

OPTIMAL CONTROL FOR AN ELLIPTIC SYSTEM WITH POLYGONAL STATE CONSTRAINTS*

KARL KUNISCH[†], KEWEI LIANG[‡], AND XILIANG LU[§]

Abstract. This work is devoted to stationary optimal control problems with polygonal constraints on the components of the state. Existence of Lagrange multipliers, of different regularities, is verified for the cases with and without Slater condition holding. For the numerical realization a semismooth Newton method is proposed for an appropriately chosen family of regularized problems. The asymptotic behavior of the regularized problem class is studied, and numerical feasibility of the method is shown.

Key words. optimal control, state constraints, semismooth Newton method, polygonal constraints, Slater condition

AMS subject classifications. 49M15, 65J15

DOI. 10.1137/090758155

1. Introduction. In recent years a significant amount of attention has been paid to open loop optimal control problems governed by partial differential equations with state and/or control constraints; see, for example, [4, 5, 9, 11, 13, 8, 14] and further references cited therein. Most of this work, especially in the context of efficient numerical methods, considered the case of unilateral or bilateral constraints. We focus on polygonal state constraints in this paper. For the model problem, we consider a system of second order elliptic differential equations

$$(1.1) \quad \Lambda \vec{y} = \vec{u} \text{ in } \Omega, \quad \vec{y}|_{\partial\Omega} = \vec{0},$$

with vector-valued state variable \vec{y} and control variable \vec{u} . For simplicity here we consider only stationary problems posed on a bounded domain $\Omega \subset \mathbb{R}^n$ ($n = 2, 3$) with Lipschitz continuous boundary $\partial\Omega$. As cost we take the quadratic functional

$$J(\vec{y}, \vec{u}) = \frac{1}{2} \|\vec{y} - \vec{y}_d\|^2 + \frac{\alpha}{2} \|\vec{u}\|^2,$$

where $\alpha > 0$ and \vec{y}_d are fixed. Consider then the following optimal control problem.

PROBLEM 1.1.

$$\min J(\vec{y}, \vec{u}) \quad \text{such that (1.1) holds and } M\vec{y}(\mathbf{x}) \leq \vec{\psi} \text{ for all } \mathbf{x} \in \Omega.$$

*Received by the editors May 6, 2009; accepted for publication (in revised form) August 3, 2010; published electronically October 19, 2010.

<http://www.siam.org/journals/sicon/48-8/75815.html>

[†]Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria (karl.kunisch@uni-graz.at). This author's work was supported in part by the Austrian Science Foundation under SFB032, "Mathematical Optimization and Applications in Biomedical Sciences."

[‡]Department of Mathematics, Zhejiang University, Hangzhou 310027, People's Republic of China (matlkw@zju.edu.cn).

[§]Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040 Linz, Austria. Current address: Department of Mathematics, Wuhan University, Wuhan 430072, People's Republic of China (xiliang.lu@ricam.oeaw.ac.at).

Here $\vec{\psi} \in \mathbb{R}^m$, M is a matrix of appropriate dimension, and $M\vec{y}(\mathbf{x}) \leq \vec{\psi}$ describes a pointwise polygonal bound that needs to be satisfied by the state \vec{y} of the system (1.1).

To solve Problem 1.1, a Lagrangian approach will be used. It is by now well known that the Lagrange multiplier corresponding to the state constraint is only a measure; see [4, 5]. Therefore a penalized formulation is introduced. Specifically, for $\gamma > 0$, we consider the following family of regularized problems.

PROBLEM 1.2.

$$\min J(\vec{y}, \vec{u}) + \frac{\gamma}{2} \|(M\vec{y} - \vec{\psi})^+\|^2 \quad \text{such that (1.1) holds,}$$

where $(M\vec{y} - \vec{\psi})^+$ is defined coordinatewise: $(M\vec{y} - \vec{\psi})_i^+ = \max(M_i^T \vec{y} - \psi_i, 0)$ for $i = 1, \dots, m$.

Throughout $\|\cdot\|$ denotes the L^2 on Ω of appropriate dimensions and $\|\cdot\|_2$ stand for the norm in H^2 .

There are three instances which involve a dimension concept in this paper. First, there is the spatial domain Ω and we assume that it is of dimension 2 or 3. The dimension of the system, in particular the dimension of \vec{y} , is chosen to be 2. This is mainly for transparency of the proofs, and higher dimensions can be treated by the same techniques. Finally, there is the image space of M , which we assume to be of dimension m . This is the number of constraints on the state \vec{y} . Throughout it will be essential that the regularity and dimension assumptions be such that for the state of the system we have $\vec{y} \in C(\Omega, \mathbb{R}^2)$ for any admissible control \vec{u} .

While the linear-quadratic model problem considered here is rather standard, it should be noted that it arises in the inner loop of the sequential quadratic programming approach to general nonlinear optimal control problems. After an implicit Euler discretization step, parabolic PDE-constrained optimal control problems also fall into this class of model problem.

Let us also compare the problem considered here to coordinatewise unilaterally constrained problems. In this case the matrix M is the identity matrix and, in particular, it is surjective. Here we consider the case where M is not surjective. This requires a different treatment of the Lagrange multipliers, in particular, to argue their uniqueness and their convergence properties in the semismooth Newton method. A very particular case involves bilateral constraints of the form $\vec{\phi} \leq \vec{y} \leq \vec{\psi}$, with $\vec{\phi} \in C(\bar{\Omega}, \mathbb{R}^n)$, $\vec{\psi} \in C(\bar{\Omega}, \mathbb{R}^n)$, $\vec{\phi} < \vec{\psi}$. In this case one can separate the active sets from below, and the active sets from above, and in each coordinate these closed sets have empty intersection. Then the Lagrange multipliers can essentially be introduced as in the unilateral case, having opposite signs on the active sets from above (respectively, from below). The case of optimal control problems related to the Navier–Stokes equations with bilateral state constraints was considered in [7].

2. Preliminaries and problem setting.

2.1. Problem setting. We define the spaces $W = H^2 \cap H_0^1$ and $\vec{W} = (H^2 \cap H_0^1)^2$. Let $C(\bar{\Omega})$ be the space of continuous functions on $\bar{\Omega}$, endowed with the maximum norm, and denote by $\vec{C}_0(\Omega)$ the subspace of $C(\bar{\Omega})$ with vanishing trace on the boundary. It is known that $\vec{W} \hookrightarrow (C_0(\Omega))^2 = C_0(\Omega, \mathbb{R}^2)$ for $1 \leq n \leq 3$.

We assume that for every $\vec{u} \in L^2(\Omega, \mathbb{R}^2)$ system (1.1) has a unique solution $\vec{y} \in \vec{W}$ satisfying

$$(2.1) \quad \|\vec{y}\|_2 \leq C\|\vec{u}\|.$$

We also consider a polygonal domain $D \subset \mathbb{R}^2$, which is defined as the intersection of m half-spaces $M_i^T \vec{y} \leq \psi_i$, $i = 1, \dots, m$. Defining the matrix

$$M : \mathbb{R}^2 \rightarrow \mathbb{R}^m$$

and the vector $\psi \in \mathbb{R}^m$ by

$$(2.2) \quad M = \begin{pmatrix} M_1^T \\ \vdots \\ M_m^T \end{pmatrix}, \quad \vec{\psi} = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_m \end{pmatrix},$$

the polygon can be expressed as $D = \{\vec{y} : M\vec{y} \leq \vec{\psi}\}$; see Figure 2.1.

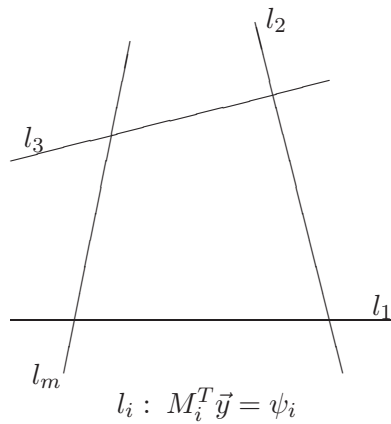


FIG. 2.1. Polygonal constraint.

Remark 2.1. Note that the dimensions of the composite mapping $\mathbf{x} \rightarrow M\vec{y}(\mathbf{x})$ are given by

$$\mathbf{x} \in \Omega \in \mathbb{R}^{2(3)} \rightarrow \vec{y}(\mathbf{x}) \in \mathbb{R}^2 \rightarrow M\vec{y}(\mathbf{x}) \in \mathbb{R}^m.$$

As already remarked, considering $\vec{y}(\mathbf{x}) \in \mathbb{R}^2$ simplifies the explanation, but it is not essential. The results can be extended to $\vec{y}(\mathbf{x}) \in \mathbb{R}^l$, $l > 2$.

2.2. Measure theory. We review some basic results in measure theory. By Riesz's representation theorem, the regular Borel measures on the compact set $\bar{\Omega}$ can be identified with the dual space of $C(\bar{\Omega})$ (see [6]). An analogous result holds for vector-valued regular Borel measures. We shall not distinguish in notation these two concepts; i.e., we use $\lambda(U)$ to denote the measure of a Borel measurable set U , and we use the duality pairing $\langle \lambda, f \rangle_{C^*, C}$ for $f \in C(\bar{\Omega})$ to represent a measure. The positivity of a measure can equivalently be checked by following relation:

$$\lambda \geq 0 \Leftrightarrow \langle \lambda, f \rangle_{C^*, C} \geq 0 \quad \text{for all } f \in C(\bar{\Omega}), \text{ with } f \geq 0.$$

Given $\lambda \in C^*$, the measure can be defined as follows (see [6]). For any open subset $U \in \Omega$,

$$(2.3) \quad \lambda(U) = \sup_{f \in C_0(\Omega), 0 \leq f \leq \chi_U} \langle \lambda, f \rangle_{C^*, C}.$$

Subsequently the property of the regular Borel measure allows us to extend the above calculation to any Borel measurable set A :

$$\lambda(A) = \inf_{A \subset U, U \text{ is open}} \lambda(U).$$

The following relation can be obtained by the above characterization (see [6]). Suppose that f is a continuous function; then

$$(2.4) \quad \begin{aligned} \chi_A \leq f &\Rightarrow \lambda(A) \leq \langle \lambda, f \rangle_{C^*, C} \quad \text{for all measurable sets } A, \\ 0 \leq f \leq \chi_A &\Rightarrow \lambda(A) \geq \langle \lambda, f \rangle_{C^*, C} \quad \text{for all compact sets } A. \end{aligned}$$

3. First order optimality condition. Let K be the set

$$(3.1) \quad K = \{\vec{z} \in C(\bar{\Omega}, \mathbb{R}^m) : \vec{z}(\mathbf{x}) \leq \vec{\psi} \text{ for all } \mathbf{x} \in \bar{\Omega}\},$$

and let I_K be the indicator functional of K :

$$(3.2) \quad I_K(\vec{z}) = \begin{cases} +\infty, & \vec{z} \notin K, \\ 0, & \vec{z} \in K. \end{cases}$$

It is clear that K is a closed convex set in $C(\bar{\Omega}, \mathbb{R}^m)$, and therefore the indicator functional I_K is convex.

We denote the map $\vec{u} \rightarrow \vec{y}(\vec{u})$ by T , where $T : L^2(\Omega, \mathbb{R}^2) \rightarrow C_0(\Omega, \mathbb{R}^2)$. Hence cost functional $J(\vec{y}, \vec{u})$ can be equivalently represented by the reduced functional

$$\hat{J}(\vec{u}) = J(T\vec{u}, \vec{u}),$$

and Problem 1.1 can be rewritten in the following equivalent way:

$$(3.3) \quad \inf_{\vec{u} \in L^2(\Omega, \mathbb{R}^2)} \hat{J}(\vec{u}) + I_K(MT\vec{u}).$$

THEOREM 3.1. *There exists a unique solution for Problem 1.1.*

Proof. We can easily check that $\hat{J} + I_K \circ MT$ is a lower semicontinuous convex functional. Then standard arguments imply the existence of an optimal solution. Strict convexity implies uniqueness of the solution. \square

3.1. Optimality system. Using the notation of subdifferential calculus (see [3]), \vec{u}^* is a solution to (3.3) if and only if

$$0 \in \partial(\hat{J}(\vec{u}^*) + I_K(MT\vec{u}^*)).$$

Assumption 3.2. $\vec{\psi} > 0$.

With Assumption 3.2, holding the Slater condition is satisfied; i.e.,

$$\text{there exists } \vec{u} \text{ such that } MT\vec{u} \in \text{int}(K).$$

In fact, for the choice $\vec{u} = 0$, we have $\vec{y} = T\vec{u} = 0$, and hence $M\vec{y} = 0$ is an interior point of K .

Remark 3.1. The assumption $\vec{\psi} > 0$ ensures that the Slater condition is satisfied. It implies that $\mathbf{0}$ is the interior point of the polygon $D \subset \mathbb{R}^m$. In practice this assumption may not be true; e.g., $\mathbf{0}$ is a vertex of D in our numerical example in section 5. The following convex analysis approach relies on the Slater condition, and

thus it is not applicable to the case where we have only $\vec{\psi} \geq 0$. For this case, however, we can still study the existence of a Lagrange multiplier in a slightly weaker sense, namely, with $\vec{\lambda}^* \in W^*(\Omega, \mathbb{R}^m)$. This will be considered in Proposition 3.4.

If Assumption 3.2 is satisfied, then (see [12])

$$0 \in \partial \hat{J}(\vec{u}^*) + T^* M^T \partial I_k(MT\vec{u}^*).$$

Hence we can find $\vec{\lambda}^* \in \partial I_k(MT\vec{u}^*)$ such that $\vec{\lambda}^* \in C^*(\bar{\Omega}, \mathbb{R}^m)$ and

$$0 \in \partial \hat{J}(\vec{u}^*) + T^* M^T \vec{\lambda}^*.$$

Since $\hat{J}(\vec{u}^*) = \frac{1}{2} \|T\vec{u}^* - \vec{y}_d\|^2 + \frac{\alpha}{2} \|\vec{u}^*\|^2$, we have

$$(3.4) \quad 0 = (T^*(T\vec{u}^* - \vec{y}_d) + \alpha\vec{u}^* + T^* M^T \vec{\lambda}^*, \vec{v}) = (\alpha\vec{u}^* - \vec{p}^*, \vec{v}) \quad \text{for all } \vec{v} \in L^2(\Omega, \mathbb{R}^2),$$

where the adjoint variable \vec{p}^* is given by $\vec{p}^* = -T^*(T\vec{u}^* - \vec{y}_d) - T^* M^T \vec{\lambda}^*$. Hence

$$(3.5) \quad \begin{aligned} (\vec{p}^*, \vec{v}) &= (-T^* M^T \vec{\lambda}^* - T^*(T\vec{u}^* - \vec{y}_d), \vec{v}) \quad \text{for all } \vec{v} \in L^2(\Omega, \mathbb{R}^2), \\ (\vec{p}^*, \Lambda \vec{z}) &= -\langle \vec{\lambda}^*, M\vec{z} \rangle_{C^*, C} - (\vec{y}^* - \vec{y}_d, \vec{z}) \quad \text{for all } \vec{z} \in \vec{W}. \end{aligned}$$

Last, $\vec{\lambda}^* \in \partial I_k(MT\vec{u}^*)$ is equivalent to

$$M\vec{y}^* \in K, \quad \langle \vec{\lambda}^*, \vec{z} - M\vec{y}^* \rangle_{C^*, C} \leq 0 \quad \text{for all } \vec{z} \in K.$$

By a standard argument, this is also equivalent to

$$(3.6) \quad \vec{\lambda}^* \geq 0, \quad M\vec{y}^* \leq \vec{\psi}, \quad \langle \vec{\lambda}^*, M\vec{y}^* - \vec{\psi} \rangle_{C^*, C} = 0.$$

Combining these arguments we have the following theorem.

THEOREM 3.3. *If Assumption 3.2 is satisfied, then there exist an optimal solution \vec{u}^* , \vec{y}^* , an associated dual variable $\vec{p}^* \in L^2(\Omega, \mathbb{R}^2)$, and a Lagrange multiplier $\vec{\lambda}^* \in C^*(\bar{\Omega}, \mathbb{R}^m)$ such that the first order optimality system*

$$(3.7) \quad \begin{cases} \Lambda \vec{y}^* = \vec{u}^*, \\ (\vec{p}^*, \Lambda \vec{z}) + \langle \vec{\lambda}^*, M\vec{z} \rangle_{C^*, C} = (\vec{y}_d - \vec{y}^*, \vec{z}) \quad \text{for all } \vec{z} \in \vec{W}, \\ \alpha \vec{u}^* = \vec{p}^*, \\ \vec{\lambda}^* \geq 0, \quad M\vec{y}^* \leq \vec{\psi}, \quad \langle \vec{\lambda}^*, M\vec{y}^* - \vec{\psi} \rangle_{C^*, C} = 0 \end{cases}$$

holds.

Remark 3.2. The dual variable \vec{p}^* has more regularity than $L^2(\Omega, \mathbb{R}^2)$. The second equation of the optimality system has the form

$$\Lambda^T \vec{p}^* = \vec{y}_d - \vec{y}^* - M^T \vec{\lambda}^*.$$

This is an elliptic equation where the right-hand side is given by a Radon measure. From [15], we have $\vec{p}^* \in W^{1,p}(\Omega, \mathbb{R}^2)$ for any $1 \leq p < 2$, where $\Omega \subset \mathbb{R}^n$, $n = 2$ (or $\vec{p}^* \in W^{1,p}(\Omega, \mathbb{R}^2)$ for any $1 \leq p < \frac{3}{2}$ when $n = 3$). By the Sobolev embedding theorem this implies that $\vec{p}^* \in L^q(\Omega, \mathbb{R}^2)$ for all $q < \infty$ when $n = 2$ (or $\vec{p}^* \in L^q(\Omega, \mathbb{R}^2)$ for all $q < 3$ when $n = 3$). Then for sufficiently smooth $\partial\Omega$, use of standard regularity theory for elliptic equations implies that the regularity of the state can be improved

for \bar{y}^* from $H^2 \cap H_0^1(\Omega, \mathbb{R}^2)$ to $W^{2,q} \cap H_0^1(\Omega, \mathbb{R}^2)$ for any $q < \infty$ when $n = 2$ (or from $H^2 \cap H_0^1(\Omega, \mathbb{R}^2)$ to $W^{2,q} \cap H_0^1(\Omega, \mathbb{R}^2)$ for any $q < 3$ when $n = 3$). Here we use the first and the third equations in (3.7).

PROPOSITION 3.4. *Let $\bar{\psi} \geq 0$. Then there exist an optimal solution \bar{u}^* , \bar{y}^* , an associated dual variable $\bar{p}^* \in L^2(\Omega, \mathbb{R}^2)$, and a Lagrange multiplier $\bar{\lambda}^* \in W^*(\Omega, \mathbb{R}^m)$, which satisfy the first order optimality system*

$$(3.8) \quad \begin{cases} \Lambda \bar{y}^* = \bar{u}^*, \\ (\bar{p}^*, \Lambda \bar{z}) + \langle \bar{\lambda}^*, M \bar{z} \rangle_{W^*, W} = (\bar{y}_d - \bar{y}^*, \bar{z}) \text{ for all } \bar{z} \in \bar{W}, \\ \alpha \bar{u}^* = \bar{p}^*, \\ \bar{\lambda}^* \geq 0, \quad M \bar{y}^* \leq \bar{\psi}, \quad \langle \bar{\lambda}^*, M \bar{y}^* - \bar{\psi} \rangle_{W^*, W} = 0. \end{cases}$$

The proof is given in the appendix. It depends on a regularization technique which is developed in section 4. The assumption $\bar{\psi} > 0$ is not essential in most parts of this paper, including the convergence of the solution to Problem 1.2 and the superlinear convergence of the semismooth method. But we need to assume $\bar{\psi} > 0$ in section 3.2 to ensure the uniqueness of the Lagrange multiplier.

3.2. Uniqueness of optimal solution. From Theorem 3.1, we know that the optimal solution pair (\bar{u}^*, \bar{y}^*) is unique, and hence \bar{p}^* is unique. The main task is to prove the uniqueness of the Lagrange multiplier $\bar{\lambda}^*$. We already know that the Lagrange multiplier $\bar{\lambda}^* \in C^*(\bar{\Omega}, \mathbb{R}^m)$, and we henceforth investigate some of its properties. First, it is clear that the linear complementarity condition (3.6) holds componentwise, i.e.,

$$(3.9) \quad \lambda_i^* \geq 0, \quad M_i^T \bar{y}^* \leq \psi_i, \quad \langle \lambda_i^*, M_i^T \bar{y}^* - \psi_i \rangle_{C^*, C} = 0, \quad i = 1, \dots, m.$$

LEMMA 3.5. *There exists a decomposition Ω_i , $i = 1, \dots, m$, of Ω , which satisfies*

$$\bigcup_i \Omega_i = \Omega, \quad \Omega_i \text{ is open,} \quad \lambda_j^*(\Omega_i) = 0 \text{ for all } j \neq i, i+1 \text{ (where we let } m+1 \text{ be 1)}.$$

Proof. From the definition, the function \bar{y}^* maps from $\bar{\Omega}$ to the polygonal domain D . It is continuous and hence uniformly continuous. More precisely, for any given positive constant δ , there exists another positive constant ϵ such that

$$|\mathbf{x}_1 - \mathbf{x}_2| \leq \epsilon \Rightarrow |\bar{y}^*(\mathbf{x}_1) - \bar{y}^*(\mathbf{x}_2)| \leq \frac{\delta}{2}.$$

We can decompose the polygon D into m parts such that every part D_i^o has at least distance $\delta > 0$ to the constraint whose index is different from $i, i+1$. Then define the subdomains Ω_i^o , Ω_i , and D_i as

$$\begin{aligned} \Omega_i^o &= \{\mathbf{x} : \bar{y}^*(\mathbf{x}) \in D_i^o\}, \\ \Omega_i &= \{\mathbf{x} \in \Omega : \text{dist}(\mathbf{x}, \Omega_i^o) < \epsilon\}, \\ D_i &= \bar{y}^*(\Omega_i). \end{aligned}$$

For this kind of partition, see Figure 3.1.

Clearly Ω_i are open subsets and $\Omega = \cup_i \Omega_i$. By the above definition and from the uniform continuity of \bar{y}^* on $\bar{\Omega}$, we have that every D_i has at least distance $\frac{\delta}{2}$ to the

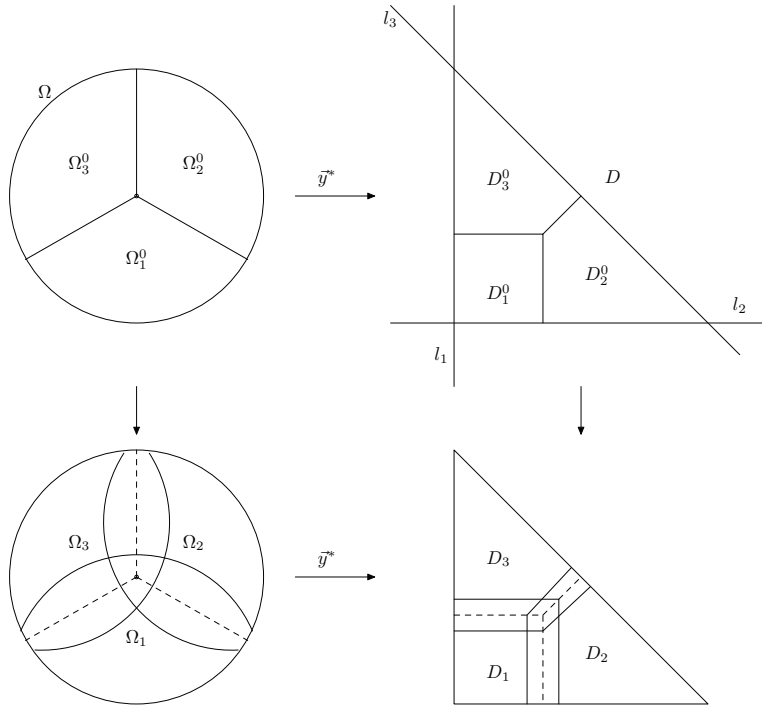


FIG. 3.1. Partition of the domain and the polygonal.

constraint whose index is not $i, i + 1$. In fact, without loss of generality, we consider only the case $i = 1$. Since D_1 has strict positive distance at least $\frac{\delta}{2}$ to the constraints with index $i \neq 1, 2$, we have

$$\max_{i \neq 1, 2} \sup_{\mathbf{x} \in \Omega_1} (M_i \bar{y}^* - \psi_i) \leq -\frac{\delta}{2} < 0.$$

Let $f_i = \psi_i - M_i \bar{y}^*$. Then $f_i \geq \frac{\delta}{2} \chi_{\Omega_1}$ and from (2.4) we have

$$0 \leq \lambda_i^*(\Omega_1) \leq \frac{2}{\delta} \langle \lambda_i^*, f_i \rangle_{C^*, C} = 0 \quad \text{for all } 2 < i \leq m. \quad \square$$

LEMMA 3.6. *There exist an open set $\Omega_0 \subset \Omega$ and a function $f_0 \in C_0(\Omega)$ such that the following hold:*

1. $\text{dist}(\Omega_0, \partial\Omega) > 0$, and $\vec{\lambda}^*(\overline{\Omega} \setminus \Omega_0) = 0$;
2. for all $\vec{\phi} \in C(\overline{\Omega}, \mathbb{R}^m)$, we have

$$\langle \vec{\lambda}^*, \vec{\phi} \rangle_{C^*, C} = \langle \vec{\lambda}^*, f_0 \vec{\phi} \rangle_{C^*, C}.$$

Proof. By Assumption 3.2, and since \bar{y}^* satisfies a homogeneous boundary condition, $M\bar{y}^* - \vec{\psi}$ has strictly negative values on $\partial\Omega$; i.e., there exists a positive constant δ such that

$$M_i \bar{y}^*(\mathbf{x}) - \psi_i \leq -\delta < 0 \quad \text{for all } i = 1, \dots, m, \text{ and } \mathbf{x} \in \partial\Omega.$$

Since $M\bar{y}^* - \vec{\psi}$ is uniformly continuous on $\overline{\Omega}$, there exists a positive constant ϵ such that

$$|\mathbf{x}_1 - \mathbf{x}_2| \leq \epsilon \Rightarrow |g_i(\mathbf{x}_1) - g_i(\mathbf{x}_2)| \leq \frac{\delta}{2} \quad \text{for all } i,$$

where $g_i(\mathbf{x}) = M_i \bar{y}^*(\mathbf{x}) - \psi_i$. We define the open set

$$(3.10) \quad \Omega_0 = \{\mathbf{x} : \mathbf{x} \in \Omega, \text{dist}(\mathbf{x}, \partial\Omega) > \epsilon\}.$$

Uniform continuity of the function $M\bar{y}^* - \vec{\psi}$ implies that

$$M_i \bar{y}^*(\mathbf{x}) - \psi_i \leq -\frac{\delta}{2} < 0 \quad \text{for all } i, \text{ for all } \mathbf{x} \in \bar{\Omega} \setminus \Omega_0.$$

The linear complementarity condition (3.9) implies that $\lambda_i(\bar{\Omega} \setminus \Omega_0) = 0$. We define the function f_0 by

$$f_0(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \Omega_0, \\ \frac{1}{\epsilon} \text{dist}(\mathbf{x}, \partial\Omega), & \mathbf{x} \in \bar{\Omega} \setminus \Omega_0. \end{cases}$$

Clearly f_0 is continuous and for all $\vec{\phi} \in C(\bar{\Omega}, \mathbb{R}^m)$ we have

$$\langle \vec{\lambda}^*, \vec{\phi} \rangle_{C^*, C} = \int_{\bar{\Omega}} \vec{\phi} d\vec{\lambda}^* = \int_{\Omega_0} \vec{\phi} d\vec{\lambda}^* = \int_{\bar{\Omega}} f_0 \vec{\phi} d\vec{\lambda}^* = \langle \vec{\lambda}^*, f_0 \vec{\phi} \rangle_{C^*, C}. \quad \square$$

We are now prepared for the proof of uniqueness of the Lagrange multiplier. Suppose that there exist two measures $\vec{\lambda}^1$ and $\vec{\lambda}^2$ satisfying the optimality system. By Lemma 3.6,

$$\vec{\lambda}^1(\bar{\Omega} \setminus \Omega_0) = \vec{\lambda}^2(\bar{\Omega} \setminus \Omega_0) = \vec{0}.$$

Since $\vec{\lambda}^1$ and $\vec{\lambda}^2$ are regular Borel measures, we need only show that they coincide on any open subset $U \subset \Omega_0$. By the next lemma, the problem can be further reduced.

LEMMA 3.7. *If two measures λ^1 and λ^2 coincide on any open subset $U \subset \Omega_i$, $i = 1, \dots, m$, then $\lambda^1(V) = \lambda^2(V)$ for any open set $V \in \Omega_0$.*

Proof. Let $V_i = V \cap \Omega_i$ for $i = 1, \dots, m$. Since $\Omega_0 \subset \cup_{i=1}^m \Omega_i$, we have $V = \cup_{i=1}^m V_i$. By the inclusion-exclusion principle (see [2]), we find for $i = 1$ and $i = 2$

$$\begin{aligned} \lambda^i(V) &= \lambda^i\left(\bigcup_{j=1}^m V_j\right) = \sum_{j=1}^m \lambda^i(V_j) - \sum_{1 \leq j_1 < j_2 \leq m} \lambda^i(V_{j_1} \cap V_{j_2}) \\ &\quad + \sum_{1 \leq j_1 < j_2 < j_3 \leq m} \lambda^i(V_{j_1} \cap V_{j_2} \cap V_{j_3}) - \dots + (-1)^m \lambda^i\left(\bigcap_{j=1}^m V_j\right). \end{aligned}$$

From the definition of V_j , every term on the right-hand side coincides for $i = 1$ and $i = 2$. This implies that $\lambda^1(V) = \lambda^2(V)$. \square

From Lemma 3.7, it is enough to show for any open subset $U \subset \Omega_i$, $\vec{\lambda}^1(U) = \vec{\lambda}^2(U)$. Without loss of generality, we consider only the case $i = 1$. Lemma 3.5 implies that

$$(3.11) \quad \lambda_i^1(U) = \lambda_i^2(U) = 0 \quad \text{for all } 2 < i \leq m.$$

For any $\vec{z} \in \vec{W}$, by (3.7), we have

$$\langle M^T \vec{\lambda}^1, \vec{z} \rangle_{C^*, C} = \langle M^T \vec{\lambda}^2, \vec{z} \rangle_{C^*, C},$$

which implies that

$$\langle M_{1,2}^T \vec{\lambda}_{1,2}^1, \vec{z} \rangle_{C^*, C} = \langle M_{1,2}^T \vec{\lambda}_{1,2}^2, \vec{z} \rangle_{C^*, C},$$

where the square matrix $M_{1,2}$ and the vectors $\vec{\lambda}_{1,2}^1, \vec{\lambda}_{1,2}^2$ are given by

$$M_{1,2} = \begin{pmatrix} M_1^T \\ M_2^T \end{pmatrix}, \quad \vec{\lambda}_{1,2}^1 = \begin{pmatrix} \lambda_1^1 \\ \lambda_2^1 \end{pmatrix}, \quad \vec{\lambda}_{1,2}^2 = \begin{pmatrix} \lambda_1^2 \\ \lambda_2^2 \end{pmatrix}.$$

Since the line segments 1 and 2 are not parallel (they intersect at vertex 1), the matrix $M_{1,2}$ is invertible. Since the space \vec{W} is dense in $C_0(\Omega, \mathbb{R}^2)$, (2.3) implies that

$$M_{1,2}^T \vec{\lambda}_{1,2}^1(U) = M_{1,2}^T \vec{\lambda}_{1,2}^2(U).$$

Hence $\vec{\lambda}^1(U) = \vec{\lambda}^2(U)$ is obtained by the invertibility of the matrix $M_{1,2}$.

The above argument implies the uniqueness of the corresponding Lagrange multiplier. Moreover, we have following theorem.

THEOREM 3.8. *There exists a unique solution $(\vec{y}^*, \vec{u}^*, \vec{p}^*, \vec{\lambda}^*)$ satisfying the optimality system (3.7).*

Proof. To complete the argument, we need only show the uniqueness for $(\vec{y}^*, \vec{u}^*, \vec{p}^*)$. Let $(\vec{y}^1, \vec{u}^1, \vec{p}^1, \vec{\lambda}^1)$ and $(\vec{y}^2, \vec{u}^2, \vec{p}^2, \vec{\lambda}^2)$ be two solutions which satisfy (3.7). For $i = 1, 2$, the complementarity conditions

$$\vec{\lambda}^i \geq 0, \quad M\vec{y}^i \leq \vec{\psi}, \quad \langle \vec{\lambda}^i, M\vec{y}^i - \vec{\psi} \rangle_{C^*, C} = 0$$

are equivalent to the variational inequalities

$$M\vec{y}^i \leq \vec{\psi}, \quad \langle \vec{\lambda}^i, \vec{z} - M\vec{y}^i \rangle_{C^*, C} \leq 0 \quad \text{for all } \vec{z} \leq \vec{\psi}.$$

Choosing $\vec{z} = M\vec{y}^1$ when $i = 2$ and $\vec{z} = M\vec{y}^2$ when $i = 1$, we have

$$\langle \vec{\lambda}^1, M\vec{y}^2 - M\vec{y}^1 \rangle_{C^*, C} \leq 0, \quad \langle \vec{\lambda}^2, M\vec{y}^1 - M\vec{y}^2 \rangle_{C^*, C} \leq 0.$$

Hence

$$(3.12) \quad \langle \vec{\lambda}^1 - \vec{\lambda}^2, M\vec{y}^1 - M\vec{y}^2 \rangle_{C^*, C} \geq 0.$$

Letting $(\delta\vec{y}, \delta\vec{u}, \delta\vec{p}) = (\vec{y}^1, \vec{u}^1, \vec{p}^1) - (\vec{y}^2, \vec{u}^2, \vec{p}^2)$, we have

$$\begin{cases} \Lambda\delta\vec{y} = \delta\vec{u}, \\ (\delta\vec{p}, \Lambda\vec{z}) + \langle \vec{\lambda}^1 - \vec{\lambda}^2, M\vec{z} \rangle_{C^*, C} = -(\delta\vec{y}, \vec{z}) \quad \text{for all } \vec{z} \in \vec{W}, \\ \alpha\delta\vec{u} = \delta\vec{p}. \end{cases}$$

Multiplying the first equation with $\delta\vec{p}$ and choosing $\vec{z} = \delta\vec{y}$ in the second equation and then subtracting them, we have

$$0 = (\delta\vec{u}, \delta\vec{p}) + \langle \vec{\lambda}^1 - \vec{\lambda}^2, M\delta\vec{y} \rangle_{C^*, C} + \|\delta\vec{y}\|^2.$$

Using $\delta\vec{p} = \alpha\delta\vec{u}$, $\delta\vec{y} = \vec{y}^1 - \vec{y}^2$, and inequality (3.12), we find

$$\|\delta\vec{y}\|^2 + \alpha\|\delta\vec{u}\|^2 \leq 0.$$

This implies that $(\delta\vec{y}, \delta\vec{u}) = 0$, which leads to the uniqueness of $(\vec{y}^*, \vec{u}^*, \vec{p}^*)$. \square

Remark 3.3. Theorem 3.3 gives only first order necessary condition. Together with Theorem 3.8 this optimality system (3.7) also provides a sufficient condition.

4. Semismooth Newton method. The optimal control problem (Problem 1.2) without constraints on the state admits an optimal solution which satisfies the following first order optimality system:

$$(4.1) \quad \begin{cases} \Lambda \vec{y}_\gamma = \vec{u}_\gamma, \\ (\vec{p}_\gamma, \Lambda \vec{z}) + (\vec{\lambda}_\gamma, M \vec{z}) = (\vec{y}_d - \vec{y}_\gamma, \vec{z}) \text{ for all } \vec{z} \in \vec{W}, \\ \alpha \vec{u}_\gamma = \vec{p}_\gamma, \\ \vec{\lambda}_\gamma = \gamma(M \vec{y}_\gamma - \psi)^+. \end{cases}$$

We will use a semismooth Newton method to solve system (4.1). Before, however, we consider the asymptotic behavior as the penalty parameter γ tends to infinity.

4.1. Convergence with respect to γ . In this subsection, we study the convergence when the parameter $\gamma \rightarrow +\infty$.

THEOREM 4.1. *Let $(\vec{u}^*, \vec{y}^*, \vec{p}^*, \vec{\lambda}^*)$ and $(\vec{u}_\gamma, \vec{y}_\gamma, \vec{p}_\gamma, \vec{\lambda}_\gamma)$ satisfy the optimality systems (3.7) and (4.1), respectively. Then we have*

$$\begin{aligned} \vec{y}_\gamma &\rightarrow \vec{y}^* && \text{in } \vec{W}, \\ \vec{p}_\gamma &\rightarrow \vec{p}^* && \text{in } L^2(\Omega, \mathbb{R}^2), \\ \vec{\lambda}_\gamma &\rightarrow \vec{\lambda}^* && \text{in } W^*(\Omega, \mathbb{R}^m). \end{aligned}$$

Proof. Since $\vec{\lambda}_\gamma \geq 0$ and $M \vec{y}_\gamma - \vec{\psi} \leq 0$, it follows that

$$(4.2) \quad (\vec{\lambda}_\gamma, M(\vec{y}_\gamma - \vec{y}^*)) = (\vec{\lambda}_\gamma, M \vec{y}_\gamma - \vec{\psi}) - (\vec{\lambda}_\gamma, M \vec{y}^* - \psi) \geq \frac{1}{\gamma} \|\vec{\lambda}_\gamma\|^2.$$

From (4.1), we deduce that

$$(4.3) \quad \alpha(\Lambda \vec{y}_\gamma, \Lambda(\vec{y}_\gamma - \vec{y}^*)) = (\vec{p}_\gamma, \Lambda(\vec{y}_\gamma - \vec{y}^*)) = (\vec{y}_d - \vec{y}_\gamma, \vec{y}_\gamma - \vec{y}^*) - (\vec{\lambda}_\gamma, M(\vec{y}_\gamma - \vec{y}^*)).$$

Inequality (4.2) and equality (4.3) lead to

$$\alpha \|\Lambda \vec{y}_\gamma\|^2 - \alpha(\Lambda \vec{y}_\gamma, \Lambda \vec{y}^*) \leq -\|\vec{y}_\gamma\|^2 + (\vec{y}_\gamma, \vec{y}^*) + (\vec{y}_d, \vec{y}_\gamma - \vec{y}^*) - \frac{1}{\gamma} \|\vec{\lambda}_\gamma\|^2.$$

Hence

$$\begin{aligned} \alpha \|\Lambda \vec{y}_\gamma\|^2 + \frac{1}{\gamma} \|\vec{\lambda}_\gamma\|^2 + \|\vec{y}_\gamma\|^2 &\leq \alpha(\Lambda \vec{y}_\gamma, \Lambda \vec{y}^*) + (\vec{y}_\gamma, \vec{y}^* + \vec{y}_d) - (\vec{y}_d, \vec{y}^*) \\ &\leq \frac{\alpha}{2} \|\Lambda \vec{y}_\gamma\|^2 + \frac{\alpha}{2} \|\Lambda \vec{y}^*\|^2 + \frac{1}{2} \|\vec{y}_\gamma\|^2 + \frac{1}{2} \|\vec{y}^* + \vec{y}_d\|^2 + \|\vec{y}^*\| \|\vec{y}_d\|. \end{aligned}$$

Therefore, $\alpha \|\Lambda \vec{y}_\gamma\|^2 + \frac{1}{\gamma} \|\vec{\lambda}_\gamma\|^2 \leq C(\|\vec{u}^*\|^2 + \|\vec{y}^*\|^2 + \|\vec{y}_d\|^2)$ for a constant independent of γ . This implies uniform boundedness of $\|\vec{y}_\gamma\|_2^2 + \frac{1}{\gamma} \|\vec{\lambda}_\gamma\|^2$ with respect to $\gamma \geq 1$. Hence $\|\vec{p}_\gamma\| + \|M^T \lambda_\gamma\|_{C^*}$ is also uniformly bounded. After passage to a subsequence,

$$\begin{aligned} \vec{y}_\gamma &\rightharpoonup \hat{y} && \text{in } \vec{W}, \\ \vec{p}_\gamma &\rightharpoonup \hat{p} && \text{in } L^2(\Omega, \mathbb{R}^2). \end{aligned}$$

From the definition $\vec{\lambda}_\gamma = \gamma(M\vec{y}_\gamma - \vec{\psi})^+$, it can be shown that

$$\|(M\vec{y}_\gamma - \vec{\psi})^+\| = \frac{1}{\gamma^2} \|\vec{\lambda}_\gamma\|^2 \rightarrow 0.$$

Since \vec{W} is compactly imbedded into $C(\bar{\Omega}, \mathbb{R}^2)$, we have $\vec{y}_\gamma \rightarrow \hat{y}$ in $C(\bar{\Omega}, \mathbb{R}^2)$ after passage to a subsequence. Then we have $(M\vec{y}_\gamma - \vec{\psi})^+ \rightarrow (M\hat{y} - \vec{\psi})^+$ in $C(\bar{\Omega}, \mathbb{R}^m)$, and hence

$$\|(M\vec{y}_\gamma - \vec{\psi})^+\| \rightarrow \|(M\hat{y} - \vec{\psi})^+\| = 0.$$

This implies $M\hat{y} \leq \vec{\psi}$. Using (3.7), (4.1), and (4.2), we have the following estimate:

$$\begin{aligned} (4.4) \quad \alpha \|\Lambda(\vec{y}_\gamma - \vec{y}^*)\|^2 &= (\vec{p}_\gamma - \vec{p}^*, \Lambda(\vec{y}_\gamma - \vec{y}^*)) \\ &= \langle -\vec{y}_\gamma + \vec{y}^* + M^T(\vec{\lambda}^* - \vec{\lambda}_\gamma), \vec{y}_\gamma - \vec{y}^* \rangle_{C^*, C} \\ &= -\|\vec{y}_\gamma - \vec{y}^*\|^2 + \langle \vec{\lambda}^* - \vec{\lambda}_\gamma, M(\vec{y}_\gamma - \vec{y}^*) \rangle_{C^*, C} \\ &\leq -\|\vec{y}_\gamma - \vec{y}^*\|^2 - \frac{1}{\gamma} \|\vec{\lambda}_\gamma\|^2 + \langle \vec{\lambda}^*, M(\vec{y}_\gamma - \vec{y}^*) \rangle_{C^*, C}. \end{aligned}$$

Hence the complementarity condition (3.6) implies that

$$\alpha \|\Lambda(\vec{y}_\gamma - \vec{y}^*)\|^2 + \|\vec{y}_\gamma - \vec{y}^*\|^2 + \frac{1}{\gamma} \|\vec{\lambda}_\gamma\|^2 \leq \langle \vec{\lambda}^*, M(\vec{y}_\gamma - \vec{y}^*) \rangle_{C^*, C} = \langle \vec{\lambda}^*, M\vec{y}_\gamma - \vec{\psi} \rangle_{C^*, C}.$$

Taking the limit $\gamma \rightarrow \infty$ and using that $M\hat{y} \leq \vec{\psi}$, we have $\hat{y} = \vec{y}^*$ and

$$\begin{aligned} \vec{y}_\gamma &\rightarrow \vec{y}^* \quad \text{in } \vec{W}, \\ \vec{p}_\gamma &\rightarrow \vec{p}^* \quad \text{in } (L^2)^2, \\ \frac{1}{\gamma} \|\vec{\lambda}_\gamma\|^2 &\rightarrow 0. \end{aligned}$$

Now we consider convergence of the Lagrange multipliers. Since

$$M^T(\vec{\lambda}_\gamma - \vec{\lambda}^*) = \Lambda^T(\vec{p}^* - \vec{p}_\gamma) + (\vec{y}^* - \vec{y}_\gamma),$$

the convergence results for \vec{y}_γ and \vec{p}_γ imply that

$$(4.5) \quad M^T \vec{\lambda}_\gamma \rightarrow M^T \vec{\lambda}^* \quad \text{in } \vec{W}^*.$$

Moreover, we also have

$$\begin{aligned} \langle \vec{\lambda}_\gamma, \vec{\psi} \rangle_{C^*, C} &= \langle \vec{\lambda}_\gamma, \vec{\psi} - M\vec{y}_\gamma \rangle_{C^*, C} + \langle \vec{\lambda}_\gamma, M\vec{y}_\gamma \rangle_{C^*, C} \\ &= -\frac{1}{\gamma} \|\vec{\lambda}_\gamma\|^2 + \langle M^T \vec{\lambda}_\gamma, \vec{y}_\gamma \rangle_{C^*, C} \rightarrow \langle M^T \vec{\lambda}^*, \vec{y}^* \rangle_{C^*, C} = \langle \vec{\lambda}^*, \vec{\psi} \rangle_{C^*, C}. \end{aligned}$$

Since M is not a square matrix, the investigation of convergence of the Lagrange multipliers $\vec{\lambda}_\gamma$ requires extra care. Let Ω_i , $i = 0, 1, \dots, m$, be the open sets defined in Lemmas 3.5 and 3.6, and recall that

$$(4.6) \quad \vec{\lambda}_\gamma = \gamma(M\vec{y}_\gamma - \vec{\psi})^+.$$

Our goal is to show that $\vec{\lambda}_\gamma \rightharpoonup \vec{\lambda}^*$ in $W^*(\Omega, \mathbb{R}^m)$, i.e.,

$$(\vec{\lambda}_\gamma, \vec{\phi}) \rightarrow \langle \vec{\lambda}^*, \vec{\phi} \rangle_{W^*, W} \quad \text{for all } \vec{\phi} \in W(\Omega, \mathbb{R}^m).$$

Without loss of generality, we need only show that

$$(4.7) \quad (\lambda_{\gamma,1}, \phi) \rightarrow \langle \lambda_1^*, \phi \rangle_{W^*, W} \quad \text{for all } \phi \in W.$$

By the construction of Ω_0 it follows that $\Omega_0 \Subset \Omega$. Using a partition of unity argument (cf. [1]) there exist smooth functions $\xi_i \in C_0^\infty(\Omega)$ with support in Ω_i , and

$$(4.8) \quad 1 = \sum_{i=1}^m \xi_i(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \Omega_0.$$

We already have $\vec{y}_\gamma \rightarrow \vec{y}^*$ in $C(\bar{\Omega}, \mathbb{R}^2)$, and $M\vec{y}^* - \vec{\psi} \leq -\frac{\delta}{2}$ in $\bar{\Omega} \setminus \Omega_0$. This implies that for γ large enough we have $M\vec{y}_\gamma - \vec{\psi} \leq 0$ on $\bar{\Omega} \setminus \Omega_0$, and hence $\vec{\lambda}_\gamma|_{\bar{\Omega} \setminus \Omega_0} = \gamma(M\vec{y}_\gamma - \vec{\psi})^+|_{\bar{\Omega} \setminus \Omega_0} = 0$. Together with $\vec{\lambda}^*(\bar{\Omega} \setminus \Omega_0) = 0$ this implies that

$$\begin{aligned} (\lambda_{\gamma,1}, \phi) &= (\lambda_{\gamma,1}, \phi)_{\Omega_0} = \sum_i (\lambda_{\gamma,1}, \phi \xi_i)_{\Omega_0} = \sum_i (\lambda_{\gamma,1}, \phi \xi_i), \\ \langle \lambda_1^*, \phi \rangle_{W^*, W} &= \langle \lambda_1^*, \phi \rangle_{C^*, C} = \int_{\Omega_0} \phi d\lambda_1^* \\ &= \sum_i \int_{\Omega_0} \phi \xi_i d\lambda_1^* = \sum_i \langle \lambda_1^*, \phi \xi_i \rangle_{C^*, C} = \sum_i \langle \lambda_1^*, \phi \xi_i \rangle_{W^*, W}. \end{aligned}$$

Hence (4.7) follows by applying Lemma 4.2. \square

LEMMA 4.2. *Let $\phi \in H^2 \cap H_0^1$, with support in Ω_i , $i = 1, \dots, m$. Then we have $(\lambda_{\gamma,1}, \phi) \rightarrow \langle \lambda_1^*, \phi \rangle_{W^*, W}$.*

Proof. If $i = 2, \dots, m - 1$, then by the argument below (4.8), we find $\lambda_{\gamma,1}|_{\Omega_i} = 0$ for γ large enough. Hence we have $(\lambda_{\gamma,1}, \phi) = 0 = \langle \lambda_1^*, \phi \rangle_{W^*, W}$ for $\text{supp } \phi \subset \Omega_i$, $i = 2, \dots, m - 1$. Next let $i = 1$. The case $i = m$ can be treated analogously. Since $\vec{y}_\gamma \rightarrow \vec{y}^*$ in $C(\bar{\Omega}, \mathbb{R}^2)$ and $(M\vec{y}^* - \vec{\psi})_j \leq -\frac{\delta}{2}$ in Ω_1 for all $2 < j \leq m$, we have $\lambda_{\gamma,j}|_{\Omega_1} = 0$ for all $2 < j \leq m$ and sufficiently large γ . This implies that for any $\vec{f} \in \vec{W}$ which vanishes outside of Ω_1

$$(M^T \vec{\lambda}_\gamma, \vec{f}) = (\vec{\lambda}_\gamma, M\vec{f}) = (\vec{\lambda}_\gamma, M\vec{f})_{\Omega_1} = (\lambda_{\gamma,1}, M_1^T \vec{f}) + (\lambda_{\gamma,2}, M_2^T \vec{f}),$$

where we recall the notation introduced in (2.2).

Applying Lemma 3.5, we find

$$\langle M^T \vec{\lambda}^*, \vec{f} \rangle_{W^*, W} = \langle \lambda_1^*, M_1^T \vec{f} \rangle_{W^*, W} + \langle \lambda_2^*, M_2^T \vec{f} \rangle_{W^*, W}.$$

Using the notation $M_{1,2}$ as in section 3.2 and choosing $\vec{f} = \phi M_{1,2}^{-1} \mathbf{e}$ (where $\mathbf{e} = (1, 0)^T$), we deduce from the above two equalities that

$$(M^T \vec{\lambda}_\gamma, \vec{f}) = (\lambda_{\gamma,1}, \phi), \quad \langle M^T \vec{\lambda}^*, \vec{f} \rangle_{W^*, W} = \langle \lambda_1^*, \phi \rangle_{W^*, W}.$$

Therefore (4.5) implies that for ϕ with support in Ω_1

$$(\lambda_{\gamma,1}, \phi) \rightarrow \langle \lambda_1^*, \phi \rangle_{W^*, W}. \quad \square$$

4.2. Algorithm. For convenience we recall the optimality system (4.1) for Problem 1.2:

$$(4.9) \quad \begin{cases} \alpha \Lambda \vec{y}_\gamma = \vec{p}_\gamma, \\ \Lambda^T \vec{p}_\gamma + M^T \vec{\lambda}_\gamma + \vec{y}_\gamma = \vec{y}_d, \\ \vec{\lambda}_\gamma = \gamma(M \vec{y}_\gamma - \psi)^+, \end{cases}$$

where $\gamma > 0$ is fixed. The semismooth Newton algorithm (or primal-dual active set algorithm) to solve the optimality system (4.9) is given in Algorithm 1 (see [10, 11]).

ALGORITHM 1. PRIMAL-DUAL ACTIVE SET ALGORITHM

- 1: Set $k = 0$, initialize \vec{y}^0 ;
- 2: let $A_i^k = \{\mathbf{x} : M_i \vec{y}^k > \psi_i\}$, and set

$$M^k = \begin{pmatrix} M_1^T \chi_{A_1^k} \\ \vdots \\ M_m^T \chi_{A_m^k} \end{pmatrix}, \quad \vec{\psi}^k = \begin{pmatrix} \psi_1 \chi_{A_1^k} \\ \vdots \\ \psi_m \chi_{A_m^k} \end{pmatrix};$$

- 3: solve for $(\vec{y}^{k+1}, \vec{p}^{k+1})$

$$(4.10) \quad \begin{cases} \alpha \Lambda \vec{y}^{k+1} = \vec{p}^{k+1}, \\ \Lambda^T \vec{p}^{k+1} + \gamma M^T (M^k \vec{y}^{k+1} - \vec{\psi}^k) = \vec{y}_d - \vec{y}^{k+1}; \end{cases}$$

- 4: define $\vec{\lambda}^{k+1} = \gamma(M^k \vec{y}^{k+1} - \vec{\psi}^k)$;
 - 5: stop or update $k = k + 1$, and go to 2.
-

PROPOSITION 4.3. *There exists a unique solution $(\vec{y}^{k+1}, \vec{p}^{k+1})$ for system (4.10).*

Proof. Since, for fixed γ , (4.10) is an elliptic system, a standard argument implies existence and uniqueness of the solution. \square

PROPOSITION 4.4. *If $A_i^k = A_i^{k+1}$ for $i = 1, \dots, m$, then $(\vec{y}^{k+1}, \vec{p}^{k+1}, \vec{\lambda}^{k+1})$ solves system (4.9).*

Proof. The proof is based on the following key observation: if $A_i^k = A_i^{k+1}$, then $M^k = M^{k+1}$ and $M^k \vec{y}^{k+1} - \vec{\psi}^k = (M \vec{y}^{k+1} - \vec{\psi})^+$. \square

By the above proposition, it is reasonable to choose $A_i^k = A_i^{k+1}$ for $i = 1, \dots, m$ as the stopping criterion. Numerical results show that this stopping criterion is typically reached.

4.3. Superlinear convergence of the semismooth Newton method.

THEOREM 4.5. *Let $\vec{y}^0 \in L^p(\Omega, \mathbb{R}^2)$ for some $p > 2$, and let $\|\vec{y}^0 - \vec{y}_\gamma\|_{L^p}$ be sufficiently small. Then $(\vec{y}^k, \vec{p}^k, \vec{\lambda}^k)$ converges to $(\vec{y}_\gamma, \vec{p}_\gamma, \vec{\lambda}_\gamma)$ superlinearly.*

Proof. We define the error between the $(k + 1)$ st iteration and the solution to the γ problem by

$$\delta \vec{y} = \vec{y}^{k+1} - \vec{y}_\gamma, \quad \delta \vec{p} = \vec{p}^{k+1} - \vec{p}_\gamma, \quad \delta \vec{\lambda} = \vec{\lambda}^{k+1} - \vec{\lambda}_\gamma.$$

These expressions satisfy

$$(4.11) \quad \begin{cases} \alpha \Lambda \delta \vec{y} = \delta \vec{p}, \\ \Lambda^T \delta \vec{p} + M^T \delta \vec{\lambda} = -\delta \vec{y}, \\ \frac{1}{\gamma} \delta \vec{\lambda} = (M^k \vec{y}^{k+1} - \vec{\psi}^k) - (M \vec{y}_\gamma - \tilde{\psi})^+. \end{cases}$$

Since

$$(M^k \vec{y}^{k+1} - \vec{\psi}^k) - (M \vec{y}_\gamma - \tilde{\psi})^+ = M^k (\vec{y}^{k+1} - \vec{y}^k) + (M \vec{y}^k - \psi)^+ - (M \vec{y}_\gamma - \psi)^+,$$

we find for $i = 1$ ($i = 2, \dots, m$ can be treated similarly)

$$\begin{aligned} \frac{1}{\gamma} \delta \lambda_1 &= M_1^T (\vec{y}^{k+1} - \vec{y}^k) \chi_{A_1^k} + \max(0, M_1^T \vec{y}^k - \psi_1) - \max(0, M_1^T \vec{y}_\gamma - \psi_1) \\ &= (M_1^T \delta \vec{y} + M_1^T \vec{y}_\gamma - M_1^T \vec{y}^k) \chi_{A_1^k} + \max(0, M_1^T (\vec{y}^k - \psi_1)) - \max(0, M_1^T \vec{y}_\gamma - \psi_1). \end{aligned}$$

For $g = M_1^T \vec{y}_\gamma - \psi_1$, $h = M_1^T \vec{y}^k - M_1^T \vec{y}_\gamma$, we have $g + h = M_1^T \vec{y}^k - \psi_1$. It is well established [10] that the mapping $\phi \rightarrow \max(0, \phi)$ is Newton differentiable from L^p to L^2 with $p > 2$, with the Newton derivative given by the characteristic function $\chi_{\{x: \phi(x) > 0\}}$. As a consequence we have (see, e.g., [10])

$$(4.12) \quad \|\max(0, g + h) - \max(0, g) - h \chi_{A_1^k}\| = o(\|h\|_{L^p}).$$

Let $R_1^k = \max(0, g + h) - \max(0, g) - h \chi_{A_1^k}$, and R_2^k, \dots, R_m^k can be defined similarly. Hence

$$\delta \vec{\lambda} = \gamma M^k \delta \vec{y} + R^k,$$

where $R^k = (R_1^k, \dots, R_m^k)^T$. Then we multiply $\Lambda \delta \vec{y}$ in the first equation of (4.11):

$$0 = \alpha \|\Lambda \delta \vec{y}\|^2 - (\delta \vec{p}, \Lambda \delta \vec{y}) = \alpha \|\Lambda \delta \vec{y}\|^2 + \|\delta \vec{y}\|^2 + (M \delta \vec{y}, \delta \vec{\lambda}).$$

Since $(M \delta \vec{y}, M^k \delta \vec{y}) = \|M^k \delta \vec{y}\|^2$, we have

$$\|\delta \vec{y}\|_2 = o(\|\vec{y}^k - \vec{y}_\gamma\|_{L^p}),$$

and hence the desired estimate follows:

$$\|\delta \vec{y}\|_2 + \|\delta \vec{\lambda}\| + \|\delta \vec{p}\|_2 = o(\|\vec{y}^k - \vec{y}_\gamma\|_{L^p}). \quad \square$$

5. Numerical results. Here we present a numerical example by utilizing a finite difference discretization to the following elliptic system in the unit square with the homogeneous Dirichlet boundary conditions:

$$(5.1) \quad -\Delta y_1 + y_2 = u_1,$$

$$(5.2) \quad -\Delta y_2 - y_1 = u_2.$$

Define the elliptic operator $\Lambda = -\Delta I + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. It can be shown that Λ satisfies $\|\vec{y}\|_2 \leq C \|\vec{u}\|$.

The state variable \vec{y} has to satisfy the constraint

$$(5.3) \quad y_1 \geq 0, \quad y_2 \geq 0, \quad y_1 + y_2 \leq 1.$$

It can be equivalently represented in matrix form as $M\vec{y} \leq \vec{\psi}$, where

$$M = \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \end{pmatrix}, \quad \vec{\psi} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Recall that the cost functional J is

$$J(\vec{y}, \vec{u}) = \frac{1}{2} \|\vec{y} - \vec{y}_d\|^2 + \frac{\alpha}{2} \|\vec{u}\|^2.$$

Let $\vec{y}_d = (4 \sin(4\pi xy), 2 \sin(2\pi x) + 2 \cos(2\pi y))^T$, $\gamma = 10^4$, $\alpha = 0.001$, $h = \frac{1}{64}$. For this choice of y_d all three constraints are active on some part of the domain. We refer to Figures 5.1 and 5.2 for the optimal state and control. The active sets associated with the three different constraints are depicted in Figure 5.3. We note that the active set can be a “slim set” (as for the first constraint) or a set which clearly has an open interior (as for the second constraint). The corresponding Lagrange multipliers are depicted in Figure 5.4, and as expected they are oscillatory near the boundary of the active set.

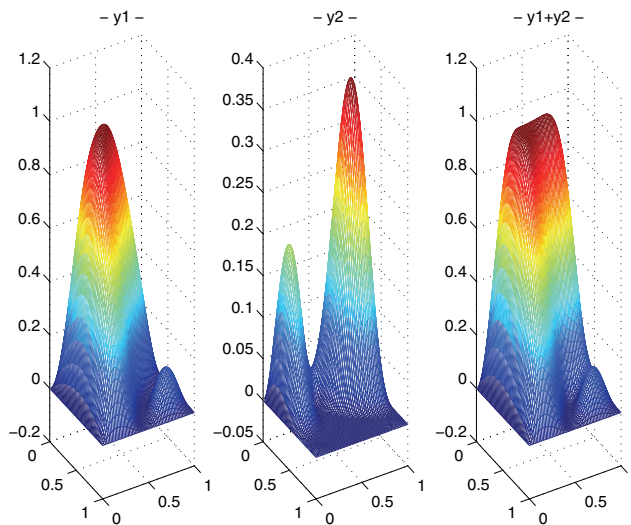


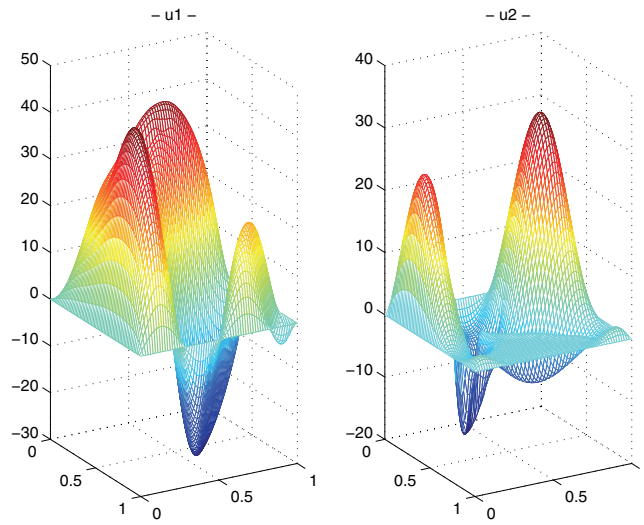
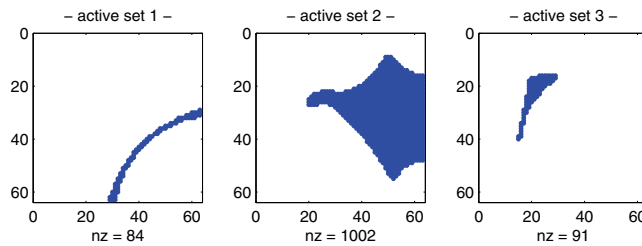
FIG. 5.1. *Optimal state.*

Superlinear convergence can be observed numerically. In fact, denoting by \vec{y}_h^* the solution to the discretized problem, we compute the ratios

$$r_k = \frac{\|\vec{y}_h^{k+1} - \vec{y}_h^*\|}{\|\vec{y}_h^k - \vec{y}_h^*\|}$$

for a fixed $\gamma = 1000$ and $\alpha = 0.01$. The result for iterations $k = 11, \dots, 17$ is given in Table 5.1.

The above test problem has the same diffusion parameter in the underlying equations, and the constraint set is a triangle with two equally long sides. We then

FIG. 5.2. *Optimal control.*FIG. 5.3. *Active set.*

change the governing equation and the constraints to be

$$(5.4) \quad -\Delta y_1 + y_2 = u_1,$$

$$(5.5) \quad -\beta \Delta y_2 - y_1 = u_2,$$

and

$$(5.6) \quad y_1 \geq 0, \quad y_2 \geq 0, \quad \theta y_1 + y_2 \leq 1.$$

As before we set $\vec{y}_d = (4 \sin(4\pi xy), 2 \sin(2\pi x) + 2 \cos(2\pi y))^T$ and fix $\gamma = 10^6$, $\alpha = 0.001$, $h = \frac{1}{64}$. We test the performance of the algorithm under changes of θ and β . It can be noted from Tables 5.2 and 5.3 that it converges for a wide range of values for θ and β , and that the number of iterations before convergence is achieved does not depend on these parameters significantly. As expected, taking into consideration the shape of y_d , the number of components as well as the associated active sets increase with θ .

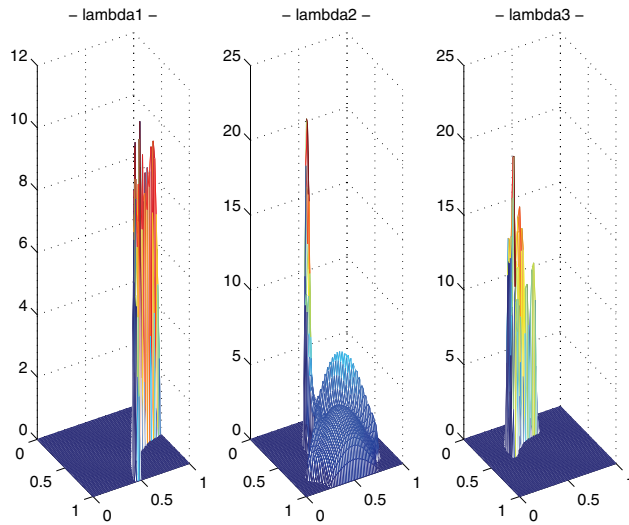


FIG. 5.4. Lagrange multiplier.

TABLE 5.1
Superlinear convergence.

k	11	12	13	14	15	16	17
r_k	0.9449	0.8910	0.7685	0.6405	0.4670	0.0118	0.0

TABLE 5.2
 $\gamma = 10^6, \alpha = 0.001, \beta = 1.$

θ	0.1	0.5	1	10
Iteration number	36	23	15	18
Active components	1	2	3	3

TABLE 5.3
 $\gamma = 10^6, \alpha = 0.001, \theta = 1.$

β	0.01	1	100
Iteration number	23	15	17
Active components	3	3	3

Appendix. *Proof of Proposition 3.4.* Let $(\vec{y}^*, \vec{u}^*) \in \vec{W} \times L^2(\Omega, \mathbb{R}^2)$ denote the unique optimal solution. Since the Slater condition is not assumed, convex analysis is not directly applicable. To prove the existence of a Lagrange multiplier, we consider

the convergence of the solutions to the γ problems. Let $(\vec{y}_\gamma, \vec{u}_\gamma, \vec{p}_\gamma, \vec{\lambda}_\gamma)$ satisfy

$$(5.7) \quad \begin{cases} \Lambda \vec{y}_\gamma = \vec{u}_\gamma, \\ (\vec{p}_\gamma, \Lambda \vec{z}) + (\vec{\lambda}_\gamma, M \vec{z}) = (\vec{y}_d - \vec{y}_\gamma, \vec{z}) \text{ for all } \vec{z} \in \vec{W}, \\ \alpha \vec{u}_\gamma = \vec{p}_\gamma, \\ \vec{\lambda}_\gamma = \gamma(M \vec{y}_\gamma - \psi)^+. \end{cases}$$

By arguments similar to those in section 4.1, we obtain

$$\begin{aligned} \vec{y}_\gamma &\rightharpoonup \hat{y} \quad \text{in } \vec{W}, \\ \vec{u}_\gamma &\rightharpoonup \hat{u} \quad \text{in } L^2(\Omega, \mathbb{R}^2), \end{aligned}$$

and $M\hat{y} \leq \vec{\psi}$. For $J_\gamma(\vec{y}, \vec{u}) = J(\vec{y}, \vec{u}) + \frac{\gamma}{2}\|(M\vec{y} - \psi)^+\|^2$, we have

$$J_\gamma(\vec{y}_\gamma, \vec{u}_\gamma) \leq J_\gamma(\vec{y}^*, \vec{u}^*) = J(\vec{y}^*, \vec{u}^*).$$

The weak lower semicontinuity of J implies that

$$J(\hat{y}, \hat{u}) \leq \liminf J(\vec{y}_\gamma, \vec{u}_\gamma) \leq \liminf J_\gamma(\vec{y}_\gamma, \vec{u}_\gamma) \leq J(\vec{y}^*, \vec{u}^*).$$

Since \hat{y} satisfies the state constraints and (\hat{y}, \hat{u}) satisfies $\Lambda \hat{y} = \hat{u}$, uniqueness of the optimal solution (\vec{y}^*, \vec{u}^*) implies that $(\vec{y}^*, \vec{u}^*) = (\hat{y}, \hat{u})$. From the second equation in system (5.7), we deduce that $\{\|M^T \vec{\lambda}_\gamma\|_{\vec{W}^*}\}_{\gamma \geq 1}$ is bounded. To prove that $\{\vec{\lambda}_\gamma\}_{\gamma \geq 1}$ is bounded in space $(W^*)^m$, we need to guarantee that for some constant $C > 0$ independent of $\gamma \geq 1$

$$\langle \vec{\lambda}_\gamma, \vec{\phi} \rangle_{W, W^*} \leq C \|\vec{\phi}\|_2 \quad \text{for all } \vec{\phi} \in W(\Omega, \mathbb{R}^m).$$

We will follow the notation and the arguments in section 4.1 and note that the construction of Ω_i for $i = 1, \dots, m$ does not use Assumption 3.2. Since we use a homogeneous Dirichlet boundary condition, \vec{y}^* maps from $\partial\Omega$ to $\vec{0}$. Without loss of generality, we assume that $\vec{0}$ has strictly positive distance to D_i for $i = 2, \dots, m$ (see Figure 5.5). Hence a neighborhood of $\partial\Omega$ is in Ω_1 . Therefore there exists a compact subset $\Omega_c \Subset \Omega$ such that $\bar{\Omega} \setminus \Omega_c \subset \Omega_1$. Since Ω_c is covered by open sets Ω_i , by partition of unity (cf. [1]), there are smooth functions $\xi_i \in C_0^\infty(\Omega)$ with support in Ω_i , and

$$1 = \sum_{i=1}^m \xi_i(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \Omega_c.$$

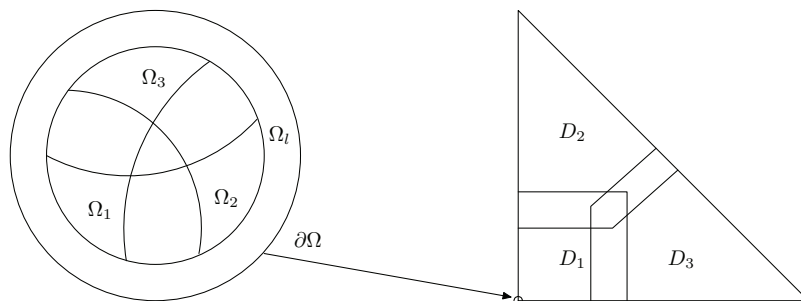


FIG. 5.5. Partition with a homogeneous Dirichlet boundary condition.

Define $\xi_0 = 1 - \sum_{i=1}^m \xi_i$, which has support in Ω_1 . Then any fixed function $\phi \in H^2 \cap H_0^1(\Omega)$ can be expressed as

$$\phi = \phi \left(\xi_0 + \sum_{i=1}^m \xi_i \right).$$

Let $\phi_i = \phi \xi_i$ for $i = 0, 1, \dots, m$. Clearly ϕ_i has support in Ω_i for $i = 1, \dots, m$, and ϕ_0 has support in Ω_1 . Moreover, the domain Ω_i depends on the optimal solution \bar{y}^* but not on γ . Therefore there exists a constant \bar{C} which does not depend on γ such that

$$(5.8) \quad \|\phi_i\|_2 \leq \bar{C} \|\phi\|_2, \quad i = 0, 1, \dots, m.$$

Using the same technique as in Lemma 4.2 and noticing the uniform boundedness of $\|M^T \bar{\lambda}_\gamma\|_{\bar{W}^*}$, we obtain for $\bar{\varphi} \in W^m$ and with support in Ω_i

$$\langle \bar{\lambda}_\gamma, \bar{\varphi} \rangle_{W, W^*} \leq C_i \|\bar{\varphi}\|_2,$$

where C_i depends only on domain Ω_i and the bound on $\{\|M^T \bar{\lambda}_\gamma\|_{\bar{W}^*}\}_{\gamma \geq 1}$. Combining this with estimate (5.8), we have

$$\langle \bar{\lambda}_\gamma, \bar{\phi} \rangle_{W, W^*} = \sum_{i=0}^m \langle \bar{\lambda}_\gamma, \bar{\phi}_i \rangle_{W, W^*} \leq \sum_i C_i \|\phi_i\|_2 \leq (m+1) \bar{C} \max_i \{C_i\} \|\phi\|_2.$$

Hence $\{\bar{\lambda}_\gamma\}_{\gamma \geq 1}$ is uniformly bounded in $(W^*)^m$. Then there exist $\bar{\lambda}^* \in W^*(\Omega, \mathbb{R}^m)$ and a subsequence, still denoted by $\bar{\lambda}_\gamma$, such that

$$\bar{\lambda}_\gamma \rightharpoonup \bar{\lambda}^* \quad \text{in } W^*(\Omega, \mathbb{R}^m).$$

This implies the existence of a Lagrange multiplier of Problem 1.1. In fact, to obtain the optimality system, we need to define the dual variable $\bar{p}^* = \alpha \bar{u}^*$. By passing to the limit in the first three equations of system (5.7), $(\bar{y}^*, \bar{u}^*, \bar{p}^*, \bar{\lambda}^*)$ satisfies

$$\begin{aligned} \Lambda \bar{y}^* &= \bar{u}^*, \\ (\bar{p}^*, \Lambda \bar{z}) + (\bar{\lambda}^*, M \bar{z}) &= (\bar{y}_d - \bar{y}^*, \bar{z}) \quad \text{for all } \bar{z} \in \bar{W}, \\ \alpha \bar{u}^* &= \bar{p}^*. \end{aligned}$$

We need to check the linear complementarity condition to close the system. Positivity of $\bar{\lambda}^*$ follows from weak convergence of λ_γ . Utilizing the estimate (4.4), we find

$$\alpha \|\Lambda(\bar{y}_\gamma - \bar{y}^*)\|^2 \leq -\|\bar{y}_\gamma - \bar{y}^*\|^2 - \frac{1}{\gamma} \|\bar{\lambda}_\gamma\|^2 + \langle \bar{\lambda}^*, M(\bar{y}_\gamma - \bar{y}^*) \rangle_{W^*, W}.$$

Since $\bar{y}_\gamma \rightharpoonup \bar{y}^*$ in \bar{W} , and $M^T \bar{\lambda}^* \in \bar{W}^*$, we obtain

$$\frac{1}{\gamma} \|\bar{\lambda}_\gamma\|^2 \rightarrow 0, \quad \bar{y}_\gamma \rightarrow \bar{y}^* \quad \text{in } \bar{W}.$$

Using that $\bar{\lambda}_\gamma = \gamma(M\bar{y}_\gamma - \bar{\psi})^+$, we have

$$(\bar{\lambda}_\gamma, M\bar{y}_\gamma - \bar{\psi}) = \frac{1}{\gamma} \|\bar{\lambda}_\gamma\|^2.$$

Strong convergence $\vec{y}_\gamma \rightarrow \vec{y}^*$ in \vec{W} and weak convergence $M^T \vec{\lambda}_\gamma \rightharpoonup M^T \vec{\lambda}^*$ in \vec{W}^* imply that

$$\langle \vec{\lambda}_\gamma, M\vec{y}_\gamma - \vec{\psi} \rangle \rightarrow \langle \vec{\lambda}^*, M\vec{y}^* - \vec{\psi} \rangle_{W^*, W}.$$

Combining the above estimates leads to the linear complementarity condition

$$\vec{\lambda}^* \geq 0, \quad M\vec{y}^* \leq \vec{\psi}, \quad \langle \vec{\lambda}^*, M\vec{y}^* - \vec{\psi} \rangle_{W^*, W} = 0.$$

This completes the proof. \square

REFERENCES

- [1] R. ADAMS AND J. FOURNIER, *Sobolev Spaces*, Elsevier/Academic Press, Amsterdam, 2003.
- [2] M. AIGNER, *A Course in Enumeration*, Springer, Berlin, 2007.
- [3] V. BARBU AND T. PRECUPANU, *Convexity and Optimization in Banach Spaces*, D. Reidel, Dordrecht, The Netherlands, 1986.
- [4] M. BERGOUNIOUX AND K. KUNISCH, *On the structure of Lagrange multipliers for state-constrained optimal control problems*, Systems Control Lett., 48 (2002), pp. 169–176.
- [5] E. CASAS, *Control of an elliptic problem with pointwise state constraints*, SIAM J. Control Optim., 24 (1986), pp. 1309–1318.
- [6] D. COHN, *Measure Theory*, Birkhäuser, Boston, 1980.
- [7] J. C. DE LOS REYES AND K. KUNISCH, *A semi-smooth Newton method for regularized state-constrained optimal control of the Navier-Stokes equations*, Computing, 78 (2006), pp. 287–309.
- [8] J. C. DE LOS REYES, C. MEYER, AND B. VEXLER, *Finite element error analysis for state-constrained optimal control of the Stokes equations*, Control Cybernet., 37 (2008), pp. 251–284.
- [9] K. DECKELNICK, A. GÜNTHER, AND M. HINZE, *Finite element approximation of elliptic control problems with constraints on the gradient*, Numer. Math., 111 (2009), pp. 335–350.
- [10] M. HINTERMÜLLER, K. ITO, AND K. KUNISCH, *The primal-dual active set strategy as a semi-smooth Newton method*, SIAM J. Optim., 13 (2003), pp. 865–888.
- [11] K. ITO AND K. KUNISCH, *Semi-smooth Newton methods for state-constrained optimal control problems*, Systems Control Lett., 50 (2003), pp. 221–228.
- [12] P. NEITTAANMAKI, J. SPREKELS, AND D. TIBA, *Optimization of Elliptic Systems: Theory and Applications*, Springer, New York, 2006.
- [13] U. PRÜFERT, F. TRÖLTZSCH, AND M. WEISER, *The convergence of an interior point method for an elliptic control problem with mixed control-state constraints*, Comput. Optim. Appl., 39 (2008), pp. 183–218.
- [14] A. RÖSCH AND F. TRÖLTZSCH, *On regularity of solutions and Lagrange multipliers of optimal control problems for semilinear equations with mixed pointwise control-state constraints*, SIAM J. Control Optim., 46 (2007), pp. 1098–1115.
- [15] G. STAMPACCHIA, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble), 15 (1965), pp. 189–258.