SIAM J. OPTIMIZATION
Vol. 2, No. 1, pp. 43–54, February 1992

004

ERROR BOUND AND CONVERGENCE ANALYSIS OF MATRIX SPLITTING ALGORITHMS FOR THE AFFINE VARIATIONAL INEQUALITY PROBLEM*

ZHI-QUAN LUO† AND PAUL TSENG‡

Abstract. Consider the affine variational inequality problem. It is shown that the distance to the solution set from a feasible point near the solution set can be bounded by the norm of a natural residual at that point. This bound is then used to prove linear convergence of a matrix splitting algorithm for solving the symmetric case of the problem. This latter result improves upon a recent result of Luo and Tseng that further assumes the problem to be monotone.

Key words. affine variational inequality, linear complementarity, error bound, matrix splitting, linear convergence

AMS(MOS) subject classifications. 49, 90

1. Introduction. Let M be an $n \times n$ matrix and let q be a vector in $\Re^n$, the n-dimensional Euclidean space. Let X be a polyhedral set in $\Re^n$. We consider the following *affine* variational inequality problem associated with M, q, and X:

$$\text{find an } x^* \in X \text{ satisfying } \langle x - x^*, Mx^* + q \rangle \geq 0 \quad \forall x \in X. \tag{1.1}$$

The problem (1.1) is well known in optimization and contains as special cases linear (and quadratic) programming, bimatrix games, etc. (see Cottle and Dantzig [CoD68]). When X is the nonnegative orthant in $\Re^n$, it is called the linear complementarity problem (LCP). We will not attempt to survey the literature on this problem, which is vast. Expository articles on the subject include [CoD68], [Eve71], [CGL80], and [Mur88]. For a discussion of variational inequality problems in general, see [Aus76], [BeT89], [CGL80], and [KiS80].

Let X^* denote the set of solutions of the affine variational inequality problem (1.1), which we assume from here on to be nonempty. It is well known (and not difficult to see from the convexity of X) that X^* is precisely the set of fixed points of the nonlinear mapping $x \mapsto [x - Mx - q]^+$, where $[\cdot]^+$ denotes the orthogonal projection onto X, i.e., $[x]^+ = \arg\min_{z \in X} \|x - z\|$ and $\|\cdot\|$ denotes the usual Euclidean norm in $\Re^n$. In other words, we have

$$X^* = \{ \, x^* \in \Re^n \mid x^* = [x^* - Mx^* - q]^+ \, \}. \tag{1.2}$$

(Our notation for the projection operator is nonstandard but has the advantage of simplicity.) Although in general X^* is not convex, it can be shown that X^* is the union of a finite collection of polyhedral sets (see (3.10)).

An important topic in the study of variational inequalities and complementarity problems concerns *error bounds* for estimating the closeness of a point to X^* (see [Pan87], [MaD88], [MaS86]). Such error bounds can serve as termination criteria for

* Received by the editors July 5, 1990; accepted for publication (in revised form) February 27, 1991. This work was supported by U.S. Army Research Office contract DAAL03-86-K-0171 (Center for Intelligent Control Systems), National Science Foundation grant NSF–DDM–8903385, and by a grant from the Science and Engineering Research Board of McMaster University.

† Room 225, Communications Research Laboratory, Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, L8S 4K1, Canada (luozq@sscvax.cis.mcmaster.ca).

‡ Department of Mathematics, GN-50, University of Washington, Seattle, Washington 98195 (tseng@math.washington.edu).

iterative algorithms and can be used to estimate the amount of error allowable in an inexact computation of the iterates (see [Pan86b]). Recently the authors [LuT90] showed that one such bound, based on the norm of the natural residual function

$$\|x - [x - Mx - q]^+\| \tag{1.3}$$

is also useful for analyzing the *rate* of convergence of iterative algorithms for solving (1.1). In particular, they showed that, for the problem of minimizing a certain convex essentially smooth function over a polyhedral set, a bound analogous to the above can be used as the basis for proving the linear convergence of a number of well-known iterative algorithms (applied to solve this problem).

The contribution of this paper is twofold: (i) we show that the error bound (1.3) holds locally for the affine variational inequality problem (1.1) for general M, thus extending a result of [LuT90, §2] for the case where M is symmetric positive semidefinite, (ii) we show, by using the above error bound, that if M is symmetric, then any matrix splitting algorithm using regular Q-splitting, applied to solve (1.1), is linearly convergent. (Here, by linear convergence, we mean linear convergence in the root sense of [OrR70].) This latter result extends the one in [LuT90, §5], which proved linear convergence for the same algorithm under the additional assumption that M is positive semidefinite. It also improves upon the results of Pang [Pan84, §4], [Pan86a, §2], which showed convergence (respectively, weak convergence) for a special case of the algorithm, i.e., one that solves LCP, under the additional assumption that M is nondegenerate (respectively, strictly copositive). Matrix splitting algorithms using regular Q-splitting represent an important class of algorithms for solving affine variational inequality problems and LCPs (see [LiP87]), so the resolution of their convergence (and their rate of convergence) is of great interest. (See §3 for a more detailed discussion of the subject.)

This paper proceeds as follows. In §2, we prove that an error bound based on (1.3) holds for all points in X near X^*. In §3, we consider the special case of (1.1) where M is symmetric and we use the bound of §2 to prove the linear convergence of matrix splitting algorithms using regular Q-splitting, applied to solve this problem. Finally, in §4, we give our conclusion and discuss possible extensions.

We adopt the following notations throughout. For any $x \in \Re^n$ and $y \in \Re^n$, we denote by $\langle x, y\rangle$ the Euclidean inner product of x with y. For any $x \in \Re^n$, we denote by $\|x\|$ the usual Euclidean norm of x, i.e., $\|x\| = \sqrt{\langle x, x\rangle}$. For any two subsets C_1, C_2 of $\Re^n$, we denote by $d(C_1, C_2)$ the usual Euclidean distance between the sets C_1 and C_2, that is,

$$d(C_1, C_2) = \inf_{x \in C_1, y \in C_2} \|x - y\|.$$

For any $k \times l$ matrix A, we denote by A^T the transpose of A, by $\|A\|$ the matrix norm of A induced by the vector norm $\|\cdot\|$ (i.e., $\|A\| = \max_{\|x\|=1} \|Ax\|$), by A_i the ith row of A and, for any subset $I \subseteq \{1, \cdots, k\}$, by A_I the submatrix of A obtained by removing all rows $i \notin I$ of A. Analogously, for any vector $x \in \Re^k$, we denote by x_i the ith coordinate of x and, for any subset $I \subseteq \{1, \cdots, k\}$, by x_I the vector with components x_i, $i \in I$ (with the x_i's arranged in the same order as in x).

2. A local error bound.

In this section we show that $d(x, X^*)$ can be bounded from above by the norm of $x - [x - Mx - q]^+$, the natural residual at x, whenever the latter quantity is small. Our proof, like the proof of Theorem 2.3 in [LuT90], exploits heavily the affine structure of the problem.

Since X is a polyhedral set, we can for convenience express it as

$$X = \{ x \in \Re^n \mid Ax \geq b \},$$

for some $m \times n$ matrix A and some $b \in \Re^m$. Then, for any $x \in X$, the vector $[x - Mx - q]^+$ is simply the unique vector z which, together with some multiplier vector $\lambda \in \Re^m$, satisfies the Kuhn–Tucker conditions

$$z - x + Mx + q - A^T\lambda = 0, \quad Az \geq b, \quad \lambda \geq 0, \tag{2.1}$$

$$A_i z = b_i \quad \forall i \in I(x), \quad \lambda_i = 0 \quad \forall i \notin I(x), \tag{2.2}$$

where we denote

$$I(x) = \{ i \in \{1, \cdots, n\} \mid A_i[x - Mx - q]^+ = b_i \}.$$

We say that an $I \subseteq \{1, \cdots, m\}$ is *active* at a vector $x \in X$ if $z = [x - Mx - q]^+$, together with some $\lambda \in \Re^m$, satisfies (2.1) and

$$A_i z = b_i \quad \forall i \in I, \quad \lambda_i = 0 \quad \forall i \notin I. \tag{2.3}$$

(Clearly, $I(x)$ is active at x for all $x \in X$.)

The following lemma, due originally to Hoffman [Hof52] (also see [Rob73], [MaS87]), will be used extensively in the analysis to follow.

LEMMA 2.1. *Let B be a $k \times l$ matrix, let C be an $h \times l$ matrix, and let d be a vector in $\Re^h$. There exists a scalar $\tau > 0$ depending on B and C only such that, for any $\bar{x}$ satisfying $C\bar{x} \geq d$ and any $e \in \Re^k$ such that the linear system $By = e$, $Cy \geq d$ is consistent, there is a point $\bar{y}$ satisfying $B\bar{y} = e$, $C\bar{y} \geq d$ with $\|\bar{x} - \bar{y}\| \leq \tau\|B\bar{x} - e\|$.*

We next have the following lemma, which roughly says that if $x \in X$ is sufficiently close to X^*, then those constraint indices that are active at x are also active at some element of X^*.

LEMMA 2.2. *There exists a scalar $\epsilon > 0$ such that, for any $x \in X$ with $\|x - [x - Mx - q]^+\| \leq \epsilon$, $I(x)$ is active at some $x^* \in X^*$.*

Proof. We argue by contradiction. If the claim does not hold, then there would exist an $I \subseteq \{1, \cdots, m\}$ and a sequence of vectors $\{x^1, x^2, \cdots\}$ in X satisfying $I(x^r) = I$ for all r and $x^r - z^r \to 0$, where we let $z^r = [x^r - Mx^r - q]^+$ for all r, and yet there is no $x^* \in X^*$ for which I is active at x^*.

For each r, consider the following linear system in x, z, and λ:

$$x - z - Mx + A^T\lambda = q, \quad Az \geq b, \quad \lambda \geq 0,$$

$$A_i z = b_i \quad \forall i \in I, \quad \lambda_i = 0 \quad \forall i \notin I,$$

$$x - z = x^r - z^r.$$

The above system is consistent since, by $I(x^r) = I$ and (2.1)–(2.2), (x^r, z^r) together with some $\lambda^r \in \Re^m$ is a solution of it. Then, by Lemma 2.1, it has a solution $(\hat{x}^r, \hat{z}^r, \hat{\lambda}^r)$ whose size is bounded by some constant (depending on A and M only) times the size of the right-hand side. Since the right-hand side of the above system is clearly bounded as $r \to \infty$, we have that $\{(\hat{x}^r, \hat{z}^r, \hat{\lambda}^r)\}$ is bounded. Moreover, every one of its cluster points, say $(x^\infty, z^\infty, \lambda^\infty)$, satisfies (cf. $x^r - z^r \to 0$)

$$x^\infty - z^\infty - Mx^\infty + A^T\lambda^\infty = q, \quad Az^\infty \geq b, \quad \lambda^\infty \geq 0,$$

$$A_i z^\infty = b_i \quad \forall i \in I, \quad \lambda_i^\infty = 0 \quad \forall i \notin I,$$

$$x^\infty - z^\infty = 0.$$

This shows that $x^\infty = [x^\infty - Mx^\infty - q]^+$ (cf. (2.1), (2.2)) and that I is active at x^∞ (cf. (2.1), (2.3)), a contradiction of our earlier hypothesis on I. □

By using Lemma 2.2, we can now establish the main result of this section.

THEOREM 2.3. *There exist scalars* $\epsilon > 0$ *and* $\tau > 0$ *such that*

$$d(x, X^*) \leq \tau \|x - [x - Mx - q]^+\|$$

for all $x \in X$ *with* $\|x - [x - Mx - q]^+\| \leq \epsilon$.

Proof. Let ϵ be the scalar given in Lemma 2.2. Consider any $x \in X$ satisfying the hypothesis of Lemma 2.2, and let $z = [x - Mx - q]^+$. Then, by (2.1) and (2.2), there exists some $\lambda \in \Re^m$ satisfying, together with x and z,

$$x - z - Mx + A^T\lambda = q, \quad Az \geq b, \quad \lambda \geq 0,$$

$$A_i z = b_i \quad \forall i \in I(x), \quad \lambda_i = 0 \quad \forall i \notin I(x).$$

By Lemma 2.2, there exists an $x^* \in X^*$ such that $I(x)$ is active at x^*, so the following linear system in (x^*, z^*, λ^*)

$$x^* - z^* - Mx^* + A^T\lambda^* = q, \quad Az^* \geq b, \quad \lambda^* \geq 0,$$

$$A_i z^* = b_i \quad \forall i \in I(x), \quad \lambda_i^* = 0 \quad \forall i \notin I(x), \quad x^* - z^* = 0,$$

is consistent. Moreover, every solution (x^*, z^*, λ^*) of this linear system satisfies $x^* = [x^* - Mx^* - q]^+$ (cf. (2.1), (2.2)) so, by (1.2), $x^* \in X^*$. Upon comparing the above two systems, we see that, by Lemma 2.1, there exists a solution (x^*, z^*, λ^*) to the second system such that

$$\|(x^*, z^*, \lambda^*) - (x, z, \lambda)\| \leq \tau \|x - z\|,$$

where τ is some scalar constant depending on A and M only. Hence

$$\|x^* - x\| \leq \tau \|x - z\|.$$

Since $x^* \in X^*$, so $d(x, X^*) \leq \|x^* - x\|$; this then completes the proof. □

Error bounds for estimating the distance from a point to the solution set, similar to that given in Theorem 2.3, have been fairly well studied. In fact, the same bound had been demonstrated by Pang [Pan87] and by Mathias and Pang [MaP90] to hold globally on X for the special cases of an LCP where M is, respectively, positive *definite* and a P-matrix. The bound has also been demonstrated by the authors [LuT90] to hold locally on X for the special case where M is symmetric and positive semidefinite. (This bound also extends to strongly monotone variational inequality problems [Pan87] and to problems of minimizing a certain convex, essentially smooth, function over a polyhedral set [LuT90].)

Alternative bounds have also been proposed by Mangasarian and Shiau [MaS86] for the special case of an LCP where M is positive semidefinite and for strongly convex programs [MaD88]. These alternative error bounds have the advantage that they hold globally *everywhere* (even for points outside X), whereas the bound of Theorem 2.3 holds only locally on X. Might the latter bound hold globally also? For general matrices M, the answer unfortunately is "no," as shown by an example of a nonsymmetric LCP furnished in [MaS86] (see Example 2.10 therein). What if M is symmetric? The answer is still "no," as shown by the following modification of Example 2.10 in [MaS86].

Example 2.1. Let

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad q = \begin{pmatrix} -1 \\ -2 \end{pmatrix}, \quad X = [0, \infty)^2.$$

It is easily checked that $X^* = \{ (1,1), (0,2) \}$. Let $x(\theta) = (\theta, 1)$, where $\theta \in [0, \infty)$. Then, as $\theta \to \infty$, we have $d(x(\theta), X^*) \to \infty$ but $\|x(\theta) - [x(\theta) - Mx(\theta) - q]^+\|$ remains bounded.

Subsequent to the writing of this paper, we learned that Theorem 2.3 can also be deduced from a result of Robinson [Rob81] on a locally upper Lipschitzian property of polyhedral multifunctions. More precisely, let $R : \Re^n \to \Re^n$ be the natural residual function given by

$$R(x) = x - [x - Mx - q]^+.$$

Then, the inverse of R is a polyhedral multifunction and thus, by Robinson's result [Rob81, Prop. 1], is locally upper Lipschitzian at the origin, that is, there exist scalars $\epsilon > 0$ and $\tau > 0$ such that

$$R^{-1}(z) \subseteq R^{-1}(0) + \tau \|z\| \mathcal{B},$$

for all z with $\|z\| \le \epsilon$, where $\mathcal{B}$ denotes the unit Euclidean ball in $\Re^n$. This statement is entirely equivalent to Theorem 2.3.

3. Linear convergence of matrix splitting algorithm for the symmetric case.

In this section we further assume that M is symmetric, in which case the variational inequality problem (1.1) may be formulated as a quadratic program of the form

$$\begin{aligned} &\text{minimize} \ \ f(x) \\ &\text{subject to } x \in X, \end{aligned} \tag{3.1}$$

where f is the quadratic function in $\Re^n$ given by

$$f(x) = \tfrac{1}{2}\langle x, Mx \rangle + \langle q, x \rangle. \tag{3.2}$$

It is easily seen that the set of stationary points for (3.1) is precisely X^* (cf. (1.2)), which, by assumption, is nonempty. Note, however, that f may not be bounded from below on X.

Let (B, C) be a *regular splitting* of M (see, e.g., [OrR70], [Kel65], [LiP87]), i.e.,

$$M = B + C, \qquad B - C \text{ is positive definite.} \tag{3.3}$$

Consider the following well-known iterative algorithm for solving (3.1), based on the splitting (B, C).

MATRIX SPLITTING ALGORITHM. At the rth iteration we are given an $x^r \in X$ (with $x^0 \in X$ chosen arbitrarily), and we compute a new iterate x^{r+1} in X satisfying

$$x^{r+1} = [x^{r+1} - Bx^{r+1} - Cx^r - q + h^r]^+, \tag{3.4}$$

where h^r is some n-vector.

The problem of finding an x^{r+1} satisfying (3.4) may be viewed as an affine variational inequality problem, whereby x^{r+1} is the vector in X that satisfies the variational inequality

$$\langle Bx^{r+1} + Cx^r + q - h^r, z - x^{r+1} \rangle \geq 0 \quad \forall z \in X. \tag{3.5}$$

In general, such an x^{r+1} need not exist, in which case the above algorithm would break down. To ensure that this does not happen, we will, following [LiP87], assume that

$$(B, C) \quad \text{is a } Q\text{-splitting} \tag{3.6}$$

or, equivalently, an x satisfying

$$x = [x - Bx - Cx^r - q]^+ \tag{3.7}$$

exists for all r. (For example, (B, C) is a Q-splitting if B is positive definite (see [BeT89], [KiS80]).)

The vector x^{r+1} may be viewed as an inexact solution of (3.7) with h^r as the associated "error" vector (so $h^r = 0$ corresponds to an exact solution of (3.7)). The idea of introducing an error vector in this manner is adopted from Mangasarian [Man90]. Let γ denote the smallest eigenvalue of the symmetric part of $B - C$ (which by hypothesis is positive) and let ϵ be a fixed scalar in $(0, \gamma/2]$. We will consider the following restriction on h^r governing how fast h^r tends to zero:

$$\|h^r\| \leq (\gamma/2 - \epsilon)\|x^r - x^{r+1}\| \quad \forall r. \tag{3.8}$$

The above restriction on h^r provides a *finite* termination criterion for any iterative method used to solve (3.7). To illustrate, fix r and suppose that we have a sequence of points converging to a solution of (3.7). (If B is positive definite, then such a sequence can be generated, for example, by applying the projection iteration $y := [y - \alpha(By + Cx^r + q)]^+$, with α a suitably chosen positive stepsize.) Suppose that the limit is not x^r (otherwise x^r is already in X^*) and let $F : \Re^n \mapsto \Re^n$ be the continuous function given by $F(y) = [y - By - Cx^r - q]^+$. Then, for all points y sufficiently far along in this sequence we have

$$\|(I - B)(y - F(y))\| \leq (\gamma/2 - \epsilon)\|x^r - F(y)\|.$$

(This is because the limit, say $\bar{y}$, is not equal to x^r and satisfies $\bar{y} = F(\bar{y})$.) Take any such point y and set

$$h^r = (I - B)(y - F(y)), \qquad x^{r+1} = F(y).$$

Then, h^r and x^{r+1} satisfy (3.4) and (3.8).

The above matrix splitting algorithm was first proposed by Pang [Pan82], based on the works of Hildreth [Hil57], Cryer [Cry71], Mangasarian [Man77], and others. (Actually, Pang considered the somewhat simpler case of an LCP with no error vector, i.e., X is the nonnegative orthant in $\Re^n$ and $h^r = 0$ for all r.) This algorithm has been studied extensively (see [LiP87], [LuT90], [LuT91], [Man77], [Man90], [Pan82], [Pan84], [Pan86a], and references therein), but, owing to the possible unboundedness of the set of stationary points, its convergence was very difficult to establish and was typically shown under restrictive assumptions on the problem (such as that the stationary point is unique). It was shown only recently that, if M is positive semidefinite (in addition to being symmetric) and f given by (3.2) is bounded from below on X, then the iterates generated by this algorithm converge to a stationary point [LuT91] with a rate of convergence that is at least linear [LuT90, §5]. In this section we show that the same linear convergence result holds for *any* symmetric M, and thus we resolve the issue of convergence (and rate of convergence) for this algorithm on symmetric problems. The convergence of this algorithm for the special case of a symmetric LCP has been studied by Pang (see [Pan84, §4] and [Pan86a, §2]). However, Pang did not analyze the rate of convergence of the algorithm and his convergence results require restrictive assumptions on the problem, such as that the set of stationary points be finite.

The line of our analysis follows that outlined in [LuT90] (also see [LuT92] for a similar analysis) and is based on using the error bound of Theorem 2.3 to show that, asymptotically, the objective function value, evaluated at the new iterate x^{r+1} and at some stationary point, differ by only an order of $\|x^{r+1} - x^r\|^2$ (see (3.19)). This then enables us to show that the objective function values converge at least linearly, from which one can deduce that the iterates converge at least linearly. (This is the main motivation for considering the symmetric case, so that an objective function exists and can be used to monitor the progress of the algorithm. The algorithm itself is well defined whether M is symmetric or not.) On the other hand, because f is not convex and the set of stationary points X^* is not necessarily convex or even connected, a new analysis, different from that in [LuT90], is needed to show the above relation.

We begin our analysis by giving, in the lemma below, a characterization of the connected components of X^* and the behaviour of f over these connected components.

LEMMA 3.1. *Suppose that M is symmetric. Let $C_1, C_2, \cdots, C_t$ denote the connected components of X^*, where t is some positive integer. Then,*

$$X^* = \bigcup_{i=1}^{t} C_i,$$

and the following hold:

(a) *Each C_i is the union of a finite collection of polyhedral sets.*

(b) *The C_i's are properly separated from one another, that is, $d(C_i, C_j) > 0$ for all $i \neq j$.*

(c) *f given by (3.2) is constant on each C_i.*

Proof. Since X is a polyhedral set, we can express it as

$$X = \{\, x \in \Re^n \mid Ax \geq b \,\}$$

for some $m \times n$ matrix A and some $b \in \Re^m$. For each $I \subseteq \{1, 2, \cdots, m\}$, let

$$X_I = \{ \, x \mid Ax \geq b, \; A_I x = b_I, \; Mx + q = A^T \lambda \quad \text{for some } \lambda \in [0,\infty)^m \text{ with } \lambda_i = 0 \;\; \forall i \notin I \, \}. \tag{3.9}$$

Then, each X_I simply comprises those elements of X^* at which I is active (see (2.1) and (2.2)), so it readily follows that

$$X^* = \bigcup_{I \subseteq \{1,\cdots,m\}} X_I. \tag{3.10}$$

Moreover, each X_I, if nonempty, is a polyhedral set. We claim that f is *constant* on each nonempty X_I. To see this, fix any $I \subseteq \{1,\cdots,m\}$ for which X_I is nonempty. Let x and y be any two elements of X_I (possibly equal). Since $x \in X_I$ and $y \in X_I$, we have from (3.9) that $A_I(x-y) = 0$ and there exists some $\lambda \in [0,\infty)^m$ with $My + q = (A_I)^T \lambda_I$. Then we have from (3.2) that

$$\begin{aligned} f(x) - f(y) &= \langle My + q, x - y \rangle + \tfrac{1}{2}\langle x - y, M(x-y) \rangle \\ &= \langle (A_I)^T \lambda_I, x - y \rangle + \tfrac{1}{2}\langle x - y, M(x-y) \rangle \\ &= \langle \lambda_I, A_I(x-y) \rangle + \tfrac{1}{2}\langle x - y, M(x-y) \rangle \\ &= \tfrac{1}{2}\langle x - y, M(x-y) \rangle. \end{aligned}$$

By symmetry, we also have

$$f(y) - f(x) = \tfrac{1}{2}\langle x - y, M(x-y) \rangle,$$

and thus $f(x) = f(y)$. Since the above choice of x and y was arbitrary, then $f(y) = f(x)$ for all $x \in X_I, y \in X_I$.

Since each X_I is connected, it follows from (3.10) that each C_i is the union of a finite collection of nonempty X_I's. Since the nonempty X_I's are polyhedral and the C_i's are, by definition, mutually disjoint, this then proves parts (a) and (b). Since f is constant on each X_I, this also proves part (c). □

(Lemma 3.1(c) is quite remarkable, since the gradient of f does not need to be constant on each C_i, as can be seen from an example.)

By using Theorem 2.3 and Lemma 3.1, we can now prove the main result of this section. (The first third of our proof follows closely that of Theorem 5.1 in [LuT90].)

THEOREM 3.2. *Suppose that M is symmetric and that f given by* (3.2) *is bounded from below on X. Let $\{x^r\}$ be iterates generated by the matrix splitting algorithm* (3.3), (3.4), (3.6), (3.8). *Then $\{x^r\}$ converges at least linearly* (*in the root sense*) *to an element of X^*.*

Proof. First we claim that

$$f(x^{r+1}) - f(x^r) \leq -\epsilon \|x^{r+1} - x^r\|^2 \quad \forall r. \tag{3.11}$$

To see this, fix any r. Since the variational inequality (3.5) holds, then, by plugging in x^r for z in (3.5), we obtain

$$\langle Bx^{r+1} + Cx^r + q - h^r, x^{r+1} - x^r \rangle \leq 0.$$

Also, from $M = B + C$ (cf. (3.3)) and the definition of f (cf. (3.2)), we have that

$$f(x^{r+1}) - f(x^r) = \langle Bx^{r+1} + Cx^r + q, x^{r+1} - x^r \rangle + \langle x^{r+1} - x^r, (C - B)(x^{r+1} - x^r) \rangle / 2.$$

Combining the above two relations then gives

$$\begin{aligned} f(x^{r+1}) - f(x^r) &\le \langle h^r, x^{r+1} - x^r\rangle + \langle x^{r+1} - x^r, (C-B)(x^{r+1} - x^r)\rangle/2 \\ &\le \|h^r\|\|x^{r+1} - x^r\| - \gamma\|x^{r+1} - x^r\|^2/2 \\ &\le -\epsilon\|x^{r+1} - x^r\|^2, \end{aligned}$$

where the last inequality follows from (3.8). Thus, (3.11) holds.

Next we claim that there exists a scalar constant $\kappa_1 > 0$ for which

$$\|x^r - [x^r - Mx^r - q]^+\| \le \kappa_1\|x^{r+1} - x^r\| \quad \forall r. \tag{3.12}$$

To see this, fix any r. From (3.4) we have that

$$\begin{aligned} \|x^r - [x^r - Mx^r - q]^+\| &= \|x^r - [x^r - Mx^r - q]^+ - x^{r+1} \\ &\qquad + [x^{r+1} - Bx^{r+1} - Cx^r - q + h^r]^+\| \\ &\le \|x^r - x^{r+1}\| + \|[x^r - Mx^r - q]^+ \\ &\qquad - [x^{r+1} - Bx^{r+1} - Cx^r - q + h^r]^+\| \\ &\le 2\|x^r - x^{r+1}\| + \|Mx^r - Bx^{r+1} - Cx^r + h^r\| \\ &\le 2\|x^r - x^{r+1}\| + \|B(x^r - x^{r+1})\| + \|h^r\| \\ &\le (2 + \|B\| + \gamma/2)\|x^r - x^{r+1}\|, \end{aligned}$$

where the second inequality follows from the nonexpansive property of the projection operator $[\cdot]^+$, the third inequality follows from $M = B + C$, and the last inequality follows from (3.8). This shows that (3.12) holds with $\kappa_1 = 2 + \|B\| + \gamma/2$.

Since f is bounded from below on X, (3.11) implies

$$\|x^{r+1} - x^r\| \to 0. \tag{3.13}$$

Then we have from (3.12) that $\|x^r - [x^r - Mx^r - q]^+\| \to 0$, so, by Theorem 2.3 (and using (3.12)), there exist a scalar constant $\kappa_2 > 0$ and an index r_1 such that

$$d(x^r, X^*) \le \kappa_2\|x^{r+1} - x^r\| \quad \forall r \ge r_1.$$

For each r, let y^r be any element of X^* attaining $\|y^r - x^r\| = d(x^r, X^*)$. Then the above relation implies

$$\|y^r - x^r\| \le \kappa_2\|x^{r+1} - x^r\| \quad \forall r \ge r_1, \tag{3.14}$$

which, when combined with (3.13), yields

$$y^r - x^r \to 0. \tag{3.15}$$

Let $C_1, C_2, \cdots, C_t$ denote the connected components of X^*, where t is some positive integer. By Lemma 3.1 (b), the C_i's are properly separated from one another. Since $y^r \in X^*$ for all r and, by (3.13) and (3.15), $y^r - y^{r+1} \to 0$, this implies that the sequence $\{y^r\}$ eventually settles down at one of the C_i's. In other words, there exists a $k \in \{1, \cdots, t\}$ and a scalar $r_2 \ge r_1$ such that

$$y^r \in C_k \quad \forall r \ge r_2.$$

By Lemma 3.1 (c), f is constant on C_k. Let us denote this constant by f^∞. Then the above relation implies

$$f(y^r) = f^\infty \qquad \forall r \ge r_2. \tag{3.16}$$

For any $r \geq r_2$ we have from $y^r \in X^*$ and $x^r \in X$ that $\langle My^r + q, x^r - y^r \rangle \geq 0$ and from the Mean Value Theorem (also using (3.2)) that $f(y^r) - f(x^r) = \langle M\psi^r + q, y^r - x^r \rangle$, for some n-vector ψ^r lying on the line segment joining y^r with x^r. Upon summing these two relations and using (3.16), we obtain

$$\begin{aligned} f^\infty - f(x^r) &\leq \langle M\psi^r - My^r, y^r - x^r \rangle \\ &\leq \|M\| \|\psi^r - y^r\| \|y^r - x^r\| \\ &\leq \|M\| \|y^r - x^r\|^2. \end{aligned}$$

This, together with (3.15), yields

$$\liminf_{r \to \infty} f(x^r) \geq f^\infty. \tag{3.17}$$

We now show that $f(x^r) \to f^\infty$ and estimate the speed at which this convergence takes place. Fix any $r \geq r_2$. Since $r \geq r_1$ (cf. $r_2 \geq r_1$) so that (3.14) holds, this implies

$$\begin{aligned} \langle My^r + q, x^{r+1} - y^r \rangle &\leq \langle My^r + q, x^{r+1} - y^r \rangle + \langle Bx^{r+1} + Cx^r + q - h^r, y^r - x^{r+1} \rangle \\ &= \langle B(x^{r+1} - x^r) + M(x^r - y^r) - h^r, y^r - x^{r+1} \rangle \\ &\leq \left(\|B\| \|x^{r+1} - x^r\| + \|M\| \|x^r - y^r\| + \|h^r\| \right) \|y^r - x^{r+1}\| \\ &\leq \left(\|B\| \|x^{r+1} - x^r\| + \|M\| \kappa_2 \|x^{r+1} - x^r\| + \|h^r\| \right) \\ &\quad \times (\kappa_2 + 1) \|x^{r+1} - x^r\| \\ &\leq \left(\|B\| + \|M\| \kappa_2 + \gamma/2 \right) (\kappa_2 + 1) \|x^{r+1} - x^r\|^2, \end{aligned} \tag{3.18}$$

where the first inequality follows from (3.5) with z set to y^r, the equality follows from $C = M - B$ (cf. (3.3)), the third inequality follows from (3.14), and the last inequality follows from (3.8). For convenience, let κ_3 denote the scalar constant on the right-hand side of (3.18). Then we obtain from (3.16) that

$$\begin{aligned} f(x^{r+1}) - f^\infty &= f(x^{r+1}) - f(y^r) \\ &= \langle My^r + q, x^{r+1} - y^r \rangle + \tfrac{1}{2} \langle x^{r+1} - y^r, M(x^{r+1} - y^r) \rangle \\ &\leq \kappa_3 \|x^{r+1} - x^r\|^2 + \tfrac{1}{2} \|M\| \|x^{r+1} - y^r\|^2 \\ &\leq \left(\kappa_3 + \tfrac{1}{2} \|M\| (\kappa_2 + 1)^2 \right) \|x^{r+1} - x^r\|^2, \end{aligned} \tag{3.19}$$

where the second equality follows from (3.2), the first inequality follows from (3.18), and the last inequality follows from (3.14).

Let κ_4 denote the scalar constant on the right-hand side of (3.19). Then (3.11) and (3.19) yield

$$\begin{aligned} f(x^{r+1}) - f^\infty &\leq \kappa_4 \|x^{r+1} - x^r\|^2 \\ &\leq \frac{\kappa_4}{\epsilon} (f(x^r) - f(x^{r+1})) \quad \forall r \geq r_2. \end{aligned}$$

Upon rearranging terms, we find that

$$\left(1 + \frac{\kappa_4}{\epsilon} \right) (f(x^{r+1}) - f^\infty) \leq \frac{\kappa_4}{\epsilon} (f(x^r) - f^\infty) \quad \forall r \geq r_2.$$

On the other hand, we have from (3.17) and the fact that $f(x^r)$ is monotonically decreasing with r (cf. (3.11)) that $f(x^r) \geq f^\infty$ for all r, so the above relation implies that $\{f(x^r)\}$ converges at least linearly (in the root sense) to f^∞. By (3.11), $\{x^r\}$ also converges at least linearly (in the root sense). Since $d(x^r, X^*) \to 0$ (cf. (3.15)), the point to which $\{x^r\}$ converges is an element of X^*. □

Note that we can allow the matrix splitting (B, C) to vary from iteration to iteration, provided that the eigenvalues of the symmetric part of $B - C$ are bounded away from zero and that $\|B\|$ is bounded.

Also note that because f is not convex, the point to which the iterates converge need not be an optimal solution of (3.1). (Finding such an optimal solution is certainly desirable.) On the other hand, it is easily seen from Lemma 3.1(c) and the fact that the f value of the iterates are monotonically decreasing that *local* convergence to an optimal solution holds. In other words, if the initial iterate (namely, x^0) is sufficiently close to the optimal solution set of (3.1), then the point to which the iterates converge is an optimal solution of (3.1).

4. Concluding remarks. In this paper, we have shown that a certain error bound holds locally for the affine variational inequality problem. By using this bound, we are able to prove the linear convergence of matrix splitting algorithms using regular Q-splitting for the symmetric case of the problem.

There are a number of open questions raised by our work. The first question concerns whether the error bound studied here holds globally. Example 2.1 shows that it does not hold globally even when M is symmetric. But what if M, in addition, is positive semidefinite? A "yes" answer to this question would allow us to show *global* linear convergence for the matrix splitting algorithm of §3 on symmetric monotone problems. Also, our convergence result (Theorem 3.2) asserts convergence only when f given by (3.2) is bounded from below on X. If this were not the case, could something meaningful about convergence still be said? Another question concerns whether other error bounds, such as those proposed in [MaD88] and [MaS86], can be used to analyze the convergence of an iterative algorithm, as is done here. Also, can the analysis of §3 be extended to the nonsymmetric case by finding an appropriate "objective function" to work with? Or to the simpler case of a nonsymmetric LCP? (It is well known that any LCP can be converted to a quadratic program. However, except under certain conditions (see [CPV89]), the set of solutions for the former does not need to coincide with the set of stationary points for the latter.)

It would also be worthwhile to find other problem classes for which the error bound studied here holds. Then, we can be hopeful of proving linear convergence results for these other problems.

Acknowledgment. We thank Professor O. L. Mangasarian for pointing out to us the references [MaD88], [MaS87] on error bounds and the examples therein. We thank Professor J.-S. Pang for bringing to our attention the paper [Rob81].

REFERENCES

[Aus76] A. AUSLENDER, *Optimisation: Méthodes Numériques*, Masson, Paris, New York, Barcelona, Milan, 1976.

[BeT89] D. P. BERTSEKAS AND J. N. TSITSIKLIS, *Parallel and Distributed Computation: Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[CoD68] R. W. COTTLE AND G. B. DANTZIG, *Complementary pivot theory of mathematical programming*, Linear Algebra Appl., 1 (1968), pp. 103–125.

[CGL80] R. W. COTTLE, F. GIANNESSI, AND J.-L. LIONS, *Variational Inequalities and Complementarity Problems: Theory and Applications*, John Wiley, New York, NY, 1980.

[CPV89] R. W. COTTLE, J.-S. PANG, AND V. VENKATESWARAN, *Sufficient matrices and the linear complementarity problem*, Linear Algebra Appl., 114/115 (1989), pp. 231–249.

[Cry71] C. W. CRYER, *The solution of a quadratic programming problem using systematic overrelaxation*, SIAM J. Control Optim., 9 (1971), pp. 385–392.

[Eve71] B. C. EAVES, *The linear complementarity problem*, Management Sci., 17 (1971), pp. 612–635.

[Hil57] C. HILDRETH, *A quadratic programming procedure*, Naval Res. Logist. Quart., 4 (1957), pp. 79–85; see also *Erratum*, Naval Res. Logist. Quart., 4 (1957), p. 361.

[Hof52] A. J. HOFFMAN, *On approximate solutions of systems of linear inequalities*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 263–265.

[Kel65] H. B. KELLER, *On the solution of singular and semidefinite linear systems by iteration*, SIAM J. Numer. Anal., 2 (1965), pp. 281–290.

[KiS80] D. KINDERLEHRER AND G. STAMPACCHIA, *An Introduction to Variational Inequalities and Applications*, Academic Press, New York, 1980.

[LiP87] Y. Y. LIN AND J.-S. PANG, *Iterative methods for large convex quadratic programs: A survey*, SIAM J. Control Optim., 25 (1987), pp. 383–411.

[LuT90] Z.-Q. LUO AND P. TSENG, *On the linear convergence of descent methods for convex essentially smooth minimization*, Laboratory for Information and Decision Systems Report No. P–1979, Massachusetts Institute of Technology, Cambridge, MA (June 1990); SIAM J. Control Optim., 30 (1992), to appear.

[LuT91] ———, *On the convergence of a matrix-splitting algorithm for the symmetric monotone linear complementarity problem*, SIAM J. Control Optim., 29 (1991), pp. 1037–1060.

[LuT92] ———, *On the convergence of the coordinate descent method for convex differentiable minimization*, J. Optim. Theory Appl., 72 (1992), pp. 7–35.

[Man77] O. L. MANGASARIAN, *Solution of symmetric linear complementarity problems by iterative methods*, J. Optim. Theory Appl., 22 (1977), pp. 465–485.

[Man90] ———, *Convergence of iterates of an inexact matrix splitting algorithm for the symmetric monotone linear complementarity problem*, SIAM J. Optimization, 1 (1991), pp. 114–122.

[MaD88] O. L. MANGASARIAN AND R. DE LEONE, *Error bounds for strongly convex programs and (super)linearly convergent iterative schemes for the least 2-norm solution of linear programs*, Appl. Math. Optim., 17 (1988), pp. 1–14.

[MaS86] O. L. MANGASARIAN AND T.-H. SHIAU, *Error bounds for monotone linear complementarity problems*, Math. Programming, 36 (1986), pp. 81–89.

[MaS87] ———, *Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems*, SIAM J. Control Optim., 25 (1987), pp. 583–595.

[MaP90] R. MATHIAS AND J.-S. PANG, *Error bounds for the linear complementarity problem with a P-matrix*, Linear Algebra Appl., 132 (1990), pp. 123–136.

[Mur88] K. G. MURTY, *Linear Complementarity, Linear and Nonlinear Programming*, Helderman-Verlag, Berlin, 1988.

[OrR70] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[Pan82] J.-S. PANG, *On the convergence of a basic iterative method for the implicit complementarity problem*, J. Optim. Theory Appl., 37 (1982), pp. 149–162.

[Pan84] ———, *Necessary and sufficient conditions for the convergence of iterative methods for the linear complementarity problem*, J. Optim. Theory Appl., 42 (1984), pp. 1–17.

[Pan86a] ———, *More results on the convergence of iterative methods for the symmetric linear complementarity problem*, J. Optim. Theory Appl., 49 (1986), pp. 107–134.

[Pan86b] ———, *Inexact Newton methods for the nonlinear complementarity problem*, Math. Programming, 36 (1986), pp. 54–71.

[Pan87] ———, *A posteriori error bounds for the linearly-constrained variational inequality problem*, Math. Oper. Res., 12 (1987), pp. 474–484.

[Rob73] S. M. ROBINSON, *Bounds for error in the solution set of a perturbed linear program*, Linear Algebra Appl., 6 (1973), pp. 69–81.

[Rob81] ———, *Some continuity properties of polyhedral multifunctions*, Math. Programming Stud., 14 (1981), pp. 206–214.