NUMER. FUNCT. ANAL. AND OPTIMIZ., 12(5&6), 429–442 (1991)

APPROXIMATING OPTIMAL CONTROLS FOR ELLIPTIC OBSTACLE PROBLEM BY MONOTONE ITERATION SCHEMES

Viorel Barbu

University of Iasi and
Department of Mathematical Sciences
University of Cincinnati
Cincinnati, Ohio 45221-0025

Philip Korman

Institute for Dynamics and
Department of Mathematical Sciences
University of Cincinnati
Cincinnati, Ohio 45221-0025

1. INTRODUCTION

We begin by discussing the numerical solution of the obstacle problem with quasilinear nonlinearity

$$\begin{cases} -\Delta u \leq f(x, u, \nabla u) & \\ u \leq q(x) & \text{a.e. in } \Omega \\ (\Delta u + f(x, u, \nabla u))(u - q) = 0 & \end{cases} \tag{1}$$
$$u = g(x) \text{ on } \partial\Omega$$

Here the unknown function $u(x)$ is required to be below the obstacle $q(x)$; Ω is a smooth domain of any shape in R^d, and $-\Delta$ may be replaced throughout the paper by a uniformly elliptic operator of second order. We shall sometimes assume for definiteness that $d = 2$, although all our results hold for any spatial dimension.

We replace the domain Ω by a uniform square mesh Δ_h of step size h, denoting $n = (n_1, n_2)$, $x_n = (n_1 h, n_2 h)$ and $u_n = u(x_n)$. The finite difference version of (1) is then

$$\begin{cases} -\Delta_h u_n \leq f\left(x_n, u_n, \dfrac{u_{n+e_1} - u_{n-e_1}}{2h}, \dfrac{u_{n+e_2} - u_{n-e_2}}{2h}\right) & \\ u_n \leq q_n \equiv q(x_n) & x_n \in \Omega_h \\ -\Delta_h u_n = f\left(x_n, u_n, \dfrac{u_{n+e_1} - u_{n-e_1}}{2h}, \dfrac{u_{n+e_2} - u_{n-e_2}}{2h}\right) & \text{when } u_n < q_n \end{cases} \tag{2}$$
$$u_n = g_n \quad \text{for} \quad x_n \in \partial\Omega_h$$

Here Δ_h is the central difference approximation of the Laplacian, and the boundary conditions on $\partial\Omega_h$ are defined in any standard way (see e.g., [10]); $e_1 = (1, 0)$ and $e_2 = (0, 1)$.

To solve the discrete version of the equation $-\Delta u = f(x, u, \nabla u)$ with Dirichlet boundary conditions, it is natural to look for the steady-state of the corresponding parabolic problem, and use the Crank-Nicholson scheme

$$\frac{u_n^{p+1} - u_n^p}{\tau} = \alpha \Delta_h u_n^{p+1} + (1 - \alpha)\Delta_h u_n^p$$
$$+ f\left(x_n, u_n^p, \frac{u_{n+e_1}^p - u_{n-e_1}^p}{2h}, \frac{u_{n+e_2}^p - u_{n-e_2}^p}{2h}\right) \tag{3}$$

with $0 \leq \alpha \leq 1$, and properly chosen initial guess u_n^0 and the time step τ. This approach was used by P. Korman [6], and earlier by C.U. Huy, P.J. McKenna and W. Walter [5] in case f is independent of ∇u. The authors of [5] had found that in order for (3) to give a monotone in p scheme, one needs to impose a restriction on size of τ of the type $\tau \leq \text{const} \cdot h^2$ for all $0 \leq \alpha < 1$ (i.e. except for fully implicit scheme). Since obstacle problems in particular include equations, we restrict our attention only to the fully explicit and fully implicit schemes for (2).

Starting with $u_n^0 = q_n$, we obtain iterations u_n^p decreasing in p for all n. In case of explicit scheme this requires that h and $\tau = \tau(h)$ be small enough (no restriction on h if f is independent of ∇u). Monotone schemes have the advantage that convergence (possibly to $-\infty$) is guaranteed. If one assumes the existence of a subsolution, then the scheme converges to the solution of (2), and, moreover, one can obtain a posteriori error bounds.

In Section 3 we use this approach to compute the optimal control associated with a controlled obstacle problem studied in [2]. In Section 4 one derives the maximum principle and a numerical algorithm of monotone type for optimal controls of the elliptic obstacle problem with not smooth pay-off functional of the form $\|y - y_0\|_{C(\bar{\Omega})}$.

2. MONOTONE SCHEMES FOR THE OBSTACLE PROBLEM

We begin with the explicit scheme in case f is independent of ∇u. The algorithm for solving (2) in case $d = 2$ is the following (for all $x_n \in \Omega_h$)

$$\begin{cases} u_n^0 := q_n \\ \bar{u}_n^{p+1} := u_n^p\left(1 - \dfrac{4\tau}{h^2}\right) + \dfrac{\tau}{h^2}(u_{n+e_1}^p + u_{n-e_1}^p + u_{n+e_2}^p + u_{n-e_2}^p) + \tau f(x_n, u_n^p) \\ \text{If } \bar{u}_n^{p+1} < q_n \text{ then } u_n^{p+1} := \bar{u}_n^{p+1} \text{ else } u_n^{p+1} := q_n \end{cases} \tag{4}$$

Proposition 1: Assume f is Lipschitz continuous in u uniformly in $x \in \Omega$. Then for τ sufficiently small we have

$$q_n \geq u_n^1 \geq u_n^2 \geq \cdots \geq u_n^p \cdots \qquad \text{for all } p \text{ and } n \tag{5}$$

If the limit $\lim_{p\to\infty} u_n^p \equiv u_n$ exists, then it gives a solution of (2).

Proof: Clearly $u_n^1 \leq u_n^0 = q_n$ for all n. We prove by induction that $u_n^{p+1} \leq u_n^p$ for all p (and n). Let $w_n^p = u_n^{p+1} - u_n^p$, and assume $w_n^{p-1} \leq 0$. Then from (4),

$$w_n^p \leq w_n^{p-1}\left(1 - \frac{4\tau}{h^2} - \tau L\right) + \frac{\tau}{h^2}(w_{n+e_1}^{p-1} + w_{n-e_1}^{p-1} + w_{n+e_2}^{p-1} + w_{n-e_2}^{p-1}) \leq 0 \quad (6)$$

if τ is so small that $1 - 4\tau/h^2 - \tau L > 0$ (L is the Lipschitz constant). This proves (5).

Assume now that $\lim_{p\to\infty} u_n^p = u_n$ exists. We consider two cases.

(i) $u_{n_0} = q_{n_0}$. Then $\bar{u}_{n_0}^{p+1} \geq u_{n_0}^p$ for all p, as it is clear from (6) that if the iterates go down from the obstacle, they continue to go down (and hence cannot come back). Then

$$0 \leq \frac{\bar{u}_{n_0}^{p+1} - u_{n_0}^p}{\tau} = \Delta_h u_{n_0}^p + f(x_0, u_{n_0}^p)$$

Passing to the limit as $p \to \infty$, we obtain the first inequality in (2).

(ii) $u_{n_0} < q_{n_0}$. By the same reasoning as above, we conclude that $\bar{u}_{n_0}^p < q_{n_0}$ for all $p \geq \bar{p}$ with some $\bar{p} \geq 1$. It follows that $u_{n_0}^p$ satisfies the second equation in (4) (i.e., truncation never occurs). Passing to the limit as $p \to \infty$, we see that the third equation in (2) is satisfied at $n = n_0$.

One way to ensure convergence of u_n^p is to assume existence of a subsolution, i.e., a grid function φ_n, such that $\varphi_n < q_n$ in Ω_h and

$$-\Delta_h\varphi_n \leq f(x_n, \varphi_n) \text{ in } \Omega_h \qquad \varphi_n = g_n \quad \text{on } \partial\Omega_h \quad (7)$$

Define iterates v_n^p by using (4) with $v_n^0 = \varphi_n$.

Proposition 2: Assume f is Lipschitz continuous in u uniformly in $x \in \Omega$. Then for τ sufficiently small we have

$$\varphi_n \leq v_n^1 \leq v_n^2 \leq \cdots \leq u_n^2 \leq u_n^1 \leq q_n \quad (8)$$

Call $\lim_{p\to\infty} u_n^p = u_n$, $\lim_{p\to\infty} v_n^p = v_n$. Then both u_n and v_n are solutions of (2).

Proof: The inequalities in (8) are proved by induction, using an argument similar to the one in Proposition 1. Then for all p and n

$$0 \leq \frac{v_n^{p+1} - v_n^p}{\tau} = \Delta_h v_n^p + f(x_n, v_n^p)$$

Passing to the limit as $p \to \infty$, we obtain the first inequality in (2). If $v_n < q_n$, then passing to the limit in the second line of (4), we obtain the third equation in (2).

Next we describe the implicit scheme for (2), corresponding to $\alpha = 1$ in (3). Starting with $u_n^0 = q_n$, we solve at each step an obstacle problem with a known force term (I denotes the identity operator)

$$\begin{cases} (I - \tau\Delta_h)u_n^{p+1} \leq u_n^p + \tau f(x_n, u_n^p) & \\ u_n^{p+1} \leq q_n & x_n \in \Omega_h \\ (I - \tau\Delta_h)u_n^{p+1} = u_n^p + \tau f(x_n, u_n^p) \qquad \text{when } u_n^{p+1} < q_n & \end{cases} \quad (9)$$
$$u_n^{p+1} = g_n \text{ on } \partial\Omega_h$$

We solve (9) using Gauss–Seidel iteration with truncation (see [4]), taking u_n^p for an initial guess.

Proposition 3: Assume that f is Lipschitz continuous in u uniformly in $x \in \Omega$ with Lipschitz constant L. Then for $\tau < 1/L$ the iterates defined by (9) will satisfy (5). Moreover, if there exists a subsolution φ_n defined by (7), and we define the iterates v_n^p by using (9) with $v_n^0 = \varphi_n$, then (8) holds.

Proof: One easily sees that the map from u_n^p to u_n^{p+1}, defined by (9), is monotone. The rest of the proof is similar to the above.

Example: Consider the obstacle problem (1) for the equation $-u'' = u(5 - u)$ for $0 < x < 2$, $u(0) = u(2) = 0$, and the obstacle $q(x) = (x - 1)^2 + 1$. The problem describes steady-state density of a certain species, which obeys the logistic population model with diffusion and an obstacle for growth. For the explicit scheme we took $h = 0.1$, $\tau = 0.0025$ and obtained stabilization of six decimal digits after 300 time steps. For the implicit scheme we took $h = 0.1$, $\tau = 0.5$, and obtained the same accuracy after 150 time steps. We checked our result by a program based on the usual monotone iterations, developed in [7].

Next we discuss the case $f = f(x, u, \nabla u)$, which can be treated using the explicit scheme. We omit the proof of the following result, since it is a simple combination of Proposition 1 above and Theorem 3 in [6].

Proposition 4: Assume that the function $f(x, u, p_1, p_2)$ is continuously differentiable in u, p_1, p_{22} and for $x \in \Omega$ and $|u| \leq K$ satisfies

$$|f_{p_i}| \leq c(1 + |p_1| + |p_2|)^\alpha, \quad i = 1, 2, \quad 0 \leq \alpha < 1, \quad c = c(K) \tag{10}$$

Assuming existence of a subsolution, define the sequences u_n^p and v_n^p as in the Propositions 1 and 2. Then for h and $\tau = \tau(h)$ sufficiently small the conclusions of the Proposition 2 hold.

Remark: It is natural to expect that one may allow $0 \leq \alpha \leq 1$ in (10).

Example: Consider the obstacle problem for the equation $-u'' = 1 + u'^2$ for $0 < x < 2$, $u(0) = u(2) = 0$, and the obstacle $q(x) = 0.5$. Integrating the equation, we see that the solution of class $H^2(0, 2)$ is

$$u(x) = \begin{cases} \ln \left|\dfrac{\cos(-x + \alpha_1)}{\cos \alpha_1}\right| & 0 \leq x \leq p \\ 0.5 & p \leq x \leq q \\ \ln \left|\dfrac{\cos(-x + \alpha_2)}{\cos(-2 + \alpha_2)}\right| & q \leq x \leq 2 \end{cases}$$

where p and q are determined by the conditions $u(p) = u(q) = 0.5$, $u'(p) = u'(q) = 0$.

One calculates $\alpha_1 = p \simeq 0.919$ and $\alpha_2 = q \simeq 2 - \alpha_1$. We computed the solution with $h = 0.1$. The contact set consisted of $x = 0.9$, 1.0 and 1.1, and the values at other mesh points corresponded to the above formulas.

3. CONTROL OF THE OBSTACLE PROBLEM IN L^2-NORM

In [2] the following control problem was considered (we specialize):

$$\text{Minimize } \|y - y_0\|_{L^2(\Omega)} + h(u) \tag{11}$$

on all $(y, u) \in (H^2(\Omega) \cap H_0^1(\Omega)) \times L^2(\Omega)$, subject to

$$-\Delta y + \beta(y - \psi) \ni u + f(x) \text{ in } \Omega \quad y = 0 \text{ on } \partial\Omega \tag{12}$$

where

$$\beta(r) = \begin{cases} 0 & \text{for } r > 0 \\ R^- & \text{for } r = 0 \\ \emptyset & \text{for } r < 0 \end{cases}$$

$$h(u) = \begin{cases} 0 & \text{if } -\rho \le u(x) \le \rho \text{ a.e. in } \Omega \\ +\infty & \text{otherwise} \end{cases}$$

and $f \in L^2(\Omega)$, $\psi \in H^2(\Omega)$, $y_0 \in L^2(\Omega)$ are known functions with $\psi \le 0$ on $\partial\Omega$; $\rho > 0$ is a given constant.

The optimality system derived in [2], with an additional (not very restrictive) assumption that $\Delta\psi \neq \pm\rho$ a.e. in Ω, is (with a dual function $p(x) \in H_0^1(\Omega)$)

$$\begin{aligned} -\Delta y &= \rho \operatorname{sign} p \text{ in } \Omega_+ = \{x \in \Omega; y > \psi(x)\} \\ -\Delta y &\ge \rho \operatorname{sign} p \text{ a.e. in } \Omega_0 = \Omega \backslash \Omega_+ \\ y &\ge \psi \text{ a.e. in } \Omega \\ y &= 0 \text{ on } \partial\Omega \\ -\Delta p &= y^0 - y \text{ in } \Omega_+ \\ p &= 0 \text{ on } \Omega_0 \cup \partial\Omega \end{aligned} \tag{13}$$

where

$$\operatorname{sign} p = \begin{cases} \dfrac{p}{|p|} & \text{for } p \neq 0 \\ [-1, 1] & \text{for } p = 0 \end{cases}$$

The problem (13) can be thought of as an obstacle problem for y coupled with a "Poisson equation" in p. However, the domain of the equation for p depends on the values of y. The problem (13) can be numerically solved similarly to (1). We present here the explicit version of the algorithm. Although the Propositions of the previous section do not carry over to (13), the algorithm is quite effective numerically in one and two dimensions.

We start with an arbitrary y_n^0, p_n^0. We compute as follows ($y_{0_n} = y_0(x_n)$, $\psi_n = \psi(x_n)$)

$$y_n^{k+1} := y_n^k + \tau\Delta_h y_n^k + \tau\rho \operatorname{sign}_\epsilon p_n^k \tag{14}$$

$$\text{If } y_n^{k+1} > \psi_n \text{ then}$$

$$p_n^{k+1} := p_n^k + \tau\Delta_h p_n^k + \tau y_{on} - \tau y_n^k \tag{15}$$

$$\text{If } y_n^{k+1} \le \psi_n \text{ then } y_n^{k+1} := \psi_n,\ p_n^{k+1} := 0 \tag{16}$$

Here

$$\operatorname{sign}_\epsilon p = \begin{cases} 1 & \text{if } \dfrac{1}{\epsilon} p > 1 \\ \dfrac{1}{\epsilon} p & \text{if } -1 \le \dfrac{1}{\epsilon} p \le 1 \\ -1 & \text{if } \dfrac{1}{\epsilon} p < -1 \end{cases}$$

The algorithm converged for all one- or two-dimensional examples that we tried, with six decimal digits of $y(x)$ stabilizing uniformly in ϵ for $\epsilon < 10^{-5}$.

Example: Assume $y_0 < \psi < 0$ for all $x \in \Omega$. Then since $y \geq \psi$, it follows by the maximum principle that $p < 0$ in Ω_+, i.e., $-\Delta y = -\rho$ in Ω_+. Let now $\Omega = (0, 2)$, $\rho = 8$, $\psi(x) = -1$, $y_0(x) = -5$. Then one easily calculates that

$$u(x) = \begin{cases} 4(x^2 - x) & \text{for } 0 \leq x \leq \frac{1}{2} \\ -1 & \text{for } \frac{1}{2} \leq x \leq \frac{3}{2} \\ 4[(x-2)^2 + x - 2] & \text{for } \frac{3}{2} \leq x \leq 2 \end{cases}$$

is a solution of (13). We ran the algorithm (14) with $h = 0.1$, $\tau = 0.0025$, and obtained a close approximation of the solution at the grid points.

"Sticky" obstacle: The algorithm above had the disadvantage that on the contact set the iterates would go on and slightly off the obstacle. The following computational artifice was used to overcome the problem: after a certain time t_0, once the solution $y(x)$ gets into contact with the obstacle at some grid point, the solution is forced to stay in contact there for all later times. I.e., we replace (14) by

$$\begin{gathered} \text{If } (y_n^k > \psi_n) \text{ or } (t < t_0) \quad \text{then} \\ y_n^{k+1} := y_n^k + \tau\Delta_h y_n^k + \tau\rho \operatorname{sign}_\epsilon p_n^k \end{gathered} \tag{17}$$

We note that initially the iterates must be allowed to evolve freely, so that t_0 must be chosen carefully. When this modified algorithm was run for the above example with $t_0 = 2.5$, we obtained the exact values of the solution at all grid points.

Example: We computed the solution of (13) on $\Omega = (0, 2)$, with $y_0 = -2x + 1.5$, $\psi = -1.2$ and $\rho = 8$. We took $h = 0.1$, $\tau = 0.0025$. Figure 1 shows the result after 2000 time steps.

4. OPTIMAL CONTROL OF THE OBSTACLE PROBLEM IN SUPP-NORM

We shall study here the optimal control problem

$$\text{Minimize } \|y - y_0\|_{C(\bar{\Omega})} + h(u) \tag{18}$$

on all $(y, u) \in (H^2(\Omega) \cap H_0^1(\Omega)) \times U$ subject to

$$\begin{aligned} -\Delta y + \beta(y - \psi) \ni Bu + f \quad & \text{in } \Omega \\ y = 0 \quad & \text{on } \partial\Omega \end{aligned} \tag{19}$$

Here Ω is a bounded and open domain of R^3 with a smooth boundary (of class $C^{1,1}$ for instance)

$$\beta(r) = \begin{cases} 0 & \text{for } r > 0 \\ R^- & \text{for } r = 0 \\ \varnothing & \text{for } r < 0 \end{cases} \tag{20}$$

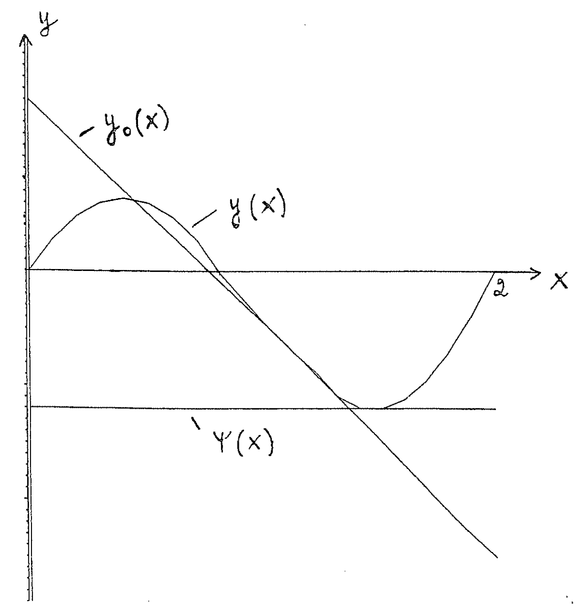

Figure 1

$y_0 \in C(\overline{\Omega})$ $f \in L^2(\Omega)$ and $\psi \in H^2(\Omega)$ are given functions such that $\psi \leq 0$ on $\partial\Omega$.

The controllers space U is a real Hilbert space with the norm $\|\cdot\|$ and scalar product $\langle\cdot, \cdot\rangle$ and B is a linear continuous operator from U to $L^2(\Omega)$.

The function $h: U \to \overline{\mathbf{R}} =]-\infty, +\infty]$ is convex, lower semicontinuous and

$$\lim_{\|u\|\to\infty} h(u) = +\infty \tag{21}$$

We have denoted by $C(\overline{\Omega})$ the space of all continuous real valued functions on $\overline{\Omega}$ with the usual supremum norm $\|\cdot\|_{C(\overline{\Omega})}$; $H^2(\Omega)$ and $H_0^1(\Omega)$ are usual Sobolev spaces on Ω.

It is well known that under our assumptions problem (2) has a unique solution $y \in H^2(\Omega) \cap H_0^1(\Omega)$ which satisfies the complementarity system (the obstacle problem)

$$\begin{aligned} (\Delta y + Bu + f)(y - \psi) &= 0 \quad \text{a.e. in } \Omega \\ y \geq \psi \quad \Delta y + Bu + f &\leq 0 \quad \text{a.e. in } \Omega \\ By &= 0 \quad \text{on } \partial\Omega \end{aligned} \tag{22}$$

Moreover, since $H^2(\Omega) \subset C(\overline{\Omega})$ the functional $u \to \|y^u - y_0\|_{C(\overline{\Omega})} + h(u)$ is weakly lower semicontinuous on U (y^n is the solution to (19)) and so by (21) it follows that the problem (18) has at least one solution u.

The main result of this section, Theorem 1 below is concerned with first order optimality conditions (the maximum principle) for this problem. Such a result has been established in [1], [2] (see also [3], [4]) for problems of the form

$$\text{Minimize } g(y) + h(u) \text{ subject to (19)} \tag{23}$$

where g is a locally Lipschitzian function on $L^p(\Omega)$, $1 \leq p < \infty$ and does not cover the present situation.

However, a payoff of the form (18) represents a more realistic model for penalizing the deviation of the system response y from a given state y_0.

Throughout in the sequel we shall denote by $M(\overline{\Omega})$ the space of all bounded Radon measures on $\overline{\Omega}$ i.e., the dual space of $C(\overline{\Omega})$, and by $[z, \varphi]_+$,

$$[z, \varphi]_+ = \lim_{\lambda \downarrow 0} \frac{1}{\lambda} (\|z + \lambda\varphi\|_{C(\overline{\Omega})} - \|z\|_{C\Omega)}) \tag{24}$$

for all $z, \varphi \in C(\overline{\Omega})$.

We set $\phi(y) = \|y - y_0\|_{C(\overline{\Omega})}$ and denote by $\partial\phi(y) \subset M(\overline{\Omega})$ the subdifferential of ϕ at y, i.e.,

$$\partial\phi(y) = \{\mu \in M(\overline{\Omega});\ \mu(\varphi) \leq [y - y_0, \varphi]_+ \ \forall\, \varphi \in C(\overline{\Omega})\} \tag{25}$$

We have (see [9])

$$\begin{aligned}[z, \varphi]_+ &= \max\{\varphi(x_0) \operatorname{sign} z(x_0);\ x_0 \in M_z\} \qquad \text{if } z \not\equiv 0 \\ &= \|\varphi\|_{C(\overline{\Omega})} \qquad \text{if } z = 0,\end{aligned} \tag{26}$$

where $M_z = \{x_0 \in \overline{\Omega};\ |z(x_0)| = \|z\|_{C(\overline{\Omega})}\}$. Hence

$$\begin{aligned}\partial\phi(y) = \{\mu \in M(\overline{\Omega});\ \mu(\varphi) \leq \max\{\varphi(x_0) \operatorname{sign}(y - y_0)(x_0);\ x_0 \in M_{y-y_0}\} \\ \forall\, \varphi \in C(\overline{\Omega})\}\end{aligned} \tag{27}$$

Theorem 1: Let (y^*, u^*) be any optimal pair in problem (18). Then there exist $p \in W_0^{1,q}(\Omega)$, $1 < q < \frac{3}{2}$, $\Delta p \in M(\overline{\Omega})$ and $\mu \in \partial\phi(y^*)$ such that

$$\Delta p = \mu \text{ in } \{x \in \Omega;\ y^*(x) > \psi(x)\} \tag{28}$$

$$p(x)(\Delta y^*(x) + Bu^*(x) + f(x)) = 0 \qquad \text{a.e. } x \in \Omega \tag{29}$$

$$B^*p \in \partial h(u) \tag{30}$$

Here ∂h is the subdifferential of h.

According to (27), equation (28) reduces to

$$\begin{aligned}-\int_{[y>\psi]} \nabla p \cdot \nabla\varphi\, dx \leq \max\{\varphi(x_0) \operatorname{sign}(y^*(x_0) - y_0(x_0));\ x_0 \in M_{y^*-y_0}\} \\ \forall\, \varphi \in C_0^\infty(y^* > \psi)\end{aligned} \tag{31}$$

We note that if $M_{y^*-y_0} = \{x_0\}$ then $\mu = \operatorname{sign}(y^* - y_0)(x_0)\delta(x_0)$ and so equation (28) becomes

$$\Delta p = \operatorname{sign}(y^* - y_0)(x_0)\delta(x_0) \text{ in } \{x \in \Omega;\ y^*(x) > \psi(x)\} \tag{32}$$

where $\delta(x_0)$ is the Dirac measure concentrated at x_0.

Proof of Theorem 1: We shall use the method developed in [1] and [2]. If (y^*, u^*) is optimal in problem (18), consider the approximating control problem

$$\min\left\{\phi_\epsilon(y) + h(u) + \frac{1}{2}\|u - u^*\|^2\right\} \tag{33}$$

subject to $(y, u) \in (H_0^1(\Omega) \cap H^2(\Omega)) \times U$ and

$$-\Delta y + \beta^\epsilon(y - \psi) = Bu + f \quad \text{in } \Omega \tag{34}$$

where β^ϵ is a smooth approximation of β, i.e.,

$$\begin{aligned}\beta^\epsilon(r) &= -\epsilon^{-1} \int_{-\infty}^{\infty} ((r - \epsilon^2\theta)^- - \epsilon^2\theta^-)\rho(\theta)\, d\theta \\ &= \epsilon^{-1} \int_{\epsilon^{-2}r}^{\infty} (r - \epsilon^2\theta)\rho(\theta)\, d\theta + \epsilon \int_0^1 \theta\rho(\theta)\, d\theta \end{aligned}\tag{35}$$

Here $\rho \in C_0^\infty(\Omega)$ is such that $\rho(r) = 0$ for $|r| > 1$, $\rho(r) = \rho(-r)$ and $\int \rho(r)\, dr = 1$. The function $\phi_\epsilon : L^2(\Omega) \to R$ is defined by

$$\phi_\epsilon(y) = \inf\left\{\frac{1}{2\epsilon} \|y - z\|^2_{L^2(\Omega)} + \overline{\phi}(z);\ z \in L^\infty(\Omega))\right\} \tag{36}$$

where $\overline{\phi}(z) = \|z - y_0\|_{L^\infty(\Omega)}$. It is well known that ϕ_ϵ is Fréchet differentiable and

$$\nabla\phi_\epsilon(y) = \epsilon^{-1}(y - (1 + \epsilon\partial\overline{\phi})^{-1}y) \tag{37}$$

where $\partial\overline{\phi} : L^2(\Omega) \to L^2(\Omega)$ is the subdifferential of $\overline{\phi}$.

Let (y_ϵ, u_ϵ) be optimal in problem (33). Then arguing as in [2], p. 68, it follows that

$$\begin{aligned} u_\epsilon &\to u^* \text{ strongly in } U \\ y_\epsilon &\to y^* \text{ strongly in } H_0^1(\Omega) \cap H^2(\Omega) \\ \beta^\epsilon(y_\epsilon - \psi) &\to Bu^* + \Delta y^* + f \quad \text{weakly in } L^2(\Omega) \end{aligned}\tag{38}$$

Moreover, there are $p_\epsilon \in H_0^1(\Omega) \cap H^2(\Omega)$ such that

$$\Delta p_\epsilon - (\beta^\epsilon)'(y_\epsilon - \psi)p_\epsilon = \nabla\phi_\epsilon(y_\epsilon) \text{ in } \Omega \tag{39}$$

$$B^*p_\epsilon \in \partial h(u_\epsilon) + u_\epsilon - u^* \tag{40}$$

On the other hand, by (37) we see that,

$$\|\nabla\phi_\epsilon(y_\epsilon)\|_{L^1(\Omega)} \le c \qquad \forall\, \epsilon > 0 \tag{41}$$

because

$$\sup\{\|\xi\|_{L^1(\Omega)};\ \xi \in \partial\overline{\phi}(y)\} \le 1$$

Now we multiply equation (39) by sign p_ϵ (or more exactly by $\zeta(p_\epsilon)$ where ζ is a smooth approximation of sign such that $\zeta(0) = 0$). We get the estimate

$$\int_\Omega |(\beta^\epsilon)'(y_\epsilon - \psi)p_\epsilon|\, dx \le 1 \qquad \forall\, \epsilon > 0 \tag{42}$$

Hence

$$\|\Delta p_\epsilon\|_{L^1(\Omega)} \le c \quad \forall\, \epsilon > 0 \tag{43}$$

Now let $h_i \in L^\alpha(\Omega)$, $i = 0, 1$, $\alpha > 3$. According to a well-known result due to G. Stampacchia [11] the boundary value problem

$$-\Delta\theta = h_0 + \sum_{i=1}^{2} \frac{\partial h_i}{\partial x_i} \quad \text{in } \Omega, \qquad \theta = 0 \text{ on } \partial\Omega \tag{44}$$

has a unique solution $\theta \in H_0^1(\Omega) \cap L^\infty(\Omega)$ and

$$\|\theta\|_{L^\infty(\Omega)} \leq c \sum_{i=0}^{2} \|h_i\|_{L^\alpha(\Omega)} \qquad \forall\ h_i \in L^\alpha(\Omega) \tag{45}$$

This yields

$$-\int_\Omega \theta \Delta p_\epsilon \, dx = \int_\Omega h_0 p_\epsilon - \sum_{i=1}^{2} \int_\Omega h_i (p_\epsilon)_{x_i} \, dx$$

and by (43) and (45) we infer that

$$\|p_\epsilon\|_{W_0^{1,q}(\Omega)} \leq c \qquad \forall\ \epsilon > 0 \tag{46}$$

where $1/q + 1/\alpha = 1$, i.e., $1 < q < 3/2$.

Selecting a subsequence we may assume that

$$p_\epsilon \to p \text{ weakly in } W_0^{1,q}(\Omega) \tag{47}$$

$$\nabla\phi(y_\epsilon) \to \mu \text{ vaguely in } M(\overline{\Omega}) \tag{48}$$

$$\dot{\beta}^\epsilon(y_\epsilon - \psi)p_\epsilon \to \nu \text{ vaguely in } M(\overline{\Omega}) \tag{49}$$

on a generalized sequence $\{\epsilon\} \to 0$.

By (39) we see that

$$\Delta p = \nu + \mu \text{ in } \mathcal{D}'(\Omega) \tag{50}$$

$$B^* p \in \partial h(u^*) \tag{51}$$

Since $y_\epsilon \to y^*$ uniformly in $C(\overline{\Omega})$ (because $H^2(\Omega) \subset C(\overline{\Omega})$ compactly) we have

$$\nu = 0 \text{ in } \{x \in \Omega;\ y^*(x) > \psi(x)\} \tag{52}$$

On the other hand, by a little calculation involving (35) we see that (see [2], p. 86)

$$|p_\epsilon \beta^\epsilon(y_\epsilon - \psi)| \leq \epsilon |p_\epsilon (\beta^\epsilon)'(y_\epsilon - \psi)| (\epsilon^{-1}|y_\epsilon - \psi| \xi_\epsilon + \epsilon^{-1}|y_\epsilon - \psi| \eta_\epsilon) + 2\epsilon |p_\epsilon| \qquad \text{a.e. in } \Omega \tag{53}$$

where

$$\xi_\epsilon(x) = \begin{cases} 0 \text{ if } |y_\epsilon(x) - \psi(x)| > \epsilon^2 \\ 1 \text{ if } |y_\epsilon(x) - \psi(x)| \leq \epsilon^2 \end{cases}$$

and

$$\eta_\epsilon(x) = \begin{cases} 0 \text{ if } y_\epsilon(x) - \psi(x) > -\epsilon^2 \\ 1 \text{ if } y_\epsilon(x) - \psi(x) \leq -\epsilon^2 \end{cases}$$

Since $\beta^\epsilon(y_\epsilon - \psi)\eta_\epsilon = \epsilon^{-1}(y_\epsilon - \psi)\eta_\epsilon + c\epsilon\eta_\epsilon$ remains in a bounded set of $L^2(\Omega)$ and

$$\epsilon^{-1}|y_\epsilon(x) - \psi(x)| \xi_\epsilon(x) \leq \epsilon \qquad \text{a.e. } x \in \Omega$$

we conclude by (42) and (53) that selecting further subsequence,

$$p_\epsilon \beta^\epsilon(y_\epsilon(x) - \psi(x)) \to 0 \qquad \text{a.e. } x \in \Omega$$

whilst by (47)

$$p_\epsilon(x) \to p(x) \qquad \text{a.e. } x \in \Omega$$

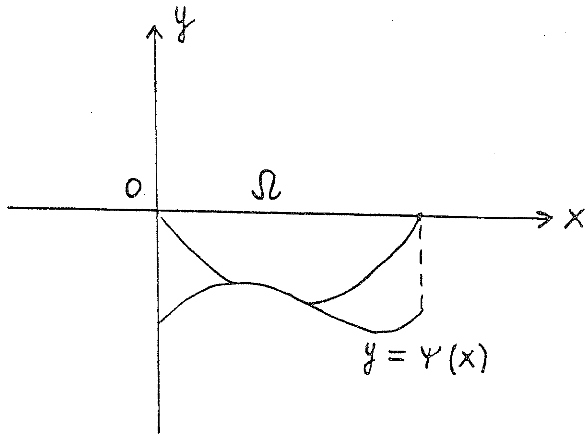

Figure 2

Since $\beta^\epsilon(y_\epsilon - \psi) \to Bu^* + f + \Delta y^*$ weakly in $L^2(\Omega)$ as $\epsilon \to 0$ we infer by the Egorov theorem that

$$p(x)(Bu^*(x) + f(x) + \Delta y^*(x)) = 0 \qquad \text{a.e. } x \in \Omega$$

Next by the obvious inequality

$$(\nabla\phi_\epsilon(y_\epsilon), y_\epsilon - \varphi) \geq \phi_\epsilon(y_\epsilon) - \phi_\epsilon(\varphi) \qquad \forall\, \varphi \in C(\overline{\Omega})$$

we infer that

$$\mu(y^* - \varphi) \geq \phi(y^*) - \phi(\varphi) \quad \forall\, \varphi \in C(\overline{\Omega}), \lim\inf \varphi_\epsilon(y_\epsilon)$$

because $\phi_\epsilon(\varphi) \leq \phi(\varphi)\ \forall\, \varphi \in C(\overline{\Omega})$ and $\lim\inf_{\epsilon\to 0} \phi_\epsilon(y_\epsilon) \geq \phi(y^*)$. Hence $\mu \in \partial\phi(y^*)$ and this completes the proof of Theorem 1.

A model problem: Consider the model of an elastic plane membrane charged along the boundary $\partial\Omega$, which is inflated from above by a vertical field of forces with density u and limited from below by a rigid obstacle $y = \psi(x) < 0$ $\forall\, x \in \Omega$ (See Figure 2) Assume we are given a desired shape of membrane defined by the deflection distribution $y = y_0(x)$. The problem consists of finding the control parameter u subject to the constraint

$$|u(x)| \leq \rho \qquad \text{a.e. } x \in \Omega$$

such that the system response y has a minimum deviation from y_0.

This leads to a problem of the form (18) where $B \equiv I$, $U = L^2(\Omega)$, $f \equiv 0$ and

$$h(u) = \begin{cases} 0 & \text{if } |u(x)| \leq \rho \text{ a.e. } x \in \Omega \\ +\infty & \text{otherwise} \end{cases}$$

This problem has of course at least one solution (y^*, u^*) and by Theorem 1 it satisfies the system

$$\begin{aligned} \Delta y^* + u^* &= 0 && \text{in } \Omega_+ = \{x \in \Omega;\, y^*(x) > \psi(x)\} \\ y^* = \psi,\ \Delta\psi + u^* &\leq 0 && \text{in } \Omega_0 = \Omega\backslash\Omega_+ \\ y^* &= 0 && \text{on } \partial\Omega \end{aligned} \tag{54}$$

$$\begin{aligned} \Delta p &= \mu \qquad \text{in } \Omega_+ \\ p(u^* + \Delta y^*) &= 0 \qquad \text{in } \Omega \qquad (55) \\ u^* &= \rho \operatorname{sign} p \quad \text{a.e.} \quad \text{in } \Omega. \qquad (56) \end{aligned}$$

Here $\mu \in M(\bar{\Omega})$ belongs to $\partial(\|y - y_0\|_{C(\bar{\Omega})})$. Let us assume now that $\Delta\psi \neq \pm\rho$ in Ω. Then $p = 0$ in Ω_0, and so $p \in W_0^{1,q}(\Omega)$ is the solution to problem

$$\begin{aligned} \Delta p &= \mu \qquad \text{in } \Omega_+ \\ p &= 0 \qquad \text{on } \partial\Omega_+ \end{aligned} \tag{57}$$

and $p = 0$ in Ω_0.

Let us assume now that $y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$ is such that

$$\Delta y_0(x) \geq \max\{\Delta\psi(x), \rho\} \qquad \text{a.e. } x \in \Omega$$

Then by equation (54) it follows that $\Delta(y^* - y_0) \leq 0$ in Ω, and so by the maximum principle $y^* - y_0 \geq 0$ in Ω. This implies that the measure μ is positive, and so by (57) it follows that $p \leq 0$ in Ω_+. (We assume that Ω_+ is connected. Otherwise we shall argue on a connected component of Ω_+.) We may conclude therefore that

$$u^* = -\rho \text{ in } \Omega_+,$$

and so y^* satisfies the variational inequality

$$\begin{aligned} \Delta y^* &= \rho \text{ in } \Omega_+, & \Delta y^* &\leq \rho \text{ in } \Omega \\ y^* &= \psi \text{ in } \partial\Omega_+; & y^* &> \psi \text{ in } \Omega_+ \\ y^* &= \psi \text{ in } \Omega_0; & y^* &= 0 \text{ on } \partial\Omega \end{aligned} \tag{58}$$

from which one might determine Ω_+.

Next we discuss a computational implementation of the above results. Let m be the number of points where $\max_\Omega |y - y_0|$ is assumed, and let $x_1, \ldots, x_m$ be these points. Define a measure μ:

$$\mu = \frac{1}{m} \sum_{k=1}^{m} \operatorname{sign}(y - y_0)(x_k)\delta(x - x_k) \tag{59}$$

Then clearly $\mu \in \partial\phi(y)$, since

$$\begin{aligned} \mu(\varphi) &= \frac{1}{m} \sum_{k=1}^{m} \operatorname{sign}(y - y_0)(x_k)\varphi(x_k) \\ &\leq \max\{\operatorname{sign}(y - y_0)(x_0)\varphi(x_0)\},\ x_0 \in M_{y-y_0} \end{aligned}$$

We used both explicit and implicit schemes to solve the system of (54), (56), and (57), with μ as in (59).

(i) Explicit scheme. Starting with an arbitrary (y_n^0, p_n^0) compute:

If $(y_n^k > \psi_n)$ or $(t < t_0)$ then $y_n^{k+1} := y_n^k + \tau\Delta_h y_n^k + \tau\rho \operatorname{sign}_\epsilon p_n^k$

If $y_n^k > \psi_n$ then $p_n^{k+1} := p_n^k + \tau\Delta_h p_n^k - \tau\mu$

If $y_n^k \leq \psi_n$ then $y_n^k := \psi_n$ and $p_n^{k+1} := 0$

(ii) Implicit scheme. Starting with an arbitrary (y_n^0, p_n^0),

(a) Solve the obstacle problem

$$(I - \tau\Delta_h)y_n^{k+1} = y_n^k + \tau\rho \operatorname{sign}_\epsilon p_n^k \quad \text{when } y_n^{k+1} > \psi_n$$
$$(I - \tau\Delta_h)y_n^{k+1} \geq y_n^k + \tau\rho \operatorname{sign}_\epsilon p_n^k$$
$$y_n^{k+1} \geq \psi_n$$

(b) Solve: $(I - \tau\Delta_h)p_n^{k+1/2} = p_n^k - \tau\mu$.

(c) If $y_n^{k+1} > \psi_n$ then $p_n^{k+1} := p_n^{k+1/2}$ else $p_n^{k+1} := 0$.

The following approximation of the delta function was used in d dimensions ($d = 1$ or 2)

$$\delta(x - x_k) = \begin{cases} \dfrac{1}{h^d} & \text{for } x = x_k \\ 0 & \text{at other mesh points} \end{cases}$$

We had implemented the explicit scheme in one dimension ($t_0 = 2.5$) and the implicit scheme in two dimensions (using Guass–Seidel iterations with projection to solve the obstacle problem, see [4]). Both schemes were convergent in the cases where the optimal control is either $u^* \equiv \rho$ or $u^* \equiv -\rho$ for the entire domain Ω. For other controls, y_n^k would oscillate slightly around the optimum state, while the values of p_n^k at some points would oscillate around zero. Actual control u^* at these mesh points can then be chosen by trial and error.

Table 1 Values at Mesh Point's

$y - y^0$				p			
			$k = 880$				
0.016	0.005	0.005	0.012	−0.008	0.027	0.018	0.010
−0.006	0.007	0.007	0.006	0.022	0.022	0.019	0.013
0.014	0.012	0.010	0.014	0.013	0.020	0.020	0.014
0.001	0.013	−0.015	−0.010	0.014	0.013	0.040	0.011
			$k = 920$				
−0.000	0.012	0.013	−0.009	0.015	0.018	0.015	0.033
0.007	0.009	0.010	0.011	0.015	0.020	0.018	0.012
0.007	0.012	0.008	−0.007	0.003	0.017	0.021	0.023
0.004	0.014	0.005	0.010	−0.005	−0.005	0.026	0.011
			$k = 960$				
−0.002	−0.010	−0.006	0.012	−0.008	0.000	0.021	0.018
0.013	0.012	0.010	0.016	−0.003	0.016	0.017	−0.000
0.014	0.010	0.010	0.009	0.006	0.015	0.019	0.015
0.004	−0.004	0.012	−0.017	0.008	0.019	0.017	−0.008
			$k = 1000$				
−0.001	0.005	0.008	−0.021	0.016	0.006	0.013	0.002
−0.007	0.008	0.010	−0.019	0.028	0.015	0.014	−0.003
−0.005	0.012	0.013	0.015	−0.005	0.014	0.016	0.012
−0.009	0.012	0.010	0.009	−0.006	0.013	−0.004	0.017

Example: We had computed the solution of (54), (55), (56) for $\Omega = (0, 1) \times (0, 1)$, $\rho = 8$, $\psi \equiv -1$ and $y_0 = 8x(1 - x)y(1 - y)$, taking $h = 0.2$ and $\tau = 0.0002$. In Table 1 we present the values of $y - y_0$ and p at the mesh points after 880, 920, 960, and 1000 time steps. One sees that y oscillates slightly around the target function (with considerably larger relative discrepancy near $\partial\Omega$). The control u^* should be set equal to 8 at the points (0.4, 0.4), (0.4, 0.6), (0.6, 0.4) and (0.6, 0.6), while at the other mesh points the values of u^* should be chosen between -8 and 8 by trial and error, so that the solution of (54) is close to y_0. The output depended on the choice of y_n^0 and p_n^0. However, all runs led to the same conclusions as above.

Finally, we remark that it is quite feasible to do computations for the above problem with $h = 0.05$.

ACKNOWLEDGEMENTS

This work has been written during the stay of the first author (V.B.) as the Otto Szasz professor at the Department of Mathematical Sciences of the University of Cincinnati. The second author (P.K.) wishes to thank D. Schmidt for many useful discussions related to computations.

REFERENCES

1. V. Barbu, Necessary conditions for nonconvex distributed control problems governed by elliptic variational inequalities, *J. Math. Anal. Appl. 80*(1981), 566–597.
2. V. Barbu, *Optimal Control of Variational Inequalities*, Pitman, London, Boston, Melbourne, 1984.
3. A. Friedman, Optimal control of elliptic variational inequalities, *SIAM J. Control & Optimiz. 25*(1987).
4. R. Glowinski, J.L. Lions, R. Tremolieres, *Numerical Analysis of Variational Inequalities*, North-Holland, 1981.
5. C.U. Huy, P.J. McKenna and W. Walter, Finite difference approximations to the Dirichlet problem for elliptic systems, *Numer. Math. 49*, (1986) 227–237.
6. P. Korman, On computation of solutions of elliptic systems, *Numer. Funct. Anal. and Optimiz. 10*(*9 & 10*), (1989), 977–990.
7. P. Korman, A.W. Leung and S. Stojanovic, Monotone iterations for nonlinear obstacle problem, *J. Austral. Math. Soc. Ser. B 31*, (1990), 259–276.
8. F. Mignot, Controle dans les inéqualities variationnelles elliptiques, *J. Funct. Anal. 22*(1976), 130–185.
9. E. Sinestrari, *Accretive differential operators*. Boll. Unione Mat. Ital. *13*(1976), 19–31.
10. G.D. Smith, *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, Oxford University Press (3rd Edition), 1989.
11. G. Stampacchia, Le Problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinues, *Ann. Inst. Fourier 15*(1965), 189–258.