A LEVEL SET APPROACH FOR THE SOLUTION OF A STATE-CONSTRAINED OPTIMAL CONTROL PROBLEM

MICHAEL HINTERMÜLLER AND WOLFGANG RING

Abstract. 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.

1. Introduction

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

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

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

$$K = \{v \in H^1_0(\Omega) |\ v \le \psi \text{ a.e. on } \Omega\} \subset H^1_0(\Omega)\,,$$

where ψ 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. [12], [10]). The problem of imposing constraints on the state has received considerable attention. The contributions in [8], [7], and [2] are concerned with theoretical aspects of deriving first and second order conditions characterizing optimal solutions. In [4], [5], [6], [13], and [19] 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 [6]. Since they are frequently slow in their practical performance, techniques based on an augmented Lagrangian approach have been introduced in [4]. These techniques typically outperform the Uzawa-based methods but the

Date: February 2001.

1991 *Mathematics Subject Classification.* 35R35, 49K20, 49Q10, 65K10.

Key words and phrases. State constrained optimal control, level set method, shape optimization, free boundary problem.

structure of the constraints is not fully exploited. The recently developed primal-dual active set strategy [5] 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 [6] the primal variables, i.e. state and control variable, are somewhat emphasized, in [4] 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 [5] 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. [21],[22]). The numerical comparison in [3], [5] 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 [5]. 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 Γ 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 [17] 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 [18] is employed.

In the present paper we focus on the theoretical aspects of combining primal-dual first order characterizations, free boundary aspects, tools from shape optimization, and level set methods, the presentation of the algorithm and its analysis together with selected numerical results. In an upcoming paper details of the implementation of the new algorithm will be addressed. Moreover a report on further numerical test runs will be given.

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. The final 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.

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

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

$$\text{subject to } \begin{cases} \Delta y + u = 0 \text{ on } \Omega \\ y \leq \psi \text{ a.e. on } \Omega \end{cases} \tag{2.1b}$$

$$\text{over } (y,u) \in H^1_0(\Omega) \times L^2(\Omega).$$

In [5] the following optimality system for (2.1) is given: The pair $(y^*, u^*) \in H^1_0(\Omega) \times L^2(\Omega)$ 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 Ω) such that

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

$$y^* \leq \psi \text{ on } \Omega \tag{2.2b}$$

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

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

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 [15, 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 Ω. Therefore $\mathcal{I}^*$ is an open subset of $\mathbb{R}^n$. We set

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

See Figure 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) $\text{int}(\mathcal{A}^*) \neq \emptyset.$

(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 Σ imply that $\Sigma \cap \Gamma^* = \emptyset$.

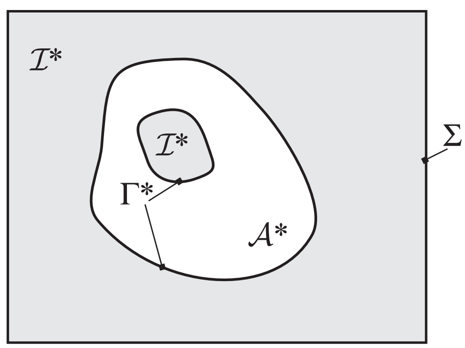

FIGURE 1. Sketch of the geometric situation at the solution to (2.1).

Next we turn towards the boundary conditions satisfied by 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

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

Now we consider u^*. It is found in [5] that the measure λ^* is concentrated on $\mathcal{A}^*$. Let $\varphi \in \mathcal{D}(\mathcal{I}^*)$ be arbitrary. Since $\varphi \in H_0^1(\Omega) \cap H^2(\Omega)$, we can use φ 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 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 $u^* \in \mathcal{W}(\mathcal{I}^*)$. Next we prove that there exists a bounded Dirichlet trace operator $w \mapsto w|_{\partial\mathcal{I}^*}$ on $\mathcal{W}(\mathcal{I}^*)$. It is known that there exists a continuous linear extension operator $Z : H^{\frac{1}{2}}(\partial\mathcal{I}^*) \to H^2(\mathcal{I}^*)$ which satisfies

$$Zf = v \in H^2(\mathcal{I}^*),$$
$$v|_{\partial\mathcal{I}^*} = 0, \text{ and } \frac{\partial v}{\partial n}\Big|_{\partial\mathcal{I}^*} = f$$

(see [20, p. 133, Thm 8.8]). Here and in the following n denotes the unit exterior normal vector on $\mathcal{I}^*$. We also used the regularity assumption (A3). For every $u \in \mathcal{W}(\mathcal{I}^*)$ we define the Dirichlet trace $\gamma_0 u$ as a functional in $H^{-\frac{1}{2}}(\partial\mathcal{I}^*)$ which satisfies

$$\langle \gamma_0 u, f\rangle_{H^{-\frac{1}{2}}(\partial\mathcal{I}^*), H^{\frac{1}{2}}(\partial\mathcal{I}^*)} = (u, \Delta(Zf))_{\mathcal{I}^*} - (\Delta u, Zf)_{\mathcal{I}^*} \tag{2.7}$$

for every $f \in H^{\frac{1}{2}}(\partial\mathcal{I}^*)$. We prove that (2.7) defines a unique functional $\gamma_0 u \in H^{-\frac{1}{2}}(\partial\mathcal{I}^*)$ which depends linearly and continuously on $u \in \mathcal{W}(\mathcal{I}^*)$, but does not depend on the particular choice of the extension operator Z. Since $Z : H^{\frac{1}{2}}(\partial\mathcal{I}^*) \to H^2(\mathcal{I}^*)$ and $\Delta : H^2(\mathcal{I}^*) \to L^2(\mathcal{I}^*)$ are bounded linear operators, the right-hand side of (2.7) defines a bounded linear functional on $H^{\frac{1}{2}}(\partial\mathcal{I}^*)$. It remains to prove that the right-hand side of (2.7) depends only on u and f, but not on the particular choice of the extension operator Z. Suppose v_1 and v_2 both satisfy $v_1, v_2 \in H^2(\mathcal{I}^*)$, $v_1|_{\partial\mathcal{I}^*} = v_2|_{\partial\mathcal{I}^*} = 0$, and $\frac{\partial v_1}{\partial n}\big|_{\partial\mathcal{I}^*} = \frac{\partial v_2}{\partial n}\big|_{\partial\mathcal{I}^*} = f$. Then $\bar v = v_1 - v_2 \in H^2_0(\mathcal{I}^*)$. Since $\mathcal{D}(\mathcal{I}^*)$ is dense in $H^2_0(\mathcal{I}^*)$ we can use $\bar v$ as a test function in the definition of the distributional Laplace operator:

$$(\Delta u, \bar v)_{\mathcal{I}^*} = (u, \Delta \bar v)_{\mathcal{I}^*}$$

for all $u \in \mathcal{W}(\mathcal{I}^*)$ and therefore

$$(\Delta u, v_1)_{\mathcal{I}^*} - (u, \Delta v_1)_{\mathcal{I}^*} = (\Delta u, v_2)_{\mathcal{I}^*} - (u, \Delta v_2)_{\mathcal{I}^*}$$

for all $u \in \mathcal{W}(\mathcal{I}^*)$. This implies that definition (2.7) is independent of the particular choice of Z. We shall use the notation $\gamma_0 u = u|_{\partial\mathcal{I}^*}$ or even $\gamma_0 u = u$ when it is clear from the context. For $y \in H^1_0(\mathcal{I}^*) \cap H^2(\mathcal{I}^*)$ and $u \in \mathcal{W}(\mathcal{I}^*)$, we obtain Green's formula

$$\Big\langle \gamma_0 u, \frac{\partial y}{\partial n}\Big\rangle_{H^{-\frac{1}{2}}(\partial\mathcal{I}^*), H^{\frac{1}{2}}(\partial\mathcal{I}^*)} = (u, \Delta y)_{\mathcal{I}^*} - (\Delta u, y)_{\mathcal{I}^*}. \tag{2.8}$$

Using an extension operator $\tilde Z : H^{\frac{3}{2}}(\partial\mathcal{I}^*) \to H^2(\mathcal{I}^*)$ which satisfies

$$\tilde Z f = \tilde v \in H^2(\mathcal{I}^*),$$
$$\tilde v|_{\partial\mathcal{I}^*} = f, \text{ and } \frac{\partial \tilde v}{\partial n}\Big|_{\partial\mathcal{I}^*} = 0,$$

we can likewise prove that a continuous linear Neumann trace operator $\gamma_1 : \mathcal{W}(\mathcal{I}^*) \to H^{-\frac{3}{2}}(\partial\mathcal{I}^*)$ exists for which the Green's formula

$$\big\langle \gamma_1 u, y\big\rangle_{H^{-\frac{3}{2}}(\partial\mathcal{I}^*), H^{\frac{3}{2}}(\partial\mathcal{I}^*)} = (\Delta u, y)_{\mathcal{I}^*} - (u, \Delta y)_{\mathcal{I}^*} \tag{2.9}$$

holds for all $u \in \mathcal{W}(\mathcal{I}^*)$ and $y \in H^2(\mathcal{I}^*)$ with $\frac{\partial y}{\partial n}\big|_{\partial\mathcal{I}^*} = 0$. We shall use the notation $\gamma_1 u = \frac{\partial u}{\partial n}\big|_{\partial\mathcal{I}^*}$.

Now let $f \in H^{\frac{1}{2}}(\Sigma)$ be arbitrarily given. We define $\tilde{f} \in H^{\frac{1}{2}}(\partial\mathcal{I}^*)$ by $\tilde{f} = f$ on Σ and $\tilde{f} = 0$ on Γ^*. (Here we used the assumption that $\Sigma \cap \Gamma^* = \emptyset$.) Let $y = Z\tilde{f} \in H^2(\mathcal{I}^*)$. Since $y = \frac{\partial y}{\partial n} = 0$ on Γ^*, y can be extended by 0 to a function (which we also denote by y) in $H_0^1(\Omega) \cap H^2(\Omega)$. 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

$$u^*|_{\Sigma} = 0. \tag{2.10}$$

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

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

Therefore, we find

$$\lambda_a^* = y_d - \psi - \alpha\Delta^2\psi \in L^2(\operatorname{int}(\mathcal{A}^*)). \tag{2.11}$$

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 $Z_2 : H^{\frac{1}{2}}(\Gamma^*) \to H^2(\operatorname{int}(\mathcal{A}^*))$ 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

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

On the other hand, using (2.8), (2.5) and Green's formula we have

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

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

$$u^*|_{\Gamma^*} = -\Delta\psi|_{\Gamma^*} \in H^{\frac{3}{2}}(\Gamma^*). \tag{2.12}$$

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 Ω we obtain

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

On the other hand, using 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 Γ^*, 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 find

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

and 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

$$\frac{\partial y^*}{\partial n}\Big|_{\Gamma^*} = \frac{\partial \psi}{\partial n}\Big|_{\Gamma^*} \in H^{\frac{5}{2}}(\Gamma^*). \tag{2.13}$$

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 $H^2(\mathrm{int}(\mathcal{A}^*))$, the Green's formula (2.9), and (2.5) we get

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

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

$$\lambda_s^*|_{\Gamma^*} = -\alpha \frac{\partial}{\partial n_{\mathcal{I}^*}}(u^* + \Delta\psi)\big|_{\Gamma^*} \in H^{-\frac{3}{2}}(\Gamma^*) \tag{2.14}$$

or, more explicitly

$$\int_{\Gamma^*} y\, d\lambda_s^* = -\alpha\Big\langle \frac{\partial}{\partial n_{\mathcal{I}^*}}(u^*+\Delta\psi), y\Big\rangle_{H^{-\frac{3}{2}}(\Gamma^*),H^{\frac{3}{2}}(\Gamma^*)} \tag{2.15}$$

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

Proposition 1. (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*

$$\Delta y^* + u^* = 0 \text{ on } \mathcal{I}^* \tag{2.16a}$$

$$y^* - \alpha\Delta u^* = y_d \text{ on } \mathcal{I}^* \tag{2.16b}$$

$$y^* < \psi \text{ on } \mathcal{I}^* \tag{2.16c}$$

$$y^*|_\Sigma = 0, \quad u^*|_\Sigma = 0, \tag{2.17a}$$

$$y^*|_{\Gamma^*} = \psi|_{\Gamma^*}, \quad u^*|_{\Gamma^*} = -\Delta\psi|_{\Gamma^*}, \tag{2.17b}$$

$$\frac{\partial y^*}{\partial n}\Big|_{\Gamma^*} = \frac{\partial \psi}{\partial n}|_{\Gamma^*} \tag{2.17c}$$

$$-\frac{\partial}{\partial n}(u^* + \Delta\psi)|_{\Gamma^*} \geq 0 \text{ as a measure on } \Gamma^* \tag{2.17d}$$

$$y_d - \psi - \alpha\Delta^2\psi \geq 0 \text{ almost everywhere on } \mathcal{A}^*. \tag{2.17e}$$

(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* (2.16) *and* (2.17) *on* $\hat{\mathcal{I}}$ *and* $\hat{\mathcal{A}}$ *respectively (i.e. all* $*$*-expressions are replaced by the corresponding* $\hat{}$ *-expressions). Then* $(\bar{y},\bar{u})$ *defined by*

$$\bar{y} = \begin{cases} \hat{y} \text{ on } \hat{\mathcal{I}} \\ \psi \text{ on } \hat{\mathcal{A}} \end{cases} \quad \text{and } \bar{u} = \begin{cases} \hat{u} \text{ on } \hat{\mathcal{I}} \\ -\Delta\psi \text{ on } \hat{\mathcal{A}} \end{cases} \tag{2.18}$$

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

Proof. We have already proved that (2.16) and (2.17a)–(2.17c) must hold for the optimal configuration $(y^*, u^*, \mathcal{I}^*, \mathcal{A}^*)$. It follows from (2.2d) that λ^* defines a positive functional on $\mathcal{C}_c(\mathcal{A}^*)$. Hence, by the Riesz representation theorem [15, p. 40. Thm 2.14] λ^* is a positive measure on $\mathcal{A}^*$. Thus, $\lambda_s^* = \lambda^*|_{\Gamma^*}$ and $\lambda_a^* = \lambda^*|_{\mathrm{int}(\mathcal{A}^*)}$ are also positive measures. Application of the characterizations (2.11) and (2.14) proves that (2.17e) and (2.17d) must hold.

Suppose conversely that (2.16) and (2.17) are satisfied by $(\hat{y}, \hat{u})$ on $\hat{\mathcal{I}}$ and $\hat{\mathcal{A}}$, respectively. Let $(\bar{y}, \bar{u})$ be defined by (2.18). We set $\bar{\lambda} = \bar{\lambda}_a + \bar{\lambda}_s \in \mathcal{M}(\Omega)$ where

$$\langle \bar{\lambda}_s, y \rangle_{\mathcal{M},\mathcal{C}_0} = -\Big\langle \frac{\partial}{\partial n}(\bar{u} + \Delta\psi), y \Big\rangle_{H^{-\frac{3}{2}}(\bar{\Gamma}), H^{\frac{3}{2}}(\bar{\Gamma})} \tag{2.19}$$

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

$$\langle \bar{\lambda}_a, y \rangle_{\mathcal{M},\mathcal{C}_0} = \int_{\hat{\mathcal{A}}} (y_d - \psi - \alpha\Delta^2\psi)\, y \, d\mathbf{x}. \tag{2.20}$$

Since $H^2(\Omega) \cap H_0^1(\Omega)$ is densely embedded in $\mathcal{C}_0(\Omega)$, the definition (2.19) 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 (2.16c) and $\bar{y} = \psi$ on $\hat{\mathcal{A}}$. Thus, (2.2b) is satisfied. Let $\varphi \in \mathcal{D}(\Omega)$. We have

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

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

Expressions (2.19) and (2.20) 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 (2.16a), (2.16b), (2.17a), (2.17b), and (2.17c). 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 (2.16), (2.17) iteratively approximate the optimal solution by a sequence $\{(y_n, u_n)\}$; see [4], [5], [6], [13], [19]. Every iterate (y_n, u_n) induces a corresponding geometric configuration given by $\mathcal{A}_n$, $\mathcal{I}_n$, Γ_n. Here $\mathcal{A}_n = \{\mathbf{x} \in \Omega : y_n(\mathbf{x}) = \psi(\mathbf{x})\}$, $\mathcal{I}_n = \Omega \setminus \mathcal{A}_n$ and Γ_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 (2.16), (2.17) 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 (2.17) and solve (2.16a) and (2.16b) with the remaining boundary conditions for (y, u). The violation of the relaxed boundary conditions together with a possible violation of the constraint $y \leq \psi$ 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 (2.16c) 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 (2.17e) 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 Γ of the (current) inactive set $\mathcal{I}$. It is therefore sufficient to consider the evolution of Γ driven by the gradient $\nabla_\Gamma K(\Gamma)$ of the cost functional with respect to Γ. It is well established (see [17], [16]) 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 (2.16),(2.17) with inactive set $\mathcal{I}_n$ induced by the actual boundary Γ_n and where one of the boundary conditions in (2.17b),(2.17c) is neglected. Moreover $\|\cdot\|$ shall denote an appropriate norm which is specified below.

Level set based algorithm.

Step 0. *Choose an appropriate initial* Γ_0*; set* $n = 0$.
Step 1. *Compute* (y_n, u_n) *from the relaxed system* $(O_r(\Gamma_n))$.
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* Ω *as speed function in the level set equation for updating the level set function.*
Step 4. *Set* Γ_{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

$$\Delta y + u = 0 \text{ on } \mathcal{I}, \tag{3.1a}$$

$$y - \alpha\Delta u = y_d \text{ on } \mathcal{I} \tag{3.1b}$$

$$y|_\Sigma = 0, \quad u|_\Sigma = 0, \tag{3.2a}$$

$$y|_\Gamma = \psi|_\Gamma, \quad u|_\Gamma = -\Delta\psi|_\Gamma \tag{3.2b}$$

Proposition 2. *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

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

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

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

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 shall also write $(y, u) = \big(y(\mathcal{I}), u(\mathcal{I})\big)$ We define the cost functional

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

Thus, the boundary condition (2.17c) and the inequality constraints (2.17d) and (2.16c) 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 (2.17e) is not represented in the cost functional. This constraint depends only on a priori given data and on the geometry. If we define

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

we can choose some subset of $\mathcal{M}$ as starting value for the active set $\mathcal{A}$. Thus, feasibility with respect to (2.17e) holds for the initial configuration. In all our numerical tests feasibility of (2.17e) is maintained for all updates of the geometry such that a representation of (2.17e) 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 Γ. Here we rely on the concepts and results given in [18]. 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 Σ. We set $\mathcal{I}_t = T_t(\mathcal{I})$ and $\Gamma_t = T_t(\Gamma)$. The Eulerian derivative of K at Γ 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 [18, Thm 2.27,p 59]) that there exists a distribution F on Γ 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 Γ 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 $F\mathbf{n}$ normal to Γ with speed function $F = F(\Gamma, \mathbf{x})$ with $\mathbf{x} \in \Gamma$.

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 Φ 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 [17] 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 Ω or, at least, on some band containing Γ. It is therefore necessary to extend F from Γ onto Ω in some appropriate (smooth) way. The extension procedure is described in detail in an upcoming paper.

4. 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 Γ be of class $\mathcal{C}^3$ and $V \in \mathcal{C}^1([0,T);\mathcal{C}^3(\Omega,\mathbb{R}^n))$ be an admissible vector field. Assume moreover that $\psi \in H^5(\Omega)$. Then the solution (y,u) to (3.1), (3.2) has a material derivative*

$$(\dot{y}(\mathcal{I};V), \dot{u}(\mathcal{I};V)) \in H^2(\mathcal{I}) \cap H_0^1(\mathcal{I}) \times H^2(\mathcal{I}) \cap H_0^1(\mathcal{I}).$$

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 strong in $H_0^1(\mathcal{I})$ and weak in $H^2(\mathcal{I})$. The analogous statement holds for $\dot{u}(\mathcal{I};V)$.

Proof. We set

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

Note that $y^t : \Omega \to \mathbb{R}$ and $u^t : \Omega \to \mathbb{R}$. 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 further set

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

On Γ we have

$$z^t(\mathbf{x}) = \frac{1}{t}(y_t \circ T_t(\mathbf{x}) - y(\mathbf{x})) = \frac{1}{t}\big(\psi(T_t(\mathbf{x})) - \psi(\mathbf{x})\big)$$

and

$$v^t(\mathbf{x}) = \frac{1}{t}(u_t \circ T_t(\mathbf{x}) - u(\mathbf{x})) = -\frac{1}{t}\big(\Delta\psi(T_t(\mathbf{x})) - \Delta\psi(\mathbf{x})\big)$$

for $\mathbf{x} \in \Gamma$. The regularity assumptions on V imply that

$$\frac{1}{t}(\psi \circ T_t - \psi) \to \nabla\psi \cdot V(0) \text{ in } H^3(\Omega)$$

and

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

as $t \to 0$. (See [18, p.71,Prop 2.37]). Therefore, we have $z^t \to \nabla\psi \cdot V(0)$ in $H^{\frac{5}{2}}(\Gamma)$ and $v^t \to -\nabla(\Delta\psi) \cdot V(0)$ in $H^{\frac{3}{2}}(\Gamma)$. At this place we need the regularity assumption $\Gamma \in \mathcal{C}^3$. Consequently, $z^t|_\Gamma$ is bounded in $H^{\frac{5}{2}}(\Gamma)$ and $v^t|_\Gamma$ is bounded in $H^{\frac{3}{2}}(\Gamma)$.

Now we investigate certain boundary value problems which are satisfied by (y,u), (y^t,u^t), and (z^t,v^t). The weak form of (3.1) for (y_t,u_t) on $\mathcal{I}_t$ is given by

$$-(\nabla y_t, \nabla\varphi_t)_{\mathcal{I}_t} + (u_t, \varphi_t)_{\mathcal{I}_t} = 0 \tag{4.1a}$$

$$(y_t, \varphi_t)_{\mathcal{I}_t} + \alpha(\nabla u_t, \nabla\varphi_t)_{\mathcal{I}_t} = (y_d, \varphi_t)_{\mathcal{I}_t} \tag{4.1b}$$

for all $\varphi_t \in H_0^1(\mathcal{I}_t)$. Pulling (4.1) back to $\mathcal{I}$ by $y^t = y_t \circ T_t$ and $u^t = u_t \circ T_t$ we obtain

$$-(A(t)\nabla y^t, \nabla\varphi)_{\mathcal{I}} + (\gamma(t)u^t, \varphi)_{\mathcal{I}} = 0 \tag{4.2a}$$

$$(\gamma(t)y^t, \varphi)_{\mathcal{I}} + \alpha(A(t)\nabla u^t, \nabla\varphi)_{\mathcal{I}} = (\gamma(t)y_d, \varphi)_{\mathcal{I}} \tag{4.2b}$$

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

At this place it is convenient to consider regularity properties of (y^t, u^t) in some detail. Using global elliptic regularity theory for uniformly elliptic operators (see e.g. [9, Sec.8.4] or [14, Sec.8.6,p.324]) it is possible to prove that $\|y^t\|_{H^2(\mathcal{I})} \leq C$ and $\|u^t\|_{H^2(\mathcal{I})} \leq C$ for all $t \in [0, T]$. We only sketch the proof at this place. We write (4.2) in the form

$$\begin{aligned} L_t y^t + \gamma(t)u^t &= 0 \\ \gamma(t)y^t - \alpha L_t u^t &= \gamma(t)y_d \end{aligned}$$

where L_t is the uniformly elliptic operator $L_t y = -\mathrm{div}(A(t)\nabla y)$. Using a continuous inverse trace operator we can find

$$\begin{aligned} r^t \in H^2(\mathcal{I}),\ r^t \to r \text{ as } t \to 0 \text{ in } H^2(\mathcal{I}),\ r^t|_\Gamma = y^t|_\Gamma \text{ and} \\ s^t \in H^2(\mathcal{I}),\ s^t \to s \text{ as } t \to 0 \text{ in } H^2(\mathcal{I}),\ s^t|_\Gamma = u^t|_\Gamma. \end{aligned}$$

With $\tilde{y}^t = y^t - r^t$ and $\tilde{u}^t = u^t - s^t$ we get

$$\begin{aligned} L_t\tilde{y}^t + \gamma(t)\tilde{u}^t &= g_1(t) \\ \gamma(t)\tilde{y}^t - \alpha L_t\tilde{u}^t &= g_2(t) \end{aligned}$$

with g_1, g_2 bounded in $L^2(\mathcal{I})$ uniformly with respect to t. Thus, we find

$$\tilde{u}^t = \frac{1}{\alpha}L_t^{-1}\big(\gamma(t)\tilde{y}^t - g_2(t)\big)$$

and therefore

$$L_t\tilde{y}^t + \frac{\gamma(t)}{\alpha}L_t^{-1}\big(\gamma(t)\tilde{y}^t\big) = g_1(t) + \frac{1}{\alpha}L_t^{-1}g_2(t) =: g(t). \tag{4.3}$$

It is easily seen that the right-hand side of (4.3) is uniformly bounded with respect to t in $L^2(\mathcal{I})$. Using arguments of the kind as described in [14, pp.322–333] we can derive an estimate of the form

$$\|\tilde{y}^t\|_{H^2(\mathcal{I})} \leq \tilde{C}\big(\|\tilde{y}^t\|_{L^2(\mathcal{I})} + \|g(t)\|_{L^2(\mathcal{I})}\big).$$

The constant $\tilde{C}$ depends on the (fixed) geometry, on $\mathcal{C}^1$-bounds for the coefficients of L_t and $\gamma(t)$, and on a coercivity bound for L_t. All these bounds can be chosen independently of t. Thus, we get a uniform H^2-bound for $\tilde{y}^t$ and consequently for y^t. The H^2-boundedness of u^t follows analogously.

Let us now consider (z^t, v^t). Using (4.2) we find

$$-(\nabla z^t, \nabla\varphi)_{\mathcal{I}} + (v^t,\varphi)_{\mathcal{I}} = -\Big(\frac{1}{t}(I-A(t))\nabla y^t, \nabla\varphi\Big)_{\mathcal{I}} + \Big(\frac{1}{t}(1-\gamma(t))u^t,\varphi\Big)_{\mathcal{I}} \tag{4.4a}$$

$$(z^t,\varphi)_{\mathcal{I}} + \alpha(\nabla v^t,\nabla\varphi)_{\mathcal{I}} = \Big(\frac{1}{t}(1-\gamma(t))y^t,\varphi\Big)_{\mathcal{I}} + \alpha\Big(\frac{1}{t}(I-A(t))\nabla u^t,\nabla\varphi\Big)_{\mathcal{I}} - \Big(\frac{1}{t}(1-\gamma(t))y_d,\varphi\Big)_{\mathcal{I}} \tag{4.4b}$$

for all $\varphi \in H_0^1(\mathcal{I})$. Using the smoothness assumptions on V we find

$$\frac{1}{t}(I-A(t)) \to -\mathrm{div}V(0)I + DV(0)^* + DV(0) \tag{4.5}$$

and

$$\frac{1}{t}(1-\gamma(t)) \to -\mathrm{div}V(0) \tag{4.6}$$

in $\mathcal{C}^1(\overline{\Omega})$. (See [18, p.64, Lem 2.31]). Consequently, $\frac{1}{t}(I-A(t))$ and $\frac{1}{t}(1-\gamma(t))$ are bounded in $\mathcal{C}^1(\overline{\Omega})$ for $t>0$. We have already seen that the families $z^t|_{\partial\mathcal{I}}$ and $v^t|_{\partial\mathcal{I}}$ are convergent in $H^{\frac{5}{2}}(\partial\mathcal{I})$ and $H^{\frac{3}{2}}(\partial\mathcal{I})$, respectively. Using appropriate inverse trace operators we find $p^t, p \in H^3(\mathcal{I})$ and $q^t, q \in H^2(\mathcal{I})$ 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 $H^3(\mathcal{I})$ and $q^t \to q$ in $H^2(\mathcal{I})$. 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 (4.4) we obtain

$$\begin{aligned} -(\nabla\tilde{z}^t,\nabla\varphi)_{\mathcal{I}} + (\tilde{v}^t,\varphi)_{\mathcal{I}} &= -\Big(\frac{1}{t}\big(I-A(t)\big)\nabla y^t,\nabla\varphi\Big)_{\mathcal{I}} \\ &\quad + \Big(\frac{1}{t}\big(1-\gamma(t)\big)u^t,\varphi\Big)_{\mathcal{I}} + (\nabla p^t,\nabla\varphi)_{\mathcal{I}} - (q^t,\varphi)_{\mathcal{I}} \\ &=: \langle f_1^t,\varphi\rangle_{H^{-1}(\mathcal{I}),H_0^1(\mathcal{I})} \end{aligned} \tag{4.7}$$

and

$$\begin{aligned} (\tilde{z}^t,\varphi)_{\mathcal{I}} + \alpha(\nabla\tilde{v}^t,\nabla\varphi)_{\mathcal{I}} &= \Big(\frac{1}{t}\big(1-\gamma(t)\big)(y^t-y_d),\varphi\Big)_{\mathcal{I}} \\ &\quad + \alpha\Big(\frac{1}{t}\big(I-A(t)\big)\nabla u^t,\nabla\varphi\Big)_{\mathcal{I}} - (p^t,\varphi)_{\mathcal{I}} - \alpha(\nabla q^t,\nabla\varphi)_{\mathcal{I}} \\ &=: \langle f_2^t,\varphi\rangle_{H^{-1}(\mathcal{I}),H_0^1(\mathcal{I})} \end{aligned} \tag{4.8}$$

for all $\varphi \in H_0^1(\mathcal{I})$. It is easily seen that the right-hand sides of (4.7) and (4.8) define continuous functionals f_1^t, f_2^t on $H_0^1(\mathcal{I})$. In fact, we have even more regularity for f_1^t and f_2^t. Using (4.5), (4.6), the boundedness of (p^t, q^t) in $H^3(\mathcal{I}) \times H^2(\mathcal{I})$ we can conclude that f_1^t and f_2^t are bounded in $L^2(\mathcal{I})$ uniformly in t. Proceeding as in the proof on proposition 2 we can combine (4.7) and (4.8) to obtain

$$(\nabla\tilde{z}^t,\nabla\varphi)_{\mathcal{I}} - \frac{1}{\alpha}(\Delta^{-1}\tilde{z}^t,\varphi)_{\mathcal{I}} = -\Big(f_1^t + \frac{1}{\alpha}\Delta^{-1}f_2^t,\varphi\Big)_{\mathcal{I}}. \tag{4.9}$$

Using $\varphi = \tilde{z}^t$ as test function together with the positivity of the second term in (4.9) we find that there exists a constant $C > 0$ such that

$$\|\tilde{z}^t\|_{H_0^1(\mathcal{I})} \leq C \tag{4.10}$$

for all t. Moreover, $\tilde{z}^t$ satisfies a uniformly elliptic equation with right-hand side bounded in $L^2(\mathcal{I})$ with respect to t. Thus, another application of global regularity theory yields

$$\|\tilde{z}^t\|_{H^2(\mathcal{I})} \leq C \tag{4.11}$$

for all $t \in [0, T]$. Since p^t is bounded in $H^2(\mathcal{I})$ it follows that z^t is bounded in $H^2(\mathcal{I})$. Furthermore, we have $y^t \to y$ in $H^2(\mathcal{I})$ as $t \to 0$.

From (4.8) we find $\tilde{v}^t = \frac{1}{\alpha}\Delta^{-1}(\tilde{z}^t - f_2^t)$. Thus, $\tilde{v}^t$ is bounded in $H^2(\mathcal{I})$ uniformly with respect to t, and consequently v^t is bounded in $H^2(\mathcal{I})$. We find also $u^t \to u$ in $H^2(\mathcal{I})$ as $t \to 0$. Estimate (4.11) 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 $H^2(\mathcal{I})$ for $t \to 0$. Analogously we find $\tilde{v}^t \rightharpoonup \tilde{v}$ weakly in $H^2(\mathcal{I})$. We set

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

and

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

for $\varphi \in H_0^1(\mathcal{I})$. Since $y^t \to y$ in $H^2(\mathcal{I})$, $u^t \to u$ in $H^2(\mathcal{I})$, $p^t \to p$ in $H^2(\mathcal{I})$, $q^t \to q$ in $H^2(\mathcal{I})$ and by (4.5) and (4.6) we find $f_1^t \to f_1$ and $f_2^t \to f_2$ strongly in $L^2(\mathcal{I})$ as $t \to 0$. This and the weak convergence of $\tilde{z}^t$ and $\tilde{v}^t$ implies

$$\begin{aligned}-(\nabla\tilde{z}, \tilde{\varphi})_{\mathcal{I}} + (\tilde{v}, \varphi)_{\mathcal{I}} = (f_1, \varphi)_{\mathcal{I}} \\ (\tilde{z}, \varphi)_{\mathcal{I}} + \alpha(\nabla\tilde{v}, \nabla\varphi)_{\mathcal{I}} = (f_2, \varphi)_{\mathcal{I}}\end{aligned} \tag{4.12}$$

for all $\varphi \in H_0^1(\mathcal{I})$. Since the right-hand side of (4.12) 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 $H^2(\mathcal{I})$ not only for some subsequence but for the whole original family $(\tilde{z}^t, \tilde{v}^t)_{t>0}$.

Since the embedding $H^2(\mathcal{I}) \cap H_0^1(\mathcal{I}) \hookrightarrow H_0^1(\mathcal{I})$ is compact we find that $\tilde{z}^t \to \tilde{z}$ and $\tilde{v}^t \to \tilde{v}$ strongly in $H_0^1(\mathcal{I})$. With $z^t = \tilde{z}^t + p^t$ and $v^t = \tilde{v}^t + q^t$ we obtain $z^t \to z = \tilde{z} + p$ and $v^t \to v = \tilde{v} + q$ strongly in $H_0^1(\mathcal{I})$ □

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

$$y'(\mathcal{I}; V) = \dot{y}(\mathcal{I}; V) - \nabla y(\mathcal{I}) \cdot V(0) \in H^1(\mathcal{I}) \tag{4.13}$$

$$u'(\mathcal{I}; V) = \dot{u}(\mathcal{I}; V) - \nabla u(\mathcal{I}) \cdot V(0) \in H^1(\mathcal{I}). \tag{4.14}$$

It follows from techniques presented in [18, p.118–119] that (y', u') is characterized as the unique solution to the boundary value problem

$$\Delta y' + u' = 0 \text{ on } \mathcal{I}, \tag{4.15a}$$

$$y' - \alpha \Delta u' = 0 \text{ on } \mathcal{I}, \tag{4.15b}$$

with boundary conditions

$$y'|_\Sigma = 0, \quad u'|_\Sigma = 0, \tag{4.16a}$$

$$y'|_\Gamma = -\frac{\partial}{\partial n}(y - \psi) v_n|_\Gamma, \quad u'|_\Gamma = -\frac{\partial}{\partial n}(u + \Delta\psi) v_n|_\Gamma \tag{4.16b}$$

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 Γ. We set

$$\zeta'(\Gamma; V) = \dot{\zeta}(\Gamma; V) - \nabla_\Gamma \zeta(\Gamma) \cdot V(0) \tag{4.17}$$

where $\nabla_\Gamma \zeta$ is the tangential gradient of ζ. 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 [18] 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

$$K_1(\Gamma) = \frac{1}{2|\Gamma|} \int_\Gamma \left| \frac{\partial}{\partial n}(y - \psi) \right|^2 d\Gamma \tag{4.18a}$$

$$K_2(\Gamma) = \frac{c_1}{2|\Gamma|} \int_\Gamma \Big(\max\Big(0, \frac{\partial}{\partial n}(u + \Delta\psi)\Big)\Big)^2 d\Gamma \tag{4.18b}$$

$$K_3(\Gamma) = \frac{c_2}{2} \int_\mathcal{I} \big(\max(0, y - \psi) \big)^2 \, d\mathbf{x}. \tag{4.18c}$$

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 [18, p.116, (2.173)] and (4.17) we find

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

Here κ denotes the mean curvature of Γ. From [18, 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 τ denotes the projection of the vector $\big(DV(0,\mathbf{x})^* n\big)$ onto the tangent space to Γ at $\mathbf{x}$. Application of the chain rule together with (4.13) gives

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

For the above formula we also used that $\psi'(\mathcal{I};\Gamma)=0$ since ψ is independent of Γ. 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 Γ by (3.2b) and $\Delta_\Gamma(y-\psi)=0$ due to $y-\psi=0$ on Γ. Here Δ_Γ denotes the Laplace-Beltrami operator on Γ. From $y-\psi=0$ on Γ it follows also that $\nabla_\Gamma(y-\psi)=0$ on Γ. Therefore, we have

$$
\nabla(y-\psi) = \frac{\partial}{\partial n}(y-\psi)\, n + \nabla_\Gamma(y-\psi) = \frac{\partial}{\partial n}(y-\psi)\, n \tag{4.19}
$$

and hence,

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

Thus, we get

$$
d\tilde{K}_1(V;\Gamma) = \int_\Gamma \Big(\frac{\partial}{\partial n}(y-\psi)\,\frac{\partial y'}{\partial n} - \frac{1}{2}\kappa\left|\frac{\partial}{\partial n}(y-\psi)\right|^2 v_n\Big)\, d\Gamma. \tag{4.20}
$$

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

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

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

$$\begin{aligned} d\tilde{K}_2(V;\Gamma) = &\int_\Gamma m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\left(\frac{\partial u'}{\partial n} + \Big(\frac{1}{\alpha}(y-y_d) + \Delta^2\psi\Big)\,v_n\right) d\Gamma \\ &- \int_\Gamma \kappa m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\,v_n\,d\Gamma. \end{aligned} \tag{4.21}$$

Applying [18, p.113, (2.168)] we find for K_3:

$$dK_3(V;\Gamma) = c_2\int_{\mathcal{I}} m'(y-\psi)\,y'\,d\mathbf{x}. \tag{4.22}$$

Finally, we set

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

Following [18, p. 80, prop.2.50 and p.93, (2.145)] we obtain

$$dK_0(V;\Gamma) = -\frac{1}{|\Gamma|^2}\int_\Gamma \kappa\,v_n\,d\Gamma. \tag{4.23}$$

Combining all preceding results we obtain

$$\begin{aligned} dK(V;\Gamma) = &\frac{1}{|\Gamma|}\int_\Gamma \left(\frac{\partial}{\partial n}(y-\psi)\,\frac{\partial y'}{\partial n} + c_1\,m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\,\frac{\partial u'}{\partial n}\right) d\Gamma \\ &+ c_2\int_{\mathcal{I}} m'(y-\psi)\,y'\,d\mathbf{x} \\ &+ \frac{c_1}{|\Gamma|}\int_\Gamma m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\,\Big(\frac{1}{\alpha}(y-y_d)+\Delta^2\psi\Big)\,v_n\,d\Gamma \\ &- \frac{1}{|\Gamma|}\int_\Gamma \kappa\,\left(\frac{1}{2}\Big|\frac{\partial}{\partial n}(y-\psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\right)\,v_n\,d\Gamma \\ &- \frac{1}{|\Gamma|^2}\int_\Gamma \kappa\,v_n\,d\Gamma\,\cdot \\ &\quad\left(\int_\Gamma\left(\frac{1}{2}\Big|\frac{\partial}{\partial n}(y-\psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\right)\,d\Gamma\right). \end{aligned} \tag{4.24}$$

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

$$\Delta\mu + \lambda = c_2 m'(y-\psi) \text{ on } \mathcal{I}, \tag{4.25a}$$

$$\mu - \alpha\Delta\lambda = 0 \text{ on } \mathcal{I} \tag{4.25b}$$

with boundary conditions

$$\mu|_\Sigma = 0, \quad \lambda|_\Sigma = 0, \tag{4.26a}$$

$$\mu|_\Gamma = \frac{1}{|\Gamma|}\,\frac{\partial}{\partial n}(y-\psi)\Big|_\Gamma, \quad \lambda|_\Gamma = -\frac{c_1}{\alpha|\Gamma|}m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\Big|_\Gamma. \tag{4.26b}$$

Using (4.25), (4.16), (4.15), and (4.26) we find

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

Hence we obtain

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

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*

$$
\begin{aligned}
dK(\Gamma, V) &= \int_\Gamma \Big(\alpha\frac{\partial\lambda}{\partial n}\frac{\partial}{\partial n}(u+\Delta\psi) - \frac{\partial\mu}{\partial n}\frac{\partial}{\partial n}(y-\psi)\Big)\, v_n\, d\Gamma \\
&\quad + \frac{c_1}{|\Gamma|}\int_\Gamma m'\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\Big(\frac{1}{\alpha}(y-y_d) + \Delta^2\psi\Big)\, v_n\, d\Gamma \\
&\quad - \frac{1}{|\Gamma|}\int_\Gamma \kappa\,\Big(\frac{1}{2}\Big|\frac{\partial}{\partial n}(y-\psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\Big)\, v_n\, d\Gamma \\
&\quad - \frac{1}{|\Gamma|^2}\int_\Gamma \kappa\, v_n\, d\Gamma\, \cdot \\
&\qquad \Big(\int_\Gamma \Big(\frac{1}{2}\Big|\frac{\partial}{\partial n}(y-\psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u+\Delta\psi)\Big)\Big)\, d\Gamma\Big)
\end{aligned}
\tag{4.27}
$$

where (λ, μ) denotes the solution to the adjoint boundary value problem (4.25), (4.26), *κ denotes the mean curvature of Γ, and $v_n = V(0)\cdot\mathbf{n}$ on Γ. We can therefore identify the gradient of K with respect to the geometry*

Γ *with the normal vector field*

$$\begin{aligned}\nabla_\Gamma K = \Bigg(& \alpha \frac{\partial \lambda}{\partial n} \frac{\partial}{\partial n}(u + \Delta\psi) - \frac{\partial \mu}{\partial n} \frac{\partial}{\partial n}(y - \psi) \\ & + \frac{c_1}{|\Gamma|} m'\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big) \Big(\frac{1}{\alpha}(y - y_d) + \Delta^2\psi\Big) \\ & - \frac{1}{|\Gamma|}\kappa \left(\frac{1}{2}\Big|\frac{\partial}{\partial n}(y - \psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big)\right) \\ & - \frac{\kappa}{|\Gamma|^2} \int_\Gamma \left(\frac{1}{2}\Big|\frac{\partial}{\partial n}(y - \psi)\Big|^2 + c_1 m\Big(\frac{\partial}{\partial n}(u + \Delta\psi)\Big)\right) d\Gamma \Bigg) n\end{aligned} \tag{4.28}$$

5. Numerical aspects

We shall now discuss numerical results attained by the level set based algorithm sketched at the beginning of section 3. Since the discretization and implementation of the algorithm are delicate issues which go far beyond the scope of the present work, we only discuss major discretization and implementation aspects and refer to an upcoming paper for more details and more numerical test results.

5.1. **Aspects of the implementation.** Step 0 of the algorithm requires an appropriate initialization respecting condition (2.17e). For this purpose we determine the set $\mathcal{M}$ (see (3.7)) and choose an initial level set function $\Phi_0(\mathbf{x}) = \Phi(t = 0, \mathbf{x})$ with the property that the zero-level set (interface) is a closed curve in $\mathcal{M}$. Moreover, Φ_0 is a signed distance function, i.e. $\Phi_0(\mathbf{x}) = \pm d$ with d the distance of $\mathbf{x}$ to the interface. The sign is chosen in such a way that $\Phi_0(\mathbf{x}) > 0$ for $\mathbf{x} \in \mathcal{I}_0$ and $\Phi_0(\mathbf{x}) < 0$ for $\mathbf{x} \in \mathcal{A}_0$. Here $\mathcal{I}_0$ and $\mathcal{A}_0$ are the initial estimates of $\mathcal{I}^*$ and $\mathcal{A}^*$, the inactive and active sets at the optimal solution. Let Φ_0^h denote the discretization of Φ_0, where h indicates the mesh size of the discretization. The domain Ω is discretized by a regular uniform grid yielding Ω^h. In the following subscripts 'n' denote quantities at iteration level n.

In step 1 the solution (y_n^h, u_n^h) of the discretization of the system (3.1),(3.2) has to be computed. 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 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 Φ_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 Γ_n^h. (See Figure 2.)

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

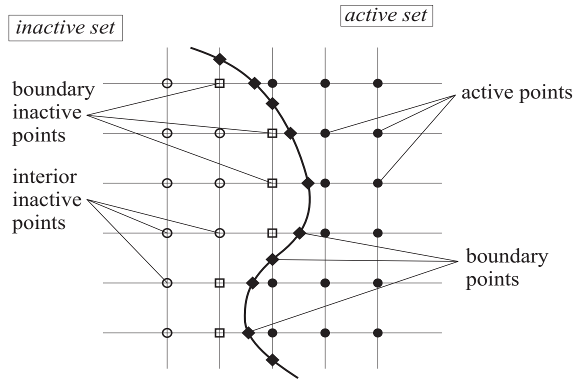

FIGURE 2. Sketch of the computational grid.

Now we turn to the integrals over Γ_n. Recall that by (4.19)

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

In the first boundary integral the squared norm of the normal derivative on Γ_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 $u_n^h + (\Delta\psi)^h$ 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 $(\nabla_\Gamma K(\Gamma_n))^h$ the discretized adjoint system (4.25) has to be solved. The realization of the boundary conditions (4.26) is done by the same techniques as described above. For the approximation κ_n^h of the curvature at iteration level n we refer to [17]. 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.

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 $\nabla_\Gamma K(\Gamma_n)$ 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 Φ_n^h . The time-step size is adjusted dynamically by relaxing the CFL-condition (see [11]) 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.

In order to obtain an efficient algorithm and good approximations to normal derivatives and some other quantities many details of the discretization and implementation become important. Also, tuning of certain parameters like c_1, c_2, or the time-step size significantly influences the number of iterations. A comprehensive description of these details will be given in an upcoming paper.

5.2. **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 $\Psi \equiv 1$, 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 Φ^h_n. 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 $K^h_{13} = 1.88$E-3. The initial zero-level set is a circle comprising the zero-level set at the solution. Thus, Φ^h_n has to evolve in a way such that the initially convex active set shrinks to the non-convex active set at the solution. The violation of the relaxed boundary conditions is of the order of 1E-6.

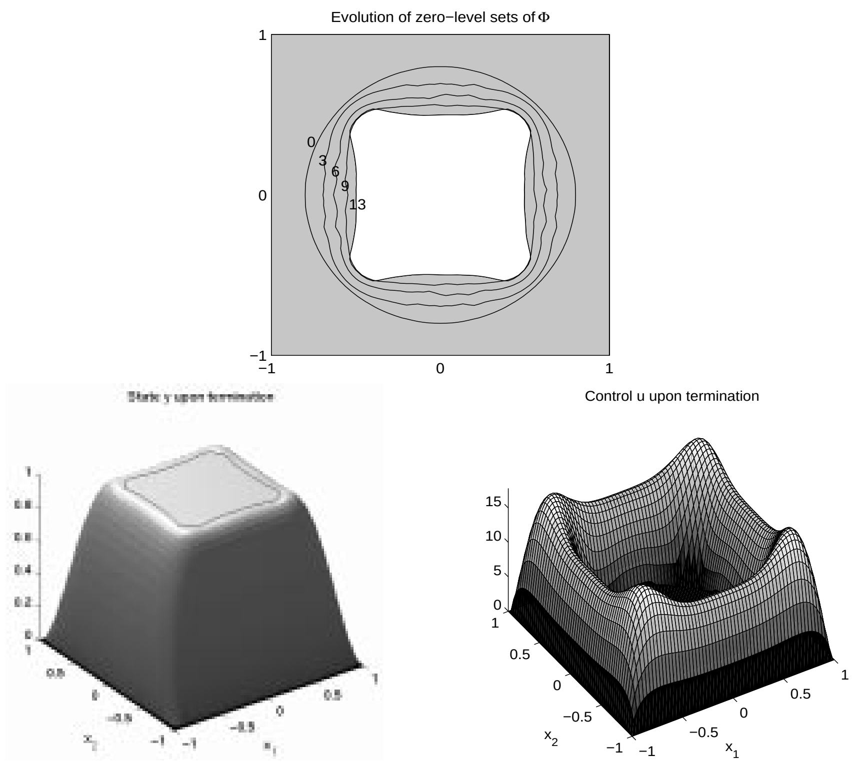

FIGURE 3. Evolution of zero level sets (upper plot); state and control upon termination (lower plots).

Example 2: The desired state y^h_d is displayed in Figure 4. From the state upon termination in 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 $c_2 = 1$E4. 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 behaviour 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 $K_{23}^h = 2.0\text{E-}2$.

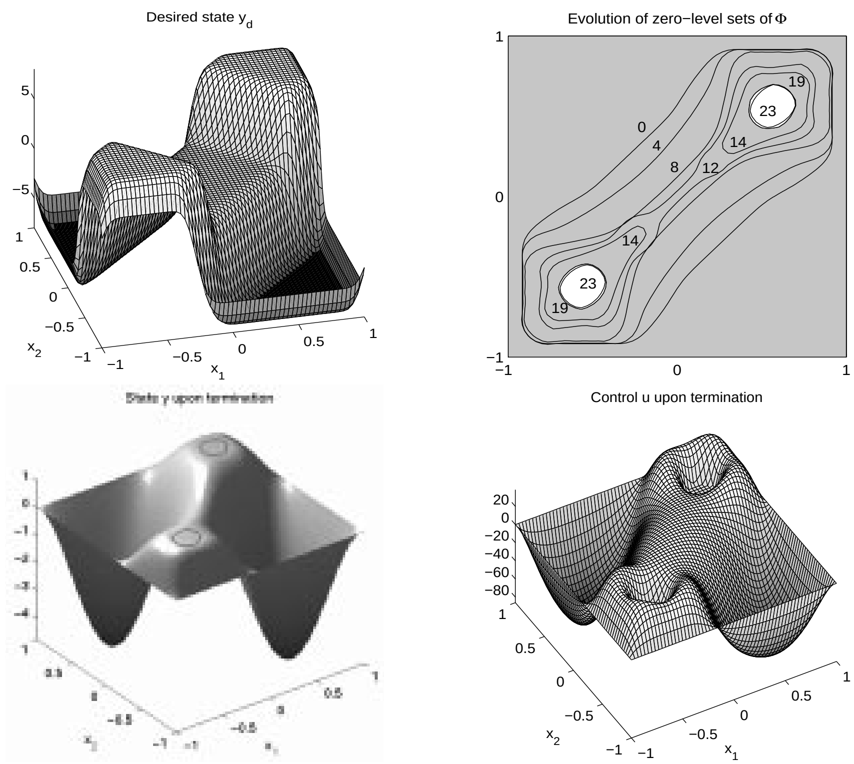

FIGURE 4. Desired state and evolution of zero level sets (upper plots); state and control upon termination (lower plots).

Example 3: In this example we demonstrate the ability of the new algorithm to expand the zero-level set of Φ^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 Figure 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 Figure 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 behaviour cannot be observed from algorithms like the primal-dual active set strategy [5] or interior point methods [3], [19].

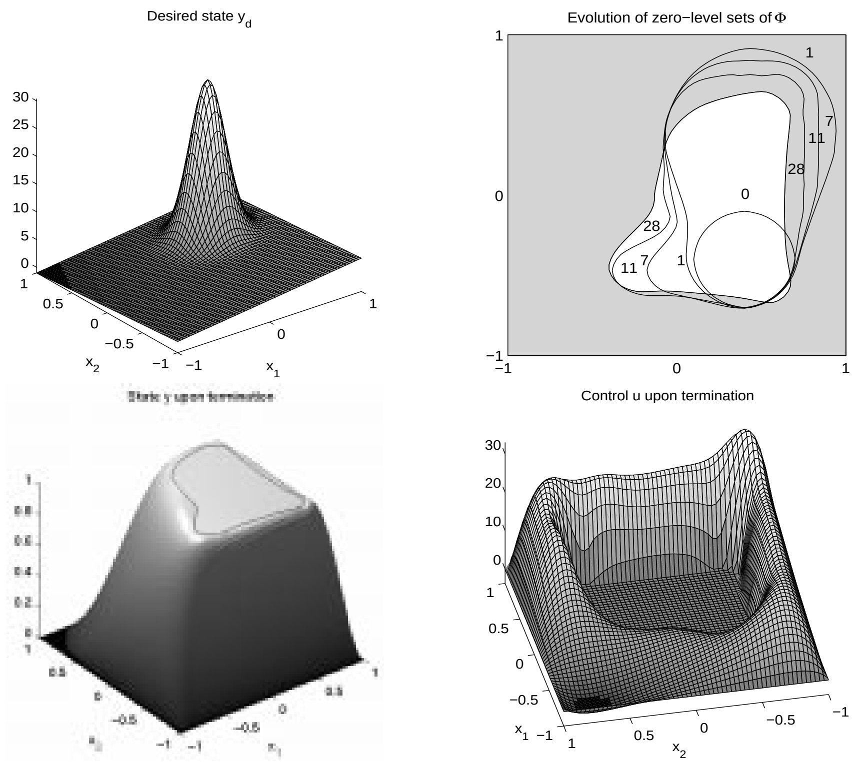

FIGURE 5. Desired state and evolution of zero level sets (upper plots); state and control upon termination (lower plots).

6. 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.

In an upcoming paper we will discuss further important aspects and details of our implementation as well as more numerical test results. In a future work we also intend to consider higher order optimization techniques.

References

[1] D. Adalsteinsson and J. A. Sethian. The fast construction of extension velocities in level set methods. *J. Comput. Phys.*, 148(1):2–22, 1999.

[2] J.-J. Alibert and J.-P. Raymond. A Lagrange multiplier theorem for control problems with state constraints. *Numer. Funct. Anal. Optim.*, 19(7-8):697–704, 1998.

[3] M. Bergounioux, M. Haddou, M. Hintermüller, and K. Kunisch. 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 (electronic), 2000.

[4] M. Bergounioux and K. Kunisch. Augmented Lagrangian techniques for elliptic state constrained optimal control problems. *SIAM J. Control Optim.*, 35(5):1524–1543, 1997.

[5] M. Bergounioux and K. Kunisch. Primal-dual strategy for state-constrained optimal control problems. *Preprint, Special research center for Optimization and Control*, 2000.

[6] M. Bergounioux. Augmented Lagrangian method for distributed optimal control problems with state constraints. *J. Optim. Theory Appl.*, 78(3):493–521, 1993.

[7] E. Casas, F. Tröltzsch, and A. Unger. Second order sufficient optimality conditions for some state-constrained control problems of semilinear elliptic equations. *SIAM J. Control Optim.*, 38(5):1369–1391 (electronic), 2000.

[8] E. Casas. Control of an elliptic problem with pointwise state constraints. *SIAM J. Control Optim.*, 24(6):1309–1318, 1986.

[9] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*, volume 224 of *Grundlehren der mathematischen Wissenschaften*. Springer Verlag, Berlin, 1977.

[10] M. Heinkenschloss. SQP interior-point methods for distributed optimal control problems. In Floudas Pardalos, editor, *Encyclopedia of Optimization*. Kluwer Academic Publishers, Boston, 2000.

[11] R. J. LeVeque. *Numerical methods for conservation laws*. Birkhäuser Verlag, Basel, second edition, 1992.

[12] E. Lunéville and F. Mignot. Un problème de contrôle avec contraintes sur l'état. In *Control of partial differential equations (Santiago de Compostela, 1987)*, pages 208–212. Springer, Berlin, 1989.

[13] H. Maurer and H. D. Mittelmann. Optimization techniques for solving elliptic control problems with control and state constraints. I. Boundary control. *Comput. Optim. Appl.*, 16(1):29–55, 2000.

[14] M. Renardy and R. C. Rogers. *An Introduction to Partial Differential Equations*. Springer Verlag, New York – Berlin – Heidelberg, 1993.

[15] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, New York, third edition, 1987.

[16] J. A. Sethian. Fast marching methods. *SIAM Rev.*, 41(2):199–235 (electronic), 1999.

[17] J. A. Sethian. *Level set methods and fast marching methods*. Cambridge University Press, Cambridge, second edition, 1999. Evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science.

[18] J. Sokołowski and J-P. Zolésio. *Introduction to shape optimization*. Springer-Verlag, Berlin, 1992. Shape sensitivity analysis.

[19] L. N. Vicente. On interior-point Newton algorithms for discretized optimal control problems with state constraints. *Optim. Methods Softw.*, 8(3-4):249–275, 1998.

[20] J. Wloka. *Partielle Differentialgleichungen*. Teubner, Stuttgart, 1982.

[21] S. J. Wright. *Primal-dual interior-point methods*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[22] Yinyu Ye. *Interior point algorithms.* John Wiley & Sons Inc., New York, 1997. Theory and analysis, A Wiley-Interscience Publication.

Special Research Center on Optimization and Control, University of Graz, Institute of Mathematics, Heinrichstrasse 36, A-8010 Graz, Austria