Technische Universität Berlin

Institut für Mathematik

Analysis of the SQP-method for optimal control problems governed by the instationary Navier-Stokes equations based on L^p-theory

Daniel Wachsmuth

Preprint 13-2004

Preprint-Reihe des Instituts für Mathematik
Technische Universität Berlin

Report 13-2004 **December 2004**

Analysis of the SQP-method for optimal control problems governed by the instationary Navier-Stokes equations based on L^p-theory [1]

DANIEL WACHSMUTH[2]

Abstract. The aim of this article is to present a convergence theory of the SQP-method applied to optimal control problems for the instationary Navier-Stokes equations. We will employ a second-order sufficient optimality condition, which requires that the second derivative of the Lagrangian is positive definit on a subspace of inactive constraints. Therefore, we have to use L^p-theory of optimal controls of the instationary Navier-Stokes equations rather than Hilbert space methods. We prove local convergence of the SQP-method. This behaviour is confirmed by numerical tests.

Key words. Optimal control, Navier-Stokes equations, control constraints, Lipschitz stability, SQP-method

AMS subject classifications. Primary 49M37, Secondary 49N60

1. Introduction. We are considering optimal control of the instationary Navier-Stokes equations. The minimization of the following quadratic objective functional serves as model problem:

$$\min J(y,u) = \frac{\alpha_T}{2}\int_\Omega |y(x,T)-y_T(x)|^2dx + \frac{\alpha_Q}{2}\int_Q |y(x,t)-y_Q(x,t)|^2dxdt \\ + \frac{\alpha_R}{2}\int_Q |\operatorname{curl} y(x,t)|^2dxdt + \frac{\gamma}{2}\int_Q |u(x,t)|^2dxdt \quad (1.1)$$

subject to the instationary Navier-Stokes equations

$$\begin{aligned} y_t - \nu\Delta y + (y\cdot\nabla)y + \nabla p &= u && \text{in } Q,\\ \operatorname{div} y &= 0 && \text{in } Q,\\ y(0) &= y_0 && \text{in } \Omega, \end{aligned} \quad (1.2)$$

and the control constraints $u \in U_{ad}$ with set of admissible controls defined by

$$U_{ad} = \{u \in L^2(Q)^2 : \ u_{a,i}(x,t) \le u_i(x,t) \le u_{b,i}(x,t) \text{ a.e. on } Q,\ i = 1..2\}.$$

Here, Ω is an open bounded subset of $\mathbb{R}^2$ with C^3-boundary Γ such that Ω is locally on one side of Γ, and Q is defined by $Q = \Omega\times(0,T)$. Further, functions $y_T \in L^2(\Omega)^2$, $y_Q \in L^2(Q)^2$, and $y_0 \in H \subset L^2(\Omega)^2$ are given. The parameters γ and ν are positive real numbers. The bounds u_a, u_b are required to be in $L^2(Q)^2$ with $u_{a,i}(x,t) \le u_{b,i}(x,t)$ a.e. on Q, $i = 1, 2$.

The aim of this article is the presentation of a convergence theory of the SQP-method to solve the optimization problem (1.1). This method is widely applied to solve finite dimensional as well as function space optimization problems. The first convergence result in the context of optimal control was given in [24]. We will prove quadratic convergence of the SQP-method in a neighborhood of a reference control, which has to

[1]This work was supported by DFG SFB 557 "Control of complex turbulent shear flows" at TU Berlin.

[2]Institut für Mathematik, Technische Universität Berlin, Str. des 17. Juni 136, D-10623 Berlin, Germany.

fulfill a second-order sufficient optimality condition. We require positive definiteness of the second derivative of the Lagrangian on a subspace of inactive constraints.

The control of instationary Navier-Stokes flow has been studied very intensively since the pioneering work [1]. Necessary as well as sufficient optimality conditions were established, cf. [7, 8, 13, 14, 25]. The optimality system can be used to derive regularity properties of optimal controls. In [27], it was proven under certain assumptions that a locally optimal control of the problem (1.1) is a continuous function in space and time. Once a sufficient optimality condition holds true, one can prove stability of optimal controls under perturbations of the reference configuration. Here, one is interested in getting maximal stability of the controls, say stability with respect to the strongest possible norm. Using Hilbert space methods, one obtain stability of optimal controls in L^q with $q < \infty$. Since under some regularity assumptions an locally optimal control is continuous, one wants to get stability in the associated L^∞-norm. To this aim, a solution theory of the Navier-Stokes equations in L^p rather than Hilbert spaces is needed. If a stability result is available, one can prove local convergence of the SQP-method using the concept of generalized equations, [12, 18, 24]. In contrast to the approaches in the literature, we require that the second derivative of the Lagrangian is positive definite only on a subspace associated with strongly active constraints.

The outline of the paper is as follows. In Section 2, we will introduce some notation and state common results concerning solvability of the instationary Navier-Stokes system (1.2). Section 3 contains a brief overview of known facts about optimality conditions including first-order necessary and second-order sufficient conditions. In Section 4, the SQP-method is considered and its local convergence is proven. Numerical results confirming the convergence theory are presented in Section 5. Some regularity results for the linearized Navier-Stokes equation and the adjoint equation are summarized in Section 6. Throughout the article, we investigate the theory of optimal controls of the instationary Navier-Stokes equations in the L^p-space context.

2. Notations and preliminary results. Here, we will restrict ourselves to the two-dimensional case, $n = 2$. First, we introduce some notations and provide some results that we need later on.

To begin with, we define the solenoidal spaces

$$H_p := \{v \in L^p(\Omega)^2 : \operatorname{div} v = 0\}, \qquad V_p := \{v \in W_0^{1,p}(\Omega)^2 : \operatorname{div} v = 0\}.$$

Here, p denotes an arbitrary exponent $p \geq 2$. These spaces are Banach spaces with their norms denoted by $|\cdot|_p$ respectively $|\cdot|_{1,p}$. For $p = 2$, we get the frequently used solenoidal spaces $H := H_2$ and $V := V_2$, which are Hilbert spaces with scalar products $(\cdot,\cdot)_H$ respectively $(\cdot,\cdot)_V$. The dual of V with respect to the scalar product of H we denote by V' with the duality pairing $\langle\cdot,\,\cdot\rangle_{V',V}$.

We shall work in the standard space of abstract functions from $[0, T]$ to a real Banach space X, $L^p(0, T; X)$, endowed with its natural norm,

$$\|y\|_{L^p(X)} := \|y\|_{L^p(0,T;X)} = \left(\int_0^T |y(t)|_X^p \, dt\right)^{1/p} \qquad 1 \leq p < \infty,$$

$$\|y\|_{L^\infty(X)} := \operatorname*{vrai\,max}_{t \in (0,T)} |y(t)|_X.$$

In the sequel, we will identify the spaces $L^p(0, T; L^p(\Omega)^2)$ and $L^p(Q)^2$ for $1 < p < \infty$, and denote their norm by $\|u\|_p := |u|_{L^p(Q)^2}$. The usual $L^2(Q)^2$-scalar product we denote by $(\cdot,\cdot)_Q$ to avoid ambiguity.

In all what follows, $\|\cdot\|$ stands for norms of abstract functions, while $|\cdot|$ denotes norms of "stationary" spaces like H and V.

To deal with the time derivative in (1.2), we introduce the common spaces of functions y whose time derivatives y_t exist as abstract functions,

$$W^\alpha(0,T;V) := \{y \in L^2(0,T;V) :\ y_t \in L^\alpha(0,T;V')\}, \quad W(0,T) := W^2(0,T;V),$$

where $1 \le \alpha \le 2$. Endowed with the norm

$$\|y\|_{W^\alpha} := \|y\|_{W^\alpha(0,T;V)} = \|y\|_{L^2(V)} + \|y_t\|_{L^\alpha(V')},$$

these spaces are Banach spaces, respectively Hilbert spaces in the case of $W(0,T)$. Every function of $W(0,T)$ is, up to changes on sets of zero measure, equivalent to a function of $C([0,T],H)$, and the imbedding $W(0,T) \hookrightarrow C([0,T],H)$ is continuous, cf. [2, 17].

Furthermore, we introduce the following space of abstract functions in the L^p-context:

$$W_p^{2,1} := \{y \in L^p(0,T;W^{2,p}(\Omega)^2 \cap V_p) :\ y_t \in L^p(0,T;L^p(\Omega)^2)\},$$

which is continuously imbedded in $C([0,T],W_0^{2-2/p,\,p}(\Omega)^2)$, [16]. Here, $W_0^{2-2/p,\,p}(\Omega)^2$ denotes the space of solenoidal $W^{2-2/p,\,p}$-functions where zero boundary values are prescribed if $p \ge 4/3$. We abbreviate $H^{2,1} = W_2^{2,1}$ for $p = 2$. Note, that in this case we have $W_0^{2-2/2,\,2}(\Omega)^2 = V$. In this article, we will use exponents $p \ge 2$.

We define the trilinear form $b : V \times V \times V \mapsto \mathbb{R}$ by

$$b(u,v,w) = ((u\cdot\nabla)v, w)_2 = \int_\Omega \sum_{i,j=1}^{2} u_i \frac{\partial v_j}{\partial x_i}\ w_j\ dx.$$

To specify the problem setting, we introduce a linear operator $A : L^2(0,T;V) \mapsto L^2(0,T;V')$ by

$$\int_0^T \langle (Ay)(t),\, v(t)\rangle_{V',V}\, dt := \int_0^T (y(t), v(t))_V\, dt,$$

and a nonlinear operator B by

$$\int_0^T \big\langle \big(B(y)\big)(t),\, v(t)\big\rangle_{V',V}\, dt := \int_0^T b(y(t), y(t), v(t))dt.$$

B is continuous for instance as operator from $W(0,T)$ to $L^2(0,T;V')$. For convenience, we will use the notation

$$b_Q(y,v,w) = \int_0^T b(y(t), v(t), w(t))dt.$$

2.1. The state equation. We begin with the notation of weak solutions for the instationary Navier-Stokes equations (1.2) in the Hilbert space setting.

DEFINITION 2.1 (Weak solution). *Let $f \in L^2(0,T;V')$ and $y_0 \in H$ be given. A function $y \in L^2(0,T;V)$ with $y_t \in L^2(0,T;V')$ is called weak solution of* (1.2) *if*

$$\begin{aligned} y_t + \nu A y + B(y) &= f, \\ y(0) &= y_0. \end{aligned} \tag{2.1}$$

Results concerning the solvability of (2.1) are standard, cf. [22] for proofs and further details.

THEOREM 2.2 (Existence and uniqueness of solutions). *For every $f \in L^2(0,T;V')$ and $y_0 \in H$, the equation* (2.1) *has a unique solution $y \in W(0,T)$. Moreover, the mapping $(y_0, u) \mapsto y$ is locally Lipschitz continuous from $H \times L^2(0,T;V')$ to $W(0,T)$.*

For more regular data, one expects more regular solutions. The next theorem states some well-known facts, see for instance [22] for the details and further regularity results.

THEOREM 2.3 (Regularity). *For the higher regularity of the weak solutions of* (2.1) *the following holds. Let $y_0 \in V$ and $f \in L^2(Q)^2$ be given. Then the weak solution of* (2.1) *fulfills $y \in H^{2,1}$. The solution mapping $(f, y_0) \mapsto y$ is locally Lipschitz continuous between $L^2(Q)^2 \times V$ and $H^{2,1}$.*

For the proof we refer again to Temam [22].

Now, we want to specify the notation of a solution of (2.1) in the L^p-context.

DEFINITION 2.4 (Strong solution in L^p). *Let $f \in L^p(Q)^2$ and $y_0 \in W_0^{2-2/p,\,p}(\Omega)^2$ be given. A function $y \in W_p^{2,1}$ is called strong solution to the exponent $p > 2$ of* (1.2) *if there holds*

$$-\int_0^T (y, \phi')dt + \nu \int_0^T (\nabla y, \nabla \phi)dt + \int_0^T b(y,y,\phi) = \int_0^T (f,\phi)dt + (y_0, \phi(0)) \tag{2.2}$$

for all test functions $\phi \in L^q(0,T;V_q)$ with $\phi_t \in L^q(0,T;L^q(\Omega)^2)$ and $\phi(T) = 0$, where q is the dual exponent to p, $1/q + 1/p = 1$.

Here the space $W_0^{2-2/p,\,p}(\Omega)^2$ is the natural trace space. Every abstract function of $L^p(0,T;W^{2,p}(\Omega)^2)$ with time derivative in $L^p(0,T;L^p(\Omega)^2)$ is - after changes on a zero measure set - continuous with values in this space, [16]. Obviously, every strong L^p-solution is a weak solution. For existence of L^p-solutions we have the following theorem.

THEOREM 2.5 (L^p-solutions). *Let $f \in L^p(Q)^2$ and $y_0 \in W_0^{2-2/p,\,p}(\Omega)^2$ be given with $p \geq 2$. Then the weak solution y of* (2.1) *in the sense of Definition 2.1 is a strong solution and satisfies $y \in W_p^{2,1}$ There exists a constant $c > 0$ such that*

$$\|y\|_{W_p^{2,1}} \leq c\,\{|y_0|_{W^{2-2/p,\,p}} + \|f\|_p\}$$

Moreover, the mapping $(f, y_0) \mapsto y$ is locally Lipschitz continuous, hence the strong solution y is unique.

If $p = 2$ this result reduces to Theorem 2.3. For the non-Hilbert space case $p > 2$, it is proven in [28].

2.2. Differentiablity of the solution mapping. So far, we provided results concerning the properties of the state equation. We denote by $G(u) = y$ the solution operator $(f, y_0) \mapsto y$ of the instationary Navier-Stokes equations (2.1).

Lemma 2.6. *The solution operator $G : L^p(Q)^2 \times W_0^{2-2/p,\,p}(\Omega)^2 \mapsto W_p^{2,1}$ is Fréchet-differentiable. The derivative G' is given by $G'(f, y_0)(h_f, h_0) = z$, where z is the solution of*

$$\begin{aligned} z_t + \nu Az + B'(\bar{y})z &= h_f, \\ z(0) &= h_0. \end{aligned} \tag{2.3}$$

with $\bar{y} = G(f, y_0)$, $(h_f, h_0) \in L^p(Q)^2 \times W_0^{2-2/p,\,p}(\Omega)^2$.

Proof. Denote by $\bar{y}$ the state associated with (f, y_0) and by y_h the one associated with $(f + f_h, y_0 + h_0)$, i.e. $\bar{y} = G(f, y_0)$ and $y_h = G(f + f_h, y_0 + h_0)$. Since

$$B(\bar{y}) - B(y_h) = B'(\bar{y})(\bar{y} - y_h) - B(\bar{y} - y_h),$$

the difference $d := \bar{y} - y_h$ solves

$$\begin{aligned} d_t + \nu Ad + B'(\bar{y})d &= h_f + B(\bar{y} - y_h), \\ d(0) &= h_0. \end{aligned}$$

Next we split this difference into functions z and r, $d = z + r$, that solve the two linear equations

$$\begin{aligned} z_t + \nu Az + B'(\bar{y})z &= h_f, & r_t + \nu Ar + B'(\bar{y})r &= B(\bar{y} - y_h), \\ z(0) &= h_0, & r(0) &= 0. \end{aligned} \tag{2.4}$$

Existence and uniqueness of z and r follow from Lemma 6.5. Let us denote the solution operator of these linear equations by $\tilde{G}(\bar{y})$, then $z = \tilde{G}(\bar{y})(h_f, h_0)$. Clearly, this operator is linear. Its boundedness is a consequence of Lemma 6.5. We arrive at

$$\bar{y} - y_h - z = G(f, y_0) - G(f + f_h, y_0 + h_0) - \tilde{G}(\bar{y})(h_f, h_0) = r.$$

To prove Fréchet-differentiability of G, we have to estimate the norm of r. We begin with the estimation of the right-hand side of the system determining r, (2.4). Since $\bar{y}, y_h$ are in $W_p^{2,1}$ we obtain

$$\begin{aligned} \|B(\bar{y} - y_h)\|_p = \|(\bar{y} - y_h) \cdot \nabla(\bar{y} - y_h)\|_p &\le \|\bar{y} - y_h\|_\infty \|\bar{y} - y_h\|_{L^p(W^{1,p}(\Omega)^2)} \\ &\le c\,\|\bar{y} - y_h\|^2_{W_p^{2,1}}. \end{aligned}$$

By subsequent application of Lemma 6.5, the previous estimate of the Navier-Stokes nonlinearity B, and the Lipschitz continuity of the solution mapping G we obtain

$$\|r\|_{W_p^{2,1}} \le c\,\|B(\bar{y} - y_h)\|_p \le c\,\|\bar{y} - y_h\|^2_{W_p^{2,1}} \le c\,(\|h_f\|_p^2 + |h_0|^2_{W^{2-2/p,\,p}}).$$

Then it follows

$$\frac{|r|_{W_p^{2,1}}}{\|h_f\|_p + |h_0|_{W^{2-2/p,\,p}}} \to 0$$

as $\|h_f\|_p + |h_0|_{W^{2-2/p,\,p}} \to 0$. In this way, the Fréchet-differentiability of G is proven, and we can identify $G'(f, y_0) := \tilde{G}(\bar{y}) = \tilde{G}(G(f, y_0))$. ∎

3. Optimality conditions. Now we return to our optimal control problem. In this section, we discuss both necessary and sufficient optimality conditions connected with the optimal control problem (1.1).

3.1. First order necessary optimality conditions. We briefly recall the necessary conditions for local optimality. For the proofs and further discussion see [1, 5, 8, 13, 25] and the references cited therein.

Definition 3.1 (Locally optimal control). *A control $\bar{u} \in U_{ad}$ is said to be locally optimal in $L^2(Q)^2$, if there exists a constant $\rho > 0$ such that*

$$J(\bar{y}, \bar{u}) \leq J(y_\rho, u_\rho)$$

holds for all $u_\rho \in U_{ad}$ with $\|\bar{u} - u_\rho\|_2 \leq \rho$. Here, $\bar{y}$ and y_ρ denote the states associated with $\bar{u}$ and u_ρ, respectively.

In the following, we denote by $B'(\bar{y})^*$ the formal adjoint of $B'(\bar{y})$, given by

$$[B'(\bar{y})^*\lambda]v = \int_Q b(\bar{y}, v, \lambda) + b(v, \bar{y}, \lambda)dt.$$

Theorem 3.2 (Necessary condition). *Let $\bar{u}$ be a locally optimal control with associated state $\bar{y} = y(\bar{u})$. Then there exists a unique solution $\bar{\lambda} \in W^{4/3}(0,T;V)$ of the adjoint equation*

$$\begin{aligned} -\bar{\lambda}_t + \nu A\bar{\lambda} + B'(\bar{y})^*\bar{\lambda} &= \alpha_Q(\bar{y} - y_Q) + \alpha_R \vec{\operatorname{curl}} \operatorname{curl} \bar{y} \\ \bar{\lambda}(T) &= \alpha_T(\bar{y}(T) - y_T). \end{aligned} \tag{3.1}$$

Moreover, the variational inequality

$$(\gamma\bar{u} + \bar{\lambda}, u - \bar{u})_{L^2(Q)^2} \geq 0 \quad \forall u \in U_{ad} \tag{3.2}$$

is satisfied.

Proofs can be found in [8, 9, 25]. The regularity of $\bar{\lambda}$ is proven in [14].

The variational inequality (3.2) can be reformulated equivalently in different forms. At first, a pointwise discussion yields the projection representation of the optimal control

$$u_i(x,t) = \operatorname{Proj}_{[u_{a,i}(x,t), u_{b,i}(x,t)]}\left(-\frac{1}{\gamma}\bar{\lambda}_i(x,t)\right) \qquad \text{a.e. on } Q,\ i = 1,2. \tag{3.3}$$

With this formula, we can see that the optimal control inherits some regularity from the adjoint state. This form is also used in connection with Lipschitz stability of optimal controls, cf. [27].

Secondly, we introduce the normal cone $N_{U_{ad}}(\bar{u})$ of the set of admissible controls given by

$$N_{U_{ad}}(\bar{u}) = \begin{cases} \{z \in L^2(Q)^2 : (z, u - \bar{u})_2 \leq 0\ \forall u \in U_{ad}\} & \text{if } \bar{u} \in U_{ad} \\ \emptyset & \text{otherwise.} \end{cases} \tag{3.4}$$

Then the variational inequality (3.2) can be written equivalently as the inclusion

$$\nu\bar{u} + \bar{\lambda} + N_{U_{ad}}(\bar{u}) \ni 0. \tag{3.5}$$

This representation fits in the context of generalized equations, see for instance [12, 27].

The adjoint state λ is the solution of a linearized adjoint equation backward in time. So it is natural, to look for its dependence of the given data. For convenience, we denote by f the right-hand side of (3.1), and by λ_T the initial value $\alpha_T(\bar{y}(T) - y_T)$.

THEOREM 3.3 (Regularity of the adjoint state).

(i) Let $\lambda_T \in H$, $f \in L^2(0,T;V')$, and $\bar{y} \in L^2(0,T;V) \cap L^\infty(0,T;H)$ be given. Then there exists a unique weak solution λ of (3.1) satisfying $\lambda \in W^{4/3}(0,T)$.

(ii) Let $\lambda_T \in V$, $f \in L^2(Q)^2$, and $\bar{y} \in L^2(0,T;H^2(\Omega)^2) \cap L^\infty(0,T;V)$ be given. Then the unique weak solution λ of (3.1) is of class $H^{2,1}$.

The mapping $(f, \lambda_T) \mapsto \lambda$ is continuous in the mentioned spaces.

It can be proven following the lines of Temam [22], see also [14].

If the state y is a weak solution of (1.2) associated to a control $u \in L^2(Q)^2$, then Theorem 2.3 yields $y \in H^{2,1}$, and the pre-requisites are met, thus we get $\lambda \in H^{2,1}$.

The existence of L^p-solutions of the adjoint equation is topic of the next Theorem.

THEOREM 3.4. *Let $f \in L^p(Q)^2$ and $\lambda_T \in W_0^{2-2/p,p}(\Omega)^2$ be given with $p \geq 2$. If $\bar{y} \in L^p(0,T;W^{2,p}(\Omega)^2 \cap V_p)$, then the weak solution λ of (3.1) is a strong solution and satisfies $\lambda \in W_p^{2,1}$. Moreover, the mapping $(f, \lambda_T) \mapsto \lambda$ is continuous, hence the weak solution λ is unique.*

The result in the case $p = 2$ is equivalent to Theorem 3.3(ii). The proof for the case $p > 2$ is sketched in Section 6.3.

Let us introduce the Lagrange function $\mathcal{L} : W(0,T) \times L^2(Q)^2 \times W^{4/3}(0,T) \mapsto \mathbb{R}$ for the optimal control problem as follows:

$$\mathcal{L}(y,u,\lambda) = J(u,y) - \left\{ \langle y_t, \lambda \rangle_{L^2(V'),L^2(V)} + \nu(y,\lambda)_{L^2(V)} + b_Q(y,y,\lambda) - (u,\lambda)_Q \right\}.$$

This function is twice Fréchet-differentiable with respect to $(y,u) \in W(0,T) \times L^2(Q)^2$, cf. [25]. The reader can readily verify that the necessary conditions can be expressed equivalently by

$$\begin{aligned} \mathcal{L}_y(\bar{y},\bar{u},\bar{\lambda})\, h = 0 \quad \forall h \in W(0,T) \text{ with } h(0) = 0, \\ \mathcal{L}_u(\bar{y},\bar{u},\bar{\lambda})(u - \bar{u}) \geq 0 \quad \forall u \in U_{ad}. \end{aligned}$$

Here, $\mathcal{L}_y$, $\mathcal{L}_u$ denote the partial Fréchet-derivative of $\mathcal{L}$ with respect to y and u.

In the sequel we denote the pair of state and control (y,u) by v for convenience. The second derivative of the Lagrangian $\mathcal{L}$ at $y \in W(0,T)$ with associated adjoint state λ in the directions $v_1 = (w_1,h_1)$, $v_2 = (w_2,h_2) \in W(0,T) \times L^2(Q)^2$ is given by

$$\mathcal{L}_{vv}(y,u,\lambda)[v_1,v_2] = \mathcal{L}_{yy}(y,u,\lambda)[w_1,w_2] + \mathcal{L}_{uu}(y,u,\lambda)[h_1,h_2] \tag{3.6}$$

with

$$\begin{aligned} \mathcal{L}_{yy}(y,u,\lambda)[w_1,w_2] = \alpha_T(w_1(T),w_2(T))_H + \alpha_Q(w_1,w_2)_Q + \alpha_R(\operatorname{curl} w_1, \operatorname{curl} w_2)_Q \\ - b_Q(w_1,w_2,\lambda) - b_Q(w_2,w_1,\lambda) \end{aligned}$$

and

$$\mathcal{L}_{uu}(y,u,\lambda)[h_1,h_2] = \gamma(h_1,h_2)_2.$$

It satisfies the estimate

$$|\mathcal{L}_{yy}(y,u,\lambda)[w_1,w_2]| \le c\left(1+\|\lambda\|_{L^2(V)}\right)\|w_1\|_{W(0,T)}\|w_2\|_{W(0,T)} \tag{3.7}$$

for all $w_1, w_2 \in W(0,T)$, confer [25].

To shorten notations, we abbreviate $[v,v]$ by $[v]^2$, i.e.

$$\mathcal{L}_{vv}(\bar v,\bar\lambda)[(w,h)]^2 := \mathcal{L}_{vv}(\bar v,\bar\lambda)[(w,h),(w,h)].$$

3.2. Regularity of extremal controls. In the sequel, we will denote controls which satisfy the first-order necessary conditions as extremal controls. We will show that under the following assumptions every extremal control is a continuous function in space and time. More precisely, we assume for a p, $2 < p < \infty$, that the following pre-requisite holds

$$\textbf{(REG)}\begin{cases} u_a,\, u_b \in C(\bar Q)^2, \\ y_0 \in W_0^{2-2/p,\,p}(\Omega)^2, \\ \textit{Either } \alpha_T = 0 \textit{ or } y_T \in W_0^{2-2/p,\,p}(\Omega)^2, \\ \textit{Either } \alpha_Q = 0 \textit{ or } y_Q \in L^p(Q)^2. \end{cases}$$

THEOREM 3.5. *Let (REG) be satisfied. Every control u which fulfills the first-order necessary conditions is continuous, i.e. $u \in C(\bar Q)^2$.*

Proof. For a detailed discussion, we refer to [27]. We only sketch the proof. For every control $u \in L^2(\Omega)^2$, we get that the associated state y as well as the adjoint state λ belong to $H^{2,1}$. The space $H^{2,1}$ is imbedded in every $L^p(Q)^2$ for $p < \infty$, so the projection formula (3.3) gives $u \in L^p(Q)^2$. Now, we can apply the strong solvability result to conclude $y, \lambda \in W_p^{2,1}$, for $p > 2$. With imbedding arguments we find $\lambda \in C(\bar Q)^2$. A second application of (3.3) finally yields $u \in C(\bar Q)^2$. □

3.3. Second-order sufficient optimality condition. Let $\bar v := (\bar y, \bar u)$ be an admissible reference pair satisfying the first-order necessary optimality conditions.

DEFINITION 3.6 (Strongly active sets). *Let $\varepsilon > 0$ and $i \in \{1,2\}$ be given. Define sets $Q_{\varepsilon,i} \subseteq Q = \Omega \times [0,T]$ by*

$$Q_{\varepsilon,i} = \{(x,t) \in Q : |\gamma \bar u_i(x,t) + \bar\lambda_i(x,t)| > \varepsilon\}.$$

We assume further that the reference pair $\bar v = (\bar y, \bar u)$ satisfies the following coercivity assumption on $\mathcal{L}''(\bar v, \bar\lambda)$, in the sequel called second-order sufficient condition:

$$\textbf{(SSC)}\begin{cases} \textit{There exist } \varepsilon > 0 \textit{ and } \delta > 0 \textit{ such that} \\ \qquad \mathcal{L}_{vv}(\bar v,\bar\lambda)[(w,h)]^2 \ge \delta\,\|h\|_2^2 \\ \textit{holds for all pairs } (w,h) \in W(0,T) \times L^2(Q)^2 \textit{ with} \\ \qquad h = u - \bar u,\ u \in U_{ad},\ h_i = 0 \textit{ on } Q_{\varepsilon,i} \textit{ for } i = 1,2, \\ \textit{and } w \in W(0,T) \textit{ being the weak solution of the linearized equation} \\ \qquad \begin{aligned} w_t + Az + B'(\bar y)w &= h \\ w(0) &= 0. \end{aligned} \end{cases}$$

THEOREM 3.7. *Let $\bar{v} = (\bar{y}, \bar{u})$ be admissible for the optimal control problem and suppose that $\bar{v}$ fulfills the first-order necessary optimality condition with associated adjoint state $\bar{\lambda}$. Assume further that (SSC) is satisfied at $\bar{v}$. Then there exist $\alpha > 0$ and $\rho > 0$ such that*

$$J(v) \geq J(\bar{v}) + \alpha \|u - \bar{u}\|_2^2$$

holds for all admissible pairs $v = (y, u)$ with $\|u - \bar{u}\|_\infty \leq \rho$.

For a proof we refer to [25]. There, Theorem 3.7 was proven in a slightly general form: Sufficiency was achieved in a L^s-neighborhood of the reference control, whereas the quadratic growth takes place in the L^q-norm with $4/3 \leq q \leq 2 \leq s \leq \infty$, $s = q/(2-q)$.

We will show that any pair satisfying (SSC) is stable under perturbations of the system equations.

4. SQP-method.

in this section, we consider the SQP-method to compute a local optimum of the control problem (1.2). It is a well-known method, applied very often to optimal control problems of partial differential equations. For the analysis of other local methods in connection with instationary Navier-Stokes equations we refer to [14, 26].

The SQP-method solves in every step a linear-quadratic optimal control problem. Given starting values y_n, u_n, λ_n, it computes the next iterates $y_{n+1}, u_{n+1}, \lambda_{n+1}$ as the solution of

$$\min J^n(y,u) = \nabla J(y_n, u_n)(y - y_n, u - u_n) + \frac{1}{2}\mathcal{L}_{vv}(y_n, u_n, \lambda_n)[(y - y_n, u - u_n)]^2 \quad (P^n)$$

subject to the linearized state equation

$$\begin{aligned} y_t + \nu A y + B'(y_n)(y - y_n) &= u - B(y_n), \\ y(0) &= y_0, \end{aligned}$$

and the control constraint $u \in U_{ad}$. The functional to be minimized we write for convenience

$$\begin{aligned} J^n(y,u) = {} & \frac{\alpha_T}{2}\int_\Omega |y(x,T) - y_T(x)|^2 dx + \frac{\alpha_Q}{2}\int_Q |y(x,t) - y_Q(x,t)|^2 dx dt \\ & + \frac{\alpha_R}{2}\int_Q |\operatorname{curl} y(x,t)|^2 dx dt + \frac{\gamma}{2}\int_Q |u(x,t)|^2 dx dt - b_Q(y - y_n, y - y_n, \lambda_n). \end{aligned}$$

In the sequel, we investigate local convergence of this method. Here, the sufficient condition (SSC) plays an essential role. As one expects, we get quadratic convergence as soon as the iterates lies in neighborhood of a local solution.

4.1. Generalized Newtons method.

We want to show, that the SQP-method can be interpreted as a Newton-method for a generalized equation of the form

$$0 \in F(x) + N(x), \tag{4.1}$$

where F is a $C^{1,1}$-mapping between to Banach spaces X and Z, while $N : X \mapsto 2^Z$ is a set-valued mapping with closed graph.

One can write down the Newton-method formally as follows: given iterate x^n, compute the next iterate x^{n+1} by solving

$$0 \in F(x_n) + F'(x_n)(x - x_n) + N(x).$$

Before we state an abstract result concerning the convergence of the generalized Newton-method, we will introduce the notation of strong regularity in the sense of Robinson [18].

Let $\bar{x}$ be a solution of (4.1). The generalized equation is said to be *strongly regular* at the point $\bar{x}$, if there are open balls $B_X(\bar{x}, \rho_x)$ and $B_Z(0, \rho_z)$ such that for all $z \in B_Z(0, \rho_z)$ the linearized and perturbed equation

$$z \in F(\bar{x}) + F'(\bar{x})(x - \bar{x}) + N(x)$$

admits a unique solution $x = x(z)$ in $B_X(\bar{x}, \rho_x)$, and the mapping $z \mapsto x$ is Lipschitz continuous $B_Z(0, \rho_z)$ from to $B_X(\bar{x}, \rho_x)$.

Then the following theorem holds, which generalizes results from the finite-dimensional case.

THEOREM 4.1. *Let $\bar{x}$ be a solution of* (4.1) *and assume that* (4.1) *is strongly regular at $\bar{x}$. Then there exist an open ball $B_X(\bar{x}, \rho'_x)$ such that for every starting element $x_1 \in B_X(\bar{x}, \rho'_x)$ the generalized Newton method generates a unique sequence $\{x_n\}_{n=1}^{\infty}$. The iterates x_n remain in $B_X(\bar{x}, \rho'_x)$, and it holds*

$$\|x_{n+1} - \bar{x}\|_X \le c_N \, \|x_n - \bar{x}\|_X^2 \quad \forall n \in \mathbb{N}, \tag{4.2}$$

where c_N is independent of n.

For the proof, we refer to [3, 6].

In order to prove local convergence of the SQP-method,we have to verify the conditions of the previous theorem.

4.2. Strong regularity - L^∞-Stability of optimal controls. Let $(\bar{y}, \bar{u}, \bar{\lambda})$ satisfy the first-order necessary optimality conditions, see Theorem 3.2, together with the second-order sufficient optimality conditions (SSC). The optimality system consisting of state equation (1.2), adjoint equation (3.1) and the inclusion (3.5), can be written in the condensed form

$$F(\bar{y}, \bar{u}, \bar{\lambda}) + \left(0, 0, 0, 0, N_{U_{ad}}(\bar{u})\right)^T \ni 0, \tag{4.3}$$

where the function F,

$$\begin{aligned} F : W_p^{2,1} \times L^\infty(Q)^2 \times W_p^{2,1} \to \\ L^p(Q)^2 \times W_0^{2-2/p,\,p}(\Omega)^2 \times L^p(Q)^2 \times W_0^{2-2/p,\,p}(\Omega)^2 \times L^\infty(Q)^2, \end{aligned} \tag{4.4}$$

is given by

$$F(y, u, \lambda) = \begin{pmatrix} y_t + \nu A y + B(y) \\ y(0) \\ -\lambda_t + \nu A \lambda + B'(y)^* \lambda \\ \lambda(T) \\ \gamma u + \lambda \end{pmatrix} - \begin{pmatrix} u \\ y_0 \\ \alpha_Q (y - y_Q) + \alpha_R \vec{\operatorname{curl}} \operatorname{curl} y \\ \alpha_T (y(T) - y_T) \\ 0 \end{pmatrix}. \tag{4.5}$$

Further, we have to re-define the normal cone $N_{U_{ad}}$ to be a subset of $L^\infty(Q)^2$,

$$N_{U_{ad}} = \begin{cases} \{z \in L^\infty(Q)^2 : (z, u - \bar u)_2 \le 0 \ \forall u \in U_{ad}\} & \text{if } \bar u \in U_{ad} \\ \emptyset & \text{otherwise.} \end{cases}$$

We will apply Theorem 4.1 to the generalized equation (4.3). To do so, we have to show strong regularity of this equation at the reference tripel $(\bar y, \bar u, \bar\lambda)$. At first, we investigate the mapping F.

Corollary 4.2. *The function F defined by* (4.5) *is continuously differentiable in the setting* (4.4).

The proof can be found in [27].

The next and largest step is the investigation of the linearized and perturbed equation

$$z \in F(\bar y, \bar u, \bar\lambda) + F'(\bar y, \bar u, \bar\lambda)(y - \bar y, u - \bar u, \lambda - \bar\lambda) + \big(0, 0, 0, 0, N_{U_{ad}}(\bar u)\big)^T. \tag{4.6}$$

Here, the perturbation vector $z = (z_y, z_0, z_Q, z_T, z_u)$ is restricted to be in the space Z given by

$$Z := L^p(Q)^2 \times W_0^{2-2/p,\,p}(\Omega)^2 \times L^p(Q)^2 \times W_0^{2-2/p,\,p}(\Omega)^2 \times L^\infty(Q)^2. \tag{4.7}$$

We equip Z with the natural norm

$$\|z\|_Z = \|(z_y, z_0, z_Q, z_T, z_u)\|_Z := \|z_y\|_p + |z_0|_{W^{2-2/p,\,p}} + \|z_Q\|_p + |z_T|_{W^{2-2/p,\,p}} + \|z_u\|_\infty.$$

To prove strong regularity of (4.3), we have to consider the linearized and perturbed generalized equation (4.6). It represents a system, which can be written in a more convenient way. The first and second component form the state equations

$$\begin{aligned} y_t + \nu A y + B'(\bar y) y &= u + B(\bar y) + z_y \\ y(0) &= y_0 + z_0, \end{aligned}$$

the third and fourth component are equivalent to the adjoint equations

$$\begin{aligned} -\lambda_t + \nu A\lambda + B'(\bar y)^*\lambda &= -B'(y - \bar y)^*\bar\lambda + \alpha_Q(y - y_Q) + \alpha_R \vec{\operatorname{curl}} \operatorname{curl} y + z_Q \\ \lambda(T) &= \alpha_T(y(T) - y_T) + z_T, \end{aligned} \tag{4.8}$$

whereas the last component contains the inclusion

$$\gamma u + \lambda + N_{U_{ad}}(u) \ni z_u.$$

It builds up the optimality system of the following perturbed linear-quadratic optimization problem:

$$\begin{aligned} \min J^{(z)}(y, u) = \frac{\alpha_T}{2}|y(T) - y_d|_H^2 + \frac{\alpha_Q}{2}\|y - y_Q\|_2^2 + \frac{\alpha_R}{2}\|\operatorname{curl} y\|_2^2 + \frac{\gamma}{2}\|u\|_2^2 \\ + (z_Q, y)_Q + (z_T, y(T))_\Omega - (z_u, u)_Q - b_Q(y - \bar y, y - \bar y, \bar\lambda) \end{aligned} \tag{P_z}$$

subject to the linearized state equation

$$\begin{aligned} y_t + \nu A y + B'(\bar y) y &= u + B(\bar y) + z_y \\ y(0) &= y_0 + z_0 \end{aligned} \tag{4.9}$$

and the control constraint

$$u \in U_{ad}.$$

The existence of a unique optimal control of the problem (P_z) can not guaranteed by the coercivity assumption (SSC). There, positivity of $\mathcal{L}_{vv}$ was assumed only for the subspace of directions where the control $\bar{u}$ is not strong active. Hence this optimization problem can be non-convex in general.

We will circumvent this difficulty in the following way: At first, we show the existence of a unique solution if we substitute the control constraint by

$$u \in \widetilde{U_{ad}} = \{v \in U_{ad} : \; v_i(x,t) = \bar{u}_i(x,t) \text{ iff } (x,t) \in Q_{\varepsilon,i}\}. \qquad (\widetilde{P_z})$$

In the sequel, we will denote by $(\widetilde{P_z})$ the linear-quadratic optimization problem (P_z) with changed set of admissible controls $\widetilde{U_{ad}}$. This problem admits a unique solution, which we will denote by (y_z, u_z, λ_z). Moreover, the mapping $z \mapsto (y_z, u_z, \lambda_z)$ is Lipschitz from Z to $[L^\infty(Q)^2]^3$. Thus, for sufficient small perturbations the control u_z is strong active on $Q_{\varepsilon,i}$ as well as $\bar{u}$. This allows us to prove that it might fulfill a second-order sufficient optimality condition. Consequently, u_z is a locally optimal solution of (4.3). At this point, we refer to [24], where similar arguments were used.

For the solvability of $(\widetilde{P_z})$, we have the following

Theorem 4.3. *Let (SSC) be satisfied for the reference solution $\bar{v} = (\bar{y}, \bar{u})$ with adjoint state $\bar{\lambda}$. Moreover, assume that $y_0, y_T \in W_0^{2-2/p,p}(\Omega)^2$, $y_Q \in L^p(Q)^2$ for some p satisfying $2 < p < \infty$, and $u_a, u_b \in L^\infty(Q)^2$.*

Then problem $(\widetilde{P_z})$ admits a unique solution (y_z, u_z, λ_z). Moreover, the solution mapping $z \to (y_z, u_z, \lambda_z)$ is Lipschitz continuous from Z to $W_p^{2,1} \times L^\infty(Q)^2 \times W_p^{2,1}$.

Proof. The claim was proven in [27] for a modified problem. There the second-order sufficient condition was used in a stronger form. Coercivity of $\mathcal{L}_{vv}$ was required for the space of all directions - not only for the inactive ones.

However, we are encountering a similar situation. Let us denote the Lagrangian associated to (P_z) by $\mathcal{L}^{(z)}$. Then it holds for all y, u, λ

$$\mathcal{L}_{vv}^{(z)}(y, u, \lambda) = \mathcal{L}_{vv}(\bar{y}, \bar{u}, \bar{\lambda}). \qquad (4.10)$$

Hence, for $u \in \widetilde{U_{ad}}$ with associated y we find using (SSC)

$$\mathcal{L}_{vv}^{(z)}(y, u, \lambda)[(y - \bar{y}, u - \bar{u})]^2 = \mathcal{L}_{vv}(\bar{y}, \bar{u}, \bar{\lambda})[(y - \bar{y}, u - \bar{u})]^2 \geq \delta \|u - \bar{u}\|_2^2.$$

Thus, the problem $(\widetilde{P_z})$ is convex on the space of admissible controls $\widetilde{U_{ad}}$, which yields the existence of a unique optimal control u_z. For a more detailed discussion of those aspects we refer to [19], where the stability analysis is done for constrained optimal control of the stationary Navier-Stokes system. Following [27], we can prove also the claimed Lipschitz continuity of the solution mapping $z \mapsto (y_z, u_z, \lambda_z)$. ∎

Now, we study the behaviour of u_z on the active set Q_ε. To this aim, we have to rely on the L^∞-stability result of the previous theorem. One should remark, that using Hilbert-space methods it is not possible to derive such a result for the constrained optimal control problem of instationary Navier-Stokes equations, cf. [27] where this issue is addressed.

Corollary 4.4. *Let the assumptions of Theorem 4.3 be fulfilled. Then there exist $\rho_z > 0$ such that for all $z \in Z$ with $\|z\|_Z < \rho_z$ the optimal control of $(\widetilde{P_z})$, u_z, is strongly active a.e. on $Q_{\varepsilon,i}$, i.e. it holds*

$$|\gamma u_{z,i}(x,t) + \lambda_{z,i}(x,t) - z_{u,i}(x,t)| > \frac{\varepsilon}{2},$$

and the signs of $(\gamma \bar{u}_i(x,t) + \bar{\lambda}_i(x,t))$ and $(\gamma u_{z,i}(x,t) + \lambda_{z,i}(x,t) - z_{u,i}(x,t))$ coincide a.e. on $Q_{\varepsilon,i}$ for $i = 1, 2$.

Proof. By Theorem 4.3, the mapping $z \mapsto (y_z, u_z, \lambda_z)$ is Lipschitz continuous from Z to $W_p^{2,1} \times L^\infty(Q)^2 \times W_p^{2,1}$. By imbedding arguments, we find that $z \mapsto \gamma u_z + \lambda_z - z_u$ is Lipschitz as mapping to $L^\infty(Q)^2$.

Let $(x,t) \in Q_{\varepsilon,i}$ such that $\gamma \bar{u}_i(x,t) + \bar{\lambda}_i(x,t) > \varepsilon$. Using this, we derive

$$\begin{aligned}
\varepsilon &< \gamma \bar{u}_i(x,t) + \bar{\lambda}_i(x,t) \\
&= \gamma \bar{u}_i(x,t) + \bar{\lambda}_i(x,t) - (\gamma u_{z,i}(x,t) + \lambda_{z,i}(x,t) - z_{u,i}(x,t)) \\
&\quad + (\gamma u_{z,i}(x,t) + \lambda_{z,i}(x,t) - z_{u,i}(x,t)) \\
&\le c\|z\|_Z + \gamma u_{z,i}(x,t) + \lambda_{z,i}(x,t) - z_{u,i}(x,t).
\end{aligned}$$

Therefore, the choice $\rho_z := c^{-1}\varepsilon/2$ yields $\gamma u_{z,i}(x,t) + \lambda_{z,i}(x,t) - z_{u,i}(x,t) > \varepsilon/2$.

Analogously, if for $(x,t) \in Q_{\varepsilon,i}$ we have $\gamma \bar{u}_i(x,t) + \bar{\lambda}_i(x,t) < -\varepsilon$, then the same value of ρ_z gives $\gamma u_{z,i}(x,t) + \lambda_{z,i}(x,t) - z_{u,i}(x,t) < -\varepsilon/2$. ∎

Corollary 4.5. *Let the assumptions of Theorem 4.3 be fulfilled. Then the control u_z associated to a perturbation $z \in Z$ with $\|z\|_Z < \rho_z$, ρ_z given by the Corollary 4.4, fulfills the variational inequality*

$$(\gamma u_z + \lambda_z - z_u,\, u - u_z) \ge 0 \quad \forall u \in U_{ad}, \tag{4.11}$$

i.e. it satisfies the first-order necessary optimality condition of (P_z).

Proof. Let $u \in U_{ad}$ be given. We begin with

$$\begin{aligned}
\int_Q (\gamma u_{z,i} + \lambda_{z,i} - z_{u,i})(u_i - u_{z,i}) &= \int_{Q \setminus Q_{\varepsilon,i}} (\gamma u_{z,i} + \lambda_{z,i} - z_{u,i})(u_i - u_{z,i}) \\
&\quad + \int_{Q_{\varepsilon,i}} (\gamma u_{z,i} + \lambda_{z,i} - z_{u,i})(u_i - \bar{u}_i),
\end{aligned} \tag{4.12}$$

since $u_z \in \widetilde{U_{ad}}$ means $u_{z,i}(x,t) = \bar{u}_i(x,t)$ a.e. on $Q_{\varepsilon,i}$. The first integral is part of the first-order necessary optimality conditions of $(\widetilde{P_z})$. Therefore, it is nonnegative.

By Corollary 4.4, $(\gamma \bar{u}_i(x,t) + \bar{\lambda}_i(x,t))$ and $(\gamma u_{z,i}(x,t) + \lambda_{z,i}(x,t) - z_{u,i}(x,t))$ have the same sign a.e. on $Q_{\varepsilon,i}$. Furthermore, $\bar{u}_i(x,t)$ is active on this set, so that $u_i(x,t) - \bar{u}_i(x,t)$ has always the same sign for all possible choices of $u_i(x,t)$. Since $\gamma \bar{u} + \bar{\lambda}$ satisfies $\int_{Q_{\varepsilon,i}} (\gamma \bar{u}_i + \bar{\lambda}_i)(u_i - \bar{u}_i) \ge 0$, the same is true for $\gamma u_z + \lambda_z - z_u$, i.e.

$$\int_{Q_{\varepsilon,i}} (\gamma u_{z,i} + \lambda_{z,i} - z_{u,i})(u_i - \bar{u}_i) \ge 0$$

is satisfied. So, we proved that both integrals in (4.12) are nonnegative. Adding them, we derived the claim (4.11). ∎

So far, we showed that (y_z, u_z, λ_z) fulfills the optimality system of the perturbed problem (P_z) or equivalently the linearized and perturbed generalized equation (4.6). We have to ask whether it might be a local minimizer of (P_z). With the previous corollary and the identity (4.10) we have all ingredients at hand to prove that (y_z, u_z, λ_z) satisfies a second-order sufficient optimality condition for the problem (P_z), i.e. it is indeed a locally optimal solution.

Theorem 4.6. *Let the assumptions of Theorem 4.3 be fulfilled. Then, there are $\rho_z, \rho_u > 0$ such that the control u_z associated to a perturbation $z \in Z$ with $\|z\|_Z < \rho_z$ is a locally optimal solution of (P_z), and it satisfies*

$$J^{(z)}(y_z, u_z) \le J^{(z)}(y, u)$$

for all $u \in U_{ad}$ with $\|u - u_z\|_\infty \le \rho_u$. Here y_z and y are the solutions of (4.9) *associated to the controls u_z and u.*

Proof. We denote by y_z and λ_z the solutions of the state respectively adjoint equations (4.9) and (4.8). By Corollary 4.5, the triple (y_z, u_z, λ_z) satisfies not only the first-order necessary optimality conditions of $(\widetilde{P_z})$ but also the necessary optimality conditions of (P_z) if the norm of the perturbation z is smaller than ρ_z, $\|z\|_Z \le \rho_z$. This amounts in the fact that the control constraints are strongly active in u_z at Q_ε.

As already mentioned above, cf. (4.10), the second derivative of the Lagrangians $\mathcal{L}^{(z)}$ and $\mathcal{L}$ associated to (P_z) and (1.1) with respect to (y, u) coincide,

$$\mathcal{L}^{(z)}_{vv}(y, u, \lambda) = \mathcal{L}_{vv}(\bar{y}, \bar{u}, \bar{\lambda}).$$

Let $u \in U_{ad}$ with $u_i = \bar{u}_i$ a.e. on $Q_{\varepsilon,i}$ be given. Denote by y the associated solution of (4.9). Set $h = u - u_z$ and $w = y - y_z$. This implies $h = 0$ a.e. on $Q_{\varepsilon,i}$. Therefore, h fits in the assumptions of (SSC). The triple $(\bar{y}, \bar{u}, \bar{\lambda})$ satisfies the second-order sufficient optimality condition (SSC), which means

$$\mathcal{L}^{(z)}_{vv}(y, u, \lambda)[(w, h)]^2 = \mathcal{L}_{vv}(\bar{y}, \bar{u}, \bar{\lambda})[(w, h)]^2 \ge \delta \|h\|_2^2. \tag{4.13}$$

Both, Corollary 4.4 and the coercivity relation (4.13) build up the second-order sufficient optimality condition connected with $(\widetilde{P_z})$. Following the lines of [25], we conclude that u_z is locally optimal: there exists a constant $\rho_u > 0$ such $J^{(z)}(y_z, u_z) \le J^{(z)}(y, u)$ for all $u \in U_{ad}$ with $\|u - u_z\|_\infty \le \rho_u$. □

Corollary 4.7. *Let the assumptions of Theorem 4.3 be fulfilled. Then the generalized equation* (4.3) *is strongly regular at $(\bar{y}, \bar{u}, \bar{\lambda})$.*

Proof. By Corollary 4.2, the function F is a $C^{1,1}$-mapping. Theorem 4.6 states that the perturbed linearized optimization Problem (P_z) has a unique optimal solution in the ball $B_{L^\infty}(\bar{u}, \rho_u)$ for perturbations from $B_Z(0, \rho_z)$. By Theorem 4.3, the associated state y lies in the ball $B_{W_p^{2,1}}(\bar{y}, c_y \rho_z)$, whereas the adjoint state λ_z is in $B_{W_p^{2,1}}(\bar{\lambda}, c_\lambda \rho_z)$. Here, c_y and c_λ are the Lipschitz constants given by Theorem 4.3. This altogether yields the unique solvability of the perturbed linearized generalized equation (4.6) in $B_{W_p^{2,1}}(\bar{y}, c_y \rho_z) \times B_{L^\infty}(\bar{u}, \rho_u) \times B_{W_p^{2,1}}(\bar{\lambda}, c_\lambda \rho_z)$ for perturbations from $B_Z(0, \rho_z)$. As already mentioned, the solution mapping $z \mapsto (y_z, u_z, \lambda_z)$ is Lipschitz. Therefore, all requirements for strong regularity are fulfilled. □

4.3. Local convergence of the SQP-type algorithm. With the help of the previous section, we are in the situation to apply the abstract convergence result

Theorem 4.1. However, we are not allowed to carry it over one-to-one. The strong regularity of the generalized equation (4.1) requires that the solution mapping of the linerized perturbed problem (P_z) is Lipschitz continuous for small perturbations z, where the solutions (y_z, u_z, λ_z) are searched in a neighborhood of the reference triple $(\bar{y}, \bar{u}, \bar{\lambda})$.

Consequently, the SQP-method has to reflect this behaviour. Although the SQP-subproblems (P^n) are uniquely solvable in a *neighborhood of the reference solution* $\bar{u}$, they need not to be uniquely solvable on the whole set of admissible controls. In other words, the global solution of (P^n) may not coincide with the locally optimal solution. See also [24], where those aspects are discussed more detailed.

Hence, we have to modify the SQP-method in the following way: Given iterates y_n, u_n, λ_n, compute the next iterates $y_{n+1}, u_{n+1}, \lambda_{n+1}$ as the solution of (P^n) subject to the control constraint

$$u \in U_{ad}^{\rho} := U_{ad} \cap \{v \in L^\infty(Q)^2 : \|v - \bar{u}\|_\infty \leq \rho\}. \tag{4.14}$$

Then Theorem 4.1 yields quadratic convergence in a neighborhood of the solution.

THEOREM 4.8. *Let the asssumptions of Theorem 4.3 be satisfied. Then there is a constant $\rho_s > 0$, such that for every starting value (y_1, u_1, λ_1) with $u_1 \in U_{ad}^{\rho_s}$ the SQP-method with control constraint* (4.14) *generates a uniquely determined sequence (y_n, u_n, λ_n) with $u_n \in U_{ad}^{\rho_s}$, and it holds*

$$\|u_{n+1} - \bar{u}\|_\infty \leq c_s \, \|u_n - \bar{u}\|_\infty^2$$

with a constant c_s independently of n. Here, y_n and λ_n are the states and adjoints associated to the control u_n.

The a-priori unknown solution $\bar{u}$ appears in the definition of U_{ad}^{ρ}, which is necessary to establish the convergence theory. To overcome this difficulty, one has to use globalization techniques. For an application of a globalized SQP-method to compute optimal controls of instationary Navier-Stokes equations, we refer to [11].

5. Numerical results. Here, we provide a computational example which confirms the convergence analysis of the SQP-method.

The following control problem is given: We want to reduce the recirculation bubble after the backward-facing step. We try this by minimization of the objective functional

$$J(y,u) = \frac{1}{2} \int_{Q_c} |y(x,t) - y_Q(x,t)|^2 dxdt + \frac{\gamma}{2} \int_{Q_c} |u(x,t)|^2 dxdt,$$

where the time horizon T and the parameter γ will be varied in several examples. The computational domain Ω is the backward-facing step. Here, observation and control take place in the same part of the domain $Q_c = \Omega_c \times (0,T)$, compare Figure 5.1.

We chose as desired flow y_Q the Stokes flow, see Figure 5.2, which is the solution of the stationary Stokes equation with the same boundary conditions as used for the instationary simulation.

At the inflow boundary Γ_{in} a parabolic velocity profile is prescribed, whereas at the boundary Γ_{out} we use the 'do-nothing' boundary condition, cf. [10]:

$$\nu \frac{\partial y}{\partial n} - pn = 0 \quad \text{a.e. on } \Gamma_{out}.$$

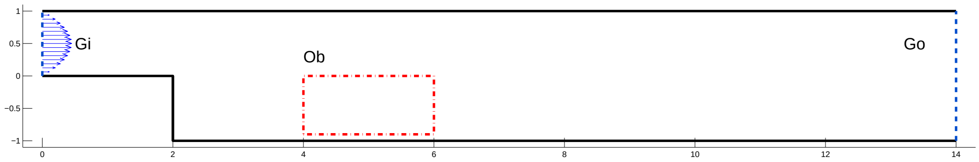

FIG. 5.1. *Flow configuration*

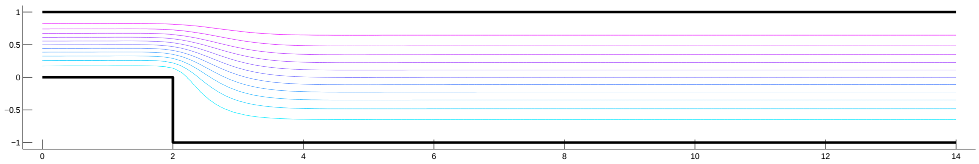

FIG. 5.2. *Desired profile is the Stokes flow*

At the rest of the boundary we use homogeneous Dirichlet conditions. All computations were done with Reynolds number $Re = 400$ which yields a viscosity parameter $\nu = 1/400$. The initial velocity profile were chosen as the stationary limit of the uncontrolled Navier-Stokes equations, cf. Figure 5.3.

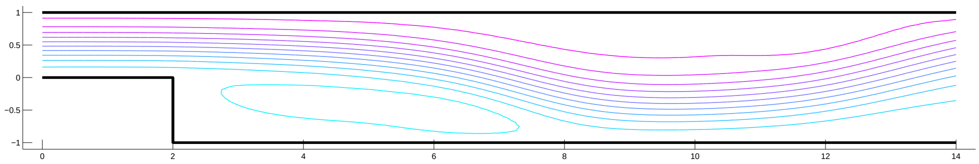

FIG. 5.3. *Initial flow profile* y_0

The continuous problem was discretized using Taylor-Hood finite elements on a grid of 1664 triangles with 3473 velocity and 905 pressure nodes. Further, we use a semi-implicit Euler scheme for time integration with a equidistant time discretization with step length $\tau = 0.005$. The computations are based on a finite element code of Michael Hinze, Dresden, see [13].

The arising discrete control problem is solved by the SQP-method without any globalization. The constraint SQP-subproblems (P^n) were solved by a primal-dual method, see for instance [15], using the CG method for the inner loop.

In all examples, the stopping criteria of the nested methods are balanced in the following way as proposed in [12]:

The outer SQP-loop was terminated if two successive iterates are close enough,

$$\|u^n - u^{n-1}\|_\infty + \|y^n - y^{n-1}\|_\infty \leq \varepsilon_{SQP}.$$

The primal-dual active set method was stopped if either the active sets of two successive control iterates coincide or the error in the variational inequality given by

$$\phi(u) = \left\| u - \mathrm{Proj}_{U_{ad}} \left(-\frac{1}{\gamma}\lambda \right) \right\|_2$$

is reduced by a factor of 0.1. The innermost iteration procedure — the CG method — was stopped if the norm of the residual was reduced by a factor of 0.01.

Example 1. In this first numerical example, we consider a situation which is nice from the optimization point of view: The parameter γ is set large enough so that the coercivity assumption (SSC) should be satisfied: $\gamma = 0.5$. And, we do not constrain the control, i.e. $U_{ad} = L^\infty(Q)^2$. Further, the time horizon was chosen to be $T = 0.25$. The starting values of the SQP-method for state and control, y^0 and u^0, were set to zero. The computed optimal objective was $J(\bar{y}, \bar{u}) = 0.02302$.

In Table 5.1, we present the results. We observe quadratical convergence of the SQP-algorithm in the iterations 1–3. In the third column, we give an estimation of the convergence speed of the SQP-algorithm by

$$q^n = \frac{\|u^n - u^{n-1}\|_\infty}{\|u^{n-1} - u^{n-2}\|_\infty^2}.$$

The iteration was stopped after iteration 3, since an adequate accuracy was achieved.

Iteration	$\|u^n - u^{n-1}\|_\infty$	q^n	$\|y^n - y^{n-1}\|_\infty$
1	$1.31 \cdot 10^{-1}$		$1.63 \cdot 10^{-2}$
2	$1.91 \cdot 10^{-2}$	1.10	$1.00 \cdot 10^{-3}$
3	$6.28 \cdot 10^{-5}$	0.17	$3.16 \cdot 10^{-6}$

TABLE 5.1. *Results of Example 1, unconstrained case*

We got similar results, if we require the control to fulfill box constraints

$$|u_i(x,t)| \le 0.05 \quad \text{a.e. on } Q,\ i = 1, 2.$$

In Table 5.2, the convergence history is presented. It turns out it that is independent of the additional constraints, although 16% of the constraints are active at the final solution.

Iteration	$\|u^n - u^{n-1}\|_\infty$	q^n	$\|y^n - y^{n-1}\|_\infty$
1	0.05		$1.63 \cdot 10^{-2}$
2	$1.90 \cdot 10^{-2}$	7.64	$9.64 \cdot 10^{-4}$
3	$6.36 \cdot 10^{-5}$	0.17	$2.87 \cdot 10^{-6}$

TABLE 5.2. *Results of Example 1, constrained case*

Example 2. Here, we extented the time horizon to $T = 1.3$. The regularization parameter γ was set to 0.1. Again, the controls has to fulfill box constraints

$$|u_i(x,t)| \le 0.3 \quad \text{a.e. on } Q,\ i = 1, 2.$$

We computed an optimal objective of $J(\bar{y}, \bar{u}) = 0.088234$. Some snapshots of the control can be seen in Figure 5.4. The streamlines of the terminal velocity $\bar{y}(T)$ field are depicted in Figure 5.5.

It seems that the behaviour of the SQP-method depends not only on γ but also on the time T. For this particular setting we encounter the situation that a larger value

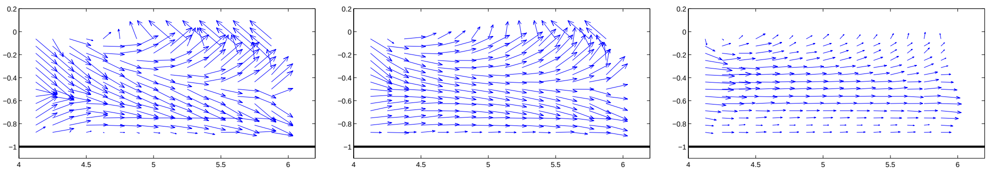

FIG. 5.4. *Control at time instances* $t_1 = 0.05$, $t_2 = 0.5$, $t_3 = T = 1.3$

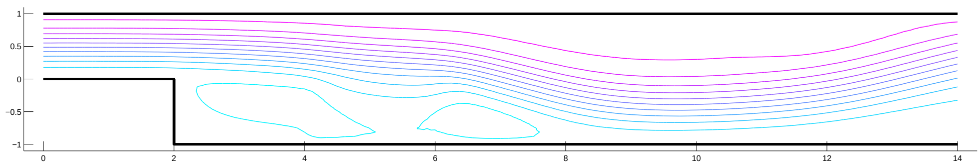

FIG. 5.5. *Flow profile* $\bar{y}(T)$

Iteration	$\|u^n - u^{n-1}\|_\infty$	q^n	$\|y^n - y^{n-1}\|_\infty$
1	$2.12 \cdot 10^{-1}$		$1.92 \cdot 10^{-1}$
2	$4.24 \cdot 10^{-1}$	9.42	$6.62 \cdot 10^{-2}$
3	$1.68 \cdot 10^{-1}$	0.93	$3.66 \cdot 10^{-2}$
4	$1.98 \cdot 10^{-2}$	0.70	$2.93 \cdot 10^{-3}$
5	$2.04 \cdot 10^{-3}$	5.21	$3.11 \cdot 10^{-4}$
6	$4.82 \cdot 10^{-5}$	11.5	$1.09 \cdot 10^{-5}$

TABLE 5.3. *Results of Example 2*

of T yields a worse convergence of the SQP-method. In Table 5.3, we listed the norm of the differences of the successive iterates.

We can explain the discrepancy between theory and numerical results in two different ways. First of all, we do not know whether the sufficient second-order optimality condition holds for the infinite-dimensional problem.

Secondly, the theory requires that the SQP-subproblems has to be solved to arbitray accuracy. However, in numerical computations this is limited by the discretization of the problem. So one can think of the situation that the discrete problem was solved to the possible accuracy before the SQP-method enters the region of quadratic convergence. This effect is reported by many authors who are dealing with optimal control of semilinear partial differential equations.

6. Proofs. In this very last section, we give an overview concerning the L^p-solution theory of the instationary Navier-Stokes equations together with its linearized and adjoint counterparts. At first, we state the fixed point theorem due to Schaefer. It is often used to prove existence of solutions of nonlinear partial differential equations.

THEOREM 6.1 (Schaefer [20]). *Let X be a normed space, H a continuous mapping of X into X which is compact on each bounded subset of X. Then either*

(i) the equation $x = \lambda H x$ has a solution for $\lambda = 1$, or
(ii) the set of all such solutions x for $0 < \lambda < 1$ is unbounded.

Secondly, consider the linear instationary Stokes system

$$\begin{aligned} y_t - \nu\Delta y + \nabla p &= f && \text{in } Q,\\ \operatorname{div} y &= 0 && \text{in } Q,\\ y(0) &= y_0 && \text{in } \Omega. \end{aligned} \tag{6.1}$$

Concerning L^p-solutions, the following result is due to Solonnikov [21] for the two- and three-dimensional case, in [28] it was generalized to arbitrary spatial dimensions. We rely in our analysis heavily on this fact.

THEOREM 6.2. *Let $p > 1$, $p \neq 3/2$, $y_0 \in W_0^{2-2/p,\,p}(\Omega)^2$, $f \in L^p(Q)^2$. Then there exists a unique weak solution y of (6.1) satisfying $y \in W_p^{2,1}$. Furthermore, there exists a constant $c > 0$ such that the estimate*

$$\|y_t\|_p + \|y\|_{L^p(W^{2,p})} \leq c\,\{\|f\|_p + |y_0|_{W^{2-2/p,\,p}}\}$$

is satisfied. Before proving the main result, we give some auxiliary lemmata which we will use in the sequel.

LEMMA 6.3. *Let for $p = 2$, $y_0 \in W_0^{2-2/p,\,p}(\Omega)^2 = V$ and $f \in L^p(Q)^2 = L^2(Q)^2$ be given. Consider the system*

$$\begin{aligned} y_t - \nu\Delta y + (y\cdot\nabla)y + \nabla p &= \sigma f && \textit{in } Q,\\ \operatorname{div} y &= 0 && \textit{in } Q,\\ y(0) &= \sigma y_0 && \textit{in } \Omega. \end{aligned} \tag{6.2}$$

with $\sigma \in [0,1]$. Denote y_σ be the associated L^2-solution of (6.2). Then it holds

$$\|y_{\sigma_1} - y_{\sigma_2}\|_{L^\infty(L^2)} \leq c|\sigma_1 - \sigma_2|\,\{\|f\|_2 + |y_0|_V\} \quad \forall \sigma_1, \sigma_2 \in [0,1].$$

Proof. The claim follows directly from the Lipschitz continuity of the solution mapping of the instationary Navier-Stokes equations, cf. Theorems 2.2 and 2.3. ∎

We need the following L^p-estimate of the nonlinearity.

LEMMA 6.4. *Let $p > 1$, $u \in W^{2,p}$. Then $(u\cdot\nabla)u \in L^p$ and*

$$|(u\cdot\nabla)u|_p \leq c\left\{\|\,|u|^3\,\|_p + \|\,|\nabla u|^{3/2}\,\|_p\right\} \leq c|u|_{W^{2,p}}\left\{|u|_2^2 + |u|_2^{1/2}\right\}.$$

For a proof, we refer to v.Wahl [28].

6.1. Proof of Theorem 2.5. We will give the proof in three steps. Although the method of proof is the same as in [28], we present it in a slightly modified form.

6.1.1. Mapping $\mathcal{T}$. For given $\sigma \in [0,1]$, we define a mapping $\mathcal{T} = \mathcal{T} : W_p^{2,1} \to W_p^{2,1}$ as $\mathcal{T}w = y$, where y is the strong solution of

$$\begin{aligned} y_t - \nu\Delta y + (w\cdot\nabla)w + \nabla p &= \sigma f && \text{in } Q,\\ \operatorname{div} y &= 0 && \text{in } Q,\\ y(0) &= \sigma y_0 && \text{in } \Omega. \end{aligned} \tag{6.3}$$

The imbedding $W_p^{2,1} \hookrightarrow C([0,T],C(\Omega))^2$ is compact for $p>2$, cf. [4]. Following [28], we find that $w \to (w\cdot\nabla)w$ is compact from $W_p^{2,1}$ to $L^p(\Omega)^2$. Now, Theorem 6.2 yields the compactness of $\mathcal{T}$.

Obviously, every fixed point of $\mathcal{T}$ for $\sigma = 1$ is a strong solution of (1.2). To apply Schaefers fixed point theorem, we have to show that the fixed points of $\tau\mathcal{T}$ for $0 < \tau < 1$ remain bounded in $W_p^{2,1}$.

6.1.2. Boundedness of fixed points. Define the set Σ as the set of all $\sigma \in [0,1]$ for which a fixed point of $\mathcal{T}$ exists. Since $\mathcal{T}$ is a mapping on $W_p^{2,1}$, the associated state y_σ is a member of that space, $y_\sigma \in W_p^{2,1}$.

Define further a set Σ^* as the set of all $\sigma \in [0,1]$ such that

$$[0,\sigma] \subset \Sigma,$$

and there exists a constant C_σ, so that the set $\{y_s\}_{s\in[0,\sigma]}$ is uniformly bounded in the following sense

$$\|y_{s,t}\|_p + \|y_s\|_{L^p(W^{2,p})} + \|\,|y_s|^3\,\|_p + \|\,|\nabla y_s|^{3/2}\,\|_p \le C_\sigma \left\{\|f\|_p + |y_0|_{W^{2-2/p,\,p}}\right\} \quad (6.4)$$

for all $s \in [0,\sigma]$. Therefore, the set Σ^* contains all those σ whose associated fixed points of $\mathcal{T}$ exists and are bounded in $W_p^{2,1}$. Clearly $0 \in \Sigma^*$.

We will show that for every $\sigma_1 \in \Sigma^*$ there exists $\delta > 0$ independently of σ_1 such that the whole interval $[\sigma_1, \sigma_1+\delta]$ belongs also to Σ^*.

To this aim, let $\sigma_1 \in \Sigma^*$, $\sigma_2 \in \Sigma$ with associated states y_{σ_i}. Their difference $y_{\sigma_1} - y_{\sigma_2}$ is the solution of an instationary Stokes equation with suitable right-hand side. By Theorem 6.2 we have

$$\begin{aligned}\|y_{\sigma_1,t} - y_{\sigma_2,t}\|_p + \|y_{\sigma_2} - y_{\sigma_1}\|_{L^p(W^{2,p})} \le C_1\Big\{&|\sigma_1-\sigma_2|\cdot\|f\|_p + |\sigma_1-\sigma_2|\cdot|y_0|_{W^{2-2/p,\,p}}\\ &+ \|(y_{\sigma_2}\cdot\nabla)y_{\sigma_2} - (y_{\sigma_1}\cdot\nabla)y_{\sigma_1}\|_p\Big\}. \quad (6.5)\end{aligned}$$

Furthermore, Lemma 6.3 gives

$$\|y_{\sigma_1} - y_{\sigma_2}\|_{L^\infty(L^2)} \le c|\sigma_1-\sigma_2|\left\{\|f\|_2 + |y_0|_V\right\}.$$

The nonlinear term is treated as in Lemma 6.4

$$\begin{aligned}&\|(y_{\sigma_2}\cdot\nabla)y_{\sigma_2} - (y_{\sigma_1}\cdot\nabla)y_{\sigma_1}\|_p \le \|(y_{\sigma_2}\cdot\nabla)y_{\sigma_2}\|_p + \|(y_{\sigma_1}\cdot\nabla)y_{\sigma_1}\|_p\\ &\le c\Big\{\|\,|y_{\sigma_2}|^3\,\|_p + \|\,|\nabla y_{\sigma_2}|^{3/2}\,\|_p + \|\,|y_{\sigma_1}|^3\,\|_p + \|\,|\nabla y_{\sigma_1}|^{3/2}\,\|_p\Big\}\\ &\le c\Big\{\|\,|y_{\sigma_2}-y_{\sigma_1}|^3\,\|_p + \|\,|\nabla(y_{\sigma_2}-y_{\sigma_1})|^{3/2}\,\|_p + \|\,|y_{\sigma_1}|^3\,\|_p + \|\,|\nabla y_{\sigma_1}|^{3/2}\,\|_p\Big\}\\ &\le C_2\left(\|y_{\sigma_2}-y_{\sigma_1}\|_2^2 + \|y_{\sigma_2}-y_{\sigma_1}\|_2^{1/2}\right)\|y_{\sigma_2}-y_{\sigma_1}\|_{L^p(W^{2,p})}\\ &\quad + C_3\left\{\|\,|y_{\sigma_1}|^3\,\|_p + \|\,|\nabla y_{\sigma_1}|^{3/2}\,\|_p\right\}\end{aligned} \quad (6.6)$$

For $|\sigma_1-\sigma_2| \le \delta$ the term $C_2\left(\|y_{\sigma_2}-y_{\sigma_1}\|_2^2 + \|y_{\sigma_2}-y_{\sigma_1}\|_2^{1/2}\right)$ is less than $1/(2C_1)$. Note, C_1, C_2, and δ are independent of σ_1 and σ_2. We thus found combining (6.5) and (6.6),

$$\begin{aligned}\|y_{\sigma_1,t} - y_{\sigma_2,t}\|_p + \|y_{\sigma_2} - y_{\sigma_1}\|_{L^p(W^{2,p})} \\ \leq 2C_1\delta\big\{\|f\|_p + |y_0|_{W^{2,p}}\big\} + 2C_1C_3\left\{\|\,|y_{\sigma_1}|^3\,\|_p + \|\,|\nabla y_{\sigma_1}|^{3/2}\,\|_p\right\} \\ \leq 2C_1(\delta + C_3C_{\sigma_1})\,\big\{\|f\|_p + |y_0|_{W^{2-2/p,\,p}}\big\}.\end{aligned}$$

Additionally, we derive

$$\begin{aligned}\|y_{\sigma_2,t}\|_p + \|y_{\sigma_2}\|_{L^p(W^{2,p})} \\ \leq \|y_{\sigma_1,t}\|_p + \|y_{\sigma_1}\|_{L^p(W^{2,p})} + \|y_{\sigma_1,t} - y_{\sigma_2,t}\|_p + \|y_{\sigma_2} - y_{\sigma_1}\|_{L^p(W^{2,p})} \\ \leq (C_{\sigma_1} + 2C_1(\delta + C_3C_{\sigma_1}))\big\{\|f\|_p + |y_0|_{W^{2-2/p,\,p}}\big\}.\end{aligned}$$

Estimate (6.6) implies

$$\begin{aligned}\|\,|y_{\sigma_2}|^3\,\|_p + \|\,|\nabla y_{\sigma_2}|^{3/2}\,\|_p \\ \leq C_2\left(\|y_{\sigma_2} - y_{\sigma_1}\|_2^2 + \|y_{\sigma_2} - y_{\sigma_1}\|_2^{1/2}\right)\|y_{\sigma_2} - y_{\sigma_1}\|_{L^p(W^{2,p})} \\ + C_3\left\{\|\,|y_{\sigma_1}|^3\,\|_p + \|\,|\nabla y_{\sigma_1}|^{3/2}\,\|_p\right\} \\ \leq \frac{1}{2C_1}\|y_{\sigma_2} - y_{\sigma_1}\|_{L^p(W^{2,p})} + C_3\left\{\|\,|y_{\sigma_1}|^3\,\|_p + \|\,|\nabla y_{\sigma_1}|^{3/2}\,\|_p\right\} \\ \leq (\delta + 2C_3C_{\sigma_1})\big\{\|f\|_p + |y_0|_{W^{2-2/p,\,p}}\big\}.\end{aligned}$$

Setting $C_{\sigma_1+\delta} = \max\{C_{\sigma_1} + 2C_1(\delta + C_3C_{\sigma_1}),\, \delta + 2C_3C_{\sigma_1}\}$, we find that for all $s \in [\sigma_1, \sigma_1 + \delta] \cap \Sigma$

$$\|y_{s,t}\|_p + \|y_s\|_{L^p(W^{2,p})} + \|\,|y_s|^3\,\|_p + \|\,|\nabla y_s|^{3/2}\,\|_p \leq C_{\sigma_1+\delta}\,\{\|f\|_p + |y_0|_{W^{2-2/p,\,p}}\} \quad (6.7)$$

is satisfied. Hence, if for $\sigma \in [\sigma_1, \sigma_1 + \delta]$ a fixed point of $\tau\mathcal{T}$, $0 < \tau < 1$, exists then it is necessarily bounded. Thus, Schaefers fixed point Theorem 6.1 yields the *existence* of an fixed point of $\mathcal{T}$ for every $\sigma \in [\sigma_1, \sigma_1 + \delta]$, which means actually $[\sigma_1, \sigma_1 + \delta] \subset \Sigma$. And, $[\sigma_1,\, \sigma_1 + \delta] \subset \Sigma^*$ is proven.

Since δ was independent of σ_1, we obtain after finite many steps $\Sigma = \Sigma^* = [0, 1]$, which proves existence of strong solutions of (1.2).

6.1.3. Conclusion. The fact that the solution mapping $(f, y_0) \mapsto y$ is bounded from $L^p(Q)^2 \times W_0^{2-2/p,\,p}(\Omega)^2 \mapsto W_p^{2,1}$ is a consequence of (6.7). The local Lipschitz continuity is a further implication of that Lemma, see Remark 6.6 below. □

6.2. Linearized equation. Let a function $\bar{y} \in W_p^{2,1}$ be given. Consider the system

$$\begin{aligned} y_t + \nu Ay + B'(\bar{y})y &= f, \\ y(0) &= y_0. \end{aligned} \quad (6.8)$$

We will show, that this system is uniquely solvable.

Lemma 6.5. *Let $\bar{y} \in W_p^{2,1}$, $f \in L^p(Q)^2$ and $y_0 \in W_0^{2-2/p,\,p}(\Omega)^2$ be given with $2 \leq p < \infty$. Then the system* (6.8) *has a unique solution $y \in W_p^{2,1}$. Moreover, there is a constant $c > 0$ independently of f and y_0 such that the following estimate is true*

$$\|y\|_{W_p^{2,1}} \leq c\,\{\|f\|_p + |y_0|_{W^{2-2/p,\,p}}\}\,. \quad (6.9)$$

Proof. The proof is carried out using bootstrapping arguments.

Step 1: $p = 2$. At first, notice that this result for $p = 2$ was proven for instance in [14]. It yields the regularity $y \in W_2^{2,1} = H^{2,1}$ and the existence of a constant $c > 0$ such that

$$\|y\|_{H^{2,1}} \le c\,\{\|f\|_2 + |y_0|_V\}$$

is fulfilled.

We consider equation (6.8) as an instationary Stokes equation,

$$\begin{aligned} y_t + \nu A y &= f - B'(\bar y)y, \\ y(0) &= y_0. \end{aligned} \tag{6.10}$$

to invoke Solonnikovs Theorem 6.2. To this aim, we have to investigate the $p-$norm auf $B'(\bar y)y = (y\cdot\nabla)\bar y + (\bar y \cdot \nabla)y$.

Step 2: $2 < p < 4$. We assume $\bar y \in W_p^{2,1}$. Then it follows that $\bar y$ is a function of class $L^\infty(0,T;W_0^{2-2/p,\,p}(\Omega)^2)$. Further, let $y \in W_2^{2,1}$ be the (weak) L^2-solution of (6.10).

At first, observe that $W_0^{2-2/p,\,p}(\Omega)^2$ is continuously imbedded in $W^{1,q}$ for $q = \frac{2p}{4-p}$, cf. [2, 23]. By $\frac{1}{p} = \frac{1}{q} + \frac{1}{q'}$, applying Hölders inequality and the imbedding $V \hookrightarrow L^{q'}$ for $q' < \infty$, we obtain

$$\|(y\cdot\nabla)\bar y\|_p \le c\|\bar y\|_{L^\infty(W^{1,q})}\|y\|_{L^\infty(L^{q'})} \le c\|\bar y\|_{L^\infty(W^{2-2/p,\,p})}|y|_{L^\infty(V)}. \tag{6.11}$$

Secondly, if $y \in L^p(W^{1,p})$ holds, we can derive

$$\|(\bar y\cdot\nabla)y\|_p \le c\|\bar y\|_\infty \|y\|_{L^p(W^{1,p})}. \tag{6.12}$$

Using the imbedding $W^{2-2/p,\,2} \hookrightarrow W^{1,p}$ and the interpolation identity $[W^{2,2}, W^{1,2}]_\theta = W^{2-2/p,\,2}$ with $\theta = 1 - 2/p$, we obtain

$$|y|^p_{W^{1,p}} \le c|y|^p_{W^{2-2/p,\,2}} \le c|y|^{p-2}_{W^{2,2}}|y|^2_{W^{1,2}}.$$

Integrating with respect to the time variable yields

$$\|y\|^p_{L^p(W^{1,p})} \le c\|y\|^{p-2}_{L^{p-2}(W^{2,2})}\|y\|^2_{L^\infty(V)} \le c\|y\|^{p-2}_{L^2(W^{2,2})}\|y\|^2_{L^\infty(V)} \le c\|y\|^p_{W_2^{2,1}} \tag{6.13}$$

provided $p \le 4$. Collecting (6.11)–(6.13),

$$\|B'(\bar y)\|_p \le c\|\bar y\|_{W_p^{2,1}}\|y\|_{W_2^{2,1}} \le c\|\bar y\|_{W_p^{2,1}}\,\{\|f\|_2 + |y_0|_V\}$$

is found. Now, we can utilize Solonnikovs result to obtain the existence of a strong solution $\tilde y$ of (6.10). Thus, $\tilde y$ is also a weak solution. Since the weak solution is unique it follows $\tilde y = y$, remember y was by definition the weak solution of (6.10). Moreover, the solution estimate

$$\begin{aligned} \|y\|_{W_p^{2,1}} &\le c\,\{\|f\|_p + |y_0|_{W^{2-2/p,\,p}} + \|B'(\bar y)\|_p\} \\ &\le c\,\{\|f\|_p + |y_0|_{W^{2-2/p,\,p}}\} + c\|\bar y\|_{W_p^{2,1}}\,\{\|f\|_2 + |y_0|_V\} \\ &\le c(1 + \|\bar y\|_{W_p^{2,1}})\,\{\|f\|_p + |y_0|_{W^{2-2/p,\,p}}\} \end{aligned}$$

is satisfied.

STEP 3: $4 \le p < \infty$. Let $\bar y \in W_p^{2,1}$. By Step 2, we find the strong solution y in $W_{4-\varepsilon}^{2,1}$, $0 < \varepsilon \le 2$. It is - after changes on a zero measure set - continuous with values in $W_0^{2-2/(4-\varepsilon),4-\varepsilon}(\Omega)^2$. Further, we the space $W_{4-\varepsilon}^{2,1}$ is continuously imbedded $L^\infty(Q)^2$. Again, we have to estimate the p-norm of $B'(\bar y)y$. We begin with

$$\|(y\cdot\nabla)\bar y\|_p \le c\|y\|_\infty\|\nabla\bar y\|_p \le c\|y\|_{W_{4-\varepsilon}^{2,1}}\|\bar y\|_{W_p^{2,1}}. \tag{6.14}$$

To estimate the second addend of $B'(\bar y)y$, we observe that for $p = \frac{8}{\varepsilon} - 2$ respectively $\varepsilon = \frac{8}{p+2}$ the imbedding

$$W_0^{2-\frac{2}{4-\varepsilon},4-\varepsilon}(\Omega)^2 = W_0^{\frac{3}{2}-\frac{1}{p},\frac{4p}{p+2}}(\Omega)^2 \hookrightarrow W_0^{1,p}(\Omega)^2$$

is continuous. Moreover, for this choice of p respectively ε we obtain

$$y \in W_{4-\varepsilon}^{2,1} \hookrightarrow L^\infty(0,T;W_0^{1,p}(\Omega)^2).$$

Hence, we arrive at

$$\|(\bar y\cdot\nabla)y\|_p \le c\|\bar y\|_\infty\|y\|_{L^\infty(W^{1,p})} \le c\|\bar y\|_{W_p^{2,1}}\|y\|_{W_{4-\varepsilon}^{2,1}}, \tag{6.15}$$

which allows as to conclude by Solonnikovs Theorem

$$\|y\|_{W_p^{2,1}} \le c(1+\|\bar y\|_{W_p^{2,1}})\left\{\|f\|_p + |y_0|_{W^{2-2/p,\,p}}\right\},$$

and the claim is proven for all p in $[2,\infty)$. □

REMARK 6.6. *The Lipschitz continuity of the solution mapping of the instationary Navier-Stokes equations can be proven using the previous Lemma. Let data $f_i \in L^p(Q)^2$ and $y_{0,i} \in W_0^{2-2/p,\,p}(\Omega)^2$ be given, $i = 1,2$. Denote the associated strong solutions by y_i, $i = 1,2$. Then the difference $d := y_1 - y_2$ satisfies*

$$\begin{aligned} d_t + \nu Ad + (y_1\cdot\nabla)d + (d\cdot\nabla)y_2 &= f_1 - f_2,\\ d(0) &= y_{0,1} - y_{0,2}. \end{aligned}$$

With analogous arguments as above, we arrive at

$$\|y_1 - y_2\|_{W_p^{2,1}} \le c(1+\|y_1\|_{W_p^{2,1}} + \|y_2\|_{W_p^{2,1}})\left\{\|f_1 - f_2\|_p + |y_{0,1} - y_{0,2}|_{W^{2-2/p,\,p}}\right\},$$

which is the claimed Lipschitz continuity.

6.3. Adjoint equation - Proof of Theorem 3.4. At last, we are going to prove the existence of a strong solution of the adjoint equation

$$\begin{aligned} -\lambda_t + \nu A\lambda + B'(\bar y)^*\lambda &= f\\ \lambda(T) &= \lambda_T. \end{aligned} \tag{6.16}$$

Via the transformations $w(t) = \lambda(T-t)$, $\hat y(t) = \bar y(T-t)$, $g(t) = f(T-t)$, $w_0 = \lambda_T$, this system is carried over in the forward-in-time equation

$$\begin{aligned} w_t + \nu Aw + B'(\hat y)^*w &= g\\ w(0) &= w_0. \end{aligned} \tag{6.17}$$

It is obvious that $\hat{y}, g, w_0$ inherited their regularity from $\bar{y}, f, \lambda_T$. Hence, the adjoint state λ has the same regularity as w.

LEMMA 6.7. *Let* $\hat{y} \in W_p^{2,1}$, $g \in L^p(Q)^2$ *and* $w_0 \in W_0^{2-2/p,p}(\Omega)^2$ *be given with* $2 \le p < \infty$. *Then the system* (6.16) *has a unique solution* $w \in W_p^{2,1}$. *Moreover, there is a constant* $c > 0$ *independently of* g *and* w_0 *such that the following estimate is true*

$$\|w\|_{W_p^{2,1}} \le c\,\{\|g\|_p + |w_0|_{W^{2-2/p,\,p}}\}\,.$$

Proof. The proof is very similar to the proof of Lemma 6.5. Therefore, we will briefly repeat its steps. At first, we observe

$$\begin{aligned}[B'(\hat{y})w]v = \int_0^T b(\hat{y}, v, w) + b(v, \hat{y}, w)dt &= \int_0^T -b(\hat{y}, w, v) + b(v, \hat{y}, w)dt \\ &= \int_0^T [-(\hat{y}\cdot\nabla)w + (\nabla\hat{y})^T w]\cdot v\,dt,\end{aligned}$$

since $\hat{y}$ is divergence-free.

STEP 1: $p = 2$. The result for $p = 2$ was proven for instance in [14]. It yields the regularity $w \in W_2^{2,1} = H^{2,1}$ and the existence of a constant $c > 0$ such that

$$\|w\|_{H^{2,1}} \le c\,\{\|g\|_2 + |w_0|_V\}$$

is satisfied.

STEP 2: $2 < p < 4$. With the help of (6.11)-(6.13), we conclude

$$\|B'(\hat{y})\|_p = \|-(\hat{y}\cdot\nabla)w+(\nabla\hat{y})^T w\|_p \le c\left\{\|\hat{y}\|_\infty\|w\|_{W_2^{2,1}} + \|\hat{y}\|_{L^\infty(W^{2-2/p,\,p})}|w|_{L^\infty(V)}\right\}.$$

Then Solonnikovs Theorem gives us the existence of a bounded strong solution in $W_p^{2,1}$.

STEP 3: $4 \le p < \infty$. Let $w \in W_{4-\varepsilon}^{2,1}$ be the strong solution of Step 2, $0 < \varepsilon < 2$. Analogously as in (6.14) and (6.15) we find

$$\|B'(\hat{y})\|_p = \|-(\hat{y}\cdot\nabla)w + (\nabla\hat{y})^T w\|_p \le c\|\hat{y}\|_{W_p^{2,1}}\|w\|_{W_{4-\varepsilon}^{2,1}},$$

and the claim follows immediately. ∎

REFERENCES

[1] F. Abergel and R. Temam. On some control problems in fluid mechanics. *Theoret. Comput. Fluid Dynam.*, 1:303–325, 1990.

[2] R. A. Adams. *Sobolev spaces.* Academic Press, San Diego, 1978.

[3] W. Alt. The Lagrange-Newton method for infinite dimensional optimization problems. *Numerical Functional Analysis and Optimization*, 11:201–224, 1990.

[4] H. Amann. Compact embeddings of vector-valued Sobolev and Besov spaces. *Glasnik Matematicki*, 35:161–177, 2000.

[5] E. Casas. An optimal control problem governed by the evolution Navier-Stokes equations. In S. S. Sritharan, editor, *Optimal control of viscous flows*, Frontiers in applied mathematics. SIAM, Philadelphia, 1993.

[6] A. L. Dontchev. Local analysis of a Newton-type method based on partial linearization. *AMS Lectures in Appl. Math.*, pages 295–306, 1996.

[7] H. O. Fattorini and S. Sritharan. Necessary and sufficient for optimal controls in viscous flow problems. *Proc. Royal Soc. of Edinburgh*, 124:211–251, 1994.

[8] M. D. Gunzburger and S. Manservisi. The velocity tracking problem for Navier-Stokes flows with bounded distributed controls. *SIAM J. Contr. Optim.*, 37:1913–1945, 1999.

[9] M. D. Gunzburger and S. Manservisi. Analysis and approximation of the velocity tracking problem for Navier-Stokes flows with distributed control. *SIAM J. Numer. Anal.*, 37:1481–1512, 2000.

[10] J. Heywood, R. Rannacher, and S. Turek. Artificial boundaries and flux and pressure conditions for the incompressible Navier-Stokes equations. *Int. J. Numer. Meth. Fluids*, 22:325–352, 1996.

[11] M. Hintermüller and M. Hinze. Globalization of SQP-methods in control of the instationary Navier-Stokes equations. *ESAIM: M2AN*, 36:725–746, 2002.

[12] 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. Technical Report TR03-11, Department of Computational and Applied Mathematics, Rice University, submitted, 2003.

[13] M. Hinze. *Optimal and instantaneous control of the instationary Navier-Stokes equations.* Habilitation, TU Berlin, revised version, 2002.

[14] M. Hinze and K. Kunisch. Second order methods for optimal control of time-dependent fluid flow. *SIAM J. Contr. Optim.*, 40:925–946, 2001.

[15] K. Kunisch and A. Rösch. Primal-dual active set strategy for a general class of constrained optimal control problems. *SIAM J. Opt.*, 13:321–334, 2002.

[16] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Uralceva. *Linear and quasilinear equations of parabolic type.* AMS Publications, Providence, 1968.

[17] J. L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications*, volume I. Springer, Berlin, 1972.

[18] S. M. Robinson. Strongly regular generalized equations. *Mathematics of Operation Research*, 5:43–62, 1980.

[19] T. Roubíček and F. Tröltzsch. Lipschitz stability of optimal controls for the steady-state Navier-Stokes equations. *Control and cybernetics. To appear*, 2002.

[20] H. Schaefer. Über die Methode der a-priori Schranken. *Math. Annalen*, 129:415–416, 1955.

[21] V. A. Solonnikov. Estimates of the solutions of a nonstationary linearized system of Navier-Stokes equations. *Am. Math. Soc., Translations*, 75:1–116, 1968. translated from Trudy Mat. Inst. Steklov, 70, 213–317, 1964.

[22] R. Temam. *Navier-Stokes equations.* North Holland, Amsterdam, 1979.

[23] H. Triebel. Function spaces in Lipschitz domains and on Lipschitz manifolds. Characteristic functions as pointwise multipliers. *Revista Matemática Complutense*, 15:475–524, 2002.

[24] F. Tröltzsch. On the Lagrange-Newton-SQP method for the optimal control of semilinear parabolic equations. *SIAM J. Control and Optimization*, 38:294–312, 1999.

[25] F. Tröltzsch and D. Wachsmuth. Second-order sufficient optimality conditions for the optimal control of Navier-Stokes equations. Preprint 30-2003, Institut für Mathematik, TU Berlin, submitted.

[26] M. Ulbrich. Constrained optimal control of Navier-Stokes flow by semismooth Newton methods. *Systems & Control Letters*, 48:297–311, 2003.

[27] D. Wachsmuth. Regularity and stability of optimal controls of instationary Navier-Stokes equations. Preprint 03-2004, Institut für Mathematik, TU Berlin, submitted, 2004.

[28] W. von Wahl. Regularitätsfragen für die instationären Navier-Stokesschen Gleichungen in höheren Dimensionen. *J. Math. Soc. Japan*, 32(2):263–283, 1980.