An International Journal
computers & mathematics with applications

PERGAMON Computers and Mathematics with Applications 40 (2000) 177-192
www.elsevier.nl/locate/camwa

A Domain Decomposition Method for Optimization Problems for Partial Differential Equations

M. D. GUNZBURGER* AND JEEHYUN LEE†
Department of Mathematics, Iowa State University
Ames, IA 50011-2064, U.S.A.

(Received and accepted January 2000)

Abstract—A nonoverlapping domain decomposition method for optimization problems for partial differential equations is presented. The domain decomposition is effected through an auxiliary optimization problem. This results in an multiobjective optimization problem involving the given functional and the auxiliary functional. The existence of an optimal solution to the multiobjective optimization problem is proved as are convergence estimates as the parameters used to regularize the problem (penalty parameters) and to combine the two objective functionals tend to zero. An optimality system for the optimal solution is derived and used to define a gradient method. Convergence results are obtained for the gradient method and the results of some numerical experiments are obtained. Then, unregularized problems having vanishing penalty parameters are discussed.

Keywords—Optimization problems for partial differential equations, Nonoverlapping domain decomposition methods, Multiobjective optimization.

1. INTRODUCTION

We introduce a nonoverlapping domain decomposition method for the solution of optimization problems having partial differential equation constraints. The given optimization problem is turned into a multiobjective optimization problem through the use of an optimization based domain decomposition algorithm.

The given optimization problem is defined with the usual ingredients:

- state variables;
- design or control variables that are available to effect the optimization;
- an objective functional to be minimized which may depend on both the state and control variables; and
- constraints in the form of partial differential equations relating the state and control variables.

*Supported in part by the National Science Foundation under Grant No. DMS-9806358.
†Supported in part by the Air Force Office of Scientific Research under Grant No. F49620-95-1-0407.

PII: S0898-1221(00)00152-8

The optimization problem to be solved is then to minimize the given objective functional over a suitable class of control variables, subject to the partial differential equation constraints being satisfied.

The partial differential equation constraints are solved by a domain decomposition algorithm. Domain decomposition methods are characterized by two features which in fact serve to define specific methods. First, one must define the subdomain problems. The only aspect of their definition that is not obvious is the coupling conditions between solutions in the different subdomains. It is through these coupling conditions that one ensures that the solution of the subdomain problems are indeed solutions of the original problem. Here, we define the coupling conditions through an optimization based strategy. (The idea of using optimization based domain decompositions has been discussed in [1–7]. The specific approach we follow here is related to that of [2,3,5,7] and especially [6].) The second feature that characterizes domain decomposition methods is the iteration strategy used to update the data of the subdomain problems in terms of the solutions in other subdomains. Here, we examine a gradient method based update strategy.

Thus, to solve the partial differential equation constraints by a domain decomposition algorithm, we set up an auxiliary optimization problem having auxiliary functional and auxiliary control variables. As a result, we have a multiobjective optimization problem to solve: first, we have to minimize the objective functional of the given optimization problem and second, we have to minimze the auxiliary functional introduced to effect the domain decomposition. We solve the multiobjective optimization problem by forming the weighted sum of the two functionals. This enables us to pose an optimization problem involving a single functional whose minimization simultaneously minimizes, in some sense, the given functional and forces the solution of the subdomain problems to satisfy the given partial differential equation constraints.

Our discussion is based on the simple model problem of Poisson's equation; this is largely done for the sake of keeping the exposition simple. Most of what we say extends in an obvious manner, especially from an algorithmic viewpoint, to more complex and realistic problems. In fact, one of the virtues of the domain decomposition strategy we use is that it extends in a straightforward way to nonlinear problems; see, e.g., [7]. Also, the functional we use to effect the domain decomposition is but one example of the many that can be chosen for this purpose; in fact, it is not necessarily the best one to use, but it is the simplest to explain. Details about the use of other functionals can be found in [2,3].

The plan of the rest of the paper is as follows. In Section 2, we discuss optimization of regularized functionals. The existence of an optimal solution to the multiobjective optimization problem is proved; also, we derive convergence estimates as the parameters used to regularize the problem (penalty parameters) and to combine the two objective functionals tend to zero. An optimality system for the optimal solution is derived and used to define a gradient method for its solution. Convergence results are obtained for the gradient method. Then, in Section 3, unregularized problems having vanishing penalty parameters are discussed. Finally, the results of some numerical experiments are discussed in Section 4.

2. OPTIMIZATION OF REGULARIZED FUNCTIONALS

2.1. The Model Problems

Let Ω be a bounded open set in $\mathbb{R}^2$ with boundary Γ. Let u and f satisfy

$$\begin{aligned} -\Delta u &= f, && \text{in } \Omega, \\ u &= 0, && \text{on } \Gamma. \end{aligned} \tag{1}$$

Given U, let

$$\mathcal{J}_\gamma(u,f) = \frac{1}{2}\int_\Omega (u-U)^2\,d\Omega + \frac{\gamma}{2}\int_\Omega f^2\,d\Omega, \tag{2}$$

where γ is a constant that can be chosen to change the relative importance of two terms in (2). The model problem we consider is as follows.

PROBLEM 1. Minimize $\mathcal{J}_\gamma(u(f), f)$ over suitable functions f subject to (1). Later, we will make clear what we mean by "suitable functions."

We intend to solve Problem 1 by a domain decomposition technique. To this end, let Ω be partitioned into two open subdomains Ω_1 and Ω_2 such that $\overline{\Omega} = \overline{\Omega}_1 \cup \overline{\Omega}_2$. The interface between the two domains is denoted by Γ_0 so that $\Gamma_0 = \overline{\Omega}_1 \cap \overline{\Omega}_2$. Let $\Gamma_1 = \overline{\Omega}_1 \cap \Gamma$ and $\Gamma_2 = \overline{\Omega}_2 \cap \Gamma$. See the sketch in Figure 1.

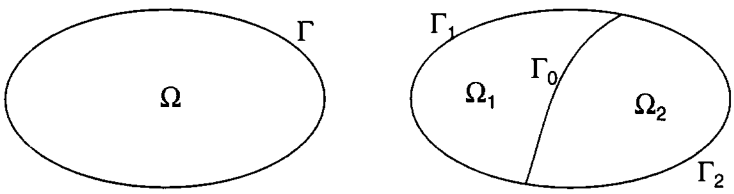

Figure 1. A subdivision of the domain Ω into two subdomains Ω_1 and Ω_2.

Instead of constraint (1), we consider the problems defined over the subdomains: for $i = 1$ and 2,

$$\begin{aligned} -\Delta u_i &= f_i, && \text{in } \Omega_i, \\ u_i &= 0, && \text{on } \Gamma_i, \\ \frac{\partial u_i}{\partial n_i} &= (-1)^{i+1} g, && \text{on } \Gamma_0, \end{aligned} \tag{3}$$

where, for $i = 1$ or 2, $\frac{\partial}{\partial n_i}$ denotes the derivative in the direction of the outer normal to Ω_i. In this paper, we refer to g as the *control*. We also rewrite $\mathcal{J}_\gamma(\cdot,\cdot)$ in terms of the two subdomains in the following form:

$$\mathcal{K}_\gamma(u_1, u_2, f_1, f_2) = \frac{1}{2}\sum_{i=1}^{2}\int_{\Omega_i}(u_i - U)^2\,d\Omega + \frac{\gamma}{2}\sum_{i=1}^{2}\int_{\Omega_i} f_i^2\,d\Omega, \tag{4}$$

where we view u_1 and u_2 as depending on f_1, f_2, and g through (3). Instead of Problem 1, we pose the following problem.

PROBLEM 2. Given g, minimize $\mathcal{K}_\gamma(u_1(f_1, f_2, g), u_2(f_1, f_2, g), f_1, f_2)$ over suitable functions f_1 and f_2 subject to (3).

For an arbitrary choice of the control g, the solutions u_1 and u_2 of Problem 2 are not the same as the solution u of Problem 1 in the respective subdomains, i.e., $u_1 \neq u|_{\Omega_1}$ and $u_2 \neq u|_{\Omega_2}$. This discrepancy is due to the fact that, for an arbitrary choice of g, we have that $u_1 \neq u_2$ along Γ_0, even in a weak sense. On the other hand, there exists a choice of g, i.e., $g = \frac{\partial u}{\partial n_1}|_{\Gamma_0} = -\frac{\partial u}{\partial n_2}|_{\Gamma_0}$, such that the solutions of Problem 2 coincide with the solution of Problem 1 on the corresponding subdomains. Thus, if we are to solve Problems 1 and 2, we must also find the "right" g so that u_1 is as close as possible to u_2 along the interface Γ_0. One way to accomplish this is to minimize the functional

$$\mathcal{G}(u_1, u_2) = \int_{\Gamma_0}(u_1 - u_2)^2\,d\Gamma. \tag{5}$$

Clearly, for given f_1 and f_2, there exists a minimizer of $\mathcal{G}(\cdot,\cdot)$ such that (3) is satisfied.

Instead of (5), we can also consider the penalized or regularized functional

$$\mathcal{G}_\delta(u_1, u_2, g) = \frac{1}{2}\int_{\Gamma_0}(u_1 - u_2)^2 d\Gamma + \frac{\delta}{2}\int_{\Gamma_0} g^2\,d\Gamma, \tag{6}$$

where δ is a constant that can be chosen to change the relative importance of the two terms in (6).

Thus, we are faced with the multiobjective optimization problem of minimizing the functionals $\mathcal{K}_\gamma(\cdot,\cdot,\cdot,\cdot)$ and $\mathcal{G}_\delta(\cdot,\cdot,\cdot)$ over suitable functions f_1, f_2, and g, subject to (3). We combine the two functionals of (4) and (6) into the functional

$$\begin{aligned}\mathcal{E}_{\sigma,\delta,\gamma}(u_1,u_2,f_1,f_2,g) &= \frac{1}{2}\sum_{i=1}^{2}\int_{\Omega_i}(u_i-U)^2\,d\Omega + \frac{\gamma}{2}\sum_{i=1}^{2}\int_{\Omega_i} f_i^2\,d\Omega \\ &+\frac{\delta}{2}\sum_{i=1}^{2}\int_{\Gamma_0} g^2\,d\Gamma + \frac{1}{2\sigma}\int_{\Gamma_0}(u_1-u_2)^2\,d\Gamma,\end{aligned} \tag{7}$$

where σ is a constant that can be chosen to change the relative importance of the contributions of the two functionals $\mathcal{K}_\gamma$ and $\mathcal{G}_\delta$. (We have redefined δ by effecting the replacement δ/σ by δ.) Then, instead of solving Problems 1 or 2, we solve the following problem.

PROBLEM 3. Minimize $\mathcal{E}_{\sigma,\delta,\gamma}(u_1(f_1,f_2,g),u_2(f_1,f_2,g),f_1,f_2,g)$ over suitable functions f_1, f_2, and g subject to (3).

In going from Problem 1 to Problem 3, we introduced the two parameters δ and σ. For any $\delta > 0$ and $\sigma > 0$, the minimizers of $\mathcal{E}_{\sigma,\delta,\gamma}(\cdot,\cdot,\cdot,\cdot,\cdot)$ do not coincide with the minimizers of $\mathcal{K}_\gamma(\cdot,\cdot,\cdot,\cdot)$. We will study this discrepancy in Section 2.3 where we show convergence of the former to the latter as $\delta \to 0$ and $\sigma \to 0$.

The outline of the rest of this section is as follows. In Section 2.2, we give precise definitions of the optimization problems and prove that optimal solutions exist. In Section 2.3, we show the convergence of optimal solutions as $\delta \to 0$ and $\sigma \to 0$. In Section 2.4, we derive an optimality system of equations from which optimal solutions may be determined. In Section 2.5, we study a gradient method for the solutions of the optimality system. Throughout this section, we consider the penalty parameter γ in the functional of (2), (4), and (7) to be a fixed, positive constant. We will consider the case of $\gamma = 0$ in Section 3.

2.2. The Existence of an Optimal Solution

For $s \ge 0$, we denote by $H^s(\Omega)$ the Sobolev space of functions having s square integrable derivatives with respect to Ω. The norm on $H^s(\Omega)$ is denoted by $\|\cdot\|_s$. We have that $H^0(\Omega) = L^2(\Omega)$. We will make use of the subspace

$$H_0^1(\Omega) = \left\{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma\right\}.$$

A weak formulation corresponding to (1) is given by: seek $u \in H_0^1(\Omega)$ such that

$$a(u,v) = (f,v)_\Omega, \qquad \forall\, v \in H_0^1(\Omega), \tag{8}$$

where

$$a(u,v) = \int_\Omega \nabla u\cdot\nabla v\,d\Omega \qquad\text{and}\qquad (f,v)_\Omega = \int_\Omega f v\,d\Omega.$$

It is well known that the bilinear form $a(\cdot,\cdot)$ is coercive and continuous. Then, the existence of a unique solution of (8) follows from the Lax-Milgram Theorem. Also, that theorem yields the continuous dependence on data, i.e., there exist a constant $C > 0$ such that

$$\|u\|_{1,\Omega} \le C\|f\|_{0,\Omega}. \tag{9}$$

Next, we give a precise definition of an optimal solution, i.e., a minimizer of $\mathcal{J}_\gamma(u,f)$. Let the *admissibility set* be defined by

$$\mathcal{U}_{\rm ad} = \left\{(u,f) \in H_0^1(\Omega)\times L^2(\Omega) \text{ such that (8) is satisfied and } \mathcal{J}_\gamma(u,f) < \infty\right\}. \tag{10}$$

Then, (u^*, f^*) is called an *optimal solution* if there exists $\epsilon > 0$ such that

$$\mathcal{J}_\gamma(u^*, f^*) \leq \mathcal{J}_\gamma(u, f), \tag{11}$$

for all $(u, f) \in \mathcal{U}_{\text{ad}}$ satisfying

$$\|u - u^*\|_1 + \|f - f^*\|_0 \leq \epsilon. \tag{12}$$

The existence and uniqueness of optimal solutions is easily proven using standard arguments.

THEOREM 1. *There exists a unique optimal solution $(u^*, f^*) \in \mathcal{U}_{\text{ad}}$ for Problem 1.*

PROOF. Clearly, $\mathcal{U}_{\text{ad}}$ is not empty. Let $\{u^{(n)}, f^{(n)}\}$ be a minimizing sequence in $\mathcal{U}_{\text{ad}}$, i.e.,

$$\lim_{n\to\infty} \mathcal{J}_\gamma\left(u^{(n)}, f^{(n)}\right) = \inf_{(u,f)\in\mathcal{U}_{\text{ad}}} \mathcal{J}_\gamma(u, f).$$

Then, from (10), we have that the sequence $\{f^{(n)}\}$ is uniformly bounded and hence, from (9), so is $\{u^{(n)}\}$. Thus, there exists a subsequence $\{u^{(n_i)}, f^{(n_i)}\}$ such that

$$\begin{aligned} u^{(n_i)} &\rightharpoonup u^*, && \text{in } H_0^1(\Omega), \\ f^{(n_i)} &\rightharpoonup f^*, && \text{in } L^2(\Omega), \\ u^{(n_i)} &\rightarrow u^*, && \text{in } L^2(\Omega), \end{aligned}$$

for some $(u^*, f^*) \in \mathcal{U}_{\text{ad}}$. Now, by the process of passing to the limit, we have that (u^*, f^*) satisfies (8). Then, the fact that the functional $\mathcal{J}_\gamma(\cdot,\cdot)$ is lower semicontinuous implies that

$$\inf_{(u,f)\in\mathcal{U}_{\text{ad}}} \mathcal{J}_\gamma(u, f) = \lim_{i\to\infty} \inf \mathcal{J}_\gamma\left(u^{(n_i)}, f^{(n_i)}\right) \geq \mathcal{J}_\gamma\left(u^*, f^*\right).$$

Hence, $\mathcal{J}_\gamma(u^*, f^*) = \inf_{(u,f)\in\mathcal{U}_{\text{ad}}} \mathcal{J}_\gamma(u, f)$ so that (u^*, f^*) is an optimal solution. Uniqueness follows from the convexity of the functional and the admissibility set and the linearity of the constraints. ∎

We will make use of an auxiliary problem involving the functional

$$\mathcal{J}_{\gamma,\delta}(u, f, g) = \mathcal{J}_\gamma(u, f) + \frac{\delta}{2}\int_{\Gamma_0} g^2 \, d\Gamma, \tag{13}$$

where $\frac{\partial u}{\partial n_1} = g$ on Γ_0. The auxiliary problem is defined as follows.

PROBLEM 1.5. Minimize $\mathcal{J}_{\gamma,\delta}(u, f, g)$ over suitable functions f subject to (1).

Now, we examine the existence of an optimal solution that minimizes $\mathcal{J}_{\gamma,\delta}(u, f, g)$. The admissibility set is now defined by

$$\mathcal{V}_{\text{ad}} = \Big\{(u, f, g) \in H_0^1(\Omega) \times L^2(\Omega) \times L^2(\Gamma_0) \text{ such that (8) is satisfied and } \mathcal{J}_{\gamma,\delta}(u, f, g) < \infty \text{ and } g = \frac{\partial u}{\partial n_1} \text{ along } \Gamma_0\Big\}. \tag{14}$$

THEOREM 2. *There exists a unique optimal solution $(u^\delta, f^\delta, g^\delta) \in \mathcal{V}_{\text{ad}}$ for Problem 1.5.*

PROOF. The proof follows along the lines of the proof of Theorem 1. ∎

Similarly, we examine the existence of an optimal solution that minimizes the functional $\mathcal{E}_{\sigma,\delta,\gamma}(u_1, u_2, f_1, f_2, g)$. Assume u_1, u_2 satisfy (3). Then, a weak formulation corresponding to (3) is given by

$$a_1(u_1, v) = (f, v)_{\Omega_1} + (g, v)_{\Gamma_0}, \qquad \forall\, v \in H^1_{\Gamma_1}(\Omega_1), \tag{15}$$

$$a_2(u_2, v) = (f, v)_{\Omega_2} - (g, v)_{\Gamma_0}, \qquad \forall\, v \in H^1_{\Gamma_2}(\Omega_2). \tag{16}$$

where, for $i = 1, 2$,

$$a_i(u, v) = \int_{\Omega_i} \nabla u \cdot \nabla v \, d\Omega.$$

The existence of a unique solution of (15) and (16) follows from the Lax-Milgram Theorem. Also, that theorem yields the continuous dependence on data, i.e., there exist constant $C > 0$ such that

$$\|u_i\|_{1,\Omega_i} \le C(\|f\|_{0,\Omega_i} + \|g\|_{0,\Gamma_0}). \tag{17}$$

Let the admissibility set be defined by

$$\begin{aligned} \mathcal{W}_{\text{ad}} = \{ &(u_1, u_2, f_1, f_2, g) \in H^1_{\Gamma_1}(\Omega_1) \times H^1_{\Gamma_2}(\Omega_2) \times L^2(\Omega_1) \times L^2(\Omega_2) \\ &\times L^2(\Gamma_0) \text{ such that (15) and (16) are satisfied and } \mathcal{E}_{\sigma,\delta,\gamma}(u_1, u_2, f_1, f_2, g) < \infty \}. \end{aligned} \tag{18}$$

THEOREM 3. *There exists a unique optimal solution* $(u_1^{\sigma,\delta,\gamma}, u_2^{\sigma,\delta,\gamma}, f_1^{\sigma,\delta,\gamma}, f_2^{\sigma,\delta,\gamma}, g^{\sigma,\delta,\gamma}) \in \mathcal{U}_{\text{ad}}$ *of Problem 3.*

PROOF. Again, the proof follows along the lines of the proof of Theorem 1. ∎

2.3. Convergence with Vanishing Parameter

Functional (13) contains the penalty parameter δ that controls the relative importance of the two terms. Clearly, for finite values of δ, solutions of Problem 1.5 will not be the same as that for Problem 1. In the next theorem, we show that, as $\delta \to 0$, optimal solutions of Problem 1.5 converge to the solution of Problem 1.

THEOREM 4. *For each* $\delta > 0$, *let* $(u^\delta, f^\delta, g^\delta)$ *denote the optimal solution satisfying of Problem 1.5. Let* (u^*, f^*) *denote the optimal solution solution of Problem 1. Then,* $\|u^\delta - u^*\|_{1,\Omega} \to 0$ *and* $\|f^\delta - f^*\|_{0,\Omega} \to 0$ *as* $\delta \to 0$.

PROOF. Let $g^* = \frac{\partial u^*}{\partial n_1}|_{\Gamma_0}$. Suppose $\{(u^\delta, f^\delta, g^\delta)\}$ is a sequence of optimal solutions and $\delta \to 0$. Then, we have that

$$\mathcal{J}_{\gamma,\delta}\left(u^\delta, f^\delta, g^\delta\right) \le \mathcal{J}_{\gamma,\delta}(u^*, f^*, g^*), \qquad \forall\, \delta,$$

i.e.,

$$\begin{aligned} &\frac{1}{2}\int_\Omega \left(u^\delta - U\right)^2 d\Omega + \frac{\gamma}{2}\int_\Omega \left(f^\delta\right)^2 d\Omega + \frac{\delta}{2}\int_{\Gamma_0} \left(g^\delta\right)^2 d\Gamma \\ &\le \frac{1}{2}\int_\Omega (u^* - U)^2 \, d\Omega + \frac{\gamma}{2}\int_\Omega (f^*)^2 \, d\Omega + \frac{\delta}{2}\int_{\Gamma_0} (g^*)^2 \, d\Gamma \\ &= \mathcal{J}_\gamma(u^*, f^*) + \frac{\delta}{2}\int_{\Gamma_0} (g^*)^2 \, d\Gamma, \qquad \forall\, \delta. \end{aligned}$$

Then, $\|f^\delta\|$ and $\|g^\delta\|$ are uniformly bounded. So is $\|u^\delta\|$ by (9). Hence, as $\delta \to 0$, there exists a subsequence which converges to some $(\hat{u}, \hat{f}, \hat{g}) \in H_0^1(\Omega) \times L^2(\Omega) \times L^2(\Gamma_0)$. By passing to the limit, we have

$$a(\hat{u}, v) = \left(\hat{f}, v\right)_\Omega, \qquad \forall\, v \in H_0^1(\Omega).$$

Then,

$$\mathcal{J}_{\gamma,\delta}\left(\hat{u}, \hat{f}, \hat{g}\right) \le \mathcal{J}_\gamma(u^*, f^*) + \frac{\delta}{2}\int_{\Gamma_0} (g^*)^2 \, d\Gamma, \qquad \forall\, \delta,$$

i.e.,

$$\mathcal{J}_\gamma\left(\hat{u}, \hat{f}\right) + \frac{\delta}{2}\int_{\Gamma_0} (\hat{g})^2 \, d\Gamma \le \mathcal{J}_\gamma(u^*, f^*) + \frac{\delta}{2}\int_{\Gamma_0} (g^*)^2 \, d\Gamma, \qquad \forall\, \delta.$$

We obtain $\mathcal{J}_\gamma(\hat{u}, \hat{f}) \le \mathcal{J}(u^*, f^*)$ as $\delta \to 0$, and thus, $\hat{u} = u^*$ and $\hat{f} = f^*$. ∎

Functional (7) also contains the penalty parameter σ. Thus, we next show that, as $\sigma \to 0$, solutions of Problem 3 converge to the solution of the auxiliary problem, Problem 1.5.

THEOREM 5. *For each $\sigma > 0$ and $\delta > 0$, let $(u_1^{\sigma,\delta,\gamma}, u_2^{\sigma,\delta,\gamma}, f_1^{\sigma,\delta,\gamma}, f_2^{\sigma,\delta,\gamma}, g^{\sigma,\delta,\gamma})$ denote an optimal solution satisfying Problem 3. Let $(u^\delta, f^\delta, g^\delta)$ denote the solution of Problem 1.5. Then, for $i = 1, 2$, $\|u_i^{\sigma,\delta,\gamma} - u^\delta|_{\Omega_i}\|_{1,\Omega_i} \to 0$ as $\sigma \to 0$.*

PROOF. Let $u_i^\delta = u^\delta|_{\Omega_i}$ for $i = 1, 2$. Suppose $(u_1^{\sigma,\delta,\gamma}, u_2^{\sigma,\delta,\gamma}, f_1^{\sigma,\delta,\gamma}, f_2^{\sigma,\delta,\gamma}, g^{\sigma,\delta,\gamma})$ is a sequence of optimal solutions of Problem 3 and $\sigma \to 0$. Then, we have that

$$\mathcal{E}_{\sigma,\delta,\gamma}\left(u_1^{\sigma,\delta,\gamma}, u_2^{\sigma,\delta,\gamma}, f_1^{\sigma,\delta,\gamma}, f_2^{\sigma,\delta,\gamma}, g^{\sigma,\delta,\gamma}\right) \le \mathcal{E}_{\sigma,\delta,\gamma}\left(u_1^\delta, u_2^\delta, f_1^\delta, f_2^\delta, g^\delta\right) = \mathcal{J}_{\gamma,\delta}\left(u^\delta, f^\delta, g^\delta\right), \qquad \forall\,\sigma,$$

so that

$$\frac{1}{2}\sum_{i=1}^{2}\int_{\Omega_i}\left(u_i^{\sigma,\delta,\gamma} - U\right)^2 d\Omega + \frac{\gamma}{2}\sum_{i=1}^{2}\int_{\Omega_i}\left(f_i^{\sigma,\delta,\gamma}\right)^2 d\Omega$$
$$+\frac{\delta}{2}\sum_{i=1}^{2}\int_{\Gamma_0}\left(g^{\sigma,\delta,\gamma}\right)^2 d\Gamma \le \mathcal{J}_{\gamma,\delta}\left(u^\delta, f^\delta, g^\delta\right),$$

$$\frac{1}{2\sigma}\int_{\Gamma_0}\left(u_1^\sigma - u_2^\sigma\right)^2 d\Gamma \le \mathcal{J}_{\gamma,\delta}\left(u^\delta, f^\delta, g^\delta\right).$$

Then, $\|f_i^{\sigma,\delta,\gamma}\|$ for $i = 1, 2$ and $\|g^{\sigma,\delta,\gamma}\|$ are uniformly bounded, and, by (17), so are $\|u_i^{\sigma,\delta,\gamma}\|$ for $i = 1, 2$. Hence, as $\sigma \to 0$, there exists a subsequence which converges to some $(\hat{u}_1, \hat{u}_2, \hat{f}_1, \hat{f}_2, \hat{g}) \in H^1_{\Gamma_1}(\Omega_1) \times H^1_{\Gamma_2}(\Omega_2) \times L^2(\Omega_1) \times L^2(\Omega_2) \times L^2(\Gamma_0)$. By passing to the limit, we have

$$a_1(\hat{u}_1, v) = \left(\hat{f}, v\right)_{\Omega_1} + (\hat{g}, v)_{\Gamma_0}, \qquad \forall\, v \in H^1_{\Gamma_1}(\Omega_1),$$
$$a_2(\hat{u}_2, v) = \left(\hat{f}, v\right)_{\Omega_2} - (\hat{g}, v)_{\Gamma_0}, \qquad \forall\, v \in H^1_{\Gamma_2}(\Omega_2).$$

We have $\hat{u}_1 = \hat{u}_2$ on Γ_0 a.e. since $\int_{\Gamma_0}(u_1^\sigma - u_2^\sigma)^2 \to 0$ as $\sigma \to 0$ so that

$$\mathcal{J}_{\gamma,\delta}\left(\hat{u}, \hat{f}\right) = \mathcal{E}_{\sigma,\delta,\gamma}\left(\hat{u}_1, \hat{u}_2, \hat{f}_1, \hat{f}_2, \hat{g}\right) \le \mathcal{J}_{\gamma,\delta}\left(u^\delta, f^\delta, g^\delta\right).$$

We thus obtain $\hat{u}_i = u_i^\delta$ for $i = 1, 2$. ∎

2.4. The Optimality System

We use the Lagrange multiplier rule to derive an optimality system of equations for the solution of Problem 3. For the linear, positive definite problem we are considering, the applicability of the Lagrange multiplier rule is easily shown; see, e.g., [8]. For simplicity of notation, throughout this section, we denote the functional and optimal solution of Problem 3 by $\mathcal{E}$ and (u_1, u_2, f_1, f_2, g), respectively, suppressing the the explicit indication of the dependence on the parameters σ, δ, and γ.

We define the Lagrangian

$$\mathcal{L}(u_1, u_2, f_1, f_2, g, \lambda_1, \lambda_2) = \mathcal{E}(u_1, u_2, f_1, f_2, g)$$
$$- \sum_{i=1}^{2}\int_{\Omega_i}\nabla u_i \cdot \nabla\lambda_i \, d\Omega + \sum_{i=1}^{2}\int_{\Omega_i} f_i\lambda_i \, d\Omega + \int_{\Gamma_0} g\lambda_1 \, d\Gamma - \int_{\Gamma_0} g\lambda_2 \, d\Gamma.$$

Next, we apply the necessary conditions for finding stationary points of $\mathcal{L}$. Setting to zero the first variations with respect to λ_1 and λ_2 yields the constraints

$$a_1(u_1, v) = (f, v)_{\Omega_1} + (g, v)_{\Gamma_0}, \qquad \forall\, v \in H^1_{\Gamma_1}(\Omega_1), \tag{19}$$
$$a_2(u_2, v) = (f, v)_{\Omega_2} - (g, v)_{\Gamma_0}, \qquad \forall\, v \in H^1_{\Gamma_2}(\Omega_2). \tag{20}$$

Setting to zero the first variations with respect to u_1 and u_2 yields the adjoint equations

$$a_1(\xi,\lambda_1) = (u_1 - U,\xi)_{\Omega_1} + \frac{1}{\sigma}(u_1 - u_2,\xi)_{\Gamma_0}, \qquad \forall \xi \in H^1_{\Gamma_1}(\Omega_1), \tag{21}$$

$$a_2(\xi,\lambda_2) = (u_2 - U,\xi)_{\Omega_2} - \frac{1}{\sigma}(u_1 - u_2,\xi)_{\Gamma_0}, \qquad \forall \xi \in H^1_{\Gamma_2}(\Omega_2). \tag{22}$$

Finally, setting to zero the first variations with respect to g, f_1, and f_2 yields the optimality conditions

$$(g,r)_{\Gamma_0} = -\frac{1}{\delta}(r,\lambda_1 - \lambda_2)_{\Gamma_0}, \qquad \forall r \in L^2(\Gamma_0), \tag{23}$$

$$(f,w)_{\Omega_1} = -\frac{1}{\gamma}(w,\lambda_1)_{\Omega_1}, \qquad \forall w \in L^2(\Omega_1), \tag{24}$$

$$(f,w)_{\Omega_2} = -\frac{1}{\gamma}(w,\lambda_2)_{\Omega_2}, \qquad \forall w \in L^2(\Omega_2). \tag{25}$$

To summarize, solutions of Problem 3 may be determined by solving the optimality system (19)–(25). Note that this system is coupled, i.e., the constraint equations for the state variables depend on the unknown controls, the adjoint equations for the Lagrange multipliers depend on the state, and optimality conditions for the controls depend on the Lagrange multipliers.

The optimality system (19)–(25) may also be derived in a direct manner using *sensitivity derivatives* instead of the Lagrange multiplier rule. Problem 3 is equivalent to the problem of determining f_1, f_2, and g such that $\mathcal{E}$ is minimized. The first derivatives $\frac{\partial\mathcal{E}}{\partial f_1}, \frac{\partial\mathcal{E}}{\partial f_2}, \frac{\partial\mathcal{E}}{\partial g}$ of $\mathcal{E}$ are defined through their actions on variations $\tilde{f}_1$, $\tilde{f}_2$, and $\tilde{g}$ as follows:

$$\left\langle \frac{\partial\mathcal{E}}{\partial g},\tilde{g}\right\rangle = \sum_{i=1}^{2}(u_i - U,\tilde{u}_i)_{\Omega_i} + \frac{1}{\sigma}(u_1 - u_2,\tilde{u}_1 - \tilde{u}_2)_{\Gamma_0} + \delta(g,\tilde{g})_{\Gamma_0}, \tag{26}$$

where $\tilde{u}_1 \in H^1_{\Gamma_1}(\Omega_1)$ and $\tilde{u}_2 \in H^1_{\Gamma_2}(\Omega_2)$ are the solutions of

$$a_1(\tilde{u}_1,v) = (\tilde{g},v)_{\Gamma_0}, \qquad \forall v \in H^1_{\Gamma_1}(\Omega_1), \tag{27}$$

$$a_2(\tilde{u}_2,v) = -(\tilde{g},v)_{\Gamma_0}, \qquad \forall v \in H^1_{\Gamma_2}(\Omega_2), \tag{28}$$

$$\left\langle \frac{\partial\mathcal{E}}{\partial f_1},\tilde{f}_1\right\rangle = (u_1 - U,\check{u}_1)_{\Omega_1} + \frac{1}{\sigma}(u_1 - u_2,\check{u}_1)_{\Gamma_0} + \gamma\left(f_1,\tilde{f}_1\right)_{\Omega_1},$$

where $\check{u}_1 \in H^1_{\Gamma_1}(\Omega_1)$ is the solution of

$$a_1(\check{u}_1,v) = \left(\tilde{f}_1,v\right)_{\Omega_1}, \qquad \forall v \in H^1_{\Gamma_1}(\Omega_1)$$

and

$$\left\langle \frac{\partial\mathcal{E}}{\partial f_2},\tilde{f}_2\right\rangle = (u_2 - U,\check{u}_2)_{\Omega_2} + \frac{1}{\sigma}(u_1 - u_2,\check{u}_2)_{\Gamma_0} + \gamma\left(f_2,\tilde{f}_2\right)_{\Omega_2},$$

where $\check{u}_2 \in H^1_{\Gamma_2}(\Omega_2)$ is the solution of

$$a_2(\check{u}_2,v) = \left(\tilde{f}_2,v\right)_{\Omega_2}, \qquad \forall v \in H^1_{\Gamma_2}(\Omega_2).$$

Now, let λ_1 and λ_2 denote the solution of (21) and (22), respectively. Set $\xi = \tilde{u}_1$ in (21), $\xi = \tilde{u}_2$ in (22), $v = \lambda_1$ in (27), and $v = \lambda_2$ in (28). Combining the results yields

$$(\tilde{g},\lambda_1 - \lambda_2)_{\Gamma_0} = (u_1 - U,\tilde{u}_1)_{\Omega_1} + (u_2 - U,\tilde{u}_2)_{\Omega_2} + \frac{1}{\sigma}(u_1 - u_2,\tilde{u}_1 - \tilde{u}_2)_{\Gamma_0}$$

so that, substituting in (26), we have that

$$\left\langle \frac{\partial \mathcal{E}}{\partial g}, \tilde{g} \right\rangle = (\tilde{g}, \lambda_1 - \lambda_2)_{\Gamma_0} + \delta(g, \tilde{g})_{\Gamma_0}, \qquad \forall \tilde{g} \in L^2(\Gamma_0). \tag{29}$$

Using similar arguments, we have

$$\left\langle \frac{\partial \mathcal{E}}{\partial f_1}, \tilde{f}_1 \right\rangle = \left(\tilde{f}_1, \lambda_1\right)_{\Omega_1} + \gamma \left(f_1, \tilde{f}_1\right)_{\Omega_1}, \qquad \forall \tilde{f}_1 \in L^2(\Omega_1), \tag{30}$$

and

$$\left\langle \frac{\partial \mathcal{E}}{\partial f_2}, \tilde{f}_2 \right\rangle + \gamma \left(f_2, \tilde{f}_2\right)_{\Omega_2}, \qquad \forall \tilde{f}_2 \in L^2(\Omega_2). \tag{31}$$

Thus, the first-order necessary conditions $\frac{\partial \mathcal{E}}{\partial g} = 0$, $\frac{\partial \mathcal{E}}{\partial f_1} = 0$, and $\frac{\partial \mathcal{E}}{\partial f_2} = 0$ yield that

$$\begin{aligned} \delta(g, \tilde{g})_{\Gamma_0} &= -(\tilde{g}, \lambda_1 - \lambda_2)_{\Gamma_0}, & \forall \tilde{g} \in L^2(\Gamma_0), \\ \gamma \left(f_1, \tilde{f}_1\right)_{\Omega_1} &= -\left(\tilde{f}_1, \lambda_1\right)_{\Omega_1}, & \forall \tilde{f}_1 \in L^2(\Omega_1), \\ \gamma \left(f_2, \tilde{f}_2\right)_{\Omega_2} &= -\left(\tilde{f}_2, \lambda_2\right)_{\Omega_2}, & \forall \tilde{f}_2 \in L^2(\Omega_2), \end{aligned}$$

which are same as (23)–(25).

REMARK. Equation (29)–(31) yield an explicit formula for the gradient of $\mathcal{E}$, i.e.,

$$\begin{aligned} \frac{\partial \mathcal{E}}{\partial f_1} &= \gamma f_1 + \lambda_1, \\ \frac{\partial \mathcal{E}}{\partial f_2} &= \gamma f_2 + \lambda_2, \\ \frac{\partial \mathcal{E}}{\partial g} &= \delta g + (\lambda_1 - \lambda_2)|_{\Gamma_0}, \end{aligned}$$

where λ_1 and λ_2 are determined from f and g through (19)–(22). Thus, one has in hand the information needed if one were to use a gradient-based method, e.g., a method that requires $\mathcal{E}$ and $\frac{\partial \mathcal{E}}{\partial f_1}, \frac{\partial \mathcal{E}}{\partial f_2}, \frac{\partial \mathcal{E}}{\partial g}$ for a given approximation of f_1, f_2, and g, to solve our optimization problem. We now consider one such method.

2.5. A Gradient Method

The simple gradient method we consider is defined as follows. Given a starting guess $f_1^{(0)}$, $f_2^{(0)}$, and $g^{(0)}$, let

$$\begin{pmatrix} f_1^{(n+1)} \\ f_2^{(n+1)} \\ g^{(n+1)} \end{pmatrix} = \begin{pmatrix} f_1^{(n)} \\ f_2^{(n)} \\ g^{(n)} \end{pmatrix} - \alpha \begin{pmatrix} \dfrac{\partial \mathcal{E}}{\partial f_1} \\ \dfrac{\partial \mathcal{E}}{\partial f_2} \\ \dfrac{\partial \mathcal{E}}{\partial g} \end{pmatrix}, \qquad \text{for } n = 1, 2, \dots, \tag{32}$$

where α is a step size. Combining with (29)–(31) yields, for $n = 1, 2, \dots,$

$$\begin{pmatrix} f_1^{(n+1)} \\ f_2^{(n+1)} \\ g^{(n+1)} \end{pmatrix} = \begin{pmatrix} f_1^{(n)} \\ f_2^{(n)} \\ g^{(n)} \end{pmatrix} - \alpha \begin{pmatrix} \gamma f_1^{(n)} + \lambda_1^{(n)} \\ \gamma f_2^{(n)} + \lambda_2^{(n)} \\ \delta g^{(n)} + \left(\lambda_1^{(n)} - \lambda_2^{(n)}\right)\Big|_{\Gamma_0} \end{pmatrix}$$

or

$$\begin{pmatrix} f_1^{(n+1)} \\ f_2^{(n+1)} \\ g^{(n+1)} \end{pmatrix} = \begin{pmatrix} (1-\gamma\alpha)f_1^{(n)} - \alpha\lambda_1^{(n)} \\ (1-\gamma\alpha)f_2^{(n)} - \alpha\lambda_2^{(n)} \\ (1-\delta\alpha)g^{(n)} - \alpha\left(\lambda_1^{(n)} - \lambda_2^{(n)}\right)\Big|_{\Gamma_0} \end{pmatrix},$$

where $\lambda_1^{(n)}$ and $\lambda_2^{(n)}$ are determined by (21) and (22) with f_1, f_2, and g replaced by $f_1^{(n)}$, $f_2^{(n)}$, and $g^{(n)}$.

In summary, the simple gradient algorithm we consider is given as follows.

ALGORITHM 1.

1. Choose α, $f_1^{(0)}, f_2^{(0)}$, and $g^{(0)}$.
2. For $n = 0, 1, 2, \dots$,
   (a) determine $u_1^{(n)}$ and $u_2^{(n)}$ from

$$a_1\left(u_1^{(n)}, v\right) = \left(f^{(n)}, v\right)_{\Omega_1} + \left(g^{(n)}, v\right)_{\Gamma_0}, \qquad \forall v \in H^1_{\Gamma_1}(\Omega_1),$$
$$a_2\left(u_2^{(n)}, v\right) = \left(f^{(n)}, v\right)_{\Omega_2} - \left(g^{(n)}, v\right)_{\Gamma_0}, \qquad \forall v \in H^1_{\Gamma_2}(\Omega_2);$$

   (b) determine $\lambda_1^{(n)}$ and $\lambda_2^{(n)}$ from

$$a_1\left(\lambda_1^{(n)}, \xi\right) = \left(u_1^{(n)} - U, \xi\right)_{\Omega_1} + \frac{1}{\sigma}\left(u_1^{(n)} - u_2^{(n)}, \xi\right)_{\Gamma_0}, \qquad \forall \xi \in H^1_{\Gamma_1}(\Omega_1),$$
$$a_2\left(\lambda_2^{(n)}, \xi\right) = \left(u_2^{(n)} - U, \xi\right)_{\Omega_2} - \frac{1}{\sigma}\left(u_1^{(n)} - u_2^{(n)}, \xi\right)_{\Gamma_0}, \qquad \forall \xi \in H^1_{\Gamma_2}(\Omega_2);$$

   (c) determine $f_1^{(n+1)}, f_2^{(n+1)}$, and $g^{(n+1)}$ from

$$\begin{pmatrix} f_1^{(n+1)} \\ f_2^{(n+1)} \\ g^{(n+1)} \end{pmatrix} = \begin{pmatrix} (1-\gamma\alpha)f_1^{(n)} - \alpha\lambda_1^{(n)} \\ (1-\gamma\alpha)f_2^{(n)} - \alpha\lambda_2^{(n)} \\ (1-\delta\alpha)g^{(n)} - \alpha\left(\lambda_1^{(n)} - \lambda_2^{(n)}\right)\Big|_{\Gamma_0} \end{pmatrix}.$$

The following result which can be found in, e.g., [9], is useful in determining sufficient conditions for the convergence of the gradient method of Algorithm 1.

THEOREM 6. *Let X be a Hilbert space equipped with the inner product $(\cdot,\cdot)_X$ and norm $\|\cdot\|_X$. Suppose $\mathcal{M}$ is a functional on X such that*

1. *$\mathcal{M}$ has a local minimum at $\hat{x}$ and is twice differentiable in an open ball B centered at $\hat{x}$;*
2. *$|\langle \mathcal{M}''(u), (x,y)\rangle| \le M\|x\|_X\|y\|_X$, $\forall u \in B, x \in X, y \in X$;*
3. *$|\langle \mathcal{M}''(u), (x,x)\rangle| \ge m\|x\|_X^2$, $\forall u \in B, x \in X$,*

where M and m are positive constants. Let R denote the Riesz map. Choose $x^{(0)}$ sufficiently close to $\hat{x}$ and choose a sequence ρ_n such that $0 < \rho_* \le \rho_n \le \rho^* < 2m/M^2$. Then the sequence $x^{(n)}$ defined by

$$x^{(n)} = x^{(n-1)} - \rho_n R\mathcal{M}'\left(x^{(n-1)}\right), \qquad \text{for } n = 1, 2, \dots,$$

converges to $\hat{x}$.

We examine the second derivatives of $\mathcal{E}$ to determine the constants M and m. We have

$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial g^2}, (\tilde{g}, g)\right\rangle = \sum_{i=1}^{2}(u_i, \tilde{u}_i)_{\Omega_i} + \frac{1}{\sigma}(u_1 - u_2, \tilde{u}_1 - \tilde{u}_2)_{\Gamma_0} + \delta(g, \tilde{g})_{\Gamma_0},$$

$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1^2}, \left(\tilde{f}_1, f_1\right)\right\rangle = (\check{u}_1, \hat{u}_1)_{\Omega_1} + \frac{1}{\sigma}(\check{u}_1, \hat{u}_1)_{\Gamma_0} + \gamma\left(f_1, \tilde{f}_1\right)_{\Omega_1},$$

$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2^2}, \left(\hat{f}_2, f_2\right) \right\rangle = (\breve{u}_2, \tilde{u}_2)_{\Omega_2} + \frac{1}{\sigma}(\breve{u}_2, \hat{u}_2)_{\Gamma_0} + \gamma \left(f_2, \tilde{f}_2\right)_{\Omega_2},$$

$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial g \partial f_1}, (\tilde{g}, f_1) \right\rangle = (\breve{u}_1, \tilde{u}_1)_{\Omega_1} + \frac{1}{\sigma}(\breve{u}_1, \tilde{u}_1 - \tilde{u}_2)_{\Gamma_0},$$

$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial g \partial f_2}, (\tilde{g}, f_2) \right\rangle = (\breve{u}_2, \tilde{u}_2)_{\Omega_2} + \frac{1}{\sigma}(-\breve{u}_2, \tilde{u}_1 - \tilde{u}_2)_{\Gamma_0},$$

$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1 \partial f_2}, \left(\tilde{f}_1, f_2\right) \right\rangle = \frac{1}{\sigma}(-\breve{u}_2, \hat{u}_1)_{\Gamma_0},$$

$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1 \partial g}, \left(\tilde{f}_1, g\right) \right\rangle = (u_1, \hat{u}_1)\Omega_1 + \frac{1}{\sigma}(u_1 - u_2, \hat{u}_1)_{\Gamma_0},$$

$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2 \partial f_1}, \left(\tilde{f}_2, f_1\right) \right\rangle = \frac{1}{\sigma}(\breve{u}_1, -\hat{u}_2)_{\Gamma_0},$$

and

$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2 \partial g}, \left(\tilde{f}_2, g\right) \right\rangle = (u_2, \hat{u}_2)_{\Omega_2} + \frac{1}{\sigma}(u_1 - u_2, -\hat{u}_2)_{\Gamma_0},$$

where

$$\begin{aligned}
a_1(\tilde{u}_1, v) &= (\tilde{g}, v)_{\Gamma_0}, && \forall v \in H^1_{\Gamma_1}(\Omega_1),\\
a_2(\tilde{u}_2, v) &= -(\tilde{g}, v)_{\Gamma_0}, && \forall v \in H^1_{\Gamma_2}(\Omega_2),\\
a_1(u_1, v) &= (g, v)_{\Gamma_0}, && \forall v \in H^1_{\Gamma_1}(\Omega_1),\\
a_2(u_2, v) &= -(g, v)_{\Gamma_0}, && \forall v \in H^1_{\Gamma_2}(\Omega_2),\\
a_1(\hat{u}_1, v) &= \left(\tilde{f}_1, v\right)_{\Omega_1}, && \forall v \in H^1_{\Gamma_1}(\Omega_1),\\
a_1(\breve{u}_1, v) &= (f_1, v)_{\Omega_1}, && \forall v \in H^1_{\Gamma_1}(\Omega_1),\\
a_2(\hat{u}_2, v) &= \left(\tilde{f}_2, v\right)_{\Omega_2}, && \forall v \in H^1_{\Gamma_2}(\Omega_2),\\
a_2(\breve{u}_2, v) &= (f_2, v)_{\Omega_2}, && \forall v \in H^1_{\Gamma_2}(\Omega_2).
\end{aligned}$$

Then,

$$\begin{aligned}
&\left(\tilde{f}_1, \tilde{f}_2, \tilde{g}\right) \begin{pmatrix} \frac{\partial^2 \mathcal{E}}{\partial f_1^2} & \frac{\partial^2 \mathcal{E}}{\partial f_1 \partial f_2} & \frac{\partial^2 \mathcal{E}}{\partial f_1 \partial g} \\ \frac{\partial^2 \mathcal{E}}{\partial f_2 \partial f_1} & \frac{\partial^2 \mathcal{E}}{\partial f_2^2} & \frac{\partial^2 \mathcal{E}}{\partial f_2 \partial g} \\ \frac{\partial^2 \mathcal{E}}{\partial g \partial f_1} & \frac{\partial^2 \mathcal{E}}{\partial g \partial f_2} & \frac{\partial^2 \mathcal{E}}{\partial g^2} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ g \end{pmatrix} \\
&= \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1^2}, \left(\tilde{f}_1, f_1\right) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1 \partial f_2}, \left(\tilde{f}_1, f_2\right) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1 \partial g}, \left(\tilde{f}_1, g\right) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2 \partial f_1}, \left(\tilde{f}_2, f_1\right) \right\rangle \\
&\quad + \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2^2}, \left(\tilde{f}_2, f_2\right) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2 \partial g}, \left(\tilde{f}_2, g\right) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial g \partial f_1}, (\tilde{g}, f_1) \right\rangle \\
&\quad + \left\langle \frac{\partial^2 \mathcal{E}}{\partial g \partial f_2}, (\tilde{g}, f_2) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial g^2}, (\tilde{g}, g) \right\rangle \\
&= (u_1, \tilde{u}_1)_{\Omega_1} + (u_2, \tilde{u}_2)_{\Omega_2} + \frac{1}{\sigma}(u_1 - u_2, \tilde{u}_1 - \tilde{u}_2)_{\Gamma_0} \\
&\quad + (\breve{u}_1, \hat{u}_1)_{\Omega_1} + (\breve{u}_2, \hat{u}_2)_{\Omega_2} + \frac{1}{\sigma}(\breve{u}_1 - \breve{u}_2, \hat{u}_1 - \hat{u}_2)_{\Gamma_0}
\end{aligned}$$

$$+ (\breve{u}_1, \tilde{u}_1)_{\Omega_1} + (\breve{u}_2, \tilde{u}_2)_{\Omega_2} + \frac{1}{\sigma}(\breve{u}_1 - \breve{u}_2, \tilde{u}_1 - \tilde{u}_2)_{\Gamma_0}$$
$$+ (u_1, \hat{u}_1)_{\Omega_1} + (u_2, \hat{u}_2)_{\Omega_2} + \frac{1}{\sigma}(u_1 - u_2, \hat{u}_1 - \hat{u}_2)_{\Gamma_0}$$
$$+ \gamma\left(\left(f_1, \tilde{f}_1\right)_{\Omega_1} + \left(f_2, \tilde{f}_2\right)_{\Omega_2}\right) + \delta\,(g, \tilde{g})_{\Gamma_0},$$

so that

$$\left| \left(\tilde{f}_1, \tilde{f}_2, \tilde{g}\right) \begin{pmatrix} \dfrac{\partial^2 \mathcal{E}}{\partial f_1^2} & \dfrac{\partial^2 \mathcal{E}}{\partial f_1 \partial f_2} & \dfrac{\partial^2 \mathcal{E}}{\partial f_1 \partial g} \\ \dfrac{\partial^2 \mathcal{E}}{\partial f_2 \partial f_1} & \dfrac{\partial^2 \mathcal{E}}{\partial f_2^2} & \dfrac{\partial^2 \mathcal{E}}{\partial f_2 \partial g} \\ \dfrac{\partial^2 \mathcal{E}}{\partial g \partial f_1} & \dfrac{\partial^2 \mathcal{E}}{\partial g \partial f_2} & \dfrac{\partial^2 \mathcal{E}}{\partial g^2} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ g \end{pmatrix} \right|$$

$$\le C^2 \sum_{i=1}^{2} \left(\|g\|_{0,\Gamma_0} \|\tilde{g}\|_{0,\Gamma_0} + \|f_i\|_{0,\Omega_i} \left\|\tilde{f}_i\right\|_{0,\Omega_i} + \|f_i\|_{0,\Omega_i} \| \, \|\tilde{g}\|_{0,\Gamma_0} + \|g\|_{0,\Gamma_0} \left\|\tilde{f}_i\right\|_{0,\Omega_i} \right)$$
$$+ \frac{1}{\sigma} C^2 \left((\|f_1\|_{0,\Omega_1} + \|f_2\|_{0,\Omega_2} + 2\|g\|_{0,\Gamma_0}) \left(\left\|\tilde{f}_1\right\|_{0,\Omega_1} + \left\|\tilde{f}_2\right\|_{0,\Omega_2} + 2\|\tilde{g}\|_{0,\Gamma_0} \right) \right)$$
$$+ \delta \|g\|_{0,\Gamma_0} \|\tilde{g}\|_{0,\Gamma_0} + \gamma \sum_{i=1}^{2} \|f_i\|_{0,\Omega_i} \left\|\tilde{f}_i\right\|_{0,\Omega_i}$$
$$\le 3\left(2C^2 + \frac{4}{\sigma}C^2 + \delta + \gamma\right) \|x\|\|y\| \le M\|x\|\|y\|,$$

where $M = 12\max\{\gamma, \delta, 2C^2, 4(C^2/\sigma)\}$, $x = (\tilde{f}_1, \tilde{f}_2, \tilde{g})^\top$, $y = (f_1, f_2, g)^\top$,

$$\|x\| = \sqrt{\|f_1\|_{1,\Omega_1}^2 + \|f_2\|_{1,\Omega_2}^2 + \|g\|_{0,\Gamma_0}^2} \qquad \text{and} \qquad \|y\| = \sqrt{\left\|\tilde{f}_1\right\|_{1,\Omega_1}^2 + \left\|\tilde{f}_2\right\|_{1,\Omega_2}^2 + \|\tilde{g}\|_{0,\Gamma_0}^2}.$$

Setting $\tilde{g} = g$, $\tilde{f}_1 = f_1$, and $\tilde{f}_2 = f_2$, we also have that

$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial g^2}, (g, g) \right\rangle = \sum_{i=1}^{2} (u_i, u_i)_{\Omega_i} + \frac{1}{\sigma}(u_1 - u_2, u_1 - u_2)_{\Gamma_0} + \delta(g, g)_{\Gamma_0},$$
$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1^2}, (f_1, f_1) \right\rangle = (\breve{u}_1, \breve{u}_1)_{\Omega_1} + \frac{1}{\sigma}(\breve{u}_1, \breve{u}_1)_{\Gamma_0} + \gamma(f_1, f_1)_{\Omega_1},$$
$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2^2}, (f_2, f_2) \right\rangle = (\breve{u}_2, \breve{u}_2)_{\Omega_2} + \frac{1}{\sigma}(\breve{u}_2, \breve{u}_2)_{\Gamma_0} + \gamma(f_2, f_2)_{\Omega_2},$$
$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1 \partial g}, (f_1, g) \right\rangle = (u_1, \breve{u}_1)_{\Omega_1} + \frac{1}{\sigma}(u_1, \breve{u}_1 - \breve{u}_2)_{\Gamma_0},$$
$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial g \partial f_1}, (g, f_1) \right\rangle = (\breve{u}_1, u_1)_{\Omega_1} + \frac{1}{\sigma}(\breve{u}_1, u_1 - u_2)_{\Gamma_0},$$
$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2 \partial g}, (f_2, g) \right\rangle = (u_2, \breve{u}_2)_{\Omega_2} + \frac{1}{\sigma}(-u_2, \breve{u}_1 - \breve{u}_2)_{\Gamma_0},$$
$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial g \partial f_2}, (g, f_2) \right\rangle = (\breve{u}_2, u_2)_{\Omega_2} + \frac{1}{\sigma}(\breve{u}_2, u_1 - u_2)_{\Gamma_0},$$
$$\left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1 \partial f_2}, (f_1, f_2) \right\rangle = \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2 \partial f_1}, (f_2, f_1) \right\rangle = \frac{1}{\sigma}(-\breve{u}_2, \breve{u}_2)_{\Gamma_0}.$$

Then,

$$
(f_1, f_2, g) \begin{pmatrix} \dfrac{\partial^2 \mathcal{E}}{\partial f_1^2} & \dfrac{\partial^2 \mathcal{E}}{\partial f_1 \partial f_2} & \dfrac{\partial^2 \mathcal{E}}{\partial f_1 \partial g} \\ \dfrac{\partial^2 \mathcal{E}}{\partial f_2 \partial f_1} & \dfrac{\partial^2 \mathcal{E}}{\partial f_2^2} & \dfrac{\partial^2 \mathcal{E}}{\partial f_2 \partial g} \\ \dfrac{\partial^2 \mathcal{E}}{\partial g \partial f_1} & \dfrac{\partial^2 \mathcal{E}}{\partial g \partial f_2} & \dfrac{\partial^2 \mathcal{E}}{\partial g^2} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ g \end{pmatrix}
$$
$$
\begin{aligned}
&= \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1^2}, (f_1, f_1) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1 \partial f_2}, (f_1, f_2) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_1 \partial g}, (f_1, g) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2 \partial f_1}, (f_2, f_1) \right\rangle \\
&\quad + \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2^2}, (f_2, f_2) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial f_2 \partial g}, (f_2, g) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial g \partial f_1}, (g, f_1) \right\rangle \\
&\quad + \left\langle \frac{\partial^2 \mathcal{E}}{\partial g \partial f_2}, (g, f_2) \right\rangle + \left\langle \frac{\partial^2 \mathcal{E}}{\partial g^2}, (g, g) \right\rangle \\
&= (u_1 + \breve{u}_1, u_1 + \breve{u}_1)_{\Omega_1} + (u_2 + \breve{u}_2, u_2 + \breve{u}_2)_{\Omega_2} \\
&\quad + \frac{1}{\sigma}(u_1 - u_2 + \breve{u}_1 - \breve{u}_2, u_1 - u_2 + \breve{u}_1 - \breve{u}_2)_{\Gamma_0} \gamma(f_1, f_1)_{\Omega_1} + \gamma(f_2, f_2)_{\Omega_2} + \delta(g, g)_{\Gamma_0} \\
&\geq m\|x\|^2,
\end{aligned}
$$

where $m = \min\{\gamma, \delta\}$. Thus, the conditions of Theorem 6 hold for Algorithm 1.

Of course, other, more practical gradient-based algorithms can be defined. For example, the step length α could be changed at each iteration, e.g., through a line search in the direction of the negative gradient, or a conjugate gradient method may be used.

3. OPTIMIZATION WITHOUT PENALIZATION

In Section 2, we introduced two regularizing or penalty parameters, γ and δ, which substantially simplified the analyses. However, the use of these parameters detracts from the performance of gradient-type methods and, in any case, must be chosen with care. Furthermore, they prevent convergence to the true goals of the optimization problem which should involve the minimization of unpenalized functionals. Therefore, in this section, we consider domain decomposition methods for the optimization of unpenalized functionals.

3.1. The Model Problems

The functional in (2) contains the penalty parameter γ. Since the goal of our model problem is to find u as close as possible to a given U, we would rather minimize the unpenalized functional

$$
\mathcal{J}(u) = \frac{1}{2} \int_\Omega (u - U)^2 \, d\Omega \tag{33}
$$

subject to (1). Thus, we consider a model problem.

PROBLEM 4. Minimize the functional $\mathcal{J}(u(f))$ defined in (33) over suitable functions f subject to the constraints in (1).

We rewrite the functional $\mathcal{J}(u)$ in terms of the two subdomains as

$$
\mathcal{K}(u_1, u_2) = \frac{1}{2} \sum_{i=1}^{2} \int_{\Omega_i} (u_i - U)^2 \, d\Omega. \tag{34}
$$

Then, instead of Problem 4, we consider the following.

PROBLEM 5. Minimize the functional $\mathcal{K}(u_1(f_1, f_2), u_2(f_1, f_2))$ over suitable functions f_1 and f_2 subject to the constraints in (3).

Similar to Section 2 and for the same reason, for an arbitrary choice of the control g, the solutions u_1 and u_2 of Problem 5 are not the same as the solution u of Problem 4 in the respective subdomains. Thus, if we are to solve Problem 4 through Problem 5, we must also find the "right" g so that u_1 is as close as possible to u_2 along the interface Γ_0. One way to accomplish this is to minimize the unpenalized functional (5).

Thus, we again face a multiobjective optimization problem in which we now wish to find f_1, f_2, and g such that both the functionals $\mathcal{G}(\cdot,\cdot)$ of (5) and $\mathcal{K}(\cdot,\cdot)$ of (34) are minimized subject to (3). We combine the two functionals into the functional

$$\mathcal{F}_\sigma(u_1, u_2) = \frac{1}{2}\sum_{i=1}^{2}\int_{\Omega_i}(u_i - U)^2\,d\Omega + \frac{1}{2\sigma}\int_{\Gamma_0}(u_1 - u_2)^2\,d\Gamma, \tag{35}$$

where σ is a constant that can be chosen to change the relative importance of the contributions of the two functionals $\mathcal{K}$ and $\mathcal{G}$. The multiobjective minimization problem is then given by the following.

PROBLEM 6. Minimize the functional $\mathcal{F}_\sigma(u_1(f_1, f_2, g), u_2(f_1, f_2, g))$ over suitable functions f_1, f_2, and g subject to the constraints in (3).

3.2. A Gradient Method

We define the gradient method

$$\begin{pmatrix} f_1^{(n+1)} \\ f_2^{(n+1)} \\ g^{(n+1)} \end{pmatrix} = \begin{pmatrix} f_1^{(n)} \\ f_2^{(n)} \\ g^{(n)} \end{pmatrix} - \alpha_n \begin{pmatrix} \dfrac{\partial \mathcal{F}_\sigma}{\partial f_1} \\ \dfrac{\partial \mathcal{F}_\sigma}{\partial f_2} \\ \dfrac{\partial \mathcal{F}_\sigma}{\partial g} \end{pmatrix}, \qquad \text{for } n = 1, 2, \dots, \tag{36}$$

to solve Problem 6, where α_n is a step size.

The first derivatives $\frac{\partial \mathcal{F}_\sigma}{\partial f_1}$, $\frac{\partial \mathcal{F}_\sigma}{\partial f_1}$, and $\frac{\partial \mathcal{F}_\sigma}{\partial f_1}$ can be defined through their actions on variations $\tilde{f}_1$, $\tilde{f}_2$, and $\hat{g}$ of the control functions:

$$\left\langle \frac{\partial \mathcal{F}_\sigma}{\partial f_1}, \tilde{f}_1 \right\rangle = (u_1 - U, \tilde{u}_1)_{\Omega_1} + \frac{1}{\sigma}(u_1 - u_2, \tilde{u}_1)_{\Gamma_0}, \tag{37}$$

$$\left\langle \frac{\partial \mathcal{F}_\sigma}{\partial f_2}, \tilde{f}_2 \right\rangle = (u_2 - U, \tilde{u}_2)_{\Omega_2} - \frac{1}{\sigma}(u_1 - u_2, \tilde{u}_2)_{\Gamma_0}, \tag{38}$$

$$\left\langle \frac{\partial \mathcal{F}_\sigma}{\partial g}, \hat{g} \right\rangle = \sum_{i=1}^{2}(u_i - U, \hat{u}_i)_{\Omega_i} + \frac{1}{\sigma}(u_1 - u_2, \hat{u}_1 - \hat{u}_2)_{\Gamma_0}, \tag{39}$$

where, for $i = 1$ and 2, $\tilde{u}_i \in H^1_{\Gamma_i}(\Omega_i)$ is the solution of

$$a_i(\tilde{u}_i, v) = \left(\tilde{f}_i, v\right)_{\Omega_i}, \qquad \forall\, v \in H^1_{\Gamma_i}(\Omega_i), \tag{40}$$

and $\hat{u}_i \in H^1_{\Gamma_i}(\Omega_i)$ is the solution of

$$a_i(\hat{u}_i, v) = (-1)^{i+1}(\hat{g}, v)_{\Gamma_0}, \qquad \forall\, v \in H^1_{\Gamma_i}(\Omega_i). \tag{41}$$

Now, for $i = 1$ and 2, consider the adjoint problems

$$-\Delta\lambda_i = u_i - U, \quad \text{in } \Omega_i, \qquad \lambda_i = 0, \quad \text{on } \Gamma_i,$$
$$\frac{\partial\lambda_i}{\partial n_i} = (-1)^{i+1}\frac{1}{\sigma}(u_1 - u_2), \qquad \text{on } \Gamma_0,$$

or, in weak form,

$$a_i(\xi, \lambda_i) = (u_1 - U, \xi)_{\Omega_i} + (-1)^{i+1}\frac{1}{\sigma}(u_1 - u_2, \xi)_{\Gamma_0}, \qquad \forall\, \xi \in H^1_{\Gamma_i}(\Omega_i). \tag{42}$$

Using (37)–(42), one can obtain

$$\left\langle \frac{\partial \mathcal{F}_\sigma}{\partial f_1}, \tilde{f}_1 \right\rangle = \left(\tilde{f}_1, \lambda_1\right)_{\Omega_1},$$
$$\left\langle \frac{\partial \mathcal{F}_\sigma}{\partial f_2}, \tilde{f}_2 \right\rangle = \left(\tilde{f}_2, \lambda_2\right)_{\Omega_2},$$
$$\left\langle \frac{\partial \mathcal{F}_\sigma}{\partial g}, \hat{g} \right\rangle = (\hat{g}, \lambda_1 - \lambda_2)_{\Gamma_0}.$$

Then, the gradient method is given as follows: starting with an initial guess $f_1^{(0)}$, $f_2^{(0)}$, and $g^{(0)}$ for the control functions, for $n = 0, 1, 2, \dots$, first solve for the states $u_i^{(n)}$, $i = 1, 2$, from

$$a_i\left(u_i^{(n)}, v\right) = \left(f_i^{(n)}, v\right)_{\Omega_i} + \left(g^{(n)}, v\right)_{\Gamma_0}, \qquad \forall\, v \in H^1_{\Gamma_i}(\Omega_i), \tag{43}$$

then solve for the adjoint states $\lambda_i^{(n)}$, $i = 1, 2$, from

$$a_i\left(\xi, \lambda_i^{(n)}\right) = \left(u_1^{(n)} - U, \xi\right)_{\Omega_i} + (-1)^{i+1}\frac{1}{\sigma}\left(u_1^{(n)} - u_2^{(n)}, \xi\right)_{\Gamma_0}, \qquad \forall\, \xi \in H^1_{\Gamma_i}(\Omega_i), \tag{44}$$

and then update the control functions according to

$$\begin{pmatrix} f_1^{(n+1)} \\ f_2^{(n+1)} \\ g^{(n+1)} \end{pmatrix} = \begin{pmatrix} f_1^{(n)} \\ f_2^{(n)} \\ g^{(n)} \end{pmatrix} - \alpha_n \begin{pmatrix} \lambda_1^{(n)} \\ \lambda_2^{(n)} \\ \left(\lambda_1^{(n)} - \lambda_2^{(n)}\right)\Big|_{\Gamma_0} \end{pmatrix}.$$

4. NUMERICAL EXPERIMENTS

We now report on some experiments with the gradient algorithms of Sections 2.5 and 3.2. Let the domain Ω be the square $\{(x, y) : 0 < x < 1,\ 0 < y < 1\}$; Ω is divided into the two subdomains $\Omega_1 = \{(x, y) : 0 < x < 1/2,\ 0 < y < 1\}$ and $\Omega_2 = \{(x, y) : 1/2 < x < 1,\ 0 < y < 1\}$ having the interface $\Gamma_0 = \{(x, y) : x = 1/2,\ 0 < y < 1\}$. The finite element spaces are chosen to consist of the standard continuous, piecewise quadratic polynomial spaces based on triangular meshes. The computations were carried out for a target solution $U = (x - 1)y \sin x \cos(\pi y/2)$ which is feasible so that the solution of Problem 4 is $u = U$ and $f = -\Delta U$. This setup allows us to test the convergence of the various algorithms with respect to various parameters.

For the gradient algorithm of Section 2.5, is necessary to adjust the parameters α, δ, γ, and σ so that satisfactory convergence results can be obtained while still obtaining good agreement with the exact solution. We choose γ and δ to be small so that we do not over penalize the functional. Then α can be chosen so that the gradient method converges. Calculations have been performed for various values for the parameters and a tolerance 10^{-5} in the stopping criterion for the gradient method. First, we examine the convergence to the exact solutions for fixed parameters $\delta = 10^{-5}$, $\gamma = 10^{-7}$, and $\sigma = 1.4$; the rates of convergence (with respect to the grid

Table 1. Convergence of solution using the penalized functional with respect to the grid size for different values of the step size α in the gradient method.

α	L^2 Rate	H^1 Rate
1	2.66	1.98
3	2.49	1.98
5	2.60	1.98
6	2.55	1.98
7	2.90	1.99

Table 2. Convergence of gradient method using the penalized functional with respect to different values of the step size α in the gradient method and for different grid sizes.

α	$h = \frac{1}{4}$	$h = \frac{1}{8}$
1	13	13
3	4	4
5	7	8
6	16	17
7	188	262
8	diverges	

size h) for different step sizes α in the gradient method are given in Table 1. We see that the H^1 rate of convergence is exactly what we expect with piecewise quadratic finite element spaces; the L^2 rate is somewhat more erratic.

Next, we examine the number of iterations required for convergence of the gradient iteration. Again, $\delta = 10^{-5}$, $\gamma = 10^{-7}$, and $\sigma = 1.4$. The number of iterations is given in Table 2 for various values of α, the step size in the gradient method, and for two values of the grid size h. We see that too large a step size results in divergence and that there is an optimal step size; moreover, the number of iterations seems unaffected by the grid sizes h.

Results for the gradient method for the unpernalized functional (see Section 3.2) are similar.

REFERENCES

1. Q. Dinh, R. Glowinski and J. Periaux, Solving elliptic problems by domain decomposition methods with application, In *Elliptic Problem Solvers II*, (Edited by G. Birkoff and A. Schoenstadt), pp. 395–426, Academic Press, New York, (1984).
2. Q. Du, Optimization based nonoverlapping domain decomposition methods and their convergence, *Mathematics of Computation* (to appear).
3. Q. Du and M. Gunzburger, A gradient method approach to optimization-based multidisciplinary simulations and nonoverlapping domain decomposition algorithms, *SIAM J. Numer. Anal.* (to appear).
4. R. Glowinski and P. Le Tallec, Augmented Lagrangian interpretation of the nonoverlapping Schwarz alternating method, In *Proc. 3rd International Symposium on Domain Decomposition Methods for Partial Differential Equations*, (Edited by T. Chan *et al.*), pp. 224–231, SIAM, Philadelphia, PA, (1990).
5. M. Gunzburger, M. Heinkenschloss and H. Lee, Solution of elliptic pdes by an optimization-based domain decomposition method, *Appl. Math. Comp.* (to appear).
6. M. Gunzburger, H. Kwon and J. Peterson, An optimization based domain decomposition method for partial differential equations, *Computers Math. Applic.* **37** (10), 77–93, (1999).
7. M. Gunzburger and H. Lee, An optimization based domain decomposition method for the Navier-Stokes equations, *SIAM J. Numer. Anal.* (to appear).
8. V. Tikhomirov, *Fundamental Principles of the Theory of Extremal Problems*, Wiley, Chichester, (1986).
9. P. Ciarlet, *Introduction to Numerical Linear Algebra and Optimization*, Cambridge University, Cambridge, (1989).