Numer. Math. (2004) 98: 135–166
Digital Object Identifier (DOI) 10.1007/s00211-004-0531-z

Numerische Mathematik

# A level set approach for the solution of a state-constrained optimal control problem

**Michael Hintermüller**[1], **Wolfgang Ring**[2]

[1] Department of Computational and Applied Mathematics, Rice University, 6100 Main Street, Houston, TX 77005, USA
[2] Special Research Center on Optimization and Control, University of Graz, Institute of Mathematics, Heinrichstrasse 36, 8010 Graz, Austria

Received February 26, 2001 / Revised version received March 19, 2003 /
Published online April 8, 2004 – 

**Summary.** State constrained optimal control problems for linear elliptic partial differential equations are considered. The corresponding first order optimality conditions in primal-dual form are analyzed and linked to a free boundary problem resulting in a novel algorithmic approach with the boundary (interface) between the active and inactive sets as optimization variable. The new algorithm is based on the level set methodology. The speed function involved in the level set equation for propagating the interface is computed by utilizing techniques from shape optimization. Encouraging numerical results attained by the new algorithm are reported on.

*Mathematics Subject Classification (1991):* 35R35, 49K20, 49Q10, 65K10

## 1 Introduction

This paper is devoted to the numerical solution of state constrained optimal control problems of the type

$$\begin{aligned} &\text{minimize} && \tfrac{1}{2}\|y-y_d\|^2_{L^2(\Omega)} + \tfrac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \\ &\text{subject to} && -\Delta y = u \quad \text{in } \Omega, \quad y = 0 \quad \text{on } \partial\Omega, \\ &&& y \in K, \end{aligned} \tag{1.1}$$

where $\alpha > 0$, $\Omega \subset \mathbb{R}^n$ and $\partial\Omega$ its sufficiently smooth boundary. The state $y$ is constrained by the requirement $y \in K$, with

*Correspondence to*: M. Hintermüller

$$K = \{v \in H_0^1(\Omega) |\ v \leq \psi \text{ a.e. on } \Omega\} \subset H_0^1(\Omega)\,,$$

where $\psi$ is sufficiently regular.

Problems of type (1.1) frequently arise in practical applications either in their own right or as sub-problems in sequential quadratic programming approaches for the numerical solution of general nonlinear optimal control problems (See e.g. [16], [9]). The problem of imposing constraints on the state has received considerable attention. The contributions in [7], [8], and [2] are concerned with theoretical aspects of deriving first and second order conditions characterizing optimal solutions. In [5], [6], [3], [17], and [22] numerical solution algorithms are introduced and analyzed. However, the development of efficient numerical schemes for (1.1) is far from being complete. Uzawa-type algorithms with or without block relaxation are considered in [3]. Since they are frequently slow in their practical performance, techniques based on an augmented Lagrangian approach have been introduced in [5]. These techniques typically outperform the Uzawa-based methods but the structure of the constraints is not fully exploited. The recently developed primal-dual active set strategy [6] is promising in the sense that only a low number of iterations is required for finding the optimal solution to the discretized counterpart of (1.1). However, in contrast to the Uzawa-type and augmented Lagrangian-based methods no infinite dimensional analysis is possible. Also a remarkable sensitivity with respect to discretization parameters can be observed and the discretization of the Lagrange multiplier poses difficulties.

Another distinct difference between the above mentioned numerical approaches is the way of how the optimality system is taken into account. While in [3] the primal variables, i.e. state and control variable, are somewhat emphasized, in [5] a dual variable, i.e. a Lagrange multiplier, for the state constraint is introduced. The augmented Lagrangian technique employed in this paper establishes a link between the primal and dual variables. Finally, the primal-dual active set strategy of [6] keeps the primal and the dual variables separate as it is done in primal-dual path following interior point techniques in finite dimensions (see e.g. [24],[25]). The numerical comparison in [4], [6] for control and state constrained optimal control problems gives strong evidence that primal-dual techniques are superior to either primal or dual approaches. In this present paper we suggest yet another idea for a numerical treatment of (1.1). Based on the primal-dual formulation of the first order optimality conditions we derive an equivalent characterization of the optimal solution to (1.1) as the solution to a free boundary problem.

The main intention of this paper is to introduce an efficient numerical algorithm that captures the specific features due to the type of constraints together with the PDE-type state equation of the underlying problem while keeping the primal-dual aspect mentioned above. Therefore, the basis for this research will be a first order characterization of the optimal solution of (1.1) involving primal and dual variables like in [6]. Given an initial guess of the optimal solution, a closer inspection of the system guides us to an iterative procedure based on a free boundary problem for finding the optimal solution to (1.1) numerically. We shall stress that this free boundary aspect is novel in the sense that the free boundary $\Gamma$ replaces $(y, u)$ in the role of the optimization variables. The new approach is well suited to constraints of the type considered here and is not included in the aforementioned papers. The numerical treatment of the free boundary problem is based on an adaptation of level set methods [20] to the present situation. The favor for using a level set based scheme comes from the fact that level set methods are numerically efficient and robust procedures for the tracking of interfaces which allow topology changes in the course of the iteration. In our case, the speed vector field which drives the propagation of the level set function is given by the Eulerian derivative of an appropriately defined cost functional with respect to the free boundary. To calculate the Eulerian derivative, shape sensitivity analysis using adjoint variables in the spirit of [21] is employed.

We shall note that the subsequently presented techniques can be applied to problems with more general smooth cost functionals and any second order elliptic differential operator instead of $-\Delta$. Moreover, the treatment of additional constraints on the control variable $u$ poses no difficulty. In order to make the subsequent ideas more apparent we consider (1.1) as a model problem where we omit control constraints right from the beginning.

In the following section 2 we start by establishing the (primal-dual) first order sufficient and necessary optimality conditions for the optimal solution of (1.1). Moreover, we give a thorough discussion of the boundary conditions satisfied by the optimal state $y^*$ and optimal control $u^*$. In section 3 we discuss the level set approach and introduce a basic algorithm for finding a solution to (1.1). The speed function necessary for the level set based algorithm is discussed in section 4 together with relevant issues concerning sensitivity. Section 5 is devoted to a brief description of the discrete algorithm and its implementation. A report on selected test examples emphasizes the feasibility and efficiency of our novel approach and is the content of section 6.

## 2 First Order Optimality Conditions

Let $\Omega \subset \mathbb{R}^n$ $(n = 2, 3)$ be a bounded, piecewise smooth domain and let $\psi \in H^4(\Omega)$ be given with $0 < \psi(x) \leq M$ on $\partial\Omega$ for some $M > 0$. Moreover, assume that $y_d \in H^2(\Omega)$. We consider the state-constrained optimal control problem

(2.1a) $$\min J(y, u) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)}$$

(2.1b) $$\text{subject to } \begin{cases} \Delta y + u = 0 \text{ on } \Omega \\ y \le \psi \text{ a.e. on } \Omega \end{cases}$$
$$\text{over } (y, u) \in H_0^1(\Omega) \times L^2(\Omega).$$

In [6] the following optimality system for (2.1) is given: The pair $(y^*, u^*) \in H_0^1(\Omega) \times L^2(\Omega)$ is the unique solution to (2.1) if and only if there exists a Lagrange multiplier $\lambda^* \in \mathcal{M}(\Omega)$ (the space of regular Borel-measures on $\Omega$) such that

(2.2a) $$\Delta y^* + u^* = 0 \text{ on } \Omega,$$

(2.2b) $$y^* \le \psi \text{ on } \Omega$$

(2.2c) $$-\alpha(u^*, \Delta y)_\Omega + \langle \lambda^*, y\rangle_{\mathcal{M},\mathcal{C}_0} = (y_d - y^*, y)_\Omega \quad \text{for all } y \in H_0^1(\Omega) \cap H^2(\Omega)$$

(2.2d) $$\langle \lambda^*, z - y^*\rangle_{\mathcal{M},\mathcal{C}_0} \le 0 \text{ for all } z \in \mathcal{C}_0(\Omega) \text{ with } z \le \psi.$$

Here $(\cdot,\cdot)_\Omega$ denotes the inner product in $L^2(\Omega)$, and $\langle\cdot,\cdot\rangle_{\mathcal{M},\mathcal{C}_0}$ denotes the duality pairing between $\mathcal{C}_0(\Omega)$ and its dual $\mathcal{M}(\Omega)$. (See Rudin [18, p. 70, Def 3.16] for the definition of $\mathcal{C}_0(\Omega)$).

We define the active and inactive sets with respect to the solution $(y^*, u^*)$ by

$$\mathcal{A}^* = \{\mathbf{x} \in \Omega : y^*(\mathbf{x}) = \psi(\mathbf{x})\}, \qquad \mathcal{I}^* = \Omega \setminus \mathcal{A}^*.$$

Elliptic regularity implies that $y^* \in H_0^1(\Omega) \cap H^2(\Omega)$ and hence due to Sobolev's lemma $y^* \in \mathcal{C}(\bar{\Omega})$. From the definition and from the continuity of $y^* - \psi$ it follows that $\mathcal{A}^*$ is closed in $\Omega$. Therefore $\mathcal{I}^*$ is an open subset of $\mathbb{R}^n$. We set

$$\Sigma = \partial\Omega \text{ and } \Gamma^* = \partial\mathcal{A}^*.$$

See Fig. 1 for a sketch of the configuration at the optimum.

Throughout the paper we invoke the following assumptions on the regularity of the geometric situation at the optimum $(y^*, u^*)$.

(A1) $\Gamma^* = \partial\mathcal{A}^* = \partial\big(\text{int}(\mathcal{A}^*)\big) = \partial\mathcal{I}^* \setminus \Sigma \subset \Omega$.

(A2) $\text{int}(\mathcal{A}^*) \ne \emptyset$.

(A3) Both $\mathcal{I}^*$ and $\text{int}(\mathcal{A}^*)$ are smooth enough to allow the existence of first order (Neumann) forward and inverse trace operators and the applicability of Green's formula.

Note that the assumptions $0 < \psi$ and $y^* = 0$ on $\Sigma$ imply that $\Sigma \cap \Gamma^* = \emptyset$.

We formulated Assumption (A3) without concrete specification of the required smoothness for the active and inactive sets. Obviously $\Gamma^* \in \mathcal{C}^{1,1}$ would be sufficient. Using geometric measure theory, we could relax this

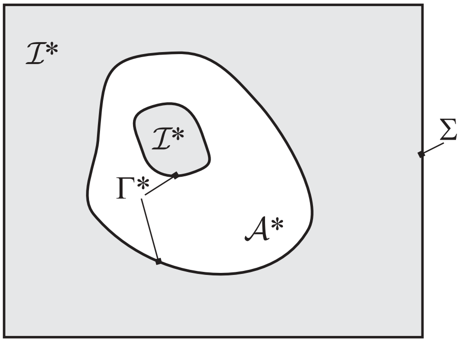


**Fig. 1.** Sketch of the geometric situation at the solution to (2.1)

condition even further. However, as the assumption shall hold for the active set with respect to the unknown optimal solution $y^*$, it cannot be checked anyway, unless, we have some kind of regularity theory for a generic (non-degenerate) solution of the state constrained optimal control problem (1.1). We therefore gave a condition which yields that the procedures of taking traces and using Green's formula are feasible, since this is what we actually need in the subsequent considerations. For the somewhat similar situation of variational inequalities, the smoothness of the contact set is a well investigated subject and several smoothness results for non-degenerate situations are known (see [13, Thm. 3.4, p.163; Sec. V.7, p178; etc.]). It would be interesting to investigate the smoothness for the active set for the solution (1.1) in the spirit of Kinderlehrer and Stampacchia [13].

The aim of the following considerations is to find an equivalent formulation for the optimality system (2.2). The main problem with (2.2) is, that the multiplier $\lambda^* \in \mathcal{M}(\Omega)$ is not a function but a measure on $\Omega$. This implies that any discretization of $\lambda^*$ (especially of the singular part of $\lambda^*$) on some grid defined on $\Omega$ is problematic. We shall replace $\lambda^*$ by certain traces of functions on the boundary $\Gamma^*$ of the active set.

We start with considering the boundary conditions which must hold for $y^*$ and $u^*$ on $\mathcal{I}^*$. By definition we have $y^* \in H_0^1(\Omega)$ and hence

$$y^* = 0 \text{ on } \Sigma. \tag{2.3}$$

Since $y^* \in \mathcal{C}_0(\Omega)$, $y^* = \psi$ on $\mathcal{A}^*$, and $\Gamma^* = \partial\mathcal{A}^*$, we obtain

$$y^*|_{\Gamma^*} = \psi|_{\Gamma^*}. \tag{2.4}$$

Let us now consider $u^*$. It is found in [6] that the measure $\lambda^*$ is concentrated on $\mathcal{A}^*$. Let $\varphi \in \mathcal{D}(\mathcal{I}^*)$ be arbitrary. Here and in the following, $\mathcal{D}(\mathcal{I}^*)$ denotes the space of $\mathcal{C}^\infty$-functions with compact support in $\mathcal{I}^*$. Since $\varphi \in H_0^1(\Omega) \cap H^2(\Omega)$, we can use $\varphi$ as test function in (2.2c). We find

$$-\alpha(u^*, \Delta\varphi)_{\mathcal{I}^*} = (y_d - y^*, \varphi)_{\mathcal{I}^*}.$$

Thus,

$$-\alpha\Delta u^* = y_d - y^* \in H^2(\mathcal{I}^*) \tag{2.5}$$

in the sense of distributions.

We now define the space

$$\mathcal{W}(\mathcal{I}^*) = \{w \in L^2(\mathcal{I}^*) \,:\, \Delta w \in L^2(\mathcal{I}^*)\}, \tag{2.6}$$

with the norm $\|w\|^2_{\mathcal{W}(\mathcal{I}^*)} = \|w\|^2_{L^2(\mathcal{I}^*)} + \|\Delta w\|^2_{L^2(\mathcal{I}^*)}$. Due to (2.5) we have $u^* \in \mathcal{W}(\mathcal{I}^*)$.

**Proposition 1** *There exist continuous (Dirichlet and Neumann) trace operators*

$$\gamma_0 : \mathcal{W}(\mathcal{I}^*) \to H^{-\frac{1}{2}}(\Gamma^*)$$

*and*

$$\gamma_1 : \mathcal{W}(\mathcal{I}^*) \to H^{-\frac{3}{2}}(\Gamma^*)$$

*with* $\gamma_0 u = u|_{\Gamma^*}$ *and* $\gamma_1 u = \frac{\partial u}{\partial n}|_{\Gamma^*}$ *for* $u \in \mathcal{C}^\infty(\overline{\mathcal{I}^*})$*, where* $n$ *denotes the unit exterior normal vector to* $\mathcal{I}^*$*. Moreover, Green's formula*

$$\begin{aligned}(\Delta v, w)_{\mathcal{I}^*} &- (\Delta w, v)_{\mathcal{I}^*} \\ &= \left\langle \gamma_0 w, \frac{\partial v}{\partial n}\right\rangle_{H^{-\frac{1}{2}}(\Gamma^*), H^{\frac{1}{2}}(\Gamma^*)} - \left\langle \gamma_1 w, v\right\rangle_{H^{-\frac{3}{2}}(\Gamma^*), H^{\frac{3}{2}}(\Gamma^*)}\end{aligned} \tag{2.7}$$

*holds for all* $w \in \mathcal{W}(\mathcal{I}^*)$ *and all* $v \in H^2(\mathcal{I}^*)$.

*Proof.* See Lions and Magenes [15, Théorème 6.5, page 187 and p. 229]. □

Now let $f \in H^{\frac{1}{2}}(\Sigma)$ be arbitrarily given. Let $y \in H^2(\mathcal{I}^*)$ be chosen such that $y|_\Sigma = y|_{\Gamma^*} = \frac{\partial y}{\partial n}|_{\Gamma^*} = 0$ and $\frac{\partial y}{\partial n}|_\Sigma = f$ (see [23, p. 133, Thm. 8.8] for the existence of such a function). Due to the homogenous boundary conditions on $\Gamma^*$, $y$ can be extended by 0 to a function (which we also denote by $y$) in $H^1_0(\Omega) \cap H^2(\Omega)$. Hence, $y$ is an admissible test function in (2.2c). We find

$$-\alpha(u^*, \Delta y)_{\mathcal{I}^*} = (y_d - y^*, y)_{\mathcal{I}^*}.$$

On the other hand, (2.5) implies

$$-\alpha(\Delta u^*, y)_{\mathcal{I}^*} = (y_d - y^*, y)_{\mathcal{I}^*}.$$

Combination of the last two expressions with (2.7) and the definition of $y$ gives

$$\langle u^*, f\rangle_{H^{-\frac{1}{2}}(\Sigma), H^{\frac{1}{2}}(\Sigma)} = 0$$

for all $f \in H^{\frac{1}{2}}(\Sigma)$. Thus, we conclude that

$$u^*|_\Sigma = 0. \tag{2.8}$$

At this point it is convenient to consider some structural properties of the multiplier $\lambda^*$. Using Lebesgue's decomposition theorem and the fact that $\lambda^*|_{\mathcal{I}^*} = 0$ we can write $\lambda^* = \lambda_s^* + \lambda_a^*$ where $\lambda_s^*$ is concentrated on $\Gamma^*$ and $\lambda_a^*$ is concentrated on $\mathrm{int}(\mathcal{A}^*)$. Let $\varphi \in \mathcal{D}(\mathrm{int}(\mathcal{A}^*))$. Using (2.2c) and $y^* = \psi$ on $\mathcal{A}^*$ we get

$$\begin{aligned}\int_{\mathrm{int}(\mathcal{A}^*)} \varphi \, d\lambda_a^* &= (y_d - \psi, \varphi)_{\mathrm{int}(\mathcal{A}^*)} - \alpha(\Delta\psi, \Delta\varphi)_{\mathrm{int}(\mathcal{A}^*)} \\ &= (y_d - \psi - \alpha\Delta^2\psi, \varphi)_{\mathrm{int}(\mathcal{A}^*)}.\end{aligned}$$

Therefore, we find

$$\lambda_a^* = y_d - \psi - \alpha\Delta^2\psi \in L^2(\mathrm{int}(\mathcal{A}^*)). \tag{2.9}$$

Let $g \in H^{\frac{1}{2}}(\Gamma^*)$ be given. Using appropriate Neumann extension operators $Z_1 : H^{\frac{1}{2}}(\Gamma^*) \to H^2(\mathcal{I}^*)$ and $Z_2 : H^{\frac{1}{2}}(\Gamma^*) \to H^2(\mathrm{int}(\mathcal{A}^*))$ we can construct a function $y_g \in H_0^1(\Omega) \cap H^2(\Omega)$ which satisfies $y_g|_{\Gamma^*} = 0$ and $\frac{\partial y}{\partial n}\big|_{\Gamma^*} = g$. Recall that $n$ is the unit *exterior* normal vector field to $\mathcal{I}^*$. With $y_g$ as a test function in (2.2c) and the fact that $y_g|_{\Gamma^*} = 0$, we obtain

$$\begin{aligned}(u^*, \Delta y_g)_\Omega &= \frac{1}{\alpha}\left(\int_{\Gamma^*} y_g \, d\lambda_s^* + \int_{\mathrm{int}(\mathcal{A}^*)} y_g \, d\lambda_a^*\right) - \frac{1}{\alpha}(y_d - y^*, y_g)_\Omega \\ &= \frac{1}{\alpha}(y_d - \psi - \alpha\Delta^2\psi, y_g)_{\mathcal{A}^*} - \frac{1}{\alpha}(y_d - y^*, y_g)_\Omega \\ &= \frac{1}{\alpha}(y^* - y_d, y_g)_{\mathcal{I}^*} - (\Delta^2\psi, y_g)_{\mathcal{A}^*}.\end{aligned}$$

On the other hand, with (2.7), (2.5) and Green's formula we find

$$\begin{aligned}(u^*, \Delta y_g)_\Omega &= (u^*, \Delta y_g)_{\mathcal{I}^*} + (u^*, \Delta y_g)_{\mathcal{A}^*} \\ &= (\Delta u^*, y_g)_{\mathcal{I}^*} + (\Delta u^*, y_g)_{\mathcal{A}^*} + \langle u^* + \Delta\psi, g\rangle_{H^{-\frac{1}{2}}(\Gamma^*), H^{\frac{1}{2}}(\Gamma^*)} \\ &= \frac{1}{\alpha}(y^* - y_d, y_g)_{\mathcal{I}^*} - (\Delta^2\psi, y_g)_{\mathcal{A}^*} \\ &\quad + \langle u^* + \Delta\psi, g\rangle_{H^{-\frac{1}{2}}(\Gamma^*), H^{\frac{1}{2}}(\Gamma^*)}.\end{aligned}$$

Subtracting the last two expressions gives

$$\langle u^* + \Delta\psi, g\rangle_{H^{-\frac{1}{2}}(\Gamma^*), H^{\frac{1}{2}}(\Gamma^*)} = 0$$

for arbitrary $g \in H^{\frac{1}{2}}(\Gamma^*)$. Thus, we obtain

$$u^*|_{\Gamma^*} = -\Delta\psi|_{\Gamma^*} \in H^{\frac{3}{2}}(\Gamma^*). \tag{2.10}$$

Now, we come back to $y^*$. Let $h \in H^{\frac{3}{2}}(\Gamma^*)$ be arbitrary and suppose $y_h \in H^2(\Omega) \cap H_0^1(\Omega)$ is constructed analogously to $y_g$ such that $y_h|_{\Gamma^*} = h$ and $\frac{\partial y_h}{\partial n}\big|_{\Gamma^*} = 0$. Using Green's formula together with the homogeneous Dirichlet boundary conditions satisfied by $y^*$ and $y_h$ on $\Omega$, we obtain

$$(\Delta y^*, y_h)_\Omega = (y^*, \Delta y_h)_\Omega.$$

On the other hand, by Green's formula on $\mathcal{I}^*$ and $\mathcal{A}^*$ with $n_{\mathcal{I}^*}$ and $n_{\mathcal{A}^*}$ denoting the unit exterior normals on $\mathcal{I}^*$ and $\mathcal{A}^*$ respectively, and the Dirichlet boundary conditions satisfied by $y^*$ and $y_h$ on $\Gamma^*$, we find

$$\begin{aligned}(\Delta y^*, y_h)_\Omega &= (\Delta y^*, y_h)_{\mathcal{I}^*} + (\Delta y^*, y_h)_{\mathcal{A}^*}\\ &= -(\nabla y^*, \nabla y_h)_{\mathcal{I}^*} + \int_{\Gamma^*} \frac{\partial y^*}{\partial n_{\mathcal{I}^*}}\, y_h\, d\Gamma^*\\ &\quad -(\nabla y^*, \nabla y_h)_{\mathcal{A}^*} + \int_{\Gamma^*} \frac{\partial y^*}{\partial n_{\mathcal{A}^*}}\, y_h\, d\Gamma^*\\ &= (y^*, \Delta y_h)_\Omega + \int_{\Gamma^*} \left( \frac{\partial y^*}{\partial n_{\mathcal{I}^*}} + \frac{\partial y^*}{\partial n_{\mathcal{A}^*}} \right) h\, d\Gamma^*.\end{aligned}$$

Since $h$ is arbitrary in $H^{\frac{3}{2}}(\Gamma^*)$, we conclude

$$\frac{\partial y^*}{\partial n_{\mathcal{I}^*}} = -\frac{\partial y^*}{\partial n_{\mathcal{A}^*}}.$$

Because $\frac{\partial y^*}{\partial n_{\mathcal{A}^*}} = \frac{\partial \psi}{\partial n_{\mathcal{A}^*}} = -\frac{\partial \psi}{\partial n_{\mathcal{I}^*}}$, we finally obtain

$$\frac{\partial y^*}{\partial n}\bigg|_{\Gamma^*} = \frac{\partial \psi}{\partial n}\bigg|_{\Gamma^*} \in H^{\frac{5}{2}}(\Gamma^*). \tag{2.11}$$

We can also use $y_h$ as a test function in (2.2c). Doing so we obtain

$$(u^*, \Delta y_h)_\Omega = \frac{1}{\alpha} \int_{\Gamma^*} h\, d\lambda_s^* + \frac{1}{\alpha}(y^* - y_d, y_h)_{\mathcal{I}^*} - (\Delta^2 \psi, y_h)_{\mathcal{A}^*}.$$

On the other hand, using the standard Green's formula for functions in $H^2(\mathrm{int}(\mathcal{A}^*))$, and the Green's formula (2.7), we get

$$\begin{aligned}(u^*, \Delta y_h)_\Omega &= (u^*, \Delta y_h)_{\mathcal{I}^*} + (u^*, \Delta y_h)_{\mathcal{A}^*}\\ &= (\Delta u^*, y_h)_{\mathcal{I}^*} - \left\langle \frac{\partial u^*}{\partial n_{\mathcal{I}^*}}, h \right\rangle_{H^{-\frac{3}{2}}(\Gamma^*), H^{\frac{3}{2}}(\Gamma^*)}\\ &\quad +(\Delta u^*, y_h)_{\mathcal{A}^*} + \int_{\Gamma^*} \frac{\partial}{\partial n_{\mathcal{A}^*}} (\Delta \psi)\, h\, d\Gamma^*\\ &= \frac{1}{\alpha}(y^* - y_d, y_h)_{\mathcal{I}^*} - (\Delta^2 \psi, y_h)_{\mathcal{A}^*}\\ &\quad -\left\langle \frac{\partial u^*}{\partial n_{\mathcal{I}^*}} + \frac{\partial}{\partial n_{\mathcal{I}^*}} (\Delta \psi), h \right\rangle_{H^{-\frac{3}{2}}(\Gamma^*), H^{\frac{3}{2}}(\Gamma^*)}\end{aligned}$$

for all $h \in H^{\frac{3}{2}}(\Gamma^*)$. Combining the last two formulas gives

$$\lambda_s^*|_{\Gamma^*} = -\alpha \frac{\partial}{\partial n_{\mathcal{I}^*}}(u^* + \Delta\psi)\big|_{\Gamma^*} \in H^{-\frac{3}{2}}(\Gamma^*) \tag{2.12}$$

or, more explicitly

$$\int_{\Gamma^*} y\, d\lambda_s^* = -\alpha \Big\langle \frac{\partial}{\partial n_{\mathcal{I}^*}}(u^* + \Delta\psi), y \Big\rangle_{H^{-\frac{3}{2}}(\Gamma^*), H^{\frac{3}{2}}(\Gamma^*)} \tag{2.13}$$

for all $y \in H^2(\Omega)$.

**Proposition 2** (Necessary conditions) *Suppose* $(y^*, u^*) \in H_0^1(\Omega) \times L^2(\Omega)$ *is the solution to* (2.1) *and the active and inactive sets* $\mathcal{I}^*$ *and* $\mathcal{A}^*$ *satisfy the regularity assumptions* (A1), (A2), *and* (A3). *Then* $u^* \in \mathcal{W}(\mathcal{I}^*)$, $y^* \in H^2(\Omega)$ *and we have*

$$\Delta y^* + u^* = 0 \text{ on } \mathcal{I}^* \tag{2.14a}$$

$$y^* - \alpha\Delta u^* = y_d \text{ on } \mathcal{I}^* \tag{2.14b}$$

$$y^* < \psi \text{ on } \mathcal{I}^* \tag{2.14c}$$

$$y^*|_\Sigma = 0, \quad u^*|_\Sigma = 0, \tag{2.15a}$$

$$y^*|_{\Gamma^*} = \psi|_{\Gamma^*}, \quad u^*|_{\Gamma^*} = -\Delta\psi|_{\Gamma^*}, \tag{2.15b}$$

$$\frac{\partial y^*}{\partial n}\Big|_{\Gamma^*} = \frac{\partial \psi}{\partial n}|_{\Gamma^*} \tag{2.15c}$$

$$-\frac{\partial}{\partial n}(u^* + \Delta\psi)|_{\Gamma^*} \geq 0 \text{ as a measure on } \Gamma^* \tag{2.15d}$$

$$y_d - \psi - \alpha\Delta^2\psi \geq 0 \text{ almost everywhere on } \mathcal{A}^*. \tag{2.15e}$$

**(Sufficient conditions).** *Suppose conversely that an open set* $\hat{\mathcal{I}} \subset \Omega$ *is found such that* $\Sigma \subset \partial\hat{\mathcal{I}}$. *We set* $\hat{\Gamma} = \partial\hat{\mathcal{I}} \setminus \Sigma$ *and we assume that the smoothness assumptions* (A1), (A2), *and* (A3) *are satisfied for* $\hat{\mathcal{I}}$ *and* $\hat{\mathcal{A}} = \Omega \setminus \hat{\mathcal{I}}$. *Suppose moreover that* $(\hat{y}, \hat{u}) \in H^2(\hat{\mathcal{I}}) \times \mathcal{W}(\hat{\mathcal{I}})$ *satisfy* (2.14) *and* (2.15) *on* $\hat{\mathcal{I}}$ *and* $\hat{\mathcal{A}}$ *respectively (i.e. all* $*$*-expressions are replaced by the corresponding -expressions). Then* $(\bar{y}, \bar{u})$ *defined by*

$$\bar{y} = \begin{cases} \hat{y} \text{ on } \hat{\mathcal{I}} \\ \psi \text{ on } \hat{\mathcal{A}} \end{cases} \quad \text{and } \bar{u} = \begin{cases} \hat{u} \text{ on } \hat{\mathcal{I}} \\ -\Delta\psi \text{ on } \hat{\mathcal{A}} \end{cases} \tag{2.16}$$

*is the unique solution to* (2.1), *i.e.* $\bar{y} = y^*$ *and* $\bar{u} = u^*$.

*Proof.* We have already proved that (2.14) and (2.15a)–(2.15c) must hold for the optimal configuration $(y^*, u^*, \mathcal{I}^*, \mathcal{A}^*)$. It follows from (2.2d) that $\lambda^*$ defines a positive functional on $\mathcal{C}_c(\mathcal{A}^*)$. Hence, by the Riesz representation theorem [18, p. 40. Thm 2.14] $\lambda^*$ is a positive measure on $\mathcal{A}^*$. Thus,

$\lambda_s^* = \lambda^*|_{\Gamma^*}$ and $\lambda_a^* = \lambda^*|_{\text{int}(\mathcal{A}^*)}$ are also positive measures. Application of the characterizations (2.9) and (2.12) proves that (2.15e) and (2.15d) must hold.

Suppose conversely that (2.14) and (2.15) are satisfied by $(\hat{y}, \hat{u})$ on $\hat{\mathcal{I}}$ and $\hat{\mathcal{A}}$, respectively. Let $(\bar{y}, \bar{u})$ be defined by (2.16). We set $\bar{\lambda} = \bar{\lambda}_a + \bar{\lambda}_s \in \mathcal{M}(\Omega)$ where

$$\langle \bar{\lambda}_s, y \rangle_{\mathcal{M},\mathcal{C}_0} = -\Big\langle \frac{\partial}{\partial n}(\bar{u} + \Delta\psi), y \Big\rangle_{H^{-\frac{3}{2}}(\hat{\Gamma}), H^{\frac{3}{2}}(\hat{\Gamma})} \tag{2.17}$$

for all $y \in H^2(\Omega) \cap H_0^1(\Omega)$ and

$$\langle \bar{\lambda}_a, y \rangle_{\mathcal{M},\mathcal{C}_0} = \int_{\hat{\mathcal{A}}} (y_d - \psi - \alpha\Delta^2\psi)\, y \, d\mathbf{x}. \tag{2.18}$$

Since $H^2(\Omega) \cap H_0^1(\Omega)$ is densely embedded in $\mathcal{C}_0(\Omega)$, the definition (2.17) can be extended to a bounded linear functional on $\mathcal{C}_0(\Omega)$.

It remains to prove that $(\bar{y}, \bar{u}, \bar{\lambda})$ defined in this way fulfill the optimality system (2.2). We have $\bar{y} \le \psi$ on $\hat{\mathcal{I}}$ by (2.14c) and $\bar{y} = \psi$ on $\hat{\mathcal{A}}$. Thus, (2.2b) is satisfied. Let $\varphi \in \mathcal{D}(\Omega)$. Then

$$\begin{aligned}
(\bar{y}, \Delta\varphi)_\Omega &= (\bar{y}, \Delta\varphi)_{\hat{\mathcal{I}}} + (\bar{y}, \Delta\varphi)_{\hat{\mathcal{A}}} \\
&= -(\nabla\hat{y}, \nabla\varphi)_{\hat{\mathcal{I}}} + \int_{\hat{\Gamma}} \psi \frac{\partial\varphi}{\partial n_{\hat{\mathcal{I}}}} \, d\hat{\Gamma} - (\nabla\hat{y}, \nabla\varphi)_{\hat{\mathcal{A}}} + \int_{\hat{\Gamma}} \psi \frac{\partial\varphi}{\partial n_{\hat{\mathcal{A}}}} \, d\hat{\Gamma} \\
&= (\Delta\hat{y}, \varphi)_{\hat{\mathcal{I}}} - \int_{\hat{\Gamma}} \frac{\partial\psi}{\partial n_{\hat{\mathcal{I}}}} \varphi \, d\hat{\Gamma} + (\Delta\hat{y}, \varphi)_{\hat{\mathcal{A}}} - \int_{\hat{\Gamma}} \frac{\partial\psi}{\partial n_{\hat{\mathcal{A}}}} \varphi \, d\hat{\Gamma} \\
&= (\bar{u}, \varphi)_\Omega
\end{aligned}$$

holds for all $\varphi \in \mathcal{D}(\Omega)$. Therefore, (2.2a) is fulfilled in the distributional sense for $(\bar{y}, \bar{u})$.

Expressions (2.17) and (2.18) define positive measures which are both concentrated on $\hat{\mathcal{A}}$. Thus, $\bar{\lambda}$ is positive and concentrated on $\hat{\mathcal{A}}$ and we have

$$\langle \bar{\lambda}, z - \bar{y} \rangle_{\mathcal{M},\mathcal{C}_0} = \int_{\hat{\mathcal{A}}} (z - \psi)\, d\bar{\lambda} \le 0$$

for all $z \in \mathcal{C}_0(\Omega)$ with $z \le \psi$. Consequently, (2.2d) holds for $\bar{\lambda}$ and $\bar{y}$.

Let us now consider (2.2c). Let $y \in H^2(\Omega) \cap H_0^1(\Omega)$ be arbitrarily given. We have

$$
\begin{aligned}
&(y_d - \bar{y}, y)_\Omega + \alpha(\bar{u}, \Delta y)_\Omega \\
&\quad = (y_d - \bar{y}, y)_\Omega + \alpha(\bar{u}, \Delta y)_{\hat{\mathcal{I}}} - \alpha(\Delta\psi, \Delta y)_{\hat{\mathcal{A}}} \\
&\quad = (y_d - \bar{y}, y)_\Omega - \alpha(\nabla\bar{u}, \nabla y)_{\hat{\mathcal{I}}} - \alpha \int_{\hat{\Gamma}} \Delta\psi \, \frac{\partial y}{\partial n_{\hat{\mathcal{I}}}} \, d\hat{\Gamma} \\
&\qquad + \alpha(\nabla\Delta\psi, \nabla y)_{\hat{\mathcal{A}}} - \alpha \int_{\hat{\Gamma}} \Delta\psi \, \frac{\partial y}{\partial n_{\hat{\mathcal{A}}}} \, d\hat{\Gamma} \\
&\quad = (y_d - \bar{y}, y)_{\hat{\mathcal{I}}} + (y_d - \psi, y)_{\hat{\mathcal{A}}} + \alpha(\Delta\bar{u}, y)_{\hat{\mathcal{I}}} - \alpha \int_{\hat{\Gamma}} \frac{\partial \bar{u}}{\partial n_{\hat{\mathcal{I}}}} \, y \, d\hat{\Gamma} \\
&\qquad - \alpha(\Delta^2\psi, y)_{\hat{\mathcal{A}}} + \alpha \int_{\hat{\Gamma}} \frac{\partial}{\partial n_{\hat{\mathcal{A}}}} \Delta\psi \, y \, d\hat{\Gamma} \\
&\quad = (y_d - \psi - \alpha\Delta^2\psi, y)_{\hat{\mathcal{A}}} - \alpha \int_{\hat{\Gamma}} \frac{\partial}{\partial n_{\hat{\mathcal{I}}}} (\bar{u} + \Delta\psi) \, y \, d\hat{\Gamma} \\
&\quad = \langle \bar{\lambda}, y \rangle_{\mathcal{M}, \mathcal{C}_0}.
\end{aligned}
$$

Here we used (2.14a), (2.14b), (2.15a), (2.15b), and (2.15c). We conclude that (2.2c) is satisfied for $(\bar{y}, \bar{u}, \bar{\lambda})$ and the proof is complete. □

## 3 Level set approach

Traditional techniques for computing the solution $(y^*, u^*)$ to the system (2.14), (2.15) iteratively approximate the optimal solution by a sequence $\{(y_n, u_n)\}$; see [5], [6], [3], [17], [22]. Every iterate $(y_n, u_n)$ induces a corresponding geometric configuration given by $\mathcal{A}_n, \mathcal{I}_n, \Gamma_n$. Here $\mathcal{A}_n = \{\mathbf{x} \in \Omega : y_n(\mathbf{x}) = \psi(\mathbf{x})\}, \mathcal{I}_n = \Omega \backslash \mathcal{A}_n$ and $\Gamma_n$ denotes the boundary between $\mathcal{A}_n$ and $\mathcal{I}_n$.

In contrast to these techniques, we shall pursue the following idea to solve the optimality system (2.14), (2.15) which we consider as a free boundary value problem. The geometric configuration of the active and inactive sets is a priori not known. An iterative approach must therefore include an update of the geometry in every step. Conceptually we consider the updates of the geometry as discrete snapshots of a continuously moving geometry. Suppose, we have a current geometric configuration $(\mathcal{I}, \mathcal{A})$ where $\mathcal{I}$ and $\mathcal{A}$ are approximations of the inactive and active sets $\mathcal{I}^*$ and $\mathcal{A}^*$, respectively. On $\mathcal{I}$, we can relax some of the boundary conditions (2.15) and solve (2.14a) and (2.14b) with the remaining boundary conditions for $(y, u)$. The violation of the relaxed boundary conditions together with a possible violation of the constraint $y \le \psi$ defines a “distance" of the actual configuration $(\mathcal{I}, y(\mathcal{I}), u(\mathcal{I}))$ to the optimal solution $(\mathcal{I}^*, y^*, u^*)$. If we quantify the violation of the relaxed boundary conditions and the violation of (2.14c) by an appropriate cost functional $K(\Gamma)$, we can use the gradient of the cost functional with respect to the geometry as a speed function for the evolution of the moving geometry. The condition (2.15e) is satisfied by an appropriate choice of the initial configuration $(\mathcal{I}_0, \mathcal{A}_0)$. For a geometric configuration where the relaxed boundary conditions is exactly satisfied and $y$ is feasible, the gradient of the cost functional is 0, hence, a steady state of the evolution problem is attained.

The geometry of the problem is uniquely defined by the boundary $\Gamma$ of the (current) inactive set $\mathcal{I}$. It is therefore sufficient to consider the evolution of $\Gamma$ driven by the gradient $\nabla_\Gamma K(\Gamma)$ of the cost functional with respect to $\Gamma$. It is well established (see [20], [19]) that level set formulations of moving interface problems possess several advantages including flexibility with respect to topology changes, the possibility to use fixed grids, low computational cost, and robustness. We propose a level set formulation for the solution of the moving interface problem in this paper.

Summarizing the introductory discussion of this section we give a sketch of the proposed algorithm. For this purpose let $(\mathrm{O}_r(\Gamma_n))$ denote the system (2.14),(2.15) with inactive set $\mathcal{I}_n$ induced by the actual boundary $\Gamma_n$ and where one of the boundary conditions in (2.15b),(2.15c) is neglected. Moreover $\|\cdot\|$ shall denote an appropriate norm which is specified below.

**Level set based algorithm.**

Step 0. Choose an appropriate initial $\Gamma_0$; set $n = 0$.
Step 1. Compute $(y_n, u_n)$ from the relaxed system $(\mathrm{O}_r(\Gamma_n))$.
Step 2. Evaluate the cost functional $K(\Gamma_n)$ and compute its derivative $\nabla_\Gamma K(\Gamma_n)$. If $\|\nabla_\Gamma K(\Gamma_n)\| = 0$ then stop; otherwise continue with step 3.
Step 3. Use an appropriate extension of $\nabla_\Gamma K(\Gamma_n)$ to $\Omega$ as speed function in the level set equation for updating the level set function.
Step 4. Set $\Gamma_{n+1}$ equal to the zero level set of the updated level set function, and put $n := n + 1$. Go to step 1.

Subsequently we elaborate the details that are necessary for obtaining a well-defined algorithm. We start by proving existence and uniqueness of the solution to an important auxiliary problem. Suppose an open set $\mathcal{I} \subset \Omega$ is given which satisfies $\Sigma \subset \partial\mathcal{I}$, $\Gamma = \partial\mathcal{I} \setminus \Sigma \neq \emptyset$ and for which the smoothness assumptions (A1), (A2), and (A3) hold. We consider the boundary value problem

(3.1a) $$\Delta y + u = 0 \text{ on } \mathcal{I},$$
(3.1b) $$y - \alpha\Delta u = y_d \text{ on } \mathcal{I}$$

(3.2a) $$y|_\Sigma = 0, \quad u|_\Sigma = 0,$$
(3.2b) $$y|_\Gamma = \psi|_\Gamma, \quad u|_\Gamma = -\Delta\psi|_\Gamma$$

**Proposition 3** *Under the assumptions on $\mathcal{I}$ described above, the boundary value problem* (3.1), (3.2) *has a unique solution* $(y, u) \in H^2(\mathcal{I}) \times H^2(\mathcal{I})$.

*Proof.* Using appropriate extensions of the boundary values (3.2a) it is easily seen that it is sufficient to consider solvability of the problem with homogeneous boundary values

$$\Delta y + u = f_1 \text{ on } \mathcal{I}, \tag{3.3a}$$

$$y - \alpha \Delta u = f_2 \text{ on } \mathcal{I} \tag{3.3b}$$

$$y|_{\partial \mathcal{I}} = 0, \quad u|_{\partial \mathcal{I}} = 0 \tag{3.3c}$$

with $f_1, f_2 \in L^2(\mathcal{I})$.

Let $\Delta^{-1} : L^2(\mathcal{I}) \to H^2(\mathcal{I}) \cap H_0^1(\mathcal{I})$ denote the solution operator to the homogeneous Dirichlet problem for the Laplace operator on $\mathcal{I}$. From (3.3b) it follows that $u = \frac{1}{\alpha}\Delta^{-1}(y - f_2)$. Inserting this in (3.3a) we get

$$\Delta y + \frac{1}{\alpha}\Delta^{-1} y = f_1 + \frac{1}{\alpha}\Delta^{-1} f_2 \text{ on } \mathcal{I}. \tag{3.4}$$

The weak formulation of (3.4) reads as: Find $y \in H_0^1(\mathcal{I})$ such that

$$(\nabla y, \nabla \varphi)_{\mathcal{I}} - \frac{1}{\alpha}(\Delta^{-1} y, \varphi)_{\mathcal{I}} = -\left(f_1 + \frac{1}{\alpha}\Delta^{-1} f_2, \varphi\right)_{\mathcal{I}} \tag{3.5}$$

for all $\varphi \in H_0^1(\mathcal{I})$. Let $w = \Delta^{-1} v \in H^2(\mathcal{I}) \cap H_0^1(\mathcal{I})$ for some given $v \in H_0^1(\mathcal{I})$. We have

$$(\Delta^{-1} v, v)_{\mathcal{I}} = (w, \Delta w)_{\mathcal{I}} = -(\nabla w, \nabla w)_{\mathcal{I}} \le 0.$$

Thus, the left hand side of (3.5) defines a uniformly elliptic bilinear form on $H_0^1(\mathcal{I})$. Application of the Lax-Milgram theorem ensures the existence of a unique solution $y \in H_0^1(\mathcal{I})$ to (3.5). Standard regularity theory implies $y \in H^2(\mathcal{I})$. Setting $u = \frac{1}{\alpha}\Delta^{-1}(y - f_2) \in H^2(\mathcal{I}) \cap H_0^1(\mathcal{I})$ completes the proof. □

Let $(y, u) = (y(\Gamma), u(\Gamma))$ denote the solution to (3.1), (3.2). Sometimes we write $(y, u) = \big(y(\mathcal{I}), u(\mathcal{I})\big)$. We define the cost functional

$$\begin{aligned} K(\Gamma) &= K(\Gamma, u(\Gamma), y(\Gamma)) \\ &= \frac{1}{2|\Gamma|}\int_\Gamma \left( \left|\frac{\partial}{\partial n}(y - \psi)\right|^2 + c_1\left( \max\left(0, \frac{\partial}{\partial n}(u + \Delta\psi)\right)\right)^2 \right) d\Gamma \\ &\quad + \frac{c_2}{2}\int_{\mathcal{I}} \big(\max(0, y - \psi)\big)^2 \, d\mathbf{x}. \end{aligned} \tag{3.6}$$

Thus, the boundary condition (2.15c) and the inequality constraints (2.15d) and (2.14c) are implemented in the cost functional, the latter ones as penalty terms with penalty parameters $c_1, c_2 > 0$. The factor $\frac{1}{|\Gamma|}$ is introduced to prevent the cost functional from vanishing if $|\Gamma|$ goes to 0. The constraint (2.15e) is not represented in the cost functional. This constraint depends only on a priori given data and on the geometry. If we define

$$\mathcal{M} = \{\mathbf{x} \in \Omega \mid y_d - \psi - \alpha\Delta^2\psi \geq 0\} \tag{3.7}$$

we can choose some subset of $\mathcal{M}$ as starting value for the active set $\mathcal{A}$. Thus, feasibility with respect to (2.15e) holds for the initial configuration. In all our numerical tests feasibility of (2.15e) is maintained for all updates of the geometry such that a representation of (2.15e) in the cost functional appears to be unnecessary.

Now we briefly recall some theoretical aspects concerning the gradient of a cost functional like (3.6) with respect to the geometry $\Gamma$. Here we rely on the concepts and results given in [21]. Let $V(t, \mathbf{x})$ be a smooth vector field defined on $[0, T] \times \Omega$ with $V(t, \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = 0$ for almost every $\mathbf{x} \in \Sigma$ and all $t \in [0, T]$. If the unit exterior normal vector $\mathbf{n} = \mathbf{n}(\mathbf{x})$ in not defined at a singular $\mathbf{x} \in \Sigma$ we assume that $V(t, \mathbf{x}) = \mathbf{0}$. We shall refer to $V$ as admissible if it satisfies the above mentioned conditions. Let $\mathbf{x} = \mathbf{x}(t, \mathbf{X})$ denote the solution to the initial value problem

$$\begin{aligned} &\frac{d}{dt}\mathbf{x}(t, \mathbf{X}) = V(t, \mathbf{x}(t, \mathbf{X})), \\ &\mathbf{x}(0, \mathbf{X}) = \mathbf{X} \end{aligned} \tag{3.8}$$

with $\mathbf{X} \in \Omega$ and $t \in [0, T]$ and we denote by $T_t : \Omega \to \Omega$ the time-$t$ map with respect to (3.8), i.e. $T_t(\mathbf{X}) = \mathbf{x}(t, \mathbf{X})$. Note that $T_t(\mathbf{X}) \in \Omega$ due to the properties of $V$ on $\Sigma$. We set $\mathcal{I}_t = T_t(\mathcal{I})$ and $\Gamma_t = T_t(\Gamma)$. The Eulerian derivative of $K$ at $\Gamma$ in direction of the vector field $V$ is defined as the limit

$$dK(\Gamma; V) = \lim_{t \downarrow 0} \frac{1}{t}(K(\Gamma_t) - K(\Gamma)). \tag{3.9}$$

It is known (see [21, Thm 2.27,p 59]) that there exists a distribution $F$ on $\Gamma$ such that $dK(\Gamma; V) = \langle F, v_n \rangle_\Gamma$ where $v_n = V(0, \mathbf{x}) \cdot \mathbf{n}(\mathbf{x})$ and $\langle \cdot, \cdot \rangle_\Gamma$ denotes an appropriate duality pairing. The correct functional analytic setting for this formula will be discussed in section 4. If $\langle \cdot, \cdot \rangle_\Gamma$ can be realized as an integral over $\Gamma$ we have

$$dK(\Gamma; V) = \int_\Gamma F(V(0, \cdot) \cdot \mathbf{n})\, d\Gamma = \int_\Gamma (F\mathbf{n}) \cdot V(0, \cdot)\, d\Gamma. \tag{3.10}$$

Thus, the Eulerian derivative is represented by a vector field $F\mathbf{n}$ which is normal to $\Gamma$ with speed function $F = F(\Gamma, \mathbf{x})$ for any $\mathbf{x} \in \Gamma$.

The speed function $F$ can now be used to define a family of propagating interfaces $\Gamma(\tau)$. We assume that a point $\mathbf{x}(\tau) \in \Gamma(\tau)$ propagates along the direction given by the negative gradient of the cost functional $K$, i.e. along the vector field $-F(\Gamma(\tau), \mathbf{x}(\tau))\,\mathbf{n}(\mathbf{x}(\tau))$. That is to say $\mathbf{x}(\tau)$ is solution to the ordinary differential equation

$$\mathbf{x}'(\tau) = -F(\Gamma(\tau), \mathbf{x}(\tau))\;\mathbf{n}(\Gamma(\tau), \mathbf{x}(\tau)) \tag{3.11}$$

with $\mathbf{x}(0) = \mathbf{x}_0 \in \Gamma(0)$ and $\Gamma(\tau)$ is defined as $\Gamma(\tau) = \{\mathbf{x}(\tau) : \mathbf{x}_0 \in \Gamma(0)\}$ for $\tau > 0$. We can consider the propagating interface $\Gamma(\tau)$ as the zero level set of a time dependent function $\Phi : [0, \infty) \times \Omega \to \mathbb{R}$. The evolution equation for the level set function $\Phi$ is then given by the hyperbolic Hamiltonian equation

$$\Phi_\tau - F|\nabla\Phi| = 0 \text{ on } \Omega \tag{3.12}$$

with $\Phi(0, \mathbf{x}) = \Phi_0(\mathbf{x})$ given. (See [20] for the connection between the moving boundary formulation (3.11) and the level set formulation (3.12)). The speed function in (3.10) is defined only on the moving boundary $\Gamma = \Gamma(\tau)$. For (3.12) we need, however, that the speed function is defined on $\Omega$ or, at least, on some band containing $\Gamma$. It is therefore necessary to extend $F$ from $\Gamma$ onto $\Omega$ in some appropriate (smooth) way. The extension procedure is described in detail in [11, 10].

## 4 Shape Sensitivity Analysis

Let $(y, u)$ and $(y_t, u_t)$ denote the solutions to (3.1), (3.2) on $\mathcal{I}$ and $\mathcal{I}_t = T_t(\mathcal{I})$ respectively. We introduce the material derivative of $y_t$ at $t = 0$ as

$$\dot{y}(\mathcal{I}; V) = \lim_{t\downarrow 0} \frac{1}{t}\big(y(\mathcal{I}_t) \circ T_t(V) - y(\mathcal{I}_0)\big)$$

and analogously for $u$. Here we use the notation $\mathcal{I}_0 = \mathcal{I}$. Before we derive an analytic expression for the Eulerian derivative $dK(\Gamma; V)$, we prove existence and regularity results for the material derivatives $\dot{y}$ and $\dot{u}$.

**Theorem 1** *Let $\Gamma$ be of class $\mathcal{C}^{1,1}$ and $V \in \mathcal{C}([0, T); \mathcal{C}^2(\Omega, \mathbb{R}^n))$ be an admissible vector field. Assume moreover that $\psi \in H^5(\Omega)$. Let $(y, u)$ denote the solution to* (3.1), (3.2). *Then, the material derivative $(\dot{y}(\mathcal{I}; V), \dot{u}(\mathcal{I}; V))$ exists in $H_0^1(\mathcal{I}) \times H_0^1(\mathcal{I})$. The convergence*

$$\frac{1}{t}\big(y(\mathcal{I}_t) \circ T_t(V) - y(\mathcal{I}_0)\big) \to \dot{y}(\mathcal{I}; V)$$

*is weak in $H_0^1(\mathcal{I})$. The analogous statement holds for $\dot{u}(\mathcal{I}; V)$.*

*Proof.* See Appendix A. □

Let us remark that the smoothness requirements of Theorem 1 are necessitated by analytical issues. On the numerical level, however, we observe that the method is successful also in cases where the smoothness assumptions on e.g. $\Gamma$ do not hold; see section 6.

It turns out to be convenient to define the *shape derivatives*

$$y'(\mathcal{I};V) = \dot{y}(\mathcal{I};V) - \nabla y(\mathcal{I})\cdot V(0) \in H^1(\mathcal{I}) \tag{4.1}$$

$$u'(\mathcal{I};V) = \dot{u}(\mathcal{I};V) - \nabla u(\mathcal{I})\cdot V(0) \in H^1(\mathcal{I}). \tag{4.2}$$

It follows from techniques presented in [21, p.118–119] that $(y', u')$ is characterized as the unique solution to the boundary value problem

$$\Delta y' + u' = 0 \text{ on } \mathcal{I}, \tag{4.3a}$$

$$y' - \alpha\Delta u' = 0 \text{ on } \mathcal{I}, \tag{4.3b}$$

with boundary conditions

$$y'|_\Sigma = 0, \quad u'|_\Sigma = 0, \tag{4.4a}$$

$$y'|_\Gamma = -\frac{\partial}{\partial n}(y-\psi)v_n|_\Gamma, \quad u'|_\Gamma = -\frac{\partial}{\partial n}(u+\Delta\psi)v_n|_\Gamma \tag{4.4b}$$

where $v_n = V(0)\cdot n$. Note that $(y', u') \in \mathcal{W}(\mathcal{I})\times\mathcal{W}(\mathcal{I})$, hence, the Neumann traces $\frac{\partial y'}{\partial n}|_\Gamma$ and $\frac{\partial u'}{\partial n}|_\Gamma$ exist in $H^{-\frac{1}{2}}(\Gamma)$. We also recall the definition of the shape derivative of a family of functions $\zeta(\Gamma) \in W(\Gamma)$, where $W(\Gamma)$ is some Sobolev space on $\Gamma$. We set

$$\zeta'(\Gamma;V) = \dot{\zeta}(\Gamma;V) - \nabla_\Gamma\zeta(\Gamma)\cdot V(0) \tag{4.5}$$

where $\nabla_\Gamma\zeta$ is the tangential gradient of $\zeta$. Note that, in the particular case where $\zeta(\Gamma) = \eta(\mathcal{I})|_\Gamma$ for some family of functions $\eta(\mathcal{I}) : \mathcal{I} \to \mathbb{R}$, we have

$$\zeta'(\Gamma;V) = \eta'(\mathcal{I};V)|_\Gamma + \frac{\partial\eta(\mathcal{I})}{\partial n}|_\Gamma v_n.$$

(See [21] for details).

Now let us consider the cost functional (3.6), i.e.

$$K(\Gamma) = \frac{1}{2|\Gamma|}\int_\Gamma \left(\left|\frac{\partial}{\partial n}(y-\psi)\right|^2 + c_1\Big(\max\Big(0, \frac{\partial}{\partial n}(u+\Delta\psi)\Big)\Big)^2\right)d\Gamma$$
$$+\frac{c_2}{2}\int_{\mathcal{I}}\big(\max(0, y-\psi)\big)^2\,d\mathbf{x},$$

where $(y, u)$ is the solution to (3.1), (3.2). We set

$$K_1(\Gamma) = \frac{1}{2|\Gamma|}\int_\Gamma \left|\frac{\partial}{\partial n}(y-\psi)\right|^2 d\Gamma \tag{4.6a}$$

$$K_2(\Gamma) = \frac{c_1}{2|\Gamma|}\int_\Gamma \Big(\max\Big(0, \frac{\partial}{\partial n}(u+\Delta\psi)\Big)\Big)^2 d\Gamma \tag{4.6b}$$

$$K_3(\Gamma) = \frac{c_2}{2}\int_{\mathcal{I}}\big(\max(0, y-\psi)\big)^2\,d\mathbf{x}. \tag{4.6c}$$

Next we derive expressions for the Eulerian derivatives of $K_1$, $K_2$, and $K_3$. We set

$$\tilde{K}_1(\Gamma) = \frac{1}{2}\int_\Gamma \left|\frac{\partial}{\partial n}(y-\psi)\right|^2 d\Gamma.$$

Using [21, p.116, (2.173)] and (4.5) we find

$$\begin{aligned} d\tilde{K}_1(\Gamma; V) = &\int_\Gamma \frac{\partial}{\partial n}(y-\psi)\Big(\big(\nabla(y-\psi)\cdot n\big)^{\boldsymbol{\cdot}} \\ &-\nabla_\Gamma\Big(\frac{\partial}{\partial n}(y-\psi)\Big)\cdot V(0)\Big)\, d\Gamma \\ &+\frac{1}{2}\int_\Gamma \kappa\left|\frac{\partial}{\partial n}(y-\psi)\right|^2 v_n\, d\Gamma \end{aligned}$$

Here $\kappa$ denotes the mean curvature of $\Gamma$. From [21, p.125, Lem.3.4] we learn that

$$\dot{n}(\Gamma; V)(\mathbf{x}) = -\big(DV(0,\mathbf{x})^* n\big)_\tau,$$

where the subscript $\tau$ denotes the projection of the vector $\big(DV(0,\mathbf{x})^* n\big)$ onto the tangent space to $\Gamma$ at $\mathbf{x}$. Application of the chain rule together with (4.1) gives

$$\begin{aligned} d\tilde{K}_1(V;\Gamma) = &\int_\Gamma \frac{\partial}{\partial n}(y-\psi)\left(\frac{\partial y'}{\partial n} + \Big(\nabla\frac{\partial}{\partial n}(y-\psi)\right. \\ &-\nabla_\Gamma\frac{\partial}{\partial n}(y-\psi)\Big)\cdot V(0) \\ &\left.-\nabla(y-\psi)\cdot\big(DV(0,\mathbf{x})^* n\big)_\tau\right) d\Gamma \\ &+\frac{1}{2}\int_\Gamma \kappa\left|\frac{\partial}{\partial n}(y-\psi)\right|^2 v_n\, d\Gamma. \end{aligned}$$

For the above formula we also used that $\psi'(\mathcal{I};\Gamma) = 0$ since $\psi$ is independent of $\Gamma$. We have

$$\begin{aligned} &\nabla\frac{\partial}{\partial n}(y-\psi) - \nabla_\Gamma\frac{\partial}{\partial n}(y-\psi) = \frac{\partial^2}{\partial n^2}(y-\psi)\, n \\ &= \Big(\Delta(y-\psi) - \Delta_\Gamma(y-\psi) - \kappa\frac{\partial}{\partial n}(y-\psi)\Big)\, n = -\kappa\frac{\partial}{\partial n}(y-\psi)\, n \end{aligned}$$

because $\Delta(y-\psi) = -u - \Delta\psi = 0$ on $\Gamma$ by (3.2b) and $\Delta_\Gamma(y-\psi) = 0$ due to $y-\psi = 0$ on $\Gamma$. Here $\Delta_\Gamma$ denotes the Laplace-Beltrami operator on $\Gamma$. From $y-\psi = 0$ on $\Gamma$ it follows also that $\nabla_\Gamma(y-\psi) = 0$ on $\Gamma$. Therefore, we have

$$\nabla(y-\psi) = \frac{\partial}{\partial n}(y-\psi)\, n + \nabla_\Gamma(y-\psi) = \frac{\partial}{\partial n}(y-\psi)\, n \tag{4.7}$$

and hence,

$$\nabla(y-\psi)\cdot\big(DV(0,\mathbf{x})^*n\big)_\tau = 0.$$

Thus, we get

$$d\tilde{K}_1(V;\Gamma) = \int_\Gamma \Big(\frac{\partial}{\partial n}(y-\psi)\,\frac{\partial y'}{\partial n} - \frac{1}{2}\kappa\,\Big|\frac{\partial}{\partial n}(y-\psi)\Big|^2 v_n\Big)\,d\Gamma. \tag{4.8}$$

With $m(x)=\frac{1}{2}\big(\max(0,x)\big)^2$ we set

$$\tilde{K}_2(\Gamma) = \int_\Gamma m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\,d\Gamma.$$

As in the above calculations we find

$$\begin{aligned}
&d\tilde{K}_2(V;\Gamma)\\
&\quad= \int_\Gamma m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\Big(\big(\nabla(u+\Delta\psi)\cdot n\big)^{\boldsymbol{\cdot}}\\
&\qquad\quad -\nabla_\Gamma\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\cdot V(0)\Big)\,d\Gamma\\
&\qquad\quad +\int_\Gamma \kappa m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\,v_n\,d\Gamma\\
&\quad= \int_\Gamma m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\Big(\frac{\partial u'}{\partial n} + \frac{\partial^2}{\partial n^2}(u+\Delta\psi)\,v_n\Big)\,d\Gamma\\
&\qquad\quad +\int_\Gamma \kappa m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\,v_n\,d\Gamma\\
&\quad= \int_\Gamma m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\Big(\frac{\partial u'}{\partial n} + \Big(\frac{1}{\alpha}(y-y_d)+\Delta^2\psi\Big)\,v_n\Big)\,d\Gamma\\
&\qquad\quad +\int_\Gamma \kappa\Big[m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\\
&\qquad\quad -m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\,\frac{\partial}{\partial n}(u+\Delta\psi)\Big]\,v_n\,d\Gamma.
\end{aligned}$$

By definition of $m$ we have $m(x)-m'(x)\,x = -m(x)$. Note that here $m'$ denotes the derivative of the function $m:\mathbb{R}\to\mathbb{R}$ and not some kind of shape derivative. Using this we get

$$\begin{aligned}
d\tilde{K}_2(V;\Gamma) &= \int_\Gamma m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\Big(\frac{\partial u'}{\partial n} + \Big(\frac{1}{\alpha}(y-y_d)+\Delta^2\psi\Big)\,v_n\Big)\,d\Gamma\\
&\quad -\int_\Gamma \kappa m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\,v_n\,d\Gamma.
\end{aligned} \tag{4.9}$$

Applying [21, p.113, (2.168)] we find for $K_3$:

$$dK_3(V;\Gamma) = c_2\int_{\mathcal{I}} m'(y-\psi)\,y'\,d\mathbf{x}. \tag{4.10}$$

Finally, we set

$$K_0(\Gamma) = \frac{1}{|\Gamma|}.$$

Following [21, p. 80, prop.2.50 and p.93, (2.145)] we obtain

$$dK_0(V;\Gamma) = -\frac{1}{|\Gamma|^2}\int_\Gamma \kappa\, v_n\, d\Gamma. \tag{4.11}$$

Combining all preceding results we obtain

$$\begin{aligned} dK(V;\Gamma) = {} & \frac{1}{|\Gamma|}\int_\Gamma \left(\frac{\partial}{\partial n}(y-\psi)\,\frac{\partial y'}{\partial n} + c_1\, m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\,\frac{\partial u'}{\partial n}\right) d\Gamma \\ & + c_2\int_{\mathcal{I}} m'(y-\psi)\, y'\, d\mathbf{x} \\ & + \frac{c_1}{|\Gamma|}\int_\Gamma m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\left(\frac{1}{\alpha}(y-y_d)+\Delta^2\psi\right) v_n\, d\Gamma \\ & - \frac{1}{|\Gamma|}\int_\Gamma \kappa\, \left(\frac{1}{2}\Big|\frac{\partial}{\partial n}(y-\psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\right) v_n\, d\Gamma \\ & - \frac{1}{|\Gamma|^2}\int_\Gamma \kappa\, v_n\, d\Gamma \\ & \cdot\left(\int_\Gamma \left(\frac{1}{2}\Big|\frac{\partial}{\partial n}(y-\psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\right) d\Gamma\right). \end{aligned} \tag{4.12}$$

For given $(y, u)$ we define the adjoint boundary value problem by

$$\Delta\mu + \nu = c_2 m'(y-\psi) \text{ on } \mathcal{I}, \tag{4.13a}$$
$$\mu - \alpha\Delta\nu = 0 \text{ on } \mathcal{I} \tag{4.13b}$$

with boundary conditions

$$\mu|_\Sigma = 0, \quad \nu|_\Sigma = 0, \tag{4.14a}$$
$$\mu|_\Gamma = \frac{1}{|\Gamma|}\,\frac{\partial}{\partial n}(y-\psi)\Big|_\Gamma, \quad \nu|_\Gamma = -\frac{c_1}{\alpha|\Gamma|} m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\Big|_\Gamma. \tag{4.14b}$$

Using (4.13), (4.4), (4.3), and (4.14) we find

$$\begin{aligned} & (c_2 m'(y-\psi), y')_{\mathcal{I}} = (\Delta\mu+\nu, y')_{\mathcal{I}} \\ & = -(\nabla\mu, \nabla y')_{\mathcal{I}} - \int_\Gamma \frac{\partial\mu}{\partial n}\,\frac{\partial}{\partial n}(y-\psi)\, v_n\, d\Gamma + (\nu, y')_{\mathcal{I}} \\ & = (\mu, \Delta y')_{\mathcal{I}} - \frac{1}{|\Gamma|}\int_\Gamma \frac{\partial}{\partial n}(y-\psi)\,\frac{\partial y'}{\partial n}\, d\Gamma - \int_\Gamma \frac{\partial\mu}{\partial n}\,\frac{\partial}{\partial n}(y-\psi)\, v_n\, d\Gamma \\ & \quad + (\nu, y')_{\mathcal{I}} + (u', \mu)_{\mathcal{I}} - (u', \mu)_{\mathcal{I}} - (\alpha\Delta u', \nu)_{\mathcal{I}} + (\alpha\Delta u', \nu)_{\mathcal{I}} \end{aligned}$$

$$
\begin{aligned}
&= -(\alpha \nabla u', \nabla v)_{\mathcal{I}} - \frac{\alpha c_1}{\alpha |\Gamma|} \int_\Gamma m'\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big) \frac{\partial u'}{\partial n}\, d\Gamma \\
&\quad -(u', \mu)_{\mathcal{I}} - \frac{1}{|\Gamma|} \int_\Gamma \frac{\partial}{\partial n}(y - \psi)\, \frac{\partial y'}{\partial n}\, d\Gamma - \int_\Gamma \frac{\partial \mu}{\partial n}\, \frac{\partial}{\partial n}(y - \psi)\, v_n\, d\Gamma \\
&= \alpha (u', \Delta v)_{\mathcal{I}} + \alpha \int_\Gamma \frac{\partial v}{\partial n}\, \frac{\partial}{\partial n}(u + \Delta\psi)\, v_n\, d\Gamma - (u', \mu)_{\mathcal{I}} \\
&\quad - \frac{c_1}{|\Gamma|} \int_\Gamma m'\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big) \frac{\partial u'}{\partial n}\, d\Gamma - \frac{1}{|\Gamma|} \int_\Gamma \frac{\partial}{\partial n}(y - \psi)\, \frac{\partial y'}{\partial n}\, d\Gamma \\
&\quad - \int_\Gamma \frac{\partial \mu}{\partial n}\, \frac{\partial}{\partial n}(y - \psi)\, v_n\, d\Gamma.
\end{aligned}
$$

Hence we obtain

$$
\begin{aligned}
&\frac{1}{|\Gamma|} \int_\Gamma \left( \frac{\partial}{\partial n}(y - \psi)\, \frac{\partial y'}{\partial n} + c_1\, m'\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big) \frac{\partial u'}{\partial n} \right) d\Gamma \\
&\qquad + c_2 \int_{\mathcal{I}} m'(y - \psi)\, y'\, d\mathbf{x} \\
&\quad = \int_\Gamma \Big( \alpha\, \frac{\partial v}{\partial n}\, \frac{\partial}{\partial n}(u + \Delta\psi) - \frac{\partial \mu}{\partial n}\, \frac{\partial}{\partial n}(y - \psi) \Big)\, v_n\, d\Gamma.
\end{aligned}
$$

We are now able to formulate the following result.

**Theorem 2** *Suppose the regularity assumptions of theorem 1 hold. Then the Eulerian derivative of the cost functional $K$ defined in* (3.6) *is given by*

$$
\begin{aligned}
dK(\Gamma, V) &= \int_\Gamma \Big( \alpha\, \frac{\partial v}{\partial n}\, \frac{\partial}{\partial n}(u + \Delta\psi) - \frac{\partial \mu}{\partial n}\, \frac{\partial}{\partial n}(y - \psi) \Big)\, v_n\, d\Gamma \\
&\quad + \frac{c_1}{|\Gamma|} \int_\Gamma m'\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big) \Big( \frac{1}{\alpha}(y - y_d) + \Delta^2 \psi \Big)\, v_n\, d\Gamma \\
&\quad - \frac{1}{|\Gamma|} \int_\Gamma \kappa\, \left( \frac{1}{2}\Big|\frac{\partial}{\partial n}(y - \psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big) \right)\, v_n\, d\Gamma \\
&\quad - \frac{1}{|\Gamma|^2} \int_\Gamma \kappa\, v_n\, d\Gamma \\
&\quad \cdot \left( \int_\Gamma \left( \frac{1}{2}\Big|\frac{\partial}{\partial n}(y - \psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big) \right) d\Gamma \right)
\end{aligned}
\tag{4.15}
$$

*where $(v, \mu)$ denotes the solution to the adjoint boundary value problem* (4.13), (4.14)*, $\kappa$ denotes the mean curvature of $\Gamma$, and $v_n = V(0) \cdot \mathbf{n}$ on $\Gamma$. We can therefore identify the gradient of $K$ with respect to the geometry $\Gamma$*

*with the normal vector field*

$$
\begin{aligned}
\nabla_\Gamma K = \Big( &\alpha \frac{\partial \nu}{\partial n} \frac{\partial}{\partial n}(u + \Delta\psi) - \frac{\partial \mu}{\partial n} \frac{\partial}{\partial n}(y - \psi) \\
&+ \frac{c_1}{|\Gamma|} m'\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big) \Big(\frac{1}{\alpha}(y - y_d) + \Delta^2\psi\Big) \\
&- \frac{1}{|\Gamma|}\kappa \left(\frac{1}{2}\Big|\frac{\partial}{\partial n}(y - \psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big)\right) \\
&- \frac{\kappa}{|\Gamma|^2} \int_\Gamma \left(\frac{1}{2}\Big|\frac{\partial}{\partial n}(y - \psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big)\right) d\Gamma \Big) n
\end{aligned}
\tag{4.16}
$$

## 5 Implementation

In this section we introduce the discretized version of the level set based algorithm considered in section 3 and discuss details of the implementation.

### *5.1 The discrete algorithm*

In the subsequent algorithm for all quantities superscript $h$ refers to the discrete counterpart of the respective continuous variable. Details of the discretization are specified in section 5.2 below.

**Discrete level set based algorithm.**

Step 0. Choose an appropriate initial $\Gamma_0^h$; set $n = 0$.
Step 1. Compute $(y_n^h, u_n^h)$ from the discrete relaxed system $(O_r^h(\Gamma_n^h))$.
Step 2. Evaluate the cost functional $K^h(\Gamma_n^h)$ and compute its derivative $\nabla_\Gamma^h K^h(\Gamma_n^h)$. If $\|\nabla_\Gamma^h K^h(\Gamma_n^h)\| = 0$ then stop; otherwise continue with step 3.
Step 3. Use an appropriate extension of $\nabla_\Gamma^h K^h(\Gamma_n^h)$ to $\Omega^h$ as speed function in the discrete level set equation for updating the level set function $\Phi_n^h$.
Step 4. Set $\Gamma_{n+1}^h$ equal to the zero level set of the updated level set function $\Phi_{n+1}^h$, and put $n := n + 1$. Go to step 1.

Let us now address the steps of the above algorithm in detail.

### *5.2 Aspects of the implementation*

The discretization of the relaxed first order system $(O_r)$, *i.e.* (2.14), (2.15a), (2.15b), (2.15e), can be based on either finite differences or finite elements, and it is–to some extent–independent of the discretization of the level set equation. Below we concentrate on a finite difference discretization for both.

We assume that the discretized domain $\Omega^h$ is given by a uniform grid with mesh size $h$. We denote the grid points by $\mathbf{x}_i$, $i = 1, \ldots, N$. For the discretization of the Laplace operator we use the standard five point stencil with an appropriate modification near the discrete interface $\Gamma^h$ (details are given below in the description of step 1). The grid functions $y^h$, $u^h$, $y_d^h$, $\psi^h$, ... are defined on the grid points. The level set equation is discretized on the same grid. By $\Phi^h$ we denote the discrete level set function defined on the grid points.

*Step 0* The algorithm requires an appropriate initialization respecting the discrete analogue of condition (2.15e). For this purpose we determine the set

$$\mathcal{M}^h := \{\mathbf{x}_i \in \Omega^h : (y_d^h - \psi^h - \alpha(\Delta^h)^2(\psi^h))(\mathbf{x}_i) \geq 0\}$$

(see (3.7)) and choose a function $\tilde{\Phi}_0^h$ with $\tilde{\Phi}_0^h(\mathbf{x}_i) = 0$ for $\mathbf{x}_i \in \Gamma_0^h$. Here $\Gamma_0^h$ is an appropriate initial interface for which we assume that $\Gamma_0$ is a closed curve in $\mathcal{M}$ on the continuous level. The level set function $\Phi_0^h$ used in the discrete algorithm is a signed distance function, i.e. $\Phi_0^h(\mathbf{x}_i) = \pm d$ with $d$ the distance of $\mathbf{x}_i$ to the interface. The sign is chosen in such a way that $\Phi_0^h(\mathbf{x}_i) > 0$ for $\mathbf{x}_i \in \mathcal{I}_0^h$ and $\Phi_0^h(\mathbf{x}_i) < 0$ for $\mathbf{x}_i \in \mathcal{A}_0^h$. Here $\mathcal{I}_0^h$ and $\mathcal{A}_0^h$ are the initial estimates of $\mathcal{I}^{h,*}$ and $\mathcal{A}^{h,*}$, the inactive and active sets at the optimal solution. The signed distance character of $\Phi_0^h$ is obtain by computing $\Phi_0^h$ as the numerical solution to the discretized version of

$$\Phi_t = \text{sign}(\Phi)(1 - |\nabla\Phi|). \tag{5.1}$$

Here sign$(\Phi)$ gives the sign of $\Phi$. The evolution equation (5.1) is initialized by $\tilde{\Phi}_0^h$ with settings $\tilde{\Phi}_0^h(\mathbf{x}_i) > 0$ for $i \in \mathcal{I}_0^h$ and $\tilde{\Phi}_0^h(\mathbf{x}_i) < 0$ for $i \in \mathcal{A}_0^h$. For more details on the numerical solution of (5.1) we refer to [20].

*Step 1* The solution $(y_n^h, u_n^h)$ of the discretization of the system (3.1),(3.2) has to be computed. As noted earlier, the discretization of the Laplace operator is based on a five-point finite difference stencil, which is regular at *interior inactive points*, i.e. points where no stencil neighbor is active. A currently inactive point with one or more active stencil neighbors is called *boundary inactive point*. For these grid points the five-point stencil is modified yielding a Shortly-Weller-type difference scheme. For the numerical realization of the boundary conditions additional *boundary points* are computed. A boundary point $\mathbf{x}_i \in \Omega$ is defined by $\bar{\Phi}_n^h(\mathbf{x}_i) = 0$ on the grid, i.e. one grid neighbor (nodal point) of $\mathbf{x}_i$ is inactive and the other one is active. Here $\bar{\Phi}_n^h$ denotes a linear interpolation of $\Phi_n^h$. On the boundary points the discretized boundary conditions (3.2b) are realized. The set of all boundary points at iteration level $n$ is denoted by $\Gamma_n^h$. (See Fig. 2.)

*Step 2* The computation of the cost functional is realized in the following way. The integral over $\mathcal{I}_n$ is approximated by means of a term resulting from

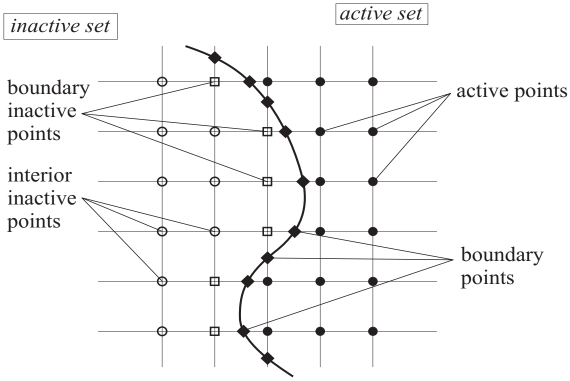


**Fig. 2.** Sketch of the computational grid

application of the trapezoidal rule on a grid shifted by $h/2$ for interior inactive points plus an interface contribution.

Now we turn to the integrals over $\Gamma_n$. Recall that by (4.7)

$$\frac{\partial}{\partial n}(y-\psi)\,n = \nabla(y-\psi).$$

In the first boundary integral the squared norm of the normal derivative on $\Gamma_n^h$ is replaced by the squared norm of the gradient on the related boundary inactive points. For the second boundary integral, due to the max operation involved we must specify a sign, which makes the numerical treatment more complicated. In our discrete algorithm the sign is determined by the value of $u_n^h + \Delta^h\psi^h$ at the boundary point minus the corresponding value at the related boundary inactive point. Again, the normal derivative is replaced by the gradient due to the same reasons as above.

For computing the gradient $\nabla_\Gamma^h K^h(\Gamma_n^h)$ the discretized adjoint system (4.13) has to be solved. The realization of the boundary conditions (4.14) is done by the same techniques as described in step 1 above. For the approximation $\kappa_n^h$ of the curvature at iteration level $n$ we refer to [20]. In our computations we use an upper bound to the curvature in order to avoid numerical instabilities resulting from huge curvature values at kinks or along edges.

*Step 3* Since the level set equation in step 3 is defined on the whole domain, an extension velocity has to be computed. This is necessary due to the fact that the continuous shape gradient $\nabla_\Gamma K(\Gamma_n)$ is defined only on the interface. Here we use a technique based on [1]. For the solution of the discrete level set equation we use an ENO-scheme for updating $\Phi_n^h$ with an explicit Euler step in time. The time-step size is adjusted dynamically by relaxing the CFL-condition (see [14]) and enlarging or reducing the time-step size according to the evolution of the cost functional. The technique applied here corresponds to an Armijo type line search and is described in detail in [11].

*Step 4* Due to the discrete update process $\Phi_{n+1}^h$ may lose its signed distance character. For this reason we incorporate a reinitialization step in our discrete algorithm. Here the aim is to keep the zero-level-set of $\Phi_{n+1}^h$ while regaining the signed distance nature of the discrete level set function. Since this aim is related to the initialization step 0 we use (5.1) with $\Phi_{n+1}^h$ as its initialization.

## 6 Numerical results

We shall now discuss several numerical test runs for different geometric situations at the solution. In all examples listed below the domain is fixed to $\Omega = (-1, 1)^2$, the bound on the state is chosen to be $\Psi \equiv 1$, $\alpha$ =5E-3, and for the discretization we use $h = 1/30$.

*Example 1* The penalty parameters have values $c_1 = 0.05$ and $c_2 = 1$. For the desired state $y_d \equiv 1.2$ is chosen yielding $\mathcal{M}^h = \Omega^h$. Figure 3 displays the state and control upon termination of the algorithm and some snapshots of the evolution of the zero-level set of $\Phi_n^h$. The white area represents the active set upon termination of the algorithm, the gray area is the corresponding inactive set. The algorithm terminated at iteration 13 with a $K$-value of $K_{13}^h = 1.88$E-3. The initial zero-level set is a circle comprising the zero-level set at the solution. Thus, $\Phi_n^h$ has to evolve in a way such that the initially convex active set estimate shrinks to the non-convex active set at the solution. The violation of the relaxed boundary conditions is of the order of 1E-6.

We also compared our new algorithm with the primal-dual active set method (pdAS) of [6]. It is known that the number of iterations required by pdAS significantly depends on the mesh size of the discretization. For the present example we found the following behavior: As can be seen from Table 1 pdAS requires approximately twice as many iterations when reducing the mesh size by a factor of $\frac{1}{2}$. On the other hand, the level set based algorithm (LSA) requires approximately the same number of iterations for all mesh sizes. This behavior can be attributed to our line search technique which allows to relax the CFL-condition and, thus, essentially decouples the time step size in the discretization of the level set equation and $h$, the mesh size for the spatial discretization; see also [12]. A similar observation holds true for the numerical examples discussed below.

**Table 1.** Comparison of pdAS and the new level set based algorithm (LSA)

| $h$ | 1/30 | 1/45 | 1/60 |
|---|---|---|---|
| pdAS | 11 | 16 | 21 |
| LAS | 8 | 7 | 9 |

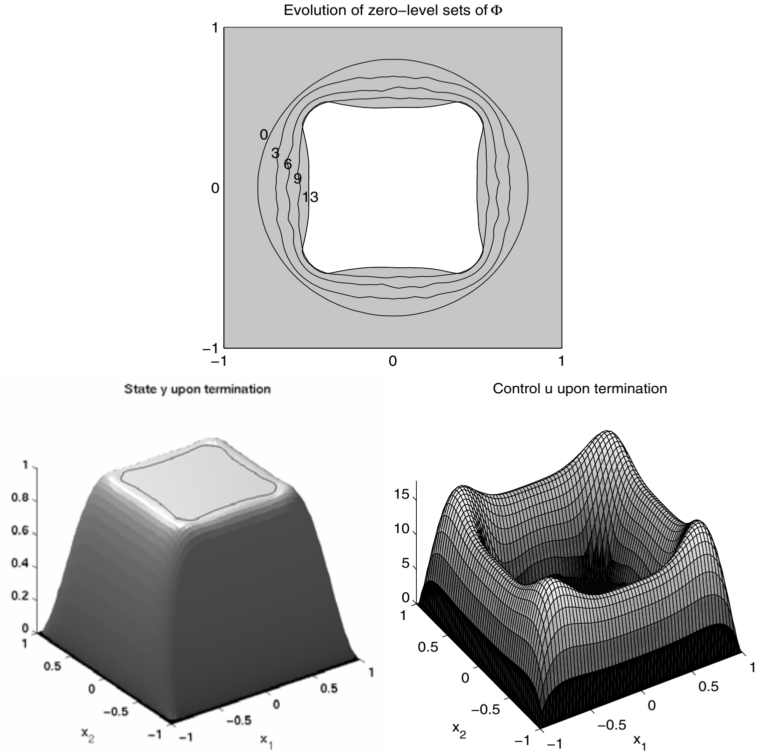


**Fig. 3.** Evolution of zero level sets (upper plot); state and control upon termination (lower plots)

*Example 2* The desired state $y_d^h$ is displayed in Figure 4. From the state upon termination and the structure of the corresponding active set (white area) in Figure 4 we observe that the active set consists of two disjoint components. For this example the initial zero-level set coincides with the boundary of $\mathcal{M}$ and comprises both components of the active set at the solution. Thus, in the course of the iterations the initial active set has to collapse onto two separated components which is a numerically challenging situation. The penalty parameters are tuned during the iteration, i.e. initially we use $c_1 = 0.1$ and $c_2 = 1\text{E}4$. From iteration 2–6 we set $c_1 = 1$, and for iteration levels greater than 6 we fix $c_1 = 5$. This particular tuning is due to stability and constraint violation reasons. If $c_1$ is chosen too large initially, then one has to reduce the time step size significantly in order to avoid unstable behavior of the evolution of the interface. In the course of the iterations the $c_1$-term in the cost functional decreases. Thus we can gradually increase the penalty parameter to force the iterates to become feasible. At iteration 23 the algorithm stops with an actual $K$-value of $K_{23}^h = 2.0\text{E-}2$.

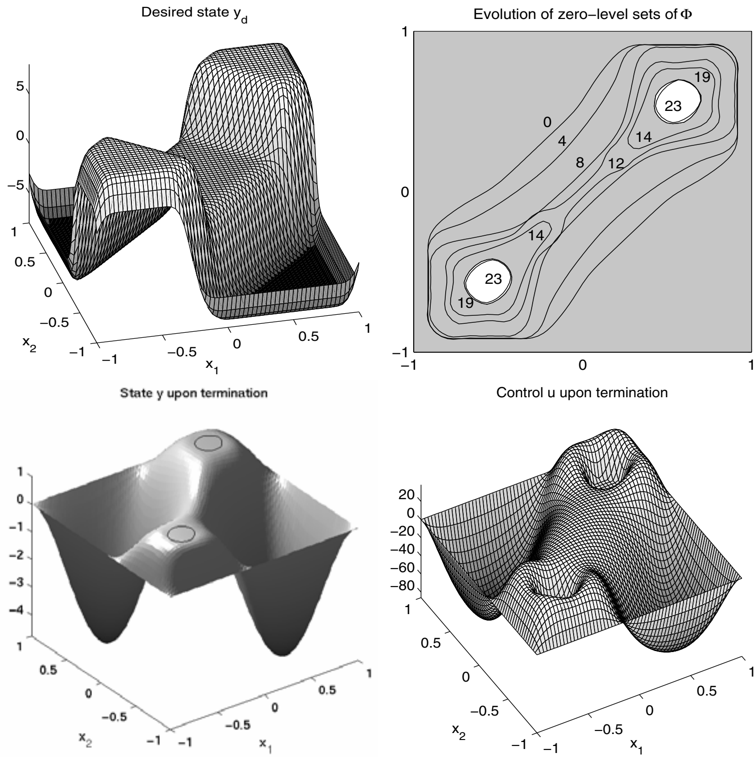


**Fig. 4.** Desired state and evolution of zero level sets (upper plots); state and control upon termination (lower plots)

*Example 3* In this example we demonstrate the ability of the new algorithm to expand the zero-level set of $\Phi^h$ from an initially symmetric to a non-symmetric shape at the optimal solution. The desired state is displayed in the upper left graph of Fig. 5. The penalty parameters are initialized to $c_1 = 0.1$ and $c_2 = 1\text{E}3$. At iteration 18 $c_2$ is increased to $c_2 = 5\text{E}3$. We also note that the $c_1$-contribution to the cost functional and gradient is already zero at iteration 18. Thus no adaptation of $c_1$ is necessary. At iteration 23 we fix $c_2 = 1\text{E}5$. The algorithm terminates at iteration 28 with $K_{28}^h = 1.02\text{E-}2$ and a constraint violation of the order of 6.01E-8. Finally we shall point out another important ability of the new algorithm. In fact, as can be seen from the evolution of the zero-level sets in the upper right graph of Fig. 5, the first iteration achieves a dramatic improvement over the initial configuration, i.e. the zero-level set moves from the initial small circle to a much larger shape such that $\mathcal{A}_1^h$ covers a big part of the active set at termination of the algorithm. This particular behavior cannot be observed from algorithms like the primal-dual active set strategy [6] or interior point methods [4], [22].

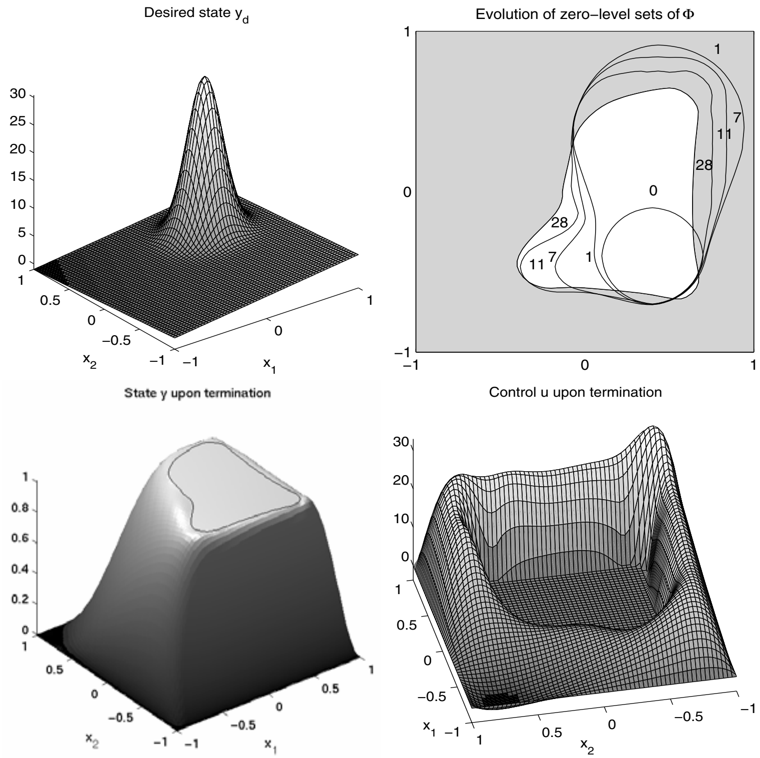


**Fig. 5.** Desired state and evolution of zero level sets (upper plots); state and control upon termination (lower plots)

## 7 Conclusions

The numerical treatment of state constrained optimal control problems represents a significant challenge. In this paper, based on a thorough analysis of the first order necessary and sufficient optimality conditions we have given a characterization of the optimal solution as the solution to a related free boundary problem. Due to the requirements and the regularity properties of the Lagrange multiplier on the boundary (interface) between the active and inactive sets it is rather natural to consider the interface as optimization variable. We have adapted level set methods to the present situation because of their efficiency, flexibility and robustness in tracking interfaces. These properties are desirable in our context since we cannot assume to have a priori knowledge of the shape of the interface at the optimal solution. It turns out that tools from shape optimization are well suited for computing the speed function needed in the level set equation for propagating the interface in the course of the iteration of our algorithm. Our numerical results are very encouraging since the newly introduced algorithm copes with topological changes and allows significant improvements from one iteration to the next. Especially the latter behavior can usually not be observed from traditional techniques like the primal-dual active set strategy and interior point methods.

## Appendix A. Proof of Theorem 1

We set

$$y^t = y_t \circ T_t \text{ and } u^t = u_t \circ T_t.$$

Note that $y^t : \mathcal{I} \to \mathbb{R}$ and $u^t : \mathcal{I} \to \mathbb{R}$. Using the regularity assumptions on $V$ it follows that there exists a constant $C > 0$ such that $\|y^t\|_{H^1(\mathcal{I})} \leq C$ and $\|u^t\|_{H^1(\mathcal{I})} \leq C$. We define

$$z^t = \frac{1}{t}(y^t - y) \text{ and } v^t = \frac{1}{t}(u^t - u).$$

Since $\psi \in H^2(\Omega)$, it follows from the results in [21, p. 64, Lem. 2.31 and p. 65, Prop. 2.32] (with slight modifications for the $L^2$-situation) that

$$\frac{1}{t}(\psi \circ T_t - \psi) \to (\nabla\psi, V(0)) \text{ in } L^2(\Omega)$$

and analogously that

$$\begin{aligned} \nabla\left(\frac{1}{t}(\psi \circ T_t - \psi)\right) &= \frac{1}{t}(\nabla\psi \circ T_t - \nabla\psi)\, DT_t + \frac{1}{t}(DT_t - I)\nabla\psi \\ &\to D^2\psi \cdot V(0) + DV(0) \cdot \nabla\psi = \nabla(\nabla\psi, V(0)) \end{aligned}$$

in $L^2(\Omega)$. Thus, $\frac{1}{t}(\psi \circ T_t - \psi) \to (\nabla\psi, V(0))$ in $H^1(\Omega)$. For the trace $z^t|_\Gamma = \frac{1}{t}(\psi \circ T_t - \psi)|_\Gamma$, it follows that $z^t|_\Gamma \to (\nabla\psi, V(0))|_\Gamma$ in $H^{\frac{1}{2}}(\Gamma)$. Analogously, we conclude that $v^t|_\Gamma \to -(\nabla(\Delta\psi), V(0))$ in $H^{\frac{1}{2}}(\Gamma)$.

The functions $(y^t, u^t)$ and $(z^t, v^t)$ are the solutions of appropriate boundary value problems on $\mathcal{I}$, which we now investigate in more detail. The weak form of (3.1) for $(y_t, u_t)$ on $\mathcal{I}_t$ is given by

$$-(\nabla y_t, \nabla\varphi_t)_{\mathcal{I}_t} + (u_t, \varphi_t)_{\mathcal{I}_t} = 0 \tag{A.1a}$$

$$(y_t, \varphi_t)_{\mathcal{I}_t} + \alpha(\nabla u_t, \nabla\varphi_t)_{\mathcal{I}_t} = (y_d, \varphi_t)_{\mathcal{I}_t} \tag{A.1b}$$

for all $\varphi_t \in H_0^1(\mathcal{I}_t)$. Pulling (A.1) back to $\mathcal{I}$ using $y^t = y_t \circ T_t$ and $u^t = u_t \circ T_t$ leads to

$$-(A(t)\nabla y^t, \nabla\varphi)_{\mathcal{I}} + (\gamma(t)u^t, \varphi)_{\mathcal{I}} = 0 \tag{A.2a}$$

$$(\gamma(t)y^t, \varphi)_{\mathcal{I}} + \alpha(A(t)\nabla u^t, \nabla\varphi)_{\mathcal{I}} = (\gamma(t)y_d, \varphi)_{\mathcal{I}} \tag{A.2b}$$

for all $\varphi \in H_0^1(\mathcal{I}_t)$. Here $\gamma(t) = \det(DT_t)$ and $A(t) = \gamma(t)(DT_t^*)^{-1}\,(DT_t)^{-1}$.

Let us now consider $(z^t, v^t)$. Using (A.2), we find

$$-(\nabla z^t, \nabla\varphi)_{\mathcal{I}} + (v^t, \varphi)_{\mathcal{I}} = -\Big(\frac{1}{t}(I - A(t))\nabla y^t, \nabla\varphi\Big)_{\mathcal{I}} + \Big(\frac{1}{t}(1-\gamma(t))u^t, \varphi\Big)_{\mathcal{I}} \tag{A.3a}$$

$$(z^t, \varphi)_{\mathcal{I}} + \alpha(\nabla v^t, \nabla\varphi)_{\mathcal{I}} = \Big(\frac{1}{t}(1-\gamma(t))y^t, \varphi\Big)_{\mathcal{I}} + \alpha\Big(\frac{1}{t}(I - A(t))\nabla u^t, \nabla\varphi\Big)_{\mathcal{I}} - \Big(\frac{1}{t}(1-\gamma(t))y_d, \varphi\Big)_{\mathcal{I}} \tag{A.3b}$$

for all $\varphi \in H_0^1(\mathcal{I})$. The smoothness assumptions on $V$ (see [21, p. 64, Lem. 2.31]) imply that

$$\frac{1}{t}(I - A(t)) \to -\mathrm{div}V(0)I + DV(0)^* + DV(0) \tag{A.4}$$

and

$$\frac{1}{t}(1-\gamma(t)) \to -\mathrm{div}V(0) \tag{A.5}$$

in $\mathcal{C}(\overline{\Omega})$. Consequently, $\frac{1}{t}(I - A(t))$ and $\frac{1}{t}(1-\gamma(t))$ are bounded in $\mathcal{C}(\overline{\Omega})$ for $t > 0$. We have already seen that the families $z^t|_{\partial\mathcal{I}}$ and $v^t|_{\partial\mathcal{I}}$ converge in $H^{\frac{1}{2}}(\partial\mathcal{I})$. Using appropriate inverse trace operators we find $p^t, p \in H^1(\mathcal{I})$ and $q^t, q \in H^1(\mathcal{I})$ such that $p^t|_{\partial\mathcal{I}} = z^t|_{\partial\mathcal{I}}$, $q^t|_{\partial\mathcal{I}} = v^t|_{\partial\mathcal{I}}$, with $p^t \to p$ in $H^1(\mathcal{I})$ and $q^t \to q$ in $H^1(\mathcal{I})$. We set $\tilde{z}^t = z^t - p^t$ and $\tilde{v}^t = v^t - q^t$. Obviously $(\tilde{z}^t, \tilde{v}^t)$ satisfy homogeneous Dirichlet boundary conditions on $\partial\mathcal{I}$. If we insert $z^t = \tilde{z}^t + p^t$ and $v^t = \tilde{v}^t + q^t$ in (A.3) we obtain

$$\begin{aligned}-(\nabla\tilde{z}^t, \nabla\varphi)_{\mathcal{I}} + (\tilde{v}^t, \varphi)_{\mathcal{I}} &= -\Big(\frac{1}{t}\big(I - A(t)\big)\nabla y^t, \nabla\varphi\Big)_{\mathcal{I}} \\ &\quad + \Big(\frac{1}{t}\big(1-\gamma(t)\big)u^t, \varphi\Big)_{\mathcal{I}} + (\nabla p^t, \nabla\varphi)_{\mathcal{I}} - (q^t, \varphi)_{\mathcal{I}} \\ &=: \langle f_1^t, \varphi\rangle_{H^{-1}(\mathcal{I}), H_0^1(\mathcal{I})}\end{aligned} \tag{A.6}$$

and

$$\begin{aligned}(\tilde{z}^t, \varphi)_{\mathcal{I}} + \alpha(\nabla\tilde{v}^t, \nabla\varphi)_{\mathcal{I}} &= \Big(\frac{1}{t}\big(1-\gamma(t)\big)(y^t - y_d), \varphi\Big)_{\mathcal{I}} \\ &\quad + \alpha\Big(\frac{1}{t}\big(I - A(t)\big)\nabla u^t, \nabla\varphi\Big)_{\mathcal{I}} - (p^t, \varphi)_{\mathcal{I}} - \alpha(\nabla q^t, \nabla\varphi)_{\mathcal{I}} \\ &=: \langle f_2^t, \varphi\rangle_{H^{-1}(\mathcal{I}), H_0^1(\mathcal{I})}\end{aligned} \tag{A.7}$$

for all $\varphi \in H_0^1(\mathcal{I})$. It is easily seen that the right-hand sides of (A.6) and (A.7) define continuous functionals $f_1^t$, $f_2^t$ on $H_0^1(\mathcal{I})$. Proceeding as in the proof of proposition 3 we can combine (A.6) and (A.7) to obtain

$$(\text{A.8})\quad (\nabla\tilde{z}^t, \nabla\varphi)_{\mathcal{I}} - \frac{1}{\alpha}(\Delta^{-1}\tilde{z}^t, \varphi)_{\mathcal{I}} = -\Big\langle f_1^t + \frac{1}{\alpha}\Delta^{-1} f_2^t, \varphi\Big\rangle_{H^{-1}(\mathcal{I}), H_0^1(\mathcal{I})}.$$

With $\varphi = \tilde{z}^t$ as test function and the positivity of the second term in (A.8) we find that there exists a constant $C > 0$ such that

$$(\text{A.9})\qquad \|\tilde{z}^t\|_{H_0^1(\mathcal{I})} \le C$$

for all $t$.

From (A.7) we find $\tilde{v}^t = \frac{1}{\alpha}\Delta^{-1}(\tilde{z}^t - f_2^t)$. Thus, $\tilde{v}^t$ is bounded in $H_0^1(\mathcal{I})$ uniformly with respect to $t$, and consequently $v^t$ is bounded in $H_0^1(\mathcal{I})$. We find also $u^t \to u$ in $H^2(\mathcal{I})$ as $t \to 0$. Estimate (A.9) allows to choose a subsequence of $\{\tilde{z}^t\}$ (which we denote again by the same expression) such that $\tilde{z}^t \rightharpoonup \tilde{z}$ weakly in $H_0^1(\mathcal{I})$ for $t \to 0$. Analogously we find $\tilde{v}^t \rightharpoonup \tilde{v}$ weakly in $H_0^1(\mathcal{I})$. We set

$$\begin{aligned}\langle f_1, \varphi\rangle_{H^{-1}(\mathcal{I}), H_0^1(\mathcal{I})} &= \big((\operatorname{div} V(0) - DV(0)^* - DV(0))\,\nabla y, \nabla\varphi\big)_{\mathcal{I}} \\ &\quad -(\operatorname{div} V(0)u, \varphi)_{\mathcal{I}} + (\nabla p, \nabla\varphi)_{\mathcal{I}} - (q, \varphi)_{\mathcal{I}}\end{aligned}$$

and

$$\begin{aligned}\langle f_2, \varphi\rangle_{H^{-1}(\mathcal{I}), H_0^1(\mathcal{I})} &= -(\operatorname{div} V(0)(y - y_d), \varphi)_{\mathcal{I}} \\ &\quad -\alpha\big((-\operatorname{div} V(0) + DV(0)^* + DV(0))\,\nabla u, \nabla\varphi\big)_{\mathcal{I}} \\ &\quad -(p, \varphi)_{\mathcal{I}} - \alpha(\nabla q, \nabla\varphi)_{\mathcal{I}}\end{aligned}$$

for $\varphi \in H_0^1(\mathcal{I})$. The boundedness of $\{\tilde{z}^t\}$ in $H_0^1(\mathcal{I})$ implies the boundedness of $\{z^t\}$ in $H^1(\mathcal{I})$. Hence, we have $y^t \to y$ strongly in $H^1(\mathcal{I})$. Likewise, we obtain $u^t \to u$ in $H^1(\mathcal{I})$. We also have $p^t \to p$ in $H^1(\mathcal{I})$, $q^t \to q$ in $H^1(\mathcal{I})$ and by (A.4) and (A.5) we find $f_1^t \to f_1$ and $f_2^t \to f_2$ strongly in $H^{-1}(\mathcal{I})$ as $t \to 0$. This and the weak convergence of $\tilde{z}^t$ and $\tilde{v}^t$ implies

$$(\text{A.10})\qquad \begin{aligned}-(\nabla\tilde{z}, \nabla\varphi)_{\mathcal{I}} + (\tilde{v}, \varphi)_{\mathcal{I}} &= (f_1, \varphi)_{\mathcal{I}} \\ (\tilde{z}, \varphi)_{\mathcal{I}} + \alpha(\nabla\tilde{v}, \nabla\varphi)_{\mathcal{I}} &= (f_2, \varphi)_{\mathcal{I}}\end{aligned}$$

for all $\varphi \in H_0^1(\mathcal{I})$. Since the right-hand side of (A.10) is independent of the subsequences chosen in the above weak-compactness argument and since the solution $(\tilde{z}, \tilde{v})$ is uniquely determined by $(f_1, f_2)$ we conclude that the limits $\tilde{z}$ and $\tilde{v}$ are independent of the chosen subsequence. An elementary argument shows that $\tilde{z}^t \rightharpoonup \tilde{z}$ and $\tilde{v}^t \rightharpoonup \tilde{v}$ weakly in $H^1(\mathcal{I})$ not only for some subsequence but for the whole original family $(\tilde{z}^t, \tilde{v}^t)_{t>0}$. With $z^t = \tilde{z}^t + p^t$ and $v^t = \tilde{v}^t + q^t$ we obtain $z^t \rightharpoonup z = \tilde{z} + p$ and $v^t \rightharpoonup v = \tilde{v} + q$ weakly in $H_0^1(\mathcal{I})$.

## References

1. Adalsteinsson, D., Sethian, J.A.: The fast construction of extension velocities in level set methods. J. Comput. Phys. **148**(1), 2–22 (1999)
2. Alibert, J.-J., Raymond, J.-P.: A Lagrange multiplier theorem for control problems with state constraints. Numer. Funct. Anal. Optim. **19**(7-8), 697–704 (1998)
3. Bergounioux, M.: Augmented Lagrangian method for distributed optimal control problems with state constraints. J. Optim. Theory Appl. **78**(3), 493–521 (1993)
4. Bergounioux, M., Haddou, M., Hintermüller, M., Kunisch, K.: A comparison of a Moreau-Yosida-based active set strategy and interior point methods for constrained optimal control problems. SIAM J. Optim. **11**(2), 495–521 (2000), (electronic)
5. Bergounioux, M., Kunisch, K.: Augmented Lagrangian techniques for elliptic state constrained optimal control problems. SIAM J. Control Optim. **35**(5), 1524–1543 (1997)
6. Bergounioux, M., Kunisch, K.: Primal-dual strategy for state-constrained optimal control problems. Preprint, Special research center for Optimization and Control, 2000
7. Casas, E.: Control of an elliptic problem with pointwise state constraints. SIAM J. Control Optim. **24**(6), 1309–1318 (1986)
8. Casas, E., Tröltzsch, F., Unger, A.: Second order sufficient optimality conditions for some state-constrained control problems of semilinear elliptic equations. SIAM J. Control Optim. **38**(5), 1369–1391 (2000), (electronic)
9. Heinkenschloss, M.: SQP interior-point methods for distributed optimal control problems. In: Floudas Pardalos, (ed.) Encyclopedia of Optimization. Boston: Kluwer Academic Publishers, 2000
10. Hintermüller, M., Ring, W.: An inexact Newton-cg-type active contour approach for the minimization of the Mumford-Shah functional. J. Math. Imag. Vision. **20**, 19–42 (2004)
11. Hintermüller, M., Ring, W.: A second order shape optimization approach for image segmentation. SIAM J. Appl. Math. **64**(2), 442–467 (2003), (electronic)
12. Hintermüller, M., Ring, W.: Numerical aspects of a level set based algorithm for state constrained optimal control problems. Computer Assisted Mechanics and Eng. Sci. **10**(2), 149–161 (2003)
13. Kinderlehrer, D., Stampacchia, G.: An introduction to variational inequalities and their applications, Vol. **31** of Classics in Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2000. Reprint of the 1980 original
14. LeVeque, R.J.: Numerical methods for conservation laws. Basel: Birkhäuser Verlag, second edition, 1992
15. Lions, J.-L., Magenes, E.: Problèmes aux limites non homogènes et applications. Vol. **1**. Travaux et Recherches Mathématiques, No. 17. Dunod, Paris, 1968
16. Lunéville, E., Mignot, F.: Un problème de contrôle avec contraintes sur l'état. In: Control of partial differential equations (Santiago de Compostela, 1987), Berlin: Springer, 1989, pp. 208–212
17. Maurer, H., Mittelmann, H.D.: Optimization techniques for solving elliptic control problems with control and state constraints. I. Boundary control. Comput. Optim. Appl. **16**(1), 29–55 (2000)
18. Rudin, W.: Real and Complex Analysis. New York: McGraw-Hill, third edition, 1987
19. Sethian, J.A.: Fast marching methods. SIAM Rev. **41**(2), 199–235 (1999), (electronic)
20. Sethian, J.A.: Level set methods and fast marching methods. Cambridge: Cambridge University Press, second edition, 1999. Evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science

21. Sokołowski, J., Zolésio, J.-P.: Introduction to shape optimization. Berlin: Springer-Verlag, 1992. Shape sensitivity analysis
22. Vicente, L.N.: On interior-point Newton algorithms for discretized optimal control problems with state constraints. Optim. Methods Softw. **8**(3-4), 249–275 (1998)
23. Wloka, J.: Partielle Differentialgleichungen. Stuttgart: Teubner, 1982
24. Wright, S.J.: Primal-dual interior-point methods. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 1997
25. Ye, Y.: Interior point algorithms. New York: John Wiley & Sons Inc., 1997. Theory and analysis, A Wiley-Interscience Publication