Technische Universität Berlin
Institut für Mathematik

Sufficient second-order optimality conditions for convex control constraints

Daniel Wachsmuth

Preprint 30-2004

Preprint-Reihe des Instituts für Mathematik
Technische Universität Berlin

Report 30-2004 October 2004

Sufficient second-order optimality conditions for convex control constraints

DANIEL WACHSMUTH[1]

Abstract. In this article sufficient optimality conditions are established for optimal control problems with pointwise convex control constraints. Here, the control is a function with values in $\mathbb{R}^n$. The constraint is of the form $u(x) \in U(x)$, where U is an set-valued mapping that is assumed to be measurable with convex and closed images. The second-order condition requires coercivity of the Lagrange function on a suitable subspace, which excludes strongly active constraints, together with first-order necessary conditions. It ensures local optimality of a reference function in a L^∞-neighborhood. The analysis is done for a model problem namely the optimal distributed control of the instationary Navier-Stokes equations.

Key words. Optimal control, sufficient second-order conditions, strongly active sets, convex control constraints, measurable set-valued functions, Navier-Stokes equations

AMS subject classifications. Primary 49K20, Secondary 26E25, 49J53, 76D05

1. Introduction. We are considering optimal control problems with convex control constraints. The abstract problem, we have in mind, reads as follows

$$\min f(y,u) \text{ subject to } E(y,u) = 0 \text{ and } u \in U_{ad}.$$

Here, the set of admissible controls U_{ad} is a subset of $L^p(D)^n$ where D is a domain in $\mathbb{R}^n$ and $p \geq 1$. The controls have to satisfy for almost all $\xi \in D$ the pointwise constraint

$$u(\xi) \in U(\xi),$$

where $U : D \rightsquigarrow \mathbb{R}^n$ is a given set-valued measurable function. The equality constraint E will be a partial differential equation. Thus, the function y will be the state of the system.

Problems of this form are often studied in literature. Optimality conditions were established for optimal control problems for partial differential equations in the last two decades. Sufficient conditions to ensure local optimality were presented for instance in [7, 8, 12]. Other type of constraints, such as state or mixed control-state constraints, are topic of current research, and the theory of sufficient optimality conditions is far from being complete. First results are proven for instance in [18]. Sufficient optimality conditions are basic requirements to prove stability of optimal controls, convergence of optimization algorithms, and even convergence of numerical schemes. See for example [21] for a convergence analysis of the SQP-method applied to optimal control of semilinear parabolic equations.

The problems studied in the mentioned articles are subject to box-constraints on the control. This is the most suitable choice in cases where the control is a scalar such as heating, cooling and so on. But in some applications the control has a vector nature. For instance, in fluid dynamics the control is a velocity, which apparently is a vector in $\mathbb{R}^2$ respectively $\mathbb{R}^3$. A second application is the control of reaction-diffusion equations. Here, the system is controlled by supply of the involved chemicals. In

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

those cases, it is more adequate to have control constraints of the form $g(\xi; u(\xi)) = g(\xi; u_1(\xi), \dots, u_n(\xi)) = 0$ or $u(\xi) \in U(\xi) \subset \mathbb{R}^n$.

Optimal control problems with such control constraints are rarely investigated in literature. Second-order necessary conditions for problems with the control constraint $u(\xi) \in U(\xi)$ were proven in [15] involving second-order admissible variations. In [4, 5, 9], second-order necessary as well as sufficient conditions are established. However, the set of admissible controls has to be polygonal and independent of ξ, i.e. $U(\xi) \equiv U$. This results were extended in [6] to the case of finitely many convex contraints $g_i(u(\xi)) = 0$, $i = 1, \dots, l$.

The aim of the present article is the following. The control constraint is treated as an inclusion $u(\xi) \in U(\xi)$. The advantage of our approach is that the analysis is based on rather elementary say geometrical arguments, hence there is no need of any constraint qualification. We will prove a second-order sufficient optimality condition. It requires the fulfillment of first-order necessary conditions together with coercivity of the Lagrange function on a suitable subspace, which excludes strongly active constraints. Here, the set of strongly active contraints is defined by geometrical terms, which means it is independent of the representation of the control constraint $U(\cdot)$. Then, the second-order condition ensures local optimality of a reference function in a L^∞-neighborhood.

As a model problem serves the optimal distributed control of the instationary Navier-Stokes equations in two dimensions. We emphasize that the restriction to two dimensions is only due to the limitation of the analysis of instationary Navier-Stokes equations. As long as there exists an applicable theory of a state equation in $\mathbb{R}^n$, all results are ready for an extension to the n-dimensional case.

To be more specific, we want to minimize the following quadratic objective functional:

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

subject to the instationary Navier-Stokes equations

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

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

$$U_{ad} = \{u \in L^2(Q)^2 : \ u(x,t) \in U(x,t) \text{ a.e. on } Q\}.$$

Here, Ω is a bounded domain in $\mathbb{R}^2$, Q denotes the time-space cylinder $Q := \Omega \times (0,T)$. Let us underline the fact that for $(x,t) \in Q$ the control $u(x,t)$ is a vector in $\mathbb{R}^2$.

The conditions imposed on the various ingredients of the optimal control problem are specified in Sections 2.1 and 2.2, see Assumptions (A) and (AU).

The plan of the article is as follows. At first, in Section 2, we supply some basic material such as the definition of functions spaces and solvability of the state equation. The pointwise control constraint is studied in Section 2.2. In Section 3 we write down briefly the well-known necessary optimality conditions of the model problem. The main result of the article is stated in Section 4 and proven afterwards in 4.2.

2. Notations and preliminary results. At first, we introduce some notations and provide some results that we need later on. To begin with, we define the spaces of solenoidal or divergence-free functions

$$H := \{v \in L^2(\Omega)^2 : \operatorname{div} v = 0\}, \qquad V := \{v \in H_0^1(\Omega)^2 : \operatorname{div} v = 0\}.$$

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

$$\|y\|_{L^p(X)} := \|y\|_{L^p(0,T;X)} = \left(\int_0^T |y(t)|_X^p dt \right)^{1/p} \qquad 1 \le p < \infty,$$
$$\|y\|_{L^\infty(X)} := \operatorname*{vrai\,max}_{t\in(0,T)} |y(t)|_X .$$

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

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

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, 13].

2.1. The state equation. Before we start with the discussion of the state equation, we specify the requirements for the various ingredients describing the optimal control problem. In the sequel, we assume that the following conditions are satisfied:

$$(\mathbf{A}) \begin{cases} \textit{1. } \Omega \textit{ has Lipschitz boundary } \Gamma := \partial\Omega, \textit{ such that } \Omega \textit{ is locally on} \\ \quad \textit{one side of } \Gamma, \\ \textit{2. } y_0,\, y_d \in H,\ y_Q \in L^2(Q)^2, \\ \textit{3. } \alpha_T,\, \alpha_Q,\, \alpha_R \ge 0, \\ \textit{4. } \gamma,\, \nu > 0. \end{cases}$$

The assumptions on the set-valued mapping U are given in the next section. Now, we will briefly summarize known facts about the solvability of the instationary Navier-

Stokes equations (1.2). First, we define the trilinear form $b : V \times V \times V \mapsto \mathbb{R}$ by

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

Its time integral is denoted by b_Q,

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

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

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

For instance, the operator B is continuous and twice Frechét-differentiable as operator from $W(0, T)$ to $L^2(0, T; V')$.

Now, we concretize 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*

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

Results concerning the solvability of (2.1) are standard, cf. [20] for proofs and further details.

Theorem 2.2 (Existence and uniqueness of solutions). *For every source term $f \in L^2(0, T; V')$ and initial value $y_0 \in H$, 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 $W(0, T)$.*

It is well-known that the control-to-state mapping is even Fréchet-differentiable, whereas the first derivative can be computed as the solution of a linearized equation, cf. [22].

Remark 2.3 (Linearized state equation). *We consider the linearized equation*

$$\begin{aligned} y_t + \nu Ay + B'(\bar{y})y &= f, \\ y(0) &= y_0, \end{aligned} \tag{2.2}$$

for a given state $\bar{y}$, which is usually the solution of the nonlinear system (2.1). *Following the lines of Temam, one can prove existence and uniqueness of a weak solution y in the space $W(0, T)$.*

2.2. Convex control constraints. In this section, we want to investigate the convex control constraint, which has to hold pointwisely

$$u(x,t) \in U(x,t) \text{ a.e. on } Q.$$

We recall the definition of the set of admissible controls U_{ad},

$$U_{ad} = \{u \in L^2(Q)^2 : \ u(x,t) \in U(x,t) \text{ a.e. on } Q\}.$$

Here, we have to make clear, which assumptions we impose on the constraint mapping $U(\cdot)$. At first, we want to introduce measurable, set-valued functions.

DEFINITION 2.4. *A set-valued mapping $F : \Omega \rightsquigarrow X$ with closed images is called* measurable*, if the inverse of each open set is measurable. In other words, for every open subset $\mathcal{O} \subset X$ the inverse image*

$$F^{-1}(\mathcal{O}) = \{\omega \in \Omega : \ F(\omega) \cap \mathcal{O} \neq \emptyset\}$$

has to be measurable.

Once and for all, we specify the requirements for the function U, which defines the control constraints.

(AU)
The set-valued function $U : Q \rightsquigarrow \mathbb{R}^2$ satisfies:
1. *U is a measurable set-valued function.*
2. *The images of U are closed and convex with non-empty interior a.e. on Q. That is, the sets $U(x,t)$ are closed and convex with non-empty interior for almost all $(x,t) \in Q$.*
3. *There exists a function $f_U \in L^2(Q)^2$ with $f_U(x,t) \in U(x,t)$ a.e. on Q.*

Assumption *(i)* and *(ii)* guarantee that there exist a measurable selection of U, i.e. a measurable single-valued function f_M with $f_M(x,t) \in U(x,t)$ a.e. on Q. However, the function f_M needs not to be square-integrable. The existence of at least one square integrable, admissible function is then ensured by the third assumption. Further, it allows us to prove that the pointwise projection on U_{ad} of a L^2-function is itself a L^2-function. Please note, we did not impose any conditions on the sets $U(x,t)$ that are beyond convexity such as boundedness or regularity.

COROLLARY 2.5. *The set of admissible controls U_{ad} defined by*

$$U_{ad} = \{u \in L^2(Q)^2 : \ u(x,t) \in U(x,t) \text{ a.e. on } Q\}$$

is non-empty, convex and closed in $L^2(Q)^2$.

Proof. By assumption (AU), we have $f_U \in U_{ad}$. It is obvious that U_{ad} is convex, since $U(x,t)$ is convex for almost all $(x,t) \in Q$. Let a sequence $\{f_n\}_{n=1}^\infty \subset U_{ad}$ converging in L^2 to f be given. Then, we can find a subsequence f_{n_k} which converges to f pointwise almost everywhere. Since $U(x,t)$ is closed, it follows $f(x,t) \in U(x,t)$ a.e. on Q. Hence, it holds $f \in U_{ad}$. ∎

COROLLARY 2.6. *Let be given a function $u \in L^2(Q)^2$. Then the function v defined pointwise a.e. by*

$$v(x,t) = \mathrm{Proj}_{U(x,t)}(u(x,t))$$

is also in $L^2(Q)^2$*. Further, if for some* $p \geq 2$ *the functions* u *and* f_U *are in* $L^p(Q)^2$*, then the projection* v *is in* $L^p(Q)^2$ *as well.*

Proof. By assumption (AU), the set-valued function U is measurable with closed and convex images, and u is a measurable single-valued function. Then the function v is measurable as well, cf. [3, Cor. 8.2.13]. By Lipschitz continuity of the pointwise projection, it holds

$$\begin{aligned}|v(x,t) - f_U(x,t)| &= |\operatorname{Proj}_{U(x,t)}(u(x,t)) - \operatorname{Proj}_{U(x,t)}(f_U(x,t))| \\ &\leq |u(x,t) - f_U(x,t)|\end{aligned}$$

almost everywhere on Q. Thus, squaring and integrating gives

$$\|v - f_U\|_2^2 \leq \|u - f_U\|_2^2 < \infty,$$

which implies $v \in L^2(Q)^2$. If in addition, u and f_U are in $L^p(Q)^2$ for some $p > 2$, then we can prove analogously that the projection is also in L^p, i.e. $v \in L^p(Q)^2$. $\square$

Let us recall some definitions from the theory of convex sets. For a convex set $C \in \mathbb{R}^n$ and an element $u \in C$, we denote by $\mathcal{N}_C(u)$ and $\mathcal{T}_C(u)$ the normal cone respectively polar cone of tangents of C at the point u, which are defined by

$$\begin{aligned}\mathcal{N}_C(u) &= \{z \in \mathbb{R}^n : z^T(v-u) \leq 0 \quad \forall v \in C\}, \\ \mathcal{T}_C(u) &= \{z \in \mathbb{R}^n : z^T v \leq 0 \quad \forall v \in N_C(u)\}.\end{aligned}$$

Further, we will need the linear subspaces

$$N_C(u) = \operatorname{cl}\operatorname{span}\mathcal{N}_C(u), \quad T_C(u) = N_C(u)^\perp.$$

Now, we want to use these notations with $C = U_{ad}$. Let be given an admissible control $u \in U_{ad}$. It is well-known, that the sets $\mathcal{N}_{U_{ad}}(u)$, $\mathcal{T}_{U_{ad}}(u)$, $N_{U_{ad}}(u)$, and $T_{U_{ad}}(u)$ admit a pointwise representation as U_{ad} itself, cf. [3, 17]. For instance, the set $\mathcal{N}_{U_{ad}}(u)$ is given by

$$\mathcal{N}_{U_{ad}}(u) = \left\{v \in L^2(Q)^2 : v(x,t) \in \mathcal{N}_{U(x,t)}(u(x,t)) \text{ a.e. on } Q\right\}.$$

We introduce the following projection operations. Let be given a function $u \in L^p(Q)^n$, $1 \leq p \leq \infty$, with $u(x,t) \in U(x,t)$ a.e. on Q. For $w \in L^p(Q)^n$ we define

$$w_N(x,t) = \operatorname{Proj}_{\left[N_{U(x,t)}(u(x,t))\right]}(w(x,t)), \tag{2.3}$$

which is the pointwise projection of $w(x,t)$ on the space of normal directions of $U(x,t)$ at $u(x,t)$. Its orthogonal counterpart is denoted by

$$w_T(x,t) = \operatorname{Proj}_{\left[T_{U(x,t)}(u(x,t))\right]}(w(x,t)). \tag{2.4}$$

Following the lines of [3, Sect. 8], it is not difficult but technical to prove that if the set-valued mapping $U : Q \rightsquigarrow \mathbb{R}^n$ is measurable, then the set-valued mappings $\mathcal{N}_U, \mathcal{T}_U, N_U, T_U : Q \rightsquigarrow \mathbb{R}^n$ are measurable as well. By [3, Cor. 8.2.13], the projection of a measurable function on the images of a measurable set-valued mapping is measurable. So, we find that the functions w_N and w_T are measurable. Since the projection is pointwise non-expansive, it holds w_n, $w_T \in L^p(Q)^n$.

3. First order necessary optimality conditions. We briefly recall the necessary conditions for local optimality. For the proofs and further discussion see [1, 10, 11, 22] 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 $L^p(Q)^2$, 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 $\|\bar{u} - u_\rho\|_p \leq \rho$. 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

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

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 weak solution $\bar{\lambda} \in W^{4/3}(0, T; V)$ of the adjoint equation*

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

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 [10, 22]. The regularity of $\bar{\lambda}$ is proven in [12].

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

$$u(x,t) = \mathrm{Proj}_{U(x,t)}\left(-\frac{1}{\gamma}\bar{\lambda}(x,t)\right) \qquad \text{a.e. on } Q. \tag{3.3}$$

See for instance [14]. Using the normal cone of the set of admissible controls, the variational inequality (3.2) can be written equivalently as the inclusion

$$-(\gamma\bar{u} + \bar{\lambda}) \in \mathcal{N}_{U_{ad}}(\bar{u}). \tag{3.4}$$

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). *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 $\lambda \in W^{4/3}(0, T)$. Moreover, the mapping $(f, \lambda_T) \mapsto \lambda$ is continuous from $L^2(0, T; V') \times H$ to $W^{4/3}(0, T)$.*

A prove is given in [12].

3.1. Regularity of locally optimal controls. Let us comment on the regularity of a locally optimal control $\bar{u}$. By (3.3), it inherits some regularity from the associated adjoint state $\bar{\lambda}$. If the inhomogeneities of the adjoint system are more regular than required in the previous theorem, one gets more regular adjoint states, see [12]. This can be applied to obtain more regular optimal controls. If $\bar{\lambda}$ is in $L^p(Q)^2$ for some $p \geq 2$ then we know from Corollary 2.6 that $\bar{u}$ is in $L^p(Q)^2$ as well.

In the presence of box constraints

$$u_{a,i}(x,t) \leq u_i(x,t) \leq u_{b,i}(x,t), \quad i = 1, \ldots, n,$$

one can prove even more. If the adjoint state $\bar{\lambda}$ is in $H^1(Q)^2$ then the control $\bar{u}$ is in $H^1(Q)^2$ as well, provided u_a, $u_b \in H^1(Q)^2$ holds, see [23]. However, in our case of convex constraints it is not clear under which assumptions on $U(\cdot)$ the regularity $\bar{\lambda} \in H^1(Q)^2$ can be carried over to $\bar{u} \in H^1(Q)^2$.

3.2. Lagrangian formulation. We introduce the Lagrange functional

$$\mathcal{L} : W(0,T) \times L^2(Q)^2 \times W^{4/3}(0,T) \mapsto \mathbb{R}$$

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. [22]. The reader can readily verify that the necessary conditions can be expressed equivalently by

$$\begin{aligned} \mathcal{L}_y(\bar{y},\bar{u},\bar{\lambda})\, h = 0 \quad \forall h \in W(0,T) \text{ with } h(0) = 0, \\ \mathcal{L}_u(\bar{y},\bar{u},\bar{\lambda})(u - \bar{u}) \geq 0 \quad \forall u \in U_{ad}. \end{aligned} \tag{3.5}$$

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

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

for all $w_1, w_2 \in W(0,T)$, confer [22].

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

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

4. Second-order sufficient optimality conditions. Before we can work with the set of strongly active constraints, we have to define some more notations.

The relative interior of a convex set is defined by

$$\operatorname{ri} C = \{x \in \operatorname{aff} C : \exists \varepsilon > 0, B_\varepsilon(x) \cap \operatorname{aff} C \subset C\},$$

its complement in C is called the relative boundary

$$\operatorname{rb} C = C \setminus \operatorname{ri} C.$$

The distance of a point $u \in \mathbb{R}^n$ to a set $C \subset \mathbb{R}^n$ is defined by

$$\operatorname{dist}(u, C) = \min_{x \in C} |u - x|.$$

For fixed $\varepsilon > 0$, we define the set of strongly active constraints by

$$Q_\varepsilon = \left\{(x,t) \in Q : \operatorname{dist}\left(-(\gamma \bar{u}(x,t) + \bar{\lambda}(x,t)), \operatorname{rb} \mathcal{N}_{U(x,t)}(\bar{u}(x,t))\right) > \varepsilon\right\}. \tag{4.1}$$

In the following, $\bar{v} = (\bar{y}, \bar{u})$ is a fixed admissible reference pair. We suppose that the first-order necessary optimality conditions (3.1)–(3.2) are fulfilled at $\bar{v}$. Furthermore, we assume 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:

(SSC)
There exist $\varepsilon > 0$ and $\delta > 0$ such that

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[(z,h)]^2 \geq \delta \, \|h\|_2^2$$

holds for all pairs $(z,h) \in W(0,T) \times L^2(Q)^2$ with

$$h \in \mathcal{T}_{U_{ad}}(\bar{u}), \;\; h_N = 0 \text{ on } Q_\varepsilon,$$

and $z \in W(0,T)$ being the weak solution of the linearized equation

$$\begin{aligned} z_t + Az + B'(\bar{y})z &= h \\ z(0) &= 0. \end{aligned}$$

The main result of the present article is the following theorem that states the sufficiency of (SSC).

THEOREM 4.1. *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 conditions 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 $|u - \bar{u}|_\infty \leq \rho$.

We will give the proof in Section 4.2 after a series of auxiliary results. At first, we consider the set of strongly active constraints. We prove its measurability, which is a non-trivial result obtained using set-valued analysis. Secondly, we derive from the strongly active contraints some positiveness in directions of test functions that are not included in (SSC).

4.1. Strongly active constraints. Before we turn to the discussion of measurability, we give some interpretation of the set of strongly active constraints. To keep the illustration as simple as possible, the following considerations are only valid for two-dimensional controls, i.e. $U(x,t) \subset \mathbb{R}^2$.

We will distinguish some cases whether $\bar{u}(x,t)$ lies in the interior, on an edge or in a corner of the admissible set $U(x,t)$.

At first, consider the case that $\bar{u}(x,t)$ lies in the interior of $U(x,t)$. Then it holds $\mathcal{N}_{U(x,t)}(\bar{u}(x,t)) = N_{U(x,t)}(\bar{u}(x,t)) = \{0\}$. Thus, the first-order necessary optimality conditions imply $\gamma\bar{u}(x,t) + \bar{\lambda}(x,t) = 0$, which is equivalent to $-(\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)) \in \mathrm{rb}\mathcal{N}_{U(x,t)}(\bar{u}(x,t))$. Hence by condition (4.1), the set of strongly active constraints can not contain points where $\bar{u}(x,t)$ lies in the interior of $U(x,t)$. This is what one expects, since no constraint is active.

Now, assume that $\bar{u}(x,t)$ lies on a smooth part of $\partial U(x,t)$, i.e. the normal cone $\mathcal{N}_{U(x,t)}(\bar{u}(x,t))$ is one-dimensional. Then, its relative boundary is the origin, $\mathrm{rb}\mathcal{N}_{U(x,t)}(\bar{u}(x,t)) = \{0\}$. Consequently, (4.1) means $|\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)| > \varepsilon$ on Q_ε. The latter relation is often used to define strongly active constraints for box-constrained optimal control problems, cf. [7, 22].

If $\bar{u}(x,t)$ is a corner of $U(x,t)$ then the dimension of $\mathcal{N}_{U(x,t)}(\bar{u}(x,t))$ is equal to the space dimension two. Here, $\mathcal{N}_{U(x,t)}(\bar{u}(x,t))$ is the convex and conical hull of two extremal vectors n_1 and n_2. We can assume that $|n_1| = |n_2| = 1$ holds. The relative boundary of the normal cone admits the representation

$$\mathrm{rb}\mathcal{N}_{U(x,t)}(\bar{u}(x,t)) = \{a_1\, n_1 |\, a_1 \geq 0\} \cup \{a_2\, n_2 |\, a_2 \geq 0\}$$

The condition (4.1) is equivalent to the fact that $-(\gamma\bar{u} + \bar{\lambda})$ lies in a cone that is the result of a shifting of the normal cone by $\sigma(n_1 + n_2)$, i.e.

$$-(\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)) \in \sigma(n_1 + n_2) + \mathcal{N}_{U(x,t)}(\bar{u}(x,t)),$$

see Figure 4.1. Here, σ is given by $\sigma = \frac{\varepsilon}{\sqrt{1-(n_1 \cdot n_2)^2}}$.

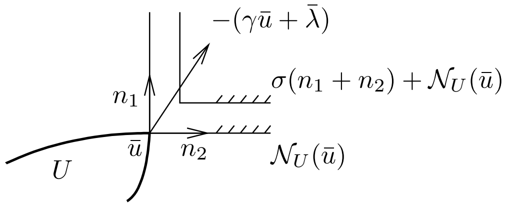

FIG. 4.1.

Now, we want to prove the measurability of the set of strongly active constraints.

LEMMA 4.2. *Suppose $\bar{u}$ and $\bar{\lambda}$ fulfill the first-order necessary optimality conditions, and $U : Q \rightsquigarrow \mathbb{R}^2$ is measurable. Then the set Q_ε defined in* (4.1) *is measurable as well.*

Proof. At first, one finds for a convex set $\mathcal{N} \subset \mathbb{R}^n$ and a vector $u \in \mathcal{N} \subset \mathbb{R}^n$ that the following

$$\mathrm{dist}\,(u,\, \mathrm{rb}\mathcal{N}) = \mathrm{dist}\,(u,\, \mathrm{span}\mathcal{N} \setminus \mathcal{N}) \tag{4.2}$$

holds. As already mentioned, the set-valued mapping $(x,t) \rightsquigarrow \mathcal{N}_{U(x,t)}(\bar{u}(x,t))$ is measurable. Using proving techniques of [3], one can check measurability of $(x,t) \rightsquigarrow \operatorname{span}\mathcal{N}_{U(x,t)}(\bar{u}(x,t))$. By [3, Cor. 8.2.13], the distance between a measurable function u and a measurable set-valued function U, which is a function defined by

$$[\operatorname{dist}(u,U)](x,t) := \operatorname{dist}(u(x,t),\, U(x,t)),$$

is also measurable. This implies that the function $d_{\mathcal{N}}$ given by

$$d_{\mathcal{N}}(x,t) = \operatorname{dist}\left(-(\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)),\, \operatorname{span}\mathcal{N}_{U(x,t)}(\bar{u}(x,t)) \setminus \mathcal{N}_{U(x,t)}(\bar{u}(x,t))\right)$$

is measurable. By assumption, $\bar{u}$, $\bar{\lambda}$ fulfill the first-order necessary optimality conditions especially relation (3.4). Therefore, we can apply (4.2) and obtain that

$$\operatorname{dist}\left(-(\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)),\, \operatorname{rb}\mathcal{N}_{U(x,t)}(\bar{u}(x,t))\right) = d_{\mathcal{N}}(x,t)$$

is a measurable function from Q to $\mathbb{R}$. Using the representation

$$Q_\varepsilon = d_{\mathcal{N}}^{-1}\left((-\infty,-\varepsilon) \cup (\varepsilon,+\infty)\right),$$

we finally find that Q_ε is a measurable set. ∎

The condition (SSC) requires coercivity of the second derivative of the Langrangian only with respect to test functions h, whose normal components are zero, i.e. $h_N = 0$. However, by the following Lemma, we gain an additional positive term that we will need in the proof of sufficiency, see Section 4.2. To this aim, we denote the L^p-norm with respect to the set of positivity for $u \in L^p(Q)^2$ and $1 \le p < \infty$ by

$$\|u\|_{L^p(Q_\varepsilon)} := \left(\int_{Q_\varepsilon} |u(x,t)|^p \,dx\,dt\right)^{1/p}.$$

The positiveness result then reads as follows.

Lemma 4.3. *For all $u \in U_{ad}$ with $\|u - \bar{u}\|_\infty < \rho$ it holds*

$$(\gamma\bar{u} + \bar{\lambda},\, u - \bar{u})_Q \ge \frac{\varepsilon}{\rho}\|(u-\bar{u})_N\|^2_{L^2(Q_\varepsilon)},$$

where $(\cdot)_N$ denotes the pointwise projection on $N_{U(x,t)}(\bar{u}(x,t))$, which is the space of normal directions of $U(x,t)$ at $\bar{u}(x,t)$.

Proof. Let $u \in U_{ad}$ be given. Since $(\bar{u},\bar{\lambda})$ fulfills the first-order necessary optimality conditions, it holds

$$\int_{Q\setminus Q_\varepsilon} (\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)) \cdot (u(x,t) - \bar{u}(x,t))\,dx\,dt \ge 0.$$

Hence, we only need to investigate the difference $u - \bar{u}$ on the set of strongly active constraints Q_ε. Now, take $(x,t) \in Q_\varepsilon$. We split the difference of both controls into parts belonging to the space of normal directions $N(x,t) = N_{U(x,t)}(\bar{u}(x,t))$ and its orthogonal complement $T(x,t) = T_{U(x,t)}(\bar{u}(x,t)) = N(x,t)^\perp$,

$$u(x,t) - \bar{u}(x,t) = (u(x,t) - \bar{u}(x,t))_N + (u(x,t) - \bar{u}(x,t))_T.$$

The necessary optimality conditions imply

$$-(\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)) \in \mathcal{N}_{U(x,t)}(\bar{u}(x,t)) \subset N_{U(x,t)}(\bar{u}(x,t)) = N(x,t),$$

which allows us to conclude

$$(\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)) \cdot (u(x,t) - \bar{u}(x,t))_T = 0 \tag{4.3}$$

almost everywhere on Q_ε. Now, we have to distinguish two cases: whether the normal component $(u(x,t) - \bar{u}(x,t))_N$ vanishes or not. If it is zero, we have trivially

$$0 = (\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)) \cdot (u(x,t) - \bar{u}(x,t))_N \geq \varepsilon|(u(x,t) - \bar{u}(x,t))_N| = 0.$$

On the other hand, suppose $(u(x,t) - \bar{u}(x,t))_N \neq 0$. By definition, the gradient $-(\gamma\bar{u}(x,t) + \bar{\lambda}(x,t))$ belongs to the relative interior of $\mathcal{N}_{U(x,t)}(\bar{u}(x,t))$. Thus, there exists $\tau > 0$, such that

$$-(\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)) + \tau(u(x,t) - \bar{u}(x,t))_N \in \mathrm{rb}\,\mathcal{N}_{U(x,t)}(\bar{u}(x,t))$$

is satisfied, which is equivalent to

$$\big(\gamma\bar{u}(x,t) + \bar{\lambda}(x,t) - \tau(u(x,t) - \bar{u}(x,t))_N\big) \cdot (u(x,t) - \bar{u}(x,t)) \geq 0. \tag{4.4}$$

But we know even more, we can estimate the norm of the correction $\tau(u - \bar{u})_N$ using (4.1) by

$$\tau|(u(x,t) - \bar{u}(x,t))_N| > \varepsilon.$$

Combining (4.3), (4.4), and the previous estimate, we obtain for $(x,t) \in Q_\varepsilon$

$$\begin{aligned}(\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)) \cdot (u(x,t) - \bar{u}(x,t)) &\geq \tau(u(x,t) - \bar{u}(x,t))_N \cdot (u(x,t) - \bar{u}(x,t))_N \\ &\geq \tau|(u(x,t) - \bar{u}(x,t))_N|^2 \\ &\geq \varepsilon|(u(x,t) - \bar{u}(x,t))_N|.\end{aligned} \tag{4.5}$$

Now, we integrate over Q and take (4.3), (4.5) into account to get

$$\begin{aligned}&\int_Q (\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)) \cdot (u(x,t) - \bar{u}(x,t))\,dx\,dt \\ &\qquad \geq \int_{Q_\varepsilon} (\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)) \cdot (u(x,t) - \bar{u}(x,t))\,dx\,dt \\ &\qquad \geq \varepsilon \int_{Q_\varepsilon} |(u(x,t) - \bar{u}(x,t))_N|\,dx\,dt \\ &\qquad = \varepsilon\|(u - \bar{u})_N\|_{L^1(Q_\varepsilon)}.\end{aligned}$$

An interpolation argument together with the pre-requisite $\|u - \bar{u}\|_\infty \leq \rho$ yields

$$\begin{aligned}&\int_Q (\gamma\bar{u}(x,t) + \bar{\lambda}(x,t)) \cdot (u(x,t) - \bar{u}(x,t))\,dx\,dt \geq \varepsilon\|(u - \bar{u})_N\|_{L^1(Q_\varepsilon)} \\ &\qquad \geq \frac{\varepsilon}{\rho}\|(u - \bar{u})_N\|_{L^1(Q_\varepsilon)}\|(u - \bar{u})_N\|_{L^\infty(Q_\varepsilon)} \geq \frac{\varepsilon}{\rho}\|(u - \bar{u})_N\|^2_{L^2(Q_\varepsilon)},\end{aligned}$$

which is the desired result. ∎

4.2. Proof of Theorem 4.1. Throughout the proof, c is used as a generic constant. Suppose that $\bar{v} = (\bar{y}, \bar{u})$ fulfills the assumptions of the theorem. Let (y, u) be another admissible pair. We have

$$J(\bar{v}) = \mathcal{L}(\bar{v}, \bar{\lambda}) \quad \text{and} \quad J(v) = \mathcal{L}(v, \bar{\lambda}),$$

since $\bar{v}$ and v are admissible. Taylor-expansion of the Lagrange-function yields

$$\mathcal{L}(v, \bar{\lambda}) = \mathcal{L}(\bar{v}, \bar{\lambda}) + \mathcal{L}_y(\bar{v}, \bar{\lambda})(y - \bar{y}) + \mathcal{L}_u(\bar{v}, \bar{\lambda})(u - \bar{u}) + \frac{1}{2}\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v - \bar{v}, v - \bar{v}]. \tag{4.6}$$

Notice that there is no remainder term due to the quadratic nature of all nonlinearities. Moreover, the necessary conditions (3.5) are satisfied at $\bar{v}$ with adjoint state $\bar{\lambda}$. Therefore, the second term vanishes. The third term is nonnegative due to the variational inequality (3.2). However, we get even more by Lemma 4.3,

$$\mathcal{L}_u(\bar{v}, \bar{\lambda})(u - \bar{u}) = \int_Q (\gamma\bar{u} + \bar{\lambda})(u - \bar{u})\, dxdt \geq \frac{\varepsilon}{\rho}\|(u - \bar{u})_N\|^2_{L^2(Q_\varepsilon)}.$$

Here, ρ is a parameter such that $\|u - \bar{u}\|_\infty \leq \rho$, which will be chosen sufficiently small in the course of the proof. So we arrive at

$$\begin{aligned} J(v) &= J(\bar{v}) + \mathcal{L}_y(\bar{v}, \bar{\lambda})(y - \bar{y}) + \mathcal{L}_u(\bar{v}, \bar{\lambda})(u - \bar{u}) + \frac{1}{2}\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v - \bar{v}]^2 \\ &\geq J(\bar{v}) + \frac{1}{2}\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[v - \bar{v}]^2 + \frac{\varepsilon}{\rho}\|(u - \bar{u})_N\|^2_{L^2(Q_\varepsilon)}. \end{aligned} \tag{4.7}$$

We set $\delta u := u - \bar{u}$. Let us define δy to be the weak solution of the linearized system

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

When we use δy instead of $y - \bar{y}$, we make a small error $r_1 := (y - \bar{y}) - \delta y$. A short calculation shows that r_1 solves the following linearized system

$$\begin{aligned} r_t + \nu A r + B'(\bar{y})r &= B''(\bar{y})[y - \bar{y}]^2, \\ r(0) &= 0. \end{aligned}$$

Thus, we can estimate the norm of the error r_1 by

$$\|r_1\|_{W(0,T)} \leq c\|B''(\bar{y})[y - \bar{y}]^2\|_{L^2(V')} \leq c\|y - \bar{y}\|^2_{W(0,T)}.$$

Since the solution mapping of the nonlinear problem is locally Lipschitz continuous, we find

$$\|r_1\|_{W(0,T)} \leq c\|y - \bar{y}\|^2_{W(0,T)} \leq c\|\delta u\|_2^2.$$

Substituting $y - \bar{y} = \delta y + r_1$, we obtain

$$\begin{aligned} \mathcal{L}_{yy}(\bar{v}, \bar{\lambda})[y - \bar{y}]^2 &= \mathcal{L}_{yy}(\bar{v}, \bar{\lambda})[\delta y]^2 + 2\mathcal{L}_{yy}(\bar{v}, \bar{\lambda})[\delta y, r_1] + \mathcal{L}_{yy}(\bar{v}, \bar{\lambda})[r_1]^2 \\ &= \mathcal{L}_{yy}(\bar{v}, \bar{\lambda})[\delta y]^2 + r_2, \end{aligned}$$

where r_2 is a remainder term satisfying

$$\frac{\|r_2\|_{W(0,T)}}{\|\delta u\|_2^2} \to 0 \quad \text{as } \|\delta u\|_2 \to 0.$$

Let us abbreviate $\delta v = (\delta y, \delta u)$. We achieved the following estimate for the difference of the objective values

$$J(v) - J(\bar{v}) \geq \frac{1}{2}\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[\delta v]^2 + \frac{\varepsilon}{\rho}\|(u-\bar{u})_N\|^2_{L^2(Q_\varepsilon)} + r_2,$$

provided $\|u - \bar{u}\|_\infty \leq \rho$ holds. In the next step, we want to apply the coercivity assumption (SSC). To this aim, we split δu in two components as follows:

$$\delta u = h_u + r_u,$$

where h_u and r_u are defined by

$$h_u = \begin{cases} \delta u_T & \text{on } Q_\varepsilon \\ \delta u & \text{on } Q \setminus Q_\varepsilon \end{cases}, \qquad r_u = \begin{cases} \delta u_N & \text{on } Q_\varepsilon \\ 0 & \text{on } Q \setminus Q_\varepsilon \end{cases}.$$

Observe, that h_u and r_u are orthogonal, i.e. $(h_u, r_u)_Q = 0$. Moreover, it follows from the definition that the identity

$$\|r_u\|_p = \|(u-\bar{u})_N\|_{L^p(Q_\varepsilon)} \tag{4.8}$$

holds. Similarly, we split $\delta y = h_y + r_y$, where h_y and r_y are solutions of the respective linearized systems with right-hand sides h_u and r_u. Further, we set $h_v := (h_y, h_u)$ and $r_v := (r_y, r_u)$. We continue the investigation of the Lagrangian,

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[\delta v]^2 = \mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[h_v]^2 + 2\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[h_v, r_v] + \mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[r_v]^2. \tag{4.9}$$

Now, we can use (SSC) to obtain

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[h_v]^2 \geq \delta\|h_u\|_2^2. \tag{4.10}$$

The following estimate is a conclusion of the inequality (3.7) and the Lipschitz continuity of the solution mapping of the linearized system

$$\begin{aligned} \left|2\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[h_v, r_v] + \mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[r_v]^2\right| &\geq -c\|r_y\|_{W(0,T)}\left(\|h_y\|_{W(0,T)} + \|r_y\|_{W(0,T)}\right) \\ &\geq -c\|r_u\|_2\left(\|h_u\|_2 + \|r_u\|_2\right) \\ &\geq -\frac{\delta}{2}\|h_u\|_2^2 - c\|r_u\|_2^2. \end{aligned} \tag{4.11}$$

Using the relation $\|h_u\|_2^2 \geq 1/2\|\delta u\|_2^2 - \|r_u\|_2^2$, we get by (4.9)–(4.11)

$$\mathcal{L}_{vv}(\bar{v}, \bar{\lambda})[\delta v]^2 \geq \frac{\delta}{2}\|h_u\|_2^2 - c\|r_u\|_2^2 \geq \frac{\delta}{4}\|\delta u\|_2^2 - c\|r_u\|_2^2.$$

So far, we proved the following estimate

$$J(v) - J(\bar{v}) \geq \frac{\delta}{8}\|\delta u\|_2^2 + \left(\frac{\varepsilon}{\rho} - c\right)\|(u-\bar{u})_N\|^2_{L^2(Q_\varepsilon)} + r_2.$$

Here, we used the identity (4.8). Choosing ρ small enough, we finally find

$$J(v) - J(\bar{v}) \geq \frac{\delta}{16}\|\delta u\|_2^2 = \frac{\delta}{16}\|u - \bar{u}\|_2^2.$$

Thus, we proved quadratic growth of the objective functional in a L^∞-neighborhood of the reference control. It implies the local optimality of the pair $(\bar{y}, \bar{u})$. ∎

4.3. Generalizations, concluding remarks.

4.3.1. General objective functional. The analysis of the proof of sufficiency in not restricted to the special quadratic nature of the objective functional J defined in (1.1). Let us consider the minimization of the functional

$$\tilde{J}(y,u) = \int_\Omega \omega(x, y(x,T))\,dx + \int_Q q(x,t,y(x,t),u(x,t))\,dxdt.$$

We have to require appropriate measurability and differentiability assumptions, which are standard in the literature, see for instance [7, 16]. Furthermore, we need L^∞-regularity of the state and control to obtain Frechét differentiability of the objective functional. Regarding instationary Navier-Stokes equations, it is known that the regularity $u \in L^p(Q)^2$ with $p > 2$ gives the regularity of the state $y \in L^\infty(Q)^2$, cf. [19, 24]. Additionally, we get an extra second-order remainder term in the Tayor expansion (4.6) of the Lagrange functional. Up to this differences, the method of proof remains the same.

4.3.2. Local optimality in L^s-neighborhood. The condition (SSC) together with the first-order necessary optimality conditions yields the local optimality of a reference control in a L^∞-neighborhood. This means more or less that jumps of the optimal control have to be known a-priorily. With minor modifications, one can proof optimality of a reference pair $(\bar{y}, \bar{u})$ in neighborhoods of the control $\bar{u}$ defined by norms weaker than L^∞.

Let be given two numbers q and s satisfying $4/3 \le q < 2$ and $1/q = 1/2 + 1/(2s)$. This implies $2 \le s < \infty$ to hold. The first assumption is needed to ensure $L^q \subset W(0,T)^*$, which yields continuity of the control-to-state mapping from L^q to $W(0,T)$. The second one allows us to estimate $\|u\|_q \le \|u\|_1^{1/2}\|u\|_s^{1/2}$, which is used in connection with strongly active constraints. Summarizing, one can proof along the lines of Section 4.2 the following:

THEOREM 4.4. *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 conditions 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*

$$J(v) \ge J(\bar{v}) + \alpha\,|u - \bar{u}|_q^2$$

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

Observe, that we achieve quadratic growth of the objective functional in the L^q-norm which is weaker than L^2, but the growth takes place in a L^s-neigborhood of the reference control. For a more detailed discussion, we refer to [22].

4.3.3. Equivalent formulation. In [4, 6], Bonnans proposed the following formulation of (SSC):

$$
(\mathbf{SSC}_0)\begin{cases}
\textit{It holds} \\
\qquad\qquad \mathcal{L}_{vv}(\bar v,\bar\lambda)[(z,h)]^2 > 0 \\
\textit{for all pairs } (z,h)\in W(0,T)\times L^2(Q)^2 \textit{ with } h\neq 0, \\
\qquad\qquad h\in \mathcal{T}_{U_{ad}}(\bar u),\ h_N = 0 \textit{ on } Q_0, \\
\textit{and } z\in W(0,T) \textit{ being the weak solution of the linearized equation} \\
\qquad\qquad z_t + Az + B'(\bar y)z = h \\
\qquad\qquad z(0) = 0.
\end{cases}
$$

Despite the fact, that (SSC_0) looks weaker than (SSC), it can be proven that both conditions are equivalent, cf. [22]. Moreover, condition (SSC_0) implies quadratic growth of the objective functional. Although in the original paper the control constraints were described by finitely many inequalities of the form

$$g_i(\, u(x,t)\,) \le 0, \quad i = 1,\dots,q,$$

the proofs carry over to the control constraints considered in the present article. However, the methods of proof are taylored to the case that $\mathcal{L}_{vv}$ is a Legendre form.

Acknowledgements. The author wants to thank Fredi Tröltzsch for several helpful hints improving the readability of the article.

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] J.-P. Aubin and H. Frankowska. *Set-valued analysis.* Birkhäuser, Boston, 1990.

[4] J. F. Bonnans. Second-order analysis for control constrained optimal control problems of semilinear elliptic equations. *Appl. Math. Optim.*, 38:303–325, 1998.

[5] J. F. Bonnans and A. Shapiro. *Perturbation analysis of optimization problems.* Springer, New York, 2000.

[6] J. F. Bonnans and H. Zidani. Optimal control problems with partially polyhedric constraints. *SIAM J. Control Optim.*, 37:1726–1741, 1999.

[7] E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for a nonlinear elliptic control problem. *J. for Analysis and its Applications*, 15:687–707, 1996.

[8] A. L. Dontchev, W. W. Hager, A. B. Poore, and B. Yang. Optimality, stability, and convergence in optimal control. *Appl. Math. Optim.*, 31:297–326, 1995.

[9] J. C. Dunn. Second-order optimality conditions in sets of L^∞ functions with range in a polyhedron. *SIAM J. Control Optim.*, 33(5):1603–1635, 1995.

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

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

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

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

[14] K. Malanowski. Convergence of approximations vs. regularity of solutions for convex, control-constrained optimal control problems. *Appl. Math. Opt.*, 8:69–95, 1981.

[15] Zs. Páles and V. Zeidan. Optimum problems with measurable set-valued constraints. *SIAM J. Opt.*, 11:426–443, 2000.
[16] J.-P. Raymond and F. Tröltzsch. Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints. *Discrete and Continuous Dynamical Systems*, 6:431–450, 2000.
[17] R. T. Rockafellar. *Conjugate duality and optimization.* SIAM, Philadelphia, 1974.
[18] A. Rösch and F. Tröltzsch. Sufficient second-order optimality conditions for a parabolic optimal control problem with pointwise state constraints. *SIAM J. Control Optim.*, 42:138–154, 2003.
[19] 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.
[20] R. Temam. *Navier-Stokes equations.* North Holland, Amsterdam, 1979.
[21] F. Tröltzsch. On the Lagrange-Newton-SQP method for the optimal control of semilinear parabolic equations. *SIAM J. Control Optim.*, 38:294–312, 1999.
[22] 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, 2003.
[23] D. Wachsmuth. Regularity and stability of optimal controls of instationary Navier-Stokes equations. Preprint 03-2004, Institut für Mathematik, TU Berlin, submitted, 2004.
[24] W. von Wahl. Instationary Navier-Stokes equations and parabolic systems. *Pacific J. Math.*, 72(2):557–569, 1977.