MATHEMATICA – REVUE D'ANALYSE NUMÉRIQUE
ET DE THÉORIE DE L'APPROXIMATION

MATHEMATICA, Tome 29(52), N° 2, 1987, pp. 157 – 167

A STEEPEST DESCENT METHOD FOR THE APPROXIMATION OF THE BOUNDARY CONTROL IN TWO-PHASE STEFAN PROBLEM

P. NEITTAANMÄKI and D. TIBA

1. Introduction

Consider the boundary control problem

$$\text{(P)}\qquad \text{Minimize}\left\{\pi(u) = \int_0^T \left[\frac{1}{2}|y - d|^2_{L^2(\Omega)} + \frac{1}{2}|u|^2_{L^2(\partial\Omega)}\right] \mathrm{d}t\right\}$$

over all $u \in L^2(\Sigma)$ and $y = y(u) \in L^2(0, T; H^1(\Omega))$
subject to

$$\frac{\partial}{\partial t} v(t, x) - \Delta y(t, x) = f(t, x) \qquad \text{in } Q$$

$$(1.1)\qquad v(t, x) \in \beta(y(t, x)) \qquad \text{in } Q,$$

$$(1.2)\qquad \frac{\partial y(t, x)}{\partial n} = u(t, x) \qquad \text{on } \Sigma,$$

$$(1.3)\qquad v(0, x) = v_0(x) \qquad \text{on } \Omega.$$

In the above $\Omega \subset \mathbb{R}^n$, $n \geqslant 1$, is a bounded domain with smooth boundary and $Q =]0, T[\times \Omega$ is a cylinder with lateral face Σ.

We assume that $v_0 \in L^2(\Omega)$, $d \in L^2(Q)$ and that β is a strongly maximal monotone graph in $\mathbb{R} \times \mathbb{R}$, bounded on bounded sets. When β is given by

$$(1.4)\qquad \beta(r) = \begin{cases} r - r_0, & r > r_0 \\ [-\delta, 0], & r = r_0 \\ k(r - r_0) - \delta, & r < r_0 \end{cases}$$

where $k,\ \delta > 0$, we obtain a two-phase Stefan problem (see [6], p. 196).

A similar control process is considered in [13] in connection with some problems arising in metallurgy. In the paper [9] the case of differentiable control is studied and in [7] necessary optimality conditions are obtained for the problem (P) with distributed control.

A finite element discretization of two-phase Stefan problems is discussed in [15, 16] and the control problem is similarly treated in [8]. For further references in connection with Stefan problems see [3, 6, 10].

In this paper we analyze a regularization of problem (P) which can be mainly compared with the works [1, 2]. It consists of replacing (P) by a family of smooth problems and afterwards tending to the limit with the approximate control (see Proposition 2.3 and Theorem 2.4).

We shall also present an algorithm for finding a computer solution for problem (P). Due to the lack of convexity the emphasis will be on the descent property, not on the convergence properties of the algorithm. To obtain the numerical solution of the state and the adjoint system finite elements in the space and finite differences in time are used.

The plan of the paper is as follows. In section 2 we briefly discuss the existence and the regularization of problem (P). Section 3 containts the main results on the descent property of the gradient method. In the last part a numerical example is discussed.

2. Existence and regularization

We will briefly outline the existence of an $L^2(\Sigma)$ optimal control for problem (P). Next, the approximation properties of the regularized controls are given. For more details, we quote [14].

Denote $V = H^1(\Omega)$, $H = L^2(\Omega)$ with scalar product $(\cdot,\cdot)$ and norm $\|\cdot\|$, V^* is the dual of V. Equation (1.1) — (1.3) can be written in an abstract form as

$$\frac{dv}{dt} + Ay = f, \quad v(t) \in \beta(y(t)) \qquad \text{a.e.}[0, T] \tag{2.1}$$

$$v(0) = v_0 \tag{2.2}$$

The function $f \in L^2(0, T; V^*)$ is given by

$$\int_0^T (f(t), \psi(t))dt = \int_0^T\int_\Gamma u \cdot \psi d\Gamma dt, \ \forall \psi \in L^2(0, T; V) \tag{2.3}$$

Operator $A : V \to V^*$ is defined by

$$(Ay, z) = \int_\Omega \text{grad } y \cdot \text{grad } z \, dx, \qquad \forall y, z \in V. \tag{2.4}$$

and operator $B : H \to H$ is the realization of β in $L^2(\Omega)$.

The existence of the solution for equation (2.1), (2.2) is studied, for example in paper [3], where A and B may be both nonlinear.

THEOREM 2.1. *Let $u_n \to u$ weakly in $L^2(\Sigma)$. Then $y_n \to y$ weakly in $L^2(0, T; V)$, where y_n, y are the solutions of* (2.1), (2.2) *corresponding to u_n, u.*

From this result, one obtains at once

THEOREM 2.2 *There is an optimal pair* $[u^*, y^*]$ *in* $L^2(\Sigma) \times L^2(0, T; V)$ *for problem* (P).

Consider the regularized problem

$$(P_\varepsilon) \qquad \text{Minimize} \left\{ \pi^\varepsilon(u) = \int_0^T \left[\frac{1}{2} |y - d|_H^2 + \frac{1}{2} |u|_H^2 \right] dt \right\}$$

subject to

$$(2.5) \qquad \frac{\partial \beta^\varepsilon(y(t, x))}{\partial t} - \Delta y(t, x) = f(t, x) \qquad \text{in } Q,$$

$$(2.6) \qquad \frac{\partial}{\partial n} y(t, x) = u(t, x) \qquad \text{on } \Sigma,$$

$$(2.7) \qquad y(0, x) = y_0(x) \qquad \text{on } \Omega,$$

where we define

$$(2.8) \qquad \beta^\varepsilon(y) = y + \int_{-\infty}^{\infty} \gamma_\varepsilon(y - \varepsilon^2\theta)\, \rho(\theta)\, d\theta$$

and γ_ε is the Yosida approximation of the maximal monotone graph $\gamma(y) = \beta(y) - y$ (we assume for convenience that $k \geq 1$ in (1.4)), and ρ is a Friedrichs mollifier, such that $\rho \in C_0^\infty(R)$, supp $\rho \subset (-1, 1)$, $\rho(-\theta) = \rho(\theta)$ and $\int_{-\infty}^{\infty} \rho(\theta) d\theta = 1$. Obviously the problem (P_ε) has an optimal pair $[y_\varepsilon, u_\varepsilon] \in L^2(Q) \times L^2(\Sigma)$.

PROPOSITION 2.3 *The subsequences converge as follows*

$$(2.9) \qquad u_\varepsilon \to u^* \quad \textit{strongly in } L^2(\Sigma),$$

$$(2.10) \qquad y_\varepsilon \to y^* \quad \textit{strongly in } L^2(Q),$$

The corresponding convergence result for the cost functional is

THEOREM 2.4. *The sequence* $\pi(u_\varepsilon) \to \pi(u^*)$, *the optimal value of problem* (P), *when* $\varepsilon \to 0$, *and therefore* $\{u_\varepsilon\}$, *is a minimizing sequence for* (P).

3. The descent property

In order to obtain a suboptimal control for (P), by Theorem 2.4 one may solve problem (P_ε). Due to the good differentiability properties in (P_ε), a gradient algorithm can be utilized to find u_ε efficiently.

We denote by $\theta_\varepsilon : L^2(\Sigma) \to L^2(Q)$ the mapping $u \to y$ given by (2.5) – (2.7).

PROPOSITION 3.1. *For all* $u \in L^2(\Sigma)$ *there exists a linear operator* $\nabla \theta_\varepsilon(u) : L^2(\Sigma) \to L^2(Q)$ *defined by*:

$$\nabla \theta_\varepsilon(u)\, v = \operatorname*{weak-lim}_{\lambda \to 0} \frac{\theta_\varepsilon(u + \lambda v) - \theta_\varepsilon(u)}{\lambda} \tag{3.1}$$

for all $v \in L^2(\Sigma)$. *Moreover*

$$\nabla \theta_\varepsilon(u) v = \frac{\partial z}{\partial t} \tag{3.2}$$

where z *is the solution of the problem*

$$\nabla \beta^\varepsilon(\theta_\varepsilon(u)) \frac{\partial z}{\partial t} + Az = h \quad \text{a.e. } [0, T] \tag{3.3}$$

$$z(0) = 0. \tag{3.4}$$

In *equation* (3.3) $h \in W^{1,2}(0, T; V^*)$ *satisfies*

$$h(t) = \int_0^t g_1(\xi) d\xi + v_0 \tag{3.5}$$

$$\int_0^T (g_1(t)\, \psi(t)) dt = \int_0^T \int_{\partial\Omega} v\, \psi\, d\sigma\, dt \text{ for every } \psi \in L^2(0, T;\, V). \tag{3.6}$$

Proof. Denote B^ε as the realization in H of β^ε and $y_\lambda = \theta_\varepsilon(u + \lambda v)$, $y = \theta_\varepsilon(u)$. Then, by the definition of solution we get

$$B^\varepsilon\left(\frac{dw_\lambda}{dt}\right) + A\, w_\lambda = g + \lambda h \qquad \text{a.e. } [0, T], \tag{3.7}$$

$$B^\varepsilon\left(\frac{dw}{dt}\right) + Aw = g \qquad \text{a.e. } [0, T], \tag{3.8}$$

$$w_\lambda(0) = w(0) = 0. \tag{3.9}$$

Here $g(t) = \int_0^t f(\xi)\, d\xi + v_0$, f given by (2.3)

and $w_\lambda(t) = \int_0^t y_\lambda(\xi)\, d\xi$, $w(t) = \int_0^t y(\xi)\, d\xi$

Subtract the two relations and multiply by $\frac{dw_\lambda}{dt} - \frac{dw}{dt}$:

$$\int_0^t \left| \frac{dw_\lambda}{dt} - \frac{dw}{dt} \right|_H^2 ds + \frac{1}{2} (A(w_\lambda(t) - w(t)),\ w_\lambda(t) - w(t))$$

$$\leq \lambda \int_0^t \left(h, \frac{dw_\lambda}{dt} - \frac{dw}{dt} \right) ds.$$

Then $\frac{dw_\lambda}{dt} \to \frac{dw}{dt}$ and $w_\lambda \to w$ strongly in $L^2(0, T; H)$, $C(0, T; H)$ respectively.

We set $z_\lambda = \frac{w_\lambda - w}{\lambda}$, that is

$$\int_0^t \left| \frac{dz_\lambda}{dt} \right|_H^2 ds + \frac{1}{2}(Az_\lambda(t), z_\lambda(t)) \leqslant \int_0^t \left(h, \frac{dz_\lambda}{dt} \right) ds.$$

Integrating by parts in the right hand side we obtain $\{z_\lambda\}$, $\left\{\frac{dz_\lambda}{dt}\right\}$ bounded in $L^\infty(0, T; V)$, $L^2(0, T; H)$. Since B^ε is Lipschitz, the Lebesque theorem shows that

$$\frac{B^\varepsilon\left(\frac{dw_\lambda}{dt}\right) - B^\varepsilon\left(\frac{dw}{dt}\right)}{\lambda} = \frac{B^\varepsilon\left(\frac{dw_\lambda}{dt}\right) - B^\varepsilon\left(\frac{dw}{dt}\right)}{\frac{dw_\lambda}{dt} - \frac{dw}{dt}} \cdot \frac{dz_\lambda}{dt}$$

is weakly convergent in $L^2(0; T; H)$ to $\nabla B^\varepsilon\left(\frac{dw}{dt}\right) \cdot \frac{dz}{dt}$, where z is such that $z_\lambda \to z$ strongly in $C(0, T; H)$. We can pass to the limit and obtain (3.2) — (3.4) to finish the proof.

Now, we can define the adjoint system for the control problem (P_ε):

(3.10) $$\Delta\beta^\varepsilon(y_\varepsilon)\frac{\partial p_\varepsilon}{\partial t} - Ap_\varepsilon = y_\varepsilon - d \qquad \text{a.e. in } [0, T]$$

(3.11) $$p_\varepsilon(T) = 0$$

The gradient algorithm for solving problem (P_ε) is obvious (for brevity we omit the subindex ε):

Algorithm 3.2.

Step 1. Choose any u_0 and set $n := 0$.
Step 2. Compute y_n by solving (2.5) — (2.7).
Step 3. Test in the pair $[y_n, u_n]$ is satisfactory; if YES then STOP; otherwise GO TO step 4.
Step 4. Compute p_n by (3.10) — (3.11).
Step 5. Compute u_{n+1} by equation
(3.12) $u_{n+1} = u_n - \rho_n(u_n - p_{n|\Sigma})$, where ρ is an appropriate real parameter.
Step 6. Set $n := n + 1$ and GO TO step 2.

The convergence test involved in step 3 is the difference $|u_n - p_{n|\Sigma}|$ which is to be smaller than a given parameter. In step 5 the parameter ρ_n can for example be selected by utilizing a line search.

It is known that without convexity assumptions, the above gradient algorithm may be convergent only to a stationary point of the functional (see [4], [12]). Since the state equation is nonlinear, the cost functional is no more convex and our result underlines the descent property of (3.12).

THEOREM 3.3

(i) *Let ε be fixed. The sequence $\pi_\varepsilon(u_n)$ is convergent, when $n \to \infty$,* (ii) *Let $\tilde{u}_\varepsilon$ be the approximate value of $\tilde{u}_\varepsilon$ as computed by* Algorithm 3.2. *The sequence $\pi_\varepsilon(\tilde{u}_\varepsilon)$ is bounded with respect to ε and every cluster point $\tilde{\pi}$ satisfies*

$$\tilde{\pi} \leqslant \pi(u_0) \tag{3.13}$$

where u_0 is the first iteration.

Proof. (i) The sequence $\{\pi_\varepsilon(u_n)\}$ decreases and it is bounded by $\pi_\varepsilon(u_0)$ and $\pi_\varepsilon(u_\varepsilon)$.
(ii) We have

$$\pi_\varepsilon(u_\varepsilon) \leqslant \pi_\varepsilon(\tilde{u}_\varepsilon) \leqslant \pi_\varepsilon(u_0) \tag{3.14}$$

and, by an easy consequence of Theorem 2.4, $\pi_\varepsilon(u_\varepsilon) \to \pi(u^*)$.
We will show that $\pi_\varepsilon(u_0) \to \pi(u_0)$ too. This is equivalent to

$$y_\varepsilon = \theta_\varepsilon(u_0) \to y \text{ strongly in } L^2(Q),$$

where y is the solution of (1.1) – (1.3) corresponding to u_0. Let $w_\varepsilon(t) = \int_0^t y_\varepsilon(\xi)\,\mathrm{d}\xi$. Then

$$B^\varepsilon\left(\frac{\mathrm{d}w_\varepsilon}{\mathrm{d}t}\right) + Aw_\varepsilon = g_0 \qquad \text{a.e. } [0, T] \tag{3.15}$$

$$w_\varepsilon(0) = 0$$

with $g_0(t) = \int_0^t f_0(\xi)\,\mathrm{d}\xi + v_0$ and

$$\int_0^T (f_0(t), \psi(t))\mathrm{d}t = \int_0^T \int_\Gamma u_0 \cdot \psi \mathrm{d}\Gamma\, \mathrm{d}t$$

for every $\psi \in L^2(0, T; V)$.

Multiply (3.15) by $\left\{\frac{\mathrm{d}w_\varepsilon}{\mathrm{d}t}\right\}$. Then we obtain $\{w_\varepsilon\}$, $\left\{\frac{\mathrm{d}w_\varepsilon}{\mathrm{d}t}\right\}$ bounded in $L^\infty(0, T; V)$, $L^2(0, T; V)$ respectively

Since B is supposed be bounded on bounded sets we get $\left\{B^\varepsilon\left(\frac{\mathrm{d}w_\varepsilon}{\mathrm{d}t}\right)\right\}$ bounded in $L^2(0, T; H)$. Next, subtract two equations (3.15) and multiply by $\frac{\mathrm{d}w_\varepsilon}{\mathrm{d}t} - \frac{\mathrm{d}w_\sigma}{\mathrm{d}t}$. By (2.8) we get :

$$\int_0^t \left|\frac{\mathrm{d}w_\varepsilon}{\mathrm{d}t} - \frac{\mathrm{d}w_\sigma}{\mathrm{d}t}\right|_H^2 + \int_0^t \int_\Omega \left(\gamma^\varepsilon\left(\frac{\mathrm{d}w_\varepsilon}{\mathrm{d}t}\right) - \gamma^\sigma\left(\frac{\mathrm{d}w_\sigma}{\mathrm{d}t}\right), \frac{\mathrm{d}w_\varepsilon}{\mathrm{d}t} - \frac{\mathrm{d}w_\sigma}{\mathrm{d}t}\right) + \frac{1}{2}\,|\nabla w_\varepsilon(t) - w_\sigma(t)|_H^2 = 0. \tag{3.16}$$

Here $\gamma^\varepsilon(y) = \beta^\varepsilon(y) - y$, i.e. the second term in (2.8) and $\left\{\gamma^\varepsilon\left(\frac{\mathrm{d}w_\varepsilon}{\mathrm{d}t}\right)\right\}$ is bounded in $L^2(Q)$.

Taking into account the properties of the Yosida approximation : $\gamma_\varepsilon(y) \in \gamma((I + \varepsilon\gamma)^{-1}(y))$, $\varepsilon\gamma_\varepsilon(y) = y - (I + \varepsilon y)^{-1}(y)$ and the above boundedness, one can infer from (3.16) that $\{w_\varepsilon\}$, $\left\{\frac{\mathrm{d}w_\varepsilon}{\mathrm{d}t}\right\}$ are Cauchy sequences in $L^2(0, T; V)$ and $L^2(Q)$ respectively. Now, it is possible to pass to the limit in (3.15) and to finish the proof.

Remark 3.4. The practical meaning of Theorem 3.3 is that in a given problem one should take the first iteration as the control u_0 already used in practice. Next the algorithm improves the performance given by it.

4. A numerical example

The regularized state problem (2.5) -- (2.8) and the adjoint state problem are discretized by applying the finite difference method in time and the finite element method in space. Concerning the convergence and stability of such a discretization method we refer to [10, 16].

To illustrate the efficiency of Algorithm 3.2 the following numerical example is considered :

$$\Omega =]0, 1[\times]0, 1[$$
$$T = 1.$$

Let

$$\beta(y) = \begin{cases} y & y < 0, \\ [0, 2] & y = 0, \\ 4y + 2 & y > 0, \end{cases} \tag{4.1}$$

$$f(t, x_1, x_2) = \begin{cases} 8(2\mathrm{e}^{-2t} - 1), & x_1^2 + x_2^2 > \mathrm{e}^{-2t} \\ 2(\mathrm{e}^{-2t} - 2), & x_1^2 + x_2^2 \leqslant \mathrm{e}^{-2t} \end{cases} \tag{4.2}$$

$$v_0 = \beta(y_0) \tag{4.3}$$

and

$$y_0 = \begin{cases} x_1^2 + x_2^2 - 1, & x_1^2 + x_2^2 < 1 \\ 2(x_1^2 + x_2^2 - 1), & x_1^2 + x_2^2 \geqslant 1 \end{cases}. \tag{4.4}$$

For the boundary control

$$u(t, x_1, x_2) = \begin{cases} 0 & \text{if } x_1 = 0, \text{ or } x_2 = 0 \\ 4 & \text{on the remaining of } \partial\Omega \end{cases} \tag{4.6}$$

The exact solution y, of (1.1) — (1.3) with given data (4.1) — (4.4) is

$$y(t, x_1, x_2) = \begin{cases} 2(x_1^2 + x_2^2 - e^{-2t}) & x_1^2 + x_2^2 > e^{-t} \\ x_1^2 + x_2^2 - e^{-t}, & x_1^2 + x_2^2 \leq e^{-t} \end{cases}.$$

Consider the cost functional

$$\pi_\lambda(u) = \frac{1}{2}\int_0^1 \left[|y|_{L^2(\Omega)} + \frac{\lambda}{2} |u|^2_{L^2(\partial\Omega)} \right] dt \quad \text{with} \quad \lambda = 0.1.$$

We shall now test the efficiency of different variants of Algorithm 3.2. The nonlinear programming methods tested are

— steepest descent Algorithm 3.2

— a conjugate gradient method with an automatic restart ([11], ZXCGR of IMSL Subroutine Library)

— a bundle algorithm due to C. Lemarechal (BCG), [5].

We have chosen $\Delta t = 1/16$ (time step) and 64 triangular linear elements in discretization of state and adjoint problem. For more details about *FE*-method and algorithms see [8, 10, 16]. For simplicity, we have replaced β^ε by a piecewise linear function such as

$$\beta_\varepsilon(y) = \begin{cases} y & , \quad y < 0 \\ \dfrac{2 + y_\varepsilon}{\varepsilon}, & y \in [0, \varepsilon] \\ 4y + 2, & y > \varepsilon \end{cases}$$

for $\varepsilon = 1/16$ (with appropriate smoothing for $y = 0$ and $y = \varepsilon$). In Table 4.1 we see the diminution of π_λ per iteration when three different gradient algorithms have been applied.

Table 4.1. Comparision of different gradient algorithms

Number of iteration	Value of $\pi_\lambda(u^n)$ for different gradient algorithms		
	steepest descent	ZXCGR	BCG
0	2.166	2.166	2.166
1	.426	.418	.935
2	.203	.148	.681
3	.124	.116	.252
4	.110	—	.208
5	.101	—	.144
6	.091	—	.142
7	.090	—	.102
CPU(seconds)	840	181	488

The optimal control found by different gradient algorithms is roughly speaking the same.

In Figures 4.2 — 4.4 we can see the boundary controls and corresponding temperature distributions obtained by Algorithm 3.2 at time levels $t = .325$, $t = .625$ and for $t = .935$.

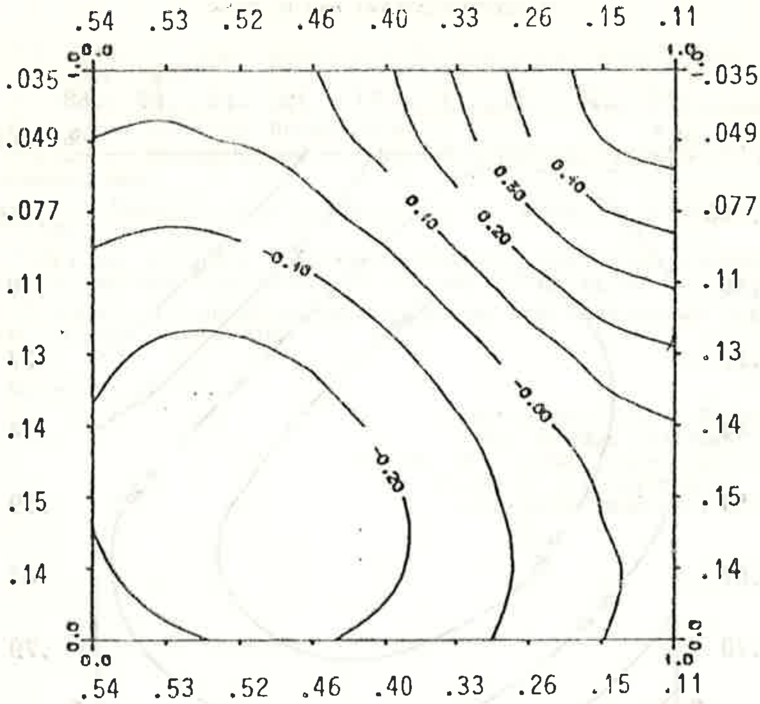

Fig. 4.2. — u_h and y_h for $t = .325$

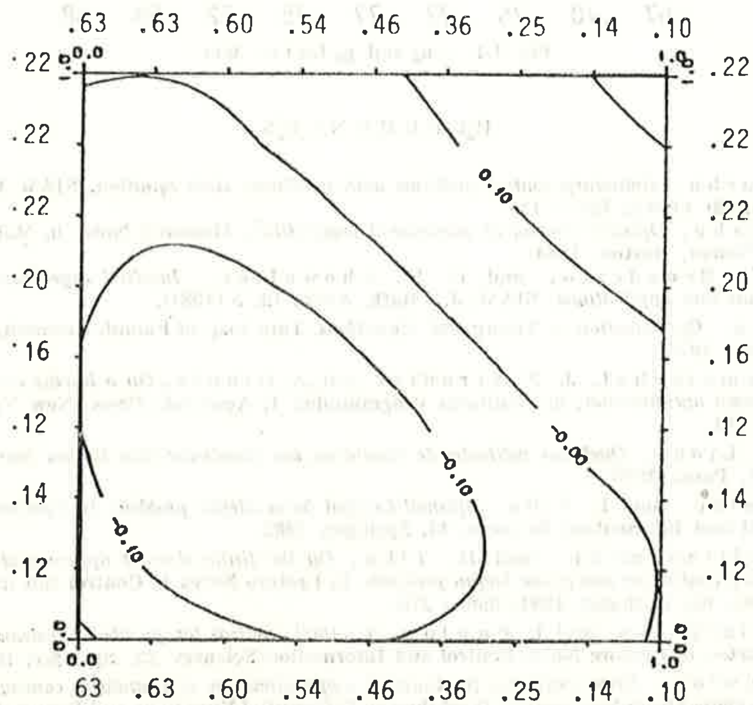

Fig. 4.3. — u_h and y_h for $t = .625$

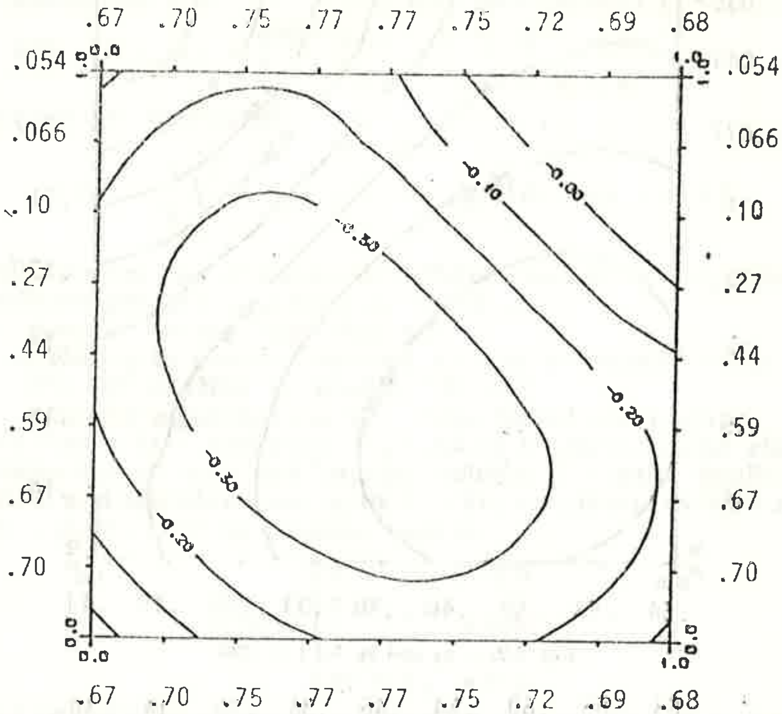

Fig. 4.4. — u_h and y_h for $t = .935$

REFERENCES

1. V. Barbu, *Boundary control problems with nonlinear state equation*, SIAM J., Control Optim., 20 (1982), 125 — 143.
2. V. Barbu, *Optimal control of variational inequalities*, Research Notes in Mathematics, 100, Pitman, Boston, 1984.
3. E. Di Benedetto, and R. E. Showalter, *Implicit degenerate evolution equations and applications*, SIAM, J., Math. Anal., 12, 5 (1981).
4. J. Cea, *Optimization — Theory and Algorithms*, Tata Inst. of Found. Research, Springer, Bombay, 1978.
5. C. Lemaréchal, J. J. Strodiot and A. Bihain, *On a bundle algorithm for nonsmooth optimization*, in Nonlinear Programming 4, Academic Press, New York, 1981, 245 — 281.
6. J. L. Lions, *Quelques méthodes de resolution des problemes aux limites non linéaires*, Dunod, Paris, 1969.
7. Z. Meike and T. Tiba, *Optimal Control for a Stefan problem*, in Lecture Notes in Control and Information Sciences, 44, Springer, 1982.
8. P. Neittaanmäki and D. Tiba, *On the finite element approximation of the boundary control for two-phase Stefan problems*, in Lecture Notes in Control and Information Sciences, 62, Springer, 1984, 356 — 370.
9. M. Niezgódka and I. Pawlow, *Optimal control for parabolic systems with free boundaries*, in Lecture Notes Control and Information Sciences, 22, Springer, 1979.
10. I. Pawlow, *Error estimates for Galerkin approximation of boundary control problems for two-phase Stefan type process*. Workshop on Differential Equations and Boundary Control, INCREST, 1983, 129 — 147.

11. M. J. D. Powell, *Restart procedures for the conjugate gradient method*, Math. Programming 12 (1977), 241 – 254.

12. B. N. Pschenichny and Yu. M. Danilin, *Numerical methods in extremal problems*. MIR Publishers, Moscow, 1978.

13. C. Saquez, *Contrôle optimal de systemès à frontiere libre*, Thèse, Univ. Technologique Compiegne, 1980.

14. D. Tiba, *Boundary control for a Stefan problem*, ISNM 68, Birhäuser, Basel, 1984, 229 – 242.

15. D. Tiba and M. Tiba, *Regularity of the boundary data and the convergence of the finite element discretization in two-phase Stefan problems*, Int. J. Eng. Sci. (to appear).

16. M. Zlamal, *A finite element solution of the nonlinear heath equation R.A.I.R.O.* Num. Anal., 14 (1980), 203 – 216.

Received 16. II. 1985

Lappeenranta Univ. of Techn.
Dept. of Phys. and Math.
Box 20, SF. — 53851 Lappeenranta Finland
INCREST — BUCUREȘTI
Dept. Math. ROMANIA