s \\ &= \int_{\Omega} (-\nu u \cdot \Delta\varphi_i - (u\cdot\nabla)\varphi_i u)\,\mathrm{d}x \\ &\quad + \int_{\Gamma} \left(\nu u\frac{\partial\varphi_i}{\partial\mathfrak{n}} + (u\cdot\mathfrak{n})(\varphi_i\cdot u)\right)\mathrm{d}s \quad - \int_{\Gamma} \left(\nu\frac{\partial u}{\partial\mathfrak{n}} - p\mathfrak{n}\right)\varphi_i \,\mathrm{d}s, \end{aligned}$$

where φ_i, $i \in \{d,l\}$ are defined as in the section before. Here, we need $\varphi_i \in \mathbf{H}^2(\Omega)$, $i \in \{d,l\}$. Using the transformation above, we can write for $i \in \{d,l\}$.

$$\begin{aligned}-\int_{\Gamma_w}\left(\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n}\right)\cdot \vec{e}_i \,\mathrm{d}s &= -\int_{\Gamma}\left(\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n}\right)\varphi_i \,\mathrm{d}s\\ &= \int_\Omega (\nu u\cdot\Delta\varphi_i + (u\cdot\nabla)\varphi_i u)\,\mathrm{d}x\\ &\qquad - \int_\Gamma\left(\nu u\frac{\partial\varphi_i}{\partial\mathfrak{n}} + (u\cdot\mathfrak{n})(\varphi_i\cdot u)\right)\mathrm{d}s.\end{aligned}$$

Taking the right-hand side of this expression, we define the functional

$$\tilde{F}_i(u) = \int_\Omega (\nu u\cdot\Delta\varphi_i + (u\cdot\nabla)\varphi_i u)\,\mathrm{d}x - \int_\Gamma\left(\nu u\frac{\partial\varphi_i}{\partial\mathfrak{n}} + (u\cdot\mathfrak{n})(\varphi_i\cdot u)\right)\mathrm{d}s.$$

This function is well-defined for $u \in \mathbf{H}^s(\Omega)$, $s > 1/2$, but not for $u \in \mathbf{U}$. To handle this problem, we substitute the state u by the control function $u = g$ on the boundary. This yields

$$\tilde{F}_{\vec{e}_i}(u,g) = \int\limits_\Omega (\nu u\cdot\Delta\varphi_i + (u\cdot\nabla)\varphi_i u)\,\mathrm{d}x - \int\limits_\Gamma\left(\nu g\frac{\partial\varphi_i}{\partial\mathfrak{n}} + (g\cdot\mathfrak{n})(\varphi_i\cdot g)\right)\mathrm{d}s. \tag{3.2.1}$$

In contrast to (1.0.2) and (1.0.3), the function $\tilde{F}_{\vec{e}_i}$, $i \in \{d,l\}$, is well-defined for states $u \in \mathbf{U}$ and controls $g \in \mathbf{H}^0(\Gamma)$, since the functions φ are very regular in comparison to the very weak solutions in $\mathbf{U}$. Their boundary values are in fact a constant vector, thus, the regularity of φ_i is only influenced by the regularity of the boundary Γ. But we have to note that the control now appears nonlinearly in the functionals which leads to additional difficulties for the optimal control problem, see Subsection 3.3.1.

The following result can be deduced from Theorem 2.11:

Lemma 3.1. *The functions φ_i, $i \in \{d,l\}$, belong to $\mathbf{H}^2(\Omega) \cap \mathbf{W}^{2,p}(\Omega)$ for all $p < \infty$.*

Then we can prove a continuity and differentiability statement for f_i.

Lemma 3.2. *The functions $\tilde{F}_{\vec{e}_i}$ are continuous and twice Fréchet- differentiable from $\mathbf{U} \times \mathbf{H}^0(\Gamma)$ to $\mathbb{R}$. Moreover, for given $u \in \mathbf{L}^q(\Omega)$ and $g \in \mathbf{L}^p(\Gamma)$, $p, q \in (1, +\infty)$, it holds $\tilde{F}'_{e_i,u}(u,g) \in \mathbf{L}^q(\Omega)$ and $\tilde{F}'_{e_i,g}(u,g) \in \mathbf{L}^p(\Omega)$.*

Proof. We get the continuity and the Fréchet-derivatives by standard arguments. Because of Lemma 3.1, φ_i, $i \in \{d,l\}$ possesses enough regularity to give $\tilde{F}'_{e_i,u}(u,g) \in \mathbf{L}^q(\Omega)$ and $\tilde{F}'_{e_i,g}(u,g) \in \mathbf{L}^p(\Omega)$. □

3.3 The optimal control problem

Now, we are able to reformulate the original optimal control problem (1.0.2)–(1.0.5) using the very weak solutions and the extended functionals $\tilde{F}_{\vec{e}_l}, \tilde{F}_{\vec{e}_d}$ (3.2.1). Let us denote the following optimal control problem (3.3.1)–(3.3.4) by **(P)**: Minimize

$$J(u,g) := -\tilde{F}_{\vec{e}_l}(u,g) + \frac{\alpha}{2}\|g\|^2_{\mathbf{L}^2(\Gamma_c)} \tag{3.3.1}$$

subject to the very weak form of

$$\begin{aligned} -\nu\Delta u + (u\cdot\nabla)u + \nabla p = f & \quad \text{in } \Omega \\ \operatorname{div} u = 0 & \quad \text{in } \Omega \\ u = g & \quad \text{on } \Gamma_c, \\ u = 0 & \quad \text{on } \Gamma \setminus \Gamma_c, \end{aligned} \tag{3.3.2}$$

the control constraints

$$\begin{aligned} g \in G_{ad} := \{v \in \mathbf{H}^0(\Gamma_c) :\ & (g_a)_i(x) \le v_i(x) \le (g_b)_i(x) \\ & \text{a.e. on } \Gamma_c,\ \forall i \in \{1,\dots,n\}\}, \end{aligned} \tag{3.3.3}$$

and the integral control-state constraint

$$\tilde{F}_{\vec{e}_d}(u,g) \le D_0. \tag{3.3.4}$$

The functions g_a and g_b are elements of $\mathbf{L}^\infty(\Gamma_c)$ with $(g_a)_i \le (g_b)_i$ for all $i \in \{1,\dots,n\}$ a.e. on Γ_c and we assume $0 \in G_{ad}$. Here, we introduced an additional regularization term $\frac{\alpha}{2}\|g\|^2_{\mathbf{L}^2(\Gamma_c)}$, where α is called the Tichonov parameter, which measures the cost of the control and is important for the optimality system, see Chapter 4. The parameter α is supposed to be positive.

3.3.1 Existence of solutions

After introducing the optimality problem, we would like to prove the existence of solutions of problem **(P)**. Unfortunately, we can not prove that the objective functional is bounded from below. This is due to the absence of a uniqueness result for the state equation for large data. Moreover, bounds on the state of the kind $\|u\|_{\mathbf{U}} \le C\|g\|_{\mathbf{L}^2(\Gamma_c)}$ can only be derived for small data. This is different to the distributed control problems for Navier-Stokes equations, where we can test the state equation with the state itself and obtain an a-priori bound without smallness assumptions.

Due to the fact that we are not able to prove the existence of solutions for the problem **(P)**, we will introduce a modification. Let us consider the minimization of

$$\tilde{J}(u,g) := \psi(-\tilde{F}_{\vec{e}_l}(u,g)) + \frac{\tilde{\alpha}}{2}\|u\|^2_{\mathbf{H}^{1/2}(\Omega)} + \frac{\alpha}{2}\|g\|^2_{\mathbf{L}^2(\Gamma_c)}. \tag{3.3.5}$$

Here, $\tilde{\alpha}$ and α are positive and small parameters. The function $\psi : \mathbb{R} \to \mathbb{R}$ is assumed to be continuous, monotone increasing and bounded from below; e.g. $\psi(r) \geq \psi_{min}$ for all $r \in \mathbb{R}$. For example, one can choose the function $\psi(r) = -\log(L_0 - r)$ with $L_0 \in \mathbb{R}$. This function additionally forbids situations where the lift is too small; that is, smaller than a prescribed value L_0, because $(L_0 - r) \in \mathbb{R}^+$ has to be fulfilled.

It appears that the functionals $\tilde{F}_{\vec{e}_i}$ are not weakly continuous with respect to $g \in \mathbf{H}^0(\Gamma_c)$. In order to overcome this difficulty, we will impose the following control constraint, where $\tilde{G}_{ad}$ is a closed and convex set such that

$$\tilde{G}_{ad} \subset \left\{ g \in G_{ad} : \int_{\Gamma_w} (g \cdot \mathfrak{n})(\varphi_i \cdot g)\, \mathrm{d}s = 0,\ i \in \{d, l\} \right\}. \tag{3.3.6}$$

If the control boundary is not part of the observation boundary, i.e. $\Gamma_c \cap \Gamma_w = \emptyset$, one can choose $\tilde{G}_{ad} = G_{ad}$. This choice is also valid in the case of pure tangential controls $g \cdot \mathfrak{n} = 0$.

Let us denote the modified minimization problem (3.3.5)–(3.3.6) by $(\tilde{\mathbf{P}})$.

Theorem 3.3. *If there is an admissible pair* $(u^0, g^0) \in \mathbf{H}^{1/2}(\Omega) \times \tilde{G}_{ad}$, *which satisfies* (3.3.2)–(3.3.4) *and the control constraint* (3.3.6), *then the problem* $(\tilde{\mathbf{P}})$ *admits at least one solution.*

Proof. The objective functional $\tilde{J}$ is bounded from below by construction. We can restrict the optimization problem to the set of all admissible pairs (u, g) with $\tilde{J}(u,g) \leq \tilde{J}(u^0, g^0)$ without changing the set of global minimizers. Let us take such an admissible pair. We then obtain

$$\begin{aligned} \frac{\tilde{\alpha}}{2}\|u\|^2_{\mathbf{H}^{1/2}(\Omega)} + \frac{\alpha}{2}\|g\|^2_{\mathbf{L}^2(\Gamma_c)} &\leq -\psi(-\tilde{F}_{\vec{e}_l}(u,g)) + J(u^0, g^0) \\ &\leq -\psi_{min} + J(u^0, g^0), \end{aligned} \tag{3.3.7}$$

which implies that the set of admissible pairs with smaller value of the objective than $J(u^0, g^0)$ is bounded.

Since $\tilde{J}$ is bounded from below, there exists a minimizing sequence $(u_n, g_n) \in \mathbf{H}^{1/2}(\Omega) \times \mathbf{H}^0(\Gamma)$. In view of (3.3.7), this sequence is bounded and we can extract a weakly converging subsequence, which is again denoted by (u_n, g_n), i.e. $u_n \rightharpoonup \bar{u}$ in $\mathbf{H}^{1/2}(\Omega)$ and $g_n \rightharpoonup \bar{g}$ in $\mathbf{H}^0(\Gamma)$.

By compact embeddings, we have $u_n \to u$ in $L^p(\Omega)^n$ for all $p < 3$ after extracting another subsequence, see [3, Theorem 6.2].

The set $\tilde{G}_{ad}$ is weakly closed by construction, $\tilde{G}_{ad}$ is defined as closed and convex, which implies $\bar{g} \in \tilde{G}_{ad}$. Together with the control constraint $g_n \in \tilde{G}_{ad}$, this allows us to pass to the limit in the functions $\tilde{F}_{\vec{e}_i}$, $\lim \tilde{F}_{\vec{e}_i}(u_n, g_n) = \tilde{F}_{\vec{e}_i}(\bar{u}, \bar{g})$.

Passing to the limit in the very weak solution is straight-forward, which implies that $\bar{u}$ is a very weak solution to $\bar{g}$. Now, standard arguments using lower semi-continuity of norms conclude the proof.

□

Let us summarize the difficulties in proving the existence of solutions:

1. The functional (3.3.1) is not bounded from below, since there is no a-priori boundary $\|u\|_U \leq C\|g\|_{\mathbf{L}^2(\Gamma_c)}$.

2. If a minimizing sequence would exist, the sequence u_n is not necessarily bounded in $\mathbf{U}$.

3. The functions $\tilde{F}_{\vec{e}_i}$ are not weakly continuous on $\mathbf{U} \times \mathbf{H}^0(\Gamma)$.

Therefore the modification of the objective and the control constraint were made to cope with these points.

1. The function ψ guarantees that the objective function is bounded from below.

2. The regularization term $\|u\|^2_{\mathbf{H}^{1/2}(\Omega)}$ gives boundedness of u_n in $\mathbf{H}^{1/2}(\Omega)$. By compact embeddings, it allows furthermore to pass to the limit in the part of $\tilde{F}_{\vec{e}_i}$ that involves the state u.

3. The control constraint $\int_\Gamma (g \cdot \mathfrak{n})(\varphi_i \cdot g)\mathrm{d}s = 0$ permits to pass to the limit in the non-linear part of $\tilde{F}_{\vec{e}_i}$ that involves the control.

Of course, there are several other possibilities to enforce existence of solutions. For instance, one could add regularization with respect to stronger norms in g. This would, however, lead to different first-order necessary optimality conditions, which are more challenging to solve numerically; see the comments below.

Another popular change would be to explicitly impose a smallness condition on the controls. However, known smallness conditions are difficult to verify, especially in the case of non-homogenous Dirichlet boundary conditions when Γ consists of more than one connected component, see [44]. But this is the case in the application we have in mind. Moreover, the smallness condition on $1/\nu = Re$ is expected to be violated for concrete applications, where typically $1/\nu = Re$ is large.

In the following, we consider the original problem without the function ψ and the term $\frac{\tilde{\alpha}}{2}\|u\|_{\mathbf{H}^{1,2}(\Omega)}$. These modifications, particularly the term with the norm of $\mathbf{H}^{1,2}(\Omega)$, would generate problems with the variational inequality, for example in the Subsection 4.4, where we consider the second-order sufficient optimality conditions.

Chapter 4

Optimality conditions

In this chapter, we will establish the first-order necessary and the second-order sufficient optimality conditions. These conditions are very important for the numerical solution methods in the next chapter. First, we will investigate the first-order necessary optimality condition for the solution in the very weak sense. We will also introduce a new space for the control function, fitting better to the numerical investigation.

4.1 First order necessary optimality conditions

Let us return to problem **(P)** stated at the beginning of Section 3.3. We will now derive necessary optimality conditions to characterize local optimal solutions. Here, we will follow the presentation in [102, Section 6.1.2] of the regularity theory of [115]. Let now $(\bar{u}, \bar{g}) \in \mathbf{U} \times G_{ad}$ be a non-singular and locally optimal pair for **(P)**. and we define the operator

$$\begin{array}{lll} G = (G_1, G_2): & \mathcal{O}(\bar{u}) \times \mathcal{O}(\bar{g}) \to \mathbf{U} \times \mathbb{R} & \\ (u, g) & \mapsto \begin{pmatrix} S(g) - u \\ \tilde{F}_{\vec{e}_d}(u, g) - D_0 \end{pmatrix} \end{array} \tag{4.1.1}$$

and the cone

$$K = \{0\} \times (-\infty, 0] \subset \mathbf{U} \times \mathbb{R}.$$

Here, D_0 denotes the drag constraint (1.0.5) and K induces a partial ordering $<_K$ on $\mathbf{U} \times \mathbb{R}$ by: $x <_K 0 \Leftrightarrow x \in K$. It is easy to show using Theorem 2.26 that the mapping G is twice Fréchet differentiable. Then, $(\bar{u}, \bar{g})$ is a local solution of the minimization problem

$$\min_{u \in \mathcal{O}(\bar{u}), g \in \mathcal{O}(\bar{g})} J(u, g) \text{ subject to } G(u, g) <_K 0, \; u \in G_{ad}.$$

Let us define the Lagrangian associated with this optimization problem

$$L(u,g;\theta,\xi) = J(u,g) + \langle S(g)-u,\theta\rangle_{\mathbf{U}',\mathbf{U}} + \xi(\tilde{F}_{\vec{e}_d}(u,g)-D_0), \quad \theta\in\mathbf{U}', \xi\in\mathbb{R}.$$

To show existence of Lagrange multipliers, we will assume the following regularity condition: there exists $\tilde{g}\in G_{ad}\cap\mathcal{O}(\bar{g})$ and the associated state $\tilde{u}\in\mathcal{O}(\bar{u})$ such that

$$G_1'(\bar{u},\bar{g})(\tilde{u},\tilde{g}) = 0,\ G_2(\bar{u},\bar{g}) + G_2'(\bar{u},\bar{g})(\tilde{u}-\bar{u},\tilde{g}-\bar{g}) < 0. \tag{4.1.2}$$

This condition is sufficient for the Zowe-Kurcyusz regularity assumption [115], see e.g. [99]. Now, we are able to establish the necessary first-order optimality conditions for **(P)**.

The optimality system

Under these assumptions, the existence of Lagrange multipliers follows by known results, see e.g. [99, 102, 115].

Theorem 4.1. *Let $(\bar{u},\bar{g})$ a non-singular local optimal solution for (P). Let us assume that there are $\tilde{g}\in G_{ad}\cap\mathcal{O}(\bar{g})$ and the associating state $\tilde{u}\in\mathcal{O}(\bar{u})$ such that the linearized Slater condition* (4.1.2) *is satisfied. Then there exists $\theta\in\mathbf{U}'$ and $\xi\geq 0$ such that the equation*

$$\theta = -\tilde{F}'_{\vec{e}_l,u}(\bar{u},\bar{g}) + \xi\tilde{F}'_{\vec{e}_d,u}(\bar{u},\bar{g}), \tag{4.1.3a}$$

the variational inequality

$$(\alpha\bar{g} - \tilde{F}'_{\vec{e}_l,g}(\bar{u},\bar{g}) + \xi\tilde{F}'_{\vec{e}_d,g}(\bar{u},\bar{g}) + S'(\bar{g})^*\theta,\ g-\bar{g})_{L^2(\Gamma_c)} \geq 0 \quad \forall g\in G_{ad}, \tag{4.1.3b}$$

and the complementarity condition

$$\xi(\tilde{F}_{\vec{e}_d}(\bar{u},\bar{g}) - D_0) = 0,\ \xi\geq 0,\ \tilde{F}_{\vec{e}_d}(\bar{u},\bar{g})\leq D_0 \tag{4.1.3c}$$

hold.

Here, the adjoint operator $S'(\bar{g})^*:\mathbf{U}'\to\mathbf{H}^0(\Gamma)'$ appears with

$$\mathbf{U}' = \mathbf{L}^{(2n/(n-1))'}(\Omega) = \mathbf{L}^{2n/(n+1)}(\Omega).$$

The dual space of $L^p(\Omega)$ can be identified with the space $L^{p'}(\Omega)$, where p' is defined by $p' := p/(p-1)$. It is related to the solution of the so-called adjoint equation. In fact, it holds [79]:

Theorem 4.2. *The action of $S'(\bar{g})^*$ can be characterized as follows: for given $\theta \in \mathbf{U}'$ it holds*

$$S'(\bar{g})^*\theta = -\Big(\nu\frac{\partial\lambda}{\partial\mathfrak{n}} - \pi\mathfrak{n}\Big)\Big|_{\Gamma_c}, \tag{4.1.4}$$

where $\lambda \in \mathbf{H}^1_0(\Omega) \cap W^{2,r}(\Omega)^n$ is the unique solution of the equation in a weak sense

$$\int_\Omega (\nu\nabla\lambda\cdot\nabla v - (\bar{u}\cdot\nabla)\lambda v - (v\cdot\nabla)\lambda\bar{u})\,\mathrm{d}x = \langle\theta, v\rangle_{H^{-1},H^1} \tag{4.1.5}$$

for all $v \in \mathbf{H}^1_0(\Omega)$ and $\pi \in W^{1,r}(\Omega)$ the associated pressure field for all $r \in [2,\infty)$.

Proof. The representation of $S'(\bar{g})^*$ is proven for instance in [79]. It remains to investigate the regularity of λ. The right-hand side $\langle\theta, v\rangle_{H^{-1},H^1}$ is given according to the previous Theorem 4.2 by

$$\langle\theta, v\rangle_{H^{-1},H^1} = \int_\Omega (-\nu\Delta\tilde{\varphi}v - (v\cdot\nabla)\tilde{\varphi}\bar{u} - (\bar{u}\cdot\nabla)\tilde{\varphi}v)\mathrm{d}x,$$

where we used the notation $\tilde{\varphi} = -\varphi_l + \xi\varphi_d$. By assumption, G_{ad} is a subset of $\mathbf{L}^\infty(\Gamma)$, hence $\bar{g} \in \mathbf{L}^\infty(\Gamma)$ and $\bar{u} \in \mathbf{L}^q(\Omega)$, $2 \le q < \infty$ for $n = 2$, $2 \le q \le \infty$ for $n = 3$, cf. (2.4.4). Due to the high regularity of $\tilde{\varphi}$, compare Lemma 3.1, we can estimate with some $\tilde{p} > n$ such that $\mathbf{W}^{2,\tilde{p}}(\Omega) \hookrightarrow \mathbf{W}^{1,\infty}(\Omega)$

$$|\langle\theta, v\rangle_{H^{-1},H^1}| \le c(1 + \|\bar{u}\|_{\mathbf{L}^q})\|\tilde{\varphi}\|_{\mathbf{W}^{2,\tilde{p}}}\|v\|_{\mathbf{L}^r}$$

with $1/q + 1/\tilde{p} + 1/r = 1$. Since q and $\tilde{p}$ can be chosen arbitrarily large (but not equal to ∞), we obtain $\theta \in \mathbf{L}^q(\Omega)^n$, for all $q \in (2,\infty)$. Let us now estimate the addend on the left-hand side of (4.1.5) that comes from the nonlinearity of the state equation:

$$\left|\int_\Omega ((\bar{u}\cdot\nabla)\lambda v + (v\cdot\nabla)\lambda\bar{u})\,\mathrm{d}x\right| \le c\|\bar{y}\|_{\mathbf{L}^q}\|\nabla\lambda\|_{\mathbf{L}^p}\|v\|_{\mathbf{L}^r} \tag{4.1.6}$$

with $1/q + 1/p + 1/r = 1$, $2 \le q < \infty$.

Since $(\bar{u}, \bar{g})$ is non-singular, the equation (4.1.5) is uniquely solvable with solution $\lambda \in \mathbf{H}^1_0(\Omega)$. That is, estimate (4.1.6) holds with $p = 2$. We can interprete the adjoint state as the weak solution of a Stokes equation, where the terms in (4.1.6) are put on the right-hand side. This allows to apply known regularity results for the Stokes equation.

With $p = 2$ the estimate (4.1.6) holds for all $r > 2$, hence the functional in (4.1.6) is in $\mathbf{L}^{r'}(\Omega)$ for all $r' < 2$. The regularity result by Galdi [44, Lemma

IV.6.1] gives in a first step the regularities $\lambda \in \mathbf{W}^{2,r'}(\Omega)$ and $\pi \in W^{1,r'}(\Omega)$ for all $r' < 2$.

By embedding arguments, we have then $\nabla\lambda \in L^p(\Omega)^{n,n}$, where p depends on n: $p \in (2,\infty)$ for $n = 2$; $p \in (2,6)$ for $n = 3$.

In the 2d-case, (4.1.6) is valid for all $p < \infty$, which allows us to chose r arbitrary small with $r > 1$. That is, the functional involving $\bar{y}$ and $\nabla\lambda$ is in $\mathbf{L}^q(\Omega)$ for all $q < \infty$. Again applying the regularity result for the Stokes equation, we find $\lambda \in \mathbf{W}^{2,q}(\Omega)$ and $\pi \in W^{1,q}(\Omega)$ for all $q < \infty$.

For the three-dimensional case, we obtain similarly $\lambda \in \mathbf{W}^{2,q}(\Omega)$ and $\pi \in W^{1,q}(\Omega)$ for all $q < 6$. By continuous imbeddings, $\nabla\lambda \in L^\infty(\Omega)^{n,n}$ follows, which gives after applying again Galdi's regularity result $\lambda \in \mathbf{W}^{2,q}(\Omega)$ and $\pi \in W^{1,q}(\Omega)$ for all $q < \infty$. □

The adjoint pressure π is only determined up to constant. This fact is usually circumvented by requiring $\int_\Omega \pi \, \mathrm{d}x = 0$. Here, it is not necessary to fix the constant, since the constant does not change the variational inequality due to the construction of $\mathbf{H}^0(\Gamma)$

$$(\pi\mathfrak{n}, v)_{L^2(\Gamma_c)} = ((\pi + c)\mathfrak{n}, v)_{L^2(\Gamma_c)} \quad \forall c \in \mathbb{R}, v \in \mathbf{H}^0(\Gamma).$$

Furthermore, the variational inequality (4.1.3b) can be written as a non-smooth equation.

Corollary 4.3. *Let the assumptions of the previous theorem be fulfilled. Then the variational inequality* (4.1.3b) *is equivalent to the following condition.*

For each connected component $\Gamma_j \in \Gamma$ with $\Gamma_j \cap \Gamma_c \neq \emptyset$ there is $\eta_j \in \mathbb{R}$ such that the following pointwise representation holds for a.a. $x \in \Gamma_c$

$$\bar{g}(x) = \mathbb{P}_{G_{ad}}\{-\frac{1}{\alpha}(-(\nu\frac{\partial\lambda}{\partial\mathfrak{n}} - \pi\mathfrak{n})(x) - \tilde{F}'_{\vec{e}_l,g}(\bar{u},\bar{g})(x) \\ + \xi\tilde{F}'_{\vec{e}_d,g}(\bar{u},\bar{g})(x) + \eta_j\mathfrak{n}(x))\} \quad (4.1.7)$$

for $x \in \Gamma_c \cap \Gamma_j$ and the zero net-mass conditions

$$\int_{\Gamma_c\cap\Gamma_j} \bar{g}\cdot\mathfrak{n}\,\mathrm{d}s = 0 \quad \forall j : \Gamma_j \cap \Gamma_c \neq \emptyset$$

are satisfied. Here, $\mathbb{P}_G : \mathbb{R}^n \to \mathbb{R}^n$ denotes the Euclidean projection in $\mathbb{R}^n$ onto the set G.

Proof. At first, the variational inequality (4.1.3b) is equivalent to

$$\bar{g} = \mathbb{P}_{G_{ad}}\left\{-\frac{1}{\alpha}\left(-\left(\nu\frac{\partial\lambda}{\partial\mathfrak{n}} - \pi\mathfrak{n}\right) - \tilde{F}'_{\vec{e}_l,g}(\bar{u},\bar{g}) + \xi\tilde{F}'_{\vec{e}_d,g}(\bar{u},\bar{g})\right)\right\},$$

where $\mathbb{P}_{G_{ad}} : \mathbf{L}^2(\Gamma) \to \mathbf{H}^0(\Gamma)$ is the projection with respect to the $\mathbf{L}^2(\Gamma)$-norm on G_{ad}. That is, $\bar{g}$ solves the minimization problem

$$\min \frac{1}{2} \left\| g + \frac{1}{\alpha} \left(- \left(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi \mathfrak{n} \right) - \tilde{F}'_{\vec{e}_l,g}(\bar{u}, \bar{g}) + \xi \tilde{F}'_{\vec{e}_d,g}(\bar{u}, \bar{g}) \right) \right\|^2_{\mathbf{L}^2(\Gamma)}$$

subject to

$$\int_{\Gamma_c \cap \Gamma_j} g \cdot \mathfrak{n} \, \mathrm{d}s = 0, \quad \forall j : \Gamma_j \cap \Gamma_c \neq \emptyset,$$
$$g(x) \in G_{ad} \text{ a.e. on } \Gamma_c.$$

Then there exist Lagrange multipliers η_j associated to the integral constraints in this auxiliary problem, and the variational inequality

$$\left(\alpha \bar{g} - \left(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi \mathfrak{n} \right) - \tilde{F}'_{\vec{e}_l,g}(\bar{u}, \bar{g}) + \xi \tilde{F}'_{\vec{e}_d,g}(\bar{u}, \bar{g}) + \big(\sum_j \eta_j \chi_j \big) \mathfrak{n}, \, g - \bar{g} \right) \geq 0 \tag{4.1.8}$$

holds for all $g \in L^2(\Gamma)$ satisfying $g(x) \in G_{ad}$ a.e. on Γ_c. By standard arguments [102], it can be proven that this variational inequality is equivalent to the projection representation as claimed. □

Remark 4.4. *The derivative of $\tilde{F}_{\vec{e}_i}$, $i \in \{d, l\}$ with respect to g is*

$$\tilde{F}'_{\vec{e}_i,g}(u(x), g(x)) = - \int_\Gamma \nu g(x) \frac{\partial \varphi_i}{\partial \mathfrak{n}}(x) + (g(x) \cdot \mathfrak{n}(x))(\varphi_i(x) \cdot g(x)) \, \mathrm{d}s.$$

Unfortunately, the argument $-\tilde{F}'_{\vec{e}_l,g}(\bar{u}, \bar{g}) + \xi \tilde{F}'_{\vec{e}_d,g}(\bar{u}, \bar{g})$ of the projection depends on the control itself. This term involves no smoothing operation. Hence, we cannot conclude higher regularity of optimal controls from the projection representation, such that $\bar{u}$ has the same regularity as $\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi \mathfrak{n}$. Moreover, the non-smooth formulation of the variational inequality is not suitable for semi-smooth Newton methods.

To handle this difficulty, we will consider in the section after the next an optimal control problem allowing more regular control functions. The idea behind this is to consider a finite-dimensional control space and the once reformulated functionals $\hat{F}_{\vec{e}_i}$, $i \in \{d, l\}$, see Chapter 3.

4.2 Second-order sufficient optimality condition

The presentation of second-order sufficient optimality conditions in this subsection follows [21, 22, 48, 103, 109].

Let $(\bar{u}, \bar{p}, \bar{g})$ be a fixed admissible pair that fulfills the first-order necessary optimality condition of Theorem 4.1 together with the adjoint state $\bar{\lambda}$ and the Lagrange multipliers $\bar{\xi}$, $\bar{\eta}$ and let us define for simplification

$$j := \alpha \bar{g} - \left(\nu \frac{\partial \bar{\lambda}}{\partial \mathfrak{n}} - \bar{\pi} \mathfrak{n} \right) - \tilde{F}'_{\vec{e}_l, g}(\bar{u}, \bar{g}) + \xi \tilde{F}'_{\vec{e}_d, g}(\bar{u}, \bar{g}) + \left(\sum_j \bar{\eta}_j \chi_j \right) \mathfrak{n}. \quad (4.2.1)$$

For $\varepsilon > 0$, we define by

$$\Gamma_{\varepsilon,i} := \{x \in \Gamma : |j_i(x)| > \varepsilon\}$$

the set of strongly active control constraints for $\bar{g}$.

Remark 4.5. *Note that the variational inequality* (4.1.8) *determines the optimal control $\bar{g}$ uniquely on $\Gamma_{\varepsilon,i}$. We obtain*

$$\bar{g}_i(x) = g_{a,i}(x), \text{ if } j_i(x) \geq \varepsilon$$

and

$$\bar{g}_i(x) = g_{b,i}(x), \text{ if } j_i(x) < -\varepsilon.$$

Following Casas, Tröltzsch and Unger [22], based on Maurer and Zowe [70], the linearized cone $L(\bar{u}, \bar{p}, \bar{g})$ is made up of all $(z, \mu, h) \in \mathbf{U} \times W^{-1,2n/(n-1)} \times \mathbf{L}^2(\Gamma)$ satisfying the following conditions (4.2.2)-(4.2.4):

$$\begin{aligned} -\nu \Delta z + (\bar{u} \cdot \nabla) z + (z \cdot \nabla) \bar{u} + \nabla \mu &= 0 \quad \text{in } \Omega \\ \operatorname{div} z &= 0 \quad \text{in } \Omega \\ z &= h \quad \text{on } \Gamma_c, \\ z &= 0 \quad \text{on } \Gamma \setminus \Gamma_c, \end{aligned} \qquad (4.2.2)$$

$$\tilde{F}'_{\vec{e}_d}(\bar{u}, \bar{p}, \bar{g})(z, \mu, h) \leq 0, \qquad (4.2.3)$$

$$h = g - \bar{g}, \ g \in G_{ad}. \qquad (4.2.4)$$

Let us denote by

$$\mathbb{P}_\varepsilon : \mathbf{L}^2(\Gamma_c) \to \mathbf{L}^2(\Gamma_c), \ g \mapsto \chi_{\Gamma \setminus \Gamma_\varepsilon} g$$

the projection operator. That means

$$(\mathbb{P}_\varepsilon g)(x) = \begin{cases} g(x) & \text{on } \Gamma \setminus \Gamma_\varepsilon \\ 0 & \text{on } \Gamma_\varepsilon \end{cases}.$$

Then, we split for all $v \in L(\bar{v})$ the control function g into $g_1 = (g - \mathbb{P}_\varepsilon g)$ and $g_2 = (\mathbb{P}_\varepsilon g)$. The solutions (u_i, p_i), $i = 1, 2$, of the linearized state equations (4.2.2) are associated to g_i, $i = 1, 2$. This means

$$v = v_1 + v_2 = (u_1, p_1, g_1) + (u_2, p_2, g_2). \tag{4.2.5}$$

We assume that $\bar{v} = (\bar{u}, \bar{p}, \bar{g})$ fulfills with the Lagrange multipliers $\bar{l} = (\bar{\lambda}, \bar{\pi}, \bar{\xi}, \bar{\eta})$ the following coercivity assumption $\mathcal{L}''(\bar{v}, \bar{l})$ called (SSC):

$$(SSC) \begin{cases} \text{There exist } \varepsilon > 0 \text{ and } \delta > 0 \text{ such that} \\ \mathcal{L}_{vv}(\bar{v}, \bar{l})[(u_2, p_2, g_2)^2] \geq \delta \|g_2\|^2_{\mathbf{L}^2(\Gamma_c)} \\ \text{holds for all pairs } (u_2, p_2, g_2) \text{ constructed by (4.2.5).} \end{cases}$$

Theorem 4.6. *Let $(\bar{v})$ be an admissible non-singular point for the optimal control problem and fulfill the first-order necessary optimality condition of Theorem 4.11 with associated λ, ξ. Assume furthermore that the coercivity assumption (SSC) is satisfied. Then there exist $\alpha > 0$ and $\tau > 0$ such that*

$$J(v) \geq J(\bar{v}) + \alpha \|g - \bar{g}\|^2_{\mathbf{L}^2(\Gamma_c)}$$

holds for all admissible pairs v with $\|g - \bar{g}\|_{\mathbf{L}^\infty(\Gamma_c)} \leq \tau$.

To prove this theorem, we establish the following two lemma.

Lemma 4.7. *For all $g \in G_{ad}$ there holds*

$$\int\limits_{\Gamma_c} (j)\,(g - \bar{g}) \mathrm{d}x \geq \varepsilon \|g - \bar{g}\|^2_{\mathbf{L}^1(\Gamma_\varepsilon)}. \tag{4.2.6}$$

Proof. Let $g \in G_{ad}$ and because $(\bar{v}, \bar{\lambda}, \bar{\pi}, \bar{\eta}, \bar{\xi})$ with $\bar{v} = (\bar{u}, \bar{p}, \bar{g})$ fulfill the first-order necessary optimality condition (4.11), we have

$$(j_i(x))\,(g_i(x) - \bar{g}_i(x)) \geq 0 \tag{4.2.7}$$

for almost all $x \in \Gamma_c$, $i = 1, \ldots, n$. Integrating (4.2.7) over Γ_c leads with the definition of $\Gamma_{\varepsilon,i}$ to

$$\begin{aligned} &\int\limits_{\Gamma_c} (j_i(x))\,(g_i(x) - \bar{g}_i(x)) \mathrm{d}x \\ &\qquad \geq \int\limits_{\Gamma_{\varepsilon,i}} (j_i(x))\,(g_i(x) - \bar{g}_i(x)) \mathrm{d}x \\ &\qquad = \int\limits_{\Gamma_{\varepsilon,i}} |j_i(x)||g_i(x) - \bar{g}_i(x)| \mathrm{d}x \\ &\qquad \geq \varepsilon \int\limits_{\Gamma_{\varepsilon,i}} |g_i(x) - \bar{g}_i(x)| \mathrm{d}x \end{aligned}$$

and the sum of $i = 1, 2$ to the statement of the theorem. □

Lemma 4.8. *For all $v = v_1 + v_2$ defined as in* (4.2.5) *in the linearized cone $L(\bar{v})$ there holds*

$$\mathcal{L}_{vv}(\bar{v}, \bar{l})[v_1, v_2] \leq c\|g_1\|_{\mathbf{L}^2(\Gamma_c)}\|g_2\|_{\mathbf{L}^2(\Gamma_c)}.$$

Proof. We get

$$\begin{aligned}\mathcal{L}_{vv}(\bar{v}, \bar{l})[v_1, v_2] = &\int_\Omega -(u_1 \cdot \nabla)\bar{\lambda}u_2) - (u_2 \cdot \nabla)\bar{\lambda}u_1) \,\mathrm{d}x \\ &+ \int_\Omega (u_1 \cdot \nabla)\varphi_l u_2 + (u_2 \cdot \nabla)\varphi_l u_1 \,\mathrm{d}x \\ &+ \int_\Omega (u_1 \cdot \nabla)\varphi_d u_2 + (u_2 \cdot \nabla)\varphi_d u_1 \,\mathrm{d}x \\ &+ \int_\Gamma (g_1 \cdot \mathfrak{n})(\varphi_l \cdot g_2) + (g_2 \cdot \mathfrak{n})(\varphi_l \cdot g_1) \,\mathrm{d}s \\ &+ \int_\Gamma (g_1 \cdot \mathfrak{n})(\varphi_d \cdot g_2) + (g_2 \cdot \mathfrak{n})(\varphi_d \cdot g_1) \,\mathrm{d}s + \alpha(g_1, g_2)_{\mathbf{L}^2(\Gamma_c)} \\ \leq\, & c(\|u_1\|_{\mathbf{U}}\|u_2\|_{\mathbf{U}} + \|g_1\|_{\mathbf{L}^2(\Gamma_c)}\|g_2\|_{\mathbf{L}^2(\Gamma_c)}).\end{aligned}$$

Because $(\bar{u}, \bar{g})$ is non-singular and u_i, $i = 1, 2$, are the solutions of the linearized Navier-Stokes equations, we obtain

$$\mathcal{L}_{vv}(\bar{v}, \bar{l})[v_1, v_2] \leq c\|g_1\|_{\mathbf{L}^2(\Gamma_c)}\|g_2\|_{\mathbf{L}^2(\Gamma_c)}.$$

□

Now, we are able to proof Theorem 4.6. This proof is based on [22, 109].

Proof of Theorem 4.6. We assume that $(\bar{u}, \bar{p}, \bar{g})$ fulfill the assumptions of the theorem and let $(u, p, g) \in \mathcal{O}(\bar{u}) \times \mathcal{O}(\bar{p}) \times \mathcal{O}(\bar{g})$ be another admissible pair. Then we have with $v = (u, p, g)$ and $\bar{l} = (\bar{\lambda}, \bar{\pi}, \bar{\xi}, \bar{\eta})$

$$\begin{aligned}J(\bar{v}) &= \mathcal{L}(\bar{v}, \bar{l}) - \bar{\xi}(F_{\vec{e}_d}(\bar{v}) - D_0) = \mathcal{L}(\bar{v}, \bar{l}), \\ J(v) &= \mathcal{L}(v, \bar{l}) - \bar{\xi}(F_{\vec{e}_d}(v) - D_0) \geq \mathcal{L}(v, \bar{l})\end{aligned}$$

due to the complementary condition. This yields

$$J(v) - J(\bar{v}) \geq \mathcal{L}(v, \bar{l}) - \mathcal{L}(\bar{v}, \bar{l}).$$

A Taylor-expansion of the Lagrangian $\mathcal{L}$ yields the following equality:

$$\begin{aligned}\mathcal{L}(v, \bar{l}) = \mathcal{L}(\bar{v}, \bar{l}) &+ \frac{\partial \mathcal{L}}{\partial(u, p)}(\bar{v}, \bar{l})(u - \bar{u}, p - \bar{p}) \\ &+ \frac{\partial \mathcal{L}}{\partial g}(\bar{v}, \bar{l})(g - \bar{g}) + \frac{1}{2}\mathcal{L}_{vv}(\bar{v}, \bar{l})[(v - \bar{v})]^2.\end{aligned} \tag{4.2.8}$$

Due to the quadratic nature of the nonlinear term, the remainder vanishes and because the first-order necessary optimality conditions are satisfied at $\bar{v}$ with the corresponding Lagrange multipliers $\bar{l}$, the second term of (4.2.8) is equal to zero. For the third term we have the inequality

$$\frac{\partial \mathcal{L}}{\partial g}(\bar{v},\bar{l})(g-\bar{g}) = \int\limits_{\Gamma_c} (j)\,(g-\bar{g})\mathrm{d}x \geq \varepsilon \|g-\bar{g}\|^2_{\mathbf{L}^1(\Gamma_\varepsilon)}$$

see Lemma 4.7. This leads to

$$\begin{aligned} J(v) = J(\bar{v}) &+ \frac{\partial \mathcal{L}}{\partial(u,p)}(\bar{v},\bar{l})(u-\bar{u},p-\bar{p}) \\ &+ \frac{\partial \mathcal{L}}{\partial g}(\bar{v},\bar{l})(g-\bar{g}) + \frac{1}{2}\mathcal{L}_{vv}(\bar{v},\bar{\lambda},\bar{\pi},\bar{\xi})[(v-\bar{v})]^2 \\ &\geq J(\bar{v}) + \varepsilon\|g-\bar{g}\|^2_{\mathbf{L}^1(\Gamma_\varepsilon)} + \frac{1}{2}\mathcal{L}_{vv}(\bar{v},\bar{l})[(v-\bar{v})]^2. \end{aligned}$$

Analogous to Casas, Tröltzsch and Unger [22, Section 7.2, proof of Theorem 4.2], we approximate $v-\bar{v}$ by $v_l = (u_l, p_l, g_l)$ of the linearized cone $L(\bar{v})$ and the remainder term $e_1 = (v-\bar{v}) - v_l$ satisfies the estimate

$$\|e_1\| \leq c\|g-\bar{g}\|^2_{\mathbf{L}^2(\Gamma_c)}.$$

Let us now take $u_l + e_1$ instead of $u - \bar{u}$, then we derive

$$\begin{aligned} \frac{\partial^2 \mathcal{L}}{\partial(u,p)^2}(\bar{v},\bar{l})[u-\bar{u},p-\bar{p}]^2 &= \frac{\partial^2 \mathcal{L}}{\partial(u,p)^2}(\bar{v},\bar{l})[u_l,p_l]^2 \\ &\quad + 2\frac{\partial^2 \mathcal{L}}{\partial(u,p)^2}(\bar{v},\bar{l})[u_l,p_l,e_1] \\ &\quad + \frac{\partial^2 \mathcal{L}}{\partial(u,p)^2}(\bar{v},\bar{l})[e_1]^2 \\ &= \frac{\partial^2 \mathcal{L}}{\partial(u,p)^2}(\bar{v},\bar{l})[u_l,p_l]^2 + e_2, \end{aligned}$$

with the remainder e_2 estimated due to the non-singularity of $(\bar{u},\bar{g})$ by

$$|e_2| \leq c(\|u_l\|_{\mathbf{U}} + \|p_l\|_{W^{-1,2n/(n-1)}} + \|e_1\|_{\mathbf{U}})\|e_1\|_{\mathbf{U}} \leq c(\|g_l\|_{\mathbf{L}^2(\Gamma_c)} + \|e_1\|_{\mathbf{U}})\|e_1\|_{\mathbf{U}}$$

and fulfilling

$$\frac{|e_2|}{\|g_l\|^2_{\mathbf{L}^2(\Gamma_c)}} \to 0, \text{ for } \|g_l\|_{\mathbf{L}^2(\Gamma_c)} \to 0.$$

Summarized, we obtain

$$J(v) - J(\bar{v}) \geq \frac{1}{2}\mathcal{L}_{vv}(\bar{v}, \bar{l})[u_l, p_l, g_l]^2 + \varepsilon\|g - \bar{g}\|^2_{\mathbf{L}^2(\Gamma_\varepsilon)} + e_2. \tag{4.2.9}$$

Now, as in (4.2.5), we split $v_l = (u_l, p_l, g_l)$ into

$$u_l = u_1 + u_2, \;\; p_l = p_1 + p_2, \;\; g_l = g_1 + g_2,$$

where (u_1, p_1) and (u_2, p_2) are the solutions of

$$\begin{aligned} -\nu\Delta u_1 + (\bar{u}\cdot\nabla)u_1 + (u_1\cdot\nabla)\bar{u} + \nabla p = 0 \quad & \text{in } \Omega \\ \operatorname{div} u_1 = 0 \quad & \text{in } \Omega \\ u_1 = g_1 \quad & \text{on } \Gamma. \end{aligned}$$

and

$$\begin{aligned} -\nu\Delta u_2 + (\bar{u}\cdot\nabla)u_1 + (u_1\cdot\nabla)\bar{u} + \nabla p = 0 \quad & \text{in } \Omega \\ \operatorname{div} u_2 = 0 \quad & \text{in } \Omega \\ u_2 = g_2 \quad & \text{on } \Gamma. \end{aligned}$$

respectively, and (u_2, p_2, g_2) fulfill $(u_2, p_2, g_2) \in L(\bar{v})$ and $(g_2)_i = 0$ on $\Gamma_{\varepsilon,i}$, $i = 1, 2$. Thus, (SSC) applies to $\mathcal{L}_{vv}(\bar{v}, \bar{l})(u_2, g_2)$.

Considering the first term of the right-hand side of (4.2.9)

$$\begin{aligned} \mathcal{L}_{vv}(\bar{v}, \bar{l})[(u_l, g_l)]^2 = \;& \mathcal{L}_{vv}(\bar{v}, \bar{l})[v_1]^2 \\ & + 2\mathcal{L}_{vv}(\bar{v}, \bar{l})[v_1, v_2] \\ & + \mathcal{L}_{vv}(\bar{v}, \bar{l})[v_2]^2, \end{aligned} \tag{4.2.10}$$

we are able to use (SSC) and obtain

$$\mathcal{L}_{vv}(\bar{v}, \bar{l})[(v_2)]^2 \geq \delta\|g_2\|^2_{L^2(\Gamma_c)}. \tag{4.2.11}$$

Lemma 4.8 leads to

$$\mathcal{L}_{vv}(\bar{v}, \bar{l})[v_1, v_2] \leq c\|g_1\|_{\mathbf{L}^2(\Gamma_c)}\|g_2\|_{\mathbf{L}^2(\Gamma_c)}.$$

We estimate $\mathcal{L}_{vv}(\bar{v}, \bar{l})[v_1]^2$ analogously to Lemma 4.8 by

$$\mathcal{L}_{vv}(\bar{v}, \bar{l})[v_1, v_2] \leq c\|g_1\|^2_{\mathbf{L}^2(\Gamma_c)}.$$

Then, we obtain

$$\begin{aligned} |2\mathcal{L}_{vv}(\bar{v}, \bar{l})[v_1, v_2] + \mathcal{L}_{vv}(\bar{v}, \bar{l})[v_2]^2| & \\ \geq -c\|g_1\|_{\mathbf{L}^2(\Gamma_c)}(\|g_1\|_{\mathbf{L}^2(\Gamma_c)} + \|g_2\|_{\mathbf{L}^2(\Gamma_c)}) & \\ \geq -\frac{\delta}{2}\|g_2\|^2_{\mathbf{L}^2(\Gamma_c)} - c\|g_1\|^2_{\mathbf{L}^2(\Gamma_c)}, & \end{aligned} \tag{4.2.12}$$

because of Lemma 4.8 and the Lipschitz continuity of the solution mapping of the linearized system. Considering the relation $\|g_2\|^2_{\mathbf{L}^2(\Gamma_c)} \geq \frac{1}{2}\|g_l\|^2_{\mathbf{L}^2(\Gamma_c)} - \|g_1\|^2_{\mathbf{L}^2(\Gamma_c)}$, (4.2.10), (4.2.11) and (4.2.12), we get

$$\begin{aligned}\mathcal{L}_{vv}(\bar{v},\bar{l})[(v_l)]^2 &\geq \frac{\delta}{2}\|g_2\|^2_{\mathbf{L}^2(\Gamma_c)} - c\|g_1\|^2_{\mathbf{L}^2(\Gamma_c)} \geq \frac{\delta}{4}\|g_l\|^2_{\mathbf{L}^2(\Gamma_c)} - c\|g_1\|^2_{\mathbf{L}^2(\Gamma_c)} \\ &= \frac{\delta}{4}\|g-\bar{g}\|^2_{\mathbf{L}^2(\Gamma_c)} - c\|g-\bar{g}\|^2_{\mathbf{L}^2(\Gamma_\varepsilon)}.\end{aligned}$$

Now, we proved

$$J(v) - J(\bar{v}) \geq \frac{\delta}{8}\|g-\bar{g}\|^2_{\mathbf{L}^2(\Gamma_c)} + \varepsilon\|g-\bar{g}\|_{\mathbf{L}^1(\Gamma_\varepsilon)} - c\|g-\bar{g}\|^2_{\mathbf{L}^2(\Gamma_\varepsilon)} + e_2.$$

Furthermore, we obtain

$$J(v) - J(\bar{v}) \geq \frac{\delta}{8}\|g-\bar{g}\|^2_{\mathbf{L}^2(\Gamma_c)} + (\varepsilon - c\|g-\bar{g}\|_{\mathbf{L}^\infty(\Gamma_\varepsilon)})\|g-\bar{g}\|_{\mathbf{L}^1(\Gamma_\varepsilon)} + e_2.$$

Choosing τ sufficiently small, it holds

$$J(v) - J(\bar{v}) \geq \frac{\delta}{16}\|g-\bar{g}\|^2_{\mathbf{L}^2(\Gamma_c)}.$$

□

4.3 Finite-dimensional control set

We will now consider a regularized version of the optimal control problem (3.3.1)–(3.3.4). In particular, the controls will now be taken from $\mathbf{H}^{1/2}(\Gamma)$, which leads to higher regularity of the associated states. For $\mathbf{H}^{1/2}(\Gamma)$-controls one has the following regularity result, see Theorem 2.15 or [44, Theorem VIII.4.1].

Lemma 4.9. *For every $g \in \mathbf{H}^{1/2}(\Gamma)$ the very weak solution u belongs to $\mathbf{H}^1(\Omega)$, and the trace of u coincides with the control g on the boundary Γ.*

Because in this case the state u is of the space $\mathbf{H}^1(\Omega)$, we are able to use for the boundary integrals (1.0.2) and (1.0.3) the reformulation

$$\hat{F}_{\vec{e}_i}(u) := -\int_\Omega (\nu\nabla u\cdot\nabla\varphi_i + (u\cdot\nabla)y\varphi_i)\,\mathrm{d}x,\ i \in \{d,l\},$$

see (3.1.2), instead of $\tilde{F}_{\vec{e}_i}$, $i = d, l$. Then $\hat{F}_{\vec{e}_i}$, $i = d, l$ is twice continuously Fréchet differentiable from $\mathbf{H}^1(\Omega)$ to $\mathbb{R}$. And it holds

$$\hat{F}_{\vec{e}_i}(u) = \int_\Gamma \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \right) \varphi_i \mathrm{d}s = \tilde{F}_{\vec{e}_i}(u, g),\ i \in \{d, l\},$$

for smooth states u associated to controls g. Then, we have to redefine $G_2(u, g)$ in (4.1.1) by

$$G_2(u, g) = (\hat{F}_{\vec{e}_d}(u) - D_0).$$

.

Additionally, let us introduce a finite-dimensional control space. Let e_i, $i = 1 \dots l$, be linearly independent functions from $\mathbf{H}^{1/2}(\Gamma)$ with support on Γ_c. Let $Q_{ad} \in \mathbb{R}^l$ be the set of admissible coefficients

$$Q_{ad} := \{q \in \mathbb{R}^l :\ q_{a,i} \leq q_i \leq q_{b,i}\},\ q_a, q_b \in \mathbb{R}^l.$$

Then, we define the set of admissible controls as

$$\hat{G}_{ad,Q} := \left\{ g \in \mathbf{H}^{1/2}(\Gamma) :\ g = \sum_{j=1}^{l} q_j e_j,\ q \in Q_{ad} \right\}.$$

The idea behind this space is that in many applications only a few parameters can be optimized. For example, in [89] the actuation in a separation control investigation consists of a loudspeaker, where the free optimization parameters were frequency and amplitude. So, only two parameters appear.

Instead of this construction, we could have added a penalization term like $\beta \|g\|_{\mathbf{H}^{1/2}}$ to the cost functional. However, this additional term is not justified physically. Moreover, the optimality system would involve a variational inequality with a non-local differential operator on Γ_c.

Now, we are considering the following optimization problem, henceforth called $(\mathbf{P}_l)$: Minimize

$$J(u, g) := -\hat{F}_{\vec{e}_l}(u) + \frac{\alpha}{2} \|g\|^2_{\mathbf{L}^2(\Gamma_c)}$$

subject to the very weak form of

$$\begin{aligned} -\nu \Delta u + (u \cdot \nabla) u + \nabla p &= 0 \quad \text{in } \Omega \\ \operatorname{div} u &= 0 \quad \text{in } \Omega \\ u &= g \quad \text{on } \Gamma_c, \\ u &= 0 \quad \text{on } \Gamma \setminus \Gamma_c, \end{aligned}$$

the control constraints

$$g \in \hat{G}_{ad,Q}$$

and the integral state constraint

$$\hat{F}_{\vec{e}_d}(u) \leq D_0.$$

This means

$$g(x) = \sum_{j=1}^{l} q_j e_j(x), \; q \in Q_{ad}$$

and

$$\hat{F}_{\vec{e}_d}(u) = -\int\limits_{\Omega} (\nu \nabla u \cdot \nabla \varphi_i + (u \cdot \nabla) u \varphi_i) \; \mathrm{d}x \leq D_0.$$

Due to the same reasons as above, existence of solutions cannot be proven directly. Here, we would have to work with similar modifications to $(\mathbf{P}_l)$ as in Section 3.3.1 above. Rather, we would like to derive a first-order optimality system. To this end, let us assume that $(\bar{u}, \bar{g})$ is a non-singular and locally optimal solution of $(\mathbf{P}_l)$. Moreover, let us assume that a linearized Slater point for the state constraint exists, similarly defined as in (4.1.2)

Then one can argue as above to obtain:

Theorem 4.10. *Let $(\bar{u}, \bar{g})$ be a non-singular local optimal solution for $(\mathbf{P}_l)$. Let us assume that there are $\tilde{g} \in G_{ad,Q} \cap \mathcal{O}(\bar{g})$ and the associated $\tilde{u} \in \mathcal{O}(\bar{u})$ such that the linearized Slater condition* (4.1.2) *is satisfied. Then there exists a multiplier $\xi \geq 0$, an adjoint state $\lambda \in \mathbf{H}_0^1(\Omega) \cap W^{2,r}(\Omega)^n$, and an adjoint pressure $\pi \in W^{1,r}(\Omega)$, $r \in [2, \infty)$, such that (λ, π) is the weak solution of*

$$\begin{aligned} -\nu \Delta \lambda + (\nabla \bar{u})^T \lambda - (\bar{u} \cdot \nabla) \lambda + \nabla \pi = (\nabla \bar{u})^T (\varphi_l - \xi \varphi_d) - \nu \Delta (\varphi_l - \xi \varphi_d) & \\ - (\bar{u} \cdot \nabla)(\varphi_l - \xi \varphi_d) \qquad & \text{in } \Omega \\ \operatorname{div} \lambda = 0 \qquad & \text{in } \Omega \\ \lambda = 0 \qquad & \text{on } \Gamma, \end{aligned} \tag{4.3.1a}$$

and such that the variational inequality

$$\left(\alpha \bar{g} - \left(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi \mathfrak{n} \right), \; g - \bar{g} \right)_{L^2(\Gamma_c)} \geq 0 \quad \forall g \in \hat{G}_{ad,Q}, \tag{4.3.1b}$$

and the complementarity condition

$$\xi(\hat{F}_{\vec{e}_d}(\bar{u}) - D_0) = 0, \, \xi \geq 0, \, \hat{F}_{\vec{e}_d}(\bar{u}) \leq D_0 \tag{4.3.1c}$$

are satisfied.

Here, the mapping $\bar{g} \mapsto (\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi \mathfrak{n})$ is differentiable from $\mathbf{L}^2(\Gamma_c)$ to $\mathbf{L}^\infty(\Gamma_c)$. Thus, we are able to apply the semi smooth Newton method to the system of Theorem 4.10. This was not possible for the system in Theorem 4.1 due to the lack of smoothing in the argument of the projection (4.1.7).

The variational inequality (4.3.1b) can be written equivalently as a variational inequality for the coefficients $\bar{q}$ of $\bar{g}$. Let us define the mass matrix M and a vector D as:

$$M \in \mathbb{R}^{l,l},\ M_{i,j} = \int\limits_{\Gamma_c} e_i e_j \,\mathrm{d}s,\ D \in \mathbb{R}^l,\ D_i = \int\limits_{\Gamma_c} \left(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi \mathfrak{n} \right) e_i \,\mathrm{d}s.$$

Then (4.3.1b) is equivalent to

$$(\alpha M \bar{q} - D)^T (q - \bar{q}) \geq 0 \ \forall q \in Q_{ad}, \tag{4.3.2}$$

which is the necessary and sufficient optimality condition of the quadratic programming problem

$$\min_{q \in Q_{ad}} \frac{\alpha}{2} q^T M q - D^T q.$$

Under some additional assumptions, we can simplify the system (4.3.1) even more. Here, we will replace the functions φ_l and φ_d with differently defined functions. Let us assume that there exists functions $(\varphi_i, \pi_i) \in \mathbf{H}^2(\Omega) \times H^1(\Omega)$, $i \in \{d, l\}$, such that it holds

$$\begin{aligned} \operatorname{div} \varphi_i &= 0 && \text{on } \Omega \\ \varphi_i &= e_i && \text{on } \Gamma_w \\ \varphi_i &= 0 && \text{on } \Gamma \setminus \Gamma_w \\ \nu \frac{\partial \varphi_i}{\partial \mathfrak{n}} - \pi_i \mathfrak{n} &= 0 && \text{on } \Gamma_c. \end{aligned} \tag{4.3.3}$$

The main advantage will be that that we need not calculate φ_l and φ_d in the numerical investigation. Of course, the φ_i cannot be chosen as solutions of a Stokes system, since the above conditions represent over-determined boundary conditions. With this choice of auxiliary functions, all result remain valid, since we have never used that (φ_i, π_i) should be solutions of a Stokes equation. Introducing a new adjoint state as the difference of $\varphi_l - \xi \varphi_d$ and the adjoint state given by Theorem 4.10, we obtain

Theorem 4.11. *Let the assumptions of Theorem 4.10 be satisfied. Assume there exists* $(\varphi_i, \pi_i) \in \mathbf{H}^2(\Omega) \times H^1(\Omega)$*,* $i \in \{d, l\}$*, such that* (4.3.3) *is satisfied.*

Then there exists a multiplier $\xi \geq 0$, an adjoint state $\lambda \in \mathbf{H}^1(\Omega) \cap \mathbf{W}^{2,r}(\Omega)$, and an adjoint pressure $\pi \in W^{1,r}(\Omega)$, $r \in [2,\infty)$, such that (λ,π) is the weak solution of

$$\begin{aligned} -\nu\Delta\lambda + (\nabla\bar{u})^T\lambda - (\bar{u}\cdot\nabla)\lambda + \nabla\pi &= 0 && \text{in } \Omega \\ \operatorname{div}\lambda &= 0 && \text{in } \Omega \\ \lambda &= \vec{e}_l - \xi\vec{e}_d && \text{on } \Gamma_w \\ \lambda &= 0 && \text{on } \Gamma \setminus \Gamma_w, \end{aligned} \tag{4.3.4a}$$

and such that the variational inequality

$$\left(\alpha\bar{g} - \left(\nu\frac{\partial\lambda}{\partial\mathfrak{n}} - \pi\mathfrak{n} + (\bar{u}\cdot\mathfrak{n})\lambda\right),\ g - \bar{g}\right)_{L^2(\Gamma_c)} \geq 0 \quad \forall g \in \hat{G}_{ad,Q}, \tag{4.3.4b}$$

and the complementarity condition

$$\xi(\hat{F}_{\vec{e}_d}(\bar{u}) - D_0) = 0,\ \xi \geq 0,\ \hat{F}_{\vec{e}_d}(\bar{u}) \leq D_0 \tag{4.3.4c}$$

are satisfied.

The additional term $(\bar{u}\cdot\mathfrak{n})\lambda$ in (4.3.4b) appears because of $\lambda \in \mathbf{H}^1(\Omega) \cap \mathbf{W}^{2,r}(\Omega)$ instead of $\mathbf{H}^1_0(\Omega) \cap \mathbf{W}^{2,r}(\Omega)$, $r \in [2,\infty)$, see e.g. Section 5.2.1 for a formal Lagrange approach.

Please note that this system does not involve the functions φ_i in the adjoint equation and in the variational inequality, which was the case for the optimality systems obtained in Theorems 4.1, 4.2, and 4.10. This makes this system favorable for computations, and it is used for the solution algorithm that we employed in our numerical experiments.

The optimality system (4.3.4a)-(4.3.4c) of the previous theorem can also be obtained formally with the help of the Lagrangian. Let us define the Lagrange functional $\mathcal{L}$ as the sum of the cost functional, Navier-Stokes equation tested with (λ,π), and state constraint

$$\begin{aligned} \mathcal{L}(u,p,g,\lambda,\pi,\xi) = \; & -\int_{\Gamma_w}\left(\nu\frac{\partial u}{\partial\mathfrak{n}_w} - p\mathfrak{n}_w\right)\cdot\vec{e}_l\,\mathrm{d}s + \frac{\alpha}{2}\int_{\Gamma_c}|g|^2\mathrm{d}s \\ & - \int_{\Omega}\lambda(-\nu\Delta u + (u\cdot\nabla)u + \nabla p)\mathrm{d}x \\ & + \int_{\Gamma_c}(g-u)\lambda_2\mathrm{d}s + \int_{\Gamma_c}u\lambda_3\mathrm{d}s \\ & + \xi\left(-\int_{\Gamma_w}\left(\nu\frac{\partial u}{\partial\mathfrak{n}} - p\mathfrak{n}\right)\cdot\vec{e}_d\,\mathrm{d}s - D_0\right). \end{aligned}$$

Then, (4.3.4a) is equivalent to

$$\begin{aligned}\mathcal{L}_u(\bar{u},\bar{p},\bar{g},\lambda,\pi,\xi)v = &\ 0,\ \forall v \in \mathbf{H}^1(\Omega),\\ \mathcal{L}_p(\bar{u},\bar{p},\bar{g},\lambda,\pi,\xi)q = &\ 0,\ \forall q \in L^2(\Omega)\end{aligned}$$

and (4.3.4b) can be obtained formally by

$$\mathcal{L}_g(\bar{u},\bar{p},\bar{g},\lambda,\pi,\xi)(g-\bar{g}) \geq 0,\ \forall g \in \hat{G}_{ad,Q}.$$

In order to obtain (4.3.4b) in a rigorous way, we had to use regular controls and to suppose the existence of φ_i satisfying (4.3.3).

We have chosen

$$\hat{G}_{ad,Q} := \left\{ g \in \mathbf{H}^{1/2}(\Gamma) :\ g = \sum_{j=1}^{l} q_j e_j,\ q \in Q_{ad} \right\}$$

with $e_i \in \mathbf{H}^{1/2}(\Gamma)$. Let us consider for simplicity in the following chapters

$$G_{ad,Q} := \left\{ g \in \mathbf{H}^{3/2}(\Gamma) :\ g = \sum_{j=1}^{l} q_j e_j,\ q \in Q_{ad} \right\} \tag{4.3.5}$$

with $e_i \in \mathbf{H}^{3/2}(\Gamma)$ and $q_a, q_b \in \mathbb{R}^l$.

Based on this set of admissible controls $G_{ad,Q}$, we get $u \in \mathbf{H}^2(\Omega)$, see Theorem 2.17.

Depending on the respective situation in the following, we will decide to apply one of the two equivalent notations (4.3.2) or (4.3.1b) with $G_{ad,Q}$ instead of $\hat{G}_{ad,Q}$.

Now, we are also able to use $F_{\vec{e}_i}(u)$ in (4.3.4c) and the cost functional instead of $\hat{F}_{\vec{e}_i}(u)$, $i \in \{d, l\}$. Additionally, we have to redefine $G_2(u,g)$ in (4.1.1) by

$$G_2(u,g) = (F_{\vec{e}_d}(u) - D_0).$$

.

Furthermore, we have to redefine the definition of non-singularity and Theorem 2.26 with $(\bar{u},\bar{g}) \in \mathbf{H}^2(\Omega) \times \mathbf{H}^{3/2}(\Gamma_c)$ or $(\bar{u},\bar{q}) \in \mathbf{H}^2(\Omega) \times \mathbb{R}^l$, respectively.

4.4 Second-order sufficient optimality conditions for the finite-dimensional case

Almost analogously to the infinite-dimensional case, we are able to consider the second-order sufficient optimality conditions with the control space Q_{ad}. We refer also to [31] for this case.

Let us therefore define $d := (\alpha M\bar{q} - D)^T$, similar to (4.3.2), with the mass matrix M and a vector D

$$M \in \mathbb{R}^{l,l},\ M_{i,j} = \int\limits_{\Gamma_c} e_i e_j \,\mathrm{d}s,\ D \in \mathbb{R}^l,\ D_i = \int\limits_{\Gamma_c} \left(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi\mathfrak{n} + (\bar{u}\cdot\mathfrak{n})\lambda\right) e_i \,\mathrm{d}s.$$

Then (4.3.4b) is equivalent to

$$(\alpha M\bar{q} - D)^T(q - \bar{q}) \geq 0\ \forall q \in Q_{ad}, \tag{4.4.1}$$

which is the necessary and sufficient optimality condition of the quadratic programming problem

$$\min_{q\in Q_{ad}} \frac{\alpha}{2} q^T M q - D^T q.$$

Additionally, we introduce $\mathcal{A}_+ := \{i : d_i > 0\}$, $\mathcal{A}_- := \{i : d_i < 0\}$ and $\mathcal{A} := \mathcal{A}_+ \cup \mathcal{A}_-$ and the critical cone associated with $\bar{q}$

$$C_{\bar{q}} := \{h \in \mathbb{R}^l :\ h_i = 0\ \forall i \in \mathcal{A} \text{ satisfying } (4.4.2) - (4.4.4)\}$$

$$\begin{aligned} -\nu\Delta z + (\bar{u}\cdot\nabla)z + (z\cdot\nabla)\bar{u} + \nabla\mu &= 0 && \text{in } \Omega\\ \operatorname{div} z &= 0 && \text{in } \Omega\\ z(x) &= \sum_{j=1}^{l} e_j(x)h_j && \text{on } \Gamma_c,\\ z &= 0 && \text{on } \Gamma\setminus\Gamma_c, \end{aligned} \tag{4.4.2}$$

$$\begin{cases} F'_{\vec{e}_d}(\bar{u})(z) = 0, & \text{if } F_{\vec{e}_d}(\bar{u}) = D_0 \text{ and } \xi > 0\\ F'_{\vec{e}_d}(\bar{u})(z) \leq 0, & \text{if } F_{\vec{e}_d}(\bar{u}) = D_0 \text{ and } \xi = 0, \end{cases}, \tag{4.4.3}$$

$$h_i = \begin{cases} \geq 0 \text{ if } & \bar{q}_i = q_{a,i}\\ \leq 0 \text{ if } & \bar{q}_i = q_{b,i} \end{cases}. \tag{4.4.4}$$

Then, we define the coercivity assumption (SSC'):

$$(SSC')\begin{cases} \text{The inequality}\\ h^T\mathcal{L}_{qq}(\bar{q},\xi)h > 0\\ \text{holds for all } h \in C_{\bar{q}}\setminus\{0\}. \end{cases}$$

and derive the following theorem

Theorem 4.12. *Let $(\bar{u}, \bar{g})$ be an admissible non-singular point for the optimal control problem and fulfill the first-order necessary optimality condition of Theorem 4.11 including the variational inequality* (4.3.2) *with associated*

λ, ξ. *Assume furthermore that the coercivity assumption (SSC') is satisfied. Then there exist* $\delta > 0$ *and* $\tau > 0$ *such that*

$$J(v) \geq J(\bar{v}) + \delta |q - \bar{q}|^2$$

holds for all admissible pairs (u, g) *with* $|q - \bar{q}| \leq \tau$.

Chapter 5

Numerical investigations

In this chapter, we want to provide numerical results for the problem under consideration. We handle the optimal control problem above numerically by direct solution of the optimality system that follows from Theorem 4.11 and is stated below. We follow a method suggested by Neitzel et al. [73].

Afterwards, we consider an SQP-method, whose convergence is proved in Chapter 6.

5.1 One-shot approach

Here, we consider a slightly different problem. We have an inflow g_∞ acting as an inhomogeneous Dirichlet boundary condition on the inflow boundary Γ_{in}. The control boundary Γ_c was modelled by a nonhomogeneous Dirichlet condition, where the limited suction and/or blowing occurs on small slot on the flap. A no-slip boundary condition, i.e. homogeneous Dirichlet condition, was used for the remaining airfoil Γ_w and the wall boundaries Γ_{wall}. At the outflow Γ_{out}, we applied a so called 'do-nothing'-condition:

$$\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} = 0.$$

For more details of the configuration see the technical report [19].

These do-nothing conditions have similar properties as Neumann boundary conditions for scalar elliptic equations. Let us briefly comment on available results for Navier-Stokes equations with mixed boundary conditions of Dirichlet and Neumann type. For the Navier-Stokes system with mixed boundary conditions, existence and uniqueness of solutions $u \in \mathbf{H}^1(\Omega)$ for Dirichlet data $g \in \mathbf{H}^{1/2}(\Gamma_c)$ were proven in [72], Theorem 5.2, and [71]. We can show, similar to Theorem 2.17, that we obatin a solution $u \in \mathbf{H}^2(\Omega)$ for

$g \in \mathbf{H}^{3/2}(\Gamma_c)$. Due to the 'do-nothing' boundary condition on the outflow boundary, the pressure $p \in \mathbf{L}^2(\Omega)$ is unique. To the best of our knowledge, similar results for low-regularity Dirichlet data in $\mathbf{L}^2(\Gamma_c)$ are still missing. If one carries out the formal procedure as described at the end of the finite-dimensional part or see (5.2.5), one finds that the adjoint equation for the problem with 'do-nothing' outflow condition is given by

$$\begin{aligned} -\nu\Delta\lambda + (\nabla\bar{u})^T\lambda - (\bar{u}\cdot\nabla)\lambda + \nabla\pi &= 0 && \text{in } \Omega \\ \operatorname{div} u &= 0 && \text{in } \Omega \\ \lambda &= \vec{e}_l - \xi\vec{e}_d && \text{on } \Gamma_w \\ \lambda &= 0 && \text{on } \Gamma \setminus (\Gamma_w \cup \Gamma_{out}) \\ \nu\frac{\partial\lambda}{\partial\mathfrak{n}} - \pi\mathfrak{n} + (\bar{u}\cdot\mathfrak{n})\lambda &= 0 && \text{on } \Gamma_{out}. \end{aligned}$$

This adjoint system is analogously to the one of Theorem 4.11.

As already mentioned above, we solved the optimality system given analogously to Theorem 4.11. Due to the presence of the 'do-nothing' boundary condition, we can drop the constraint $\int_{\Gamma_j} g\cdot\mathfrak{n}\,\mathrm{d}s = 0$, which was incorporated to guarantee existence of divergence-free solutions. With this simplification, the variational inequality and the complementarity condition in the optimality system given by Theorem 4.11 are equivalent to

$$\bar{g} = \mathbb{P}_G\left\{\frac{1}{\alpha}(\nu\frac{\partial\lambda}{\partial n} - \pi n + (\bar{u}\cdot n)\lambda)\right\} \text{ a.e. in } \Gamma_c$$

and

$$\xi = \max(0, \xi + F_{\vec{e}_d}(\bar{u}) - D_0).$$

This enables us to eliminate the control variable by means of the projection. Then we want to solve the following system consisting of the state equation with control eliminated by the projection formula, the equation for the

Lagrange multiplier ξ and the associated adjoint equation, see also [19]:

$$\begin{aligned}
-\nu\Delta u + (u\cdot\nabla)u + \nabla p &= 0 && \text{in } \Omega\\
\operatorname{div} u &= 0 && \text{in } \Omega\\
u = \mathbb{P}_G\{\frac{1}{\alpha}(\nu\frac{\partial\lambda}{\partial n} - \pi n + (u\cdot n)\lambda)\} & && \text{on } \Gamma_c\\
u &= 0 && \text{on } \Gamma_{wall}\cup\Gamma_w\setminus\Gamma_c\\
u &= g_\infty && \text{on } \Gamma_{in}\\
\nu\frac{\partial u}{\partial n} - pn &= 0 && \text{on } \Gamma_{out}\\
-\nu\Delta\lambda + (\nabla u)^T\lambda - (u\cdot\nabla)\lambda + \nabla\pi &= 0 && \text{in } \Omega\\
\operatorname{div}\lambda &= 0 && \text{in } \Omega\\
\lambda &= \vec{e}_l - \xi\vec{e}_d && \text{on } \Gamma_w\\
\lambda &= 0 && \text{on } \Gamma_{in}\cup\Gamma_{wall}\\
\nu\frac{\partial\lambda}{\partial n} - \pi n + (u\cdot n)\lambda &= 0 && \text{on } \Gamma_{out}
\end{aligned}$$

and

$$\xi = \max(0, \xi + F_{\vec{e}_d}(u) - D_0). \tag{5.1.1}$$

5.1.1 Numerical results

The computational domain, depicted in Figure 5.1, is a 2D generic high-lift configuration consists of a NACA4412 main airfoil at 8° angle of attack and a NACA4415 flap with a deflection angle of 37°. The Reynolds number was given as $Re = 85$ based on the chord length $L_{ref} = 1.275$ and the free stream velocity $g_\infty = 1$.

We used the commercial FEM-solver COMSOL Multiphysics with a build-in damped Newton method for the nonlinear system. The partial differential equations were discretized using Taylor-Hood finite elements, i.e. piecewise quadratic polynomials for the velocity and piecewise linear polynomials for the pressure.

The equation for ξ is in this form not solvable, because on the one hand we only can define variables on the whole area Ω and on the other hand COMSOL does not allow variables in $\mathbb{R}$. Therefore, we choose the following algorithm. In the first step, we solve the optimal control problem without the state constraint. If for the computed solution the state constraint is satisfied then this solution solves also the state-constrained problem. Otherwise, we

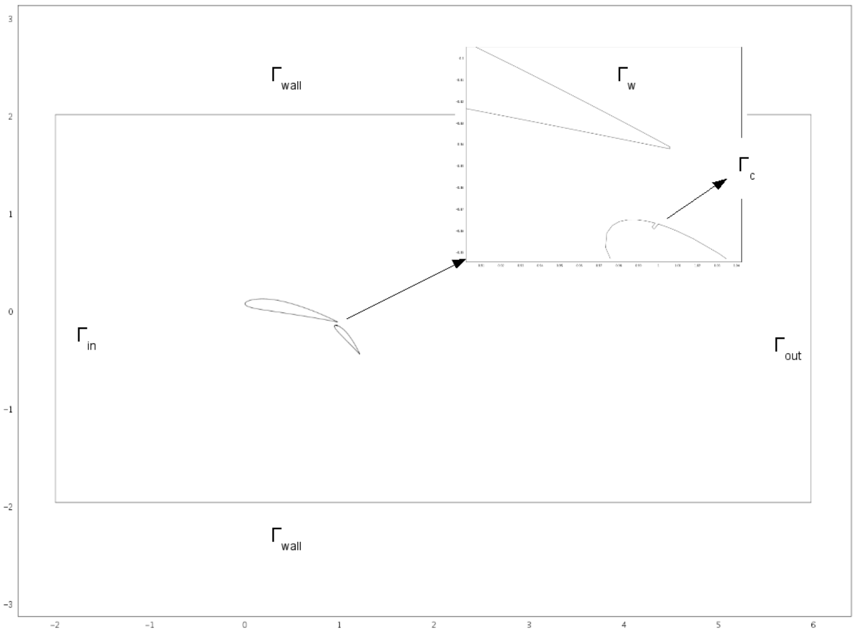

Figure 5.1: The domain

have to consider the state constraint

$$\int_{\Gamma_w} \nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \cdot \vec{e}_d \, \mathrm{d}s = D_0.$$

Integral terms in the system and COMSOL Multiphysics don't fit together. To handle this problem, we consider an augmented Lagrangian method for this problem, see [13, 46, 78]. This algorithm is related to the Penalty problem. The difference is that we here reduce the risk of ill-conditioned subproblems, because now we introduce Lagrange multiplier estimates at each step to the cost functional. The penalty term does not guarantee

$$F_{\vec{e}_d}(u,p) = D_0. \tag{5.1.2}$$

It only leads to

$$(F_{\vec{e}_d}(\bar{u},\bar{p}) - D_0) = -\frac{1}{c}\xi, \tag{5.1.3}$$

see [78], (17.45). So, we see (5.1.2) is theoretical fulfilled for $c \to \infty$, but we probably get ill conditioned and numerical problems for big values of c.

Let us ignore for a while the Navier-Stokes equations and just consider the cost functional, which we want to minimize, subject to the integral state constraint. Then the problem is

$$(P_1) \quad \min_{g \in G_{ad,Q}} J(u,p,g) \;\text{ subject to } F_{\vec{e}_d}(u,p) = D_0$$

with the optimality system

$$\nabla J(\bar{u},\bar{p}) + \xi F'_{\vec{e}_d}(\bar{u},\bar{p}) = 0. \tag{5.1.4}$$

The associated Penalty problem reads as

$$(P_2) \quad \min_{g \in G_{ad,q}} J(u,p,g) + c(F_{\vec{e}_d}(u,p) - D_0)^2$$

with the optimality system

$$\nabla J(\bar{u},\bar{p}) + 2cF'_{\vec{e}_d}(\bar{u},\bar{p})(F_{\vec{e}_d}(\bar{u},\bar{p}) - D_0) = 0 \tag{5.1.5}$$

The augmented Lagrangian function $\mathcal{L}_A(u,g,\xi)$ avoids the problem that we need very big penalty parameter c by an estimation for the Lagrange-multiplier ξ for the integral state constraint. The augment Lagrangian $\mathcal{L}_A(u,g,\xi)$ consists of the original cost functional $J(u,g)$, the penalty term and the term involving the multiplier ξ:

$$\mathcal{L}_A(u,g,\xi) := J(u,g) - \xi\left(F_{\vec{e}_d}(u,p) - D_0\right) + c\left(F_{\vec{e}_d}(u,p) - D_0\right)^2.$$

Considering $\mathcal{L}_A(u,g,\xi)$ as the new cost functional, we obtain the optimality system

$$\nabla_{(u,p)} J(u,g) - F'_{\vec{e}_d}(\bar{u},\bar{p})(\xi - 2c\left(F_{\vec{e}_d}(\bar{u},\bar{p}) - D_0\right)) = 0. \tag{5.1.6}$$

Now, we fix the penalty parameter c and the Lagrangian multiplier ξ in each step by c and ξ^k, respectively. The optimality systems (5.1.4), (5.1.5) and (5.1.6) leads to

$$\xi \approx \xi^k - 2c\left(F_{\vec{e}_d}(\bar{u},\bar{p}) - D_0\right) \tag{5.1.7}$$

which is equivalent to

$$(F_{\vec{e}_d}(\bar{u},\bar{p}) - D_0) \approx \frac{1}{2c}(\xi^k - \xi) \tag{5.1.8}$$

and we see that if ξ^k is close to the original Lagrange multiplier ξ, (5.1.8) is closer to $F_{\vec{e}_d}(\bar{u},\bar{p}) = D_0$ than (5.1.3). From (5.1.7), we get also the prescript to calculate the new Lagrange multiplier in the $k+1$th step by

$$\xi^{k+1} = \xi^k - 2c\left(F_{\vec{e}_d}(\bar{u},\bar{p}) - D_0\right).$$

Arada, Raymond and Tröltzsch proved in [13] the convergence of the augmented Lagrangian method for a class of problems.

If the integral state constraint is fulfilled, then we are able to reduce the associated Lagrange-multiplier ξ. Or otherwise, we have to increase the multiplier.

For the uncontrolled problem, we obtained a lift of $C_A = \frac{F_A}{0.5 g_\infty^2 L_{ref}} = 1.562$ and a drag of $C_D = \frac{D_A}{0.5 g_\infty^2 L_{ref}} = 0.817$, where F_A is the resulting lift force, D_A the drag force and $L_{ref} = 1.275$ is the reference length of the wing, see Figure 5.2 for a streamline plot of the velocity field, and Figure 5.4 for a plot of the absolute values of the uncontrolled velocity field and Figure 5.5 for the pressure field.

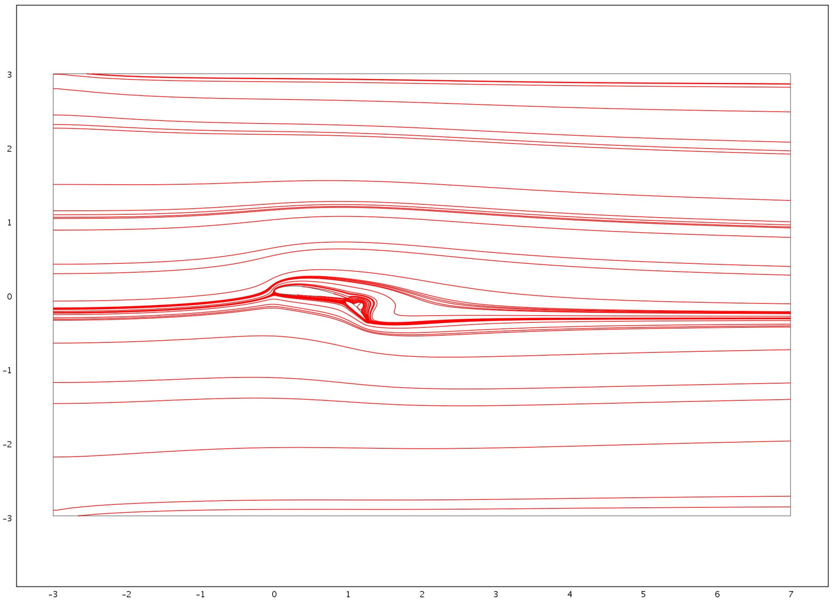

Figure 5.2: Uncontrolled case: velocity field.

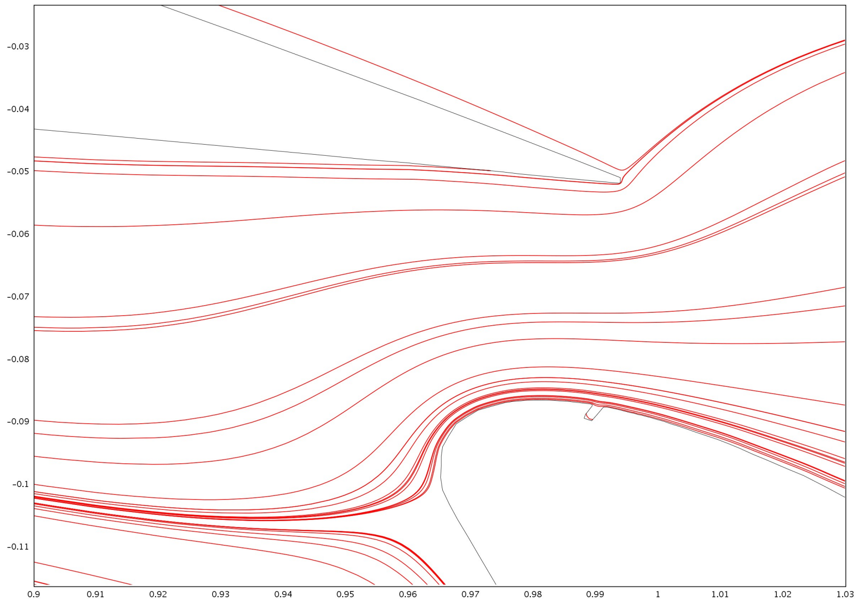

Figure 5.3: Uncontrolled case: velocity field with zoom on the wing.

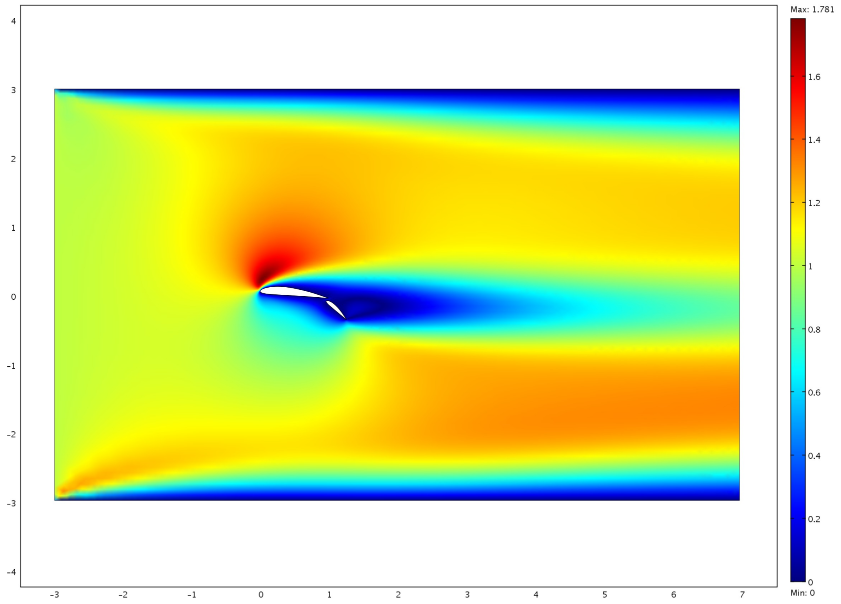

Figure 5.4: Uncontrolled case: absolute value of velocity field (left).

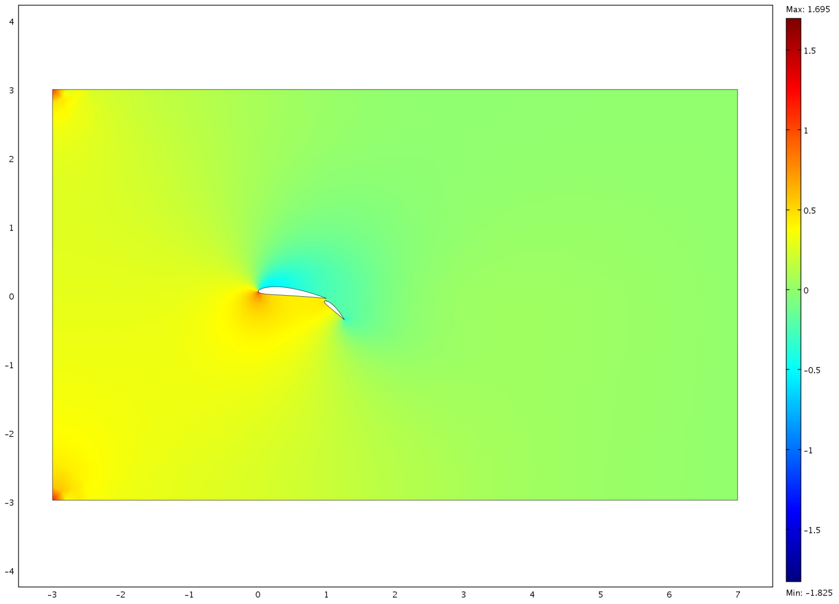

Figure 5.5: Uncontrolled case: absolute value of pressure field (right).

Now let us report about the result of the optimization. Here, we choose the control cost parameter $\alpha = 0.1$ and the control constraints as box constraints $G = [-1, +1]$.

At first, we compute the solution for the case without any drag constraint. The optimal control is given by the maximal possible suction, which is natural from a physical point of view. The obtained optimized lift is $C_A = 1.5823$ and the drag is $C_D = 0.8340$, which is a lift gain of 1.3%. The controlled velocity field can be seen in Figure 5.6. The adjoint velocity field and pressure are plotted in Figures 5.8 and 5.9.

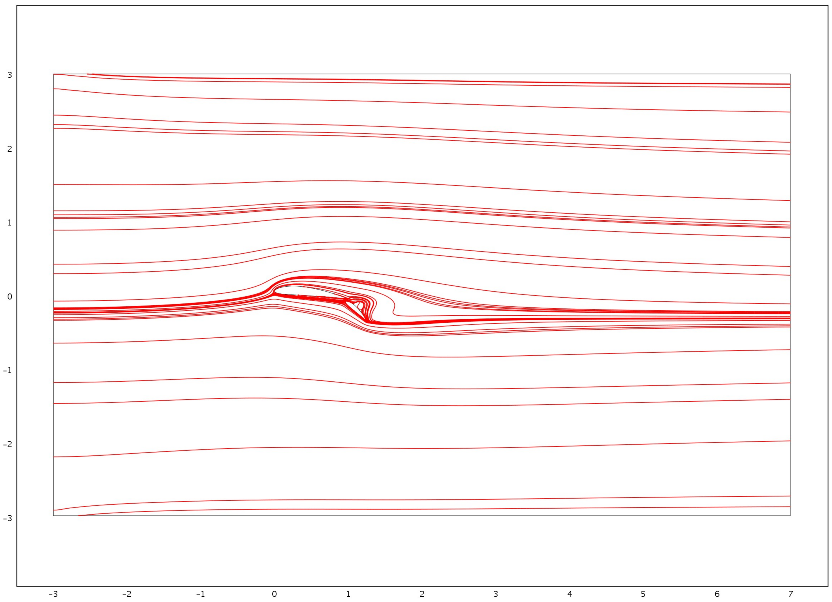

Figure 5.6: Controlled case: velocity field.

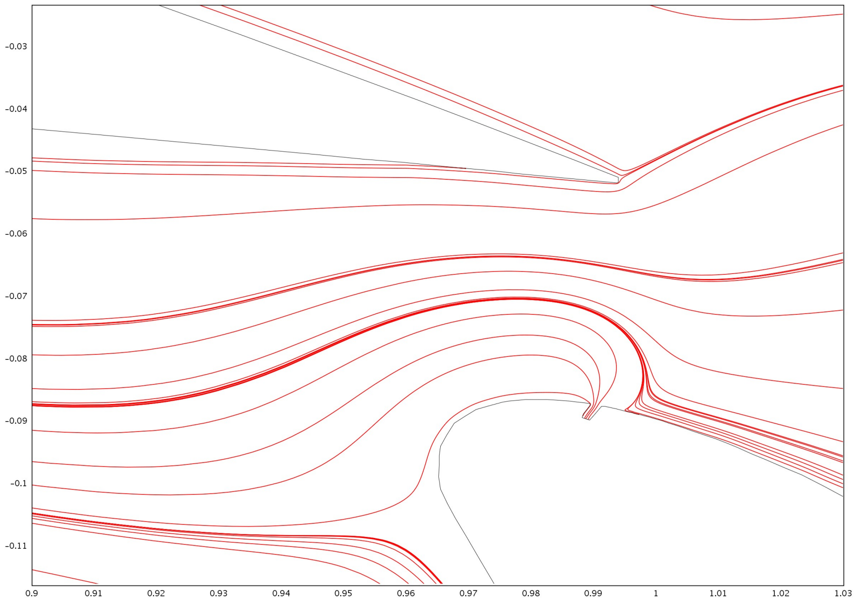

Figure 5.7: Controlled case: velocity field with zoom on the wing.

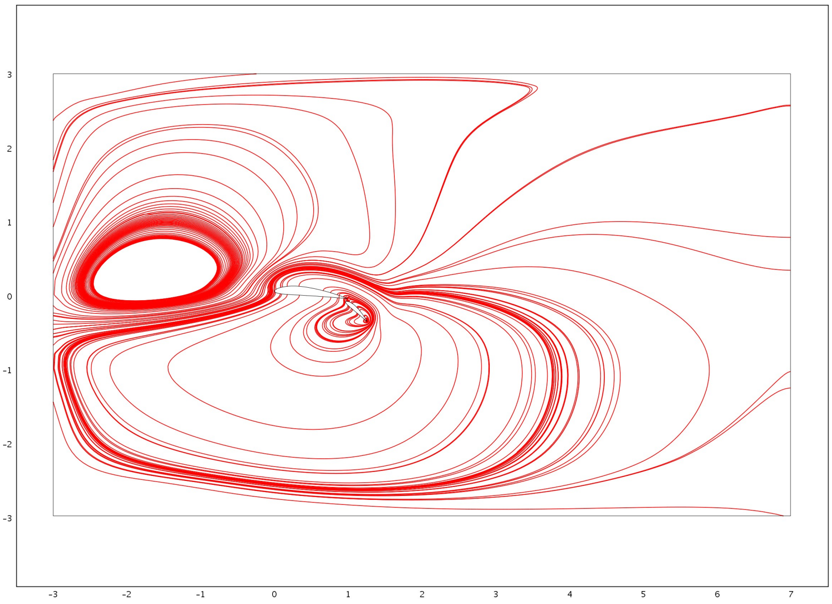

Figure 5.8: Controlled case: adjoint velocity field

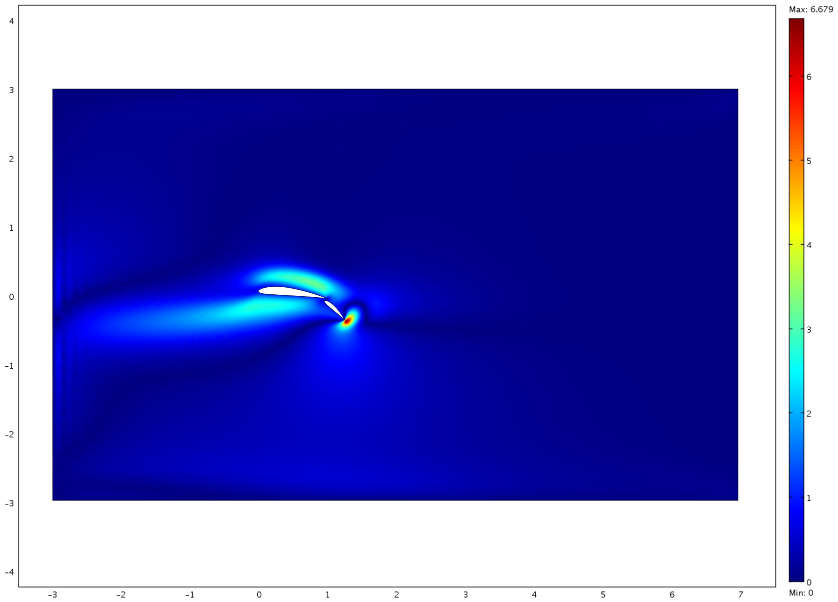

Figure 5.9: Controlled case: absolute value of adjoint velocity field (left).

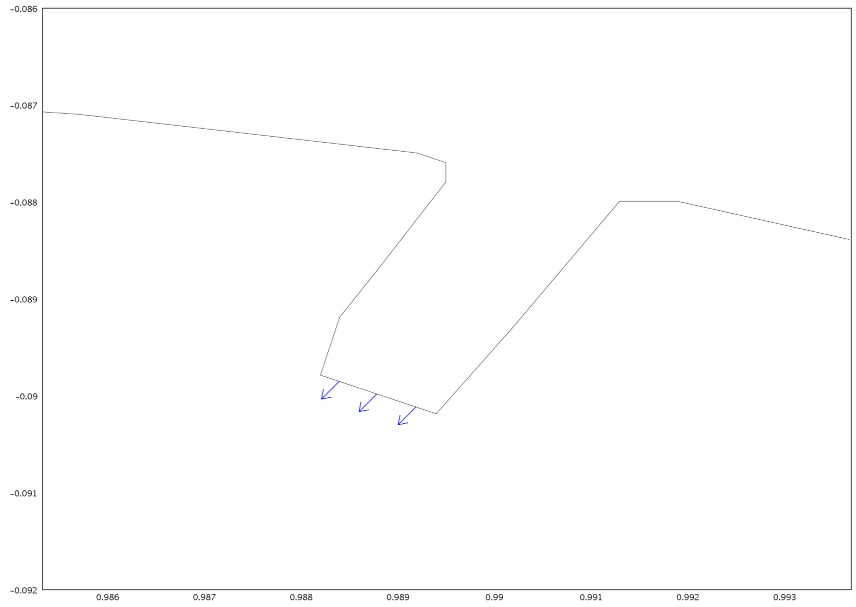

Figure 5.10: Controlled case: image of the slit.

The box constraint $G = [-0.5, +0.5]$ leads to a lift coefficient of $C_A = 1.572$ and a drag coefficient of $C_D = 0.825$, which amounts to a lift gain of about 0.65%.

In the next step, we choose $D_0 = 0.5240$ as upper bound for the drag. This equates to an constraint coefficient of $C_{D_0} = 0.8220$. Hence, we expect that this constraint will be active at the solution. In fact, for the computed solution we obtain $C_D = 0.8215$. Moreover, due to this restriction the computed lift is $C_A = 1.571$, which is less than for the case without state constraints, but which is still better than in the uncontrolled situation. An upper bound for the drag of $D_0 = 0.5230$, $C_{D_0} = 0.820$ leads to $C_d = 0.820$. and $C_a = 1.570$.

An upper bound for the drag of $D_0 = 0.5190$, $C_{D_0} = 0.814$ leads to $C_d = 0.814$. and $C_a = 1.559$.

5.2 SQP-method

In addition to the last section, where we solved the optimality system at once, we will now solve the problem by an iterative method. A widely-used method is the SQP-method.

5.2.1 The SQP-method for our problem

In this case, we do not solve the original problem. The idea is to solve a slightly different problem P_I, where the integral state constraint is added as penalty term in the cost functional. We do this, because we want to avoid the integral equation for the Lagrange multiplier ξ, see (5.1.1).

Analogously to the one-shot approach, we first solve the problem without a state constraint, this means in this case without the penalty term, and if the state constraint is satisfied for the computed solution then this solution also solves the state-constraint problem. Otherwise, it is sufficient to consider the state constraint as an equality

$$\int_{\Gamma_w} \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \right) \cdot \vec{e}_d \, \mathrm{d}s = D_0.$$

So, we can avoid the penalty term $c \left(\int\limits_{\Gamma_w} \left(\nu \dfrac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \right) \cdot \vec{e}_d \, \mathrm{d}s - D_0 \right)_+^2$, where $(s)_+ := \max\{0, s\}$, $s \in \mathbb{R}$, denotes the positive part. The problem is that this term is not twice continuously Fréchet-differentiable.

In this case, we introduce the penalty term to the cost functional and regard the following problem: Minimize

$$\begin{aligned} J^I(u,g) :=& \int\limits_{\Gamma_w} \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \right) \cdot \vec{e}_l \, \mathrm{d}s \\ &+ c \left(\int\limits_{\Gamma_w} \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \right) \cdot \vec{e}_d \, \mathrm{d}s - D_0 \right)^2 + \frac{\alpha}{2} \|g\|^2_{\mathbf{L}^2(\Gamma_c)} \end{aligned} \tag{5.2.1}$$

subject to

$$\begin{aligned} -\nu\Delta u + (u \cdot \nabla)u + \nabla p &= 0 && \text{in } \Omega \\ \operatorname{div} u &= 0 && \text{in } \Omega \\ u &= g && \text{on } \Gamma_c \\ u &= 0 && \text{on } \Gamma_{wall} \cup \Gamma_w \setminus \Gamma_c \\ u &= g_\infty && \text{on } \Gamma_{in} \\ \nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} &= 0 && \text{on } \Gamma_{out} \end{aligned} \tag{5.2.2}$$

the control constraints

$$g \in G_{ad,Q} \tag{5.2.3}$$

where $c \in \mathbb{R}$ is the penalty constant and $g_\infty \in \mathbb{R}$ the inflow. The solution u is defined in a (very) weak sense. We have shown in Section 2, Theorem 2.17, that we get for every $g \in \mathbf{H}^{3/2}(\Gamma_c)$ a solution $u \in \mathbf{H}^2(\Omega)$ for the state equation (5.2.2). The optimality system could be derived very similar to the problem with the integral state constraint. The Lagrangian looks as follows:

$$
\begin{aligned}
\mathcal{L}(u,p,g,\lambda) = & \int\limits_{\Gamma_w} \left(\nu \frac{\partial u}{\partial n} - p\mathfrak{n} \right) \cdot \vec{e}_l \, \mathrm{d}s + \frac{\alpha}{2} \int\limits_{\Gamma_c} |g|^2 \mathrm{d}s \\
& + c \left(\int\limits_{\Gamma_w} \left(\nu \frac{\partial u}{\partial n} - p\mathfrak{n} \right) \cdot \vec{e}_d \, \mathrm{d}s - D_0 \right)^2 \\
& + \int\limits_{\Omega} (-\nu \Delta u + (u \cdot \nabla) u) \lambda + \nabla p \lambda \mathrm{d}x - \int\limits_{\Omega} \pi \operatorname{div} u \mathrm{d}x + \int\limits_{\Gamma_c} (u-g)\lambda_2 \mathrm{d}s \\
& + \int\limits_{\Gamma_{wall}} u\lambda_3 \, \mathrm{d}s + \int\limits_{\Gamma_{in}} (u - g_\infty)\lambda_4 \, \mathrm{d}s + \int\limits_{\Gamma_{out}} \left(\nu \frac{\partial u}{\partial n} - pn \right) \lambda_5 \, \mathrm{d}s.
\end{aligned}
$$

with

$$
\int\limits_{\Omega} (u \cdot \nabla) u \cdot \lambda \, \mathrm{d}x = - \int\limits_{\Omega} (u \cdot \nabla)\lambda \cdot u + u \cdot \lambda \operatorname{div} u \, \mathrm{d}x + \int_{\Gamma} u \cdot \mathfrak{n} u \cdot \lambda \, \mathrm{d}\gamma.
$$

The necessary condition $\frac{\partial \mathcal{L}}{\partial(u,p)}(\bar{u},\bar{p})(u,p) = 0$ leads to

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial(u,p)}(\bar{u},\bar{p})(u,p) = & \int\limits_{\Gamma_w} \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \right) \cdot \vec{e}_l \, \mathrm{d}s \\
& + \underbrace{2c \left(\int\limits_{\Gamma_w} \left(\nu \frac{\partial \bar{u}}{\partial \mathfrak{n}} - \bar{p}\mathfrak{n} \right) \cdot \vec{e}_d \, \mathrm{d}s - D_0 \right)}_{:=K(\bar{u},\bar{p}))} \left(\int\limits_{\Gamma_w} \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \right) \cdot \vec{e}_d \, \mathrm{d}s \right) \\
& + \int\limits_{\Omega} -\nu \Delta u \lambda - (\bar{u} \cdot \nabla)\lambda \cdot u - (u \cdot \nabla)\lambda \cdot \bar{u} - \bar{u} \cdot \lambda \operatorname{div} u + \nabla p \lambda \mathrm{d}x \\
& - \int\limits_{\Omega} \pi \operatorname{div} u \, \mathrm{d}x + \int\limits_{\Gamma_c} u\lambda_2 \, \mathrm{d}s + \int\limits_{\Gamma_{wall}} u\lambda_3 \, \mathrm{d}s + \int\limits_{\Gamma_\infty} u\lambda_4 \, \mathrm{d}s \\
& + \int\limits_{\Gamma_{out}} \left(\nu \frac{\partial u}{\partial n} - pn \right) \lambda_5 \, \mathrm{d}s + \int\limits_{\Gamma_c \cup \Gamma_{out}} \bar{u} \cdot \mathfrak{n} u \cdot \lambda + u \cdot \mathfrak{n} \bar{u} \cdot \lambda \, \mathrm{d}s = 0.
\end{aligned} \tag{5.2.4}
$$

This is with partial integration and

$$K(\bar{u},\bar{p}) = 2c\left(\int\limits_{\Gamma_w}\left(\nu\frac{\partial\bar{u}}{\partial\mathfrak{n}} - \bar{p}\mathfrak{n}\right)\cdot\vec{e}_d\,\mathrm{d}s - D_0\right)$$

equivalent to

$$\begin{aligned}
&\int\limits_{\Gamma_w}\left(\nu\frac{\partial u}{\partial\mathfrak{n}} - p\mathfrak{n}\right)\cdot\vec{e}_l\,\mathrm{d}s + K(\bar{u},\bar{p})\int\limits_{\Gamma_w}\left(\nu\frac{\partial u}{\partial\mathfrak{n}} - p\mathfrak{n}\right)\cdot\vec{e}_d\,\mathrm{d}s\\
&-\int\limits_{\Omega}(\nu\nabla u\cdot\nabla\lambda - (\bar{u}\cdot\nabla)\lambda\cdot u - (u\cdot\nabla)\lambda\cdot\bar{u}\\
&\qquad\qquad + \bar{u}\cdot\lambda\operatorname{div}u - p\operatorname{div}\lambda + u\cdot\nabla\pi)\,\mathrm{d}x\\
&+\int\limits_{\Gamma_c}\left(-\nu\frac{\partial u}{\partial\mathfrak{n}}\lambda + p\mathfrak{n}\lambda - \pi\mathfrak{n}u + u\mathfrak{n}\lambda_2\right)\,\mathrm{d}s\\
&+\int\limits_{\Gamma_{wall}}\left(-\nu\frac{\partial u}{\partial\mathfrak{n}}\lambda + p\mathfrak{n}\lambda - \pi\mathfrak{n}u + u\mathfrak{n}\lambda_3\right)\,\mathrm{d}s\\
&+\int\limits_{\Gamma_{in}}\left(-\nu\frac{\partial u}{\partial\mathfrak{n}}\lambda + p\mathfrak{n}\lambda - \pi\mathfrak{n}u + u\mathfrak{n}\lambda_4\right)\,\mathrm{d}s\\
&+\int\limits_{\Gamma_{out}}\left(-\nu\frac{\partial u}{\partial\mathfrak{n}}(\lambda-\lambda_5) + p\mathfrak{n}(\lambda-\lambda_5) + \pi\mathfrak{n}u\right)\,\mathrm{d}s\\
&+\int\limits_{\Gamma_c\cup\Gamma_{out}}\bar{u}\cdot\mathfrak{n}u\cdot\lambda + u\cdot\mathfrak{n}\bar{u}\cdot\lambda\,\mathrm{d}s = 0.
\end{aligned}$$

The equality

$$\begin{aligned}
-\int\limits_{\Omega}((u\cdot\nabla)\lambda\cdot\bar{u} + \bar{u}\cdot\lambda\operatorname{div}u)\,\mathrm{d}x + \int\limits_{\Gamma_c\cup\Gamma_{out}}(u\cdot\mathfrak{n})(\bar{u}\cdot\lambda)\,\mathrm{d}s&\\
= \int\limits_{\Omega}(\nabla\bar{u})^T\lambda\cdot u\,\mathrm{d}x&
\end{aligned}$$

and another partial integration lead to

$$\begin{aligned}
&\int\limits_{\Gamma_w} \left(\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n}\right)\cdot \vec{e}_l \,\mathrm{d}s + K(\bar{u},\bar{p})\int\limits_{\Gamma_w} \left(\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n}\right)\cdot \vec{e}_d \,\mathrm{d}s \\
&-\int\limits_{\Omega} \big(-\nu\Delta\lambda + (\bar{u}\cdot\nabla)\lambda + (\nabla\bar{u})^T\lambda + \nabla\pi\big)\, u \,\mathrm{d}x - \int\limits_{\Omega} p \operatorname{div}\lambda \,\mathrm{d}x \\
&-\int\limits_{\Gamma_c} \left(\nu\frac{\partial \lambda}{\partial \mathfrak{n}}u - \nu\frac{\partial u}{\partial \mathfrak{n}}\lambda + p\mathfrak{n}\lambda - \pi\mathfrak{n}u + u\mathfrak{n}\lambda_2 + \bar{u}\cdot\mathfrak{n}u\cdot\lambda\right)\,\mathrm{d}s \\
&-\int\limits_{\Gamma_{wall}} \left(\nu\frac{\partial \lambda}{\partial \mathfrak{n}}u - \nu\frac{\partial u}{\partial \mathfrak{n}}\lambda + p\mathfrak{n}\lambda - \pi\mathfrak{n}u + u\mathfrak{n}\lambda_3\right)\,\mathrm{d}s \\
&-\int\limits_{\Gamma_{in}} \left(\nu\frac{\partial \lambda}{\partial \mathfrak{n}}u - \nu\frac{\partial u}{\partial \mathfrak{n}}\lambda + p\mathfrak{n}\lambda - \pi\mathfrak{n}u + u\mathfrak{n}\lambda_4\right)\,\mathrm{d}s \\
&-\int\limits_{\Gamma_{out}} \left(\nu\frac{\partial \lambda}{\partial \mathfrak{n}}u - \nu\frac{\partial u}{\partial \mathfrak{n}}(\lambda-\lambda_5) + p\mathfrak{n}(\lambda-\lambda_5) + \pi\mathfrak{n}u + \bar{u}\cdot\mathfrak{n}u\cdot\lambda\right)\,\mathrm{d}s \\
&= 0.
\end{aligned}$$

Taking $(u,p) \in \mathbf{H}_0^2(\Omega)\times\mathbf{H}^1(\Omega)$, we obtain

$$\begin{aligned}
&\int\limits_{\Omega} \big(-\nu\Delta\lambda + (\bar{u}\cdot\nabla)\lambda + (\nabla\bar{u})^T\lambda + \nabla\pi\big)\, u \,\mathrm{d}x - \int\limits_{\Omega} p \operatorname{div}\lambda \,\mathrm{d}x \\
&+\int\limits_{\Gamma_w} (\vec{e}_l + K(\bar{u},\bar{p})\vec{e}_d - \lambda)(\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n})\,\mathrm{d}s + \int\limits_{\Gamma_{wall}} \lambda(\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n})\,\mathrm{d}s \\
&+\int\limits_{\Gamma_{in}} \lambda(\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n})\,\mathrm{d}s + \int\limits_{\Gamma_{out}} (\lambda-\lambda_5)(\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n})\,\mathrm{d}s = 0
\end{aligned}$$

so that λ and π are the weak solutions of

$$\begin{aligned}
-\nu\Delta\lambda + (\nabla\bar{u})^T\lambda - (\bar{u}\cdot\nabla)\lambda + \nabla\pi &= 0 && \text{in } \Omega \\
\operatorname{div}\lambda &= 0 && \text{in } \Omega \\
\lambda &= \vec{e}_l + K(\bar{u},\bar{p})\vec{e}_d && \text{on } \Gamma_w \qquad (5.2.5)\\
\lambda &= 0 && \text{on } \Gamma_{in}\cup\Gamma_{wall} \\
\lambda &= \lambda_5 && \text{on } \Gamma_{out}.
\end{aligned}$$

Taking $(u,p) \in \mathbf{H}^2(\Omega) \times \mathbf{H}^1(\Omega)$, we obtain

$$\int_{\Gamma_w} \left(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} u - \pi \mathfrak{n} u + u \mathfrak{n} \lambda_2 + \bar{u} \cdot \mathfrak{n} u \cdot \lambda \right) \,\mathrm{d}s = 0,$$

$$\int_{\Gamma_{wall}} \left(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} u - \pi \mathfrak{n} u + u \mathfrak{n} \lambda_3 \right) \,\mathrm{d}s = 0,$$

$$\int_{\Gamma_{in}} \left(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} u - \pi \mathfrak{n} u + u \mathfrak{n} \lambda_4 \right) \,\mathrm{d}s = 0,$$

$$\int_{\Gamma_{out}} \left(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} u - \pi \mathfrak{n} u + \bar{u} \cdot \mathfrak{n} u \cdot \lambda \right) \,\mathrm{d}s = 0,$$

which implies

$$\begin{aligned} \lambda_2 \mathfrak{n} &= -(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi \mathfrak{n} + (\bar{u} \cdot \mathfrak{n}) \lambda) \text{ on } \Gamma_w, \\ \lambda_3 \mathfrak{n} &= -(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi \mathfrak{n}) \text{ on } \Gamma_{wall}, \\ \lambda_4 \mathfrak{n} &= -(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi \mathfrak{n}) \text{ on } \Gamma_{in}, \\ 0 &= -(\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi \mathfrak{n} + (\bar{u} \cdot \mathfrak{n}) \lambda) \text{ on } \Gamma_{out}. \end{aligned} \tag{5.2.6}$$

Now, we are able to substitute the last equation of (5.2.5) with $\nu \frac{\partial \lambda}{\partial \mathfrak{n}} + \pi \mathfrak{n} + \nu \frac{\partial \lambda}{\partial \mathfrak{n}} + \pi \mathfrak{n} + \bar{u} \cdot \mathfrak{n} u \cdot \lambda = 0$ on Γ_{out}.

The other necessary condition $\frac{\partial \mathcal{L}}{\partial g}(\bar{g})(g - \bar{g}) \geq 0$ for all $g \in G_{ad,Q}$ leads to

$$\frac{\partial \mathcal{L}}{\partial g}(\bar{g})(g - \bar{g}) = \int_{\Gamma_c} \alpha \bar{g}(g - \bar{g}) - (g - \bar{g}) \lambda_2 \,\mathrm{d}s \geq 0 \ \forall g \in G_{ad,Q}$$

which with (5.2.6) is equivalent to

$$\int_{\Gamma_c} \left(\alpha \bar{g} - (\nu \frac{\partial \lambda}{\partial \mathfrak{n}} - \pi \mathfrak{n} + (\bar{u} \cdot \mathfrak{n}) \lambda) \right) (g - \bar{g}) \,\mathrm{d}s \geq 0 \ \forall g \in G_{ad,Q}. \tag{5.2.7}$$

Altogether, we derive the adjoint system

$$\begin{aligned} -\nu \Delta \lambda + (\nabla \bar{u})^T \lambda - (\bar{u} \cdot \nabla) \lambda + \nabla \pi &= 0 && \text{in } \Omega \\ \operatorname{div} \lambda &= 0 && \text{in } \Omega \\ \lambda &= \vec{e}_l + K(\bar{u}, \bar{p}) \vec{e}_d && \text{on } \Gamma_w \\ \lambda &= 0 && \text{on } \Gamma_{in} \cup \Gamma_{wall} \\ \nu \frac{\partial \lambda}{\partial \mathfrak{n}} + \pi \mathfrak{n} + \bar{u} \cdot \mathfrak{n} u \cdot \lambda &= 0 && \text{on } \Gamma_{out} \end{aligned} \tag{5.2.8}$$

with $K(\bar{u},\bar{p}) = 2c\left(\int\limits_{\Gamma_w} \left(\nu\frac{\partial\bar{u}}{\partial\mathfrak{n}} - \bar{p}\mathfrak{n}\right)\cdot\vec{e}_d \,\mathrm{d}s - D_0\right)$ and the variational inequality

$$\int\limits_{\Gamma_c}\left(\alpha\bar{g} - (\nu\frac{\partial\lambda}{\partial\mathfrak{n}} - \pi\mathfrak{n} + (\bar{u}\cdot\mathfrak{n})\lambda)\right)(g-\bar{g})\,\mathrm{d}s \geq 0 \;\forall g \in G_{ad,Q}. \tag{5.2.9}$$

Let $(\bar{u},\bar{g},\bar{\lambda})$ satisfy the optimality system, consisting of the Navier-Stokes equations (5.2.2), the control constraint (5.2.3), the adjoint system (5.2.8) and the variational inequality (5.2.9).

First, we consider the problem without any restrictions to the control function g that means $G_{ad,Q} = \{v \in \mathbf{H}^{3/2}(\Gamma_c) : \; v(x) = \sum_{i=1}^{l} e_i(x)q_i, \; q \in \mathbb{R}^l\}$ or $Q_{ad} = \mathbb{R}^l$ and instead of the variational inequality (5.2.9), we have the equation $\alpha\bar{g} - (\nu\frac{\partial\lambda}{\partial\mathfrak{n}} + \pi\mathfrak{n} + (\bar{u}\cdot\mathfrak{n})\lambda) = 0$. So, the optimality systems reads as (5.2.2),(5.2.3),(5.2.8) and

$$\alpha\bar{g} - \left(\nu\frac{\partial\lambda}{\partial\mathfrak{n}} - \pi\mathfrak{n} + (\bar{u}\cdot\mathfrak{n})\lambda\right) = 0 \quad \forall g \in G_{ad,Q}.$$

This nonlinear system for (u,g,λ) can be solved with the Newton-method, see for instance [32]. Let us briefly sketch this method: consider the optimization problem

$$\min f(u) \;\; u \in \mathbb{R}^n$$

with $f \in C^2(\mathbb{R}^n)$.

Then, we want to solve the optimality system

$$f'(\bar{u}) = 0.$$

Solving this system with the Newton-method means that we get u_{n+1} as the solution of

$$f'(u_n) + f''(u_n)(u - u_n) = 0$$

When we transfer this idea to our problem, we obtain the following Newton-method.

Algorithm 5.1 **(NM).**

1. *Choose an initial value* $z^0 = (u_0, g_0, \lambda_0)$ *and set* $k = 0$.

2. Determine $z = (u, g, \lambda)$ by the Navier-Stokes system linearized at z^k

$$\begin{aligned}
-\nu\Delta u + (u\cdot\nabla)u^k + (u^k\cdot\nabla)u + \nabla p &= -(u^k\cdot\nabla)u^k && \textit{in } \Omega\\
\operatorname{div} u &= 0 && \textit{in } \Omega\\
u &= g && \textit{on } \Gamma_c\\
u &= 0 && \textit{on } \Gamma_{wall}\cup\Gamma_w\setminus\Gamma_c\\
u &= g_\infty && \textit{on } \Gamma_{in}\\
\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} &= 0 && \textit{on } \Gamma_{out},
\end{aligned}\tag{5.2.10a}$$

$$\begin{aligned}
-\nu\Delta\lambda + (\nabla\bar{u})^T\lambda - (\bar{u}\cdot\nabla)\lambda + \nabla\pi &= -(\nabla(\bar{u}-u^k))^T\lambda^k\\
&\quad + ((\bar{u}-u^k)\cdot\nabla)\lambda^k && \textit{in } \Omega\\
\operatorname{div}\lambda &= 0 && \textit{in } \Omega\\
\lambda &= \vec{e}_l + K(\bar{u},\bar{p})\vec{e}_d && \textit{on } \Gamma_w\\
\lambda &= 0 && \textit{on } \Gamma_{in}\cup\Gamma_{wall}\\
\nu\frac{\partial\lambda}{\partial n} - \pi n + (\bar{u}\cdot n)\lambda &= 0 && \textit{on } \Gamma_{out}.
\end{aligned}\tag{5.2.10b}$$

and

$$\alpha\bar{g} - \left(\nu\frac{\partial\lambda}{\partial\mathfrak{n}} - \pi\mathfrak{n} + (\bar{u}\cdot\mathfrak{n})\lambda\right) = 0 \tag{5.2.10c}$$

3. Set $k = k+1$ and $z^k = z$. Goto 2.

The equations (5.2.10) are equivalent to solving the quadratic problem

$$\begin{aligned}
\min J^k(u,g) := \nabla J(u^k,g^k)&(u-u^k, g-g^k)\\
&+\frac{1}{2}\mathcal{L}_{zz}(u^k,g^k,\lambda^k)[(u-u^k,g-g^k)]^2
\end{aligned}\tag{5.2.11}$$

subject to the linearized Navier-Stokes equation

$$\begin{aligned}
-\nu\Delta u + (u\cdot\nabla)u^k + (u^k\cdot\nabla)u + \nabla p &= -(u^k\cdot\nabla)u^k && \text{in } \Omega\\
\operatorname{div} u &= 0 && \text{in } \Omega\\
u &= g && \text{on } \Gamma_c\\
u &= 0 && \text{on } \Gamma_{wall}\cup\Gamma_w\setminus\Gamma_c\\
u &= g_\infty && \text{on } \Gamma_{in}\\
\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} &= 0 && \text{on } \Gamma_{out}.
\end{aligned}\tag{5.2.12}$$

So, we can also solve this system instead of (5.2.10). In contrast to the system above, the system (5.2.10) cannot be transformed to the case with

$g \in G_{ad,Q}$.
Instead, we have to solve in the n-th step (QP_n):

$$\begin{aligned}\min J^k(u,g) :=&\nabla J(u^k,g^k)(u-u^k,g-g^k)\\ &+\frac{1}{2}\mathcal{L}_{zz}(u^k,g^k,\lambda^k)[(u-u^k,g-g^k)]^2\end{aligned}$$

$$\begin{aligned}-\nu\Delta u+(u\cdot\nabla)u^k-(u^k\cdot\nabla)u+\nabla p &= -(u^k\cdot\nabla)u^k && \text{in } \Omega\\ \operatorname{div} u &= 0 && \text{in } \Omega\\ u&=g && \text{on } \Gamma_c\\ u&=0 && \text{on } \Gamma_{wall}\cup\Gamma_w\setminus\Gamma_c\\ u&=g_\infty && \text{on } \Gamma_{in}\\ \nu\frac{\partial u}{\partial \mathfrak{n}}-p\mathfrak{n}&=0 && \text{on } \Gamma_{out}\\ g&\in G_{ad,Q}.\end{aligned}$$

The cost functional, solved in the k-th step, J^k differs only by the term

$$\begin{aligned}\frac{1}{2}\mathcal{L}_{zz}(u^k,g^k,\lambda^k)[(u-u^k,g-g^k)]^2 =&(((u-u^k)\cdot\nabla)(u-u^k),\lambda^k)_{L^2(\Omega)}\\ &+2c(\nu\frac{\partial(u-u^k)}{\partial\mathfrak{n}}-(p-p^k)\mathfrak{n},\vec{e}_l)_{L^2(\Omega)}\end{aligned}$$

from the original cost functional J. Our idea is to solve the linear subproblems (QP_n) with the gradient-projection method.

Remark 5.2. *We know that it is not the best method; we solve the whole problem with the (fast) SQP-method and the subproblems with the (slow) gradient-projection method. A more appropriate choice would be for instance the active-set method. In this case, we would have to handle the integral term* $K(u,p) = 2c\left(\int_{\Gamma_w}\left(\nu\frac{\partial u}{\partial\mathfrak{n}}-p\mathfrak{n}\right)\cdot\vec{e}_d\,\mathrm{ds}-D_0\right)$. *But, we decided to solve our problems numerically with COMSOL Multiphysics. As mentioned before, this does not fit together. Therefore, we consider the gradient-projection method to calculate* $K(\bar{u},\bar{p})$ *after solving the state equation.*

5.2.2 Gradient-projection method

Let us briefly formulate the principle of this method for an optimization problem in a Hilbert space U,

$$\min_{u\in U_{ad}} f(u),$$

where $U_{ad}\subset U$ is a non-empty, bounded, convex and closed set and $f: U\to\mathbb{R}$ is a Gâteaux-differentiable functional.

The iteration steps $u_1, ..., u_n$ are finished so that u_n is the current solution. Then the algorithm look as follows:

S1 (*Direction search*) Choose the anti-gradient as descent-direction

$$v_n := -f'(u_n).$$

S2 (*Stepsize*) Choose a step size s_n, so that the equation

$$f(\mathbb{P}_{[u_a,u_b]}\{u_n + s_n v_n\}) = \min_{s>0} f(\mathbb{P}_{[u_a,u_b]}\{u_n + s v_n\})$$

is fulfilled, which guarantees the admissibility of the solution.

S3 Set $u_{n+1} = u_n + s_n v_n$, $n = n + 1$ and goto **S1**.

For the optimization subproblems (QP_n) the algorithm reads as follows

S1 Calculate (u_n, p_n) as the solution of (5.2.10a).

S2 Calculate the adjoint (λ_n, π_n) from (5.2.10b).

S3 The updated descent direction is

$$v_n := \alpha g_n - (\nu \frac{\partial \lambda}{\partial \mathfrak{n}} + \pi \mathfrak{n} + (\bar{u} \cdot \mathfrak{n})\lambda)$$

S4 Calculate the *step size* s_n from

$$\min_{s>0} f(\mathbb{P}_{[g_a(x),g_b(x)]}\{g_n + s v_n\}).$$

S5 The *updated control* g_{n+1} is

$$g_{n+1} := \mathbb{P}_{[g_a(x),g_b(x)]}\{g_n + s_n v_n\}.$$

set n:=n+1 and goto **S1**.

5.2.3 Example

Let us now consider the same example as for the one-shot approach with $g_a(x) \equiv -0.5$, $g_b(x) \equiv 0.5$ and a Reynolds number $Re = \frac{1}{\nu} L_{ref} = 85$.

With a drag constraint of $D_0 = 0.5240$ which is equal to a coefficient $C_{D_0} = 0.822$ and a penalty parameter of $c = 50$, we obtain $C_A = 1.571$ and $C_D = 0.8215$. The results are presented in the Figures 5.11 and 5.12.

An upper constraint of $D_0 = 0.5190$, $C_{D_0} = 0.814$ for the drag coefficient leads to $C_D = 0.814$ and $C_A = 1.559$.

We recognize that we get the same results for the SQP-method than for the one-shot approach.

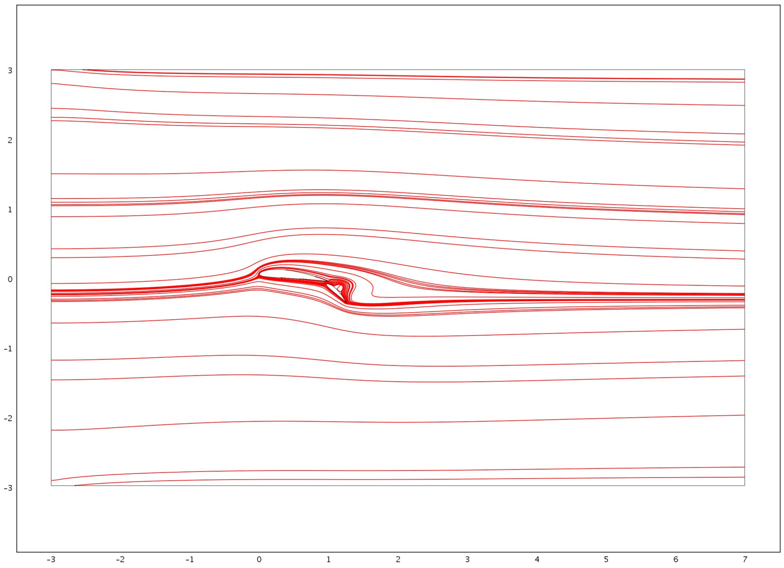

Figure 5.11: Streamlines for the optimal velocity field with an integral state constraint $D_0 = 0.5240$.

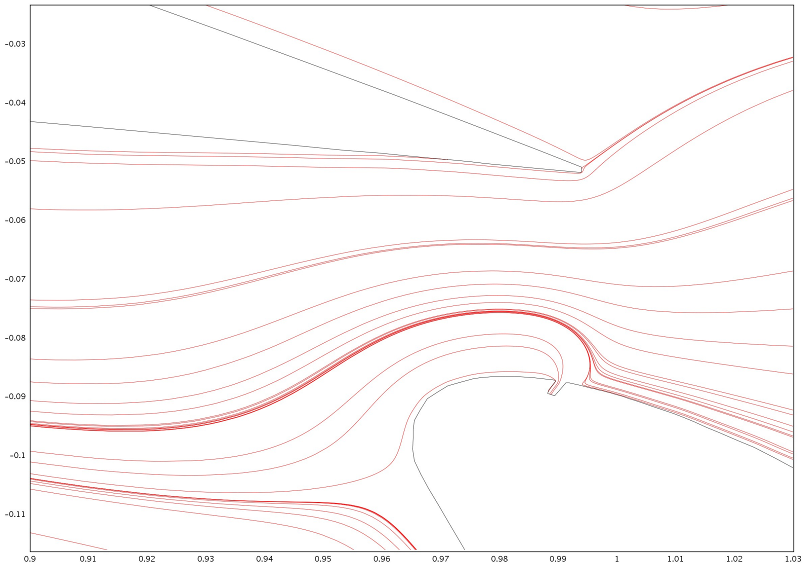

Figure 5.12: Streamlines for the optimal velocity field with an integral state constraint $D_0 = 0.5240$ an a zoom on the slit.

Chapter 6

Convergence of the SQP-method

In this chapter, we want to prove the convergence of the SQP-method, mentioned in Section 5.2. Our approach is mainly based on the theses of A. Unger [104] and D. Wachsmuth [109]. Additionally, we refer to [84].

We want to prove the convergence of the SQP-method for the following problem with penalty term: Minimize

$$
\begin{aligned}
J^I(u,q) :=& \int\limits_{\Gamma_w} \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \right) \cdot \vec{e}_l \, \mathrm{d}s + \frac{\alpha}{2} \|g\|^2_{\mathbf{L}^2(\Gamma_c)} \\
&+ c \left(\int\limits_{\Gamma_w} \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \right) \cdot \vec{e}_d \, \mathrm{d}s - D_0 \right)^2
\end{aligned}
$$

subject to the very weak form of the nonhomogenous Navier-Stokes equations

$$
\begin{aligned}
-\nu\Delta u + (u \cdot \nabla)u + \nabla p &= 0 && \text{in } \Omega \\
\operatorname{div} u &= 0 && \text{in } \Omega \\
u &= \sum_{i=1}^{l} e_i q_i && \text{on } \Gamma_c \\
u &= 0 && \text{on } \Gamma \setminus \Gamma_c
\end{aligned}
$$

and the control constraints

$$
q \in Q_{ad},
$$

where $c \in \mathbb{R}$ is the penalty constant.

The associated optimality system consists analogously to (5.2.2), (5.2.8) and (5.2.9) of the state equations

$$\begin{aligned} -\nu\Delta u + (u\cdot\nabla)u + \nabla p &= 0 && \text{in } \Omega \\ \operatorname{div} u &= 0 && \text{in } \Omega \\ u = g &= \sum_{i=1}^{l} e_i q_i && \text{on } \Gamma_c \\ u &= 0 && \text{on } \Gamma\setminus\Gamma_c \end{aligned}$$

the adjoint system

$$\begin{aligned} -\nu\Delta\lambda + (\nabla\bar u)^T\lambda - (\bar u\cdot\nabla)\lambda + \nabla\pi &= 0 && \text{in } \Omega \\ \operatorname{div}\lambda &= 0 && \text{in } \Omega \\ \lambda &= \vec e_l + K(\bar u,\bar p)\vec e_d && \text{on } \Gamma_w \\ \lambda &= 0 && \text{on } \Gamma\setminus\Gamma_w \end{aligned}$$

and the variational inequality

$$\int\limits_{\Gamma_c} \left(\alpha\bar g - (\nu\frac{\partial\lambda}{\partial\mathfrak{n}} - \pi\mathfrak{n} + (\bar u\cdot\mathfrak{n})\lambda)\right)(g-\bar g)\,\mathrm{d}s \geq 0 \;\forall g\in G_{ad,Q}$$

or equivalent

$$(\alpha M\bar q - D)^T(q-\bar q)\,\mathrm{d}s \geq 0 \;\forall q\in Q_{ad}$$

with

$$M\in\mathbb{R}^{l,l},\; M_{i,j} = \int\limits_{\Gamma_c} e_i e_j\,\mathrm{d}s,\; D\in\mathbb{R}^l,\; D_i = \int\limits_{\Gamma_c}\left(\nu\frac{\partial\lambda}{\partial\mathfrak{n}} - \pi\mathfrak{n} + (\bar u\cdot\mathfrak{n})\lambda\right)e_i\,\mathrm{d}s.$$

Consider with $d := (\alpha M\bar q - D)^T$, see (4.4.1), the sets $\mathcal{A}_+ := \{i : d_i > 0\}$, $\mathcal{A}_- := \{i : d_i < 0\}$, $\mathcal{A} := \mathcal{A}_+\cup\mathcal{A}_-$ and the critical cone associated with $\bar q$

$$C_{\bar q} := \{h\in\mathbb{R}^l :\; h_i = 0\;\forall i\in\mathcal{A} \text{ satisfying (6.0.2)}\}$$

$$h_i = \begin{cases} \geq 0 \text{ if} & \bar q_i = q_{a,i} \\ \leq 0 \text{ if} & \bar q_i = q_{b,i} \end{cases}. \tag{6.0.2}$$

Then, we define the coercivity assumption (SSC''):

$$(SSC'')\begin{cases} \text{The inequality} \\ h^T\mathcal{L}_{qq}(\bar q)h > 0 \\ \text{holds for all } h\in C_{\bar q}\setminus\{0\}. \end{cases}$$

There are several problems in handling the optimality system, especially the nonlinearity of the state equation and the control. The Newton-method is very popular to solve nonlinear systems of the form

$$0 = f(x),$$

because of the fast local convergence. Due to the variational inequality, we are not able to use it in the classical form. A loophole are generalized equations. Therefore, we have first to introduce generalized equations.

6.1 Generalized equations

Let the normal cone given by

$$\mathcal{N}_C(u) := \{z \in \mathbb{R}^n : z^T(v-u) \leq 0 \ \forall v \in C\},$$

and

$$\begin{aligned} \tilde{G}_c &: \mathbf{H}^{3/2}(\Gamma_c) \to \mathbf{H}^2(\Omega), \ g \mapsto u, \\ G_w &: \mathbf{H}^{3/2}(\Gamma_w) \to \mathbf{H}^2(\Omega), \ g \mapsto u \end{aligned}$$

and

$$S : \mathbf{L}^2(\Omega) \to \mathbf{H}^2(\Omega), \ f \mapsto u$$

denote the control-to-solution operators of the Stokes equations, see Corollary 2.12. Let us furthermore define the operators

$$H_c : \mathbb{R}^l \to \mathbf{H}^{3/2}(\Gamma_c), \ q \mapsto \sum_{i=1}^{l} e_i(x) q_i$$

and

$$G_c := \tilde{G}_c \circ H_c : \mathbb{R}^l \to \mathbf{H}^2(\Omega), \ q \mapsto u.$$

Let us recall

$$K(\bar{u}, \bar{p}) = 2c \left(\int\limits_{\Gamma_w} \left(\nu \frac{\partial \bar{u}}{\partial \mathfrak{n}} - \bar{p}\mathfrak{n} \right) \cdot \vec{e}_d \, \mathrm{d}s - D_0 \right)$$

and we denote the derivative by

$$\tilde{K}(u,p) := \frac{\partial}{\partial(u,p)} K(\bar{u}, \bar{p})(u,p) = 2c \left(\int\limits_{\Gamma_w} \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \right) \cdot \vec{e}_d \, \mathrm{d}s \right).$$

Then, we reformulate the optimality system (6.0.1)-(6.0.1) at $\bar{z} = (\bar{u}, \bar{p}, \bar{q}, \bar{\lambda}, \bar{\pi})$ to

$$\begin{aligned}
-\bar{u} + S(-(\bar{u}\cdot\nabla)\bar{u}) + G_c(\bar{q}) &= 0\\
-\bar{\lambda} + S(-(\nabla\bar{u})^T\lambda + (\bar{u}\cdot\nabla)\lambda) + G_w(\vec{e}_l + K(\bar{u},\bar{p})\vec{e}_d) &= 0\\
-(\alpha M\bar{q} - D)^T &\in \mathcal{N}_{Q_{ad}}(\bar{q})
\end{aligned}$$

and the equivalent formulation

$$\begin{aligned}
&0 \in \{-\bar{u} + S(-(\bar{u}\cdot\nabla)\bar{u}) + G_c(\bar{q})\} + \{0\}\\
&0 \in \{-\bar{\lambda} + S(-(\nabla\bar{u})^T\lambda + (\bar{u}\cdot\nabla)\lambda) + G_w(\vec{e}_l + K(\bar{u},\bar{p})\vec{e}_d)\} + \{0\}\\
&0 \in \{(\alpha M\bar{q} - D)^T\} + \mathcal{N}_{Q_{ad}}(\bar{q}).
\end{aligned}$$

We write for short

$$Z := \mathbf{H}^2(\Omega) \times \mathbb{R}^l \times \mathbf{H}^2(\Omega) \text{ and } \tilde{W} := \mathbf{H}^2(\Omega) \times \mathbf{H}^2(\Omega) \times \mathbb{R}^l$$

and define $F : Z \to \tilde{W}$ and $T : Z \to 2^{\tilde{W}}$ as follows

$$F(z) = F(u,q,\lambda) = \begin{pmatrix} -u + S(-(u\cdot\nabla)u) + G_c(q) \\ -\lambda + S(-(\nabla\bar{u})^T\lambda + (\bar{u}\cdot\nabla)\lambda) + G_w(\vec{e}_l + K(u,p)\vec{e}_d) \\ (\alpha M q - D)^T \end{pmatrix} \tag{6.1.1}$$

for the differentiable part and

$$T(z) = T(u,q,\lambda) = \begin{pmatrix} \{0\} \\ \{0\} \\ \mathcal{N}_{Q_{ad}}(q) \end{pmatrix} \tag{6.1.2}$$

for the set-valued part in the generalized equation. Now, we can formulate the optimality system as

$$0 \in F(\bar{z}) + T(\bar{z}). \tag{6.1.3}$$

The equation (6.1.3) is a generalized equation. For more details see Alt [5, 8, 9] Dontchev [33, 35, 36, 37], Goldberg and Tröltzsch [48] and Josephy [60].

The classical Newton-method is not applicable to this kind of system. The generalized Newton-method is mentioned in the works above and reads as

Algorithm 6.1 **(GNM).**

1. *Choose an initial value z^0 and set $k = 0$*

2. *Determine z by the linearized generalized equation*

$$0 \in F(z^k) + F'(z^k)(z - z^k) + T(z) \tag{6.1.4}$$

3. *Set $k = k + 1$ and $z^k = z$. Goto 2.*

Similar to the theory of the classical Newton-method, we need further assumptions to show the convergence of this generalized Newton-method, which is closely related to the notion of strong regularity of (6.1.3) and based on Robinson [82].

Definition 6.2 (Strong regularity)**.** *Let $\tilde{z} = (\tilde{u}, \tilde{q}, \tilde{\lambda}) \in Z$. The generalized equation* (6.1.3) *is said to be strongly regular at $\tilde{z}$, if there exist open balls $\mathcal{B}_{r_1}(0)$ in $\tilde{W}$ and $\mathcal{B}_{r_2}(\tilde{z})$ in Z, with positive constants $r_1 > 0$, $r_2 > 0$ and $c_L > 0$ such that for all perturbations $e \in \mathcal{B}_{r_1}(0)$ the linearized equation*

$$e \in F(\tilde{z}) + F'(\tilde{z})(z - \tilde{z}) + T(z)$$

admits a unique solution $z = z(e)$ in $\mathcal{B}_{r_2}(\tilde{z})$ and the Lipschitz property

$$\|z(e_1) - z(e_2)\|_Z \leq c_L \|e_1 - e_2\|_{\tilde{W}}$$

holds for all $e_1, e_2 \in \mathcal{B}_{r_1}(0)$.

In the original paper of Robinson [83], it was assumed that T has a closed graph. Dontchev has shown in [35] that this assumption is not needed. The next theorem provides the possibility to transfer stability results for the perturbed linearized equation to the perturbed nonlinear equation.

Theorem 6.3. *Let $\bar{z} \in Z$ be a solution of the generalized equation* (6.1.3) *such that this equation is strongly regular at $\bar{z}$. Then there exist open balls $\mathcal{B}_{r_1}(0)$ and $\mathcal{B}_{r_2}(\bar{z})$ such that, for all $e \in \mathcal{B}_{r_1}(0)$, the perturbed equation*

$$e \in F(\bar{z}) + T(\bar{z})$$

has a unique solution $z = z(e)$ in $\mathcal{B}_{r_2}(\bar{z})$ and the solution mapping $e \mapsto z(e)$ is Lipschitz-continuous from $\mathcal{B}_{r_1}(0)$ to $\mathcal{B}_{r_2}(\bar{z})$.

Based on the strong regularity, the next theorem states the convergence of the SQP-method.

Theorem 6.4. *Let $\bar{z} = (\bar{u}, \bar{p}, \bar{q}, \bar{\lambda}, \bar{\pi})$ be a solution of* (6.1.3) *and additionally let this generalized equation be strongly regular at $\bar{z}$. Then there exists an open ball $\mathcal{B}_{r_2}(\bar{z})$ such that for every starting point z_1 in $\mathcal{B}_{r_2}(\bar{z})$ the generalized Newton method generates a unique sequence $\{z_k\}_{k=1}^{\infty}$, where z_k stays in $\mathcal{B}_{r_2}(\bar{z})$ and we obtain*

$$\|z^{k+1} - \bar{z}\|_Z \leq c_G \|z^k - \bar{z}\|_Z^2.$$

with c_G independent of k.

For a proof, we refer to [5, 36].

This theorem means that the linearized generalized equation is uniquely solvable under the suitable assumptions. The sequence of solutions, generated by the generalized Newton-method, converges quadratically to the solution of the generalized equation (6.1.3). In the next part we investigate the solvability of the linearized generalized equation and the relation to the SQP-method.
Let us consider the linearized equation

$$0 \in F(z^i) + F'(z^i)(z - z^i) + T(z).$$

This means that we have due to the linearity of K

$$\begin{aligned} 0 &\in -u + S((u^i \cdot \nabla)u^i - (u^i \cdot \nabla)u - (u \cdot \nabla)u^i) + G_c(q) + \{0\} \\ 0 &\in -\lambda + S(-(\nabla u)^T \lambda^i + (u \cdot \nabla)\lambda^i - (\nabla u^i)^T \lambda + (u^i \cdot \nabla)\lambda \\ &\quad + (\nabla u^i)^T \lambda^i - (u^i \cdot \nabla)\lambda^i) + G_w(e_l + \tilde{K}(u,p)e_d) + \{0\} \\ 0 &\in (\alpha M q - D(z^i) - D'(z^i)(z - z^i))^T + \mathcal{N}_{Q_{ad}}(q). \end{aligned} \tag{6.1.5}$$

The first relation of (6.1.5) is equivalent to the boundary value problem

$$\begin{aligned} -\nu\Delta u + (u^i \cdot \nabla)u + (u \cdot \nabla)u^i + \nabla p = (u^i \cdot \nabla)u^i \quad & \text{in } \Omega \\ \operatorname{div} u = 0 \quad & \text{in } \Omega \\ u = \sum_{i=1}^{l} e_i q_i \quad & \text{on } \Gamma_c, \\ u = 0 \quad & \text{on } \Gamma \setminus \Gamma_c \end{aligned} \tag{6.1.6}$$

and the second to the problem

$$\begin{aligned} -\nu\Delta\lambda + \nabla\pi &= -(\nabla u)^T \lambda^i - (\nabla u^i)^T (\lambda - \lambda^i) \\ &\quad + (u^i \cdot \nabla)(\lambda - \lambda^i) + (u \cdot \nabla)\lambda^i && \text{in } \Omega \\ \operatorname{div} \lambda &= 0 && \text{in } \Omega \\ \lambda &= \vec{e}_l + \tilde{K}(u,p)\vec{e}_d && \text{on } \Gamma_w \\ \lambda &= 0 && \text{on } \Gamma \setminus \Gamma_{wall}. \end{aligned} \tag{6.1.7}$$

Considering $\mathcal{N}_{Q_{ad}}(q)$, we additionally get

$$(\alpha M q - D(z^i) - D'(z^i)(z - z^i))^T (\tilde{q} - q) \geq 0 \quad \forall \tilde{q} \in Q_{ad}. \tag{6.1.8}$$

We see that the linearized generalized equation (6.1.4) is equivalent to (6.1.6)-(6.1.8) which is similar to the first-order necessary optimality system, see

(6.1.9). In reality, this is the necessary optimality system for the linear-quadratic problem of the SQP-method, mentioned above. The cost functional of the linear-quadratic problem was defined by

$$J^k(u,q) = \nabla J(u^k,q^k)(u-u^k,q-q^k) + \frac{1}{2}\mathcal{L}_{zz}(u^k,q^k,\lambda^k)[(u-u^k,q-q^k)]^2.$$

After expanding, it is with $g = H_c(q) = \sum_{i=1}^{l} e_i q_i$ equal to

$$\begin{aligned}
J^k(u,q) =& \int\limits_{\Gamma_w} \left(\nu \frac{\partial(u-u^k)}{\partial\mathfrak{n}} - (p-p^k)\mathfrak{n}\right)\cdot \vec{e}_l \,\mathrm{d}s \\
&+ K(u^k,p^k)\left(\int\limits_{\Gamma_w}\left(\nu\frac{\partial(u-u^k)}{\partial\mathfrak{n}} - (p-p^k)\mathfrak{n}\right)\cdot\vec{e}_d\,\mathrm{d}s\right) \\
&+ \alpha\int\limits_{\Gamma_c} g^k(g-g^k)\,\mathrm{d}s + b(u-u^k,u-u^k,\lambda^k) \\
&+ \tilde{K}(u-u^k,p-p^k)\left(\int\limits_{\Gamma_w}\left(\nu\frac{\partial(u-u^k)}{\partial\mathfrak{n}} - (p-p^k)\mathfrak{n}\right)\cdot\vec{e}_d\,\mathrm{d}s\right) \\
=& \int\limits_{\Gamma_w} \left(\nu \frac{\partial(u-u^k)}{\partial\mathfrak{n}} - (p-p^k)\mathfrak{n}\right)\cdot \vec{e}_l \,\mathrm{d}s \\
&+ K(u,p)\left(\int\limits_{\Gamma_w}\left(\nu\frac{\partial(u-u^k)}{\partial\mathfrak{n}} - (p-p^k)\mathfrak{n}\right)\cdot\vec{e}_d\,\mathrm{d}s\right) \\
&+ \alpha\int\limits_{\Gamma_c} g^k(g-g^k)\,\mathrm{d}s + b(u-u^k,u-u^k,\lambda^k).
\end{aligned}$$

Now, we have shown the connection between the generalized Newton-method and the SQP-method. The SQP-method is a kind of a generalized Newton-method to solve the optimality system

$$\begin{aligned}
-\nu\Delta u + (u\cdot\nabla)u + \nabla p &= f && \text{in } \Omega \\
\operatorname{div} u &= 0 && \text{in } \Omega \\
u &= \sum_{i=1}^{l} e_i q_i && \text{on } \Gamma_c, \\
u &= 0 && \text{on } \Gamma\setminus\Gamma_c.
\end{aligned} \tag{6.1.9a}$$

$$\begin{aligned} -\nu\Delta\lambda + (\nabla\bar{u})^T\lambda - (\bar{u}\cdot\nabla)\lambda + \nabla\pi &= 0 && \text{in } \Omega \\ \operatorname{div}\lambda &= 0 && \text{in } \Omega \\ \lambda &= \vec{e}_l + K(\bar{u},\bar{p})\vec{e}_d && \text{on } \Gamma_w \\ \lambda &= 0 && \text{on } \Gamma\setminus\Gamma_w, \end{aligned} \tag{6.1.9b}$$

$$(\alpha M\bar{q} - D)^T(q-\bar{q}) \geq 0 \quad \forall q \in Q_{ad} \tag{6.1.9c}$$

obtained analogously to (4.3.4a)-(4.3.4c).

The optimality system of the linear-quadratic system is equivalent to the linearized generalized system (6.1.4). For the finite-dimensional case, one can find an example in [42] and for the case of infinite-dimensional optimization, see [6] and [33]. In the next sections, we want to prove the convergence of the SQP-method based on Theorem 6.4. That means that we have to show the assumptions in this theorem.

6.2 Perturbed optimization problem

There is a lot of literature about convergence of the SQP-method. Alt [6, 7], Alt and Malanowski [10, 11], Malanowski [68, 67] and Dontchev et. al. [34] proved convergence for optimal control problems subject to ODEs. Optimal control problems related to PDEs were investigated in Hinze and Hintermüller [53], Goldberg and Tröltzsch [48], Kupfer and Sachs [64], Tröltzsch [100, 101], and Volkwein [107].

In the last section, we have shown that the SQP-method for the optimal boundary control problem is a variation of the generalized Newton-method for solving the first-order necessary optimality conditions. In this section of this thesis, we follow the work of Unger [104] and Wachsmuth [109].

In the next step, we want to verify continuous differentiability of F, Lipschitz continuity of F', and strong regularity of the optimality system to use Theorem 6.4 and to show convergence of the SQP-method.

Theorem 6.5. *The function F defined by* (6.1.1) *is continuous differentiable and the derivative $B'(u)$ is Lipschitz continuous.*

The proof is analogous to [109], Corollary 5.2.

Now, let us come to the more complex assumption of strong regularity. Therefore, we want to consider the pertubed linearized at $\bar{z}$ generalized equation

$$e \in F(\bar{z}) + F'(\bar{z})(z-\bar{z}) + T(z) \tag{6.2.1}$$

as an optimality system of another optimal control problem. Let
$\bar{z} = (\bar{u}, \bar{p}, \bar{q}, \bar{\lambda}, \bar{\pi})$ fulfill the the first-order necessary optimality conditions of

(P_I) and because of this be a solution of the generalized equation (6.1.3). The perturbed equation (6.2.1) with $e = (\tilde{e}_u, \tilde{e}_\lambda, \tilde{e}_q) \in \hat{W}$ at $\bar{z}$ is equivalent to

$$\begin{aligned}
0 &\in -\tilde{e}_u - u + S(-(\bar{u}\cdot\nabla)u - (u\cdot\nabla)\bar{u} + (\bar{u}\cdot\nabla)\bar{u}) + G_c(q) + \{0\}\\
0 &\in -\tilde{e}_\lambda - \lambda + S(-(\nabla u)^T\bar{\lambda} + (u\cdot\nabla)\bar{\lambda} - (\nabla\bar{u})^T\lambda + (\bar{u}\cdot\nabla)\lambda\\
&\qquad\qquad + (\nabla\bar{u})^T\bar{\lambda} - (\bar{u}\cdot\nabla)\bar{\lambda}) + G_w(e_l + \tilde{K}(u,p)e_d) + \{0\}\\
0 &\in (-\tilde{e}_q + \alpha Mq - D(\bar{z}) - D'(\bar{z})(z-\bar{z}))^T + \mathcal{N}_{Q_{ad}}(q),
\end{aligned}$$

with $\hat{u} = u + \tilde{e}_u$, $\hat{p} = p$, $\hat{q} = q$, $\hat{\lambda} = \lambda + \tilde{e}_\lambda$ and $\hat{\pi} = \pi$ equivalent to

$$\begin{aligned}
0 &\in -\hat{u} + S(-(\bar{u}\cdot\nabla)(\hat{u}-\tilde{e}_u) - ((\hat{u}-\tilde{e}_u)\cdot\nabla)\bar{u} + (\bar{u}\cdot\nabla)\bar{u}) + G_c(\hat{q}) + \{0\}\\
0 &\in -\hat{\lambda} + S(-(\nabla(\hat{u}-\tilde{e}_u))^T\bar{\lambda} + ((\hat{u}-\tilde{e}_u)\cdot\nabla)\bar{\lambda} + (\nabla\bar{u})^T\bar{\lambda} - (\bar{u}\cdot\nabla)\bar{\lambda}\\
&\qquad\qquad - (\nabla\bar{u})^T(\hat{\lambda}-\tilde{e}_\lambda) + (\bar{u}\cdot\nabla)(\hat{\lambda}-\tilde{e}_\lambda))\\
&\qquad\qquad + G_w(\vec{e}_l + \tilde{K}(\hat{u}-\tilde{e}_u,p)\vec{e}_d) + \{0\}\\
0 &\in (-\tilde{e}_q + \alpha M\hat{q} - D(\bar{z}) - D'(\bar{z})((\hat{u}-\tilde{e}_u,\hat{p},\hat{q},\hat{\lambda}-\tilde{e}_\lambda,\hat{\pi}) - \bar{z}))^T + \mathcal{N}_{Q_{ad}}(q).
\end{aligned}$$

The last system is with

$$\begin{aligned}
e_u &= (\bar{u}\cdot\nabla)\tilde{e}_u + (\tilde{e}_u\cdot\nabla)\bar{u},\\
e_\lambda^1 &= (\nabla\tilde{e}_u)^T\bar{\lambda} - (\tilde{e}_u\cdot\nabla)\bar{\lambda} + (\nabla\bar{u})^T\tilde{e}_\lambda - (\bar{u}\cdot\nabla)\tilde{e}_\lambda,\\
e_\lambda^2 &= -K(\tilde{e}_u,0),\\
Ne_q &= -\tilde{e}_q - \tilde{D}
\end{aligned} \tag{6.2.2}$$

and

$$\begin{aligned}
\tilde{D}_i &= \int_{\Gamma_c} (\nu\frac{\partial\tilde{e}_\lambda}{\partial\mathfrak{n}} + (\bar{u}\cdot\mathfrak{n})\tilde{e}_\lambda + (\tilde{e}_u\cdot\mathfrak{n})\bar{\lambda} + (\bar{u}\cdot\mathfrak{n})\bar{\lambda})\; e_i \,\mathrm{d}s,\\
\hat{D}_i &= \int_{\Gamma_c} (\nu\frac{\partial\hat{\lambda}}{\partial\mathfrak{n}} - \hat{\pi}\mathfrak{n} + (\bar{u}\cdot\mathfrak{n})\hat{\lambda} + (\hat{u}\cdot\mathfrak{n})\bar{\lambda})\; e_i \,\mathrm{d}s\\
N \in \mathbb{N}^l,\ N_i &= \int_{\Gamma_c} e_i \,\mathrm{d}s
\end{aligned}$$

for $i = 1,\dots,l$ equivalent to

$$\begin{aligned}
0 &\in -\hat{u} + S(-(\bar{u}\cdot\nabla)\hat{u} - (\hat{u}\cdot\nabla)\bar{u} + (\bar{u}\cdot\nabla)\bar{u}) + G_c(\hat{q}) + S(e_u) + \{0\}\\
0 &\in -\hat{\lambda} + S(e_\lambda^1) + G_w(e_\lambda^2) + G_w(\vec{e}_l + \tilde{K}(\hat{u},p)\vec{e}_d) + \{0\}\\
&\quad + S(-(\nabla\hat{u})^T\bar{\lambda} + (\hat{u}\cdot\nabla)\bar{\lambda} - (\nabla\bar{u})^T\hat{\lambda} + (\bar{u}\cdot\nabla)\hat{\lambda} + (\nabla\bar{u})^T\bar{\lambda} - (\bar{u}\cdot\nabla)\bar{\lambda})\\
0 &\in (Ne_q + \alpha M\hat{q} - \hat{D})^T + \mathcal{N}_{Q_{ad}}(q).
\end{aligned} \tag{6.2.3}$$

The system (6.2.3) corresponds to the following system of the perturbed state equation

$$
\begin{aligned}
-\nu\Delta\hat{u} + (\bar{u}\cdot\nabla)\hat{u} + (\hat{u}\cdot\nabla)\bar{u} + \nabla p &= (\bar{u}\cdot\nabla)\bar{u} + e_u && \text{in } \Omega\\
\operatorname{div}\hat{u} &= 0 && \text{in } \Omega\\
\hat{u} &= \sum_{i=1}^{l} e_i\hat{q}_i && \text{on } \Gamma_c,\\
\hat{u} &= 0 && \text{on } \Gamma\setminus\Gamma_c,
\end{aligned}
\tag{6.2.4}
$$

the perturbed adjoint equation

$$
\begin{aligned}
-\nu\Delta\hat{\lambda} + (\nabla\bar{u})^T\hat{\lambda} - (\bar{u}\cdot\nabla)\hat{\lambda} + \nabla\pi &= -(\nabla\hat{u})^T\bar{\lambda} + (\hat{u}\cdot\nabla)\bar{\lambda} + e_\lambda^1\\
&\quad + (\nabla\bar{u})^T\bar{\lambda} - (\bar{u}\cdot\nabla)\bar{\lambda} && \text{in } \Omega\\
\operatorname{div}\lambda &= 0 && \text{in } \Omega\\
\lambda &= \vec{e}_l + \tilde{K}(\hat{u},p)\vec{e}_d + e_\lambda^2 && \text{on } \Gamma_w\\
\lambda &= 0 && \text{on } \Gamma\setminus\Gamma_w
\end{aligned}
\tag{6.2.5}
$$

and the associated variational inequality

$$
(\alpha M\hat{q} - \hat{D} + Ne_q)^T(q-\bar{q}) \geq 0 \quad \forall q \in Q_{ad}. \tag{6.2.6}
$$

Now, we see that this is the optimality problem of the following pertubated linear-quadratic optimization problem (P_e), where we write $(u,p,q,\lambda,\pi) = (\hat{u},\hat{p},\hat{q},\hat{\lambda},\hat{\pi})$ for simplicity:

$$
\begin{aligned}
\min J_e(u,q) :=& \int\limits_{\Gamma_w}\left(\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n}\right)\cdot\vec{e}_l\,\mathrm{d}s + \frac{c}{2}\left(\int\limits_{\Gamma_w}\left(\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n}\right)\cdot\vec{e}_d\,\mathrm{d}s - D_0\right)^2\\
&+ \int\limits_{\Gamma_c}|H_c(q)|^2\,\mathrm{d}s + \int\limits_{\Omega} e_u u + e_\lambda^1\lambda\,\mathrm{d}x + \int\limits_{\Gamma_c} e_q H_c(q)\,\mathrm{d}s\\
&+ \int\limits_{\Gamma_w} e_\lambda^2\lambda\,\mathrm{d}s + b(u-\bar{u},u-\bar{u},\bar{\lambda})
\end{aligned}
$$

subject to the pertubated Navier-Stokes equations

$$
\begin{aligned}
-\nu\Delta u + (\bar{u}\cdot\nabla)u + (u\cdot\nabla)\bar{u} + \nabla p &= (\bar{u}\cdot\nabla)\bar{u} + e_u && \text{in } \Omega\\
\operatorname{div} u &= 0 && \text{in } \Omega\\
u &= \sum_{i=1}^{l} e_i q_i && \text{on } \Gamma_c,\\
u &= 0 && \text{on } \Gamma\setminus\Gamma_c,
\end{aligned}
\tag{6.2.7}
$$

and the control constraints

$$q \in Q_{ad}.$$

Let us now consider $e := (e_u, e_\lambda^1, e_\lambda^2, e_q) \in W := \mathbf{H}^2(\Omega) \times \mathbf{H}^2(\Omega) \times \mathbb{R} \times \mathbb{R}^l$. Then, we obtain due to (6.2.2)

$$\|e\|_W \leq C\|\tilde{e}\|_{\tilde{W}}$$

with

$$\begin{aligned} \|e\|_W &= \|e_u\|_{\mathbf{H}^2(\Omega)} + \|e_\lambda^1\|_{\mathbf{H}^2(\Omega)} + |e_\lambda^2| + |e_q|, \\ \|e\|_W &= \|\tilde{e}_u\|_{\mathbf{H}^2(\Omega)} + \|\tilde{e}_\lambda\|_{\mathbf{H}^2(\Omega)} + |\tilde{e}_q|. \end{aligned}$$

6.3 A modified problem

Because there is no convexity of the cost functional J_e in directions not included in (SSC), we need

$$\tilde{Q}_{ad} := \{\tilde{q} \in Q_{ad} : \tilde{q} = \bar{q} \text{ on } \mathcal{A}\}$$

to not allow changes of the control function q on the active sets. Let us denote the new problem by $(\tilde{P}_e)$ consisting of the cost functional J_e, the state equation (6.2.7) and the control constraint $q \in \tilde{Q}_{ad}$. Finding a solution for the problem $(\tilde{P}_e)$ is equivalent to finding a solution for the linearized and perturbed generalized equation (6.2.1), with

$$T(z) = (\{0\}, \{0\}, \mathcal{N}_{\tilde{Q}_{ad}}(q))^T.$$

First, we investigate the existence of a solution for the perturbed linear-quadratic optimal control problem $(\tilde{P}_e)$. Then, we want to prove that the optimal control of $(\tilde{P}_e)$ is Lipschitz-continuous with respect to the perturbation e. The idea is to investigate strong regularity of the generalized equation

$$e \in F(z) + (0, 0, \mathcal{N}_{\tilde{Q}_{ad}}(q)) \tag{6.3.1}$$

and then transfer the result to the original problem.

6.3.1 Existence of a solution

The next theorem proves the solvability of $(\tilde{P}_e)$.

Theorem 6.6. *Assume that $\bar{z} = (\bar{u}, \bar{p}, \bar{q}, \bar{\lambda}, \bar{\pi})$ satisfy the optimality system and the coercivity condition (SSC'') at the begin of this chapter. Then $(\tilde{P}_e)$ admits a unique optimal control $\bar{u}_e$.*

Proof. Let $c \in \mathbb{R}$ be a generic constant. Denoting the Lagrangian belonging to $(\tilde{P}_e)$ by $\mathcal{L}^e$, we derive

$$\mathcal{L}^e_{qq}(q) = \mathcal{L}_{qq}(\bar{q}) \tag{6.3.2}$$

for all q, because the perturbation appears only linear.

Taking $q_1, q_2 \in \tilde{Q}_{ad}$ with associating u_1, u_2 of (6.2.7), the pair (u_1-u_2, q_1-q_2) fits to the assumption of (SSC'') and we obtain

$$(q_1 - q_2)\mathcal{L}^e_{qq}(q)(q_1 - q_2) = (q_1 - q_2)\mathcal{L}_{qq}(\bar{q})(q_1 - q_2) > 0.$$

for $q_1 \neq q_2$. Because of this, the problem $(\tilde{P}_e)$ is convex on $\tilde{Q}_{ad}$. Thus, $(\tilde{P}_e)$ is uniquely solvable as a linear-quadratic optimization problem with strongly convex cost functional with modifications described in Section 3.3. We denote the unique solution by $\bar{u}_e$. □

For more details, see [104].

6.3.2 Lipschitz stability

We still need to prove the Lipschitz-continuity of the perturbation to solution mapping $e \mapsto (u_e, g_e, \lambda_e)$ of $(\tilde{P}_e)$ to show strong regularity.

Theorem 6.7. *Let $(\bar{z})$ fulfill the coercivity condition (SSC''). Then the solution mapping $e := (e_u, e^1_\lambda, e^2_\lambda, e_q) \mapsto z_e = (u_e, p_e, q_e, \lambda_e, \pi_e)$ is Lipschitz-continuous from W to Z.*

Proof. Let z_1, z_2 be two elements of Z and q_i, $i = 1, 2$ be the optimal control functions of the optimization problem $(\tilde{P}_e)$ with the associated states u_i and adjoints λ_i. Let us furthermore define the differences $z := z_1 - z_2$, $q := q_1 - q_2$, $u := u_1 - u_2$ and $\lambda := \lambda_1 - \lambda_2$.

So, the variational inequality with the constraint $q_i \in \tilde{Q}_{ad}$ looks as

$$(\alpha M q_i - \hat{D}(u_i, \lambda_i, \pi_i) - N e_{q,i})^T (q - q_i) \geq 0, \ \forall q \in \tilde{Q}_{ad}. \tag{6.3.3}$$

Testing this inequality with q_2 for $i = 1$ and q_1 for $i = 2$ and adding both, we derive

$$\hat{D}(u, \lambda, \pi)^T q + (N e_q)^T q \geq (\alpha M q)^T q. \tag{6.3.4}$$

The difference u is the weak solution of the state equation

$$\begin{aligned}
-\nu\Delta u + (\bar{u}\cdot\nabla)u + (u\cdot\nabla)\bar{u} + \nabla p &= e_u && \text{in } \Omega \\
\operatorname{div} u &= 0 && \text{in } \Omega \\
u = g &:= \sum_{j=1}^{l} e_j q_j && \text{on } \Gamma_c, \\
u &= 0 && \text{on } \Gamma \setminus \Gamma_c,
\end{aligned} \tag{6.3.5}$$

Testing (6.3.5) with $(\lambda, \pi) = (\lambda_1 - \lambda_2, \pi_1 - \pi_2)$, we obtain

$$\begin{aligned} a(u,\lambda) - (\nu\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n}, \lambda)_{\mathbf{L}^2(\Gamma)} + (p, \operatorname{div}\lambda)_{\mathbf{L}^2(\Omega)} & \\ +\langle(\bar{u}\cdot\nabla)u + (u\cdot\nabla)\bar{u}, \lambda\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)} &= (e_u, \lambda)_{\mathbf{L}^2(\Omega)} \\ (\operatorname{div} u, \pi)_{\mathbf{L}^2(\Omega)} &= 0 \\ \tau_{\Gamma_c} u &= g \end{aligned} \qquad (6.3.6)$$

and testing the adjoint

$$\begin{aligned} -\nu\Delta\lambda + (\nabla\bar{u})^T\lambda - (\bar{u}\cdot\nabla)\lambda + \nabla\pi &= -(\nabla u)^T\bar{\lambda} + (u\cdot\nabla)\bar{\lambda} + e_\lambda^1 && \text{in } \Omega \\ \operatorname{div}\lambda &= 0 && \text{in } \Omega \\ \lambda &= \vec{e}_l + \tilde{K}(u,p)\vec{e}_d + e_\lambda^2 && \text{on } \Gamma_w \\ \lambda &= 0 && \text{on } \Gamma\setminus\Gamma_w \end{aligned} \qquad (6.3.7)$$

with $(u, p) = (u_1 - u_2, p_1 - p_2)$, we get

$$\begin{aligned} & a(\lambda,u) - (\nu\frac{\partial \lambda}{\partial \mathfrak{n}} - \pi\mathfrak{n} + (\bar{u}\cdot\mathfrak{n})\lambda + (u\cdot\mathfrak{n})\bar{\lambda}, u)_{\mathbf{L}^2(\Gamma)} \\ & + \langle(\nabla\bar{u})^T\lambda - (\bar{u}\cdot\nabla)\lambda, u\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)} \\ & + \langle(\nabla u)^T\bar{\lambda} - (u\cdot\nabla)\bar{\lambda}, u\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)} + (\pi, \operatorname{div} u)_{\mathbf{L}^2(\Omega)} = (e_\lambda^1, u)_{\mathbf{L}^2(\Omega)} \end{aligned} \qquad (6.3.8a)$$

$$(\operatorname{div}\lambda, p)_{\mathbf{L}^2(\Omega)} = 0 \qquad (6.3.8b)$$

$$\tau_{\Gamma_w}\lambda = \vec{e}_l + \tilde{K}(u,p)\vec{e}_d + e_\lambda^2. \qquad (6.3.8c)$$

The systems (6.3.6) and (6.3.8) are equivalent to

$$\begin{aligned} a(u,\lambda) - \nu(\frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n}, \vec{e}_l + \tilde{K}(u,p)\vec{e}_d + e_\lambda^2)_{\mathbf{L}^2(\Gamma_w)} & \\ + \langle(\bar{u}\cdot\nabla)u + (u\cdot\nabla)\bar{u}, \lambda\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)} &= (e_u, \lambda)_{\mathbf{L}^2(\Omega)} \end{aligned}$$

and

$$\begin{aligned} & a(\lambda,u) - \nu(\frac{\partial \lambda}{\partial \mathfrak{n}} - \pi\mathfrak{n} + (\bar{u}\cdot\mathfrak{n})\lambda + (u\cdot\mathfrak{n})\bar{\lambda}, g)_{\mathbf{L}^2(\Gamma_c)} \\ & + \langle(\bar{u}\cdot\nabla)u + (u\cdot\nabla)\bar{u}, \lambda\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)} \\ & + \langle(\nabla u)^T\bar{\lambda} - (u\cdot\nabla)\bar{\lambda}, u\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)} = (e_\lambda^1, u)_{\mathbf{L}^2(\Omega)}. \end{aligned}$$

Considering them with

$$\langle(\nabla u)^T\bar{\lambda} - (u\cdot\nabla)\bar{\lambda}, u\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)} = \langle(u\cdot\nabla)u, \bar{\lambda}\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)}$$

we obtain the equality

$$\begin{aligned}(e_u,\lambda)_{\mathbf{L}^2(\Omega)}+&(\nu\frac{\partial u}{\partial\mathfrak{n}}-p\mathfrak{n},\vec{e}_l+\tilde{K}(u,p)\vec{e}_d+e_\lambda^2)_{\mathbf{L}^2(\Gamma_w)}\\ &=(e_\lambda^1,u)_{\mathbf{L}^2(\Omega)}-\langle(u\cdot\nabla)u,\bar{\lambda}\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)}\\ &\quad+(\nu\frac{\partial\lambda}{\partial\mathfrak{n}}-\pi\mathfrak{n}+(\bar{u}\cdot\mathfrak{n})\lambda+(u\cdot\mathfrak{n})\bar{\lambda},g)_{\mathbf{L}^2(\Gamma_c)}\end{aligned}$$

or equivalent

$$\begin{aligned}\langle(u\cdot\nabla)u,\bar{\lambda}\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)}&-\tilde{K}(u,p)(\nu\frac{\partial u}{\partial\mathfrak{n}}-p\mathfrak{n},\vec{e}_d)_{\mathbf{L}^2(\Gamma_w)}\\ &=(e_\lambda^1,u)_{\mathbf{L}^2(\Omega)}-(e_u,\lambda)_{\mathbf{L}^2(\Omega)}-(\nu\frac{\partial u}{\partial\mathfrak{n}}-p\mathfrak{n},e_\lambda^2)_{\mathbf{L}^2(\Gamma_w)}\\ &\quad+(\nu\frac{\partial\lambda}{\partial\mathfrak{n}}-\pi\mathfrak{n}+(\bar{u}\cdot\mathfrak{n})\lambda+(u\cdot\mathfrak{n})\bar{\lambda},g)_{\mathbf{L}^2(\Gamma_c)}\\ &\quad-(\nu\frac{\partial u}{\partial\mathfrak{n}}-p\mathfrak{n},\vec{e}_l)_{\mathbf{L}^2(\Gamma_w)}.\end{aligned}\tag{6.3.9}$$

Let us additionally consider $\tilde{u}$ as a weak solution of (6.3.5) with $q=0$. Thus, $(u-\tilde{u},q)$ fits to the assumption of the coercivity condition (SSC''), due to $q_i=(q_e)_{1,i}-(q_e)_{2,i}=(\bar{q}_e)_i-(\bar{q}_e)_i=0$ on $\mathcal{A}$ by $(q_e)_1,(q_e)_2\in\tilde{Q}_{ad}$ and $(u-\tilde{u})$ is the solution of the associated at $\bar{q}$ linearized state equation.

Furthermore, we obtain

$$\begin{aligned}0&<\mathcal{L}_{vv}(\bar{z})[u-\tilde{u},q]^2\\ &=\underbrace{\mathcal{L}_{qq}(\bar{z})[q]^2}_{*_1}+\underbrace{\mathcal{L}_{uu}(\bar{z})[u]^2}_{*_2}\underbrace{-2\mathcal{L}_{uu}(\bar{z})[u,\tilde{u}]^2}_{*_3}+\underbrace{\mathcal{L}_{uu}(\bar{z})[\tilde{u}]^2}_{*_4}.\end{aligned}\tag{6.3.10}$$

Adding the terms $*_1$ and $*_2$ leads to

$$\begin{aligned}\mathcal{L}_{qq}(\bar{z})[q]^2+\mathcal{L}_{uu}(\bar{z})[u]^2&=\alpha\|g\|^2_{\mathbf{L}^2(\Gamma_c)}-\langle(u\cdot\nabla)u,\bar{\lambda}\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)}\\ &\quad+\tilde{K}(u,p)(\nu\frac{\partial u}{\partial\mathfrak{n}}-p\mathfrak{n},\vec{e}_d)_{L^2(\Gamma_w)}\end{aligned}$$

and with (6.3.9) and $\|g\|_{\mathbf{L}^2(\Gamma_c)}=(Mq)^Tq$ to

$$\begin{aligned}\mathcal{L}_{qq}(\bar{z})[q]^2+\mathcal{L}_{uu}(\bar{z})[u]^2&=\alpha(Mq)^Tq+(e_u,\lambda)_{\mathbf{L}^2(\Omega)}-(e_\lambda^1,u)_{\mathbf{L}^2(\Omega)}\\ &\quad+(\nu\frac{\partial u}{\partial\mathfrak{n}}-p\mathfrak{n},e_\lambda^2+\vec{e}_l)_{\mathbf{L}^2(\Gamma_w)}\\ &\quad-(\nu\frac{\partial\lambda}{\partial\mathfrak{n}}-\pi\mathfrak{n}+(\bar{u}\cdot\mathfrak{n})\lambda+(u\cdot\mathfrak{n})\bar{\lambda},g)_{\mathbf{L}^2(\Gamma_c)}.\end{aligned}$$

The inequality (6.3.4) leads to

$$\begin{aligned}\mathcal{L}_{qq}(\bar{z})[q]^2+\mathcal{L}_{uu}(\bar{z})[u]^2 &\leq (e_u,\lambda)_{\mathbf{L}^2(\Omega)}-(e^1_\lambda,u)_{\mathbf{L}^2(\Omega)}+(Ne_g)^Tq\\ &\quad+(\nu\frac{\partial u}{\partial \mathfrak{n}}-p\mathfrak{n},e^2_\lambda+\vec{e}_l)_{\mathbf{L}^2(\Gamma_w)}+\hat{D}(u,\lambda,\pi)^Tq\\ &\quad-(\nu\frac{\partial\lambda}{\partial\mathfrak{n}}-\pi\mathfrak{n}+(\bar{u}\cdot\mathfrak{n})\lambda+(u\cdot\mathfrak{n})\bar{\lambda},g)_{\mathbf{L}^2(\Gamma_c)}\\ &\leq \|e^1_\lambda\|_{\mathbf{L}^2(\Omega)}\|u\|_{\mathbf{L}^2(\Omega)}+\|e_u\|_{\mathbf{L}^2(\Omega)}\|\lambda\|_{\mathbf{L}^2(\Omega)}+(Ne_g)^Tq\\ &\quad+\|\nu\frac{\partial u}{\partial\mathfrak{n}}-p\mathfrak{n}\|_{\mathbf{L}^2(\Gamma_c)}(\|e^2_\lambda\|_{\mathbf{L}^2(\Gamma_c)}+\|\vec{e}_l\|_{\mathbf{L}^2(\Gamma_c)})\\ &\quad-(\nu\frac{\partial\lambda}{\partial\mathfrak{n}}-\pi\mathfrak{n}+(\bar{u}\cdot\mathfrak{n})\lambda+(u\cdot\mathfrak{n})\bar{\lambda},g)_{\mathbf{L}^2(\Gamma_c)}\\ &\quad+(\nu\frac{\partial\lambda}{\partial\mathfrak{n}}-\pi\mathfrak{n}+(\bar{u}\cdot\mathfrak{n})\lambda+(u\cdot\mathfrak{n})\bar{\lambda},g)_{\mathbf{L}^2(\Gamma_c)}\\ &\leq \|e\|_W(\|u\|_{\mathbf{H}^2(\Omega)}+\|p\|_{\mathbf{H}^1(\Omega)}+\|\lambda\|_{\mathbf{H}^2(\Omega)})+(Ne_g)^Tq.\end{aligned} \tag{6.3.11}$$

We get

$$\|\tilde{u}\|_{\mathbf{H}^2(\Omega)}+\|\tilde{p}\|_{H^1(\Omega)}\leq c\|e_u\|_{\mathbf{H}^2(\Omega)}\leq c\|e\|_W$$

because $\tilde{u}$ is the weak solution of the at $(\bar{u},\bar{g})$ linearized Navier-Stokes equations, similar to the regularity Assumption of non-singularity 2.25. Another assertion is that we obtain by Theorem 2.11 that if (u,p) is the solution of the nonhomogenous Stokes system (2.2.1) with a sufficiently smooth boundary, then we get

$$\|u\|_{\mathbf{H}^2(\Omega)}+\|p\|_{H^1(\Omega)}\leq c(\|f\|_{\mathbf{L}^2(\Omega)}+\|g\|_{\mathbf{H}^{3/2}(\Gamma_c)}).$$

Transfering this over to the linearized Navier-Stokes equations (6.3.5) with $g=0$, we obtain

$$\begin{aligned}\|u\|_{\mathbf{H}^2(\Omega)}+\|p\|_{H^1(\Omega)} &\leq c(\|-(\bar{u}\cdot\nabla)u-(u\cdot\nabla)\bar{u}+e_u\|_{\mathbf{L}^2(\Omega)})\\ &\leq c(\|(\bar{u}\cdot\nabla)u+(u\cdot\nabla)\bar{u}\|_{\mathbf{L}^2(\Omega)}+\|e_u\|_{\mathbf{L}^2(\Omega)})\\ &\leq c(\|u\|_{\mathbf{H}^2(\Omega)}+\|e_u\|_{\mathbf{L}^2(\Omega)})\\ &\leq c(\|e_u\|_{\mathbf{L}^2(\Omega)})\end{aligned}$$

and this also leads to

$$\|\tilde{u}\|_{\mathbf{H}^2(\Omega)}+\|\tilde{p}\|_{H^1(\Omega)}\leq c\|e\|_W.$$

Adding the terms $*_3$ and $*_4$ of (6.3.10), we obtain

$$\begin{aligned}|2\mathcal{L}_{uu}(\bar{z})[u,\tilde{u}]^2| + |\mathcal{L}_{uu}(\bar{z})[\tilde{u}]^2| &\le 2|\langle (u\cdot\nabla)\tilde{u} + (\tilde{u}\cdot\nabla)u, \bar{\lambda}\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)}| \\ &\quad + |\langle 2(\tilde{u}\cdot\nabla)\tilde{u}, \bar{\lambda}\rangle_{(\mathbf{H}^1(\Omega))',\mathbf{H}^1(\Omega)}| \\ &\quad + \tilde{K}(\tilde{u},\tilde{p})(\tilde{K}(\tilde{u},\tilde{p}) + \tilde{K}(u,p)) \\ &\le 2\|u\|^{1/2}_{\mathbf{L}^2(\Omega)}\|u\|^{1/2}_{\mathbf{H}^2(\Omega)}\|\tilde{u}\|_{\mathbf{H}^1(\Omega)}\|\bar{\lambda}\|_{\mathbf{L}^2(\Omega)} \\ &\quad + c\|\tilde{u}\|_{\mathbf{H}^2(\Omega)}\|\vec{e}_d\|_{L^2(\Omega)} + c\|\tilde{p}\|_{H^2(\Omega)}\|\vec{e}_d\|_{L^2(\Omega)} \\ &\quad + \|\tilde{u}\|^{1/2}_{\mathbf{L}^2(\Omega)}\|\tilde{u}\|^{1/2}_{\mathbf{H}^2(\Omega)}\|\tilde{u}\|_{\mathbf{H}^1(\Omega)}\|\bar{\lambda}\|_{\mathbf{L}^2(\Omega)} \\ &\quad + (\|u\|_{\mathbf{H}^2(\Omega)} + \|p\|_{\mathbf{H}^1(\Omega)})(\|\tilde{u}\|_{\mathbf{H}^2(\Omega)} + \|\tilde{p}\|_{\mathbf{H}^1(\Omega)}) \\ &\quad + (\|\tilde{u}\|_{\mathbf{H}^2(\Omega)} + \|\tilde{p}\|_{\mathbf{H}^1(\Omega)})(\|\tilde{u}\|_{\mathbf{H}^2(\Omega)} + \|\tilde{p}\|_{\mathbf{H}^1(\Omega)})\end{aligned} \tag{6.3.12}$$

with similar arguments to Lemma 2.14. Furthermore, this leads to

$$\begin{aligned}|2\mathcal{L}_{uu}(\bar{z})[u,\tilde{u}]^2| + |\mathcal{L}_{uu}(\bar{z})[\tilde{u}]^2| &\le c((\|u\|_{\mathbf{H}^2(\Omega)} + \|p\|_{\mathbf{H}^1(\Omega)} + \|e\|_W)\|e\|_W + \|e\|_W) \\ &\le c\|e\|_W(\|u\|_{\mathbf{H}^2(\Omega)} + \|p\|_{\mathbf{H}^1(\Omega)} + \|e\|_W + 1).\end{aligned} \tag{6.3.13}$$

Let us now summarize (6.3.10)-(6.3.13) to:

$$\begin{aligned}0 < c\|e\|_W(\|e\|_W + \|u\|_{\mathbf{H}^2(\Omega)} + \|p\|_{H^1(\Omega)} \\ + \|\lambda\|_{\mathbf{H}^2(\Omega)} + \|\pi\|_{H^1(\Omega)}) + (Ne_q)^T q.\end{aligned} \tag{6.3.14}$$

Because (6.3.5) and (6.3.7) are at $(\bar{u},\bar{g})$ linearized Navier-Stokes equations, we obtain from the non-singularity assumption, similar to the Definition 2.25:

$$\begin{aligned}\|u\|_{\mathbf{H}^2(\Omega)} + \|p\|_{H^1(\Omega)} &\le c(\|g\|_{\mathbf{L}^2(\Gamma_c)} + \|e\|_W) \le c(|q| + \|e\|_W), \\ \|\lambda\|_{\mathbf{H}^2(\Omega)} + \|\pi\|_{H^1(\Omega)} &\le c(\|\vec{e}_l + \tilde{K}(u,p)\vec{e}_d + e^2_\lambda\|_{\mathbf{L}^2(\Gamma_w)} + \|u\|_{\mathbf{H}^2(\Omega)} + \|e^1_\lambda\|_{\mathbf{L}^2(\Omega)}) \\ &\le c(\|u\|_{\mathbf{H}^2(\Omega)} + \|p\|_{H^1(\Omega)} + \|e^1_\lambda\|_{\mathbf{L}^2(\Omega)} + +\|e^2_\lambda\|_{\mathbf{L}^2(\Gamma_w)}) \\ &\le c(\|g\|_{\mathbf{L}^2(\Gamma_c)} + \|e\|_W) \le c(|q| + \|e\|_W).\end{aligned} \tag{6.3.15}$$

The inequality (6.3.14) yields to

$$\begin{aligned}0 &< c\|e\|_W(\|e\|_W + \|g\|_{\mathbf{L}^2(\Gamma_c)}) \\ &\le c\|e\|^2_W + \frac{\delta}{2}\|g\|^2_{\mathbf{L}^2(\Gamma_c)} \le c(|q| + \|e\|_W)\end{aligned}$$

and proves Lipschitz-continuity of $e \mapsto q$ from W to $\mathbb{R}^l$. All this leads to a Lipschitz continuity of $e \mapsto (u,q,\lambda)$ from W to $Z = \mathbf{H}^2(\Omega)\times\mathbb{R}^l\times\mathbf{H}^2(\Omega)$. □

6.4 Strong regularity of the original perturbed problem

Now, we return to the original perturbed problem (P_e). To prove strong regularity of the generalized equation (6.1.3), we have to search for a solution $q_e = q(e)$ in Q_{ad}. The first idea is to take $q_e \in \tilde{Q}_{ad}$, solving $(\tilde{P}_e)$. We will show in this section that q_e is also a solution of (P_e) for a sufficiently small perturbation $e \in W$. Therefore, we have to investigate the optimal control q_e on the active set $\mathcal{A}$. In this section, we closely follow again Wachsmuth and want to refer to [109, Chapter 5, Section 4], for the ideas of the proofs of the following theorems.

Lemma 6.8. *Let $\bar{z} = (\bar{u}, \bar{p}, \bar{q}, \bar{\lambda}, \bar{\pi})$ fulfill the coercivity condition (SSC'') and let $q_a, q_b \in \mathbb{R}^l$. Then exists $\rho_e > 0$ and $\sigma > 0$ such that the optimal control function q_e of $(\tilde{P}_e)$ is active for all $e \in W$ with $\|e\|_W \leq \rho_e$. This means*

$$\begin{aligned} (\alpha M q_e - \hat{D} + Ne_q)_i^T &> \frac{\sigma}{2} \qquad \text{on } \mathcal{A}_+, \\ (\alpha M q_e - \hat{D} + Ne_q)_i^T &< -\frac{\sigma}{2} \qquad \text{on } \mathcal{A}_- \end{aligned} \tag{6.4.1}$$

and $sign(\alpha M q - \hat{D} + Ne_q)_i^T = sign(\alpha M \bar{q} - \hat{D})_i^T$ on $\mathcal{A}$.

Proof. Theorem 6.7 garantuees that the mapping $e \mapsto (u_e, p_e, q_e, \lambda_e, \pi_e)$ is Lipschitz continuous from W to Z. Furthermore, it is easy to see that $\hat{D}(z) = \int_{\Gamma_c} (\nu \partial\lambda/\partial\mathfrak{n} - \pi\mathfrak{n} + (\bar{u}\cdot\mathfrak{n})\lambda + (u\cdot\mathfrak{n})\bar{\lambda})\ e_i\ ds$ is Lipschitz continuous from $\mathbf{L}^2(\Omega) \times L^2(\Omega) \times \mathbb{R}^l \times \mathbf{L}^2(\Omega) \times L^2(\Omega)$ to $\mathbb{R}^l$. Due to $q_e \in \mathbb{R}^l$, there exists a $\sigma > 0$ with

$$|\alpha M \bar{q} - \hat{D}(z)| > \sigma.$$

Considering $i \in \mathcal{A}_+$, we obtain

$$\begin{aligned} \sigma &< (\alpha M \bar{q} - \hat{D}(z))_i^T \\ &= (\alpha M \bar{q} - \hat{D}(z))_i^T - (\alpha M \bar{q} - \hat{D}(z) + Ne_q)_i^T + (\alpha M \bar{q} - \hat{D}(z) + Ne_q)_i^T \\ &\leq c\|e\|_W + (\alpha M \bar{q} - \hat{D}(z) + Ne_q)_i^T. \end{aligned}$$

Taking ρ_e sufficiently small yields

$$(\alpha M \bar{q} - \hat{D}(z) + Ne_q)_i^T > \frac{\sigma}{2}$$

Analogously leads $i \in \mathcal{A}_-$ to

$$(\alpha M \bar{q} - \hat{D}(z) + Ne_q)_i^T < -\frac{\sigma}{2}.$$

□

An important consequence is that the control function q_e even satisfies the variational inequality based on the admissible set Q_{ad} and not only for $\tilde{Q}_{ad} \subset Q_{ad}$.

Lemma 6.9. *With the assumptions and ρ_e as in the last Lemma 6.8, the control function q_e associated to a perturbation e, fulfilling $\|e\|_W \leq \rho_e$, satisfies*

$$(\alpha M q_e - \hat{D} + N e_q)^T (q - q_e) \geq 0 \ \forall q \in Q_{ad}.$$

Proof. Let $q \in Q_{ad}$. Then we are able to split $(\alpha M q_e - \hat{D} + N e_q)^T (q - q_e)$ into

$$\begin{aligned}(\alpha M q_e - \hat{D} + N e_q)^T (q - q_e) = &\sum_{i \in \mathcal{A}_+} (\alpha M q_e - \hat{D} + N e_q)_i^T (q - q_e)_i \\ &+ \sum_{i \in \mathcal{A}_-} (\alpha M q_e - \hat{D} + N e_q)_i^T (q - q_e)_i \\ &+ \sum_{i \notin \mathcal{A}} (\alpha M q_e - \hat{D} + N e_q)_i^T (q - q_e)_i\end{aligned}$$

with $q_e \in \tilde{Q}_{ad}$. The third sum is nonnegative due to the fact that it is part of the optimality system of $(\tilde{P}_e)$.

Additionally, we have

$$\begin{aligned} q_e = \bar{q} = q_a \quad &\text{on } \mathcal{A}_+, \\ q_e = \bar{q} = q_b \quad &\text{on } \mathcal{A}_-. \end{aligned}$$

Because of $\|e\|_W \leq \rho_e$, Lemma 6.8 leads to

$$\begin{aligned} (\alpha M q_e - \hat{D} + N e_q)_i^T > \frac{\sigma}{2} \qquad &\text{on } \mathcal{A}_+, \\ (\alpha M q_e - \hat{D} + N e_q)_i^T < -\frac{\sigma}{2} \qquad &\text{on } \mathcal{A}_-. \end{aligned}$$

Thus, we obtain

$$(\alpha M q_e - \hat{D} + N e_q)^T (q - q_e) \geq 0 \ \forall q \in Q_{ad}.$$

□

Remark 6.10. *The triple (u_e, g_e, λ_e) of Theorem 6.9 fulfills the optimality system of the optimization perturbed problem (P_e) which is equivalent to the linearized and perturbed generalized equation* (6.2.1).

The next theorem shows with the help of the equality (6.3.2) that (u_e, q_e, λ_e) minimizes the optimization problem (P_e).

Theorem 6.11. *Under the same assumptions as in the two last lemma, there exist $\rho_e, \rho_q > 0$ such that the control q_e belonging to a perturbation $e \in W$ satisfying $\|e\|_W \leq \rho_e$ is locally optimal for the optimal control problem (P_e) and fulfills*

$$J^e(u_e, q_e) \leq J^e(u, q)$$

for all $q \in Q_{ad}$ fulfilling $|q - q_e| \leq \rho_q$, where u and u_e are the weak solutions of (6.2.4) *associated to q and q_e, respectively.*

The proof is very similar to [109, Theorem 5.15].

Now, we have shown the existence of a solution of the linearized and perturbed equation (6.2.1).

Theorem 6.11 shows that q_e is the unique optimal solution of Problem (P_e) in $\mathcal{B}_{\rho_q}(\bar{q})$ with perturbations e in $\mathcal{B}_{\rho_e}(0)$. By Theorem 6.7 u_q and λ_q are in $\mathcal{B}_{c_u\rho_e}(\bar{u})$ and $\mathcal{B}_{c_\lambda\rho_e}(\bar{\lambda})$ with the Lipschitz-constants c_u and c_λ given by Theorem 6.7. This leads to the unique solvability of (6.2.1) in $\mathcal{B}_{c_u\rho_e}(\bar{u}) \times \mathcal{B}_{\rho_q}(\bar{q}) \times \mathcal{B}_{c_\lambda\rho_e}(\bar{\lambda})$ for perturbations e in $\mathcal{B}_{\rho_e}(0)$.

This yields the strong regularity of the generalized equation (6.1.3).

The investigations of this chapter have shown that we only find a local solution of the linearized subproblems of the SQP-method in a neighborhood of the reference solution $\bar{v} = (\bar{u}, \bar{q})$. The idea is now to modify Q_{ad} to

$$Q_{ad}^{\rho} := Q_{ad} \cap \{h \in \mathbb{R}^l : |h - \bar{q}| \leq \rho\}$$

to have the solution of the linearized subproblems in a close neighborhood of the reference solution. See [104] for more details.

Altogether, it follows the local convergence of the SQP-method, see Theorem 6.4.

Theorem 6.12. *Let $\bar{z} \in Z$ fulfill the coercivity condition (SSC''). Then exist $\rho > 0$ such that the SQP-method with control constraint Q_{ad}^{ρ} generates a uniquely determined sequence (u_k, q_k, λ_k), $q_k \in Q_{ad}^{\rho}$ for every starting point (u_0, q_0, λ_0), with $q_0 \in Q_{ad}^{\rho}$ and we obtain*

$$\begin{aligned}\|u_{k+1} - \bar{u}\|_{\mathbf{H}^2(\Omega)} + |q_{k+1} - \bar{q}| + \|\lambda_{k+1} - \bar{\lambda}\|_{\mathbf{H}^2(\Omega)} \\ \leq c(\|u_{k+1} - \bar{u}\|^2_{\mathbf{H}^2(\Omega)} + |q_{k+1} - \bar{q}|^2 + \|\lambda_{k+1} - \bar{\lambda}\|^2_{\mathbf{H}^2(\Omega)}).\end{aligned}$$

Chapter 7

The nonstationary case

After investigating the steady-state problem, we want to focus in this chapter on the optimization problem subject to the nonstationary Navier-Stokes equations. A simplified model, similar to the stationary case, could look as follows. We want to minimize the negative lift

$$\min_{u,g} \; J(u,g) := -\iint\limits_{\Sigma} \left(\nu \frac{\partial u}{\partial \mathfrak{n}} - p\mathfrak{n} \right) \cdot \vec{e}_l \, \mathrm{d}x\mathrm{d}t + \frac{\alpha}{2} \|g\|_{\mathbf{L}^2(\Sigma)}$$

subject to the nonstationary Navier-Stokes equations describing the motion of the fluid around the body

$$\begin{aligned} u_t + \nu\Delta u + (u \cdot \nabla)u + \nabla p = & \;\; f \text{ on } Q, \\ \operatorname{div} u = & \;\; 0 \text{ on } Q, \\ u = & \;\; g \text{ on } \Sigma, \\ u(0) = & \;\; u_0 \text{ on } \Omega \end{aligned}$$

and the control constraint

$$g \in G_{ad},$$

where $Q := \Omega \times (0,T)$ with its boundary $\Sigma := \partial Q = \Gamma \times (0,T)$ with a fixed time T. But in contrast to the stationary case, we allow high Reynolds numbers in the nonstationary situation hand consider a problem closer to the real setting of the high-lift configuration, see e.g. [19, 89, 98].

In this situation, we have to deal with turbulence's, which we simulated by a k-ω-WILCOX98 model, see [113], where the equations for k and ω are given by:

$$\begin{aligned}u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}&=\frac{\partial}{\partial y}[(\nu+\nu_T)\frac{\partial u}{\partial y}]\\ u\frac{\partial k}{\partial x}+v\frac{\partial k}{\partial y}&=\nu_T(\frac{\partial u}{\partial y})^2-\beta^*\omega k+\frac{\partial}{\partial y}[(\nu+\sigma^*\nu_T)\frac{\partial k}{\partial y}]\\ u\frac{\partial \omega}{\partial x}+v\frac{\partial \omega}{\partial y}&=\alpha\frac{\omega}{k}\nu_T(\frac{\partial u}{\partial y})^2-\beta\omega^2+\frac{\partial}{\partial y}[(\nu+\sigma^*\nu_T)\frac{\partial \omega}{\partial y}]\\ \nu_T&=\alpha^*\frac{k}{\omega},\end{aligned} \tag{7.0.2}$$

where u and v are velocity components in the streamwise x and normal y directions, ν is the kinematic molecular viscosity, ν_T is the kinematic eddy viscosity and β, β^*, σ, σ^* are parameters, which are defined in [113].

Due to the high dimension of the discretized equations, the computing times for any forward solution of the model are extremely large so that a mathematical optimization of the periodic actuation is fairly unrealistic. In [19], a generic high-lift configuration was investigated and one forward solution took about 48 hours. In the case of the SCCH configuration, the computation time was nearly twice that number.

Let us mention that in this nonstationary case, we want to consider a special example with the following setting.

Setting 7.1. *We consider the incompressible two-dimensional flow over the swept constant chord half (SCCH) high-lift configuration, see Figure 7.1. The chord length c is denoted by $L_{ref} = 1.275$ and the inflow by $u_\infty = 1$. The chord length is the length of the wing in the flow direction. The Reynolds number is $Re = u_\infty c/\nu$, where ν is the kinematic viscosity of the fluid. The leading edge slat deflection angle is 26.5°, the flap deflection is 37°and the angle of attack of the wing is 6°. The periodic actuation is introduced by a zero-net-mass-flux actuator on a small slit on the flap, where the flow fully separates. The actuation velocity is*

$$g(t) = B\cos(\Omega^a t), \tag{7.0.3}$$

where $\Omega^a = 2\pi St^a$ is the angular actuation frequency, B the actuation amplitude, $St^a = f^a c/u_\infty$ the Strouhal number and f^a is the actuation frequency. Analogously, we define $St^n = f^n c/u_\infty$ with the vortex-shedding frequency f^n. The actuation intensity is characterized by the dimensionless coefficient

$$C_\mu = \frac{H}{c}\left(\frac{B}{u_\infty}\right)^2$$

with the slot width $H = 0.001238c_{fl}$ and the relative chord length $c_{fl} = 0.254c$. The full $k - \omega$ Wilcox98 model was solved by unsteady Reynolds-averaged Navier-Stokes (URANS) equations with the ELAN code[1]*. The with c_{fl} non-dimensionalized natural Strouhal number is $St^n_{fl} = f^n c_{fl}/u_\infty = 0.32$. The actuation is described by a momentum coefficient of $C_\mu = 405 \times 10^{-5}$ and an actuation frequency of $St^a_{fl} = 0.6$.*

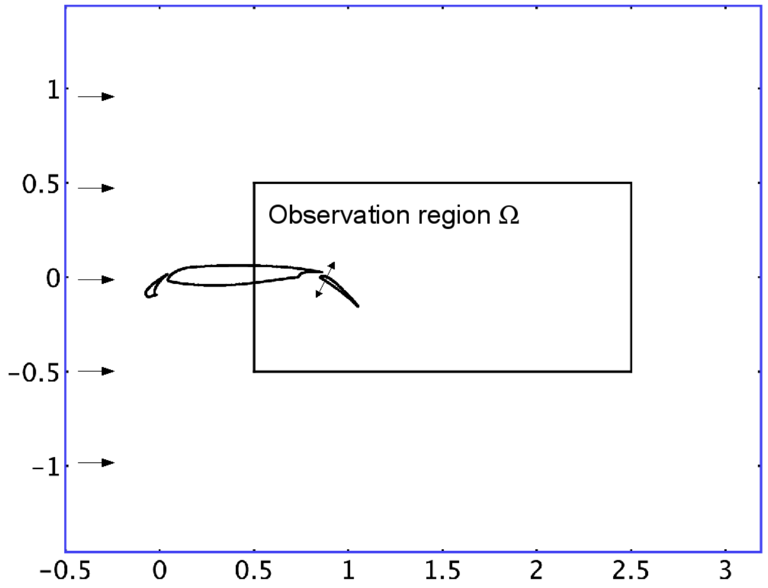

Figure 7.1: The SCCH high-lift configuration, where the periodic excitation is implemented on the flap.

We think that a model reduction is advisable, because of the reasons above. Our goal is to establish a reduced-order model (ROM) as a basis for our optimization problem.

7.1 Model reduction

The topic of model reduction is currently in great demand by engineers. A widely used method is POD [4, 56, 62, 63, 111]. In the case of the high-lift configuration, the application of standard POD does not align to the target of robust dynamical least-order models for the real flow. To establish the

[1] Developed at the Computational Fluid Dynamics and Aeroacoustics Group (Professor F. Thiele) at the TU Berlin.

reduced-order model (ROM), the computed POD basis has to be inserted as a Galerkin basis in the full WILCOX98 model, see (7.0.2). The associated implementation would be a time consuming task.

There were several approaches to deal with these problems, e.g. an extension of POD to data compression of multiple operation points, see [59] for sequential POD or [93] for DPOD. We follow an alternative approach suggested in [66, 74] of a canonical reduction with parameter identification. Here, a very small system of nonlinear ODEs is adapted to the previously computed flows in the actuated and nonactuated case. This small system is easily tractable by optimization.

In the next chapter, we will go into details of this technique, present our modification and report about first experiences in the simplified two-dimensional Setting 7.1. Our numerical results are promising for future optimization tasks.

We also refer to [96] for a similar approach.

Let us first briefly introduce the proper orthogonal decomposition (POD), which we will need to establish the reduced-order model.

7.2 Proper orthogonal decomposition POD

This section is based on the theory in [105, 106, 108, 107, 111]. In this section, we want to introduce the proper orthogonal decomposition (POD) and the way to calculate the POD basis by minimizing a least-square error formula.

There are two cases, the infinite-dimensional and the finite-dimensional one, of the POD basis. We will consider the finite-dimensional one, because we want to concentrate on real computations and there we don't have the whole trajectory $u(t)$. Therefore, we get an ensemble of snapshots. Afterwards, it is possible to prove that the POD basis is the best orthogonal system in the ensemble capturing more kinetic energy than any other one having the same basis number.

So, let $\hat{u}(t_i) \in V$ be the N snapshots computed by the full dynamical system with the Setting 7.1 at given times t_i with $i = 1, \ldots, N$, $0 = t_1 \leq t_2 \leq \ldots \leq t_{N-1} \leq t_N = T$ and $N \in \mathbb{N}$ and at least one snapshot has to be non-zero. In the next chapter, we apply the POD method for both cases, the natural and the actuated one. Let us just explain the method for one case. We define $\hat{u}_i := \hat{u}(t_i)$, $i = 1, \ldots, N$.

Furthermore, we define $\mathcal{V}_N$ as the span of the N snapshots

$$\mathcal{V}_N := span\{\hat{u}(t_1), \ldots, \hat{u}(t_N)\}$$

with $1 \leq \dim \mathcal{V}_N \leq N$.

Let $\{\Phi_k\}_{k=1}^N$ denotes an orthonormal basis for $\mathcal{V}_N$, which has still to be computed, then each snapshot $\hat{u}(t_i)$, $i = 1, \ldots, N$, can be expressed by

$$\hat{u}(t_i) = \sum_{k=1}^{N} \langle \hat{u}(t_i), \Phi_k \rangle_V \Phi_k \text{ for } i = 1, \ldots, N.$$

The idea is to expect that only some of the orthonormal basis functions $\{\Phi_k\}_{k=1}^N$ keep most of the kinetic energy so that they can represent the structure of the snapshots as good as possible. Let therefore $M \in \mathbb{R}$ with $0 < M \leq N$ be given.

Mathematically, we can formulate the problem of finding the orthonormal system $\{\Phi_k\}_{k=1}^M$ by

$$\begin{cases} \min\limits_{\Phi_k} \sum\limits_{j=1}^{N} \alpha_j \|\hat{u}(t_j) - \sum\limits_{k=1}^{M} \langle \hat{u}(t_j), \Phi_k \rangle \Phi_k \|_V^2 \\ \text{subject to } \langle \Phi_i, \Phi_j \rangle_V = \delta_{ij}, \ 1 \leq i, j \leq M \end{cases} \tag{7.2.1}$$

where α_i's stand for the trapezoidal weights

$$\alpha_1 = \frac{t_2 - t_1}{2}, \ \alpha_i = \frac{t_{i+1} - t_{i-1}}{2} \text{ for } 2 \leq i \leq N-1, \ \alpha_n = \frac{t_N - t_{N-1}}{2}.$$

For the next remark, see [111] Chapter 3.

Remark 7.2.

- *The trapezoidal approximation for the integral*

  $$\mathcal{I}(u) = \int_0^T \|\hat{u}(t) - \sum_{k=1}^{M} \langle \hat{u}(t), \Phi_k \rangle \Phi_k \|_V^2 \mathrm{d}t$$

  is

  $$\mathcal{I}_n(u) = \sum_{j=1}^{N} \alpha_j \|\hat{u}(t_j) - \sum_{k=1}^{M} \langle \hat{u}(t_j), \Phi_k \rangle \Phi_k \|_V^2$$

  for all $u \in C([0,T], V)$ and it follows that $\lim\limits_{n \to \infty} \mathcal{I}_n(u) = \mathcal{I}(u)$.

- *The least-square problem is equivalent to the largest mean square projection of the snapshot, namely*

  $$\begin{cases} \max\limits_{\Phi_k} \sum\limits_{j=1}^{N} \alpha_j \sum\limits_{k=1}^{N} |\langle \hat{u}(t_j), \Phi_k \rangle_V|^2 \\ \textit{subject to } \langle \Phi_i, \Phi_j \rangle_V = \delta_{ij}, \ 1 \leq i, j \leq M. \end{cases}$$

Consider the linear mapping $\mathcal{F}_N : \mathbb{R}^N \to \mathcal{V}_N,\ e_k \mapsto \hat{u}_k := \hat{u}(t_k)$, where e_k denotes the $k-th$ canonical basis vector of $\mathbb{R}^N$ and $\hat{u}_k$ are the snapshots. Considering the following proposition, we can see that $\mathbb{R}^N$ corresponds with V and $\mathcal{V}_N$ with W. For a proof of the following singular value decomposition (SVD), we refer for instance to [105].

Proposition 7.3. *Let $F : V \to W$ be a linear operator, where V and W denote two finite-dimensional real Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_V$ and $\langle \cdot, \cdot \rangle_W$ and $dimV = m$ and $dimW = n$ with $m \geq n$. Then there exist real numbers $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$ and orthonormal bases $\{v_k\}_{k=1}^n$ of V and $\{w_k\}_{k=1}^n$ of W, such that*

$$F(v_k) = \sigma_k w_k,\ F^*(w_k) = \sigma_k v_k,$$

for $k = 1, \cdots, n$, where the adjoint operator F^* of F is defined by the following definition.

Definition 7.4. *Let $\{V, \langle \cdot, \cdot \rangle_V\}$ and $\{W, \langle \cdot, \cdot \rangle_W\}$ be real Hilbert spaces and $F : V \to W$ a linear operator. We call F^* the adjoint operator of F if*

$$\langle w, Fv \rangle_W = \langle F^* w, v \rangle_V$$

for all $w \in W$ and all $v \in V$. F is called self-adjoint, if

$$F^* = F.$$

Then, we obtain with $\langle v, w \rangle_{\mathbb{R}^N} = \sum_{j=1}^{N} \alpha_k v_k w_k$ for all $v \in \mathbb{R}^N$

$$\mathcal{F}_N(v) = \sum_{j=1}^{N} \alpha_j \langle v, e_j \rangle_{\mathbb{R}^N} \mathcal{F}_N(e_j) = \sum_{j=1}^{N} \alpha_j \langle v, e_j \rangle_{\mathbb{R}^N} \hat{u}_j. \tag{7.2.2}$$

Assuming $\mathcal{F}_N^*$ as the adjoint of $\mathcal{F}_N$, then follows

$$\begin{aligned}
\langle \mathcal{F}_N(v), \Phi \rangle_V &= \langle \sum_{k=1}^{N} \alpha_k \langle v, e_k \rangle_{\mathbb{R}^N} \hat{u}_k, \Phi \rangle_V \\
&= \sum_{k=1}^{N} \alpha_k \langle \hat{u}_k, \Phi \rangle_V \langle e_k, v \rangle_{\mathbb{R}^N} = \sum_{k=1}^{N} \alpha_k \langle \hat{u}_k, \Phi \rangle_V v_k \\
&= \langle \begin{bmatrix} \langle \hat{u}_1, \Phi \rangle_V \\ \vdots \\ \langle \hat{u}_N, \Phi \rangle_V \end{bmatrix}, v \rangle_{\mathbb{R}^N}
\end{aligned}$$

for all $\Phi \in \mathcal{V}_N$ and so, we can interpretate the adjoint operator as

$$\mathcal{F}_N^* \Phi = \begin{bmatrix} \langle \hat{u}_1, \Phi \rangle_V \\ \vdots \\ \langle \hat{u}_N, \Phi \rangle_V \end{bmatrix} \tag{7.2.3}$$

for all $\Phi \in \mathcal{V}_N$. The idea is now to define $\mathcal{R}_N := \mathcal{F}_N \mathcal{F}_N^*$ and $\mathcal{K}_N := \mathcal{F}_N^* \mathcal{F}_N$. Together with (7.2.2) and (7.2.3), we derive with $\langle \hat{u}_k, \mathcal{F}_N(\cdot) \rangle_V = \langle \begin{bmatrix} \langle \hat{u}_k, \hat{u}_1 \rangle_V \\ \vdots \\ \langle \hat{u}_k, \hat{u}_N \rangle_V \end{bmatrix}, \cdot \rangle_{\mathbb{R}^N}$ and the following remark, see also [111], page 25,

$$\mathcal{R}_N = \sum_{j=1}^{N} \alpha_j \langle \hat{u}_j, \cdot \rangle_V \hat{u}_j$$

$$\mathcal{K}_N = \begin{bmatrix} \langle \hat{u}_1, \mathcal{F}_N(\cdot) \rangle_V \\ \vdots \\ \langle \hat{u}_N, \mathcal{F}_N(\cdot) \rangle_V \end{bmatrix}.$$

Remark 7.5.

- *The operator $\mathcal{R}_N$ is bounded, self-adjoint and non-negative.*
- *By Hilbert-Schmidt theory exists an orthonormal basis $\{\Phi_k\}_{k=1}^N$ and non-negative real numbers $\{\lambda_k\}_{k=1}^N$ such that*

$$\mathcal{R}_N \Phi_i = \lambda_i \Phi_i \textit{ and } \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N.$$

By the Lagrangian theory it follows that the first-order optimality condition for the least-square problem (7.2.1) is

$$\mathcal{R}_N \Phi_i = \lambda_i \Phi_i \text{ and } \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N. \tag{7.2.4}$$

For more details, see [106]. We are able to obtain the orthonormal basis $\{\Phi_k\}_{k=1}^N$ by solving (7.2.4), the solution of (7.2.1).

Additionally, we have the following theorem to solve (7.2.1) by choosing a fixed M.

Theorem 7.6. *Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N \geq 0$ be the non-negative eigenvalues and $\{\Phi_k\}_{k=1}^N$ the associating eigenvectors of $\mathcal{R}_N$. Let $M \ll N$, then $\{\Phi_k\}_{k=1}^N$ is orthonormal with rank M and* (7.2.1) *satisfies*

$$\sum_{j=1}^{N} \alpha_j \| \hat{u}(t_j) - \sum_{k=1}^{M} \langle \hat{u}(t_j), \Phi_k \rangle \Phi_k \|_V^2 = \sum_{j=M+1}^{N} \lambda_j.$$

For a proof, we refer to [111].

We see by Proposition 7.3 that the two orthonormal systems of the finite-dimensional spaces can be transformed into the other one by a linear mapping or its adjoint, if they are known.

Additionally, we find an orthonormal basis $\{v_k\}_{k=1}^N$ in $\mathbb{R}^N$ such that for $k = 1, \dots, N$

$$\mathcal{K}_N(v_k) = \lambda_k v_k.$$

Now, we are able to determine the optimal orthonormal basis $\{\Phi_k\}_{k=1}^M$ of (7.2.1) in $\mathcal{V}_N$ by the linear mapping $\mathcal{F}_N$

$$\mathcal{F}_N(v_k) = \sqrt{\lambda_k}\Phi_k, \text{ i.e. } \Phi_k = \frac{1}{\sqrt{\lambda_k}}\mathcal{F}_N(v_k)$$

for a fixed M, $k = 1, \dots, M$.

In the following, we consider the problem of finding the so-called 'modes' $\{u_i(x)\}_{i=0}^M$ so that the Galerkin approximations, which are defined with the corresponding mean flows $u_0(x) = 1/N \sum_{i=1}^N \hat{u}_i(x)$ by

$$u^{[M]}(x,t) := \sum_{i=0}^M a_i(t) u_i(x),$$

with $a_0 \equiv 1$ and $a_i(t) := (u - u_0, u_i)_\Omega$, minimizes the energy-related error

$$\chi_u := \frac{1}{N} \sum_{i=1}^N \|\hat{u}_i(\cdot) - u^{[M]}(\cdot, t_i)\|_{L^2(\Omega)}$$

compared to all other bases $\{w_i(x)\}_{i=1}^M$ and corresponding Galerkin approximations, i.e.

$$\chi_u \leq \chi_w.$$

For a homogeneous fluid and an incompressible flow, the flow velocities $u(x,t)$, having components u_i in the x_i coordinate direction, can be splitted into a mean part $\overline{u}(x)$ and a fluctuating part $u'(x,t)$ using the so-called Reynolds decomposition:

$$u_i = \overline{u_i} + u_i'.$$

Now, $u_0(x)$ represents the the mean part $\overline{u}(x)$ and $\sum_{i=1}^M a_i(t)u_i(x)$ without $u_0(x)$ the fluctuation part $u'(x,t)$ of the Reynolds decomposition.

This theory leads to the following algorithm to calculate the POD modes and coefficients.

Algorithm 7.7.

1. *Compute the averaged (mean) flow*

$$u_0(x) := \frac{1}{N}\sum_{i=1}^{N} \hat{u}_i(x).$$

We denote by $\tilde{u} = u - u_0$ the fluctuation of u from this mean flow.

2. *Compute the correlation matrix $C \in \mathbb{R}^{N\times N}$*

$$C_{i,j} = \frac{1}{N}(\tilde{u}_i, \tilde{u}_j).$$

3. *Compute the eigenvalues λ_i and the set of normalized eigenvectors $\vec{v}_i$, $i = 1, \ldots, N$ of C with :*

$$C\vec{v}_i = \lambda_i \vec{v}_i. \tag{7.2.5}$$

4. *Compute the POD modes*

$$u_i := \frac{1}{\sqrt{M\lambda_i}}\sum_{i=1}^{N} \vec{v}_i \tilde{u}_i.$$

5. *Compute the Fourier coefficients*

$$a_i(t_j) := (\tilde{u}_j, u_i)_\Omega.$$

For the theory in the infinite-dimensional case and the optimality of the pod basis system, we refer to [106, 111].

Let us now present the POD modes for the Setting 7.1.

First, for the unactuated system, $N = 567$ snapshots $\hat{u}_i^n(x) := \hat{u}^n(x, t_i)$ were determined at equidistant discrete times t_i, $i = 1, \cdots, N$, covering 6 convective time units. Analogously, N snapshots $\hat{u}_i^a(x) := \hat{u}^a(x, t_i)$, $i = 1, \cdots, N$, are computed for the actuated system by a URANS simulation with a WILCOX98 turbulence k-ω-model and a Reynolds number of $1.756 \cdot 10^6$.

We chose a fairly large actuation amplitude B for the actuated case to get significant differences between the frequencies of the operating conditions.

The POD method for our Setting 7.1 yields the eigenvalues in Figure 7.2, where the typical pairs of eigenvalues are demonstrated. In Figure 7.2, one can also see that the first pair of modes contains the most energy. Due to this reason, we consider only the first pair in the next chapter to introduce the reduced-order model.

The mean flows is presented in Figure 7.3, the first mode in Figure 7.4 and the second mode is presented in Figure 7.5.

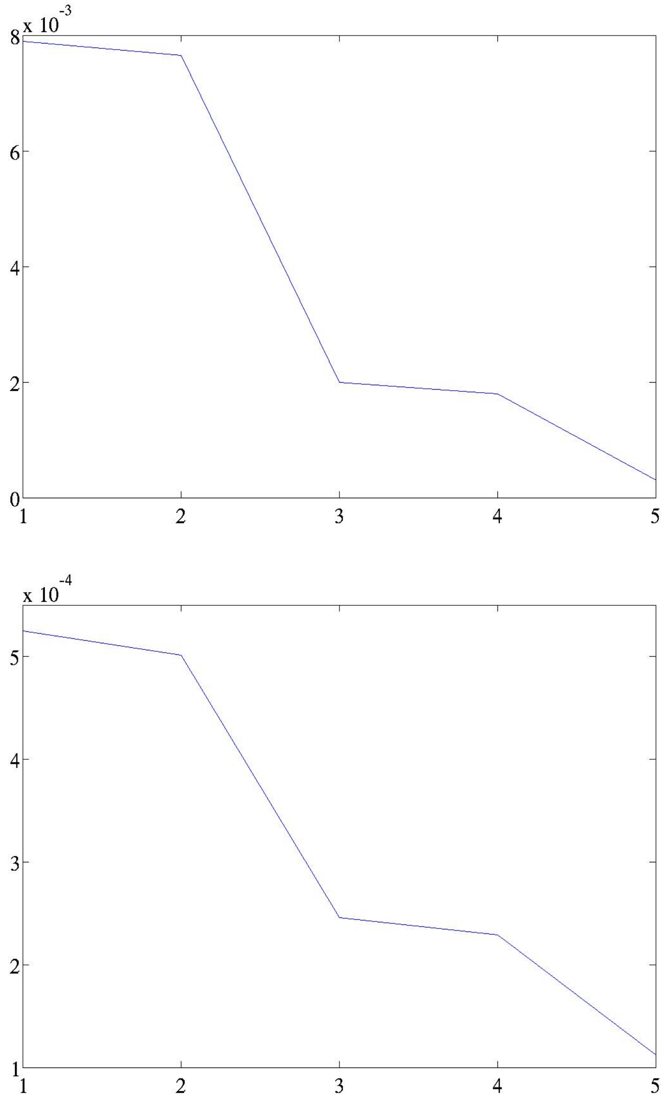

Figure 7.2: Eigenvalues of the natural (top) and the actuated (bottom) flow.

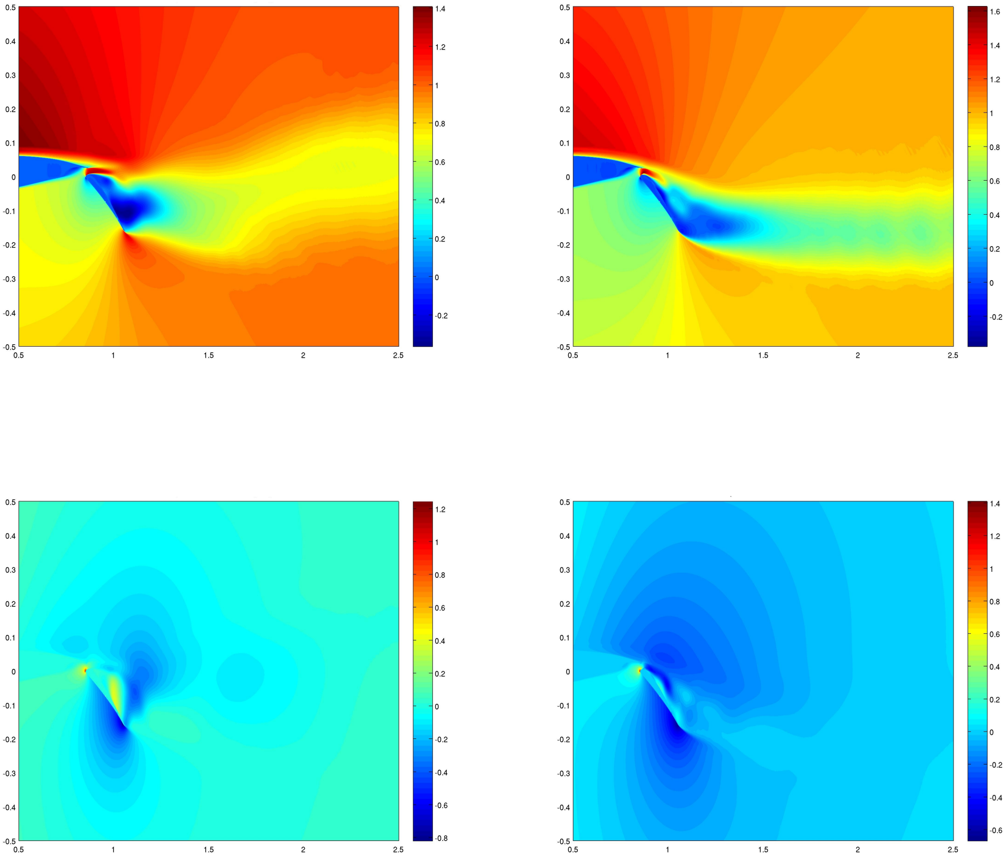

Figure 7.3: Mean flow of the natural $u_0^n = (u_0^n \; v_0^n)^T$ (u_0^n: top left, v_0^n: bottom left) and the actuated $u_0^a = (u_0^a \; v_0^a)^T$ (u_0^a: top right, v_0^a: bottom right) case.

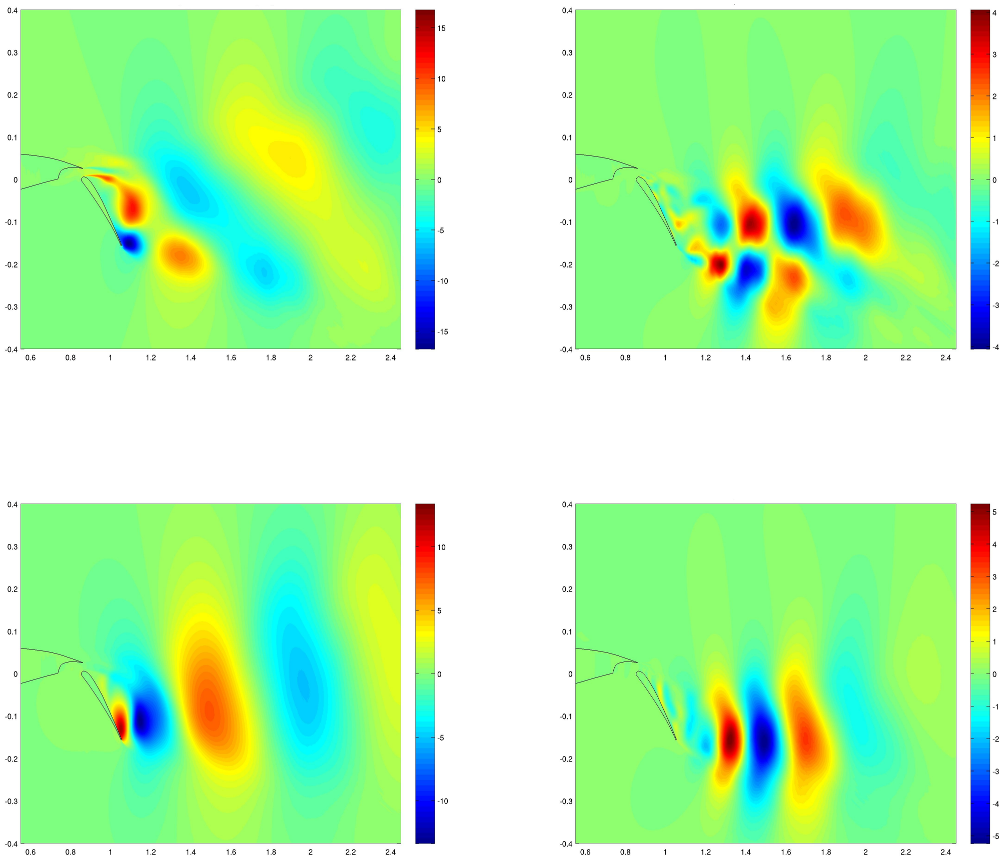

Figure 7.4: The first mode of the natural $u_1^n = (u_1^n\ v_1^n)^T$ (u_1^n: top left, v_1^n: bottom left) and the actuated $u_1^a = (u_1^a\ v_1^a)^T$ (u_1^a: top right, v_1^a: bottom right) case.

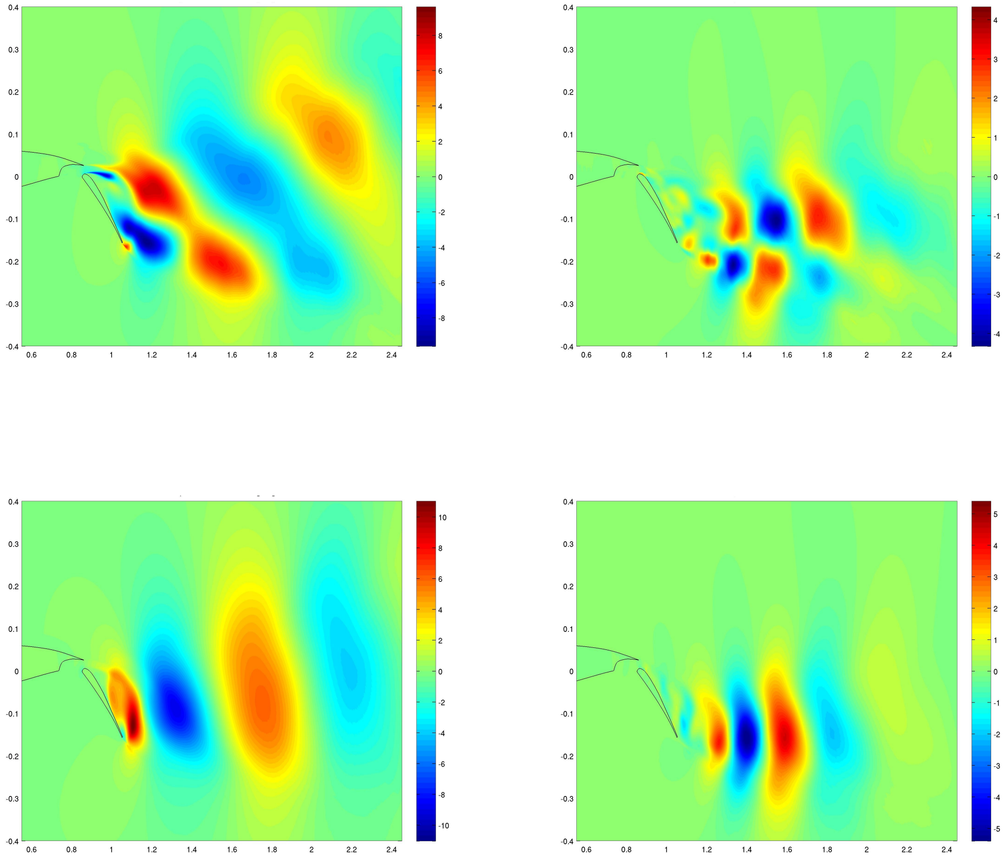

Figure 7.5: The second mode of the natural $u_2^n = (u_2^n \; v_2^n)^T$ (u_2^n: top left, v_2^n: bottom left) and the actuated $u_2^a = (u_2^a \; v_2^a)^T$ (u_2^a: top right, v_2^a: bottom right) case.

Chapter 8

Reduced-order model (ROM)

As mentioned in the last chapter, we want to introduce a low-order model describing the lift-increasing effect of high-frequency forcing.

In this chapter, we outline the approach of some engineers, see [66, 74], to set up a reduced-order model without a complete and detailed mathematical reflection. This procedure is based on many observations and is adapted to the given problem related to Navier-Stokes equations and high-lift configurations.

After a comprehensible explanation for the design of the reduced-order model, adopted almost as it stands from [66, Section 3], we will present in Section 8.1 a summary of the core statements of developing the ROM more detailed, adopted almost as it stands from [66, Section 4 and 5]. In section 8.2 and 8.3, we present our modifications on their dynamical system and a lift formula based only on the Fourier coefficients. Numerical results to compare this reduced-order model with the full turbulence model are considered at the end of this chapter. The Sections 8.2-8.4 are published by John, Noack, Schlegel, Tröltzsch and Wachsmuth in [57]. Based on this ROM, we will establish in Chapter 9 a reduced optimization problem.

The dynamical system should reflect the following behavior of the unsteady Reynolds-averaged Navier-Stokes (URANS) simulation:

(i) *von Kármán* vortex shedding without actuation: a vortex street is a phenomenon of fluid mechanics for a repeating pattern of swirling vortices behind a bluff body caused by the unsteady separation of fluid flow. It is named after the engineer, *Theodore von Kármán* (1881-1963),

(ii) lock-in shear-layer shedding under high-frequency forcing: a shear-layer is the transition region between two parallel fluid flows and a shear-layer shedding means that a boundary-layer fixed to a body separates from the body surface,

(iii) a transient behavior from the natural case (i) to the actuated one under forcing (ii): this means that the system should describe the transition from the natural flow to the actuated flow and

(iv) a transient behavior from the actuated case (ii) to the natural case (i) when forcing is turned off.

Due to the periodic actuation, we have to consider oscillatory flows, which are characterized by an amplitude A and a phase α, i.e. the argument of sinusoidal functions, if they are linear in time. We can consider them as polar coordinates of the phase-space $(a_1, a_2) = A[\cos\alpha, \sin\alpha]$.

First, we are searching for a system describing the natural flow. Noack [74] described the self-amplified amplitude-limited behavior of vortex shedding by the following Landau-equation (8.0.1). Let therefore, the superscript n stand for the unactuated natural case, σ^n for the positive growth rate, $\sigma^{n,n}$ for the positive Landau constant and $A^n = \sqrt{a_1^2 + a_2^2}$ for the amplitude. For simplicity, we assume the frequency ω^n as a constant.

$$\begin{aligned}
\dot{a}_1 &= \tilde{\sigma}^n a_1 - \omega^n a_2 \\
\dot{a}_2 &= \omega^n a_1 + \tilde{\sigma}^n a_2 \\
\tilde{\sigma}^n &= \sigma^n - \sigma^{n,n}(A^n)^2,
\end{aligned} \tag{8.0.1}$$

The superscript a stands in the following for the actuated case.

The shear-layer dynamics is stimulated by high-frequency forcing $g(t) = B\cos(\beta)$ with an amplitude B, a phase β and the frequency $\dot{\beta} = \omega^a$. The shear layer denotes in fluid mechanics the transition area between two parallel streams with different velocities in contrast to wall-bounded boundary layers. The phase difference of the actuation with respect to the oscillation of the flow is denoted by θ. That means that the oscillation flow has the phase $\theta + \beta$. The behavior is most easily modeled by a linear damped oscillator with a periodic forcing at the eigenfrequency. Here is σ^a a negative growth rate and $g_3, g_4 \in \mathbb{R}$ are parameters to describe the gain of actuation. In contrast to the natural case(8.0.1), we use in the actuated case the indices 3 and 4 instead of 1 and 2. Thus, the actuated flow is described by the system:

$$\begin{aligned}
\dot{a}_3 &= \sigma^a a_3 - \omega^a a_4 + g_3 B\cos(\theta + \beta), \\
\dot{a}_4 &= \omega^a a_3 + \sigma^a a_4 + g_4 B\sin(\theta + \beta).
\end{aligned} \tag{8.0.2}$$

In reality, every flow $\tilde{u}$ with actuation $\tilde{g}$ consists of a superposition of several frequencies $\tilde{\omega}^1, \dots, \tilde{\omega}^N$, $N \in \mathbb{N}$, such that its energy $E(\tilde{g})$ is the sum of energies associated to the frequencies $E^i(\tilde{g})$, $i = 1, \dots, N$:

$$E(\tilde{g}) = \sum_{i=1}^{N} E^i(\tilde{g}).$$

Notice that we assume in both the natural flow, $g^n = 0$, and the actuated flow, $g^a = B^a \cos(\omega^a t)$, a constant frequency ω^n and ω^a, respectively, such that the associated energies $E(g^n)$ and $E(g^a)$ are related to the only frequencies ω^n and ω^a. Thus, we characterize the energies in the natural and the actuated flow by the squares of the amplitudes $(A^n)^2$ and $(A^a)^2$, respectively:

$$E(g^n) := (A^n)^2 \text{ and } E(g^a) := (A^a)^2.$$

Let us analogously denote by $(A^i_{\tilde{g}})^2$ the energy of the flow with actuation $\tilde{g}$ associated to the frequency ω^i :

$$(A^i_{\tilde{g}})^2 := E^i(\tilde{g}). \tag{8.0.3}$$

Now, we comprise both oscillations in a four-dimensional phase space $[a_1\ a_2\ a_3\ a_4]$. With $B = 0$, i.e. $a_3 = a_4 = 0$, this system describes the natural flow, according to (i). To obey (ii), the oscillation at the natural frequency has to vanish with increasing actuation amplitude $B > 0$. To achieve this, we decrease the growth rate of the natural case σ^n with the help of the growth of the high-frequency amplitude $A^a = \sqrt{a_3^2 + a_4^2}$. Analogously to the damping term $-\sigma^{n,n}(A^n)^2$ in the Landau system, we add additionally the term $-\sigma^{n,a}(A^a)^2$ and we get

$$\tilde{\sigma}^n = \sigma^n - \sigma^{n,n}(A^n)^2 - \sigma^{n,a}(A^a)^2.$$

We see that the energy in the natural case decreases with increasing energy in the actuated case and vice versa. This guarantees $a_1 = a_2 = 0$ at the forced state, according to (ii). We substitude the terms $g_3 B \cos(\theta + \beta)$ and $g_4 B \sin(\theta + \beta)$ in (8.0.2) by $g_{31} g + g_{32} \dot{g}$ and $g_{41} g + g_{42} \dot{g}$, respectively, to guarantee more flexibility by calibrating this system to original results, see Section 8.2.

Thus, we introduced a low-order dynamical system of two coupled oscillators describing the observed behavior of the natural and the actuated flows as well as a transient behavior between them, according to the desired properties $(i) - (iv)$:

$$\begin{aligned}
\dot{a}_1 &= \tilde{\sigma}^n a_1 - \omega^n a_2 \\
\dot{a}_2 &= \omega^n a_1 + \tilde{\sigma}^n a_2 \\
\dot{a}_3 &= \tilde{\sigma}^a a_3 - \omega^a a_4 + g_{31} g + g_{32} \dot{g} \\
\dot{a}_4 &= \omega^a a_3 + \tilde{\sigma}^a a_4 + g_{41} g + g_{42} \dot{g} \\
\tilde{\sigma}^n &= \sigma^n - \beta_1 (A^n)^2 - \beta_2 (A^a)^2 \\
\tilde{\sigma}^a &= \sigma^a
\end{aligned} \tag{8.0.4}$$

with $A^n = \sqrt{a_1^2 + a_2^2}$, $A^a = \sqrt{a_3^2 + a_4^2}$. The a_i's, $i = 1, \dots, 4$ can be interpreted as the coefficients to the in Chapter 7 calculated POD modes. With the steady base flow u_0, this leads to

$$u(x,t) \approx u_0(x) + \sum_{i=1}^{4} a_i(t) u_i(x).$$

If u is a velocity field without actuation, a_3 and a_4 are almost equal to zero. Thus, u is approximated by the modes of the natural case $u(x,t) \approx u_0(x) + a_1(t)u_1(x) + a_2(t)u_2(x)$. Otherwise, u is described by modes of the actuated case and a_1 and a_2 are near to zero. One can see in the system (8.0.4) that more actuation, e.g. for instance a higher actuation amplitude B, leads to a higher weighting of the modes of the actuated case in contrast to the ones of the natural case.

Here, we see a low-order model with the control function $g(t)$. This model replace the Navier-Stokes equations in our reduced optimization problem, see Chapter 9.

After a more detailed discussion for the reduced-order model (ROM) in Section 8.1, adopted from [66], we explain in the following sections our modifications on this model (Section 8.2), introduce a lift-formula only based on the mode coefficients $\{a_i\}_{i=1}^4$ (Section 8.3) and give a numerical example to demonstrate that the ROM reproduces the nonlinear behavior of the system sufficiently well for our optimization ansatz (Section 8.4).

8.1 A generalized model

This section, where we consider the structure of the dynamical system (8.0.4) more detailed, is based on [66, Section 4 and 5].

8.1.1 Mean-field theory

We consider a computational domain $\Omega \subset \mathbb{R}^2$ with $x = (x, y) \in \Omega$. The x-axis is aligned with the flow and the y-axis with the orthogonal direction. The velocity field is denoted by $u = (u, v)$, where the components u and v are aligned with the x- and y-direction. This model will demonstrate the role of mean-field dynamics in stabilizing an attractor and as a commitment between the actuated and the natural (unactuated) case. An attractor is a set towards which a dynamical system evolves over time[1]. For the mean-field

[1] Wikipedia

theory, we refer to [74, 75, 94, 95]. In addition to the L^2-scalar product of vector fields

$$(f, g)_\Omega = \int_\Omega f \cdot g \, \mathrm{d}x,$$

let us define inner product for matrix-valued fields by

$$(A, B)_\Omega := \int\limits_\Omega A : B \, \mathrm{d}x, \text{ with } A : B := \sum_{i,j=1}^{2} A_{ij} B_{ji}$$

and the instantaneous kinetic energy of a velocity field u is given by

$$K := \frac{\|u\|_\Omega^2}{2}.$$

We consider the incompressible

$$\operatorname{div} u = 0$$

non-stationary Navier-Stokes equation

$$u_t - \nu \Delta u + (u \cdot \nabla) u + \nabla p = 0 \tag{8.1.1}$$

and the unsteady boundary condition

$$u = g$$

with boundary actuation g. Additionally to the boundary actuation, we consider a time periodic and space dependent volume force g^a. We denote the so-called ensemble average by $\overline{u}$ with its approximation

$$\overline{u}(t) := \frac{1}{T} \int_{-T/2}^{T/2} u(t + \tau) \, \mathrm{d}\tau,$$

where $T > 0$ is a set length of a time window. Next, we are formulating some assumptions. The first one is based on observed phenomenology.

(A.1) (A generalized Krylov-Bogoliubov ansatz) The velocity field u is dominated by the sum of a slowly varying base flow u^b and two oscillatory components which are nearly pure harmonics at the natural u^n and the actuation u^a frequency. Other temporal harmonics are considered as negligible. Thus, we obtain

$$u(x, t) = u^b(x, t) + u^n(x, t) + u^a(x, t), \tag{8.1.2}$$

where u^b satisfies the steady, inhomogeneous boundary condition $\overline{g}(t)$, u^n the homogenized version and u^a accounts for the residual to the unsteady boundary condition $g(t) - \overline{g}(t)$.

Due to the fact that the base flow is almost time-independent, we assume

$$\overline{u}^b = u^b.$$

Furthermore, we recognize that the ensemble averages of the fluctuations u^n and u^a vanishes

$$\overline{u}^n = 0, \ \overline{u}^a = 0$$

where u^b, u^n and u^a are averaged over the associated time intervals.

The ansatz of Dušek et. al. [38] was to establish a small parameter $\epsilon \ll 1$ and slowly varying amplitude functions u_0^b, u_i^n and u_i^a, $i = 1, 2$, such that

$$\begin{aligned} u^b(x,t) &= u_0^b(x, \epsilon t), \\ u^n(x,t) &= u_1^n(x, \epsilon t)\cos(\Omega^n t) + u_2^n(x, \epsilon t)\sin(\Omega^n t), \\ u^a(x,t) &= u_1^a(x, \epsilon t)\cos(\Omega^a t) + u_2^a(x, \epsilon t)\sin(\Omega^a t), \end{aligned}$$

expresses the assumed slow variation of the mean flow, the oscillation amplitudes, the frequencies and the phase shifts.

This ansatz implies that time derivatives of the amplitude functions are of order $O(\epsilon)$, which we want to neglect.

The second assumption is called by engineers 'a non-commensurability ansatz':

(A.2) (A non-commensurability ansatz) There is no direct interaction between u^n and u^a through the nonlinear term $(u\cdot\nabla)u$, i.e. $(u^n\cdot\nabla)u^a = (u^a \cdot \nabla)u^n = 0$.

This assumption is based on the numerically observed fact that the activity regions of these fluctuations rarely overlap. So, on each of the two attractors, we neglect fluctuations in the other frequency.

Substituting the Assumption **(A.1)** into the Navier-Stokes equations (8.1.1) and re-arranging the terms by the zeroth and the first harmonics at frequencies St^n and St^a, respectively, leads to

$$0 = -\nu\Delta u^b + (u^b\cdot\nabla)u^b + \overline{(u^n\cdot\nabla)u^n} + \overline{(u^a\cdot\nabla)u^a} + \nabla p^b, \qquad (8.1.3)$$

$$u_t^n = \nu\Delta u^n - (u^n\cdot\nabla)u^b - (u^b\cdot\nabla)u^n - \nabla p^n, \qquad (8.1.4)$$

$$u_t^a = \nu\Delta u^a - (u^a\cdot\nabla)u^b - (u^b\cdot\nabla)u^a - \nabla p^a + g^a. \qquad (8.1.5)$$

The temporal behavior of the terms $(u^k\cdot\nabla)u^k$, $k = b, n, a$, is specified by the $0th$ and second harmonics of the frequencies St^n and St^a and they are eliminated in (8.1.4) and (8.1.5) by Assumption **(A.1)**. The mean-field model in the next subsection is based on the system (8.1.3)-(8.1.5).

For a homogeneous fluid and an incompressible flow, the flow velocities $u(x,t)$, having components u_i in the x_i coordinate direction, can be splitted into a mean part $\overline{u}(x)$ and a fluctuating part $u'(x,t)$ using the so-called Reynolds decomposition:

$$u_i = \overline{u_i} + u_i'.$$

The Reynolds stress tensor τ' is then defined by $\tau_{ij}' = \overline{u_i' u_j'}$ and describes the degree of nonlinearity. The term $\nabla \tau'$ denotes the force pushing the mean flow away from the steady flow. If the Reynolds stress would be zero, then the mean flow would coincide with the steady flow.

The next and last assumption guarantees a linear relation between the Reynolds stresses τ' and the mean-field correction term u^h. Let therefore, u_s be the associating solution of the steady Navier-Stokes equations

$$\begin{aligned} -\Delta u_s + (u_s \cdot \nabla) u_s + \nabla p &= 0 && \text{in } \Omega \\ \operatorname{div} u_s &= 0 && \text{in } \Omega \\ u_s &= \overline{g}(t) && \text{on } \Gamma. \end{aligned} \tag{8.1.6}$$

(A.3) (Linearized Reynolds equation) Let us assume that (8.1.3) is linearizable at the steady solution u_s and

$$u^b = u_s + u^h. \tag{8.1.7}$$

The linearized Reynolds equation for the mean-field correction u^h is obtained by substituting (8.1.7) in (8.1.3)

$$\begin{aligned} 0 = -\nu\Delta(u_s + u^h) &+ ((u_s + u^h) \cdot \nabla)(u_s + u^h) \\ &+ \overline{(u^n \cdot \nabla) u^n} + \overline{(u^a \cdot \nabla) u^a} + \nabla(p_s + p^h) \end{aligned}$$

subtracting the steady Navier-Stokes equation

$$\begin{aligned} 0 = -\nu\Delta u^h &+ (u_s \cdot \nabla) u^h + (u^h \cdot \nabla) u_s + \nabla p^h \\ &- \nu \Delta u_s + (u_s \cdot \nabla) u^s + \nabla p_s - (-\nu\Delta u_s + (u_s \cdot \nabla) u^s + \nabla p_s) \\ &+ \overline{(u^n \cdot \nabla) u^n} + \overline{(u^a \cdot \nabla) u^a} + (u^h \cdot \nabla) u^h \end{aligned}$$

and neglecting quadratic terms in u^h:

$$\begin{aligned} 0 = -\nu\Delta u^h &+ (u_s \cdot \nabla) u^h + (u^h \cdot \nabla) u_s \\ &+ \overline{(u^n \cdot \nabla) u^n} + \overline{(u^a \cdot \nabla) u^a} + \nabla p^h. \end{aligned} \tag{8.1.8}$$

8.1.2 Mean-field Galerkin model

Let us now transform the mean-field model over to the least-order Galerkin model. For the theory of the Galerkin method, we refer to to Fletcher [41], Holmes et. al. [56] and Ladyzhenskaya [65]. Let u_0 denote the steady base flow, then the Galerkin method is based on the Galerkin expansion

$$u(x,t) = \sum_{i=0}^{M} a_i(t) u_i(x). \tag{8.1.9}$$

The time dependency is described by the Fourier coefficients a_i, $i = 1, \ldots, M$ with $a_0 = 1$. We describe the Galerkin approximation in the section *Least-order Galerkin approximation* after next.

Galerkin method

Replacing the flow u in the Navier-Stokes equations (8.1.1) with $g_a = 0$ by the Galerkin expansion (8.1.9) and projecting them onto the subspace spanned by the expansion modes:

$$(u_t - \nu \Delta u + (u \cdot \nabla) u + \nabla p, u_i)_\Omega = 0 \text{ for } i = 1, \ldots, M,$$

we obtain with $\Gamma := \partial\Omega$

1.

$$\begin{aligned}\left(\frac{\partial}{\partial t}\left(\sum_{j=0}^{M} a_j(t) u_j(x)\right), u_i(x)\right)_\Omega &= \left(\sum_{j=0}^{M} \frac{\partial a_j}{\partial t} u_j, u_i\right)_\Omega \\ &= \sum_{j=1}^{M} \left(\frac{\partial a_j}{\partial t} u_j, u_i\right)_\Omega = \sum_{j=1}^{M} \frac{\partial a_j}{\partial t} (u_j, u_i)_\Omega = \sum_{j=1}^{M} \frac{\partial a_j}{\partial t} \delta_{ij} = \frac{\partial a_i}{\partial t},\end{aligned}$$

2.

$$\begin{aligned}\nu\left(\Delta\left(\sum_{j=0}^{M} a_j(t) u_j(x)\right), u_i(x)\right)_\Omega &= \nu \sum_{j=0}^{M} (\Delta(a_j u_j), u_i)_\Omega \\ &= \nu \sum_{j=0}^{M} a_j \underbrace{(\Delta u_j, u_i)_\Omega}_{=:l_{ij}} = \nu \sum_{j=0}^{M} l_{ij} a_j,\end{aligned}$$

3.
$$\left(\left(\left(\sum_{j=0}^{M} a_j(t)u_j(x)\right)\cdot\nabla\right)\left(\sum_{k=0}^{M} a_k(t)u_k(x)\right), u_i(x)\right)_\Omega$$
$$= \sum_{j=0}^{M}\sum_{k=0}^{M} (((a_j u_j)\cdot\nabla)(a_k u_k), u_i)_\Omega$$
$$= \sum_{j=0}^{M}\sum_{k=0}^{M} a_j a_k \underbrace{((u_j\cdot\nabla)u_k, u_i)_\Omega}_{=:q^u_{ijk}} = \sum_{j,k=0}^{M} a_j a_k q^u_{ijk},$$

4. By Noack et. al. [76], we obtain a pressure expansion $\nabla p(x,t) = p^{[0\ldots M]}(x,t) = \sum\limits_{j=0}^{M}\sum\limits_{k=0}^{M} p_{jk}(x)a_j(t)a_k(t)$. This leads to

$$(\nabla p(x,t), u_i(x))_\Omega = \left(p^{[0\ldots M]}(x,t), u_i(x)\right)_\Gamma$$
$$= \left(\sum_{j=0}^{M}\sum_{k=0}^{M} p_{jk}(x)a_j(t)a_k(t), u_i(x)\right)_\Gamma = \sum_{j=0}^{M}\sum_{k=0}^{M} a_j a_k \underbrace{(p_{jk}, u_i)_\Gamma}_{=:q^p_{ijk}}$$
$$= \sum_{j,k=0}^{M} q^p_{ijk} a_j a_k.$$

Summarized, this leads with $q_{ijk} := q^u_{ijk} + q^p_{ijk}$ to a simplified ordinary differential equation system

$$\frac{\partial a_i}{\partial t} = \nu \sum_{j=0}^{M} l_{ij} a_j + \sum_{j,k=0}^{M} q_{ijk} a_j a_k$$

$i = 1, ..., M$, to define the Fourier coefficients $a_i(t)$. Together with

$$(\mathcal{M}(a,a))_i := \sum_{j,k=1}^{M} q_{ijk} a_j a_k,\ (\mathcal{F}(a))_i := \sum_{j=1}^{M} (\nu l_{ij} + q_{i0j} + q_{ij0}) a_j,$$
$$(\mathcal{C})_i := \nu l_{i0} + q_{i00}$$

and $a_0 = 1$, we obtain the system

$$\frac{\partial a}{\partial t} = \mathcal{M}(a,a) + \mathcal{F}(a) + \mathcal{C}. \tag{8.1.10}$$

We have the following difficulty: by a standard POD method, boundary values is prescribed by the dynamical system and can not be installed as a actuation command. The problem is that boundary actuation is generally not derivable from the Galerkin projection, because the Galerkin projection is composed to ignore boundary perturbations defined over a set of measure zero. But it is possible to simulate boundary effects by additional actuation modes, see for instance [16, 49, 77].

Instead, we decided to add an control term to the system (8.1.10) that is state dependent and includes the control command and its time derivative. We consider an oscillatory control with slowly varying periodic characteristics.

Then the influence on the flow can be modelled by an actuation term $\mathcal{B}\vec{g}(t)$ with actuation command $\vec{g}(t)$ and a matrix $\mathcal{B}$, see [85, 86, 92]. As mentioned in (7.0.3), we consider a periodic actuation $g(t) = B\cos(\beta(t))$, fulfilling $\partial\beta(t)/\partial t = -\Omega^a$ with an actuation amplitude B and a phase shift of $\beta(t) - \Omega^a t$. The acceleration is $\partial g(t)/\partial t = -\Omega^a B\sin(\beta(t))$.

We combine the actuation command and its derivative with respect to t to

$$\vec{g} = [g,\ \dot{g}]^T = B[\cos(\beta),\ -\Omega^a sin(\beta)]^T$$

and obtain the Galerkin system with actuation

$$\frac{\partial a}{\partial t} = \mathcal{M}(a,a) + \mathcal{F}(a) + C + \mathcal{B}\vec{g}. \tag{8.1.11}$$

The term $\mathcal{B}\vec{g}(t)$ replaces the boundary actuator in the Navier-Stokes equations, because we have no boundary terms in the dynamical system (8.1.11).

Considering the system (8.0.4), we have to define $\mathcal{B}$ by

$$\mathcal{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ g_{31} & g_{32} \\ g_{41} & g_{42} \end{bmatrix}.$$

Let us consider in the next section the Galerkin approximation for our problem.

Least-order Galerkin approximation

The least-order Galerkin approximation is based on Assumption **(A.1)**; we are interested in modes resolving u^n, u^a and u^b in (8.1.2). Let therefore u_i^n and u_i^a, $i = 1, 2$, be the two dominant POD modes of the natural and the actuated attractors. The modes u_1^n and u_2^n are orthonormal and u_1^n and u_2^n are orthonormal by construction.

Considering the computed POD mode pairs, see Figure 7.3-7.5, we recognize that the actuated modes u_1^a and u_2^a have their main fluctuations over and near the airfoil, whereas the fluctuations of the natural POD modes u_1^n and u_2^n are concentrated further downstream. This fact and the different wavelengths shows that POD mode pairs are nearly orthogonal. This observations supported by very small values $(u_1^n, u_1^a)_\Omega$, $(u_1^n, u_2^a)_\Omega$, $(u_2^n, u_1^a)_\Omega$ and $(u_2^n, u_2^a)_\Omega$.

Thus, we merge them into an orthonormal basis, using the Gram-Schmidt normalization and obtain (u_1, u_2, u_3, u_4) associated to $(u_1^n, u_2^n, u_1^a, u_2^a)$ and we shall maintain the approximation of the fluctuations u^n and u^a by

$$u^n = \sum_{i=1}^{2} a_i(t) u_i(x), \text{ and } u^a = \sum_{i=3}^{4} a_i(t) u_i(x).$$

Now, we are looking for a representation for u^b. Let u_s be the steady solution of (8.1.6), u_0^n is the mean flow of the natural attractor and $x \propto y$ means that x is proportional to y. Then, following [74], the effect of the Reynolds stress due to the natural oscillations is described by a shift-mode $u_\Delta^n \propto u_0^n - u_s,$. Analogously, we define u_Δ^a.

Assume that u_5 and u_6 are obtained by a Gram-Schmidt orthogonalization from u_Δ^n and u_Δ^a, removing any projections over $u_1, \cdots, u_4$. Thus, $u_1, \ldots, u_6$ are orthonormal. In [74], they approximate the time-varying base flow $u^b(x,t) = u_s(x) + u^h(x,t)$, see Assumption **(A.3)**, with the two shift-modes u_5 and u_6 corresponding to the two attractors of the natural and the actuated case

$$u^b(x,t) = u_s(x) + u^h(x,t) = u_s(x) + a_5(t)u_5(x) + a_6(t)u_6(x), \qquad (8.1.12)$$

This means that the fluctuations $u_1, \ldots, u_4$ are negligible to approximate the mean-field correction u^h and the base flow u^b.

A transient describes the crossing behavior into another attractor without reaching him. Then, the mean flows u_0^n and u_0^a are approximated by the associated initial and final values of the base flow u^b trajectory in transients connecting the two attractors and we obtain with $u_\Delta := (u_0^n - u_0^a)/\|u_0^n - u_0^a\|_\Omega$ the approximation

$$u^b(x,t) = u_0^n(x) + a_\Delta(t) u_\Delta(x). \qquad (8.1.13)$$

See Figure 8.1 for a visual description of the relation between the steady solution u_s and the mean flows u_0^n and u_0^a.

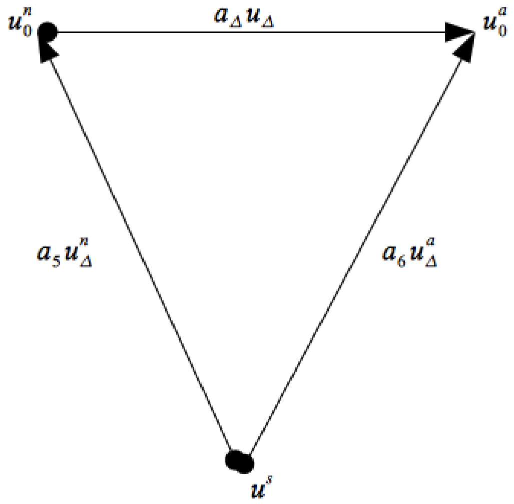

Figure 8.1: The relation between the steady solution u_s and the mean flows u_0^n and u_0^a, see [66] Figure 5.

Considering (8.1.13) instead of (8.1.12), the velocity field of the URANS data was approximated by

$$u(x,t) = u_0^n(x) + a_\Delta(t)u_\Delta(x) + \sum_{i=1}^{4} a_i(t)u_i(x). \tag{8.1.14}$$

Thus, we do not have to extract a steady solution u_s.

In (8.1.14), we approximated the system of the natural flow, the actuated flow and the states between them by the two mean flows u_0^n and u_0^a, the associated shift mode u_Δ with Fourier coefficient a_Δ, as transition between them, the two modes u_1, u_2 of the natural and the two modes u_3, u_4 of the actuated flow, with associated Fourier coefficients $a_1, ..., a_4$.

However, in the next section, we use (8.1.12) to obtain the equivalent Galerkin expansion

$$u(x,t) = u_s(x) + \sum_{i=1}^{6} a_i(t)u_i(x). \tag{8.1.15}$$

This approximation will be used in the Galerkin system (8.1.16)-(8.1.18).

The main advantages of (8.1.15) in contrast to (8.1.14) appear by calibrating the Galerkin system parameter with empirical data, see [66] for more details. The advantage of (8.1.14) was that we do not have to calculate the steady solution u_s.

Least-order Galerkin system

Inserting the equation (8.1.15) into the mean-field Navier-Stokes equations (8.1.4), (8.1.5) and (8.1.8), followed by the Galerkin projection of these equations on the expansion modes lead to the least-order Galerkin system, consisting of the Fourier coefficients a_i. Subsequently, we have to enforce a Galerkin projection of these equations on the expansion modes $\{u_i\}_{i=1}^6$.

The Galerkin system, associated to the projection of (8.1.4), is

$$\underbrace{\frac{\partial a_i}{\partial t}}_{\widehat{=}\ \partial_t u^n} = \underbrace{\sum_{j=1}^{2}\sum_{k=5}^{6} q_{ijk}a_j a_k}_{\widehat{=}\ -(u^n\cdot\nabla)u^b} + \underbrace{\sum_{j=5}^{6}\sum_{k=1}^{2} q_{ijk}a_j a_k}_{\widehat{=}\ -(u^b\cdot\nabla)u^n} + \underbrace{\nu\sum_{j=1}^{2} l_{ij}a_j}_{\widehat{=}\ \nu\Delta u^n} + \underbrace{0}_{\widehat{=}\ \nabla p^n},$$

$i \in \{1, 2\}$. For a proof, see the derivation of(8.1.10). The other equations follow analogously.

Let e_i be the unit-vector of $\mathbb{R}^M$, then we get with

$$a^n = a_1e_1 + a_2 + e_2,\ a^a = a_3e_3 + a_4e_4,\ \text{and}\ a^b = a_5e_5 + a_6e_6$$

the full Galerkin system by

$$\frac{\partial a^n}{\partial t} = \mathcal{M}(a^b, a^n) + \mathcal{M}(a^n, a^b) + \mathcal{F}(a^n), \tag{8.1.16}$$

$$\frac{\partial a^a}{\partial t} = \mathcal{M}(a^b, a^a) + \mathcal{M}(a^a, a^b) + \mathcal{F}(a^a) + \mathcal{B}\vec{g}, \tag{8.1.17}$$

$$0 = \overline{\mathcal{M}(a^n, a^n)} + \overline{\mathcal{M}(a^a, a^a)} + \mathcal{F}(a^b). \tag{8.1.18}$$

In [66, Appendix B], it is shown that the Galerkin system has the following equivalent form

$$\frac{\partial}{\partial t}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} \tilde{\sigma}^n & -\tilde{\omega}^n & 0 & 0 \\ \tilde{\omega}^n & \tilde{\sigma}^n & 0 & 0 \\ 0 & 0 & \tilde{\sigma}^a & -\tilde{\omega}^a \\ 0 & 0 & \tilde{\omega}^a & \tilde{\sigma}^a \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \kappa & -\eta \\ \eta & \kappa \end{bmatrix}\vec{g}, \tag{8.1.19}$$

with

$$\begin{aligned} \tilde{\sigma}^n &= \sigma^n - \sigma^{n,n}(A^n)^2 - \sigma^{n,a}(A^a)^2, \\ \tilde{\omega}^n &= \omega^n + \omega^{n,n}(A^n)^2 + \omega^{n,a}(A^a)^2, \\ \tilde{\sigma}^a &= \sigma^a - \sigma^{a,n}(A^n)^2 - \sigma^{a,a}(A^a)^2, \\ \tilde{\omega}^a &= \omega^a + \omega^{a,n}(A^n)^2 + \omega^{a,a}(A^a)^2, \end{aligned} \tag{8.1.20}$$

where $A^n = \|a^n\|$, $A^a = \|a^a\|$ are the respective oscillation amplitudes. The Assumption **(A.3)** avoids a Taylor series in $(A^n)^2$ and $(A^a)^2$.

The parameter $\sigma^{n,n}$ describes the decreasing growth rate of the natural attractor with increasing energy in the natural attractor and $\sigma^{n,a}$ the decreasing growth rate of the natural attractor with increasing energy in the actuated attractor. The parameter $\sigma^{a,n}$ and $\sigma^{a,a}$ can be interpreted analogously.

The parameters $\omega^{n,n}, \omega^{n,a}, \omega^{a,n}$ and $\omega^{a,a}$ describe the changes of the amplitude-dependent frequency. The equation (8.1.18) leads to a linear dependence of a^b on a^n and a^a. Consequences of this dependency are that the system (8.1.19) only consists of $a_1, \dots, a_4$ and the equations (8.1.20), for more details see [66, Appendix B].

Let us have a closer look at (8.1.19) and (8.1.20) to identify some parameters. A first ansatz is that we assume for reasons of simplifications that the oscillation frequencies are independent of A^n and A^a and a constant frequency for the natural and the actuated flow, which yields

$$\omega^{n,n} = \omega^{n,a} = \omega^{a,a} = \omega^{a,n} = 0$$

and consequently

$$\begin{aligned} \tilde{\omega}^n &= \omega^n = \Omega^n, \\ \tilde{\omega}^a &= \omega^a = \Omega^a, \end{aligned}$$

see [66, Section 5.4] for more details. In [66, Section 5.5] is the simplification

$$\tilde{\sigma}^a = \sigma^a,$$

described that means $\sigma^{a,n} = \sigma^{a,a} = 0$.

Finally, this leads to the following dynamical system

$$\begin{aligned} \dot{a}_1 &= \tilde{\sigma}^n a_1 - \Omega^n a_2 \\ \dot{a}_2 &= \Omega^n a_1 + \tilde{\sigma}^n a_2 \\ \dot{a}_3 &= \tilde{\sigma}^a a_3 - \Omega^a a_4 + \kappa g - \eta \dot{g} \\ \dot{a}_4 &= \Omega^a a_3 + \tilde{\sigma}^a a_4 + \eta g + \kappa \dot{g} \\ \tilde{\sigma}^n &= \sigma^n - \sigma^{n,n}(A^n)^2 - \sigma^{n,a}(A^a)^2 \\ \tilde{\sigma}^a &= \sigma^a \end{aligned} \tag{8.1.21}$$

and one has to calibrate the remaining parameters to the numerical data, i.e. the preliminarily calculated snapshots.

Similar to the least-order Galerkin model (8.1.19) and (8.1.20), we want to establish a new reduced-order model (ROM) calibrated with our numerical data. In the next section, we present our modifications.

8.2 Modifications on the reduced-order model

In this section, we want to demonstrate our modifications and calibrations on the reduced-order model to handle the optimization problem (P_N) in the next chapter.

Similarly to POD, see Section 7.2, all snapshots are processed. For this purpose, we consider only the velocity field $u = (u, v)$ in a certain reference domain, where the actuation has the main influence on velocity and lift, see Figure 7.1. Therefore, the snapshot velocity data are weighted by the size of their area. We select the first two POD modes of the actuated (u_1^a, u_2^a) and non-actuated (u_1^n, u_2^n) system, carrying the highest energy.

But in contrast to the ansatz in the last Section 8.1; we filter the POD modes and the associated mode coefficients by eliminating certain fluctuations to emphasize the dominant harmonic structure, see the following subsection for more details.

8.2.1 Filtering of the POD modes and coefficients

The snapshots $\hat{u}_i^n$ and $\hat{u}_i^a$ are given on a time interval $[0, T]$. Due to the fact that the natural flow and the actuated flow have different wavelengths, we search for the maximal $k^a, k^n \in \mathbb{N}$ such that the times

$$T^n = 2\pi k^n/\Omega^n \text{ and } T^a = 2\pi k^a/\Omega^a$$

fulfill $T^n < T$ and $T^a < T$. Therefore, the time T has to be big enough such that $T > 2\pi/\Omega^n$ and $T > 2\pi/\Omega^a$.

Let $(a_1^n(t), a_2^n(t))$ and $(a_1^a(t), a_2^a(t))$ are the first POD mode coefficient pairs of the natural and the actuated case, respectively. We recognize that they are very similar to trigonometric functions or oscillations and so, we want to approximate them as well as possible by oscillations. Therefore, we calculate the phases $\varphi^n(t)$, $\varphi^a(t)$ and radii $\tilde{r}^n(t)$, $\tilde{r}^a(t)$ by

$$\begin{aligned} a_1^n(t) + ia_2^n(t) &= \tilde{r}^n(t)e^{i\varphi^n(t)}, \\ a_1^a(t) + ia_2^a(t) &= \tilde{r}^a(t)e^{i\varphi^a(t)}. \end{aligned}$$

To extract the dominant harmonic oscillations from these POD coefficients, we smooth in (8.2.1) perturbations of both, the radii $\tilde{r}^n(t), \tilde{r}^a(t)$ and the phases $\varphi^n(t), \varphi^a(t)$, in anticipation of a small deformation of the corresponding modes. This holds, because the POD method does not extract pure frequencies and radii, but 'deformed' modes with the highest energy.

With the average values

$$\begin{aligned} r^n &= \overline{\tilde{r}^n}(t) := 1/T^n \int_{-T^n/2}^{T^n/2} \tilde{r}^n(t)\,\mathrm{d}t, & r^a &= \overline{\tilde{r}^a}(t), \\ \omega^n &= \overline{\partial_t \varphi^n}(t), & \omega^a &= \overline{\partial_t \varphi^a}(t), \end{aligned} \tag{8.2.1}$$

we approximate our *filtered coefficients*

$$\begin{aligned} a_1(t) + ia_2(t) &= r^n e^{i\omega^n t}, \\ a_3(t) + ia_4(t) &= r^a e^{i\omega^a t}. \end{aligned}$$

Let u_0^n and u_0^a are the mean flows

$$u_0^n(x) = 1/N \sum_{i=1}^{N} \hat{u}_i^n(x)$$

in the unactuated and

$$u_0^a(x) = 1/N \sum_{i=1}^{N} \hat{u}_i^a(x)$$

in the actuated case.

Then, the associated *filtered modes* are determined by the Fourier ansatz

$$u_i(x) = \left(\hat{u}^n(x,\cdot) - u_0^n, \frac{1}{r^n} a_i(\cdot)\right)_{T^n} := \frac{1}{T^n} \int\limits_{-T^n/2}^{T^n/2} \left(\hat{u}^n(x,t) - u_0^n(x)\right) \frac{1}{r^n} a_i(t)\, dt,$$

for $i = 1, 2$ and

$$u_i(x) = \left(\hat{u}^a(x,\cdot) - u_0^a, \frac{1}{r^a} a_i(\cdot)\right)_{T^a} = \frac{1}{T^a} \int\limits_{-T^a/2}^{T^a/2} \left(\hat{u}^a(x,t) - u_0^a(x)\right) \frac{1}{r^a} a_i(t)\, dt,$$

for $i = 3, 4$. The coefficients a_1^n and a_2^n are orthonormal and a_1^a and a_2^a are orthonormal by construction. But the coefficients a_i^n are not necessarily orthonormal to a_i^a for $i = 1, 2$. Thus, we finally orthonormalize these frequence-filtered modes u_i, $i = 1, \dots, 4$ by Gram-Schmidt and denote the orthonormal modes for simplicity by u_i, $i = 1, \dots, 4$.

The Figures 8.6 and 8.7 show the associated filtered mode coefficients a_1, $i = 1, \dots, 4$ in contrast to the original ones.

In the Figures 8.2 and 8.3, we present the filtered modes u_1 and u_2 of the natural flow and in 8.4 and 8.5, we present the filtered modes u_3 and u_4 of the actuated flow in comparison to the original POD modes. We see that the filtered modes of the actuated case u_3 and u_4 have their main fluctuations over and near the airfoil, whereas the fluctuations of the filtered modes of the natural case u_1 and u_2 are concentrated further downstream. Because we approximated the mode coefficients a_1, $i = 1, \dots, 4$ to almost pure trigonometric functions, we see that some original modes are smoother than the filtered ones.

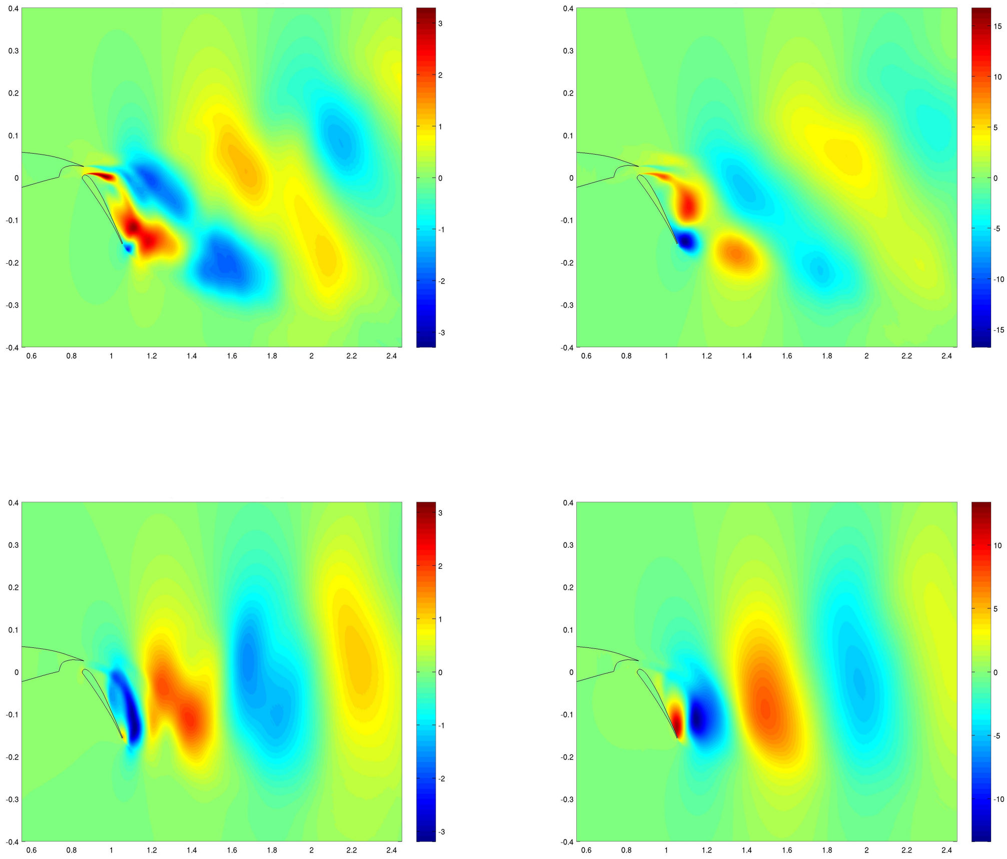

Figure 8.2: The first filtered mode $u_1 = (u_1 \; v_1)^T$ (u_1: top left, v_1: bottom left) and the original first mode $u_1^n = (u_1^n \; v_1^n)$ (u_1^n: top right, v_1^n: bottom right) of the natural case.

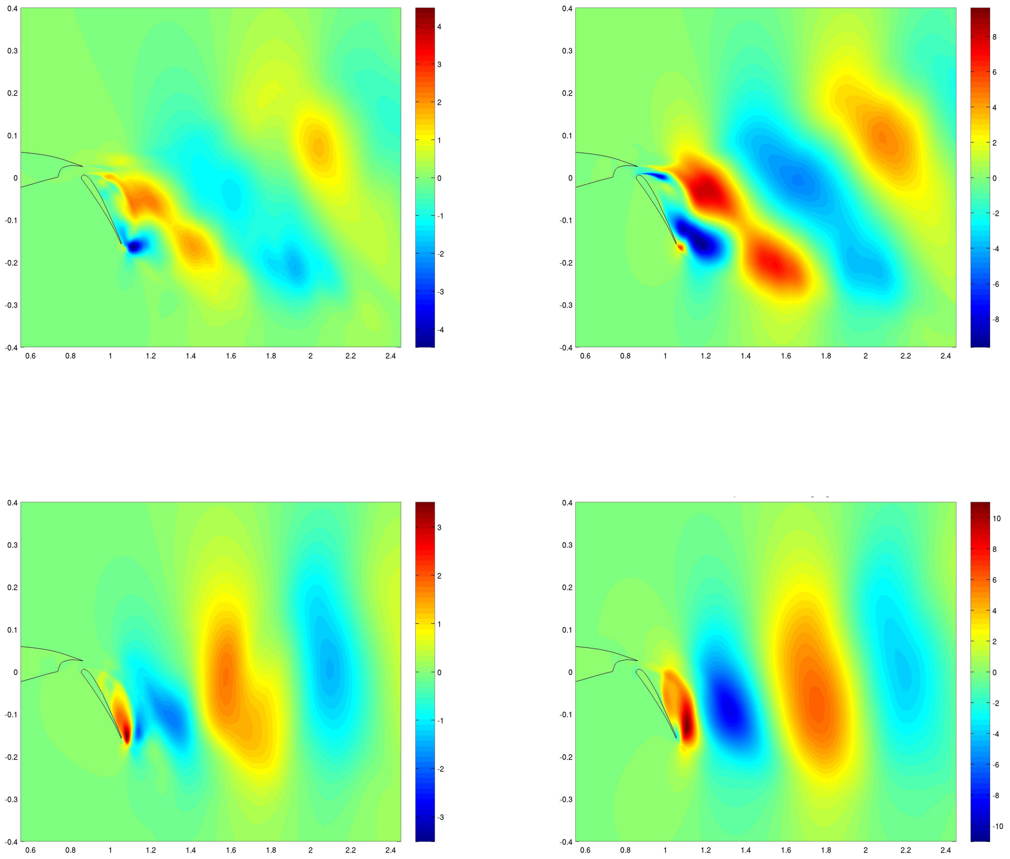

Figure 8.3: The second filtered mode $u_2 = (u_2\ v_2)^T$ (u_2: top left, v_2: bottom left) and the original first mode $u_2^n = (u_2^n\ v_2^n)$ (u_2^n: top right, v_2^n: bottom right) of the natural case.

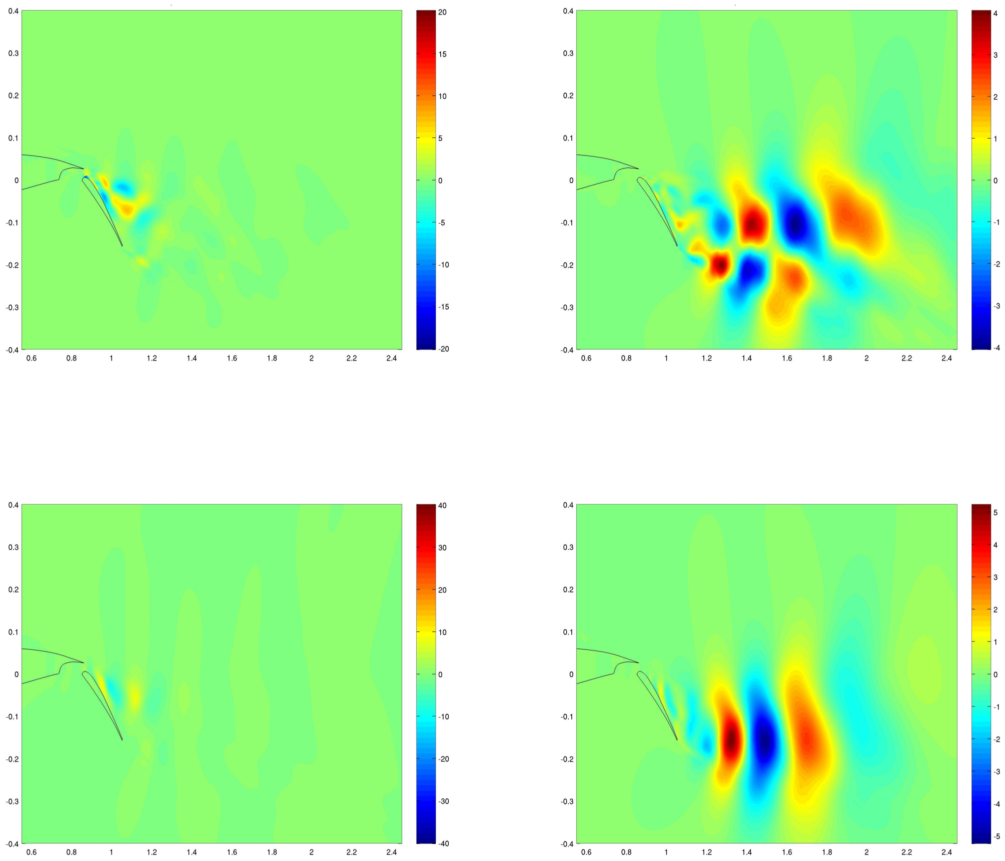

Figure 8.4: The first filtered mode $u_3 = (u_3\ v_3)^T$ (u_3: top left, v_3: bottom left) and the original first mode $u_1^a = (u_1^a\ v_1^a)$ (u_1^a: top right, v_1^a: bottom right) of the actuated case.

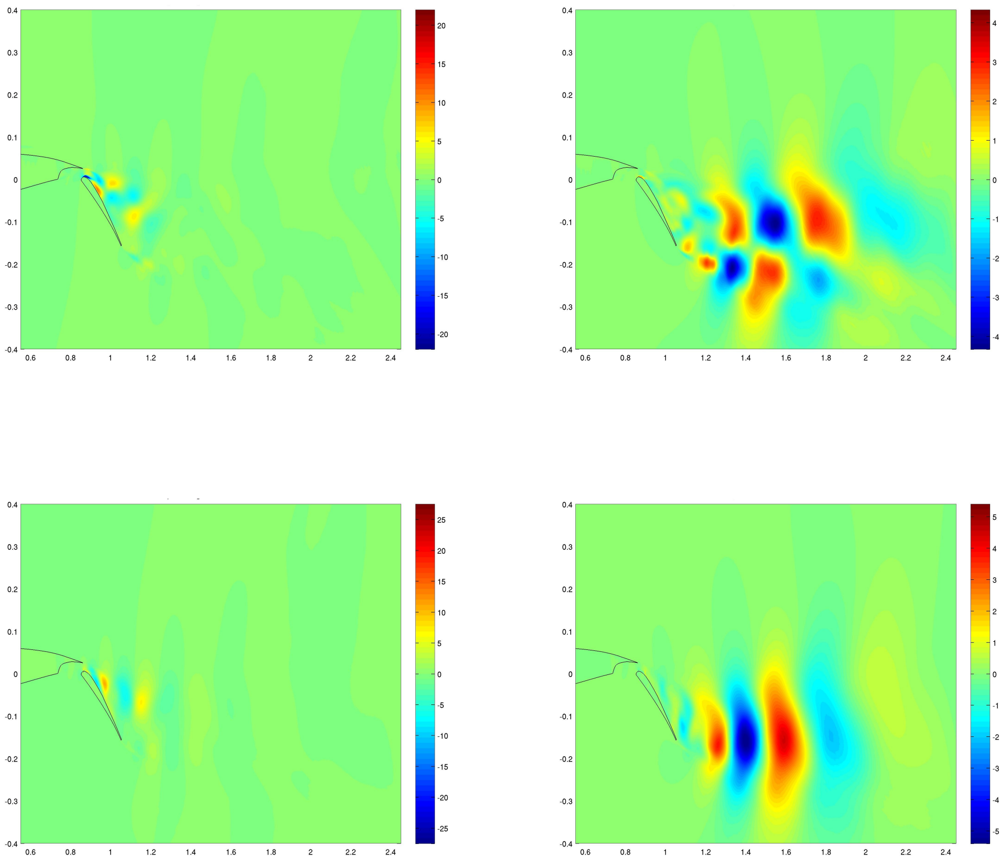

Figure 8.5: The second filtered mode $u_4 = (u_4\ v_4)^T$ (u_4: top left, v_4: bottom left) and the original first mode $u_2^a = (u_2^a\ v_2^a)$ (u_2^a: top right, v_2^a: bottom right) of the actuated case.

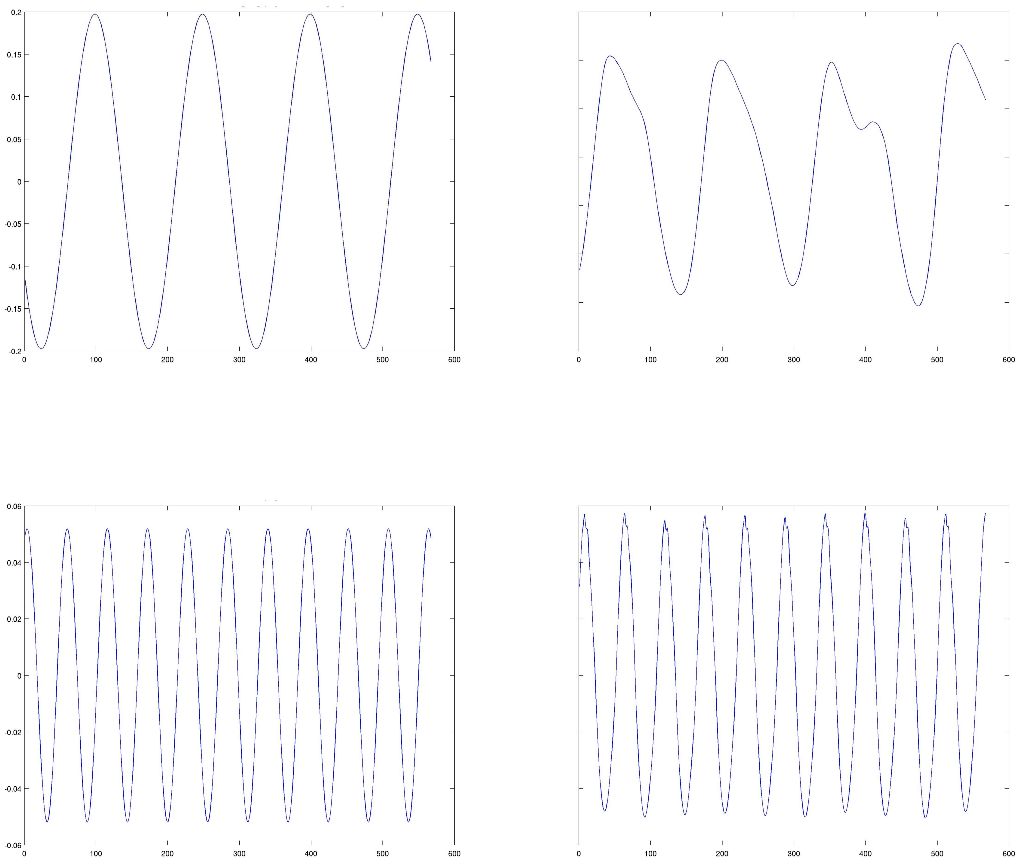

Figure 8.6: The filtered mode coefficients (natural case a_1: top left, actuated case a_3: bottom left) and the original mode coefficients (natural case a_1^n: top right, actuated case a_1^a: bottom right) of the associated first mode over the snapshots.

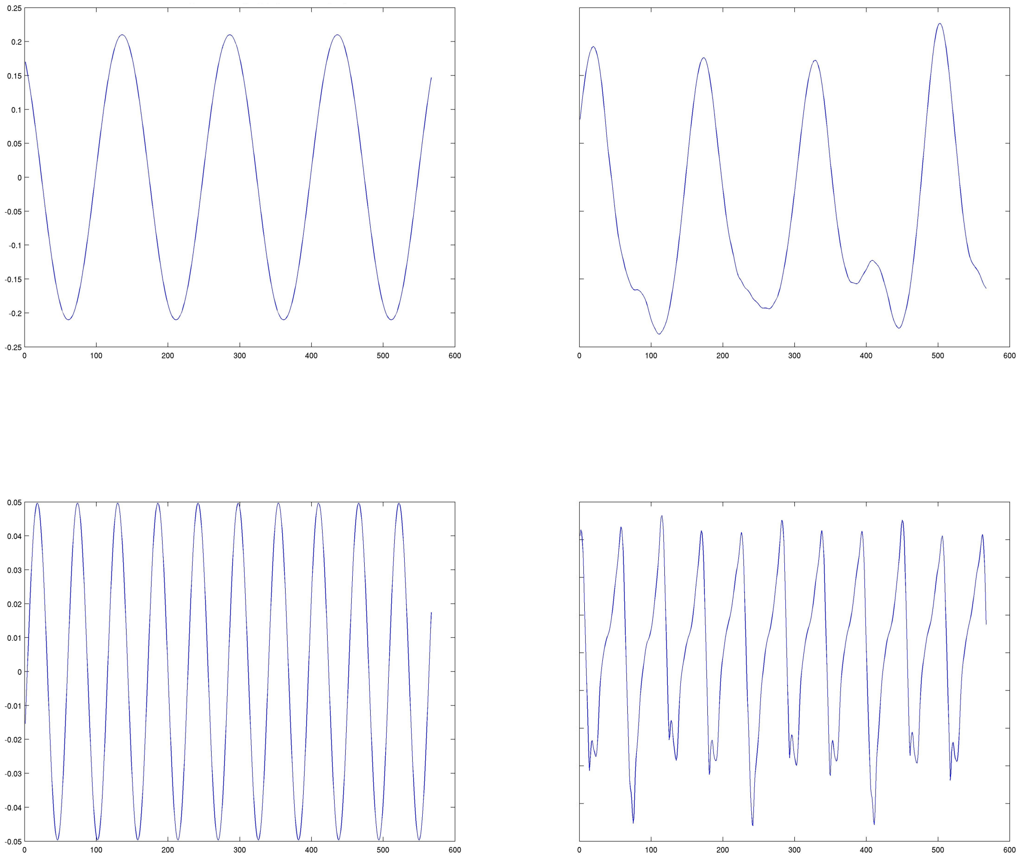

Figure 8.7: The filtered mode coefficients (natural case a_2: top left, actuated case a_4: bottom left) and the original mode coefficients (natural case a_2^n: top right, actuated case a_2^a: bottom right) of the associated second mode over the snapshots.

8.2.2 Parameter calibration

The filtered modes $(u_1, \dots u_4)$ contain significant information gained from the URANS solution by the k-ω-model. Considering the dynamical system (8.1.21), we redefined $\beta_1 := \sigma^{n,n}$ and $\beta_2 := \sigma^{n,a}$ and decided to replace κ and ν by $g_{31},\ g_{42}$ and $g_{32},\ g_{41}$, respectively, to have more degrees of freedom to calibrate the reduced-order model to the original URANS data.

This leads to the following reduced-order model:

$$\begin{aligned}
\dot{a}_1 &= \tilde{\sigma}^n a_1 - \omega^n a_2\\
\dot{a}_2 &= \omega^n a_1 + \tilde{\sigma}^n a_2\\
\dot{a}_3 &= \tilde{\sigma}^a a_3 - \omega^a a_4 + g_{31} g + g_{32}\dot{g}\\
\dot{a}_4 &= \omega^a a_3 + \tilde{\sigma}^a a_4 + g_{41} g + g_{42}\dot{g}\\
\tilde{\sigma}^n &= \sigma^n - \beta_1 (A^n)^2 - \beta_2 (A^a)^2\\
\tilde{\sigma}^a &= \sigma^a,
\end{aligned} \tag{8.2.2}$$

with $A^n = \sqrt{a_1^2 + a_2^2}$, $A^a = \sqrt{a_3^2 + a_4^2}$ and $g = B\cos(\omega^a t)$, where B is the amplitude of the actuation signal and ω^a is the associated angular frequency.

Our snapshots contain no data or information about the transient behavior, because the snapshots are taken while the natural flow and the actuated one, respectively. Because of this reason, we have to select the amplification rates σ^n, σ^a as follows by empirical values:

- $\sigma^n = 0.15$ is an empirical value, if the cord-length of the wing is 1. Because the flap is the active part of the configuration, we choose $\sigma^n = 0.15\frac{U_\infty}{c_{fl}}$.
- $\sigma^a = -\frac{1}{T_{con}}$ where T_{con} is the time that one vortex needs to pass the flap-length c_{fl}. We read off this value from the snapshots.

Now, we want to calibrate the parameters β_1 and β_2 of the system (8.2.2). They are determining the growth rates of both oscillations. That means that they are responsible for the increasing or decreasing rate of the energies in both oscillations with decreasing or increasing energy in the other one. If the fluid flow is in the unactuated state, then no energy should be in the coefficients a_3, a_4, i.e. $a_3 = a_4 = 0$, hence $A^a = 0$. Moreover, we require that

$$\tilde{\sigma}^n = \sigma^n - \beta_1 (A^n)^2 = 0$$

holds for unactuated flow dynamics. This expresses the fact that there is no additional energy contribution to the natural oscillatory behavior of a_1, a_2. Thus β_1 can be determined by

$$\beta_1 = \sigma^n \frac{1}{(A^n)^2} = \sigma^n \frac{1}{(r^n)^2}.$$

There are many possibilities to calibrate β_2. For instance, we can determine β_2 analogously to β_1 by

$$\beta_2 = \sigma^n \frac{1}{(A^a)^2} = \sigma^n \frac{1}{(r^a)^2}.$$

The problem of this ansatz is that the energy belonging to the natural frequency $A^n_{g^a}$, see (8.0.3), does not vanish completely in the actuated case. That means $A^n_{g^a} > 0$.

Let us assume that the energy in the whole system is constant over all actuation amplitudes B :

$$\sigma^n = \beta_1 (A^n)^2 + \beta_2 (A^a)^2. \tag{8.2.3}$$

Then, another possibility is to determine β_2 by

$$\beta_2 = \frac{\sigma^n - \beta_1 (A^n)^2}{(A^a)^2}.$$

We decided to follow an alternative approach. Let us consider for example the actuation amplitude $\tilde{B}$ with associated flow $\tilde{u}$. Till now, we assumed a constant frequency in both the natural and the actuated flow. To determine the parameters β_1 and β_2, we ignore this. Instead, we assume that every flow consists of a combination of the natural and the actuation frequency. Thus, the energy in this flow $\tilde{u}$ is the sum of the energies associated to the two frequencies. Then, we denote analogously to (8.0.3) by $(A^n_{\tilde{B}})^2$ and $(A^a_{\tilde{B}})^2$ the energies associated to the natural and the actuation frequency.

Let B^a be the actuation amplitude related to the actuated flow u^a. Considering the to the natural and the actuated flow associated energies $(A^n_0)^2$, $(A^a_0)^2$, $(A^n_{B^a})^2$ and $(A^a_{B^a})^2$, we are able to draw the energies over the actuation amplitude B, see Figure 8.8.

Considering Figure 8.8, we only obtain a linear relation without constraints between the energies of the natural ω^n and the actuation frequency ω^a, respectively, over the actuation amplitude B. But, this does not correspond to reality. Therefore, we need an additional set of snapshots with a small actuation amplitude B^* to identify the parameters β_1 and β_2 in the system (8.2.2). We select B^* such that the associated energy belonging to the natural frequency $(A^n_{B^*})^2$ should be significant greater than zero. For this actuation amplitude, we compute the filtered coefficients $\{a^*_i\}^4_{i=1}$ and the associated energies $(A^n_{B^*})^2$ and $(A^a_{B^*})^2$.

We draw the results in Figure 8.9 and assume that $(A^n_{B^a})^2$ is the lower bound for the energy of the natural frequency ω^n and $(A^a)^2$ is the upper bound for the energy associated with the actuation frequency ω^a.

Then, we determine the parameters β_1 and β_2 by assuming that the energy in the whole system remains constant over all amplitudes B (8.2.3) by the system

$$\begin{aligned} \beta_1 (A^n_0)^2 + \beta_2 (A^a_0)^2 &= \sigma^n, \\ \beta_1 (A^n_{B^*})^2 + \beta_2 (A^a_{B^*})^2 &= \sigma^n. \end{aligned} \tag{8.2.4}$$

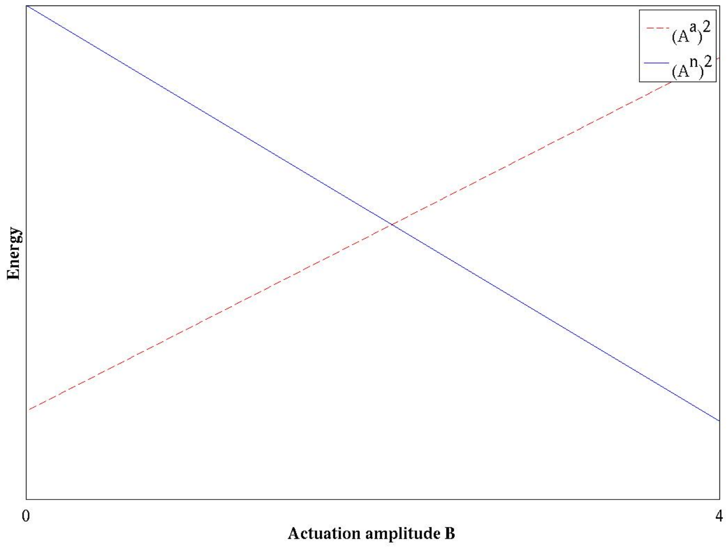

Figure 8.8: Energies $(A^n)^2$ (continuous lines) and $(A^a)^2$ (dashed lines) over the actuation amplitude B. Here, we calculated the energies only for $B = 0$ and the actuation amplitude B^a.

The reduced-order model (8.2.2) contains additionally free parameters g_{31}, g_{32}, g_{41} and g_{42} to calibrate the selected actuation to the dynamical system. Recall therefore that the actuation g and its derivative $\dot{g}$ are $g = B\cos(\omega^a t)$ and $\dot{g} = -B\omega^a \sin(\omega^a t)$, where the actuation amplitude B is our optimization variable.

Therefore, we multiply the third and fourth equation of (8.2.2) by g and $\dot{g}$, respectively, and integrate over $[0, T^a]$. This eliminates g_{32}, g_{42} and g_{31}, g_{41}, respectively. For instance

$$(\dot{a_3}, g)_{T^a} = \sigma_a(a_3, g)_{T^a} - \omega_a(a_4, g)_{T^a} + g_{31}(g, g)_{T^a}$$

leads to

$$g_{31} = \big((\dot{a_3}, g)_{T^a} - \sigma^a(a_3, g)_{T^a} + \omega^a(a_4, g)_{T^a}\big)/(g, g)_{T^a}. \qquad (8.2.5)$$

Note that $(\dot{g}, g)_{T^a}$ vanishes in the long term average.

An example of the phase portraits for the coefficients $a_1, \cdots, a_4$ of the system (8.2.2) is presented in Figure 8.10. This figure presents the solution of (8.2.2) starting on the natural attractor with actuation. We see in this figure on the left side in the phase portrait of the first oscillator (a_1, a_2), describing

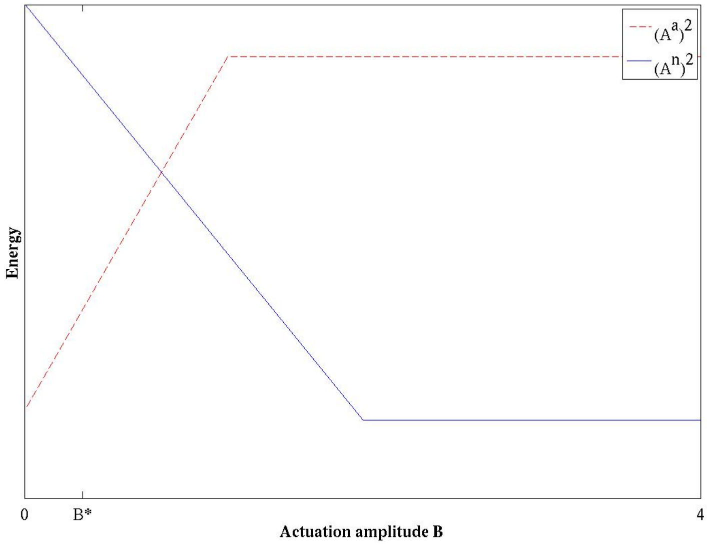

Figure 8.9: Energies $(A^n)^2$ (continuous lines) and $(A^a)^2$ (dashed lines) over the actuation amplitude B. Here, we calculated the energies for $B = 0$, B^* and the actuation amplitude B^a.

the natural flow, that the energy vanishes and transfers over to the phase portrait of the second oscillator (a_3, a_4), describing the actuated flow.

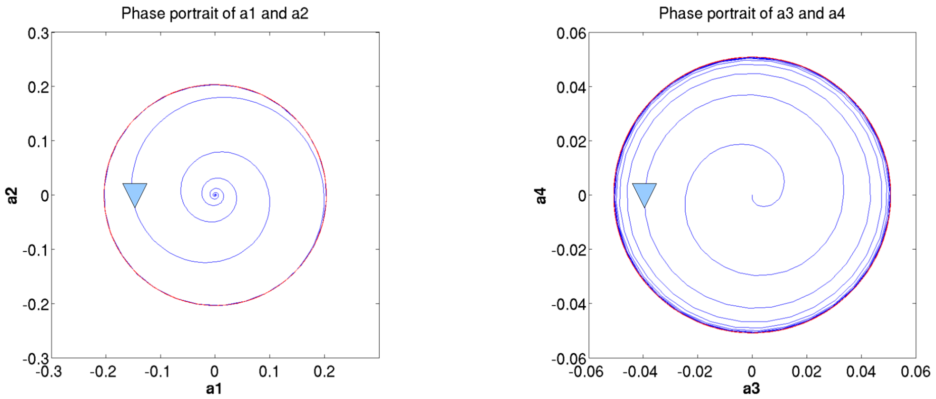

Figure 8.10: Phase portraits of (a_1, a_2) (left) and (a_3, a_4) (right) of the system (8.2.2) with full actuation, starting with (a_1, a_2) on the natural limit cycle and $(a_3, a_4) = (0, 0)$ with actuation.

The next Figure 8.11 presents $(a_1, \ldots, a_4)$ obtained by the dynamical system (8.2.2). We start without actuation and switched on the actuation after 10 seconds. First, we see a_1 and a_2 reaching the natural attractor. After switching on the actuation, a_1 and a_2 decrease to zero and a_3 and a_4 arises to the actuated attractor.

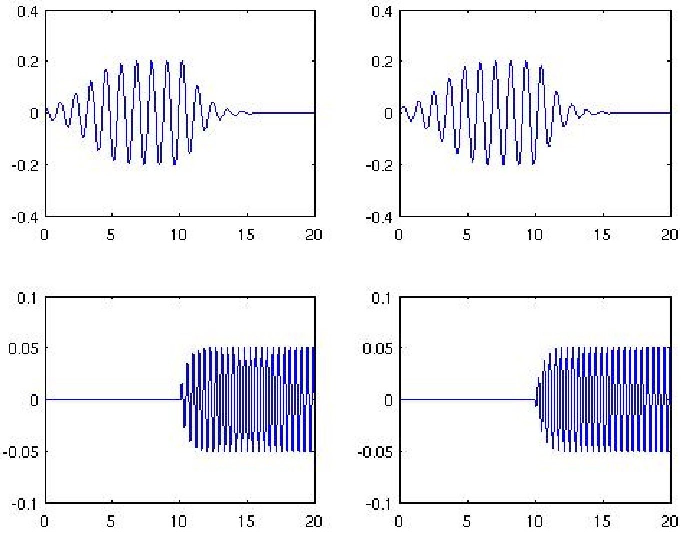

Figure 8.11: The state a (a_1: top left, a_2: top right, a_3: bottom left and a_4: bottom right) gained by the dynamical system (8.2.2), starting without an actuation. After 10 seconds, we switched on the actuation.

This figure shows that the dynamical system 8.2.2 represents the behavior of the natural and the actuated flow as expected.

8.3 Computation of lift

Based on the dynamical system $(\dot{a}_i)_i$ (8.2.2), we calculate the lift by the following ansatz with unknown coefficients c_{ij}

$$C_L(a_1, a_2, a_3, a_4) = c_{00} + \sum_{i=1}^{4} c_{1i} a_i + c_{20}(A^n)^2 + c_{40}(A^n)^4. \qquad (8.3.1)$$

There is no limitation on the energy A^a of the actuated case with respect to increasing B, hence A^a is not included in (8.3.1). The ansatz (8.3.1) is motivated by a global momentum balance equation and the constant and linear terms in (8.3.1) are related to this equation. The lift effect of base-flow variation can be lumped in $c_{20}(A^n)^2+c_{40}(A^n)^4$ assuming slow transients, see [90]. We obtain the parameters c_{ij} by a least-squares fit of $C_L((a_i^n))$ and $C_L((a_i^a))$ to the original lift values of the URANS simulation. The (a_i^n) are the filtered coefficients of the unactuated case and (a_i^a) are the filtered coefficients of the actuated case. Our goal is to fit the parameters in the sense that the simulated lift values are reproduced by the lift formula (8.3.1) in the unactuated case C_{LU} based on (a_i^n) as well as the lift values in the actuated case C_{LA} based on (a_i^a). This leads to the problem

$$\min_{c_{ij}} F(c_{ij}) \quad = \quad \|C_{LA}(\cdot) - C_L(c_{ij})((a_i^a(\cdot)))\|^2 + \|C_{LN}(\cdot) - C_L(c_{ij})((a_i^n(\cdot)))\|^2.$$

Finally, after the coefficients c_{ij} have been determined the optimization problem is formulated as

$$\max_{B>0} C_L(a_1, a_2, a_3, a_4) \tag{8.3.2}$$

subject to the ODE system (8.0.4).

Figure 8.12 presents the original mean lift by the URANS simulation (continuous lines) compared with the calculated mean lift by (8.3.1) (dashed lines).

8.4 Numerical investigation

We consider in this section a high-lift configuration with observation region Ω presented in Figure 7.1, see Setting 7.1 and [66] for more details.

The actuation amplitude to determine the set of the actuated snapshots was $B = 3.5888$ and we worked with the parameters $\sigma^n = 0.5906$, $\sigma^a = -2.0042$, $\omega^n = 5.5407$ and $\omega^a = 14.8412$. Analogously to (8.2.5), we calculated $g_{31} = 0.0284$, $g_{32} = 0.0000$, $g_{41} = 0.0000$ and $g_{42} = -0.0019$. The parameters $\beta_1 = 14.75$ and $\beta_2 = 654.0806$ are calculated by (8.2.4) with an actuation amplitude of $B^* = 1.19$.

Calibrating the parameters c_{ij} of the lift formula (8.3.1) to this data, we get $c_{10} = 2.2238$, $c_{11} = 0.2295$, $c_{12} = -0.6858$, $c_{13} = 1.6717$, $c_{14} = -0.2963$, $c_{20} = -8.3606$, $c_{40} = 39.7410$. Figure 8.13 shows the agreement of $C_L(a_i)$ with the lift-values of the URANS simulation, where the a_i are the filtered coefficients.

The mean values differ in both cases, the unactuated and the actuated one, between the original lift values and the values calculated with C_L not more than 0.5%, see Figure 8.12. This is negligible, because in contrast to our

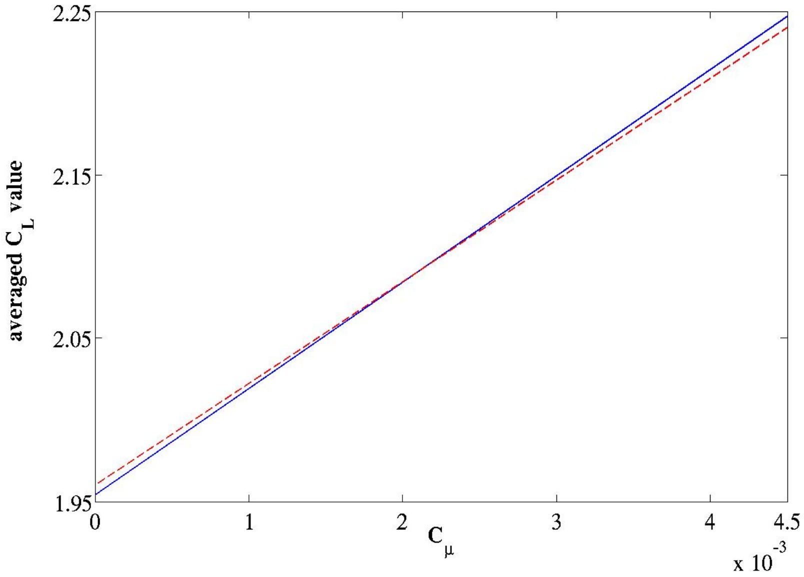

Figure 8.12: Comparison of the original mean lift values by the URANS simulation (continuous lines) and the mean lift computed by (8.3.1) (dashed lines).

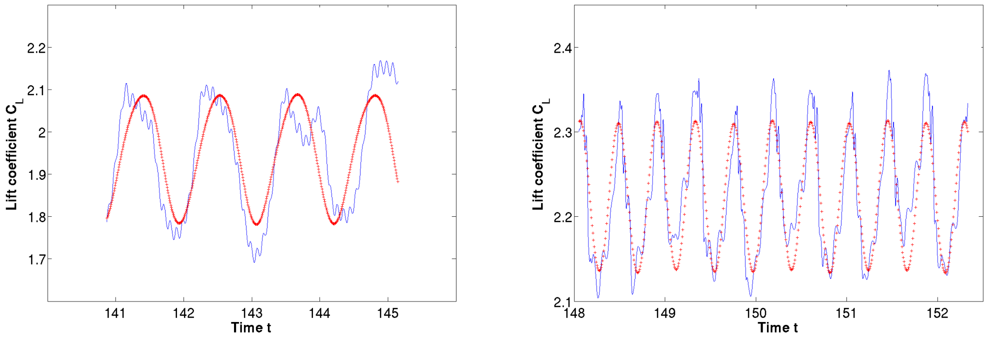

Figure 8.13: Comparison of the lift values of the URANS simulation (continuous lines) with those obtained by the lift formula based on the filtered coefficients $a_1, \cdots, a_4$ (dashed lines): Natural flow (left), actuated flow (right).

stationary case without turbulence, we achieve a lift gain of more than 14% in the full problem with an actuation in contrast to the case without actuation. Evaluating C_L with the a_i's as the solutions of the dynamical system, once computed with $B = 0$ and once with the full actuation $B = 3.5888$, we get mean values of around 1.96 respectively 2.24 and the results presented in Figure 8.14.

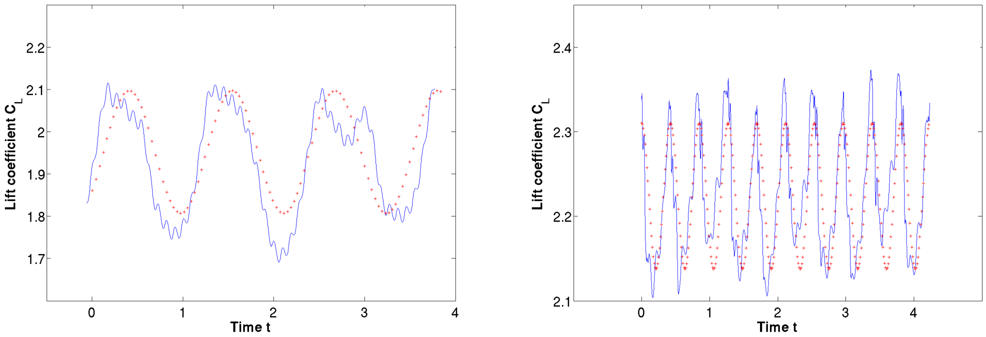

Figure 8.14: Comparison of the lift values of the URANS simulation (continuous lines) with those obtained by the lift formula based on $a_1, \cdots, a_4$ of the dynamical system (dashed lines): Natural flow (left), actuated flow (right).

Solving this dynamical system with several actuation amplitudes $B = 0$ to $B = 3.6$, we resolve the average lift values presented in Figure 8.15; for $B = 0$ an average lift of 1.96 and for $B = 3.5888$ an average lift of 2.22. The optimization problem (8.3.2) yields a lift gain of more than 13%. The maximal lift is achieved at an actuation amplitude of around $B_{opt} = 2.4$, agreeing with the results of the URANS simulation.

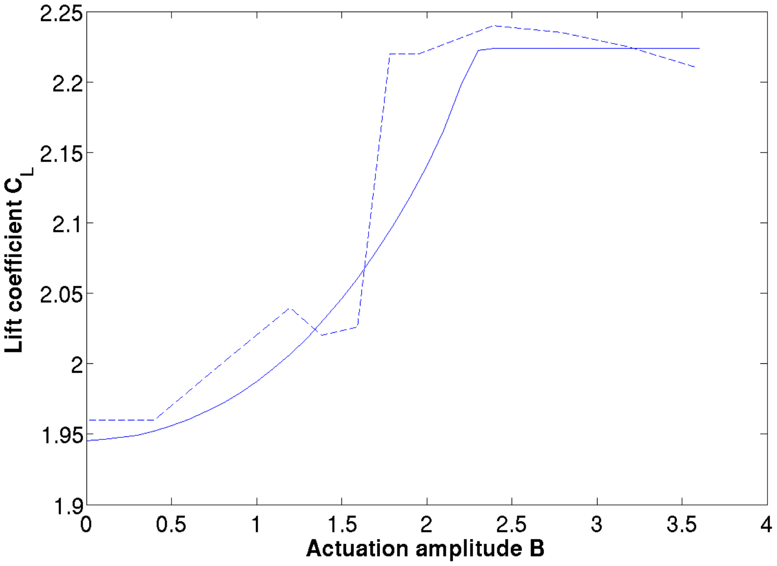

Figure 8.15: Comparison of the lift coefficients calculated by C_L (continuous lines) with those obtained by the URANS simulation (dashed lines) with the Wilcox98 $k-\omega-$turbulence model for some chosen actuation amplitudes B.

Chapter 9

The optimal control problem

In this chapter, we want to investigate an optimal control problem (P_N) based on the optimization problem (8.3.2) of the last chapter. Our goal is to reach as much lift as possible where the actuation amplitude B is the optimization parameter.

Let $\hat{B} \in \mathbb{R}$ denote an upper bound and 0 the natural lower bound for the actuation amplitude B, $a_{10},\ a_{20},\ a_{30},\ a_{40}$ the initial values for the state $a(t)$, $T_\Delta = T_e - T_a$ and $g(t) = B\cos(\omega^a t)$. Then the optimization problem (P_N) looks with $a(t) = (a_1(t), a_2(t), a_3(t), a_4(t))$ as follows:

$$\min J_N(a(t), B) := \frac{1}{T_\Delta} \int\limits_{T_a}^{T_e} -C_L(a_1(t), a_2(t), a_3(t), a_4(t))\ \mathrm{d}t + \frac{\alpha_N}{2} B^2$$

subject to the reduced order model

$$\begin{aligned}
\dot{a}_1(t) &= \tilde{\sigma}^n a_1(t) - \omega^n a_2(t) \\
\dot{a}_2(t) &= \omega^n a_1(1) + \tilde{\sigma}^n a_2(t) \\
\dot{a}_3(t) &= \tilde{\sigma}^a a_3(1) - \omega^a a_4(t) + g_{31} g(t) + g_{32} \dot{g}(t) \\
\dot{a}_4(t) &= \omega^a a_3(t) + \tilde{\sigma}^a a_4(t) + g_{41} g(t) + g_{42} \dot{g}(t) \\
a_1(0) &= a_{10} \\
a_2(0) &= a_{20} \\
a_3(0) &= a_{30} \\
a_4(0) &= a_{40}
\end{aligned} \tag{9.0.1}$$

with the amplifier rates

$$\begin{aligned}
\tilde{\sigma}^n(t) &= \sigma^n - \beta_1 (A^n(t))^2 - \beta_2 (A^a(t))^2 \\
\tilde{\sigma}^a(t) &= \sigma^a,
\end{aligned}$$

the control constraint

$$B \in B_{ad} := \{B \in \mathbb{R} : 0 \leq B \leq \hat{B}\}.$$

Note that it seems as if this system is linear, but it is nonlinear due to $\tilde{\sigma}^n$ with $A^n(t) = \sqrt{a_1^2(t) + a_2^2(t)}$ and $A^a(t) = \sqrt{a_3^2(t) + a_4^2(t)}$.

In the next section, we introduce the first-order optimality system, which we will need for the numerical investigation.

9.1 First-order necessary optimality conditions

We directly apply the formal Langrange technique to this problem (P_N) to derive the first-order necessary optimality conditions. We do this in a formal way without considering the exact function spaces providing the background of this theory.

Following the Lagrange technique, we want to substitute the ODEs in (9.0.1) by Lagrange multiplication functions $\lambda(t) = (\lambda_1(t), \lambda_2(t), \lambda_3(t), \lambda_4(t))$ while the box constraint $B \in B_{ad}$ for the control parameter and the initial values are not eliminated.

Then, we obtain the Lagrange-functional $\mathcal{L}_N(a(t), B, \lambda(t))$ with the state variable $a(t)$, the adjoint state $\lambda(t)$ and the actuation amplitude B:

$$\begin{aligned}
\mathcal{L}_N(a, B, \lambda) :=& \frac{1}{T_\Delta} \int_{T_a}^{T_e} -C_L(a)\ \mathrm{d}t + \frac{\alpha_N}{2} B^2 - \int_{T_a}^{T_e} (\dot{a}_1 - \tilde{\sigma}^n a_1 + \omega^n a_2) \lambda_1 \mathrm{d}t \\
&- \frac{1}{T_\Delta} \int_{T_a}^{T_e} (\dot{a}_2 - \omega^n a_1 - \tilde{\sigma}^n a_2) \lambda_2 \mathrm{d}t \\
&- \frac{1}{T_\Delta} \int_{T_a}^{T_e} (\dot{a}_3 - \sigma^a a_3 + \omega^a a_4 - g_{31} g - g_{32} \dot{g}) \lambda_3 \mathrm{d}t \\
&- \frac{1}{T_\Delta} \int_{T_a}^{T_e} (\dot{a}_4(t - \omega^a a_3 - \sigma^a a_4 - g_{41} g - g_{42} \dot{g}) \lambda_4 \mathrm{d}t.
\end{aligned}$$

Following the Lagrange principle, the optimal control $\bar{B}$ together with the associating optimal state $\bar{a}(t)$ has to fulfill the necessary first-order optimality condition of the problem including the minimization of the Lagrange functional $\mathcal{L}_N$ with respect to $a(t), B$ and the box constraint $B \in B_{ad}$ for the control but without the state equations (9.0.1).

Thus, the Lagrange-functional $\mathcal{L}_N$ has to fulfill

$$\frac{\partial \mathcal{L}_N}{\partial a}(\bar{a}, \bar{B}, \lambda)h = 0 \text{ for all possible } h(\cdot) \text{ with } h(0) = 0 \tag{9.1.1}$$

and with respect to the control parameter B the variational inequality

$$\frac{\partial \mathcal{L}_N}{\partial B}(\bar{a}, \bar{B}, \lambda)(B - \bar{B}) \geq 0 \text{ for all } B \in B_{ad}. \tag{9.1.2}$$

The equality (9.1.1) is with $h = (h_1, h_2, h_3, h_4)$ equivalent to

$$\begin{aligned}
\frac{\partial \mathcal{L}_N}{\partial a}(\bar{a}, \bar{B}, \lambda)(h) = & \int_{T_a}^{T_e} (\sum_{i=1}^{4}(c_{1i}h_i) + 2c_{20}a_1h_1 + 2c_{20}a_2h_2 + 4c_{40}(A^n)^2a_1h_1 \\
& \quad + 4c_{40}(A^n)^2a_2h_2) \, \mathrm{d}t \\
& - \int_{T_a}^{T_e} (\dot{h}_1 - 2(\beta_1a_1h_1 + \beta_1a_2h_2 + \beta_2a_3h_3 + \beta_2a_4h_4)a_1 \\
& \quad - \sigma^n h_1 + \omega^n h_2 - (\beta_1a_2^2 + \beta_2a_3^2 + \beta_2a_4^2)h_1)\lambda_1 \, \mathrm{d}t \\
& - \int_{T_a}^{T_e} (\dot{h}_2 - 2(\beta_1a_1h_1 + \beta_1a_2h_2 + \beta_2a_3h_3 + \beta_2a_4h_4)a_2 \\
& \quad - \omega^n h_1 - \sigma^n h_2 - (\beta_1a_1^2 + \beta_2a_3^2 + \beta_2a_4^2)h_2)\lambda_2 \, \mathrm{d}t \\
& - \int_{T_a}^{T_e} (\dot{h}_3 - \tilde{\sigma}^a h_3 + \omega^a h_4)\lambda_3 \, \mathrm{d}t \\
& - \int_{T_a}^{T_e} (\dot{h}_4 - \omega^a h_3 - \tilde{\sigma}^a h_4)\lambda_4 \, \mathrm{d}t
\end{aligned}$$

and after integration by parts, we derive

$$\begin{aligned}
\frac{\partial \mathcal{L}_N}{\partial a}(\bar{a},\bar{B},\lambda)(h) = \int\limits_{T_a}^{T_e} &((c_{11}+2c_{20}a_1+4c_{40}(A^n)^2a_1)h_1+(c_{13})h_3 \\
&+(c_{12}+2c_{20}a_2+4c_{40}(A^n)^2a_2)h_2+(c_{14})h_4)\,\mathrm{d}t \\
&-\int\limits_{T_a}^{T_e}((-\dot{\lambda}_1-\sigma^n-2(\beta_1a_1\lambda_1+\beta_1a_1\lambda_2)a_1 \\
&\quad-(\beta_1a_2^2+\beta_2a_3^2+\beta_2a_4^2)+\omega^n\lambda_2)h_1)\,\mathrm{d}t \\
&-\int\limits_{T_a}^{T_e}((-\dot{\lambda}_2-\omega^n\lambda_1-\sigma^n-2(\beta_1a_2\lambda_1+\beta_1a_2h_2)a_2 \\
&\quad-(\beta_1a_1^2+\beta_2a_3^2+\beta_2a_4^2))h_2)\,\mathrm{d}t \\
&-\int\limits_{T_a}^{T_e}((-\dot{\lambda}_3+\beta_2a_3(\lambda_1+\lambda_2)-\tilde{\sigma}^a\lambda_3+\omega^a\lambda_4)h_3)\,\mathrm{d}t \\
&-\int\limits_{T_a}^{T_e}((-\dot{\lambda}_4+\beta_2a_4(\lambda_1+\lambda_2)-\omega^a\lambda_3-\tilde{\sigma}^a\lambda_4)h_4)\,\mathrm{d}t \\
&-\int\limits_{T_a}^{T_e}(h_1(T)\lambda_1(T)+h_2(T)\lambda_2(T)+h_3(T)\lambda_3(T) \\
&\quad+h_4(T)\lambda_4(T))\mathrm{d}t.
\end{aligned}$$

Since $h(T)$ and $h(\cdot)$ can be arbitrarily, $\lambda=(\lambda_1,\lambda_2,\lambda_3,\lambda_4)$ is the weak solution of

$$\begin{aligned}
-\dot{\lambda}_1-\sigma^n-2(\beta_1a_1\lambda_1+\beta_1a_1\lambda_2)a_1+\omega^n\lambda_2-(\beta_1a_2^2+\beta_2a_3^2+\beta_2a_4^2)&=f_1\\
-\dot{\lambda}_2-\omega^n\lambda_1-\sigma^n-2(\beta_1a_2\lambda_1+\beta_1a_2\lambda_2)a_2-(\beta_1a_1^2+\beta_2a_3^2+\beta_2a_4^2)&=f_2\\
-\dot{\lambda}_3+\beta_2a_3(\lambda_1+\lambda_2)-\tilde{\sigma}^a\lambda_3+\omega^a\lambda_4&=f_3\\
-\dot{\lambda}_4+\beta_2a_4(\lambda_1+\lambda_2)-\omega^a\lambda_3-\tilde{\sigma}^a\lambda_4&=f_4\\
\lambda_1(T)=\lambda_2(T)=\lambda_3(T)=\lambda_4(T)&=0
\end{aligned} \tag{9.1.3}$$

with

$$\begin{aligned}
f_1&=c_{11}+2c_{20}a_1+4c_{40}(A^n)^2a_1\\
f_2&=c_{12}+2c_{20}a_2+4c_{40}(A^n)^2a_2\\
f_3&=c_{13}\\
f_4&=c_{14}
\end{aligned}$$

which we define as the adjoint system. The solution is the adjoint state λ.

The second requirement (9.1.2) leads to the variational inequality

$$\frac{\partial \mathcal{L}_N}{\partial B}(B-\bar{B}) = \alpha_N \bar{B}(B-\bar{B}) + \frac{1}{T_\Delta}\int\limits_{T_a}^{T_e} [(g_{31}\cos(\omega^a t) - g_{32}\omega^a \sin(\omega^a t))\lambda_3$$
$$+ (g_{41}\cos(\omega^a t) - g_{42}\omega^a \sin(\omega^a t)\lambda_4)](B-\bar{B})\,\mathrm{d}t \geq 0$$

for all $\hat{B} \geq B \geq 0$. The pointwise analysis of this inequality leads to the standard projection formula

$$\bar{B} = \mathbb{P}_{[0,\hat{B}]}\{\frac{1}{\alpha_N T_\Delta}\int\limits_{T_a}^{T_e} (g_{31}\cos(\omega^a t) - g_{32}\omega^a \sin(\omega^a t))\lambda_3$$
$$+ (g_{41}\cos(\omega^a t) - g_{42}\omega^a \sin(\omega^a t)\lambda_4)\,\mathrm{d}t\}. \tag{9.1.4}$$

9.2 Numerical investigation

Now, let us research the optimization problem (P_N) numerically based on the optimality system. In this case, we decided to use the gradient-projection method, see 5.2.2, because, as mentioned before, COMSOL and the integral term (9.1.4) don't fit together. One can see at the end of this chapter that we need about 5 iterations to get the optimal actuation amplitude. That means that we have to solve both the nonlinear state equations (9.0.1) and the linear adjoint system (9.1.3) 5 times. To approximate the optimal the optimal actuation amplitude $\bar{B}$ by just solving the state equation with different amplitudes would take probably more iterations of the nonlinear state equation, due to the fact that the upper bound $\hat{B}$ is free to select. For our optimization problem (P_N) the algorithm reads as follows with given B_n:

S1 Calculate $a_n = ((a_1)_n, (a_2)_n, (a_3)_n, (a_4)_n)$ as the solution of (9.0.1) with the current control B_n.

S2 Calculate the adjoint $\lambda_n = ((\lambda_1)_n, (\lambda_2)_n, (\lambda_3)_n, (\lambda_4)_n)$ from (9.1.3) with the current state a_n.

S3 The updated descent direction is

$$D_n = \alpha_N B_n + \int\limits_0^T (g_{31}\cos(\omega^a t) - g_{32}\omega^a \sin(\omega^a t))\lambda_3$$
$$+ (g_{41}\cos(\omega^a t) - g_{42}\omega^a \sin(\omega^a t)\lambda_4)\mathrm{d}t \in \mathbb{R}.$$

S4 Calculate the *stepsize* s_n from

$$\min_{s>0} f(\mathbb{P}_{[0,\hat{B}]}\{B_n + sD_n\}).$$

S5 The *updated control* B_{n+1} is

$$B_{n+1} := \mathbb{P}_{[0,\hat{B}]}\{B_n + s_n D_n\}.$$

Set n:=n+1and goto **S1**.

Let us now investigate the Setting 7.1 on page 90 with the parameters σ^n, σ^a, ω^n, ω^a, g_{31}, g_{32}, g_{41}, g_{42}, β_1, β_2, c_{10}, c_{11}, c_{12}, c_{13}, , c_{14}, c_{20} and c_{40} as selected Section in 8.4 numerically. Therefore, we chose $\alpha_N = 0.1$ and we decide to optimize this problem in the time interval $[T_a, T_e] = [56.8612, 75.8150]$.

The first reason of this interval is that we want to optimize the lift and not the transient oscillation, so we select $T_a \neq 0$. Additionally, we have to choose T_a and T_e in the way that they are multiples of both wavelengths, the natural and the actuated one.

We started the gradient-projection method with initial values $((a_1)_0, (a_2)_0, (a_3)_0, (a_4)_0) = (r^n, 0, 0, 0)$, where r^n denotes the radius of the natural attractor, $B_0 = 0.5$ and a mesh size of 0.0132 in the time direction.

The optimal calculated actuation amplitude is $B_{opt} = 2.2573$ with an associated averaged lift coefficient of

$$\frac{1}{T_\Delta}\int\limits_{T_a}^{T_e} C_L((a_1)_{opt}, (a_2)_{opt}, (a_3)_{opt}, (a_4)_{opt})\,\mathrm{d}t = 2.2238,$$

see Figure 9.1, and $J_N(a_{opt}, B_{opt}) = -1.9690$ as the value of the cost functional. The calculated optimal actuation amplitude B_{opt} differs slightly from the optimal value in Figure 8.15, due to the term $\frac{\alpha_N}{2}B^2$ in the cost functional. The Figure 9.1 presents the optimal lift coefficient over the time interval $[T_a, T_e]$ and the Figures 9.2 and 9.3 the optimal states $(a_1, \cdots, a_4)$ and the associating adjoint $(\lambda_1, \cdots, \lambda_4)$, respectively.

Unfortunately, we have no simulations of the full $k-\omega$ turbulence system with our optimal actuation amplitude B_{opt}. The simulation with an amplitude nearest to B_{opt} we have is a simulation with an actuation amplitude of 2.39. The results of the lift values are compared in Figure 9.5.

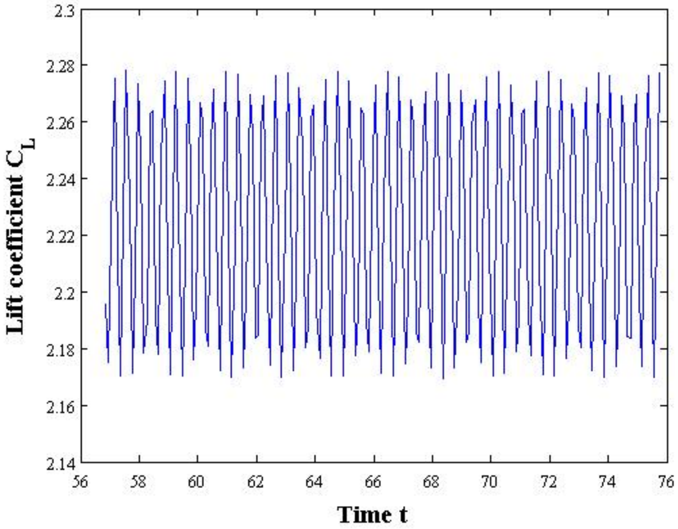

Figure 9.1: The lift coefficient C_L of the calculated optimal state over the interval $[T_a, T_e]$.

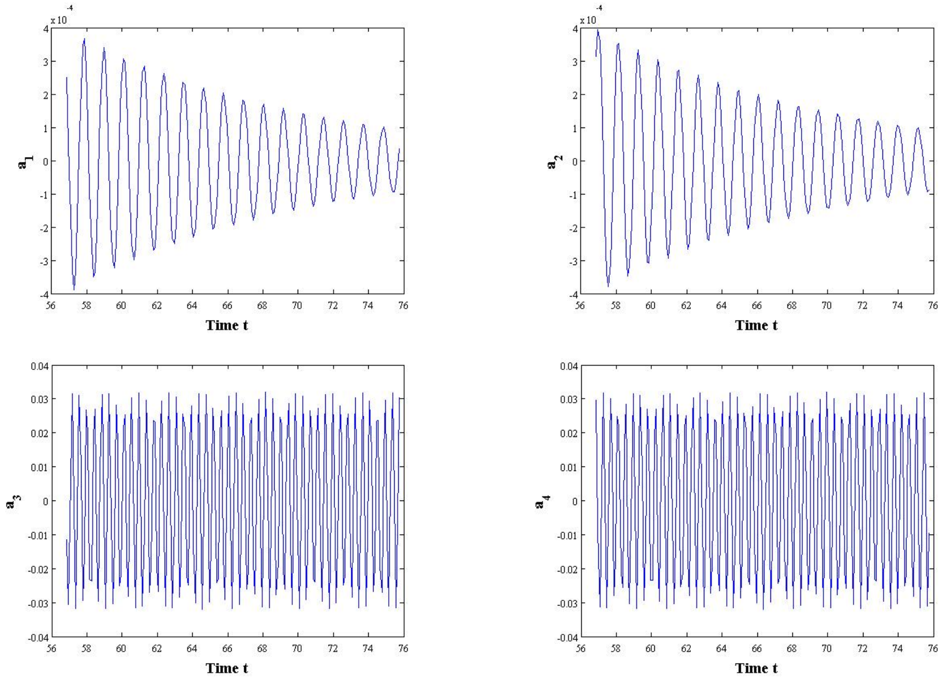

Figure 9.2: The calculated optimal state (a_1: top left, a_2: top right, a_3: bottom left, a_4: bottom right).

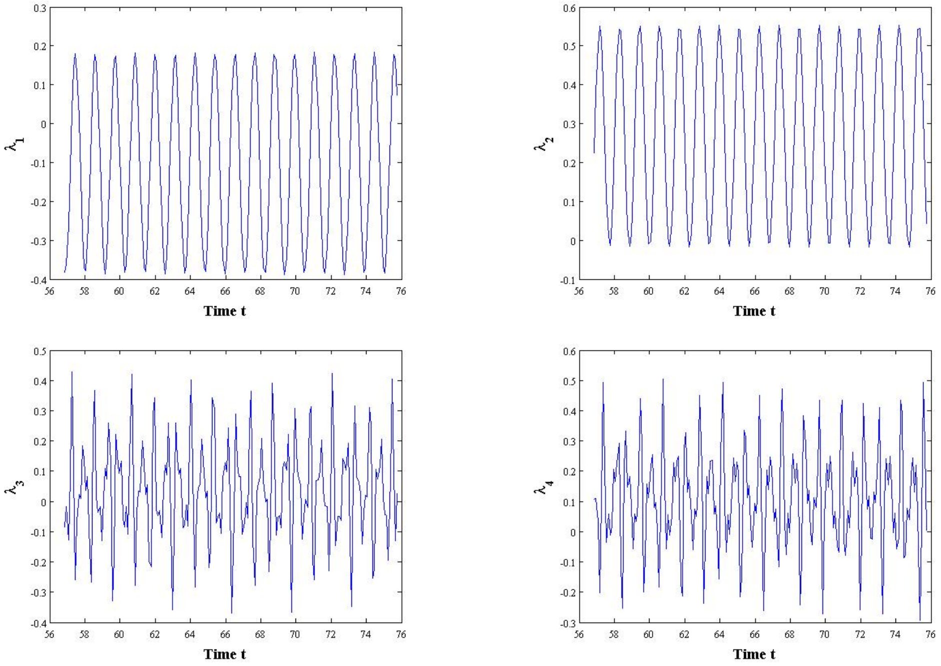

Figure 9.3: The calculated adjoint state (λ_1: top left, λ_2: top right, λ_3: bottom left, λ_4: bottom right).

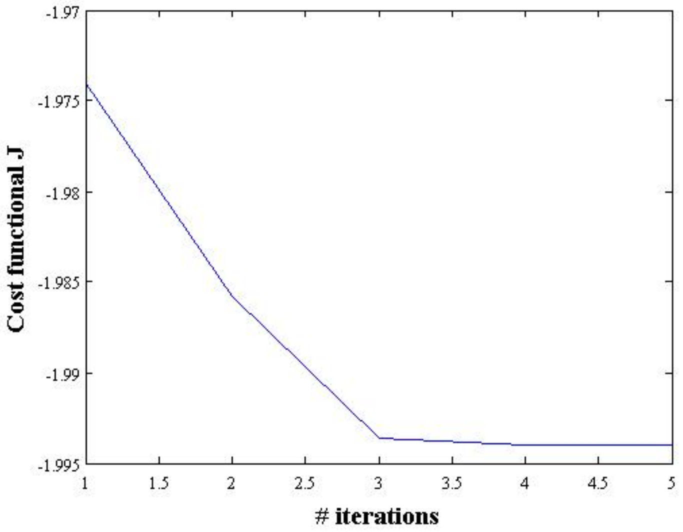

Figure 9.4: The cost functional J_N over the number of iterations.

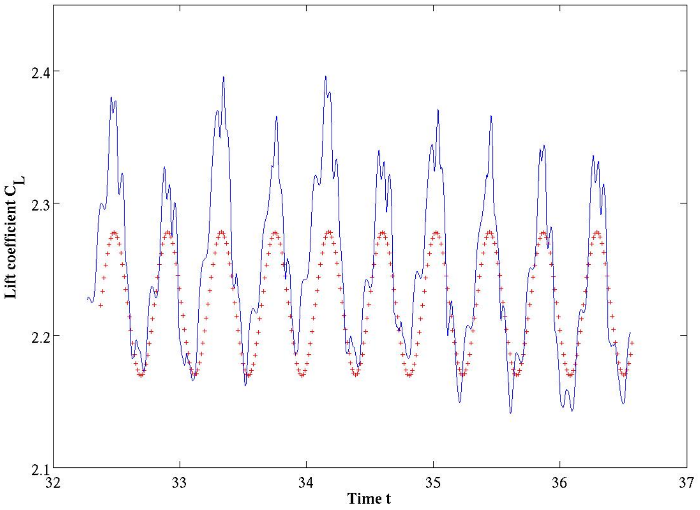

Figure 9.5: The cost functional J_N over the number of iterations.

Chapter 10

Conclusion

In this thesis, we considered two settings of optimal control problems for high-lift configurations. In the case of steady-state Navier-Stokes equations with low Reynolds number, we established first-order necessary optimality conditions for a problem with an integral state constraint on the drag. The main theoretical difficulty was the appearance of low regularity controls in a Dirichlet boundary condition. Afterwards, we considered the second-order sufficient optimality conditions for the infinite and a finite-dimensional control space. The optimal control is obtained by direct numerical solution of the established optimality system and by an SQP-method, where the integral state constraint was handled by a Penalty term in the cost functional.

An associated nonstationary case with high Reynolds number was investigated with a WILCOX98 turbulence model. To handle the problem of the high dimension, a robust reduced order model (ROM) was established fitting best to the snapshots computed by the full system in the natural and the actuated state. The ROM reproduces the nonlinear behavior of the system sufficiently well so that the subsequently optimization problem of periodic actuation leads to reasonable results. We are now able to solve our optimal control problem in about 20 minutes by 5 iterations and 4 minutes for one forward and adjoint system together. Without the model reduction, just one forward iterations would take about 4 days.

In particular, the application of trust-region proper orthogonal decomposition (TRPOD) could be considered to develop an improved reduced-order model. In [15] a ROM was used to minimize the total mean drag for a circular cylinder wake flow by updating the ROM during a (TRPOD) approach, we refer also to [14].

Chapter 11

Zusammenfassung

In dieser Arbeit haben wir zwei Optimalsteuerungsprobleme für Hochauftriebskonfigurationen untersucht.

Im ersten Fall, der stationären Navier-Stokes Gleichungen mit Kontroll-, integralen Zustandsbeschränkungen und kleinen Reynolds-Zahlen, haben wir zunächst die notwendigen Optmalitätsbedingungen erster Ordnung aufgestellt um das Problem numerisch zu untersuchen. Dabei war die gewünschte niedrige Regularität der Dirichlet-Randsteuerungen das größte theoretische Problem. Anschliessend haben wir die hinreichenden Optimalitätsbedingungen zweiter Ordnung für unendlich und endlich dimensionale Steuerungen aufgestellt. Numerisch haben wir das Problem einerseits als direkte Lösung des Optimalitätssystems und andererseits mit Hilfe der SQP-Methode untersucht. Zum Abschluss dieses Themenbereichs wurde noch die Konvergenz der SQP-Methode bewiesen.

Der instationäre Fall wurde mit grossen Reynolds-Zahlen und zugehörigen Turbulenzen betrachtet. Die Turbulenzen wurden durch das WILCOX98 Modell beschrieben, was zu einem riesigen Rechenaufand führt. Alleine eine Vorwärtsrechnung der Zustandsgleichung hat bei vergleichbaren Problemen mehr als 4 Tage gedauert. Zur Lösung dieses Problems haben wir eine Modellreduktion durchgeführt und ein reduced-order model (ROM) aufgestellt, welches am besten zu vorher berechneten Snapshots passt. Wir haben es geschafft, dass dieses ROM die nichtlineare Struktur des Systems hinreichend gut widerspiegelt, so dass eine Optimierung auf dessen Basis möglich ist und sinnvolle Resultate liefert. Desweiteren gelang es uns das Optimalsteuerungsproblem innerhalb von etwa 20 Minuten zu lösen, bei 5 Iterationen und 4 Minuten Dauer für eine Vorwärts- und eine adjungierte Gleichung.

Bibliography

[1] F. Abergel and E. Casas. Some optimal control problems associated to the stationary Navier-Stokes equation. In *Mathematics, climate and environment (Madrid, 1991)*, volume 27 of *RMA Res. Notes Appl. Math.*, pages 213–220. Masson, Paris, 1993.

[2] F. Abergel and R. Temam. On some control problems in fluid mechanics. *Theoret. Comput. Fluid Dynam.*, 1:303–325, 1990.

[3] R. A. Adams. *Sobolev Spaces.* Academic Press, Boston, 1978.

[4] K. Afanasiev and M. Hinze. Adaptive control of a wake flow using proper orthogonal decomposition. In *Lecture Notes in Pure and Applied Mathematics 216, 317-332, Shape Optimization and Optimal Design, Marcel Dekker.*

[5] W. Alt. The Lagrange-Newton method for infinite-dimensional optimization problems. *Numer. Funct. Anal. and Optim.*, 11:201–224, 1990.

[6] W. Alt. Parametric optimization with applications to optimal control and sequential quadratic programming. *Bayreuther Mathematische Schriften*, 1991.

[7] W. Alt. Sequential quadratic programming in Banach spaces. In *Advances in Optimization*, volume 382 of *Lecture Notes in Economics and Mathematical systems*, pages 281–301, New York, 1992. Springer–Verlag.

[8] W. Alt. The Lagrange-Newton method for infinite-dimensional optimization problems. *Control and Cybernetics*, 23:87–106, 1994.

[9] W. Alt. Discretization and mesh independence of Newton's method for generalized equations. *Preprint*, 1997.

[10] W. Alt and K. Malanowski. The Lagrange-Newton Method for Nonlinear Optimal Control Problems. *Computational Optimization and Applications*, 2:77–100, 1993.

[11] W. Alt and K. Malanowski. The Lagrange-Newton method for state constrained optimal control problems. *Comp. Optimization and Appl.*, 4(3):217–239, 1995.

[12] C. Amrouche and V. Girault. On the existence and regularity of the solution of Stokes problem in arbitrary dimension. *Proc. Japan Acad. Ser. A Math. Sci.*, 67(5):171–175, 1991.

[13] Raymond J.-P.. Arada, N. and F. Tröltzsch. On an augmented Lagrangian SQP method for a class of optimal control problems in Banach spaces. *Computational Optimization and Applications*, 22:369–398, 2002.

[14] E. Arian, M. Fahl, and E.W. Sachs. Trust-region proper-orthogonal decomposition for flow control. Technical Report 2000-25, Universität Trier, 2000.

[15] M. Bergmann and L. Cordier. Optimal control of the cylinder wake in the laminar regime by trust-region methods and pod reduced-order models. *J. Comput. Phys.*, 227(16):7813–7840, 2008.

[16] M. Bergmann, L. Cordier, and J.-P.. Brancher. Optimal rotary control of the cylinder wake using proper orthogonal decomposition reduced order model. *Phys. Fluids*, 17:1–21, 2005.

[17] T. Bewley, P. Moin, and R. Temam. DNS-based predictive control of turbulence: an optimal benchmark for feedback algorithms. *J. Fluid Mech.*, 447:179–225, 2001.

[18] M. Braack and T. Richter. Solutions of 3d Navier-Stokes benchmark problems with adaptive fnite elements. *Computers and Fluids.*, 35(4):372–392, 2006.

[19] A. Carnarius, B. Günther, F. Thiele, D. Wachsmuth, F. Tröltzsch, and J.C. Reyes. Numerical study of the optimization of separation control. In *Proceedings of the 45th AIAA Aerospace Sciences Meeting and Exhibit, Reno, 8-11 January 2007, AIAA 2007-58.*

[20] E. Casas, M. Mateos, and J.-P. Raymond. Error estimates for the numerical approximation of a distributed control problem for the steady-state Navier-Stokes equations. *SIAM J. Control Optim.*, 46(3):952–982, 2007.

[21] E. Casas and F. Tröltzsch. Second-order necessary and sufficient optimality conditions for optimization problems and applications to control theory. *SIAM J. Optim.*, 13(2):406–431, 2002.

[22] E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for some state-constrained control problems of semilinear elliptic equations. *SIAM J. Control and Optimization*, 38:1369–1391, 2000.

[23] Eduardo Casas. The Navier-Stokes equations coupled with the heat equation: analysis and control. *Control Cybernet.*, 23(4):605–620, 1994. Modelling, identification, sensitivity analysis and control of structures.

[24] Eduardo Casas. An optimal control problem governed by the evolution Navier-Stokes equations. In *Optimal control of viscous flow*, pages 79–95. SIAM, Philadelphia, PA, 1998.

[25] L. Cattabriga. Su un problema al contorno relativo al sistema di equazioni di Stokes. *Rend. Sem. Mat. Univ. Padova*, 31:308–340, 1961.

[26] R. Dautray and J.L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology*. Springer-Verlag, address = Berlin, year = 2000.

[27] R. Dautray and J.L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology, Vol. 6*. Springer Verlag, Berlin, 2000.

[28] J. C. de los Reyes and K. Kunisch. A semi-smooth Newton method for control constrained boundary optimal control of the Navier-Stokes equations. *Nonlinear Anal.*, 62(7):1289–1316, 2005.

[29] J. C. de los Reyes and F. Tröltzsch. Flow control with regularized state constraints. *in: Active Flow Control, Notes on Numerical Fluid Mechanics and Multidisciplinary Design (NNFM)*, 95:353–366, 2007.

[30] J. C. de los Reyes and I. Yousept. Regularized state-constrained boundary optimal control of the navier-stokes equations. *Journal of Mathematical Analysis and Applications.*, 356:257–279, 2009.

[31] J.C. de los Reyes, P. Merino, J. Rehberg, and F. Tröltzsch. Optimality conditions for state-constrained pde control problems with time-dependent controls. *To appear in Control and Cybernetics.*

[32] P. Deuflhard. *Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms.* Band 35 der Reihe Springer Series in Computational Mathematics. Springer-Verlag, Berlin, 2004.

[33] A. L. Dontchev. Uniform convergence of the Newton method for Aubin continuous maps. *Serdica Math. J.*, 22(3):385–398, 1996.

[34] A. L. Dontchev, W. W. Hager, A. B. Poore, and B. Yang. Optimality, stability, and convergence in nonlinear control. *Applied Math. and Optimization*, 31:297–326, 1995.

[35] Asen L. Dontchev. Implicit function theorems for generalized equations. *Math. Program.*, 70(1):91–106, 1995.

[36] Asen L. Dontchev. Local analysis of a Newton-type method based on partial linearization. volume 32, pages 295–306, 1996.

[37] Asen L. Dontchev and William W. Hager. Implicit functions, lipschitz maps, and stability in optimization. *Math. Oper. Res.*, 19(3):753–768, 1994.

[38] J. Dušek, P. Le Gal, and P. Fraunié. A numerical und theoretical study of the Hopf bifurcation in a cylinder wake. *J. Fluid Mech.*, 264:59–80, 1994.

[39] R. Farwig, G. P. Galdi, and H. Sohr. A new class of weak solutions of the Navier-Stokes equations with nonhomogeneous data. *J. Math. Fluid Mech.*, 8(3):423–444, 2006.

[40] R. Farwig, G. P. Galdi, and H. Sohr. Very weak solutions and large uniqueness classes of stationary Navier-Stokes equations in bounded domains of $\mathbb{R}^2$. *J. Differential Equations*, 227(2):564–580, 2006.

[41] C.A.J. Fletcher. *Computational Galerkin Methods.* Springer-Verlag, 1984.

[42] R. Fletcher. *Practical Methods of Optimization.* Wiley, 1987.

[43] M. Gad-el Hak. *Flow Control. Passive, Active, and Reactive Flow Management.* Cambridge University Press, 2000.

[44] G. P. Galdi. *An introduction to the mathematical theory of the Navier-Stokes equations.* Springer, New York, 1992.

[45] G. P. Galdi, C. G. Simader, and H. Sohr. A class of solutions to stationary Stokes and Navier-Stokes equations with boundary data in $W^{-1/q,q}$. *Math. Ann.*, 331(1):41–74, 2005.

[46] P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization.* Academic Press, London, 1981.

[47] V. Girault and P. A. Raviart. *Finite Element Methods for Navier Stokes Equations.* Springer Verlag Berlin, 1986.

[48] H. Goldberg and F. Tröltzsch. On a Lagrange-Newton Method for a nonlinear parabolic boundary control problem. *Optimization Methods and Software*, 8:225–247, 1998.

[49] W.R. Graham, J. Peraire, and K.Y. Tang. Optimal control of vortex shedding using low-order models. Part I:Open-loop model development. *Int. Num. Meth. Eng.*, 44:945–972, 1999.

[50] R. Griesse and J. C. de los Reyes. State-constrained optimal control of the three-dimensional stationary Navier-Stokes equations. *J. Math. Anal. Appl.*, 343(1):257–272, 2008.

[51] M. D. Gunzburger, L. S. Hou, and Th. P. Svobodny. Analysis and finite element approximation of optimal control problems for the stationary Navier-Stokes equations with Dirichlet controls. *RAIRO Modél. Math. Anal. Numér.*, 25(6):711–748, 1991.

[52] M. D. Gunzburger and S. Manservisi. The velocity tracking problem for Navier-Stokes flows with boundary control. *SIAM J. Control Optim.*, 39:594–634, 2000.

[53] M. Hintermüller and M. Hinze. A SQP-Semi-Smooth Newton-type Algorithm applied to Control of the instationary Navier-Stokes System Subject to Control Constraints. *Siam J. Optim.*, 16:1177–1200, 2006.

[54] M. Hinze and K. Kunisch. Second-order methods for optimal control of time-dependent fluid flow. *SIAM J. Control Optim.*, 40:925–946, 2001.

[55] M. Hinze and K. Kunisch. Second order methods for boundary control of the instationary Navier-Stokes system. *ZAMM*, 84:171–187, 2004.

[56] P. Holmes, J. L Lumley, and G. Berkooz. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry.* Cambridge University Press, Cambridge, 1998.

[57] C. John, B. R. Noack, M. Schlegel, F. Tröltzsch, and D. Wachsmuth. Optimal Boundary Control Problems Related to High-Lift Configurations. *Active Flow Control II. Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, 108/2010.

[58] C. John and D. Wachsmuth. Optimal Dirichlet boundary control of Navier-Stokes equations with state constraint. *Numerical Functional Analysis and Optimization*, 30(11&12).

[59] B.H. Jørgensen, J.N. Sørensen, and M. Brøns. Low-dimensional modelling of a driven cavity flow with two free parameters. *Theoret. Comput. Fluid Dynamics*, 16:299–317, 2003.

[60] N. H. Josephy. Newton's method for generalized equations. Technical report, 1979.

[61] K. Kunisch and B. Vexler. Constrained Dirichlet boundary control in L^2 for a class of evolution equations. *SIAM J. Control Optim.*, 46(5):1726–1753, 2007.

[62] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM Journal on Numerical Analysis*, 40:492–515, 2002.

[63] K. Kunisch and L. Xie. Suboptimal feedback control of flow separation by POD model reduction. *SIAM, DOI: 10.1137/1.9780898718935.ch12*, 2006.

[64] F.-S. Kupfer and E. W. Sachs. Numerical solution of a nonlinear parabolic control problem by a reduced SQP method. *Computational Optimization and Applications*, 1:113–135, 1992.

[65] O.A. Ladyzhenskaya. *The Mathematical Theory of Viscous Incompressible Flow.* Gordan and Breach, 1963.

[66] D.M. Luchtenburg, B. Günther, B.R. Noack, R. King, and G. Tadmor. A generalized mean-field model of the natural and high-frequency actuated flow around a high-lift configuration. *J. Fluid Mech.*, 623:283–316, 2009.

[67] K. Malanowski. Two-norm approach in stability and sensitivity analysis of optimization and optimal control problems. *Adv. Math. Sci. Appl.*, 2(2):397–443, 1993.

[68] K. Malanowski. Sufficient optimality conditions in optimal control. Technical report, 1994.

[69] E. Marušić-Paloka. Solvability of the Navier-Stokes system with L^2 boundary data. *Appl. Math. Optim.*, 41(3):365–375, 2000.

[70] H. Maurer and J. Zowe. First- and second-order conditions in infinite-dimensional programming problems. *Math. Programming*, 16:98–110, 1979.

[71] V. Maz'ya and J. Rossmann. Mixed boundary value problems for the Navier-Stokes system in polyhedral domains. *Mathematical Physics, arXiv:math-ph/0602054v1*, 2006.

[72] V. Maz'ya and J. Rossmann. L_p estimates of solutions to mixed boundary value problems for the Stokes system in polyhedral domains. *Math. Nachr.*, 280(7):751–793, 2007.

[73] I. Neitzel, U. Prüfert, and T. Slawig. Strategies for time-dependent pde control with inequality constraints using an integrated modeling and simulation environment. *Numericl Algorithms*, 50(3):241–269, 2009.

[74] B.R. Noack, K. Afanasiev, M. Morzyński, G Tadmor, and F. Thiele. A hierarchy of low-dimensional models for the transient and post-transient cylinder wake. *J. Fluid Mech.*, 497:335–363, 2008.

[75] B.R. Noack and G.S. Copeland. On a stability property of ensemble-averaged flow. *Tech. report 03/2000. Hermann-Föttinger-Institut für Strömungsmechanik.*, 3, 2000.

[76] B.R. Noack, P. Papas, and P.A. Monkewitz. The need for a pressure-term representation in empirical Galerkin models of incompressible shear flows. *J. Fluid Mech.*, 523:339–365, 2005.

[77] B.R. Noack, G. Tadmor, and M. Morzyński. Actuation models and dissipative control in empirical Galerkin models of fluid flows. In *In Proceedings of the 2004 American Control Conference*, pages 5722–5727, Daytona, OH, USA, 2004. American Automatic Control Council (AACC).

[78] J. Nocedal and S. J. Wright. *Numerical Optimization.* Springer-Verlag, New York, 1999.

[79] J.-P. Raymond. Stokes and Navier-Stokes equations with nonhomogeneous boundary conditions. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 24(6):921–951, 2007.

[80] J. C. de los Reyes and K. Kunisch. A semi-smooth Newton method for regularized state-constrained optimal control of the Navier-Stokes equations. *Computing*, 78(4):287–309, 2006.

[81] J. C. de los Reyes and F. Tröltzsch. Optimal control of the stationary Navier-Stokes equations with mixed control-state constraints. *SIAM J. Control Optim.*, 46(2):604–629, 2007.

[82] S. M. Robinson. Stability theory for systems of inequalities, part ii: Differentiable nonlinear systems. *SIAM J. Numer. Anal.*, 1979.

[83] S. M. Robinson. Strongly regular generalized equations. *Mathematics of Operation Research*, 1980.

[84] T. Roubíček and F. Tröltzsch. Lipschitz stability of optimal controls for the steady-state Navier-Stokes equations. *Control and Cybernetics*, 32(3):683–705, 2002.

[85] C.W. Rowley and V. Juttijudata. Model-based control and estimation of cavity flow oscillations. In *In Proceedings of the 44th IEEE Conference on Decision and Control and European Control Conference*, pages 512–517, Saint-Martin d'Hères, France., 2005. European Union Control Association (EUCA).

[86] M. Samimy, M. Debiasi, E. Caraballo, A. Serrani, X. Yuan, J. Little, and J. Myatt. Feedback control of subsonic cavity flows using reduced-order models. *J. Fluid Mech.*, 579:315–346, 2007.

[87] Günther B. Schatz, M. and F. Thiele. Computational investigation of separation control for high-lift airfoil flows. *Active Flow Control, Springer*, 2006.

[88] M. Schatz and F. Thiele. Numerical Study of High-Lift Flow with Separation Control by Periodic Excitation. *AIAA paper*, 2001-0296, 2001.

[89] M. Schatz, F. Thiele, R. Petz, and W. Nitsche. Separation control by periodic excitation and its application to a high-lift configuration. *AIAA*, (2507), 2004.

[90] Luchtenburg D.M. Noack B.R. Aleksic K. Pastoor M. King R. Schlegel, M. and G Tadmor. Turbulence Control Based on Reduced-Order Models and Nonlinear Control Design. In *Active Flow Control II.*

[91] Zh. W. Shen. A note on the Dirichlet problem for the Stokes system in Lipschitz domains. *Proc. Amer. Math. Soc.*, 123(3):801–811, 1995.

[92] S. Siegel, K. Cohen, and T. McLaughlin. Feedback control of a circular cylinder wake in experiment and simulation. *AIAA Paper*, pages 2003–3569, 2003.

[93] S.G. Siegel, J. Seidel, C. Fagley, D.M. Luchtenburg, K. Cohen, and T. McLaughlin. Low-dimensional modelling of a transient cylinder wake using double proper orthogonal decomposition. *J. Fluid Mech.*, 610:1–42, 2008.

[94] J.T. Stuart. On the non-linear mechanics of hydrodynamics stability. *J. Fluid Mech.*, 4:1–21, 1958.

[95] J.T. Stuart. Nonlinear stability theory. *Annu. Rev.Fluid Mech.*, 3:347–370, 1971.

[96] G. Tadmor, O. Lehmann, B. R. Noack, L. Cordier, J. Delville, j. P. Bonnet, and M. Morzyński. Reduced order models for closed-loop wake control. *Phil. Trans. R. Soc. A*, 369:1513 – 1624, 2011.

[97] R. Temam. *Navier-Stokes equations.* North Holland, Amsterdam, 1979.

[98] F.H. Tinapp and W. Nitsche. On active control of high-lift flow. In *In W. Rodi and D. Laurence, editors, Proc. 4th Int. Symposium on Engineering Turbulence Modelling and Measurements, Corsica, Elsevier Science*, 1999.

[99] F. Tröltzsch. *Optimality conditions for parabolic control problems and applications*, volume 62 of *Teubner Texte zur Mathematik.* B.G. Teubner Verlagsgesellschaft, Leipzig, 1984.

[100] F. Tröltzsch. An SQP method for the optimal control of a nonlinear heat equation. *Control and Cybernetics*, 23(1/2):267–288, 1994.

[101] F. Tröltzsch. Convergence of an SQP–Method for a class of nonlinear parabolic boundary control problems. In W. Desch, V. Kappel, and K. Kunisch, editors, *Control and Estimation of Distributed Parameter Systems: Nonlinear Phenomena. Int. Series of Num. Mathematics*, volume 118, pages 343–358, 1994.

[102] F. Tröltzsch. *Optimal Control of Partial Differential Equations. Theory, Methods and Applications. Graduate Studies in Mathematics, Volume 112.* American Mathematical Society, Providence, 2010.

[103] F. Tröltzsch and D. Wachsmuth. Second-order sufficient optimality conditions for the optimal control of Navier-Stokes equations. *ESAIM: COCV*, 12:93–119, 2006.

[104] A. Unger. *Hinreichende Optimalitätsbedingungen 2. Ordnung und Konvergenz des SQP-Verfahrens für semilineare elliptische Randsteuerprobleme.* PhD thesis, Technische Universität Chemnitz, 1997.

[105] S. Volkwein. Proper orthogonal decomposition and singular value decomposition. *SFB-Preprint*, 153, 1999.

[106] S. Volkwein. Optimal control of a phase-field model using proper orthogonal decomposition. *ZAMM*, 81:83–97, 2001.

[107] S. Volkwein. Lagrange-SQP techniques for the control constrained optimal boundary control problems for the Burgers equation. *Computational Optimization and Applications*, 26:253–284, 2003.

[108] S. Volkwein and F. Tröltzsch. Pod a-posteriori error estimates for linear-quadratic optimal control problems. *Computational Optimization and Applications*, 44:83–115, 2009.

[109] D. Wachsmuth. *Optimal control of the unsteady Navier-Stokes equations.* PhD thesis, Technische Universität Berlin, 2006.

[110] D. Wachsmuth. Sufficient second-order optimality conditions for convex control constraints. *J. Math. Anal. App.*, 319:228–247, 2006.

[111] Y. Wang, F. Tröltzsch, and G. Bärwolff. The POD Dirichlet Boundary Control of the Navier-Stokes Equations: A Low-dimensional Approach to Optimal Control with High Smoothness. Technical Report 23-2009, Technische Universität Berlin, 2009.

[112] E. Wassen and F. Thiele. Active control of a model vehicle wake. *Journal of Turbulence*, 2006.

[113] D.C. Wilcox. Simulation of Transition with a Two-Equation Turbulence Model. *AIAA JOURNAL*, 32, 1994.

[114] D.C. Wilcox. *Turbulence Modeling for CFD.* DCW Industries, La Cañada, 2006.

[115] J. Zowe and S. Kurcyusz. Regularity and stability for the mathematical programming problem in Banach spaces. *Appl. Math. Optim.*, 5(1):49–62, 1979.