Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique

U. Mosco
F. Scarpini

Complementarity systems and approximation of variational inequalities

Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique, tome 9, n° R1 (1975), p. 83-104
<http://www.numdam.org/item?id=M2AN_1975__9_1_83_0>

R.A.I.R.O.
(9e année, R-1, 1975, p. 83 à 104)

COMPLEMENTARITY SYSTEMS AND APPROXIMATION OF VARIATIONAL INEQUALITIES

by U. MOSCO and F. SCARPINI (*)

Communiqué par E. MAGENES.

Abstract. *We discuss the numerical approximation of some complementarity systems in Sobolev spaces, which occur in the variational Dirichlet problem with unilateral constraints.*

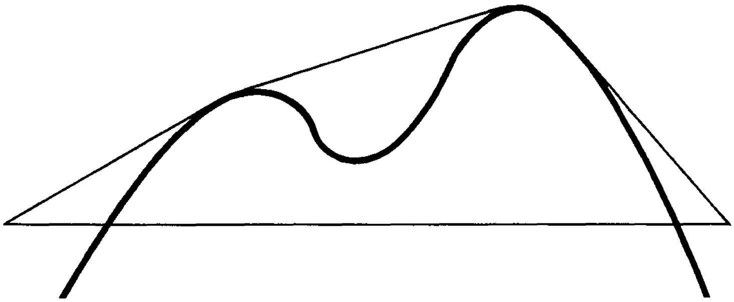
I – THE UNILATERAL DIRICHLET PROBLEM AS A COMPLEMENTARITY SYSTEM.

Let Ω be a bounded open subset of $\mathbb{R}^n$ and ψ a given real valued function defined in Ω, with $\psi \leqslant o$ on the boundary $\delta\,\Omega$ of Ω.

We want to find a function u, vanishing on $\delta\,\Omega$, which is superharmonic and greater or equal than ψ in Ω and is harmonic in the region of Ω where it does not "touch" the "obstacle" ψ, that is, where $u > \psi$.

Simple one dimensional examples show that, even if ψ is very smooth, a function u with all the properties above will have in general discontinuous *second order* derivatives at the boundary of the "contact set" $u = \psi$, as it can be intuitively checked on the following figure.:

(*) Istituto Matematico, Università di Roma.
This paper has been partially supported by GNAFA-CNR.



Thus, if $-\Delta$ is the Laplace operator, we cannot expect $-\Delta u$ to be defined pointwise in Ω in the classical sense. However, we can always think of $-\Delta u$ as a *measure* μ in Ω, as we shall see more precisely below.

The problem we have in mind can then be formally stated as follows :

$$(1) \qquad \begin{aligned} &u - \psi \geqslant 0 \text{ and } \mu = -\Delta u \geqslant 0 \text{ in } \Omega, \\ &u = 0 \quad \text{on } \delta\Omega, \\ &(u - \psi, \mu) = 0 \quad , \end{aligned}$$

where the vanishing of the pairing between the function $u - \psi$ and the measure μ has been set just to impose that u be harmonic, i.e. $\mu = 0$, where $u > \psi$.

If ψ is continuous on $\overline{\Omega}$ the problem above can be dealt with in the framework of the classical theory of superharmonic functions, see [30] .

Here, however, we shall follow the *variational inequality* approach, due to J.L. Lions and G. Stampacchia, [24] , [39] .

The function ψ and the solution u are now required to have a finite energy integral and problem (1) is then formulated in a weak sense in the Sobolev space $H^1(\Omega)$ (1). Namely, if we introduce the Dirichlet bilinear form :

$$a(u,v) = \sum_{i,j}^{1,n} \int_\Omega u_{x_i} \, v_{x_j} \, dx$$

(1) $H^1(\Omega)$ is the space of all functions $v \in L^2(\Omega)$, whose distribution derivatives v_{x_i} also belong to $L^2(\Omega)$, normed by :

$$\|v\|_{H^1} = \left(\|v\|^2_{L^2} + \sum_{i=1}^{n} \|v_{x_i}\|^2_{L^2}\right)^{1/2}$$

The closure of $C_0^1(\Omega)$ in $H^1(\Omega)$ is the Sobolev space $H_0^1(\Omega)$.

and define K to be the convex cone

$$K = \{ v \in H_0^1 (\Omega) \ / \ v \geqslant \psi \quad \text{a e in } \Omega \},$$

then our problem becomes

$$(2) \qquad u \in K \quad a(u, v-u) \geqslant 0, \quad \forall v \in K$$

The existence and uniqueness of the solution u of such a variational inequality is by now a well known result, see ref [24] , [39] , quoted above

We may wonder, however, in which sense the solution u of (2) is the solution of our original problem (1) Let us also note that (1) (as (1') below) may be regarded as a continuous analogue of so called complementarity systems that will play a basic role in the numerical approximation we shall later discuss Therefore, it is perhaps convenient to describe the equivalence between (1) and (2) with some more details (2)

If ψ itself vanishes on $\partial \Omega$, then it is not difficult, at the same time, to give a precise meaning to the conditions (1) and show that they actually characterize the solution u of (2) Indeed, by assuming $\psi \in H_0^1 (\Omega)$ it can be easily proved that a function $u \in H_0^1 (\Omega)$ is a solution of (2) if and only if u and $\mu = -\Delta u$ are solution of (1) Here $\mu = -\Delta u$ is well defined as an element of the dual space $H^{-1}(\Omega)$ of $H_0^1 (\Omega)$, hence the pairing between $u - \psi$ and μ, appearing in (1), has the natural meaning of that duality

When the obstacle does not vanish on $\partial\Omega$, some attention must be paid to the precise meaning of this pairing

If we assume that

$$\psi \in C(\bar{\Omega}) \cap H^1 (\Omega) \quad , \quad \psi < 0 \text{ on } \partial\Omega$$

then it is easy to show that the measure ψ has a compact support in Ω This clearly gives a well defined meaning to the pairing $(u - \psi, \mu)$ even if now the function $u - \psi$ dos not vanish on the boundary of Ω (3) Again it can be easily proved that $u \in H_0^1 (\Omega)$ is a solution of (2) if and only if the pair $u \in H_0^1 (\Omega)$, $\mu \in H^{-1}(\Omega)$, μ with a compact support, is a solution of (1)

When $\psi \leqslant 0$ on Ω, then the support of μ may well reach the boundary of Ω In this case, some regularity of u is needed in order to well define the pairing $(u - \psi, \mu)$

---

(2) The relation between variational inequalities and complementarity systems has been studied in a more general setting [34] and widely investigated in [11] , [18] , [19]

(3) Indeed, for any $v \in H^1 (\Omega)$ we may define (v,μ) as $(\alpha v,\mu)$, where α is some smooth function, with a compact support in Ω, which is $\equiv 1$ on the support of μ

Let us recall at this point the following well known regularity result for problem (2), see [3] , [22] : If $\psi \in C(\bar{\Omega}) \cap H^2(\Omega)$, $\psi \leqslant 0$ on a smooth $\partial\Omega$, then $u \in H_0^1(\Omega) \cap H^2(\Omega)$ and

$$\|u\|_{H^2} \leqslant C \quad \|\psi\|_{H^2} \tag{3}$$

If we assume that :

$$\psi \in C(\bar{\Omega}) \cap H^2(\Omega) \quad , \quad \psi \leqslant 0 \quad \text{on } \partial\Omega \quad ,$$

then in consequence of the above result, both $u - \psi$ and $\mu = -\Delta u$ "are" L^2 functions in Ω and the pairing $(u - \psi, \mu)$ has an abvious meaning. Again it can be shown, in this case, that (1) and (2) are equivalent, in the sense that a function $u \in H_0^1(\Omega) \cap H^2(\Omega)$ is the solution of (2) if and only if u and $\mu = -\Delta u \in L^2(\Omega)$ are solutions of (1) (4) :

Let us finally remark that by introducing the function

$$U = u - \psi$$

problem (1) can be also written as :

$$\begin{aligned} & U \geqslant 0 \ , \quad \mu = -\Delta U + \nu \geqslant 0 \quad \text{in } \Omega \\ & U = -\psi \quad \text{on } \partial\Omega \\ & (U, \mu) = 0 \quad , \end{aligned} \tag{1'}$$

where

$$\nu = -\Delta\psi \ .$$

2 – A QUALITATIVE IDEA OF AN ALGORITHM FOR SOLVING PROBLEM (2).

Let us suppose that $\psi \in C(\bar{\Omega}) \cap H^1(\Omega)$ $\psi \leqslant 0$ on $\partial\Omega$, and that

$$\nu = -\Delta\psi$$

is a (signed) measure in Ω. Then H. Lewy and G. Stampacchia [23] have proved that the measure $\mu = -\Delta u$, solution of (1) satisfies :

$$0 \leqslant \mu(F) \leqslant \nu^+(F) \qquad \forall \text{ Borel set } F \subset \Omega \tag{4}$$

where ν^+ is the positive part of ν, see also [36]. This result is stronger than the estimate (3), which clearly follows from (4) and the classical regularity results for the Dirichlet problem once we know that $\psi \in H^2(\Omega)$.

Let us now compare the solution u with the given obstacle ψ, by taking the estimate (4) into account. We see that in the regions of Ω where the function ψ is subharmonic (i.e., where $\nu^+ = 0$), there the function u is harmonic (i.e., $\mu = 0$), while in the remaining region, where ψ and u are both superharmonic, the positive measure μ is majorized by $\nu = \nu^+$.

---

(4) We can replace ψ by $\psi - \varepsilon$, $\varepsilon > 0$, use the preceding result and then let $\varepsilon \to 0$.

Let us now try to devise a procedure yielding the solution u by means of progressive modifications of the obstacle itself. By the remark above, we are induced to set out an algorithm that gradually reduces the *negative* part of ν to zero, while "releasing" at the same time ν^+ just to reach the solution μ. Such a reduction of ν to μ, which is to say, of ψ to u, could be conceivably achieved by an iterative process that, starting with an initial configuration $u_o = \psi$ brings to successive configurations $u_1, u_2, \ldots\ldots$, obtained at each step by moving from a given u_{n-1}, harmonic in a region Q_{n-1}, subharmonic in a larger region Q_n and actually *strictly* subharmonic somewhere in $P_n = Q_n - Q_{n-1}$, to a new configuration u_n *harmonic* in the whole region, Q_n. To be more definite, let us suppose that $\psi = 0$ on $\partial\Omega$ and let us define initially :

$$u_o = \psi \quad , \text{ i.e. } \quad , \quad \mu_o = \nu .$$

If $\mu_o \geqslant 0$, then $u = u_o$ is the solution. Let us suppose, instead that u_o is strictly subharmonic on a subset of on open region P_1, i.e., $\mu_o \leqslant 0$ on P_1 with $\mu_o(F) < 0$ for some borel subset $F \subset P_1$. Let us also suppose that P_1 has a smooth boundary. We can then solve the Dirichlet problem :

$$(5) \qquad -\Delta u_1 = 0 \text{ in } Q_1 = P_1 \ , \ u_1 = \psi \text{ in } \Omega - Q_1 \ .$$

If u_1, as a function defined on the whole of Ω, is such that :

$$\mu_1 = -\Delta u_1 \geqslant 0 \quad \text{in } \Omega ,$$

then $u = u_1$, $\mu = \mu_1$ is the solution we looked for [indeed, we have $u_1 - \psi \geqslant 0$ in Ω by maximum principle ; $\mu_1 \geqslant 0$ by our hypothesis ; while $(u_1 - \psi, \mu_1) = 0$, for supp μ_1, $(\mu_1 = -\Delta u_1)$, is contained-by (5) — in the region where $u_1 = \psi$, supp $\mu_1 = \psi$, denoting the support of μ_1] .

Otherwise, let $Q_2 = Q_1 \cup P_2$ be a smoothly bounded open region in Ω such that

$$\mu_1 \leqslant 0 \quad \text{on } Q_2 \ , \quad \mu_1(F) < 0 \quad \text{for some } F \subset P_2$$

Then we replace u_1 with the solution u_2 of the Dirichlet problem :

$$-\Delta u_2 = 0 \quad \text{in} \quad Q_2 = Q_1 \cup P_2 \ , \ u_2 = \psi \ \text{in}\, \Omega - Q_2$$

and so on, solving at the n^{th} step a Dirichlet problem like :

$$-\Delta u_n = 0 \quad \text{in } Q_n = Q_{n-1} \cup P_n \ , \ u_n = \psi \text{ in } \Omega - Q_n$$

with $Q_n = Q_{n-1} \cup P_n$ a smoothly bounded open region in Ω , such that :

$$\mu_{n-1} = -\Delta u_{n-1} \leqslant 0 \quad \text{on } Q_n \ , \quad \mu_{n-1}(F) < 0 \ \text{ for some borel } F \subset P_n .$$

The functions $u_{n-1} - u_n$ and $u_n - u$, u being the solution of (1), are both subharmonic in Q_n ; moreover, $u_{n-1} - u_n = \psi - \psi = 0$ and $u_n - u = \psi - u \leqslant 0$ in $\Omega - Q_n$, for all $n = 1, 2, \ldots$. By the maximum principle, we then find :

$$(6) \qquad \psi \leqslant u_{n-1} \leqslant u_n \leqslant u \qquad \text{in } \Omega \text{ for every } n \ .$$

If $\mu_n \geqslant 0$, then $u = u_n$, $\mu = \mu_n$ is the solution of (1).

If the algorithm above actually produces a *sequence* u_n, we can then consider the function :

$$\widetilde{u} = \lim u_n$$

Clearly, $\widetilde{u} \leqslant u$. If $\widetilde{\mu} = -\Delta \widetilde{u} \geqslant 0$ in Ω , then $u = \widetilde{u}$, $\mu = \widetilde{\mu}$ is the solution of (1), as it follows from the characterization of the solution u, see [22] : u is the *smallest* superharmonic function in Ω, vanishing on $\partial\Omega$, which is $\geqslant \psi$ in Ω.

It should be remarked at this point that the limit function $\widetilde{u}$ may depend on the choice of the region Q_n which is done at each step.

However, we shall not further discuss how these regions might be chosen in order to end up with a limit $\widetilde{u}$ such that $\widetilde{\mu} = -\Delta \widetilde{u} \geqslant 0$ in Ω .Indeed, before working out the algorithm just described, we shall first approximate problem (1) by replacing it with a suitable discrete problem with a *finite* number of degrees of freedom. Applied to the discrete problem, the algorithm will come to an end after a finite number of steps.

3 – A DISCRETIZATION OF THE UNILATERAL DIRICHLET PROBLEM

We shall sketch here an internal approximation of problem (2) by means of triangular affine elements. For sake of simplicity, we shall suppose that Ω is a bounded convex open subset of $\mathbb{R}^2$, with a smooth boundary $\partial\Omega$.

Given h, $0 < h < 1$, we first inscribe a polygon Ω_h in Ω, whose vertices belong to $\partial\Omega$ and whose sides have a length which does not exceed h. We then decompose Ω_h into triangles in such a way that :

(7) $$0 < l < h, \; l'/l'' \leqslant C_1 \quad , \; 0 < C_2 < \alpha \leqslant \pi/2$$

where C_1 and C_2, $C_2 < \pi/2$, are given positive costants, l, l', l'' are lenghts of arbitrary sides of the triangulation, α an arbitrary angle of our triangles.

We shall denote by I the set of all indices q associated with the **internal** nodes x_q of the triangulation (x_q is an internal node if the union of all triangles which have x_q as a vertex is contained in Ω). We shall denote by ∂I the set of all indices q associated with the **boundary** nodes of the triangulation (x_q is a boundary node if it is a vertex of some of the triangles that decompose Ω_h and x_q is not internal).

For each $q \in \mathrm{I} \cup \partial\mathrm{I}$, we shall consider the function :

$$\varphi_q^h(x), \; x \in \Omega_h$$

which is affine in each triangular element of the decomposition is $= 1$ at x_q and $= 0$ in all $x_p \neq x_q$, $p \in \mathrm{I} \cup \partial \mathrm{I}$.

We shall now consider the piecewise affine functions $v_h(x)$ defined by :

$$(8)\qquad \begin{aligned} v_h(x) &= \sum_{q \in \mathrm{I}} v_q^h \ \varphi_q^h(x) \qquad && x \in \Omega_h \\ v_h(x) &= 0 && x \in \Omega - \Omega_h \end{aligned}$$

and the cone :

$$K_h = \left\{ v_h(x) \quad / \quad v_h(x) \text{ given by (8)}, v_q^h \geqslant \psi(x_q) \ \forall q \in \mathrm{I} \right\}$$

The approximate problem is obtained by replacing K with K_h in problem (2):

$$(9)\qquad u_h \in K_h \ : a(u_h, v_h - u_h) \geqslant 0 \ \forall v_h \in K_h$$

It can be shown, see [14] and [35], that the solution $u_h(x)$ of (9) converges to the solution u of problem (2) of order h in the energy norm as $h \to 0$: if $\psi \in H^2(\Omega)$, the following estimate is indeed obtained in the papers quoted above :

$$\| u - u_h \|_{H^1} \leqslant Ch \| \psi \|_{H^2}$$

We refer to [14] and [35] for more details on this estimate. Let us also notice, incidentally, that special results on the approximation of the "contact set" $u = \psi$ have been given in [2]. For a general discussion of the convergence of approximate solutions of variational inequalities such as (2) see also [32] , [33].

Let us now write the discrete problem obtained by replacing the expression (8) of $u_h(x)$, that is :

$$(10)\qquad u_h(x) = \sum_{q \in \mathrm{I}} u_q^h \varphi_q^h(x), x \in \Omega_h \ ; u_h(x) = 0 \text{ in } \Omega - \Omega_h$$

into (9) and by choosing, for every $q \in \mathrm{I}$, $v_h(x) = u_h(x) + \varphi_q^h(x)$ with $\varphi_q^h(x) \equiv 0$ in $\Omega - \Omega_h$. We find the system :

$$(11)\qquad \begin{aligned} & u_q^h - \psi_q^h \geqslant 0 \\ & \mu_q^h = \sum_{p \in \mathrm{I}} a_{pq}^h \ u_p^h \geqslant 0 \\ & (u_q^h - \psi_q^h)\, \mu_q^h = 0, \ q \in \mathrm{I} \ ; \end{aligned}$$

where

$$\psi_q^h = \psi(x_q) \text{ for every } q \in \mathrm{I}$$

and

(12) $$a_{pq}^h = a(\varphi_q^h , \varphi_p^h) \ , \ p, q \in \mathrm{I}$$

The last condition in (11) is obtained by putting, for each node x_q where $u_h(x_q) = u_q^h > \psi_q^h$, $v_h(x) = u_h(x) - \epsilon\,\varphi^h(x)$ into (9), with $\epsilon > 0$ small enough so that $v_h \in K_h$: this gives $\mu_q^h \leqslant 0$, hence $\mu_q^h = 0$ by the second inequality in (11).

It is easy to check, in turn, that if we take the coefficients u_q^h of the function (10) to be the solution of (11) then $u_h(x)$ is the solution of (9). So the approximate problem (9) can be replaced by the equivalent discrete system (11).

If we introduce the vectors :

$$U_q^h = u_q^h - \psi_q^h \ , \ q \in \mathrm{I}$$

and

$$v_g^h = \sum_{p \in \mathrm{I}} a_{pq}^h \ \psi_p^h \ , q \in \mathrm{I}$$

then system (11) can be also written as :

(13) $$\begin{aligned} &U_q^h \geqslant 0 \\ &\mu_q^h = \sum_{p \in \mathrm{I}} a_{pq}^h \ U_p^h + v_q^h \geqslant 0 \\ &U_q^h \ \mu_q^h = 0 \quad , \quad q \in \mathrm{I}. \end{aligned}$$

Let us remark that we could have also obtained the systems (11) and (13) by directly discretizing the continuous problems (1) and (1') respectively (and by putting $U_q^h = - \psi_q^h$ for every $q \in \partial\, \mathrm{I}$ in order to eliminate the inhomogeneous condition in (1')).

The equivalence of (11) or (13) with (9) is indeed the discrete analogue of the equivalence of (1) or (1') with (2) discussed in section 1.

A system of inequalities such as (13) above is known in the literature as a *complementary system* and several algorithms are known for its solution, see for instance [9], [20].

We shall now sketch an algorithm for solving system (13), which is directly modelled on the procedure described in section 2. In the following section we shall give a more detailed description of this algorithm and we shall see that it essentially reduces to an algorithm introduced by Chandrasekaran [4] to solve complementarity systems involving matrices with non positive off diagonal entries.

We shall omit the superscript h and we shall write instead U^0, U^1,... to denote the vectors $(U_q^0)_{q \in I}$, $(U_q^1)_{q \in I}$, . . yielded by the algorithm. Similarly, we shall write $\mu^0, \mu^1, \ldots$ and so on.

At the initial step we define U^0, μ^0 by :

$$U_q^0 = 0 \ , \ \mu_q^0 = \nu_q \ , \ q \in I$$

If $\nu_q \geqslant 0 \ \forall q$, then the pair :

$$U = U^0 \ , \ \mu = \mu^0$$

is the solution of (13).

Let us suppose, instead, that there is a set P_1 of indices such that :

$$\phi \neq P_1 \subset \{ q \in I / \mu_q^0 < 0 \}$$

Then we define U_q^1 to be the solution of the system :

$$\mu_q^1 = \sum_{p \in I} a_{pq} \, U_p^1 + \nu_q = 0 \, , q \in Q_1 \ (Q_1 = P_1)$$

$$U_q^1 = 0 \qquad\qquad q \in I - Q_1$$

If

$$\mu_q^1 = \sum_{p \in I} a_{pq} \, U_p^1 + \nu_q \geqslant 0 \, , \forall \, q \in I - Q_1 \, ,$$

then the pair :

$$U = U^1 \ , \ \mu = \mu^1$$

is the solution of (13).

If this is not the case, we choose a set P_2 of indices such that :

$$\phi \neq P_2 \subset \{ q \in I \ / \mu_q < 0 \}$$

and we define the vector $U_q^2, q \in I$, by solving the system :

$$\mu_q^2 = \sum_{p \in I} a_{pq} \, U_p^2 + \nu_q = 0 \, , q \in Q_2 = Q_1 \cup P_2$$

$$U_q^2 = 0 \qquad , \qquad q \in \mathrm{I} - Q_2 \quad .$$

The algorithm ends if

$$\mu_q^2 = \sum_{p \in \mathrm{I}} a_{pq} \, U_p^2 + \nu_q \geqslant 0 \quad \forall q \in \mathrm{I} - Q_2$$

and then

$$U = U^2 \quad , \quad \mu = \mu^2$$

is the solution of (13).

Otherwise, we go on by choosing a new set of indices :

$$P_3 \subset \{ q \in \mathrm{I} \, / \, \mu_q^2 < 0 \}$$

and putting $Q_3 = Q_2 \cup P_3$.

In the following section we shall prove that the algorithm comes to an end after a finite number of steps, at most N if N is the number of the internal nodes. The proof exploits the special properties of the matrix a_{pq} which are inherited from the continuous boundary value problem of which systems (13) is the discrete analogue. Indeed, the matrix (12) is positive definite, in particular, all principal minors have positive determinants, and in consequence of condition (7), it is easy to see that $a_{pq} \leqslant 0 \; \forall \, p \neq q$.

4. – THE DISCRETE COMPLEMENTARITY SYSTEM

Let $A = (a_{pq})_{p,q \, \in \, \mathrm{I}}$ be an NxN matrix with the following properties : (i) A belongs to the class (P), which is to say, all principal minors $A_{pp} = (a_{p,q})_{p,q \, \in Q}$, $Q \subset \mathrm{I}$, have a positive determinant ;

(ii) A belongs to the class (Z), that is, $a_{p,q} \leqslant 0 \quad \forall p \neq q$.

Let us notice that we need not assume A to be symmetric.

Let $\nu = (\nu_q)_{q \in \mathrm{I}}$ be a given N vector. The complementarity system we are dealing with can be written, with standard notation, as follows (5) :

(5) If $x = x_{\mathrm{I}} = (x_q)_{q \, \in \, \mathrm{I}}$, we write $x \geqslant 0$ if $x_q \geqslant 0 \; \forall q \in \mathrm{I}$.
Moreover, $x.y = \sum_{q \in \mathrm{I}} x_q \, y_q$.

$$U \geqslant 0, \quad \mu \geqslant 0, \quad U \cdot \mu = 0$$

Algorithm I

At each step of the algorithm we choose a family $Q \subset \mathrm{I}$ of indices. The system in (14) can be accordingly partitioned as follows :

(15)
$$\mu_Q = A_{QQ} U_Q + A_{QQ'} U_{Q'} + \nu_Q ;$$
$$\mu_{Q'} = A_{Q'Q} U_Q + A_{Q'Q'} U_{Q'} + \nu_{Q'} ;$$

where
$$Q' = \mathrm{I} - Q$$ [6] .

At this point the following conditions are imposed :

(16) $$U_{Q'} = 0 \ , \ \mu_Q = 0$$

and then U_Q is determined by solving the system :

(17) $$A_{QQ} U_Q + \nu_Q = 0.$$

Let us notice that there is a unique solution of this system, since $\det A_{QQ} > 0$ by assumption (i).

(14)
$$\mu = A\, U + \nu ;$$

where $U = (U)_{q\ q \in \mathrm{I}}$, $\mu = (\mu)_{q\ q \in \mathrm{I}}$

This step is the final one, and the solution of (14) is given by :

(18) $$U_Q = - A_{QQ}^{-1} \nu_Q \ , U_{Q'} = 0$$

(19) $$\mu_Q = 0 \quad , \mu_{Q'} = A_{Q'Q} U_Q + \nu_{Q'} \quad ,$$

provided the following positivity test is satisfied :

(20) $$\mu_{Q'} = A_{Q'Q} U_Q + \nu_{Q'} \geqslant 0$$

In the opposite case, i.e. $\mu_{Q'} \ngeqslant 0$, we choose a set :

(21) $$P \subset \{ q \in Q' / \mu_q < 0 \}$$

(6) In the following, for any set $Q \subset \mathrm{I}$ we shall always put $Q' = \mathrm{I} - Q$.

and take a

$$\text{new } Q = \text{old } Q \cup P \ . \tag{22}$$

In the initial step we simply take $Q = \phi$, which is to say, we put $U = U_{\mathrm{I}} = 0$ and the positivity test (19) becomes :

$$\mu_{\mathrm{I}} = \nu_{\mathrm{I}} \geqslant 0 \ .$$

Thus, if the given ν is $\geqslant 0$, then system (14) is trivially solved by the pair :

$$U = 0 \ , \ \mu = \nu \ .$$

If $\nu \not\geqslant 0$, then we choose :

$$P \subset \{ q \in \mathrm{I} \, / \, \nu_q < 0 \}$$

and take

$$Q = P$$

and go on as above.

We shall now prove that going from the $(n-1)^{\text{th}}$ step to the n^{th} step, thus replacing Q_{n-1} with

$$Q_n = Q_{n-1} \cup P_n$$

where P_n has been so chosen as to be :

$$P_n \subset \left\{ q \in Q'_{n-1} \, / \mu_q < 0 \right\} \ ,$$

we find :

$$0 \leqslant U_q^{n-1} \leqslant U_q^n \leqslant U_q \quad , \ \forall \, q \in \mathrm{I} \tag{23}$$

$$0 \leqslant \mu_q^n \leqslant \mu_q^{n-1} \leqslant \mu_q \quad , \ \forall q \in Q'_n ; \tag{24}$$

while

$$\mu_q^n = 0 \quad , \ \forall \, q \in Q_n$$

These relations are clearly the discrete analogue of the monotonicity relations (6), which were obtained in section 2 as a consequence of the maximum principle. The role of the maximum principle is now taken by the following lemma :

LEMMA 1. **Let $A = A_{\mathrm{I}} \in (P) \cap (Z)$ and $x = x_{\mathrm{I}}$ satisfy**

$$A_{QI} \, x_I \leqslant 0 \quad , \ x_{Q'} \leqslant 0$$

for some $Q \subset \mathrm{I}$. Then, $x \leqslant 0$.

Proof. We have :

$$A_{QQ}x_Q + A_{QQ'}x_{Q'} \leqslant 0$$

$$x_{Q'} \leqslant 0$$

Since $A \in (P) \cap (Z)$, then $A_{QQ}^{-1} \geqslant 0$ (see [15]). Therefore

$$x_Q + A_{QQ}^{-1} A_{QQ'} x_{Q'} \leqslant 0$$

Since $A_{QQ'} \leqslant 0$, this implies :

$$x_Q \leqslant 0 \qquad \text{Q.E.D.}$$

COROLLARY. $\quad 0 \leqslant U^{n-1} \leqslant U^n \leqslant U$

Proof. We have $U^0 = 0$, U^{n-1} satisfies (16) and (17) with $Q = Q_{n-1}$, U^n satisfies (16) and (17) with $Q = Q_n = Q_{n-1} \cup P_n$ and $P_n \subset Q'_{n-1}$ is such that :

$$\mu_{P_n}^{n-1} = A_{P_n Q_{n-1}} U_{Q_{n-1}}^{n-1} + \nu_{P_n} < 0^{(7)} .$$

Therefore, the vector :

$$x = U^{n-1} - U^n$$

satisfies the hypothesis of Lemma 1 with $Q = Q_n$. In fact, we have :

$$A_{Q_{n-1}\mathrm{I}} x_{\mathrm{I}} = A_{Q_{n-1}\mathrm{I}} U_{\mathrm{I}}^{n-1} - A_{Q_{n-1}\mathrm{I}} U_{\mathrm{I}}^{n} = -\nu_{Q_{n-1}} + \nu_{Q_{n-1}} = 0,$$

$$A_{P_n\mathrm{I}} x_{\mathrm{I}} = A_{P_n\mathrm{I}} U_{\mathrm{I}}^{n-1} - A_{P_n\mathrm{I}} U_{\mathrm{I}}^{n} < -\nu_{P_n} + \nu_{P_n} = 0 .$$

and since $U_{Q'_n}^n = 0$, $U_{Q'_{n-1}}^{n-1} = 0$, we also have :

$$x_{Q'_n} = U_{Q'_n}^{n-1} - U_{Q'_n}^{n} = 0$$

(7) If $x = x_{\mathrm{I}}$ is a vector, we write $x > 0$ if $x_q > 0 \ \forall_q \in \mathrm{I}$.

Hence, by Lemma 1, $x < 0$, that is :

$$U^{n-1} \leqslant U^n \quad .$$

To show that $U^n \leqslant U$ we apply Lemma 1 to the vector $x = U^n - U$, again with $Q = Q_n$. Q.E.D.

Algorithm I, as we already said, essentially reproduces the algorithm given in [4]. It may be summarized in the following cyclical scheme, where $A^0 = A^0_{\mathrm{II}}$ and $b^0 = b^0_{\mathrm{I}}$ denote the initial data

STEP 0

Put $b_{\mathrm{I}} = b^0_{\mathrm{I}}$, $Q = \phi$.

Go to STEP 1.

STEP 1

Put $\nu_{\mathrm{I}} = (b_Q, 0)$, $\mu_{\mathrm{I}} = (0, b_{Q'})$.

If $b_{Q'} \geqslant 0$, stop : $\nu = \nu_{\mathrm{I}}$, $\mu = \mu_{\mathrm{I}}$ are the solutions of (14).

If $b_{Q'} \ngeqslant 0$, choose $P \subset \{ q \in Q' / b_p < 0 \}$

Go to STEP 2.

STEP 2

Put $Q = Q \cup P$.

Solve $A^0_{QQ} \, x^0_Q + b_Q = 0$

Put $b_{\mathrm{I}} = (x_Q, \ A^0_{Q'Q} \, x_Q + b^0_Q)$.

Go to STEP 1.

Since at each intermediate cycle Q increases at least of one more index, the algorithm stops after at most N cycles have been done, N being the size of I.

Let us point out that a choice must be made at each step : the choice of the set $P \subset Q' : P \subset \{ q \in Q' / b_q < 0 \}$. Let us remark in this respect that the matrix (12) is sparse, due to the fact that the support of each φ_q^h is localized around the vertex x_q. This is indeed a typical feature of any finite element method. There are, therefore, quite efficient methods for solving a system like (17), even if the order of Q is very large. For instance, iterative methods may work very well, see e.g. [42]. Thus, it seems the natural choice, in this case, to take *P maximal,* which to say :

$$P = \{ q \in Q' / \mu_q < 0 \}$$

However, the apposite *minimal* choice, consisting in taking at each step :

$$P = \{q\} \quad , \mu_q < 0 \ ,$$

may be the natural one if we want to solve system (14) by pivotal methods of Gauss-Seidel type as we shall describe in the following Algorithm II.

Algorithm II

Let us firts remark that in the Algorithm I, once we have chosen the set of indices Q as in (22) and partitioned the linear system, appearing in (14), in the form (15), we essentially make two further steps : first, we impose (16), that is the vanishing of $U_{Q'}$ and μ_Q ; then we evaluate U_Q by solving the system (17) and use it to make the positivity test (20) on $\mu_{Q'}$.

In the algorithm we shall now describe, these two steps are essentially taken in reverse order : *first* we evaluate U_P by making a pivot transform on A_{PP} and replace it in the remaining equations, in particular in the expression of $\mu_{Q'}$ (recall that new Q = old $Q \cup P$) : then we impose the vanishing of $U_{Q'}$ and μ_Q and verify the positivity of $\mu_{Q'}$.

More precisely, let initially be :

$$Q = P$$

with

$$P \subset \{ q \in I / \nu_q < 0 \} .$$

We carry out a principal pivot transform (p.p.t.) T_Q in the system (15) with block-pivot on A_{QQ}. We thus replace system (15) with :

$$(25) \qquad \begin{aligned} U_Q &= \bar{A}_{QQ}\,\mu_Q + \bar{A}_{QQ'}\,U_{Q'} + \bar{\nu}_Q \ ; \\ \mu_{Q'} &= \bar{A}_{Q'Q}\,\mu_Q + \bar{A}_{Q'Q'}\,U_{Q'} + \bar{\nu}_{Q'} \ ; \end{aligned}$$

where :

$$\bar{A}_{QQ} = A_{QQ}^{-1} \qquad \bar{A}_{QQ'} = -A_{QQ}^{-1} A_{QQ'}$$

$$\bar{A}_{Q'Q} = A_{Q'Q} A_{QQ}^{-1} \qquad \bar{A}_{Q'Q'} = A_{Q'Q'} - A_{Q'Q} A_{QQ}^{-1} A_{QQ'}$$

and

$$\bar{\nu}_Q = -A_{QQ}^{-1} \nu_Q \qquad \bar{\nu}_{Q'} = \nu_{Q'} - A_{Q'Q} A_{QQ}^{-1} \nu_Q \, .$$

Note that, since we had $\nu_Q < 0$, now we have $\bar{\nu}_Q \geqslant 0$. We now impose the vanishing of $U_{Q'}$ and μ_Q and we obtain from system (25) :

$$U_Q = \bar{\nu}_Q \quad , \quad \mu_{Q'} = \bar{\nu}_{Q'} \quad .$$

The positivity test (20) now becomes :

(26) $$\bar{\nu}_{Q'} \geqslant 0$$

If (26) is satisfied, then clearly :

$$U = (\bar{\nu}_Q \, , 0) \qquad \mu = (0 \, , \bar{\nu}_{Q'})$$

is the solution of (14). If not, we choose, as in Algorithm I, a new set of indices :

$$P \subset \{q \in Q' / \mu_q < 0\}$$

and we go on by making a block pivot transform with pivot on $\bar{A}_{PP}$.

At each intermediate n^{th} step of the algorithm we are then in presence of a system like this :

$$U_Q = B_{QQ} \mu_Q + B_{QP} U_P + B_{Q(newQ)'} U_{(newQ)'} + b_Q$$

(27) $$\mu_P = B_{PQ} \mu_Q + B_{PP} U_P + B_{P(newQ)'} U_{(newQ)'} + b_P$$

$$\mu_{(newQ)'} = B_{(newQ)'Q} \mu_Q + B_{(newQ)'P} U_P + $$
$$+ B_{(newQ)'(newQ)'} U_{(newQ)'} + b_{(newQ)'}$$

$newQ = Q \cup P$, which has been obtained by carrying out successive p.p.t., indexed by $P_1, P_2, \ldots, P_{n-1}$ on the initial system (15). Thus, at the n^{th} step we have :

$$Q = \text{old } Q = P_1 \cup P_2 \ldots\ldots \cup P_{n-1}$$

while $P \subset Q'$ is a chosen set of indices such that :

$$b_P < 0 \ .$$

The following lemma describes what happens to system (27) above when we carry out a p.p.t. T_P with block pivot B_{PP}. Let the transformed system be :

(28)
$$U_Q = \bar{B}_{QQ}\,\mu_Q + \bar{B}_{QP}\,\mu_P + \bar{B}_{Q(newQ)'}\,U_{(newQ)'} + \bar{b}_Q$$

$$U_P = \bar{B}_{PQ}\,\mu_Q + \bar{B}_{PP}\,\mu_P + \bar{B}_{P(newQ)'}\,U_{(newQ)'} + \bar{b}_P$$

$$\mu_{(newQ)'} = \bar{B}_{(newQ)'Q}\,\mu_Q + \bar{B}_{(newQ)'P}\,\mu_P + \bar{B}_{(newQ)'(newQ)'}\,U_{(newQ)'} + \bar{b}_{(newQ)'} \ .$$

Thus :

$$\bar{B}_{QQ} = B_{QQ} - B_{QP}\,B_{PP}^{-1}\,B_{PQ}$$

$$\bar{B}_{QP} = B_{QP}\,B_{PP}^{-1}$$

$$\bar{B}_{Q(newQ)'} = B_{Q(newQ)'} - B_{QP}\,B_{PP}^{-1}B_{P(newQ)'}$$

$$\bar{B}_{PQ} = -\,B_{PP}^{-1}B_{PQ}$$

$$\bar{B}_{PP} = B_{PP}^{-1}$$

$$\bar{B}_{P(newQ)'} = -\,B_{PP}^{-1}B_{P(newQ)'}$$

$$\bar{B}_{(newQ)'Q} = B_{(newQ)'Q} - B_{(newQ)'P}\,B_{PP}^{-1}\,B_{PQ}$$

$$\bar{B}_{(newQ)'P} = B_{(newQ)'P}\,B_{PP}^{-1}$$

$$\bar{B}_{(newQ)'(newQ)'} = B_{(newQ)'(newQ)'} - B_{(newQ)'P}\,B_{PP}^{-1}B_{P(newQ)'}$$

while

$$\overline{b}_Q = b_Q - B_{QP} B_{PP}^{-1} b_P$$

$$\overline{b}_P = - B_{PP}^{-1} b_P$$

$$\overline{b}_{(newQ)'} = b_{(newQ)'} - B_{(newQ)'P} B_{PP}^{-1} b_P \ .$$

We have :

LEMMA 2. **Let $B = B_{II} \in (P)$ and $Q \subset I$ such that :**

(29) $$B_{QI} \geqslant 0, B_{Q'I'} \in (Z) \quad b_q \geqslant 0 \ .$$

Let $P \subset Q'$ be such that :

$$b_P < 0$$

and put

$$\text{new } Q = Q \cup P \ .$$

Then, if a principal pivot transform, with block pivot B_{PP}, is carried out in system (27), *the transformed system* (28) *is such that $\overline{B} \in (P)$ and*

(30) $$\overline{B}_{(newQ)I} \geqslant 0 \ , \overline{B}_{(newQ)'I'} \in (Z) \ , \overline{b}_{(newQ)} \geqslant b_{(newQ)} \geqslant 0$$

Moreover,

(31) $$\overline{b}_P > 0 \text{ and } \overline{b}_{(newQ)'} \leqslant b_{(newQ)'}$$

Proof. Ilt is well known that p.p.t. leave the class (P) invariant, see [41], hence $\overline{B} \in (P)$. Moreover, as we have already noticed, since $B_{PP} \in (P) \cap (Z)$, then

(32) $$B_{PP}^{-1} \geqslant 0 \ .$$

Therefore, properties (30) of $\overline{B}$ and $\overline{b}$ can be directly checked on their expressions listed above, by taking (32) and (29) into account. In addition, since we know that det $(B_{PP}^{-1}) \neq 0$ and $b_P < 0$, then $\overline{b}_P = - B_{PP}^{-1} b_P \geqslant 0$. Finally, since $B_{(newQ)'P}$ $B_{PP}^{-1} b_P \geqslant 0$, we have $b_{(newQ)'} = \overline{b}_{(newQ)'} - B_{(newQ)'P} B_{PP}^{-1} b_P \leqslant b_{(newQ)'}$. Q.E.D.

The monotonicity conditions in (30), (31) are clearly the analogous of relations (23), (24). The new algorithm ends up when, putting $U_{Q'} = \mu_Q = 0$ in system (27) the positivity test :

$$b_{(new Q)'} \geqslant 0$$

(i.e. $\mu_{Q'} \geqslant 0$ in older notation) is satisfied.

Let us summarize Algorithm II by a cyclical scheme. We have :

STEP 0

Put $A_{II} = A^{\circ}_{II}, b_I = b^{\circ}_I, Q = \phi$

Go to STEP 1.

STEP 1

Put $\nu_I = (b_Q, 0), \mu_I = (0, b_{Q'})$.

If $b_{Q'} \geqslant 0$, stop : $\nu = \nu_I, \mu = \mu_I$ are solutions.

If $b_{Q'} \not\geqslant 0$, choose $P \subset \{q \subset Q' / b_q < 0\}$.

Go to STEP 2.

STEP 2

Make a p.p.t. with block pivot A_{PP} in the system

$$y = A x + b$$

Put $(A, b) = T_P (A, b), Q = Q \cup P$.

Go to STEP 1.

Clearly, by LEMMA 2, the algorithm stops after at most N cycles, if N is the size of I. Finally, let us notice that here too a choice of P must be made at each step. As we already remarked, the minimal choice, that is, to take P consisting of a single index q such that $b_q < 0$, has the advantage that all the p.p.t. that must be carried out, are elementary pivot transformations and there is no problem of evaluating the inverse matrix A_{PP}^{-1}.

In the case we are interested in, however, we know that the initial matrix (12) is sparse and thus special 'ad hoc' techniques could perhaps conveniently used in doing the block p.p.t. on A_{PP}. It should be also noticed that as successive p.p.t. are carried out, the matrix involved becomes less and less sparse.

The numerical approximation of problem (1) by means of the algorithms described above has been investigated in more details in [21].

Let us also mention in this regard that numerical approximations of problem (1), based on different methods, have already been considered by various authors, see for instance [2], [5], [6], [12], [13], [16], [17], [31], [37], [38], [40]. In particular, the approximation method based on Algorithm I is similar to the "conditioned harmonization" described in [12], [31].

However, it yields non-decreasing approximations of the solution of the discrete problem, whereas in [12], [31] non-increasing approximations are obtained.

We refer to [21] for further comments on all methods quoted above.

REFERENCES

[1] BAIOCCHI C. *Su un problema di frontiera libera connesso a questioni di idraulica.* Ann. Mat. Pura ed Appl. 92 (1972), 107-127.

[2] BAIOCCHI C. – COMINCIOLI V – GUERRI L. – VOLPI G *Free boundray problems in the theory of fluid flow through porous media : a numerical approach.* Calcolo 10 (1973) 1-86.

[3] BREZIS H – STAMPACCHIA G *Sur la régularité de la solution d'inéquations elliptiques.* Bull. Soc. Math. France 96 (1968) 153-180.

[4] CHANDRASEKARAN R. *A special case of the complementarity pivot problem.* Opsearch (1970), 263-268.

[5] COMINCIOLI V. – GUERRI R. – VOLPI G *Analisi numerica di un problema di frontiera libera connesso al moto di un fluido attraverso un mezzo poroso* L A.N., C.N.R. 17 (1971) Pavia

[6] COMINCIOLI V. *Metodi di rilassamento per la minimizzazione in uno spazio prodotto.* L.A N, C.N.R, 20 (1971) Pavia.

[7] COTTLE R W. – HABETLER G I – LEMKE C.E. *Quadratic forms semidefinite over convex cones.* R.P.I. Math. Rep (1967) Reusselaer Polytechnic Inst Troy, New-York, 551-565.

[8] COTTLE R. W. – DANTZIG G.B. *Positive (semi) definite programming* Non Linerar Programming (Abadie J. Ed.) North-Holland Amsterdam (1967), 55-73.

[9] COTTLER R. W. – DANTZIG G B. *Complementarity pivot theory of mathematical programming.* Linear Algebra and its Appl 1 (1968) 103-125

[10] DE ANGELIS V – FUSCIARDI A. – SCHIAFFINO A. *Some methods for the numerical solution of variational inequalities with two obstacles.* Calcolo (to appear).

[11] DOLCETTA I. *Sistemi di complementarità e diseguaglianze variazionali.* Tesi Univ. di Roma (1972).

[12] DURAND J.F *Resolution numerique des problèmes aux limites sous-harmoniques.* Thèse, Univ. de Montpellier (1968).

[13] DURAND J F *L'algorithme de Gauss Seidel applique à un problème unilateral non symetrique* R A I R O , R-2 (1972), 23-30

[14] FALK R S *Error estimates for the approximation of a class of variational inequalities* Thesis, Rutger Univ New Bruswich

[15] FIEDLER M. – PTAK V *On matrices with non positive off diagonal elements and positive principal minors* Czech Math J 12 (1962), 382-400

[16] FUSCIARDI A. – MOSCO U – SCARPINI F – SCHIAFFINO A *A dual method for the numerical solution of some variational inequalities* J Math Anal Appl 40 (1972), 471-493

[17] GLOWINSKI R *La methode de relaxation* Quaderni di Matematica Ist Mat Univ di Roma, 14 (1971)

[18] KARAMARDIANS *The nonlinear complementarity problem with application* JOTA, 4 (1969), 87-98

[19] KARAMARDIANS *The complementaring problem* Mathematical programming 2 (1972), 107-129

[20] LEMKE C E *Recent results on complementarity problems* Proceedings of the Princeton Symposium on Mathematical Programming (Kuhn H W Ed) Princeton Univ Press (1970), 349 384

[21] LEVATI C – SCARPINI F – VOLPI G *Sul trattamento numerico di alcuni problemi variazionali di tipo unilaterale* L A N , C N R 82 (1974) Pavia

[22] LEWY H – STAMPACCHIA G *On the regularity of a solution of a variational inequality* Comm Pure Appl Math 22 (1969) 153-188

[23] LEWY H – STAMPACCHIA G *On the smoothness of superharmonics which solve a minimum problem* J d'Anal Math (1970), 227-236

[24] LIONS J L – STAMPACCHIA G *Variational inequalities* Comm Pure Appl Math 20 (1967), 493-519

[25] MANCINO O G – STAMPACCHIA G *Convex programming and variational inequalities*, JOTA 9 (1972), 3-23

[26] MARZULLI P *Risoluzione alle differenze finite di equazioni alle derivate parzialli di tipo ellittico con condizioni su un contorno libero* Calcolo Suppl 1,5 (1968), 1-22

[27] MORE J J *The application of variational inequalities to complementarity problems and existence theorems* Tech Rep 71-110 Cornell Univ Ithaca

[28] MORÈ J J *Coercitivity conditions in nonlinear complementarity problems* Tech Rep 72-146 Cornell Univ Ithaca

[29] MORE J J *Classes of functions and feasibility conditions in non-linear complementarity problems* Tech Rep 73-174 Cornell Univ Ithaca

[30] MOREAU J J *Majorantes sur-harmoniques minimales d'une fonction continue* Sem d'Anal Unilat 5 (1968) Univ de Montpellier

[31] MOREAU J J *Traitement numerique d'un problème aux derivees partielles de type unilateral* Sem d'Anal Num 15 (1969) Univ de Montreal

[32] MOSCO U *Convergence of convex sets and of solutions of variational inequalities* Adv in Math 3 (1969), 510-585

[33] MOSCO U *An introduction to the approximate solution of variational inequalities* Constructive Aspects of Functional Analysis Corso C I M E 1971, Ed G Geymonat Cremonese, Roma (1973)

[34] MOSCO U. *Dual variational inequalities,* J. Math. Anal. Appl. 39 (1972), 202-206.

[35] MOSCO U. – STRANG G. *One sided approximation and variational inequalitites.* Bull. A.M.S. 80, (1974), 308-312.

[36] MOSCO U. – TROIANIELLO G. *On the smoothness of solutions of unilateral Dirichlet problems.* Boll. U.M.I. 8 (1973), 57-67.

[37] SIBONY M. *Sur l'approximation d'équations aux dérivées partielles non linéaires de type monotone.* J. Math. Anal. Appl. 34 (1971), 502-564.

[38] SIBONY M. *Méthodes itératives pour les équations et inéquations aux dérivées partielles non linéaires de type monotone.* Calcolo 2 (1970), 65-183.

[39] STAMPACCHIA G. *Formes bilinéaires coercitives sur les ensembles convexes.* C.R. Acad. Sc. Paris 258 (1964), 4413-4416.

[40] STAMPACCHIA G. *On a problem of numerical analysis connected with the theory of variational inequalities.* I.N.A.M. 10 Symposia Math. (1972), 281-293.

[41] TUCKER A.W. *Principal pivot transforms of square matrices.* SIAM 5 (1963) 305.

[42] VARGA R.S. *Matrix iterative analysis.* Prentice - Hall (1962) New Jersey.