LE MATEMATICHE
Vol. XLIX (1994) – Fasc. II, pp. 359-377

CONVERGENCE PROPERTIES OF A PROJECTION METHOD FOR AFFINE VARIATIONAL INEQUALITY PROBLEMS

LUCA ZANNI

This paper is concerned with the convergence properties of projection method proposed by S. Dafermos (1980) for solving asymmetric variational inequality problem $h \in \mathscr{K}$, $\langle C(h), f - h\rangle \geq 0$, $\forall f \in \mathscr{K}$, where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^n$, $\mathscr{K}$ is a nonempty closed convex subset of $\mathbb{R}^n$ and $C : \mathscr{K} \to \mathbb{R}^n$ is a continuously differentiable and strongly monotone function. The convergence rate of this iterative method is linear and depends on the asymmetry level of C and on the spectral condition number $K\left(\left[\frac{\partial C}{\partial f}(f)\right]_s\right)$ of the symmetric part of the Jacobian matrix $\left[\frac{\partial C}{\partial f}(f)\right]$. In particular, if $K\left(\left[\frac{\partial C}{\partial f}(f)\right]_s\right) \approx 1 \; \forall f \in \mathscr{K}$ doesn't hold, it has been proved that a rapid convergence of the method can not be guaranteed with the parameters' choices suggested by S. Dafermos. When C is an affine function, we propose a criterion to select a parameter of the method. If the asymmetry level of C is *sufficiently low*, this criterion ensures a rapid convergence even if it is not $K\left(\left[\frac{\partial C}{\partial f}(f)\right]_s\right) \approx 1$. The behaviour of the method in correspondence to this parameter's choice is also analysed experimentally on a class of randomly generated test problems.

Work supported by M.U.R.S.T. 40% project *Analisi Numerica e Matematica Computazionale*.

1. Introduction.

We consider the variational inequality

$$h \in \mathscr{K}, \qquad \langle C(h), f - h \rangle \geq 0, \qquad \forall\, f \in \mathscr{K}, \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^n$, $\mathscr{K}$ is a nonempty closed convex subset of $\mathbb{R}^n$ and $C : \mathscr{K} \to \mathbb{R}^n$ is a continuous function.
This inequality has many applications, especially in the study of equilibrium problems arising in economics, operations research and transportation sciences. If the set $\mathscr{K}$ is compact, the variational inequality has at least one solution, which is unique if C is strictly monotone [19]. In the case where the set $\mathscr{K}$ is not bounded, one can impose particular conditions on the function C in order to establish the existence of a solution.
For a survey on theory, algorithms and applications of variational inequalities, refer to [10].

The various methods proposed for solving (1.1) can be gathered in three main classes: (i) methods that follow an iterative approach based on solving a sequence of simpler problems, (ii) methods that solve an optimization problem equivalent to the variational inequality and (iii) direct methods.
The first class includes for example Newton, symmetrized Newton and linearized Jacobi methods that are locally convergent under certain conditions, and projection methods that result globally convergent. The rates of convergence of these methods are typically linear, except for Newton method that converges quadratically [2],[3],[4],[17].
The methods that follow the optimization approach generate a sequence by minimizing an appropriate nonnegative function (gap function) that is zero if and only if its argument is a solution of (1.1) [9],[13],[14],[15],[21],[22],[24]. A line search strategy allows these descent algorithms to converge globally from any starting point. In particular, in [15] and [21] Newton method is modified to obtain a globally convergent algorithm that, under suitable assumptions, has quadratic convergence rate. Recently, it has been studied [24] a very general gap function by means of which it is possible to obtain an *extended* descent algorithm that includes, as special cases, the methods proposed in [9],[21],[22].
For what concerns the relationships between the methods of these two classes, it has been observed [9],[22],[24] that the above iterative algorithms can be viewed as various realizations of particular descent methods, although they do not resort to minimization of any gap function.
When the convex set $\mathscr{K}$ is defined by linear inequality constraints, the direct methods [5],[12],[16] determine the solution of (1.1) by searching for it over linear manifolds obtained with appropriate combinations of the constraints. This

procedure implies the solution of a (generally nonlinear) system of equations at each step and, in the worst case, when there are many steps, it might not supply the solution in a short time. However, the steps of these methods can be executed simultaneously and this feature makes the methods very suited for implementation on a multiprocessor system. In a subsequent paper we will describe a parallel implementation on distributed memory systems of direct methods suggested in [12] and [16].

In this work we present a study on the convergence properties of projection method proposed in [3].
Projection method is one of the most popular methods for variational inequality problems and a good knowledge of its convergence properties can also be helpful to understand the behaviour of descent methods based on the use of a projective gap function [9],[21],[24]. When C is continuously differentiable and strongly monotone with modulus $\alpha > 0$ on $\mathcal{K}$:

$$(1.2)\quad \langle C(f^1) - C(f^2), f^1 - f^2\rangle \geq \alpha\langle f^1 - f^2, f^1 - f^2\rangle, \quad \forall\, f^1, f^2 \in \mathcal{K},$$

given $f^0 \in \mathcal{K}$, the projection method generates a sequence $\{f^i\}$, $f^i \in \mathcal{K}$, in which the iterate f^{i+1} is obtained by solving the variational inequality

$$(1.3)\qquad f^{i+1} \in \mathcal{K}, \qquad \langle \tilde{C}^i(f^{i+1}), f - f^{i+1}\rangle \geq 0, \qquad \forall\, f \in \mathcal{K},$$

where $\tilde{C}^i(f)$ is defined by

$$(1.4)\qquad \tilde{C}^i(f) = \rho C(f^i) + G(f - f^i),$$

with G symmetric positive definite matrix and $\rho > 0$. In [3] it is proved that if the parameter ρ is such that

$$(1.5)\qquad 0 < \rho < \frac{2\alpha}{\nu},$$

where ν is the maximum over $\mathcal{K}$ of the maximum eigenvalue of the symmetric positive definite matrix $\left[\frac{\partial C}{\partial f}\right]^t G^{-1}\left[\frac{\partial C}{\partial f}\right]$, then for any $f^0 \in \mathcal{K}$ the sequence $\{f^i\}$ is convergent to the solution of (1.1) and there exists a constant $\lambda \in]0, 1[$ such that for all i,

$$(1.6)\qquad \|f^{i+1} - h\|_G \leq \lambda\|f^i - h\|_G,$$

where $\|f\|_G = \langle f, Gf\rangle^{\frac{1}{2}} \quad \forall\, f \in \mathbb{R}^n$ (*).

(*) In this study the Euclidean vector norm and the matrix norm induced by Euclidean vector norm are denoted by $\|\cdot\|$.

For what concerns the convergence rate of the sequence $\{f^i\}$, in [23] it is shown that, independently of the matrix G and the parameter $\rho \in]0, \frac{2\alpha}{\nu}[$, the constant λ satisfies the following inequality:

$$\lambda \geq \left(1 - \left(K\left(\left[\frac{\partial C}{\partial f}(f)\right]_s\right)\right)^{-2}\right)^{\frac{1}{2}}, \qquad \forall f \in \mathcal{K}, \tag{1.7}$$

where $\left[\frac{\partial C}{\partial f}(f)\right]_s$ is the symmetric part of the matrix $\left[\frac{\partial C}{\partial f}(f)\right]$ and $K\left(\left[\frac{\partial C}{\partial f}(f)\right]_s\right)$ is the spectral condition number of $\left[\frac{\partial C}{\partial f}(f)\right]_s$.

Thus, except for $K\left(\left[\frac{\partial C}{\partial f}(f)\right]_s\right) \approx 1 \ \forall f \in \mathcal{K}$, the constant λ does not assume the values that would guarantee a rapid convergence of the sequence $\{f^i\}$ to the solution $h \in \mathcal{K}$.

Moreover, in [7] the constant λ is related to the asymmetry level of the function C and it is proved that $\lambda \to 1$ when the asymmetry level increases.

In Section 2, in the case C is an affine function, we study the convergence properties of the method when it is applied to solve the equivalent variational inequality [2]

$$x \in \mathcal{S}, \qquad \langle AC(A^t x), y - x\rangle \geq 0, \qquad \forall y \in \mathcal{S}, \tag{1.8}$$

where A is an $n \times n$ nonsingular matrix and

$$\mathcal{S} = A^{-t}\mathcal{K} = \left\{y \in \mathbb{R}^n \quad / \quad y = A^{-t} f, \quad f \in \mathcal{K}\right\}. \tag{1.9}$$

If A is chosen in an appropriate way, we prove that the method generates a sequence $\{y^k\}$, $y^k \in \mathcal{S}$ which converges faster than the sequence $\{f^i\}$ in $\mathcal{K}$. This result is used to introduce a criterion to select the parameter ρ that, when the asymmetry level of C is *sufficiently low*, allows the method to generate a sequence $\{\tilde{f}^i\}$ in $\mathcal{K}$ quickly convergent to $h \in \mathcal{K}$ even if $K\left(\left[\frac{\partial C}{\partial f}(f)\right]_s\right) \approx 1$ doesn't hold.

In Section 3, we compare the numerical results obtained by applying the projection method with the value of ρ proposed in [3] and with the value suggested in Section 2. We consider the example illustrated in [3] and some randomly generated test problems.

2. On the convergence properties of Dafermos' projection method.

We consider the case when $C : \mathscr{K} \to \mathbb{R}^n$ is an affine function,

$$C(f) = Jf + b,$$

where J is a nonsymmetric positive definite $n \times n$ matrix and b is an n-vector. Since J is positive definite the function C is strongly monotone and the variational inequality

$$(2.1) \qquad h \in \mathscr{K}, \qquad \langle Jh + b, f - h \rangle \geq 0, \qquad \forall f \in \mathscr{K},$$

has a unique solution.

Let J_s be the symmetric part of J, $J_s = \frac{1}{2}(J + J^t)$, and consider the Cholesky decomposition of J_s, $J_s = LL^t$, where L is a lower triangular matrix with positive diagonal entries.
If f and h are two vectors of $\mathbb{R}^n$ we have

$$\begin{aligned}\langle Jh + b, f - h \rangle &= \langle (LL^{-1})(J(L^{-t}L^t)h + b), f - h \rangle \\ &= \langle L^{-1}JL^{-t}(L^t h) + L^{-1}b, L^t f - L^t h \rangle.\end{aligned}$$

Thus, Problem (2.1) is equivalent to

$$(2.2) \quad h \in \mathscr{K}, \qquad \langle L^{-1}JL^{-t}(L^t h) + L^{-1}b, L^t f - L^t h \rangle \geq 0, \qquad \forall f \in \mathscr{K}.$$

Let

$$(2.3) \qquad \mathscr{S} = L^t \mathscr{K} = \{ y \in \mathbb{R}^n \;/\; y = L^t f, \quad f \in \mathscr{K} \},$$

$$(2.4) \qquad B = L^{-1}JL^{-t} \quad (B \text{ positive definite matrix}), \qquad c = L^{-1}b,$$

$$(2.5) \qquad x = L^t h,$$

then $h \in \mathscr{K}$ is a solution of (2.1) if and only if $x \in \mathscr{S}$ is a solution of the following variational inequality

$$(2.6) \qquad x \in \mathscr{S}, \qquad \langle Bx + c, y - x \rangle \geq 0, \qquad \forall\, y \in \mathscr{S}.$$

Now, we suppose to solve (2.6) by the projection method proposed in [3].
We observe that

$$B_s = \frac{1}{2}(B + B^t) = L^{-1}J_s L^{-t} = I,$$

$$B_a = \frac{1}{2}(B - B^t) = L^{-1}J_aL^{-t}, \qquad J_a = \frac{1}{2}(J - J^t),$$

$$\|B\| \geq \|B_s\| = 1.$$

By choosing the method's parameters $\tilde{G}$ and $\tilde{\rho}$ as suggested in [3] we have:

$$(2.7) \quad \tilde{G} = B_s = I, \qquad \tilde{\rho} = \frac{\text{minimum eigenvalue of } B_s}{\text{maximum eigenvalue of } B^t\tilde{G}^{-1}B} = \frac{1}{\|B\|^2}.$$

Given a vector $y^0 \in \mathscr{S}$, the method generates a sequence $\{y^k\}$, $y^k \in \mathscr{S}$, which is convergent to the solution $x \in \mathscr{S}$ of (2.6) and such that

$$(2.8) \qquad \|y^k - x\| \leq \frac{\tilde{\lambda}^k}{1 - \tilde{\lambda}}\|y^1 - y^0\|,$$

where

$$(2.9) \qquad \tilde{\lambda} = \left(1 - \frac{1}{\|B\|^2}\right)^{\frac{1}{2}}, \qquad 0 < \tilde{\lambda} < 1.$$

In order to establish whether the constant $\tilde{\lambda}$ assumes values close to zero we consider the following result.

Theorem 2.1. *Let J be a nonsymmetric positive definite matrix and J_s be its symmetric part. Let $J_s = LL^t$ be the Cholesky decomposition of J_s and $B = L^{-1}JL^{-t}$. It results*

$$\|B\| \leq 1 + K(J_s)d_{J,J_s},$$

where

$$(2.10) \qquad d_{J,J_s} = \frac{\|J - J_s\|}{\|J_s\|}.$$

Proof. It results

$$(2.11) \qquad \|B\| = \|L^{-1}J_sL^{-t} + L^{-1}J_aL^{-t}\| = \|I + L^{-1}J_aL^{-t}\| \leq$$
$$\leq 1 + \|L^{-1}J_aL^{-t}\| \leq 1 + \|L^{-1}\|^2\|J_a\|$$

and, since $\|L^{-1}\|^2 = \|J_s^{-1}\| = K(J_s)/\|J_s\|$, we have

$$(2.12) \qquad \|B\| \leq 1 + K(J_s)\|J_a\|/\|J_s\| = 1 + K(J_s)\|J - J_s\|/\|J_s\|. \qquad \square$$

The number d_{J,J_s} can be viewed as a measure of the *relative distance* of J from its symmetric part; i.e., a measure of the asymmetry level of the function C.
Using Theorem 2.1 and (2.9) we have

$$\tilde{\lambda} \leq \left(1 - \frac{1}{\left(1 + K(J_s)d_{J,J_s}\right)^2}\right)^{\frac{1}{2}}. \tag{2.13}$$

From the bound (2.13) and Figure 2.1 in which is illustrated the graph of the function

$$F\left(K(J_s)d_{J,J_s}\right) = \left(1 - \frac{1}{\left(1 + K(J_s)d_{J,J_s}\right)^2}\right)^{\frac{1}{2}},$$

we conclude that the constant $\tilde{\lambda}$ assumes values close to zero if the product $K(J_s)d_{J,J_s}$ is sufficiently small.

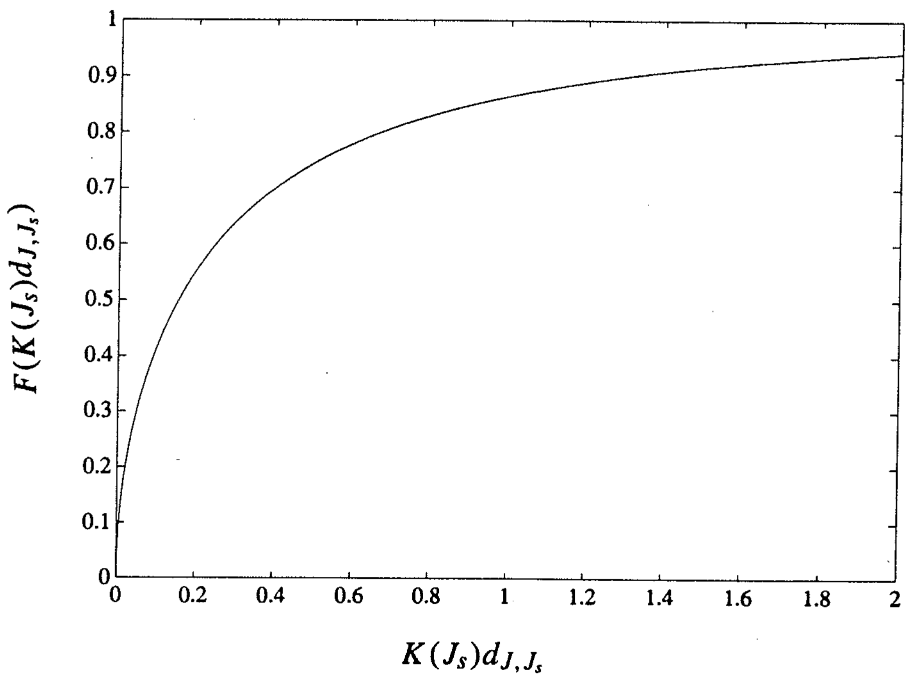

Fig. 2.1. Graph of the function $F\left(K(J_s)d_{J,J_s}\right) = \sqrt{1 - \left(1 + K(J_s)d_{J,J_s}\right)^{-2}}$

Thus, when J is sufficiently close to its symmetric part, the sequence $\{y^k\}$ obtained by applying the method proposed in [3] to the transformed Problem (2.6) is quickly convergent even if it is not $K(J_s) \approx 1$.

Let's now consider the sequence $\{\tilde{f}^k\}$, $\tilde{f}^k \in \mathscr{K}$, which corresponds to the sequence $\{y^k\}$, $y^k \in \mathscr{S}$,

$$\tilde{f}^k = L^{-t} y^k, \qquad k = 0, 1, 2, \ldots . \tag{2.14}$$

We proceed to show that $\{\tilde{f}^k\}$ can be obtained by applying projection method to (2.1) with an appropriate choice of parameters G and ρ.

Theorem 2.2. *The sequence $\{\tilde{f}^k\}$, $\tilde{f}^k \in \mathscr{K}$, defined in* (2.14), *can be generated by applying projection method to* (2.1) *with starting point $\tilde{f}^0 = L^{-t} y^0$ and parameters*

$$G = J_s, \qquad \rho = \tilde{\rho} = \frac{1}{\|B\|^2},$$

where B is the matrix defined in (2.4).

Proof. We proceed by induction. We proof that $\tilde{f}^1 = L^{-t} y^1$ is the vector generated at the first step of the method; i.e., it is the solution of the following variational inequality:

$$f \in \mathscr{K}, \quad \langle J_s f + \tilde{\rho}(J\tilde{f}^0 + b) - J_s \tilde{f}^0, g - f\rangle \geq 0, \quad \forall g \in \mathscr{K}. \tag{2.15}$$

Given $\overline{g} \in \mathscr{K}$ we show that

$$\langle J_s(L^{-t}y^1) + \tilde{\rho}(J\tilde{f}^0 + b) - J_s\tilde{f}^0, \overline{g} - L^{-t}y^1\rangle \geq 0. \tag{2.16}$$

Since $J_s = LL^t$ and $\tilde{f}^0 = L^{-t}y^0$ it results

$$\begin{aligned}
\langle J_s(L^{-t}y^1) &+ \tilde{\rho}(J\tilde{f}^0 + b) - J_s\tilde{f}^0, \overline{g} - L^{-t}y^1\rangle = \\
&= \langle LL^t(L^{-t}y^1) + \tilde{\rho}(JL^{-t}y^0 + b) - LL^t(L^{-t}y^0), \overline{g} - L^{-t}y^1\rangle \\
&= \langle L(y^1 + \tilde{\rho}(L^{-1}JL^{-t}y^0 + L^{-1}b) - y^0), \overline{g} - L^{-t}y^1\rangle \\
&= \langle y^1 + \tilde{\rho}(L^{-1}JL^{-t}y^0 + L^{-1}b) - y^0, L^t\overline{g} - y^1\rangle \\
&= \langle y^1 + \tilde{\rho}(By^0 + c) - y^0, L^t\overline{g} - y^1\rangle.
\end{aligned}$$

Set $\overline{y} = L^t\overline{g}$, $\overline{y} \in \mathscr{S}$, we obtain

$$\begin{aligned}
\langle J_s(L^{-t}y^1) + \tilde{\rho}(J\tilde{f}^0 + b) - J_s\tilde{f}^0, \overline{g} - L^{-t}y^1\rangle &= \\
= \langle y^1 + \tilde{\rho}(By^0 + c) - y^0, \overline{y} - y^1\rangle&.
\end{aligned} \tag{2.17}$$

On the other hand, it follows from the definition of the sequence $\{y^k\}$ that the vector y^1 is the solution of the variational inequality

$$y^1 \in \mathscr{S}, \quad \langle \tilde{G}y^1 + \tilde{\rho}(By^0 + c) - \tilde{G}y^0, y - y^1\rangle \geq 0, \quad \forall y \in \mathscr{S},$$

and, since $\tilde{G} = I$, we have

$$\langle y^1 + \tilde{\rho}(By^0 + c) - y^0, \overline{y} - y^1 \rangle \geq 0.$$

Thus, the inequality (2.16) follows from (2.17).
Now let us suppose that $\tilde{f}^i = L^{-t}y^i$ is the vector generated at the step i, for $i = 2, 3, \ldots, k$; with an analogous technique it is possible to prove that $\tilde{f}^{k+1} = L^{-t}y^{k+1}$ is the $(k+1)$-th vector generated by applying the method to (2.1) with $\tilde{f}^0 = L^{-t}y^0$, $G = J_s$ and $\rho = \tilde{\rho} = \frac{1}{\|B\|^2}$. □

The next Theorem establishes the convergence of the sequence $\{\tilde{f}^k\}$ to the solution of (2.1) and will be useful in the subsequent discussions.

Theorem 2.3. *Let B be defined as in* (2.4). *Given an arbitrary vector $\tilde{f}^0 \in \mathscr{K}$, projection method applied to* (2.1) *with parameters $G = J_s$ and $\rho = \tilde{\rho} = \frac{1}{\|B\|^2}$ generates a sequence $\{\tilde{f}^k\}$, $\tilde{f}^k \in \mathscr{K}$, such that*

$$\|\tilde{f}^k - h\| \leq \frac{\tilde{\lambda}^k}{1-\tilde{\lambda}} K(J_s)^{\frac{1}{2}} \|\tilde{f}^1 - \tilde{f}^0\|, \tag{2.18}$$

where $h \in \mathscr{K}$ is the solution of (2.1) *and*

$$\tilde{\lambda} = \left(1 - \frac{1}{\|B\|^2}\right)^{\frac{1}{2}} \leq \left(1 - \frac{1}{\left(1 + K(J_s)d_{J,J_s}\right)^2}\right)^{\frac{1}{2}}. \tag{2.19}$$

Proof. Since $h \in \mathscr{K}$ is the solution of (2.1) the vector $x = L^t h$ is the solution of (2.6). From Theorem 2.2 it results $\tilde{f}^k = L^{-t}y^k$ where $\{y^k\}$, $y^k \in \mathscr{S}$, is the sequence generated by applying projection method to (2.6) with starting point $y^0 = L^t \tilde{f}^0$ and parameters $\tilde{G}$ and $\tilde{\rho}$ as in (2.7). By (2.8) we have

$$\begin{aligned}
\|\tilde{f}^k - h\| &= \|L^{-t}y^k - L^{-t}x\| = \|L^{-t}(y^k - x)\| \\
&\leq \|L^{-t}\|\|y^k - x\| \leq \|L^{-t}\| \frac{\tilde{\lambda}^k}{1-\tilde{\lambda}} \|y^1 - y^0\| \\
&= \|L^{-t}\| \frac{\tilde{\lambda}^k}{1-\tilde{\lambda}} \|L^t(\tilde{f}^1 - \tilde{f}^0)\| \leq \|L^t\|\|L^{-t}\| \frac{\tilde{\lambda}^k}{1-\tilde{\lambda}} \|\tilde{f}^1 - \tilde{f}^0\| \\
&= K(L^t) \frac{\tilde{\lambda}^k}{1-\tilde{\lambda}} \|\tilde{f}^1 - \tilde{f}^0\| = \frac{\tilde{\lambda}^k}{1-\tilde{\lambda}} K(J_s)^{\frac{1}{2}} \|\tilde{f}^1 - \tilde{f}^0\|,
\end{aligned}$$

where $\tilde{\lambda}$ is defined in (2.9). □

Remark 2.1. The choice of parameter ρ examined in Theorem 2.3 is introduced when projection method [3] is applied to the Problem (2.6). This problem is obtained by preconditioning the original Problem (2.1) through the Cholesky decomposition of J_s. It is interesting to observe that analogous results can be obtained by using the decomposition $J_s = J_s^{\frac{1}{2}} J_s^{\frac{1}{2}}$, instead of Cholesky decomposition, to transform Problem (2.1). In this case, if we select $G = J_s$, the value of ρ is

$$\rho = \| J_s^{-\frac{1}{2}} J J_s^{-\frac{1}{2}} \|^{-2}$$

and it is easy to show that it satisfies the convergence condition obtained by considering projection method as a special case of the iterative scheme described in [4].

Now we may compare the convergence rate of $\{\tilde{f}^k\}$ with the convergence rate of the sequence $\{f^k\}$ obtained by applying to (2.1) the projection method with G and ρ as suggested in [3] ($G = J_s$, $\rho = \frac{\alpha}{\nu}$).
For the sequence $\{f^k\}$ we have

$$\|f^k - h\|_G \leq \frac{\lambda^k}{1-\lambda} \|f^1 - f^0\|_G, \tag{2.20}$$

with

$$\lambda \geq \left(1 - \frac{1}{K(J_s)^2}\right)^{\frac{1}{2}}, \tag{2.21}$$

and, taking into account that $G = J_s$, it results

$$\|f^k - h\| \leq \frac{\lambda^k}{1-\lambda} K(J_s)^{\frac{1}{2}} \|f^1 - f^0\|. \tag{2.22}$$

Thus, since the inequalities (2.19) and (2.21) imply

$$\tilde{\lambda} \leq \lambda \quad \text{when} \quad d_{J,J_s} \leq \frac{K(J_s) - 1}{K(J_s)},$$

by comparing (2.18) and (2.22) we may conclude that, when $d_{J,J_s} \leq \frac{K(J_s)-1}{K(J_s)}$, from the view point of the convergence rate it is better to select $\rho = \|B\|^{-2}$ than $\rho = \frac{\alpha}{\nu}$.
Moreover, from (2.19) it follows that, if J is sufficiently close to its symmetric part, the choices $G = J_s$ and $\rho = \|B\|^{-2}$ also guarantee a rapid convergence when it is not $K(J_s) \approx 1$.

The above analysis shows that the convergence rate of projection method is strictly dependent on the value of parameter ρ. In particular, we remark that when ρ is chosen as suggested in [3] the estimate (1.7) holds but, as proved in Theorem 2.3, there are problems in which a different and appropriate choice of parameter ρ ensures more promising estimates.

3. Numerical results.

We apply projection method to some problem of type (2.1).
We fix $f^0 \in \mathscr{K}$, $G = J_s$ and compare the number of iterations required by the method to approximate the solution in correspondence to the following choices of parameter ρ:

$$\rho = \frac{\alpha}{\nu} = \frac{\text{minimum eigenvalue of } J_s}{\text{maximum eigenvalue of } J^t G^{-1} J}, \tag{3.1}$$

$$\rho = \tilde{\rho} = \frac{1}{\|B\|^2}, \quad B = L^{-1} J L^{-t}, \quad L L^t = J_s. \tag{3.2}$$

The following criterion is used to stop the iterative process:

$$\frac{\|f^{i+1} - f^i\|}{\|f^{i+1}\|} < 10^{-6}, \tag{3.3}$$

where f^i and f^{i+1} are two successive iterates.
The algorithm is implemented in FORTRAN and all the experiments are performed on HP 9000-730 in double precision (eps $\approx 2.22 \times 10^{-16}$).
The matrix L and the quantities α, ν, $\|B\|$ are computed by using LINPACK library [6]. At each step, the variational inequality Problem (1.3) is reduced to a convex quadratic programming problem that is solved by using the subroutine VE07AD described in [11].

3.1. Dafermos' example

This problem is taken from [3]. Let $\mathscr{K}$ be the following convex subset of $\mathbb{R}^5$:

$$\mathscr{K} = \{f = (f_1, \dots, f_5)^t, \ f_i \geq 0, \ i = 1, \dots, 5; \\ f_1 + f_2 + f_3 = 210; \ f_4 + f_5 = 120\}$$

and let C be the affine function from $\mathscr{K}$ into $\mathbb{R}^5$ given by

$$C(f) = Jf + b,$$

where

$$J = \begin{bmatrix} 10 & 0 & 0 & 5 & 0 \\ 0 & 15 & 0 & 0 & 5 \\ 0 & 0 & 20 & 0 & 0 \\ 2 & 0 & 0 & 20 & 0 \\ 0 & 1 & 0 & 0 & 25 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 1000 \\ 950 \\ 3000 \\ 1000 \\ 1300 \end{bmatrix}.$$

The variational inequality

$$h \in \mathscr{K}, \quad \langle Jf + b, f - h \rangle \geq 0, \quad \forall f \in \mathscr{K},$$

admits the unique solution

$$h = (120, 90, 0, 70, 50)'.$$

The results obtained by applying the method with

$$f^0 = (70, 70, 70, 60, 60)', \quad G = J_s,$$

and the values of ρ considered in (3.1) and (3.2) are shown in Table 3.1.

Parameter ρ		$\rho = \frac{\alpha}{\nu} \approx 0.3406975$	$\rho = \frac{1}{\|B\|^2} \approx 0.9881579$
Number of iterations		28	7
Solution	f_1	120.000154680595	119.999999392942
	f_2	89.9998453194047	90.0000006070576
	f_3	0	0
	f_4	69.9999219948576	70.0000020844423
	f_5	50.0000780051424	49.9999979155577

Table 3.1

In this example we have $K(J_s) \approx 2.9$, $d_{J,J_s} \approx 0.077$, $\|B\| \approx 1.006$ and by (2.21) and (2.19) it results $\lambda > 0.93$ and $\tilde{\lambda} \approx 0.109$. Therefore, the parameters' choice suggested in Theorem 2.3 allows to generate a sequence which is quickly convergent to the solution $h \in \mathcal{K}$.
From the view point of convergence rate the result obtained with $\rho = \frac{1}{\|B\|^2}$ is like that obtained in [8] by updating ρ in each iteration.

3.2. *Test problems*

We construct test problems of the form

$$h \in \mathcal{K}, \quad \langle Jh + b, f - h\rangle \geq 0, \quad \forall f \in \mathcal{K}, \tag{3.4}$$

where J is a nonsymmetric positive definite $n \times n$ matrix, $b \in \mathbb{R}^n$ and $\mathcal{K}$ is given by

$$\mathcal{K} = \{ f \in \mathbb{R}^n \ / \ A_1 f = r, \quad A_2 f \geq g \}$$

where A_1 is an $m \times n$ matrix ($m \leq n$), $r \in \mathbb{R}^m$, A_2 is a $p \times n$ matrix and $g \in \mathbb{R}^p$.
We proceed as follows.

i) *Construction of* $\mathcal{K}$

We assume that the following data are given:

a) the dimension n of the problem,
b) the solution $h \in \mathbb{R}^n$,
c) the number m of equality constraints, $m \leq n$,
d) the number p of inequality constraints,
e) the number nva of inequality constraints that become active in the solution, $nva \leq n - m$.

Then we randomly generate the matrices A_1 and A_2 [18],[20] and, given the decomposition

$$A_2 = \begin{bmatrix} A_{21} \\ A_{22} \end{bmatrix} \begin{matrix} \} & nva \\ \} & p-nva \end{matrix}$$

we determine r, $g = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \begin{matrix} \} & nva \\ \} & p-nva \end{matrix}$ and a feasible point $f^0 \in \mathbb{R}^n$ such that

$$A_1 h = r, \quad A_1 f^0 = r, \quad A_2 f^0 > g, \quad A_{21} h = g_1, \quad A_{22} h > g_2.$$

ii) *Construction of* J *and* b

We suppose that the following data are given:

a) the maximum eigenvalue of the symmetric part of J, $\tau(J_s)$, and its spectral condition number $K(J_s)$,

b) the maximum singular value of the skew-symmetric part of J, $\sigma(J_a)$, its spectral condition number $K(J_a)$, and the number l, $l < n$, of the zero singular values of J_a (if n is even (odd) then l is even (odd)).

We obtain J as

$$J = J_s + J_a$$

where J_s and J_a are generated as follows.

Let V be a random orthogonal $n \times n$ matrix [18], [20] and Z be the diagonal $n \times n$ matrix of the eigenvalues of J_s:

$$\begin{aligned}
Z &= \mathrm{diag}(z_1, z_2, \dots, z_n),\\
z_1 &= \frac{\tau(J_s)}{K(J_s)}, \quad z_n = \tau(J_s),\\
z_j &= z_{j+1}\left(\frac{1}{K(J_s)}\right)^{\frac{1}{n-1}} \quad j = n-1, \dots, 2;
\end{aligned}$$

J_s is given by

$$J_s = VZV^t.$$

In order to obtain the matrix J_a [1] we generate a random orthogonal $n \times n$ matrix U, and the following $n \times n$ matrix:

$$R = \begin{bmatrix}
0 & & & & & & 0\\
& \ddots & & & & & \\
& & 0 & & & & \\
& & & \begin{pmatrix} 0 & \mu_1 \\ -\mu_1 & 0 \end{pmatrix} & & & \\
& & & & \begin{pmatrix} 0 & \mu_2 \\ -\mu_2 & 0 \end{pmatrix} & & \\
& & & & & \ddots & \\
0 & & & & & & \begin{pmatrix} 0 & \mu_k \\ -\mu_k & 0 \end{pmatrix}
\end{bmatrix}$$

in which

$$\begin{aligned}
k &= (n-l)/2,\\
\mu_1 &= \frac{\sigma(J_a)}{K(J_a)}, \quad \mu_k = \sigma(J_a),\\
\mu_j &= \mu_{j+1}\left(\frac{1}{K(J_a)}\right)^{\frac{1}{k-1}} \quad j = k-1, \dots, 2.
\end{aligned}$$

The matrix J_a is given by

$$J_a = URU^t.$$

Finally, the vector $b \in \mathbb{R}^n$ is obtained by using the Kuhn-Tucker conditions. Since $A_1 h = r$, $A_{21} h = g_1$ and $A_{22} h > g_2$, we set $u_i = 1$, $i = 1, \ldots, (m + nva)$, and consider

$$b = -Jh + [\,A_1^t \quad A_{21}^t\,] \begin{bmatrix} u_1 \\ \vdots \\ u_{m+nva} \end{bmatrix}.$$

The numerical experiments are conducted by fixing the convex $\mathscr{K}$ and considering the matrices J that correspond to several choices of $K(J_s)$ and $\sigma(J_a)$. We assume

$$n = 30, \quad m = 15, \quad p = 30, \quad nva = 10, \quad l = 0, \quad h = (1, \ldots, 1)^t,$$

$$K(J_a) = 10, \quad \|J_s\| = \tau(J_s) = 10, \quad K(J_s) = 1.1, 10, 30, 50,$$

and, for each value of $K(J_s)$,

$$\|J_a\| = \sigma(J_a) = 0.1, 0.5, 0.9, 1.3, 1.7, 2.1, 2.5.$$

For all test problems we apply projection method with starting point f^0, $G = J_s$ and the values (3.1) and (3.2) for ρ.
The average number of iterations (we generate 15 test functions for each value of $d_{J,J_s} = \frac{\|J_a\|}{\|J_s\|}$) required by the method in correspondence to the two values of ρ is shown in Figure 3.1, 3.2, 3.3, 3.4.

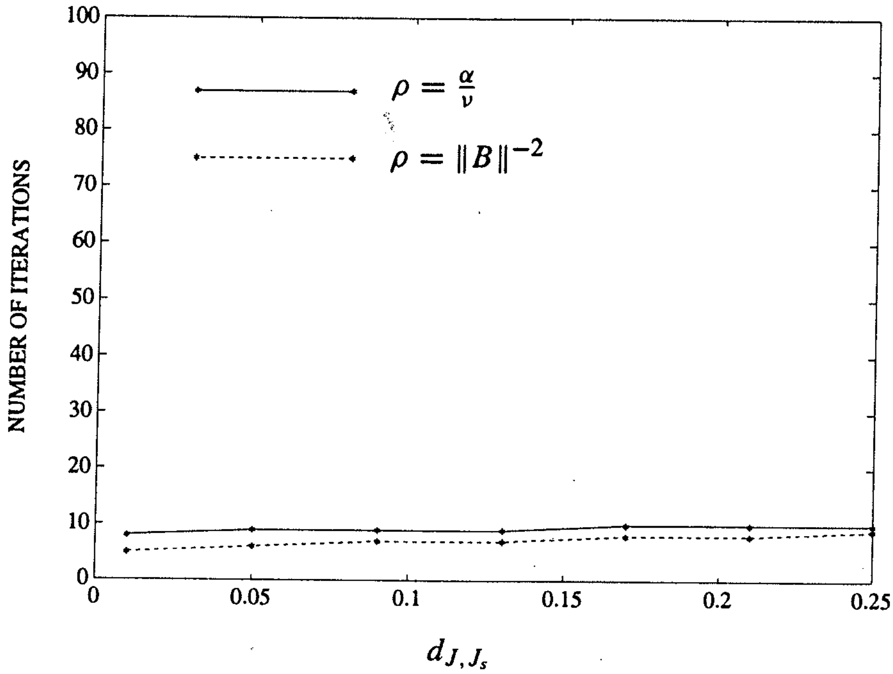

Fig. 3.1. $K(J_s) = 1.1$.

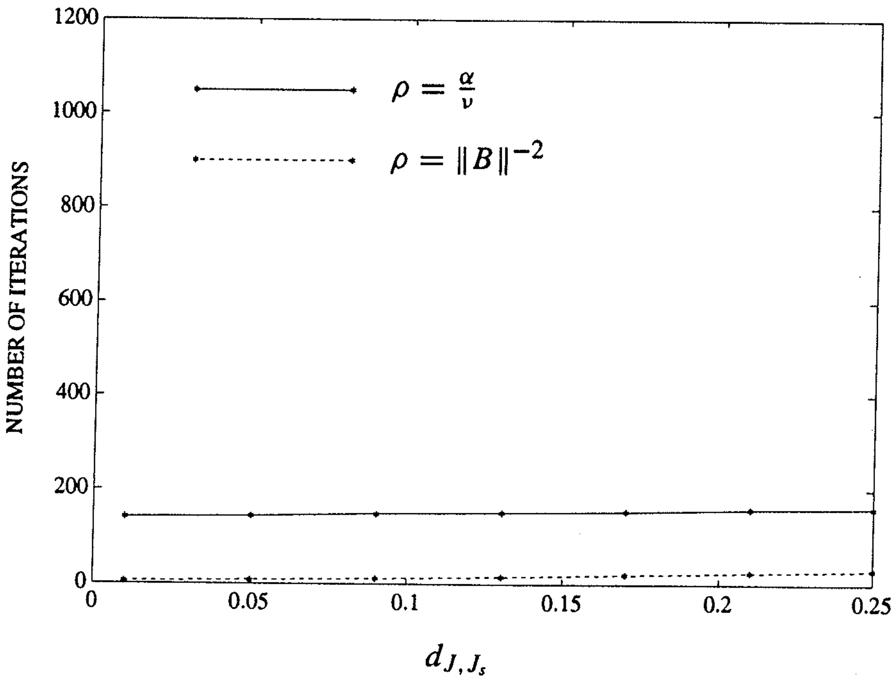

Fig. 3.2. $K(J_s) = 10$

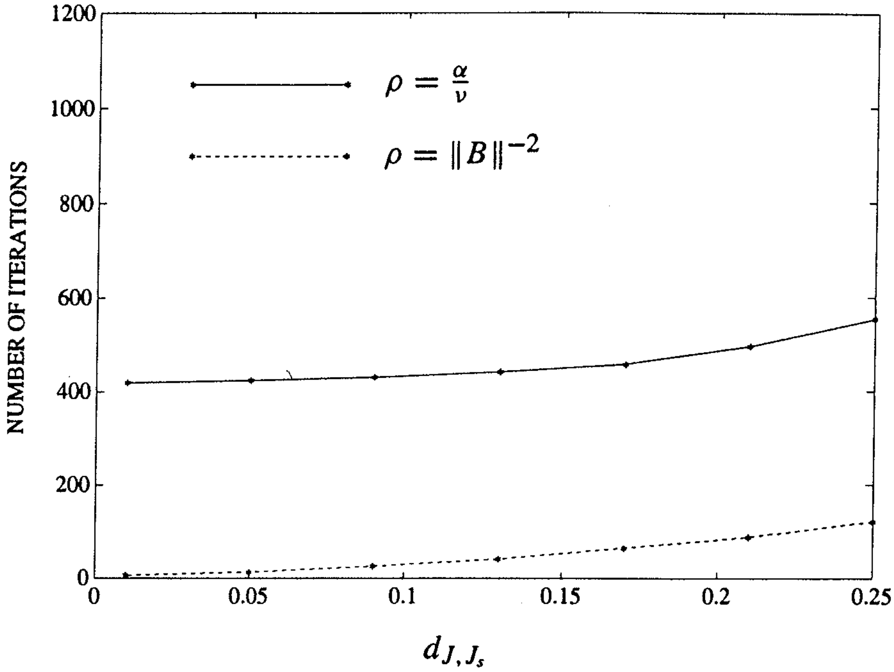

Fig. 3.3. $K(J_s) = 30$

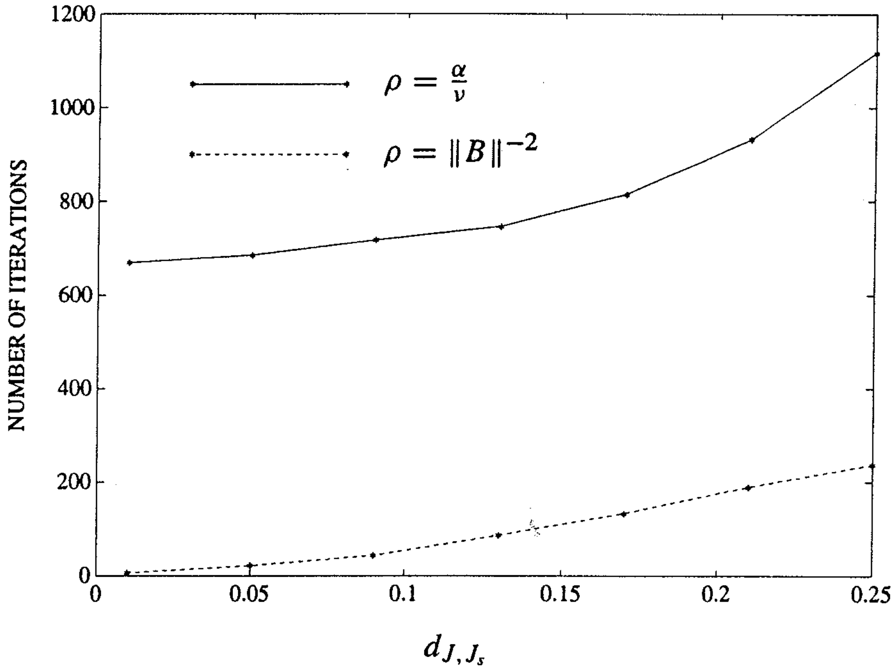

Fig. 3.4. $K(J_s) = 50$

REFERENCES

[1] R. Bellman, *Introduction to Matrix Analysis,* McGraw-Hill, New York, 1970.

[2] D.P. Bertsekas - E.M. Gafni, *Projection Methods for Variational Inequalities with Application to the Traffic Assignment Problem,* Math. Programming Study, 17 (1982), pp. 139-159.

[3] S. Dafermos, *Traffic Equilibrium and Variational Inequalities,* Transportation Sc., 14 (1980), pp. 42-54.

[4] S. Dafermos, *An Iterative Scheme for Variational Inequalities,* Math. Programming, 26 (1983), pp. 40-47.

[5] M. De Luca - A. Maugeri, *Quasi-Variational Inequalities and Applications to Equilibrium Problems with Elastic Demand,* in "Non-smooth Optimization and Related Topics", F.H. Clarke et al. editors, Plenum, New York, 1989, pp. 61-77.

[6] J.J. Dongarra - J.R. Bunch - C.B. Moler - G.W. Stewart, *Linpack Users' Guide,* SIAM, Philadelphia, 1979.

[7] C. Dupuis - J.M. Darveau, *The Convergence Conditions of Diagonalization and Projection Methods for Fixed Demand Asymmetric Network Equilibrium Problems,* Operations Research Letters, 5 (1986), pp. 149-155.

[8] C. Fisk - S. Nguyen, *Solution Algorithms for Network Equilibrium Models with Asymmetric User Costs,* Transportation Sci., 16 (1982), pp. 361-381.

[9] M. Fukushima, *Equivalent Differentiable Optimization Problems and Descent Methods for Asymmetric Variational Inequality Problems,* Math. Programming, 53 (1992), pp. 99-110.

[10] P.T. Harker - J.S. Pang, *Finite-Dimensional Variational Inequality and Nonlinear Complementarity Problems: A Survey of Theory, Algorithms and Applications,* Math. Programming, 48 (1990), pp. 161-220.

[11] Harwell Subroutine Library - A Catalogue of Subroutines (1985), *Tech. Report, Computer Science and Systems Division, Harwell Laboratory, HL85/1561 (C13), 1985.*

[12] O.G. Mancino - G. Stampacchia, *Convex Programming and Variational Inequalities,* J. Optim. Theory Appl., 9 (1972), pp. 3-23.

[13] P. Marcotte, *A New Algorithm for Solving Variational Inequalities with Application to the Traffic Assignment Problem,* Math. Programming, 33 (1985), pp. 339-351.

[14] P. Marcotte - J.P. Dussault, *A Sequential Linear Programming Algorithm for Solving Monotone Variational Inequalities,* SIAM Journal on Control and Optimization, 27 (1989), pp. 1260-1278.

[15] P. Marcotte - J.P. Dussault, *A Note on a Globally Convergent Newton Method for Solving Monotone Variational Inequalities,* Operations Research Letters, 6 (1987), pp. 35-42.

[16] A. Maugeri, *Convex Programming, Variational Inequalities, and Applications to the Traffic Equilibrium Problem,* Appl. Math. Optim., 16 (1987), pp. 169-185.

[17] J.S. Pang - D. Chan, *Iterative Methods for Variational and Complementarity problems,* Math. Programming, 24 (1982), pp. 284-313.

[18] V. Ruggiero - F. Zama, *Randomly Generated Equality-Constrained Quadratic Programming Test Problems,* CNR, Tech. Report on Project "Sistemi Informatici e Calcolo Parallelo" n. 1/82, Roma, 1991.

[19] G. Stampacchia, *Variational Inequalities,* Theory and Applications of Monotone Operators, Proceedings of the NATO Advanced Study Institute, Venice, Italy, Edizioni Oderisi, Gubbio, 1968, pp. 101-192.

[20] G.W. Stewart, *The Efficient Generation of Random Orthogonal Matrices whit an Application to Condition Estimators,* SIAM J. Numer. Anal. 17 (1980), pp. 403-409.

[21] K. Taji - M. Fukushima - T. Ibaraki, *A Globally Convergent Newton Method for Solving Strongly Monotone Variational Inequalities,* Math. Programming, 58 (1993), pp. 369-383.

[22] J.H. Wu - M. Florian - P. Marcotte, *A General Descent Framework for the Monotone Variational Inequality Problem,* Math. Programming, 61 (1993), pp. 281-300.

[23] L. Zanni, *On the Convergence Rate of two Projection Methods for Variational Inequalities in* $\mathbb{R}^n$, Calcolo, 29 (1992), pp. 193-212.

[24] D.L. Zhu - P. Marcotte, *An Extended Descent Framework for Variational Inequalities,* J. Optim. Theory Appl., 80 (1994), pp. 349-366.

Dipartimento di Matematica,
Università di Modena,
Via Campi 213/b,
41100 Modena (ITALY)
e-mail: zanni@dipmat.unimo.it