A penalty approach to optimal control of Allen-Cahn variational inequalities: MPEC-view

M. Hassan Farshbaf-Shaker

Abstract
A scalar Allen-Cahn-MPEC problem is considered and a penalization technique is applied to show the existence of an optimal control. We show that the stationary points of the penalized problems converge to some stationary points of the limit problem, which however are weaker than $C$-stationarity conditions.

Key words. Allen-Cahn system, parabolic obstacle problems, MPECs, mathematical programs with complementarity constraints, optimality conditions.

AMS subject classification. 34G25, 35K86, 35R35, 49J20, 65K10

1 Introduction
In a Mini-Workshop Control of Free Boundaries in 2007 in Oberwolfach, see [27], the following paradigm optimal control problem involving free boundaries was formulated. Control the interface evolution law

$$V = -H + u,$$  \hfill (1.1)

where $V$ is the normal velocity and $H$ is the mean curvature of the interface. The space and time dependent quantity $u$ can be used to control the interface. The above formulation is a sharp interface description of the interface. As this is well-known, one drawback of such a description is that it is difficult to handle topological changes, specially if one is interested in
numerical simulations. One way to omit these difficulties is to use suitable approximations of (1.1). Such approximations like diffuse interface models and specially Allen-Cahn models

$$\varepsilon \partial_t y = \varepsilon \Delta y - \frac{1}{\varepsilon} \psi'(y) + u,$$

(1.2)

with the smooth double well potential $\psi(u) = \frac{9}{32} (1 - u^2)^2$ are used extensively in the phase field community, see [9, 10] and references therein. The approximative models (1.2) are constructed in such a way that they converge to the evolution law (1.1) as $\varepsilon \searrow 0$ and have the advantage that topology changes can be dealt with implicity, see [16]. Here an interface in which a phase field or order parameter rapidly changes its value, is modeled to have a thickness of order $\varepsilon$, where $\varepsilon > 0$ is a small parameter. The model is based on a non-convex Ginzburg-Landau energy $E$ which has the form

$$E(y) := \int_\Omega \left( \frac{\varepsilon^2}{2} |\nabla y|^2 + \frac{1}{\varepsilon} \psi(y) \right) dx,$$

(1.3)

where $\Omega \subset \mathbb{R}^d$ is a bounded domain and $y : \Omega \to \mathbb{R}$ is the phase field, also called order parameter. The potential function $\psi$ is assumed to have two global minima at the points $\pm 1$ and the values $\pm 1$ describe the pure phases. In order to have the Ginzburg-Landau energy $E(y)$ of moderate size, $y$ favors the values $\pm 1$ due to the potential function. On the other hand given the gradient term $\int_\Omega |\nabla y|^2$ oscillations between the values $\pm 1$ are energetically not favorable. Given an initial distribution the interface motion can be modeled by the steepest decent of $E$ with respect to the $L^2$-norm which results then in $\varepsilon \partial_t y = \varepsilon \Delta y - \frac{1}{\varepsilon} \psi'(y)$. The space and time dependent quantity $u$ enables then to control the interface motion and we end with (1.2). An approach according to the above formulated paradigm problem is now as follows:

$$\min J(y, u) := \int_\Omega \nu_T \frac{1}{2} (y(T, x) - y_T(x))^2 dx + \int_{\Omega_T} \nu_d \frac{1}{2} (y(t, x) - y_d(t, x))^2 dx dt$$

$$+ \int_{\Omega_T} \frac{\nu_u u^2}{2\varepsilon} dx dt,$$

where $\nu_T, \nu_d, \nu_u > 0$,

such that (1.2) and suitable initial and boundary conditions hold. Here the goal is to transform an initial phase distribution $y_0 : \Omega \to \mathbb{R}$ to some desired phase pattern $y_T : \Omega \to \mathbb{R}$ at a given final time $T$. Moreover throughout the entire time interval the distribution additionally remains close to $y_d$. In the
formulation (1.2) the potential $\psi$ is a smooth polynomial. Hence, $y$ attains values different from $\pm 1$ in the whole domain $\Omega$ and this is a disadvantage from the numerical point of view, where the solution has to be computed on the whole domain instead on the interface. Thus, to overcome this drawback we plan to use an Allen-Cahn variational inequality instead, i.e. using the obstacle potential

$$\psi(y) = \begin{cases} \frac{1}{2}(1 - y^2) & \text{if } |y| \leq 1, \\ \infty & \text{if } |y| > 1. \end{cases}$$

Introducing $\psi_0(y) := \frac{1}{2}(1 - y^2)$ and the indicator function

$$I_{[-1,1]}(y) := \begin{cases} 0 & \text{if } |y| \leq 1, \\ \infty & \text{if } |y| > 1, \end{cases}$$

we obtain

$$\psi(y) = \psi_0(y) + I_{[-1,1]}(y).$$

Then the object is given by values identical to 1. The interface $|y| < 1$ now has a small finite thickness proportional to $\varepsilon$. An additional advantage will be that as a consequence one only has to compute the solution in a narrow band around the interface.

**Notations and general assumptions** In the sequel we always denote by $\Omega \subset \mathbb{R}^d$ a bounded domain (with spatial dimension $d$) with boundary $\Gamma = \partial \Omega$. The outer unit normal on $\Gamma$ is denoted by $n$. We denote by $L^p(\Omega), W^{k,p}(\Omega)$ for $1 \leq p \leq \infty$ the Lebesgue- and Sobolev spaces of functions on $\Omega$ with the usual norms $\| \cdot \|_{L^p(\Omega)}, \| \cdot \|_{W^{k,p}(\Omega)}$, and we write $H^k(\Omega) = W^{k,2}(\Omega)$, see [1]. For a Banach space $X$ we denote its dual by $X^*$, the dual pairing between $f \in X^*, g \in X$ will be denoted by $\langle f, g \rangle_{X^*,X}$. If $X$ is a Banach space with the norm $\| \cdot \|_X$, we denote for $T > 0$ by $L^p(0,T;X) \ (1 \leq p \leq \infty)$ the Banach space of all (equivalence classes of) Bochner measurable functions $u : (0,T) \rightarrow X$ such that $\|u(\cdot)\|_X \in L^p(0,T)$. We set $\Omega_T := (0,T) \times \Omega$, $\Gamma_T := (0,T) \times \Gamma$. ”Generic” positive constants are denoted by $C$. Furthermore we define following time dependent Sobolev spaces by

$$W(0,T) := L^2(0,T;H^1(\Omega)) \cap H^1(0,T;H^1(\Omega)^*),$$

$$V := L^2(0,T;H^2(\Omega)) \cap L^\infty(0,T;H^1(\Omega)) \cap H^1(\Omega_T).$$

Moreover specially for $\dim \Omega \leq 3$ we will use following Sobolev embeddings

$$H^1(\Omega) \hookrightarrow L^{p^*}(\Omega), \quad p^* \in [1,6],$$

(1.4)
and
\[ W^{3/2,2}(\Omega) \hookrightarrow W^{1,q}(\Omega), \quad q \in [1,3]. \] (1.5)
Besides we also will use following embedding
\[ H^1(0,T;H^1(\Omega)) \cap L^2(0,T;H^2(\Omega)) \hookrightarrow C([0,T];W^{3/2,2}(\Omega)). \] (1.6)

For the rest of the paper we make the following assumption:
\( (\text{H}) \) Assume \( \Omega \subset \mathbb{R}^d \) is a bounded domain and either convex or has a \( C^{1,1} \) boundary and let \( T > 0 \) be a positive time.

Hence, given an initial phase distribution \( y(0,\cdot) = y_0 : \Omega \to [-1,1] \) at time \( t = 0 \) the interface motion can be modeled by the steepest descent of \( E \) with respect to the \( L^2 \)-norm which results, after suitable rescaling of time, in the following Allen-Cahn equation
\[ \varepsilon \partial_t y = -\text{grad}_{L^2}E(y) = \varepsilon \Delta y - \frac{1}{\varepsilon} (\psi'_0(y) + \zeta^*), \]
where \( \zeta^* \in \partial I_{[-1,1]} \) and \( \partial I_{[-1,1]} \) denotes the subdifferential of \( I_{[-1,1]} \). This equation leads to the following variational inequality
\[ \varepsilon (\partial_t y, \eta - y)_{L^2(\Omega)} + \varepsilon (\nabla y, \nabla (\eta - y))_{L^2(\Omega)} + \frac{1}{\varepsilon} \psi'_0(y, \eta - y)_{L^2(\Omega)} \geq 0, \] (1.7)
which has to hold for almost all \( t \in [0,T] \) and all \( \eta \in H^1(\Omega) \) with \( |\eta| \leq 1 \) a.e. in \( \Omega \).

1.1 Allen-Cahn MPEC

Our overall optimization problem is now stated as
\[
\left\{ \begin{array}{ll}
\min & J(y,u), \\
\text{over} & y : [0,T] \times \Omega \to [-1,1]; \ u : [0,T] \times \Omega \to \mathbb{R}, \\
\text{s.t.} & \varepsilon (\partial_t y, \eta - y) + \varepsilon (\nabla y, \nabla (\eta - y)) \geq \left( -\frac{1}{\varepsilon} \psi'_0(y) + u, \eta - y \right), \\
& y(0) = y_0 : \Omega \to [-1,1], \\
& \text{for almost all } t \in [0,T] \text{ and all } \eta : \Omega \to [-1,1].
\end{array} \right. \]

The resulting optimization problem \( (\mathcal{P}) \) belongs to the problem class of so-called MPECs (Mathematical Programs with Equilibrium Constraints) which
are hard to handle for several reasons. Indeed, it is well known that the variational inequality condition (or equivalently in MPCC case the complementarity conditions) occurring as constraints in the minimization problem violates all the known classical NLP (nonlinear programming) constraint qualifications. Hence, the existence of Lagrange multipliers cannot be inferred from standard theory. Many authors (for example, Barbu [2], Mignot [25], Mignot-Puel [26], Bonnans-Tiba [8], Friedman [14, 15], Bermudez-Saguez [7], Bergounioux [4], Bergounioux-Zidani [6], Bergounioux-Lenhart [5], He [17]) have already considered control problems for elliptic and parabolic variational inequalities and different mathematical methods have been used and developed to tackle these problems. Barbu [2, 3] studies approximations (penalization) of the variational inequality which lead to optimal control problems governed by variational equations. Then he gets existence results and optimality conditions using a passage to the limit in the approximation process. Other authors have been considered in many different scenarios venues aspects, for example relaxation-regularization [4], Pontryagin’s principle [8, 6], Ekeland’s principle with diffuse perturbations [6], conical derivatives [25]. Nevertheless, the optimal control of variational inequalities is still a very active field of research especially concerning their numerical treatment, see e.g. the recent publications Ito-Kunisch [21, 22], Hintermüller-Kopacka [19], Hintermüller-Tber [20]. These recent works have in common that they apply mathematical methods and proof-steps which highly motivate numerical algorithms. Following these ideas here we apply a smoothed penalization approach to our problem $(P)$. We introduce approximate problems with their first-order necessary optimality conditions and show that in the limit of a vanishing approximation parameter certain generalized first-order necessary conditions of optimality are derived. The difference of this paper to [22] is as follows: In [22] a Moreau-Yosida regularization to the variational inequality is used to get a monotone equation, which fulfills the weak maximum principle. The first-order optimality conditions of the approximated problems are derived and by using additional regularity conditions certain generalized first-order necessary conditions of optimality are derived in the limit. Our problem differs in some key aspects from the problem treated in [22]. In our case having a bilevel, bi-obstacle problem, we use a different penalization, see [11] and obtain a non-monotone semi-linear parabolic equation, which does not fulfill the weak maximum principle. In fact we can show without additional assumptions that in the limit optimality conditions which are stronger than weak optimality conditions but weaker than $C-$stationarity conditions (for different notions of stationarity for MPECs we refer to [23]) arise. Moreover to get convergence individually for the dual multipliers associated to the bi-obstacles we use additional regularity assumptions and obtain generalized
first-order necessary conditions.

Our work is organized as follows. In section 2 we analyse our state equation. Most of the results of this section can be found in different papers with different penalizations, see e.g. [9], so the results are not new. But the penalization functions are different from the ones used in [9]. So we decided to keep our work self-contained and for convenience of the reader, we proved once again well-known results for our special penalization functions. In section 3 we introduce the penalized optimal control problem, prove the existence of minimizers and establish for the case when the spatial dimension is less than three the first-order optimality system. In the last section 4 we show that in the limit of the vanishing penalization parameter certain generalized optimality conditions appear.

2 Allen-Cahn variational inequality

In this section we collect and extend known results about the Allen-Cahn variational inequality, see e.g. [9] and references therein. The Allen-Cahn variational inequality is given by:

\[(ACVI)\] Assume \((H)\) hold. Then for given initial data \(y_0 \in H^1(\Omega)\) with \(|y_0| \leq 1\) a.e. in \(\Omega\) and \(u \in L^2(\Omega_T)\) find \(y \in V\) such that \(y(0) = y_0, |y| \leq 1\) a.e. in \(\Omega_T\) and

\[
\varepsilon (\partial_t y, \eta - y)_{L^2(\Omega)} + \varepsilon (\nabla y, \nabla (\eta - y))_{L^2(\Omega)} + \frac{1}{\varepsilon} (\psi'_0(y), \eta - y)_{L^2(\Omega)} \geq (u, \eta - y)_{L^2(\Omega)},
\]

which has to hold for almost all \(t\) and all \(\eta \in H^1(\Omega)\) with \(|\eta| \leq 1\) a.e. in \(\Omega\).

Due to [9] the problem \((ACVI)\) can be reformulated with the help of Lagrange multipliers \(\mu^\oplus\) and \(\mu^\ominus\) corresponding to the inequality constraints \(y \leq 1\) and \(y \geq -1\).

**Lemma 1.** Assume \((H)\) hold. Let be given \(y_0 \in H^1(\Omega)\) with \(|y_0| \leq 1\) a.e. in \(\Omega\) and \(u \in L^2(\Omega_T)\). A function \(y \in V\) solves \((ACVI)\) if there exist
\( \mu^\oplus, \mu^\ominus \in L^2(\Omega_T) \) such that

\[
\varepsilon \partial_t y - \varepsilon \Delta y + \frac{1}{\varepsilon} \psi'_0(y) + \frac{1}{\varepsilon} \mu^\oplus - \frac{1}{\varepsilon} \mu^\ominus = u \quad \text{a.e. in } \Omega_T, \tag{2.1}
\]

\[
y(0) = y_0 \quad \text{a.e. in } \Omega, \quad n \cdot \nabla y = 0 \quad \text{a.e. on } \Gamma_T, \tag{2.2}
\]

\[
|y| \leq 1 \quad \text{a.e. in } \Omega_T, \tag{2.3}
\]

\[
\mu^\oplus(y - 1) = 0, \mu^\ominus(y + 1) = 0 \quad \text{a.e. in } \Omega_T, \tag{2.4}
\]

\[
\mu^\oplus \geq 0, \mu^\ominus \geq 0 \quad \text{a.e. in } \Omega_T. \tag{2.5}
\]

The proof of Lemma 1 for \( u \equiv 0 \) can be found in [9]. The extension of the proof to our case \( u \not\equiv 0 \) is straightforward.

**Theorem 1.** Assume \( (H) \) hold. Let be given \( y_0 \in H^1(\Omega) \) with \( |y_0| \leq 1 \) a.e. in \( \Omega \) and \( u \in L^2(\Omega_T) \). Then there exists a unique solution \( (y, \mu^\oplus, \mu^\ominus) \in V \times L^2(\Omega_T) \times L^2(\Omega_T) \) of (2.1)-(2.5).

**Remark 1.** We show the existence of a solution \( y \) together with unique Lagrange multipliers \( \mu^\oplus \) and \( \mu^\ominus \) by a penalty approach for the inequality constraint \( |y| \leq 1 \). In particular, we replace the indicator function in \( \psi \) by terms penalizing deviations of \( y \) from the interval \([-1, 1]\).

**Proof of Theorem 1.**

**Step 1: Penalization.** We utilize the ideas in [11]. Given \( 0 < \sigma < 1 \) we introduce the penalized potential function \( \psi_{\sigma} \in C^2(\mathbb{R}) \) by

\[
\psi_{\sigma}(r) := \psi_0(r) + (\psi^\oplus_{\sigma}(r) + \psi^\ominus_{\sigma}(r)), \tag{2.6}
\]

where the convex functions \( \psi^\oplus_{\sigma}, \psi^\ominus_{\sigma} \in C^2(\mathbb{R}) \) are defined by

\[
\psi^\oplus_{\sigma}(r) := \begin{cases} 
\frac{1}{27} \left( r - \left( 1 + \frac{\sigma}{2} \right) \right)^2 + \frac{\sigma}{24} & \text{for } r \geq 1 + \sigma, \\
\frac{1}{6\sigma^2} (r - 1)^3 & \text{for } 1 < r < 1 + \sigma, \\
0 & \text{for } r \leq 1,
\end{cases}
\]

\[
\psi^\ominus_{\sigma}(r) := \begin{cases} 
0 & \text{for } r \geq -1, \\
-\frac{1}{6\sigma^2} (r + 1)^3 & \text{for } -1 - \sigma < r < -1, \\
\frac{1}{27} \left( r + \left( 1 - \frac{\sigma}{2} \right) \right)^2 + \frac{\sigma}{24} & \text{for } r \leq -1 - \sigma.
\end{cases}
\]

It is easy to show, see [11], that for \( \sigma < \frac{1}{4} \)

\[
\psi_{\sigma}(r) \geq -C_0 \sigma, \tag{2.7}
\]
where $C_0$ is a positive constant bounded independently of $\sigma$. Moreover we define $eta_{\sigma}^{\oplus}, \beta_{\sigma}^{\ominus}, \beta_{\sigma} \in C^1(\mathbb{R})$ by

$$\beta_{\sigma}^l := \sigma (\psi_{\sigma}^l)'(r), \quad l \in \{\oplus, \ominus\}, \quad (2.8)$$

and $\beta_{\sigma} := \beta_{\sigma}^{\oplus} + \beta_{\sigma}^{\ominus}$. We note that $\beta_{\sigma}^{\oplus}, \beta_{\sigma}^{\ominus}, \beta_{\sigma}$ are Lipschitz continuous functions, where

$$0 \leq (\beta_{\sigma}^l)' \leq 1, \quad l \in \{\oplus, \ominus\}. \quad (2.9)$$

Furthermore because of (2.8) we have

$$0 \leq (\psi_{\sigma}^l)'' \leq \frac{1}{\sigma}, \quad l \in \{\oplus, \ominus\}. \quad (2.9)$$

Inserting (2.6) into (1.3) gives the penalized energy

$$E_{\sigma}(y) := \int_{\Omega} \left( \frac{\varepsilon}{2} |\nabla y|^2 + \frac{1}{\varepsilon} \psi_{\sigma}(y) \right) dx, \quad 0 < \sigma < \frac{1}{4}, \quad (2.10)$$

and the following penalized semi-linear parabolic problem:

$$\varepsilon \partial_t y_{\sigma} - \varepsilon \Delta y_{\sigma} + \frac{1}{\varepsilon} \psi_{\sigma}'(y_{\sigma}) + \frac{1}{\varepsilon \sigma} (\beta_{\sigma}^{\oplus}(y_{\sigma}) + \beta_{\sigma}^{\ominus}(y_{\sigma})) = u_{\sigma} \quad \text{in } \Omega_T, \quad (2.10)$$

$$y_{\sigma}(0) = y_0 \quad \text{in } \Omega, \quad n \cdot \nabla y_{\sigma} = 0 \quad \text{on } \Gamma_T. \quad (2.11)$$

**Step 2: Existence of solutions to (2.10)-(2.10) and a priori estimates.**

**Lemma 2.** Assume (H) hold and $\sigma \in (0, 1/4)$. Let be given $y_0 \in H^1(\Omega)$ with $|y_0| \leq 1$ a.e. in $\Omega$ and $u_\sigma \in L^2(\Omega_T)$. Then there exists a unique solution $y_\sigma \in V$ of (2.10)-(2.11). Moreover for a sequence $\{y_\sigma\}$ bounded in $L^2(\Omega_T)$ uniformly in $\sigma \in (0, 1/4)$ we have

$$y_\sigma \quad \text{bounded in } V \quad \text{uniformly in } \sigma \in (0, 1/4),$$

$$\frac{1}{\sigma} \beta_{\sigma}^l \quad \text{bounded in } L^2(\Omega_T) \quad \text{uniformly in } \sigma \in (0, 1/4),$$

where $l \in \{\oplus, \ominus\}$.

**Proof of Lemma 2.** The existence of a solution to (2.10)-(2.11) follows by using a standard Galerkin approximation and then passing to the limit, see [9, 11]. The a priori estimates (uniformly in $\sigma \in (0, 1/4)$) are derived by testing (2.10) by suitable testfunctions like $y_\sigma, \partial_t y_\sigma, -\Delta y_\sigma$ and $\frac{1}{\sigma} \beta_{\sigma}^l, l \in \{\oplus, \ominus\}$. The key a priori estimate is the energy estimate, which we get by
testing (2.10) with \( \partial_t y_\sigma \), carrying out partial integration and using \( E(y_0) = E_\sigma(y_\sigma(0)) \):

\[
\varepsilon \| \partial_t y_\sigma \|_{L^2(\Omega_t)}^2 + E_\sigma(y_\sigma(t)) = E(y_0) + \int_0^t (\partial_t y_\sigma, u_\sigma)_{L^2(\Omega)} ds.
\] (2.12)

By (2.7) we have \( E_\sigma(y_\sigma(t)) \geq \frac{\varepsilon}{2} \| \nabla y_\sigma(t) \|_{L^2(\Omega)}^2 - \frac{C_0(\Omega)}{4} \). Since \( y_0 \in H^1(\Omega) \) there exists a positive \( C \) independent of \( \sigma \) such that \( E(y_0) < C \). Moreover using Young’s inequality for the last term and the uniform boundedness of \( \{u_\sigma\} \) in \( L^2(\Omega_T) \) we end up with

\[
\| \partial_t y_\sigma \|_{L^2(\Omega_T)} + \| \nabla y_\sigma \|_{L^\infty(0,T;L^2(\Omega))} \leq C,
\] (2.13)

where \( C \) is independent of \( \sigma \) and \( T \). Furthermore we test (2.10) by \( y_\sigma \) and note that \( (\beta^\sigma_\sigma(y_\sigma) + \beta_\sigma^\sigma(y_\sigma))y_\sigma \geq 0 \), hence we get by standard calculations

\[
\| y_\sigma(t) \|_{L^2(\Omega)}^2 + \| \nabla y_\sigma \|_{L^2(\Omega_t)}^2 \leq \| y_0 \|_{L^2(\Omega)}^2 + \frac{2}{\varepsilon^2} \| y_\sigma \|_{L^2(\Omega_t)}^2 + \frac{2}{\varepsilon} \int_0^t (y_\sigma, u_\sigma)_{L^2(\Omega)} ds.
\]

Using Young’s inequality for the last term and that \( \{u_\sigma\} \) is uniformly bounded in \( L^2(\Omega_T) \) and a Gronwall argument there exists a positive constant \( C(T) \) which depends on \( T \) but not on \( \sigma \) such that

\[
\| y_\sigma \|_{L^\infty(0,T;L^2(\Omega))} + \| \nabla y_\sigma \|_{L^2(\Omega_T)} \leq C(T).
\] (2.14)

Hence, (2.12) and (2.14) give that \( \{y_\sigma\} \) is uniformly bounded in \( L^\infty(0,T;H^1(\Omega)) \cap H^1(\Omega_T) \). Moreover we multiply (2.10) by \( -\Delta y_\sigma \) and integrate. After integration by parts we obtain

\[
\| \nabla y_\sigma(t) \|_{L^2(\Omega)}^2 + \| \Delta y_\sigma \|_{L^2(\Omega_t)}^2 + \int_0^T \int_{\Omega} \frac{1}{\sigma \varepsilon^2} (\beta^\sigma_\sigma(y_\sigma) + \beta_\sigma^\sigma(y_\sigma)) \| \nabla y_\sigma \|^2 dx =
\]

\[
= \| \nabla y_0 \|_{L^2(\Omega)}^2 + \frac{2}{\varepsilon} \| \nabla y_\sigma \|_{L^2(\Omega_t)}^2 - \frac{1}{\varepsilon} \int_0^t (\Delta y_\sigma, u_\sigma)_{L^2(\Omega)} ds.
\]

By virtue of \( (\beta^\sigma_\sigma(y_\sigma) + \beta_\sigma^\sigma(y_\sigma))' \geq 0 \), a Gronwall argument and elliptic regularity theory we obtain that \( \{y_\sigma\} \) is uniformly bounded in \( L^2(0,T;H^2(\Omega)) \). Hence, \( \{y_\sigma\} \) is uniformly bounded in \( \mathcal{V} \). For details, see e.g. [9, 11]. Moreover since \( \beta^\sigma_\sigma(y_\sigma) \cdot \beta_\sigma^\sigma(y_\sigma) = 0 \) we obtain from (2.10) and the a priori estimates on \( y_\sigma \) that

\[
\| \frac{1}{\sigma} \beta^\sigma_\sigma \|_{L^2(\Omega_T)} + \| \frac{1}{\sigma} \beta_\sigma^\sigma \|_{L^2(\Omega_T)} \leq C.
\] (2.15)
End of Proof of Lemma 2.

Step 3: Convergence result. Defining

$$\mu^\oplus := \frac{1}{\sigma} \beta^\oplus(y_\sigma) \quad \text{and} \quad \mu^\ominus := -\frac{1}{\sigma} \beta^\ominus(y_\sigma),$$

we reformulate (2.10)-(2.11) and obtain

$$\varepsilon \partial_t y_\sigma - \varepsilon \Delta y_\sigma + \frac{1}{\varepsilon} \psi'_0(y_\sigma) + \frac{1}{\varepsilon} \mu^\oplus - \frac{1}{\varepsilon} \mu^\ominus = u_\sigma \quad \text{in } \Omega_T, \quad (2.16)$$

$$y_\sigma(0) = y_0 \quad \text{in } \Omega, \quad n \cdot \nabla y_\sigma = 0 \quad \text{on } \Gamma_T. \quad (2.17)$$

**Lemma 3.** Let the assumption of Lemma 2 hold and let \(\{u_\sigma\}\) be a sequence in \(L^2(\Omega_T)\), \(u \in L^2(\Omega_T)\) such that \(u_\sigma \to u\) weakly in \(L^2(\Omega_T)\). Furthermore let \((y_\sigma, \mu^\oplus, \mu^\ominus) \in V \times L^2(\Omega_T) \times L^2(\Omega_T)\) denote the solution of (2.16)-(2.17). Then there exist \((y, \mu^\oplus, \mu^\ominus) \in V \times L^2(\Omega_T) \times L^2(\Omega_T)\) and a subsequence still denoted by \(\{(y_\sigma, \mu^\oplus, \mu^\ominus)\}\), such that as \(\sigma \searrow 0\) we have

- \(y_\sigma \rightharpoonup y\) strongly in \(L^2(0,T;H^2(\Omega))\),
- \(y_\sigma \rightharpoonup y\) weakly in \(H^1(\Omega_T)\),
- \(y_\sigma \rightharpoonup y\) weakly-star in \(L^\infty(0,T;H^1(\Omega))\),
- \(\mu^\oplus_\sigma \rightharpoonup \mu^\oplus\) weakly in \(L^2(\Omega_T)\),
- \(\mu^\ominus_\sigma \rightharpoonup \mu^\ominus\) weakly in \(L^2(\Omega_T)\).

The limit element \((y, \mu^\oplus, \mu^\ominus, u)\) satisfies (2.1)-(2.5).

**Proof of Lemma 3.** The convergence results are direct consequences of the estimates given by Lemma 2. Moreover we get from the above estimates

- \(y_\sigma \to y\) strongly in \(L^2(\Omega_T)\),
- \(y_\sigma \to y\) a.e. in \(\Omega_T\).

Because of the convergence results (2.16)-(2.17) converge to (2.1)-(2.2). For \(l \in \{\oplus, \ominus\}\) the set \(\{\mu^l_\sigma \in L^2(\Omega_T) \mid \mu^l_\sigma \geq 0\ \text{a.e. in } \Omega_T\}\) is convex and closed and hence weakly closed and we obtain \(\mu^l \geq 0\) a.e. in \(\Omega_T\) and (2.5) is proven.

To prove (2.3) we define \(\beta^l(r) : \mathbb{R} \to \mathbb{R}, l \in \{\oplus, \ominus\}\) as follows:

$$\beta^\oplus(r) := \lim_{\sigma \searrow 0} \beta^\oplus_\sigma(r) = \begin{cases} r - 1 & \text{for } r \geq 1, \\ 0 & \text{for } r \leq 1, \end{cases}$$

$$\beta^\ominus(r) := \lim_{\sigma \searrow 0} \beta^\ominus_\sigma(r) = \begin{cases} 0 & \text{for } r \geq -1, \\ r + 1 & \text{for } r \leq -1. \end{cases}$$
Furthermore we define $\beta := \beta^\oplus + \beta^\ominus$. We note that $\beta^l, l \in \{\oplus, \ominus\}$ and $\beta$ are Lipschitz continuous functions, and that for $l \in \{\oplus, \ominus\}$

$$|\beta^l(r) - \beta^l_s(r)| \leq \frac{\sigma}{2} \quad \forall r \in \mathbb{R}, \text{ and } |\beta^l(r) - \beta^l(s)| \leq |r - s| \quad \forall r, s \in \mathbb{R}.$$  

(2.18)

Using (2.18), the Lipschitz continuity of $\beta$, and (2.15) we obtain for every $\eta \in L^2(\Omega_T)$

$$\int_0^T \langle (\beta(y), \eta) \rangle_{L^2(\Omega)} dt \leq \int_0^T (\|\beta(y) - \beta(y_\sigma)\|_{L^2(\Omega)} + \|\beta(y_\sigma) - \beta(1)\|_{L^2(\Omega)} + \|\beta(1)\|_{L^2(\Omega)}) \|\eta\|_{L^2(\Omega)} dt \leq C \|\eta\|_{L^2(\Omega_T)} + \sigma \|\eta\|_{L^2(\Omega_T)},$$

so that from the result that $y_\sigma$ converges strongly to $y$ in $L^2(\Omega_T)$, $\beta(y) = 0$ a.e. in $\Omega_T$ and hence $|y| \leq 1$ a.e. in $\Omega_T$. In addition using the monotonicity of $\beta^\oplus$ and $\beta^\ominus(1) = 0$ we obtain

$$\mu^\oplus(y_\sigma - 1) = \frac{1}{\sigma} \beta^\oplus(y_\sigma)(y_\sigma - 1) = \frac{1}{\sigma} [\beta^\oplus(y_\sigma) - \beta^\oplus(1)](y_\sigma - 1) \geq 0.$$  

Since $y_\sigma \rightarrow y$ strongly in $L^2(\Omega_T)$ and $\mu^\oplus \rightarrow \mu^\oplus$ weakly in $L^2(\Omega_T)$ we get

$$\int_{\Omega_T} \mu^\oplus(y - 1) dx dt = \lim_{\sigma \searrow 0} \int_{\Omega_T} \mu^\oplus(y_\sigma - 1) dx dt \geq 0.$$  

Since $(y - 1) \leq 0$ a.e. in $\Omega_T$ and $\mu^\oplus \geq 0$ a.e. in $\Omega_T$ we hence deduce

$$\mu^\oplus(y - 1) = 0 \quad \text{a.e. in } \Omega_T.$$  

An analogue argumentation gives $\mu^\ominus(y + 1) = 0$ a.e. in $\Omega_T$ and (2.4) is proven.

End of Proof of Lemma 3.

Step 4: Uniqueness. It remains to show uniqueness. Assume that there are two solutions $(y_i, \mu^\oplus_i, \mu^\ominus_i)$, $i = 1, 2$. Defining $\overline{y} := y_1 - y_2$, $\overline{\mu} := \mu^\oplus_1 - \mu^\oplus_2$ for $l \in \{\oplus, \ominus\}$ and multiplying the difference of the equation (2.1) for $y_1$ and $y_2$ with $\overline{y}$ gives after integration

$$\varepsilon \frac{d}{dt} \|\overline{y}\|_{L^2(\Omega)}^2 + \varepsilon \|\nabla \overline{y}\|_{L^2(\Omega)}^2 + \frac{1}{\varepsilon} \int_{\Omega} \mu^\oplus \overline{y} dx - \frac{1}{\varepsilon} \int_{\Omega} \mu^\ominus \overline{y} dx = \frac{1}{\varepsilon} \|\overline{y}\|_{L^2(\Omega)}^2.$$
3 PENALIZED OPTIMAL CONTROL PROBLEM

The complementary condition (2.4) imply that the terms $\mu \oplus y$ and $-\mu \ominus y$ are non-negative. We hence deduce
\[ \varepsilon \frac{d}{dt} \|y\|^2_{L^2(\Omega)} + \varepsilon \|\nabla y\|^2_{L^2(\Omega)} \leq \frac{1}{\varepsilon} \|y\|^2_{L^2(\Omega)}. \]
A Gronwall argument now gives uniqueness of $y$. 

By virtue of Lemma 1 we can reformulate our overall optimization problem $(\mathcal{P})$ as a mathematical program with complementarity constraints (MPCC).

\[
\begin{aligned}
    \min_{(y,u)} & \quad J(y,u), \\
    \text{over} & \quad (y,u) \in \mathcal{V} \times L^2(\Omega_T), \\
    \text{s.t.} & \quad (2.10) - (2.11).
\end{aligned}
\]

3 Penalized optimal control problem

For $\sigma \in (0, 1/4)$ we define the penalized optimal control problem by

\[
\begin{aligned}
    (\mathcal{CP})_{\sigma} : \min_{(y,u)} & \quad J(y,u), \\
    \text{over} & \quad (y,u) \in \mathcal{V} \times L^2(\Omega_T), \\
    \text{s.t.} & \quad (2.10) - (2.11).
\end{aligned}
\]

3.1 Existence of an optimal control

**Theorem 2.** Assume $(H)$ hold and $\sigma \in (0, 1/4)$. The penalized optimal control problem $(\mathcal{CP})_{\sigma}$ has at least a minimizer.

**Proof.** For $\sigma \in (0, 1/4)$ let
\[ D_{\sigma} := \{(y_{\sigma}, u_{\sigma}) \in \mathcal{V} \times L^2(\Omega_T) : (y_{\sigma}, u_{\sigma}) \text{ satisfy } (2.10) - (2.11)\} \]
denote the feasible set of $(\mathcal{CP})_{\sigma}$. Let $\overline{u}_{\sigma} \in L^2(\Omega_T)$ be arbitrary but fixed and $y_{\sigma}(\overline{u}_{\sigma}) \in \mathcal{V}$ be the solution of (2.10)-(2.11) given by Lemma 2. Then $(y_{\sigma}(\overline{u}_{\sigma}), \overline{u}_{\sigma}) \in D_{\sigma}$. Hence the feasible set is nonempty. Furthermore, the
cost functional $J$ is bounded from below. Now let $\{(y_{\sigma,k}, u_{\sigma,k})\} \subset D_\sigma$ be a minimizing sequence such that
\[
\lim_{k \to \infty} J(y_{\sigma,k}, u_{\sigma,k}) = \inf_{(y_{\sigma}, u_\sigma) \in D_\sigma} J(y_{\sigma}, u_\sigma) := d < \infty.
\]
Then, we get
\[
\begin{align*}
&u_{\sigma,k} \text{ bounded in } L^2(\Omega_T) \text{ uniformly in } k, \\
y_{\sigma,k} \text{ bounded in } L^2(\Omega_T) \text{ uniformly in } k,
\end{align*}
\]
Moreover by using Lemma 2 it follows that $\{y_{\sigma,k}\}$ is bounded in $V$ uniformly in $k$. Hence, there exist $(y_{\sigma}, u_\sigma) \in V \times L^2(\Omega_T)$
such that on a subsequence (denoted the same) $u_{\sigma,k} \to u_\sigma$ weakly in $L^2(\Omega_T)$ and as $k \to \infty$
\[
\begin{align*}
y_{\sigma,k} &\to \overline{y}_{\sigma} \text{ weakly in } L^2(0,T; H^2(\Omega)), \\
y_{\sigma,k} &\to \overline{y}_{\sigma} \text{ weakly in } H^1(\Omega_T), \\
y_{\sigma,k} &\to \overline{y}_{\sigma} \text{ strongly in } L^2(\Omega_T), \\
y_{\sigma,k} &\to \overline{y}_{\sigma} \text{ weakly-star in } L^\infty(0,T; H^1(\Omega)), \\
y_{\sigma,k} &\to \overline{y}_{\sigma} \text{ a.e. in } \Omega_T.
\end{align*}
\]
Besides applying interpolation arguments we obtain that $V$ embeds continuously into $C([0,T]; H^1(\Omega))$. Hence, the evaluation of $y_{\sigma,k}$ at the final time $y_{\sigma,k}(T) \in H^1(\Omega)$ is well-defined. By Rellich-Kondrachov theorem it follows that $H^1(\Omega)$ compactly embeds into $L^2(\Omega)$. Hence, it follows that as $k \to \infty$
y_{\sigma,k}(T) \to \overline{y}_{\sigma}(T) \text{ strongly in } L^2(\Omega). \quad (3.1)
Because of the Lipschitz continuity of $\frac{1}{\sigma} \beta^l_\sigma = (\psi^l_\sigma)'$, $l \in \{\oplus, \ominus\}$ for fixed $\sigma \in (0, 1/4)$, we have as $k \to \infty$
\[
\mu^l_{\sigma,k} \to \overline{\mu}^l_{\sigma} \text{ strongly in } L^2(\Omega_T),
\]
for $l \in \{\oplus, \ominus\}$. Therefore,
\[
\epsilon \partial_t \overline{y}_{\sigma} - \epsilon \Delta \overline{y}_{\sigma} + \frac{1}{\epsilon} \psi_0' (\overline{y}_{\sigma}) + \frac{1}{\epsilon} \nu \overline{\mu}^l_{\sigma} - \frac{1}{\epsilon} \bar{\nu} \overline{\mu}^l_{\sigma} = \overline{u}_{\sigma} \quad \text{in } \Omega_T, \\
y_{\sigma}(0) = y_0, \quad n \cdot \nabla \overline{y}_{\sigma} = 0 \quad \text{a.e. on } \Gamma_T.
\]
The weakly lower semi-continuity of $J$ finally yields
\[
J(\overline{y}_{\sigma}, \overline{u}_{\sigma}) \leq \lim_{k \to \infty} J(y_{\sigma,k}, u_{\sigma,k}) = d.
\]
Hence \((\bar{y}_{\sigma}, \bar{u}_{\sigma})\) is a minimizer of \((C\!P)_{\sigma}\). \(\square\)

As far as globally optimal points are concerned, we find that solutions of the penalized optimal control problem \((C\!P)_{\sigma}\) converge to a solution of the problem \((C\!P)\), as the following theorem shows.

**Theorem 3.** Assume \((H)\) hold and \(\sigma \in (0, 1/4)\). Denote by \((y_{\sigma}, u_{\sigma})\) the minimizers of the penalized optimal control problems \((C\!P)_{\sigma}\). Then there exists a minimizer \((\bar{y}, \bar{u}) \in \mathcal{V} \times L^2(\Omega_T)\) for the problem \((C\!P)\) such that on a subsequence of minimizers (still denoted by \((y_{\sigma}, u_{\sigma})\)) as \(\sigma \searrow 0\)

\begin{align*}
  u_{\sigma} &\rightarrow \bar{u} \quad \text{strongly in } L^2(\Omega_T), \\
  y_{\sigma} &\rightarrow \bar{y} \quad \text{weakly in } L^2(0, T; H^2(\Omega)), \\
  y_{\sigma} &\rightarrow \bar{y} \quad \text{weakly in } H^1(\Omega_T), \\
  y_{\sigma} &\rightarrow \bar{y} \quad \text{strongly in } L^2(\Omega_T), \\
  y_{\sigma} &\rightarrow \bar{y} \quad \text{weakly-star in } L^\infty(0, T; H^1(\Omega)), \\
  y_{\sigma} &\rightarrow \bar{y} \quad \text{a.e. in } \Omega_T.
\end{align*}

Furthermoe we have

\begin{equation}
  y_{\sigma}(T) \rightarrow y(T) \quad \text{strongly in } L^2(\Omega). \tag{3.3}
\end{equation}

**Proof.** Let \(\hat{u} \in L^2(\Omega_T)\) be fixed, and denote by \(y_{\sigma}(\hat{u}) \in \mathcal{V}\) the solution to (2.10)-(2.11). Hence, the estimate

\begin{equation}
  J(y_{\sigma}, u_{\sigma}) \leq J(y_{\sigma}(\hat{u}), \hat{u}) \tag{3.4}
\end{equation}

holds true for \(\sigma \in (0, 1/4)\). The boundedness of \(\{y_{\sigma}(\hat{u})\}\) given by Lemma 2 implies the boundedness of \(\{J(y_{\sigma}(\hat{u}), \hat{u})\}\). Using (3.4), we conclude that also \(\{u_{\sigma}\}\) is uniformly bounded in \(L^2(\Omega_T)\), and there exists \(\bar{u} \in L^2(\Omega_T)\) such that on a subsequence (also denoted by \(\{u_{\sigma}\}\)) as \(\sigma \searrow 0\)

\begin{equation*}
  u_{\sigma} \rightarrow \bar{u} \quad \text{weakly in } L^2(\Omega_T).
\end{equation*}

Then by Lemma 3 there exists \(\bar{y} \in \mathcal{V}\) and a subsequence still denoted by \(\{y_{\sigma}\}\) such that (3.2) holds. Moreover \(\mathcal{V}\) embeds continuously into \(C([0, T]; H^1(\Omega))\). Hence, the evaluation of \(y_{\sigma}\) at final time \(y_{\sigma}(T) \in H^1(\Omega)\) is well-defined. By Rellich-Kondrachov theorem it follows that \(H^1(\Omega)\) compactly embeds into \(L^2(\Omega)\). Hence (3.3) follows. Because of Lemma 3 the limit element \((\bar{y}, \bar{u})\) is feasible for \((C\!P)\). Now let \((y^*, u^*) \in \mathcal{V} \times L^2(\Omega_T)\) be a minimizer of \((C\!P)\). Due to the lower semi-continuity of the norm, (3.4) and Lemma 3, we find that

\begin{align*}
  J(y^*, u^*) &\leq J(\bar{y}, \bar{u}) \leq \liminf_{\sigma \searrow 0} J(y_{\sigma}, u_{\sigma}) \leq \limsup_{\sigma \searrow 0} J(y_{\sigma}, u_{\sigma}) \\
  &\leq \limsup_{\sigma \searrow 0} J(y_{\sigma}(u^*), u^*) = J(y^*, u^*). \tag{3.5}
\end{align*}
Therefore, \((\bar{y}, \bar{u})\) is optimal for \((CP)\). Furthermore, we see that as \(\sigma \searrow 0\)
\[ J(y_\sigma, u_\sigma) \to J(\bar{y}, \bar{u}), \]

hence \(\|u_\sigma\|_{L^2} \to \|\bar{u}\|_{L^2}\), which together with the weak convergence of \(\{u_\sigma\}\)
implies strong convergence of \(\{u_\sigma\}\) in \(L^2(\Omega_T)\).

\[ \square \]

### 3.2 Differentiability of the control-to-state mapping

For the derivation of first-order optimality conditions, it is essential to show the Gâteaux-differentiability of the control-to-state operator, mapping \(u_\sigma\) to \(y_\sigma\). But before we need a stability result.

**Definition 1.** Assume \((H)\) hold and \(\sigma \in (0, 1/4)\). Based on Lemma 2, we introduce the control-to-state operator \(S_\sigma : L^2(\Omega_T) \to V\), where \(y_\sigma := S_\sigma(u_\sigma)\) denotes the solution of (2.10)-(2.11) associated to \(u_\sigma\).

**Lemma 4.** Assume \((H)\) hold and \(\sigma \in (0, 1/4)\). Let \(u_\sigma^1 \in L^2(\Omega_T)\) and \(y_\sigma^i = S_\sigma(u_\sigma^i) \in V\), \((i = 1, 2)\). Then there exists a positive \(c_\sigma(\sigma)\) which depends on \(\sigma\) such that the following stability estimate hold:

\[ \|y_\sigma^1 - y_\sigma^2\|_V \leq c_\sigma(\sigma)\|u_\sigma^1 - u_\sigma^2\|_{L^2(\Omega_T)}. \]  

**Proof.** We define \(\tilde{u}_\sigma := u_\sigma^1 - u_\sigma^2\) and \(\tilde{y}_\sigma := y_\sigma^1 - y_\sigma^2\) and remark that \(\tilde{y}_\sigma\) satisfies the following initial-boundary value problem:

\[ \varepsilon \partial_t \tilde{y}_\sigma - \varepsilon \Delta \tilde{y}_\sigma - \frac{1}{\varepsilon} \tilde{y}_\sigma + \frac{1}{\varepsilon} \sum_{l=\oplus, \ominus} (\beta_{\sigma}^l(y_\sigma^1) - \beta_{\sigma}^l(y_\sigma^2)) = \tilde{u}_\sigma \quad \text{in} \quad \Omega_T, \]

\[ \tilde{y}_\sigma(0) = 0 \quad \text{in} \quad \Omega, \quad n \cdot \nabla \tilde{y}_\sigma = 0 \quad \text{on} \quad \Gamma_T. \]

Testing the differential equation by \(\tilde{y}_\sigma, \partial_t \tilde{y}_\sigma\) and \(-\Delta \tilde{y}_\sigma\) and using the Lipschitz continuity of \(\beta_{\sigma}^l, l \in \{\oplus, \ominus\}\), and applying analogue techniques like in the proof of Lemma 2 we get the desired result. \( \square \)

Suppose \(u_\sigma \in L^2(\Omega_T)\) and consider a perturbation \(h_\sigma \in L^2(\Omega_T)\). In preparation of the corresponding theorem, we now consider the following linearized version of (2.10)-(2.11):

\[ \varepsilon \partial_t y_\sigma^* - \varepsilon \Delta y_\sigma^* - \frac{1}{\varepsilon} \psi_{\sigma}'(y_\sigma) y_\sigma^* = h_\sigma \quad \text{in} \quad \Omega_T, \]  

\[ y_\sigma^*(0) = 0 \quad \text{in} \quad \Omega, \quad n \cdot \nabla y_\sigma^* = 0 \quad \text{on} \quad \Gamma_T, \]  

with given functions $y_\sigma, h_\sigma$. We remark that $y_\sigma = S_\sigma(u_\sigma)$ is the unique solution of the nonlinear state system (2.10)-(2.11) associated to reference control $u_\sigma$. In the following we will show that (3.6)-(3.7) admits a unique solution $y_\sigma^* \in \mathcal{V}$. This result is then used to establish the Gâteaux-differentiability of the solution operator $S_\sigma$ associated to (2.10)-(2.11).

**Lemma 5.** Assume $(H)$ hold and $\sigma \in (0,1/4)$. Then problem (3.6)-(3.7) admits a unique solution $y_\sigma^* \in \mathcal{V}$.

**Proof.** Since for every fixed $\sigma \in (0,1/4)$ the function $\psi''_\sigma(y_\sigma) \in L^\infty(\Omega_T)$, see (2.9), the existence of a unique solution $y_\sigma^* \in \mathcal{V}$ to (3.6)-(3.7) is a classical result, (see [12], Chapter 7, Theorem 3 and Theorem 5). □

We continue the derivation of first-order conditions with the Gâteaux differentiability of the control-to-state operator $S_\sigma$, which is one of the crucial points of the first-order analysis for $(CP)_\sigma$. However, using the analysis for the linearized equation yields the desired differentiability of $S_\sigma$. Afterwards, we reformulate the derivative of the objective functional by introducing an adjoint PDE system which leads to the first-order necessary optimality conditions in form of a Karush-Kuhn-Tucker (KKT) type optimality system.

**Theorem 4.** Assume $(H)$ hold, $\sigma \in (0,1/4)$ and dim $\Omega \leq 3$. The control-to-state operator $S_\sigma : L^2(\Omega_T) \to \mathcal{V}$ is Gâteaux-differentiable. The directional derivative $y_\sigma^* = S'_\sigma(u_\sigma)h_\sigma$ in direction $h_\sigma \in L^2(\Omega_T)$ is defined as the solution to the linearized problem (3.6)-(3.7) in $y_\sigma := S_\sigma(u_\sigma)$.

**Proof.** Denote $y_\tau^* = S_\sigma(u_\sigma + \tau h_\sigma), y_\sigma = S_\sigma(u_\sigma), r_\tau := y_\tau^* - y_\sigma - \tau y_\sigma^*, \tau > 0$. We have to prove

$$\|r_\tau^*\|_\mathcal{V} = o(\tau) \quad \text{as} \quad \tau \searrow 0.$$

(3.8)

The function $r_\tau^*$ fullfills

$$\varepsilon \partial_t r_\tau^* - \varepsilon \Delta r_\tau^* + \frac{1}{\varepsilon} G^\tau = 0 \quad \text{in} \quad \Omega_T,$$

(3.9)

$$r_\tau^*(0) = 0 \quad \text{in} \quad \Omega, \quad n \cdot \nabla r_\tau^* = 0 \quad \text{on} \quad \Gamma_T,$$

(3.10)

where

$$G^\tau := \psi'_\sigma(y_\tau^*) - \psi'_\sigma(y_\sigma) - \psi''_\sigma(y_\sigma)\tau y_\sigma^*.$$  

(3.11)

For almost every $(t,x) \in \Omega_T$ and $\xi \in [0,1]$ we define $g(\xi) := \psi'_\sigma(y_\sigma(t,x) + \xi(y_\sigma(t,x) - y_\sigma(t,x)))$. It is clear that $g'(\xi) = \psi''_\sigma(y_\sigma(t,x) + \xi(y_\sigma(t,x) - y_\sigma(t,x))) - \psi''_\sigma(y_\sigma(t,x)) - \xi \psi''_\sigma(y_\sigma(t,x))$. Theorem 4 follows from the Gâteaux-differentiability of $S_\sigma$, (3.8) and the definition of $g'(\xi)$.
3 PENALIZED OPTIMAL CONTROL PROBLEM

\[ y_\sigma(t, x))(y_\sigma^*(t, x) - y_\sigma(t, x)). \] Hence, by the mean value theorem there exists a \( \xi \in (0, 1) \) such that \( g(1) = g(0) + g'(\xi) \), which means

\[
\psi'_\sigma(y_\sigma^*(t, x)) - \psi'_\sigma(y_\sigma(t, x)) = \psi''(y_\sigma(t, x)) + \xi(y_\sigma^*(t, x) - y_\sigma(t, x))(y_\sigma^*(t, x) - y_\sigma(t, x)).
\]

Inserting (3.12) into (3.11) we get

\[
G^\tau = [\psi''(y_\sigma + \xi(y_\sigma^* - y_\sigma)) - \psi''(y_\sigma)](y_\sigma^* - y_\sigma) + \psi''(y_\sigma)r_\tau^\sigma \quad \text{a.e. in } \Omega_T
\]

and because of the Lipschitz continuity of \( \psi'' \) and (2.9) we obtain

\[
|G^\tau| \leq L|y_\sigma^* - y_\sigma|^2 + c(\sigma)|r_\tau^\sigma| \quad \text{a.e. in } \Omega_T,
\]

where \( L \) is the Lipschitz constant and \( c(\sigma) > 0 \) a constant depending on \( \sigma \).

Now we test (3.9) by \( r_\tau^\sigma \) and carry out partial integration

\[
\frac{1}{2}\|r_\tau^\sigma(t)\|^2_{L^2(\Omega)} + \|\nabla r_\tau^\sigma\|^2_{L^2(\Omega_T)} + \frac{1}{\varepsilon^2}(G^\tau, r_\tau^\sigma)_{L^2(\Omega_T)} = 0.
\]

Moreover using (3.13), Young’s inequality and the Sobolev embedding \( V \subset L^4(\Omega_T) \) for \( \dim(\Omega) \leq 3 \) we get

\[
|(G^\tau, r_\tau^\sigma)_{L^2(\Omega_T)}| \leq c(L)\|y_\sigma^* - y_\sigma\|^4_{L^4(\Omega_T)} + c(\sigma)\|r_\tau^\sigma\|^2_{L^2(\Omega_T)} \\
\leq c(L)\|y_\sigma^* - y_\sigma\|^4_V + c(\sigma)\|r_\tau^\sigma\|^2_{L^2(\Omega_T)},
\]

where \( c(L) > 0 \) is a constant depending on \( L \). Utilizing Lemma 4 we have

\[
|(G^\tau, r_\tau^\sigma)_{L^2(\Omega_T)}| \leq c(L, c_s(\sigma))\tau^4\|h_\sigma\|^4_{L^2(\Omega_T)} + c(\sigma)\|r_\tau^\sigma\|^2_{L^2(\Omega_T)} \\
\leq c(L, c_s(\sigma))\tau^4 + c(\sigma)\|r_\tau^\sigma\|^2_{L^2(\Omega_T)},
\]

where \( c(L, c_s(\sigma)) > 0 \) is a constant depending on \( L \) and \( c_s(\sigma) \) and finally

\[
\frac{1}{2}\|r_\tau^\sigma(t)\|^2_{L^2(\Omega)} + \|\nabla r_\tau^\sigma\|^2_{L^2(\Omega_T)} \leq c(L, c_s(\sigma), \varepsilon)\tau^4 + c(\sigma)\|r_\tau^\sigma\|^2_{L^2(\Omega_T)},
\]

where \( c(L, c_s(\sigma), \varepsilon) > 0 \) is a positive constant additionally depending on \( \varepsilon \).

A Gronwall argument gives

\[
\|r_\tau^\sigma\|_{L^\infty(0, T; L^2(\Omega))} \leq c(T, L, c_s(\sigma), \varepsilon, u_0, \sigma)\tau^2,
\]

where \( c(T, L, c_s(\sigma), \varepsilon, u_0, \sigma) > 0 \) depends on the given data of the problem. Testing (3.9) by \( \partial_\tau r_\tau^\sigma \) and carrying out partial integration we obtain

\[
\|\partial_\tau r_\tau^\sigma\|^2_{L^2(\Omega_T)} + \frac{1}{2}\|\nabla r_\tau^\sigma(t)\|^2_{L^2(\Omega)} + \frac{1}{\varepsilon^2}(G^\tau, \partial_\tau r_\tau^\sigma)_{L^2(\Omega_T)} = 0.
\]
Moreover by using (3.13), Young's inequality and again the Sobolev embedding $V \subset L^4(\Omega_T)$ for $\dim(\Omega) \leq 3$ and Lemma 4 we have

$$
\| (G^\tau, \partial_t r^\tau) \|_{L^2(\Omega_T)} \leq \frac{1}{2} \| G^\tau \|_{L^2(\Omega_T)}^2 + \frac{1}{2} \| \partial_t r^\tau \|_{L^2(\Omega_T)}^2 
\leq c(L) \| y^\tau - y_\sigma \|_{L^4(\Omega_T)}^4 + c(\sigma) \| r^\tau \|_{L^2(\Omega_T)}^2 + \frac{1}{2} \| \partial_t r^\tau \|_{L^2(\Omega_T)}^2 
\leq c(T, L, c_\sigma(\sigma), \varepsilon, u_0, \sigma) \tau^4 + \frac{1}{2} \| \partial_t r^\tau \|_{L^2(\Omega_T)}^2, \quad (3.16)
$$

where $c(T, L, c_\sigma(\sigma), \varepsilon, u_0, \sigma) > 0$. Inserting (3.16) into (3.15) we obtain

$$
\| r^\tau \|_{H^1(0,T;L^2(\Omega)) \cap L^\infty(0,T;H^1(\Omega))} \leq c(T, L, c_\sigma(\sigma), \varepsilon, u_0, \sigma) \tau^2. \quad (3.17)
$$

Continuing testing (3.9) by $-\Delta r^\tau$ and arguing like in the last steps we end up with

$$
\| r^\tau \|_{L^2(0,T;H^2(\Omega))} \leq c(T, L, c_\sigma(\sigma), \varepsilon, u_0, \sigma) \tau^2
$$

and because of (3.14) and (3.17) finally with (3.8).

### 3.3 First-order necessary optimality conditions

In the previous section, our analysis required minimizers or global solutions of the penalized problems. However, finding globally optimal solutions (in particular by means of numerical algorithms) is difficult in practice. Often, one rather has to rely on stationary points, i.e. points satisfying first-order optimality conditions, or on local solutions. Moreover, our proof steps motivate numerical algorithms [19].

Now we are in the position to state the first-order necessary optimality conditions for $(CP)_\sigma$. Defining

$$
\lambda_\sigma^\oplus := (\psi_\sigma^\oplus)'(y_\sigma)p_{\sigma}, \quad \lambda_\sigma^\ominus := (\psi_\sigma^\ominus)''(y_\sigma)p_{\sigma},
$$

we have:

**Theorem 5.** Assume $(H)$ hold, $\sigma \in (0, 1/4)$ and $\dim \Omega \leq 3$. Then there exist functions $(y_\sigma, u_\sigma, p_\sigma) \in V \times L^2(\Omega_T) \times W(0,T)$ such that the following
first-order optimality system holds

\[ \varepsilon \partial_t y_{\sigma} - \varepsilon \Delta y_{\sigma} + \frac{1}{\varepsilon} \psi'_{0}(y_{\sigma}) + \frac{1}{\varepsilon} \mu_{\sigma}^{\oplus} - \frac{1}{\varepsilon} \mu_{\sigma}^{\ominus} = u_{\sigma} \quad \text{in } \Omega_T, \]

\[ y_{\sigma}(0) = y_0 \quad \text{in } \Omega, \quad n \cdot \nabla y_{\sigma} = 0 \quad \text{on } \Gamma_T, \]

\[ \frac{\nu}{\varepsilon} u_{\sigma} - p_{\sigma} = 0 \quad \text{in } \Omega_T, \]

\[ -\varepsilon \partial_t p_{\sigma} - \varepsilon \Delta p_{\sigma} + \frac{1}{\varepsilon} \psi''_{0}(y_{\sigma}) p_{\sigma} + \frac{1}{\varepsilon} \lambda_{\sigma}^{\oplus} + \frac{1}{\varepsilon} \lambda_{\sigma}^{\ominus} = \nu_d(y_{\sigma} - y_d) \quad \text{in } \Omega_T, \]

\[ p_{\sigma}(T, \cdot) = \nu_T(y_{\sigma}(T, \cdot) - y_T) \quad \text{in } \Omega, \quad n \cdot \nabla p_{\sigma} = 0 \quad \text{on } \Gamma_T. \]

**Proof.** Let \((u_{\sigma}, y_{\sigma})\) be an optimal solution of \((CP)_{\sigma}\). From Theorem 4 we know that \(S_{\sigma}\) is Gâteaux-differentiable from \(L^2(\Omega_T)\) to \(V\). Therefore

\[
\frac{d}{d\tau} J(S_{\sigma}(u_{\sigma} + \tau h_{\sigma}), u_{\sigma} + \tau h_{\sigma})_{|\tau=0} = \\
\nu_T \int_{\Omega} (y_{\sigma}(T, \cdot) - y_T) y_{\sigma}^{*}(T, \cdot) dx + \nu_d \int_{\Omega_T} (y_{\sigma} - y_d) y_{\sigma}^{*} dxdt + \frac{\nu}{\varepsilon} \int_{\Omega_T} u_{\sigma} h_{\sigma} dxdt, \\
(3.23)
\]

where \(y_{\sigma}^{*} = S_{\sigma}'(u_{\sigma}) h_{\sigma}\) is the weak solution of the linearized problem (3.6)-(3.7) in \(y_{\sigma} := S_{\sigma}(u_{\sigma})\), see Theorem 4. We transform (3.23) into another form by introducing the formally adjoint system to (3.6)-(3.7). The adjoint variable \(p_{\sigma}\) is the solution of the following adjoint problem:

\[ -\varepsilon \partial_t p_{\sigma} - \varepsilon \Delta p_{\sigma} + \frac{1}{\varepsilon} \psi''_{0}(y_{\sigma}) p_{\sigma} = \nu_d(y_{\sigma} - y_d) \quad \text{in } \Omega_T, \]

\[ n \cdot \nabla p_{\sigma} = 0 \quad \text{on } \Gamma_T, \]

\[ p_{\sigma}(T, \cdot) = \nu_T(y_{\sigma}(T, \cdot) - y_T) \quad \text{in } \Omega. \]

To prove existence of solutions to (3.24)-(3.26), we introduce the transformation \(\tau := T - t\) and \(p_{\sigma}(t) := \tilde{p}_{\sigma}(\tau)\). Hence, we get the following system

\[ \varepsilon \partial_{\tau} \tilde{p}_{\sigma} - \varepsilon \Delta \tilde{p}_{\sigma} + \frac{1}{\varepsilon} \psi''_{0}(y_{\sigma}) \tilde{p}_{\sigma} = \nu_d(y_{\sigma} - y_d) \quad \text{in } \Omega_T \]

\[ n \cdot \nabla \tilde{p}_{\sigma} = 0, \quad \text{on } \Gamma_T \]

\[ \tilde{p}_{\sigma}(0, \cdot) = \nu_T(y_{\sigma}(T, \cdot) - y_T) \quad \text{in } \Omega. \]

Since the function \(y_T \in L^2(\Omega)\), we have \(\tilde{p}_{\sigma}(0, \cdot) \in L^2(\Omega)\) and the existence of a solution \(\tilde{p}_{\sigma} \in W(0, T)\) and \(p_{\sigma} \in W(0, T)\) is a classical result, (see [12], Chapter 7). To prove (3.20) we test (3.21) by \(y_{\sigma}^{*}\), which is the solution of the linearized problem (3.6)-(3.7) in \(y_{\sigma} := S_{\sigma}(u_{\sigma})\). Integration by parts gives

\[ \int_{\Omega_T} p_{\sigma} h_{\sigma} dxdt = \frac{\nu}{\varepsilon} \int_{\Omega_T} u_{\sigma} h_{\sigma} dxdt. \]
4 Optimality conditions for the limit problem

In this section, based on stationary points for the penalized problems established in (3.18)-(3.22) we investigate the behaviour of accumulation points of sequences of such stationary points. Defining

\[ \lambda_{\sigma} := \lambda_{\sigma}^\ominus + \lambda_{\sigma}^\oplus, \]

we have:

**Lemma 6.** Assume \((H)\) and \(\text{dim } \Omega \leq 3\). Let \(\{u_{\sigma}\}\) be bounded in \(L^2(\Omega_T)\) uniformly in \(\sigma \in (0, 1/4)\) and \((y_{\sigma}, u_{\sigma}, p_{\sigma}) \in V \times L^2(\Omega_T) \times W(0, T)\) be a solution of the optimality system (3.18)-(3.22). Then the following estimates hold

1.) \(y_{\sigma}\) uniformly bounded in \(V\),
2.) \(\mu_{\sigma}^\ominus\) uniformly bounded in \(L^2(\Omega_T)\),
3.) \(\mu_{\sigma}^\oplus\) uniformly bounded in \(L^2(\Omega_T)\),
4.) \(p_{\sigma}\) uniformly bounded in \(L^2(0, T; H^1(\Omega)) \cap L^\infty(0, T; L^2(\Omega))\),
5.) \(\partial_t p_{\sigma}\) uniformly bounded in \(W_0(0, T)^*\),
6.) \(\lambda_{\sigma}\) uniformly bounded in \(W_0(0, T)^*\).

**Proof.** 1.), 2.) and 3.) are direct consequences of Lemma 2.

Now we prove 4.). We introduce the transformation \(\tau := T - t\) and \(p_{\sigma}(t) := e^{\alpha \tau} \tilde{p}_{\sigma}(\tau)\). Hence, we get the following system

\[
\varepsilon \partial_\tau \tilde{p}_{\sigma} - \varepsilon \Delta \tilde{p}_{\sigma} + \frac{1}{\varepsilon} [\psi''_{\sigma}(y_{\sigma}) + \alpha \varepsilon^2] \tilde{p}_{\sigma} = \nu_d e^{-\alpha \tau} (y_{\sigma} - y_d) \quad \text{in } \Omega_T, \quad (4.1)
\]
\[
n \cdot \nabla \tilde{p}_{\sigma} = 0 \quad \text{on } \Gamma_T, \quad (4.2)
\]
\[
\tilde{p}_{\sigma}(0, \cdot) = \nu_T (y_{\sigma}(T, \cdot) - y_T) \quad \text{in } \Omega. \quad (4.3)
\]

Now testing (4.1) by \(\tilde{p}_{\sigma}\) and choosing \(\alpha > 0\) such that \(\psi''_{\sigma}(y_{\sigma}) + \alpha \varepsilon^2 \geq C_0 > 0\) we get by standard calculations the existence of a constant \(C(\tau) > 0\), independent of \(\sigma\), such that

\[
\frac{\varepsilon}{2} \frac{d}{dt} \|\tilde{p}_{\sigma}\|^2_{L^2(\Omega)} + \varepsilon \|\nabla \tilde{p}_{\sigma}\|^2_{L^2(\Omega)} \leq C(\tau) \|\tilde{p}_{\sigma}\|^2_{L^2(\Omega)}.
\]

Now by a Gronwall argument we get \(\|\tilde{p}_{\sigma}\|_{L^2(0, T; H^1(\Omega))} \leq C(\tau)\), hence

\(\|p_{\sigma}\|_{L^2(0, T; H^1(\Omega)) \cap L^\infty(0, T; L^2(\Omega))} \leq C\).

To prove 5.) let \(v \in W_0(0, T) := \{v \in W(0, T) \mid v(0, \cdot) = 0\}\). Using integration by parts we obtain

\[
\langle \partial_t p_{\sigma}, v \rangle = -\langle \partial_t v, p_{\sigma} \rangle + \nu_T (y_{\sigma}(T) - y_T, v(T))_{L^2(\Omega)}.
\]
The continuous injection of $W(0, T)$ into $C([0, T]; L^2(\Omega))$ yields

$$|\langle \partial_t p_\sigma, v \rangle| \leq (\|p_\sigma\|_{L^2(0, T; H^1(\Omega))} + \nu_T \|y_\sigma(T) - y_T\|_{L^2(\Omega)}) \|v\|_{W_0(0, T)}.$$  

Hence from 1.) and 4.) we deduce 5.).

The boundedness of $\lambda_\sigma$ in $W_0(0, T)^*$ follows from the adjoint equation (3.21) and 5.).

Now we can state the main result of this section.

**Theorem 6.** Let the assumption of Lemma 6 hold and let $(y_\sigma, u_\sigma, p_\sigma) \in \mathcal{V} \times L^2(\Omega_T) \times W(0, T)$ be a solution of the optimality system (3.18)-(3.22). Then there exist

$$\begin{pmatrix} y^* \\ \mu_\sigma^\oplus \\ \mu_\sigma^\ominus \\ u^* \\ p^* \\ \lambda^* \end{pmatrix} \in \begin{pmatrix} \mathcal{V} \\ L^2(\Omega_T) \\ L^2(\Omega_T) \\ L^2(\Omega_T) \\ L^2(0, T; H^1(\Omega)) \cap L^\infty(0, T; L^2(\Omega)) \\ W_0(0, T)^* \end{pmatrix}$$

and a subsequence still denoted by $\{y_\sigma, u_\sigma, p_\sigma\}$ such that $\sigma \downarrow 0$

$$\begin{array}{ll} y_\sigma & \longrightarrow y^* \quad \text{weakly} \quad \text{in} \quad L^2(0, T; H^2(\Omega)) \cap H^1(\Omega_T), \\
y_\sigma & \longrightarrow y^* \quad \text{weakly-star} \quad \text{in} \quad L^\infty(0, T; H^1(\Omega)), \\
\mu_\sigma & \longrightarrow \mu_\sigma^\ominus \quad \text{weakly} \quad \text{in} \quad L^2(\Omega_T), \\
\mu_\sigma & \longrightarrow \mu_\sigma^\ominus \quad \text{weakly} \quad \text{in} \quad L^2(\Omega_T), \\
u_\sigma & \longrightarrow u^* \quad \text{weakly} \quad \text{in} \quad L^2(\Omega_T), \\
p_\sigma & \longrightarrow p^* \quad \text{weakly} \quad \text{in} \quad L^2(0, T; H^1(\Omega)), \\
p_\sigma & \longrightarrow p^* \quad \text{weakly-star} \quad \text{in} \quad L^\infty(0, T; L^2(\Omega)), \\
\lambda_\sigma & \longrightarrow \lambda^* \quad \text{weakly} \quad \text{in} \quad W_0(0, T)^*. \\
\end{array}$$

The limit element $\{y^*, \mu_\sigma^\oplus, \mu_\sigma^\ominus, u^*, p^*, \lambda^*\}$ satisfies the following optimality system

$$\begin{align*}
\frac{1}{\varepsilon} \langle \lambda^*, v \rangle_{W_0(0, T)^*, W_0(0, T)} + \langle p^*, \partial_t v \rangle_{L^2(0, T; H^1(\Omega)), L^2(0, T; H^1(\Omega)^*)} + \\
\varepsilon \langle \nabla p^*, \nabla v \rangle_{L^2(\Omega_T)} + \nu_d(y^* - y_d, v)_{L^2(\Omega_T)} - \frac{1}{\varepsilon} \langle p^*, v \rangle_{L^2(\Omega_T)} = \\
\nu_T(y_T - y^*(T, \cdot), v(T, \cdot))_{L^2(\Omega)} \quad \forall v \in W_0(0, T), \\
\frac{1}{\varepsilon} \langle u^* - p^* \rangle = 0 \quad \text{a.e. in} \quad \Omega_T, \\
\varepsilon \partial_t y^* - \varepsilon \Delta y^* - \frac{1}{\varepsilon} y^* + \frac{1}{\varepsilon} \mu_\sigma^\oplus - \frac{1}{\varepsilon} \mu_\sigma^\ominus = u^* \quad \text{a.e. in} \quad \Omega_T, \\
y^*(0) = y_0 \quad \text{in} \quad \Omega, \quad n \cdot \nabla y^* = 0 \quad \text{a.e. on} \quad \Gamma_T,
\end{align*}$$

(4.4) (4.5) (4.6) (4.7)
with the complementarity conditions

\[ |y^*| < 1 \quad \text{a.e. in } \Omega_T, \quad (4.8) \]
\[ \mu^\sigma + (y^* - 1) = 0, \mu^\sigma - (y^* + 1) = 0 \quad \text{a.e. in } \Omega_T, \quad (4.9) \]
\[ \mu^\sigma_* \geq 0, \mu^\sigma_- \geq 0 \quad \text{a.e. in } \Omega_T, \quad (4.10) \]
\[ \lim_{\sigma \searrow 0} (\lambda_\sigma, \max(-1, \min(y_\sigma, 1)))_{L^2(\Omega_T)} = 0, \quad (4.11) \]
\[ \lim \inf_{\sigma \searrow 0} (\lambda_\sigma, p_\sigma)_{L^2(\Omega_T)} \geq 0, \quad (4.12) \]

where \( W_0(0, T) = \{ v \in W(0, T) : v(0, \cdot) = 0 \} \). Furthermore, for every \( \omega > 0 \), there exists a subset \( Q_\omega \subset \{(t, x) \in \Omega_T : |y^*(t, x)| < 1 \} \) with \( \text{meas} \{(t, x) \in \Omega_T : |y^*(t, x)| < 1 \} \setminus Q_\omega \leq \omega \), such that as \( \sigma \searrow 0 \)
\[ \lambda_\sigma \rightarrow 0 \quad \text{uniformly in } Q_\omega. \quad (4.13) \]

Proof. The convergence results are direct consequences of the estimates given by Lemma 6. We multiply the adjoint equation (3.21) by \( v \in W_0(0, T) \) and use integration by parts. Passing to the limit \( \sigma \searrow 0 \), then yields the weak formulation of the adjoint equation as given in (4.4). To show (4.6)-(4.10) we proceed like in the proof of Lemma 2. Because of
\[ (\psi^\sigma + \psi^\sigma_\sigma y^*(y_\sigma) \cdot \max(-1, \min(y_\sigma, 1)) = 0 \]
we easily obtain (4.11). Furthermore we have
\[ (\lambda_\sigma, p_\sigma)_{L^2(\Omega_T)} = \int_{\Omega_T} (\psi^\sigma_\sigma + \psi^\sigma_\sigma y^*(y_\sigma)p_\sigma |y_\sigma|^2 dx dt \geq 0 \]
for \( \sigma \in (0, 1/4) \). Hence, we obtain (4.12).

By Lemma 6 we know that there exists a subsequence (denoted the same) such that \( y_\sigma \rightarrow y^* \) a.e. in \( \Omega_T \). Hence for almost every \( \{(t, x) \in \Omega_T : |y^*(t, x)| < 1 \} \) we have that \( |y_\sigma(t, x)| < 1 \) for \( \sigma \) sufficiently small. Therefore, as \( \sigma \searrow 0 \)
\[ \lambda_\sigma \rightarrow 0 \quad \text{a.e. in } \{(t, x) \in \Omega_T : |y^*(t, x)| < 1 \}. \]

Due to Egorov’s theorem, the quantity \( \lambda_\sigma |\{(t, x) \in \Omega_T : |y^*(t, x)| < 1 \} \) then converges uniformly with respect to the underlying measure to zero, i.e., for every \( \omega > 0 \), there exists a subset \( Q_\omega \subset \{(t, x) \in \Omega_T : |y^*(t, x)| < 1 \} \) with \( \text{meas} \{(t, x) \in \Omega_T : |y^*(t, x)| < 1 \} \setminus Q_\omega \leq \omega \), such that as \( \sigma \searrow 0 \)
\[ \lambda_\sigma \rightarrow 0 \quad \text{uniformly in } Q_\omega. \]
Hence, (4.13) is proven. \( \square \)
Remark 2. (a) Our convergence results are based on the fact that $u_\sigma$ stays inside some uniformly bounded set as $\sigma \downarrow 0$. The optimality conditions are hence derived for accumulation points of stationarity point of the penalized subproblems, only.

(b) We have to remark that the optimality system stated in Theorem 6 does not contain a complementary slackness condition of the form

$$\lim_{\sigma \downarrow 0} (p_\sigma, \mu^\oplus_\sigma - \mu^\ominus_\sigma)_{L^2(\Omega_T)} = 0,$$

which would be a part of C-stationary conditions, (see [23] for different definitions of stationarity). It is not possible to prove this issue, because it is not possible to show that $(\text{meas}\{y_\sigma + 1 \leq 0\}) \to 0$ and $(\text{meas}\{y_\sigma - 1 \geq 0\}) \to 0$ as $\sigma \downarrow 0$. In this case we would be able to use an idea from [26].

(c) The optimality conditions (4.4)-(4.13) of Theorem 6 are weaker than C-stationarity conditions. The result of Theorem 6 can be interpreted in the following way: The accumulation points of stationary points of the penalized subproblems satisfy weaker C-stationarity conditions.

(d) The weakness of the result is due to the low regularity of $\lambda^\oplus + \lambda^\ominus$. Speaking laxy the reason therefore is that the "parabolic operator $\partial_t - \Delta$ is not coercive in $W(0,T)$". In fact, if $\lambda^\oplus + \lambda^\ominus$ is bounded in $L^2(0,T;H^1(\Omega)^*)$, then the result can be strengthened as the next subsection shows.

4.1 Optimality conditions for the limit problem under additional regularity assumptions

In this subsection our aim is to obtain estimates for $\lambda^\oplus_\sigma$ and $\lambda^\ominus_\sigma$ individually, which however require better regularity estimates of $y_\sigma$. To achieve this, for the rest of the paper we make use of the following assumptions:

(A1) $\Omega$ has smooth boundary and $y_0 \in H^2(\Omega)$ with $|y_0| \leq 1$ a.e. in $\Omega$,

(A2) Let $\{u_\sigma\}$ be bounded in $H^1(0,T;L^2(\Omega))$ uniformly in $\sigma \in (0,1/4)$ and $u_0 \in L^2(\Omega)$.

Remark 3. Assume (H) and (A1)-(A2) hold, $\sigma \in (0,1/4)$ then classical results about the regularity of parabolic equations give that $y_\sigma \in \hat{V} := V \cap H^1(0,T;H^1(\Omega)) \cap W^{1,\infty}(0,T;L^2(\Omega))$, (see [12], Chapter 7, Theorem 5 and Theorem 7).
Lemma 7. Assume (H) and (A1)-(A2) hold, \( \sigma \in (0,1/4) \) and \( \dim \Omega \leq 3 \). Let \((y_\sigma,u_\sigma,p_\sigma) \in \mathcal{V} \times L^2(\Omega_T) \times W(0,T)\) be a solution of the optimality system (3.18)-(3.22). Then the following estimates hold

1.) \( y_\sigma \) uniformly bounded in \( \mathcal{V} \),
2.) \( y_\sigma \) uniformly bounded in \( H^1(0,T;H^1(\Omega)) \cap W^{1,\infty}(0,T;L^2(\Omega)) \),
3.) \( \mu_\sigma^\oplus \) uniformly bounded in \( L^2(\Omega_T) \),
4.) \( \mu_\sigma^\ominus \) uniformly bounded in \( L^2(\Omega_T) \),
5.) \( p_\sigma \) uniformly bounded in \( W(0,T) \),
6.) \( \lambda_\sigma \) uniformly bounded in \( L^2(0,T;H^1(\Omega)^*) \),
7.) \( \lambda_\sigma^\ominus \) uniformly bounded in \( L^2(0,T;H^1(\Omega)^*) \),
8.) \( \lambda_\sigma^\oplus \) uniformly bounded in \( L^2(0,T;H^1(\Omega)^*) \).

Proof. 1.), 3.) and 4.) are direct consequences of Lemma 6. To prove 2.) we formally differentiate (3.18) with respect to time and obtain

\[
\varepsilon \partial_t y_\sigma - \varepsilon \Delta (\partial_t y_\sigma) + \frac{1}{\varepsilon \sigma} (\beta_\sigma^\oplus + \beta_\sigma^\ominus)'(y_\sigma) \partial_t y_\sigma = \frac{1}{\varepsilon} \partial_t y_\sigma + \partial_t u_\sigma \quad \text{in } \Omega_T, \quad (4.14)
\]

\[
y_\sigma(0) = y_0 \quad \text{in } \Omega, \quad n \cdot \nabla (\partial_t y_\sigma) = 0 \quad \text{on } \Gamma_T. \quad (4.15)
\]

Now formally testing (4.14) by \( \partial_t y_\sigma \) and noting that \((\beta_\sigma^\oplus + \beta_\sigma^\ominus)'(y_\sigma) \geq 0\) it follows

\[
\frac{\varepsilon}{2} \frac{d}{dt} \| \partial_t y_\sigma \|_{L^2(\Omega)}^2 + \varepsilon \| \nabla (\partial_t y_\sigma) \|_{L^2(\Omega)}^2 \leq C(\varepsilon) (\| \partial_t y_\sigma \|_{L^2(\Omega)}^2 + \| \partial_t u_\sigma \|_{L^2(\Omega)}^2),
\]

(4.16)

where \( C(\varepsilon) \) is a positive constant depending on \( \varepsilon \) but not on \( \sigma \). Integrating with respect to \( t \), using (A2) and 1.) we get

\[
\frac{\varepsilon}{2} \| \partial_t y_\sigma(t) \|_{L^2(\Omega)}^2 + \varepsilon \| \nabla (\partial_t y_\sigma) \|_{L^2(\Omega)}^2 \leq C(\varepsilon) \| \partial_t y_0 \|_{L^2(\Omega)}^2. \quad (4.17)
\]

Using (3.18)-(3.19) and noting that \( \beta_\sigma^\ominus(y_0) = \beta_\sigma^\oplus(y_0) = 0 \) we can estimate the right hand side of (4.17) by

\[
\| \partial_t y_0 \|_{L^2(\Omega)}^2 \leq C(\| \Delta y_0 \|_{L^2(\Omega)}^2 + \frac{1}{\varepsilon^2} \| y_0 \|_{L^2(\Omega)}^2 + \| u_0 \|_{L^2(\Omega)}^2).
\]

(4.18)

Inserting (4.18) into (4.17) and using (A1) we get 2.) We have to remark here that the previous calculations can be done rigorously be using standard Galerkin technique, see e.g. [12].

By (A2), (3.20) and with 4.) of Lemma 6 we immediately obtain 5.). The boundedness of \( \lambda_\sigma = \lambda_\sigma^\oplus + \lambda_\sigma^\ominus \) in \( L^2(0,T;H^1(\Omega)^*) \) follows from the adjoint
To prove (4.19) we define a smooth cutoff function \( \Phi^\oplus \in C^\infty(\mathbb{R}) \), \( 0 \leq \Phi^\oplus(r) \leq 1 \), \( r \in \mathbb{R} \), \( \Phi^\oplus \equiv 1 \) on \( \{ r \geq 1 \} \), \( \Phi^\oplus \equiv 0 \) on \( \{ r \leq 0 \} \) and \( |(\Phi^\oplus)'| \leq 2 \) and get for a \( v \in L^2(0,T;H^1(\Omega)) \)

\[
\|\Phi^\oplus(y_\sigma)v\|_{L^2(0,T;H^1(\Omega))} \leq C\|v\|_{L^2(0,T;H^1(\Omega))}.
\] (4.19)

We want to prove (4.19). We get

\[
\|\nabla(\Phi^\oplus(y_\sigma)v)\|_{L^2(\Omega_T)} \leq \|(\Phi^\oplus)'(y_\sigma)\nabla y_\sigma v\|_{L^2(\Omega_T)} + \|\Phi^\oplus(y_\sigma)\nabla v\|_{L^2(\Omega_T)}.
\]

For the first summand on the right hand side of the above inequality we have using the Hölder inequality

\[
\|(\Phi^\oplus)'(y_\sigma)\nabla y_\sigma v\|_{L^2(\Omega_T)} \leq C\|\nabla y_\sigma\|_{L^\infty(0,T;L^3(\Omega))}\|v\|_{L^2(0,T;L^6(\Omega))},
\]

where \( C \) is a positive constant independent of \( \sigma \). By (1.4), (1.5), (1.6) and 2.) we get

\[
\|(\Phi^\oplus)'(y_\sigma)\nabla y_\sigma v\|_{L^2(\Omega_T)} \leq C\|v\|_{L^2(0,T;H^1(\Omega))}
\]

and in conclusion, the assertion (4.19) is proved. To complete the proof of 8.) we have for a \( v \in L^2(0,T;H^1(\Omega)) \)

\[
|\langle \lambda_\sigma^\oplus, v \rangle_{L^2(0,T;H^1(\Omega)^*),L^2(0,T;H^1(\Omega))}| = |\langle \lambda_\sigma^\oplus, \Phi^\oplus(y_\sigma)v \rangle_{L^2(0,T;H^1(\Omega)^*),L^2(0,T;H^1(\Omega))}|
\]

\[
= |\langle \lambda_\sigma, \Phi^\oplus(y_\sigma)v \rangle_{L^2(0,T;H^1(\Omega)^*),L^2(0,T;H^1(\Omega))}|
\]

\[
\leq C\|v\|_{L^2(0,T;H^1(\Omega))},
\]

where we used 6.) and (4.19) for the last inequality. Hence, we get

\[
\|\lambda_\sigma^\oplus\|_{L^2(0,T;H^1(\Omega)^*)} \leq C.
\]

Analogously by

\[
\|\Phi^\oplus(y_\sigma)v\|_{L^2(0,T;H^1(\Omega))} \leq C\|v\|_{L^2(0,T;H^1(\Omega))},
\]

where we have \( \Phi^\oplus \in C^\infty(\mathbb{R}) \), \( -1 \leq \Phi^\oplus(r) \leq 0 \), \( r \in \mathbb{R} \), \( \Phi^\oplus \equiv -1 \) on \( \{ r \leq -1 \} \), \( \Phi^\oplus \equiv 0 \) on \( \{ r \geq 0 \} \) and \( |(\Phi^\oplus)'| \leq 2 \), we get \( \|\lambda_\sigma^\ominus\|_{L^2(0,T;H^1(\Omega)^*)} \leq C \). □

Defining the functions

\[
[y_\sigma + 1]^\oplus := \max(y_\sigma + 1, 0), \quad [y_\sigma - 1]^\ominus := \min(y_\sigma - 1, 0),
\]

we have:
Corollary 1. Let the assumptions of Lemma 7 be satisfied and let \((y_\sigma, u_\sigma, p_\sigma) \in \mathcal{V} \times H^1(0, T; L^2(\Omega)) \times W(0, T)\) be a solution of the optimality system (3.18)-(3.22). Then there exist

\[
\begin{pmatrix}
y^* \\
\mu_{\sigma}^\oplus \\
\mu_{\sigma}^\ominus \\
u^* \\
p^* \\
\lambda_{\sigma}^\ominus \\
\lambda_{\sigma}^\oplus 
\end{pmatrix} \in \begin{pmatrix}
\mathcal{V} \\
L^2(\Omega_T) \\
L^2(\Omega_T) \\
H^1(0, T; L^2(\Omega)) \\
W(0, T) \\
L^2(0, T; H^1(\Omega)^*) \\
L^2(0, T; H^1(\Omega)^*) 
\end{pmatrix}
\]

and a subsequence still denoted by \(\{y_\sigma, u_\sigma, p_\sigma\}\) such that as \(\sigma \downarrow 0\)

\[
y_\sigma \rightharpoonup y^* \quad \text{weakly in} \quad L^2(0, T; H^2(\Omega)),
y_\sigma \rightharpoonup y^* \quad \text{weakly in} \quad H^1(0, T; H^1(\Omega)),
y_\sigma \rightharpoonup y^* \quad \text{weakly-star in} \quad W^{1,\infty}(0, T; L^2(\Omega)),
\mu_{\sigma}^\oplus \rightharpoonup \mu_{\sigma}^\oplus \quad \text{weakly in} \quad L^2(\Omega_T),
\mu_{\sigma}^\ominus \rightharpoonup \mu_{\sigma}^\ominus \quad \text{weakly in} \quad L^2(\Omega_T),
u_{\sigma} \rightharpoonup u^* \quad \text{weakly in} \quad H^1(0, T; L^2(\Omega)),
p_{\sigma} \rightharpoonup p^* \quad \text{weakly in} \quad W(0, T),
\lambda_{\sigma}^\ominus \rightharpoonup \lambda_{\sigma}^\ominus \quad \text{weakly in} \quad L^2(0, T; H^1(\Omega)^*),
\lambda_{\sigma}^\oplus \rightharpoonup \lambda_{\sigma}^\oplus \quad \text{weakly in} \quad L^2(0, T; H^1(\Omega)^*). \tag{4.20}
\]

The limit element \(\{y^*, \mu_{\sigma}^\oplus, \mu_{\sigma}^\ominus, u^*, p^*, \lambda_{\sigma}^\ominus, \lambda_{\sigma}^\oplus\}\) satisfies the following optimality system

\[
\left\langle \frac{1}{\varepsilon}(\lambda_{\sigma}^\ominus + \lambda_{\sigma}^\oplus) - \varepsilon \partial_t p^* - p^* - \frac{1}{\varepsilon} (\nabla p^*, v)_{L^2(\Omega_T); L^2(0, T; H^1(\Omega))} \right\rangle_{L^2(0, T; H^1(\Omega)^*), L^2(0, T; H^1(\Omega))}
+ \varepsilon(\nabla p^*, \nabla v)_{L^2(\Omega_T)} - \frac{1}{\varepsilon} (p^*, v)_{L^2(\Omega_T)} = 
= \nu_d(y_\sigma - y^*, v)_{L^2(\Omega_T)} \quad \forall v \in L^2(0, T; H^1(\Omega)), \tag{4.21}
\]

\[
p^*(T, \cdot) = \frac{\nu_T}{\varepsilon}(y^*(T, \cdot) - y_T) \quad \text{a.e. in} \quad \Omega, \tag{4.22}
\]

\[
\frac{\nu_{\sigma}}{\varepsilon} u_{\sigma} - p^* = 0 \quad \text{a.e. in} \quad \Omega_T, \tag{4.23}
\]

\[
\varepsilon \partial_t y^* - \varepsilon \Delta y^* - \frac{1}{\varepsilon} y^* - \frac{1}{\varepsilon} \mu_{\sigma}^\oplus - \frac{1}{\varepsilon} \mu_{\sigma}^\ominus = u^* \quad \text{a.e. in} \quad \Omega_T, \tag{4.24}
\]

\[
y^*(0) = y_0 \quad \text{in} \quad \Omega, \quad n \cdot \nabla y^* = 0 \quad \text{a.e. on} \quad \Gamma_T. \tag{4.25}
\]
with the complementarity conditions
\[
|y^*| < 1 \quad \text{a.e. in } \Omega_T,
\]
\[
\mu^\ominus(y^* - 1) = 0, \mu^\ominus(y^* + 1) = 0 \quad \text{a.e. in } \Omega_T,
\]
\[
\mu^\ominus \geq 0, \mu^\ominus \geq 0 \quad \text{a.e. in } \Omega_T,
\]
\[
\langle \lambda^\ominus, [y^* + 1] \rangle_{L^2(0,T;H^1(\Omega)^*)},L^2(0,T;H^1(\Omega)) = 0,
\]
\[
\langle \lambda^\ominus, [y^* - 1] \rangle_{L^2(0,T;H^1(\Omega)^*)},L^2(0,T;H^1(\Omega)) = 0.
\]
Furthermore, for every \( \omega > 0 \), there exists a subset \( Q_\omega \subset \{(t,x) \in \Omega_T : |y^*(t,x)| < 1\} \) such that
\[
\liminf_{\sigma \searrow 0} \sigma \uparrow \lambda^\ominus + \lambda^\ominus \rightarrow 0 \quad \text{uniformly in } Q_\omega.
\]
Proof. The convergence results are direct consequences of the estimates given by Lemma 7. Because of
\[
(\psi^{\ominus}_\sigma)'(y_\sigma)[y_\sigma + 1]^\ominus = 0 \quad \text{and} \quad (\psi^{\ominus}_\sigma)'(y_\sigma)[y_\sigma - 1]^\ominus = 0
\]
we easily obtain
\[
\lim_{\sigma \searrow 0} \langle \lambda^{\ominus}_\sigma, [y_\sigma + 1]^\ominus \rangle_{L^2(\Omega_T)} = 0 \quad \text{and} \quad \lim_{\sigma \searrow 0} \langle \lambda^{\ominus}_\sigma, [y_\sigma - 1]^\ominus \rangle_{L^2(\Omega_T)} = 0.
\]
Furthermore for \( \dim \Omega \leq 3 \) the space \( H^1(0,T;H^1(\Omega)) \cap L^2(0,T;H^2(\Omega)) \) compactly embeds into \( L^2(0,T;H^1(\Omega)) \), which gives strong convergence of \( \{y_\sigma\} \) in \( L^2(0,T;L^1(\Omega)) \). Together 7.), 8.) of Lemma 7 and (4.34) imply (4.29) and (4.30). Moreover we have
\[
(\lambda^l_\sigma,p_\sigma)_{L^2(\Omega_T)} = \int_{\Omega_T} (\psi^l_\sigma(y_\sigma))^\prime(y_\sigma)|p_\sigma|^2 dx dt \geq 0
\]
for \( l \in \{\ominus, \ominus\} \) and \( \sigma \in (0,1/4) \). Hence, we obtain (4.31)-(4.32).

Remark 4. (a) We have to remark that also the optimality system stated in Corollary 1 does not contain a complementary slackness condition of the form
\[
\lim_{\sigma \searrow 0} (p_\sigma, \mu^{\ominus}_\sigma - \mu^{\ominus}_\sigma)_{L^2(\Omega_T)} = 0,
\]
which would be a part of C-stationary conditions, (see [23] for different definitions of stationarity).
Acknowledgment. The author thanks the anonymous referee for a very careful reading of the manuscript. Moreover the author wishes to thank Luise Blank and Harald Garcke for many fruitful hints and discussions. This work was supported by the SPP 1253 “Optimization with Partial Differential Equations” of the German Science Foundation (DFG) through the grant GA 695/5-2.

References


REFERENCES


REFERENCES


