MATHEMATICS OF OPERATIONS RESEARCH
Vol. 21, No. 1, February 1996
Printed in U.S.A.

A PIVOTAL METHOD FOR AFFINE VARIATIONAL INEQUALITIES

MENGLIN CAO AND MICHAEL C. FERRIS

We explain and justify a path-following algorithm for solving the equations $A_C(x) = a$, where A is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$, C is a polyhedral convex subset of $\mathbb{R}^n$, and A_C is the associated normal map. When A_C is coherently oriented, we are able to prove that the path following method terminates at the unique solution of $A_C(x) = a$, which is a generalization of the well known fact that Lemke's method terminates at the unique solution of LCP (q, M) when M is a $\mathbf{P}$ = matrix. Otherwise, we identify two classes of matrices which are analogues of the class of copositive-plus and L-matrices in the study of the linear complementarity problem. We then prove that our algorithm processes $A_C(x) = a$ when A is the linear transformation associated with such matrices. That is, when applied to such a problem, the algorithm will find a solution unless the problem is infeasible in a well specified sense.

1. Introduction. This paper is concerned with the Affine Variational Inequality problem. The problem can be described as follows. Let C be a polyhedral set and let A be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$. We wish to find $z \in C$ such that

$$\text{(AVI)} \qquad \langle A(z) - a, y + z \rangle \geq 0, \quad \forall y \in C.$$

This problem has appeared in the literature in several disguises. The first is the linear generalized equation, that is

$$\text{(GE)} \qquad 0 \in A(z) - a + \partial\psi_C(z),$$

where $\psi_C(\cdot)$ is the indicator function of the set C defined by

$$\psi_C(z) := \begin{cases} 0 & \text{if } z \in C, \\ \infty & \text{if } z \notin C. \end{cases}$$

It can be easily shown that $\partial\psi_C(z) = N_C(z)$, the normal cone to C at z, if $z \in C$ and is empty otherwise, and hence (AVI) is equivalent to (GE). The solutions of such problems arise for example in the determination of a Newton-type method for generalized equations.

The problem has also been termed the linear stationary problem and we refer the reader to the work of Yamamoto (1987), Talman and Yamamoto (1989), and Dai, van der Laar, Talman and Yamamoto (1991) for several methods for the solution of this problem either over a bounded polyhedron or a pointed convex polyhedron. These methods are simplicial in nature and require a triangularization of the set C in order to general the path. Our method does not require such a procedure and is applicable to any polyhedral set C, not just the bounded polyhedra or pointed convex polyhedra. Another algorithm for this problem is given in Dai and Talman (1993). This is closer

Received March 8, 1993; revised November 1, 1993.
AMS 1991 subject classification. Primary: 90C33; Secondary: 49J40, 65K10.
OR/MS Index 1978 subject classification. Primary: 622 Programming/Complementarity.
Key words. Affine variational inequality, normal map, path-following algorithm.

0364-765X/96/2101/0044/$01.25

to our method in that only one pivot step is required to generate each segment of the path. However, as described, their path lies entirely within the set C, whereas the path generated by our algorithm moves through $\mathbb{R}^n$ (on the "normal manifold"), not through the feasible set. Further, the analysis of their algorithm is not as general as will be given for our algorithm, in that our algorithm processes problems generated from L-matrices, a new class of matrices defined in this paper. This class includes as subclasses all those processed by the above algorithms, as well as some nontrivial classes not included in the analysis of these algorithms. Other related methods for finding stationary points of affine functions on polyhedral sets are given in Eaves (1978a and b). In these papers, either the set C is assumed to have an extreme point and the set C has to lie in the positive orthant, or the feasible set C is perturbed. Our algorithm does not require either of these assumptions, although it does perform preliminary steps to modify C so that it has an extreme point. These preliminary steps are easily implementable and essentially factor out any lines in the set C (see §3) and can be used to generalize the applicability of several of the algorithms mentioned above.

In this work we will use the notion of a normal map due to Robinson (1992). The normal map, relating to a function $F: \mathbb{R}^n \to \mathbb{R}^n$ and a nonempty, closed, convex set C, is defined as

$$F_C(x) := F(\pi_C(x)) + x - \pi_C(x),$$

where $\pi_C(x)$ is the projection (with respect to the Euclidean norm) of x onto the set C. Throughout this paper, we will be concerned with solving affine normal maps, that is, $F \equiv A$ is a linear map, C is a polyhedral set and the solution x satisfies

$$\text{(NE)} \qquad A_C(x) = a.$$

Note that (NE) is equivalent to (AVI), since if $A_C(x) = a$, then $z := \pi_C(x)$ is a solution of (AVI). Furthermore, if z is a solution of (AVI), then $x := z + a - A(z)$ satisfies $A_C(x) = a$. We shall use this equivalence throughout this paper without further reference. This equivalence was originally introduced by Eaves in (1971).

A very familiar special case of (GE) is when $C \equiv K$ is a polyhedral convex cone. Then it is easy to show that (GE) is equivalent to the generalized complementarity problem (Karamardian 1976)

$$z \in K, A(z) - a \in K^D, \qquad \langle A(z) - a, z \rangle = 0,$$

where $K^D := \{z^* | \langle z^*, k \rangle \geq 0, \ \forall k \in K\}$ is the dual cone associated with K. The pivotal technique that we describe here can be thought of as a generalization of Lemke's complementary pivot algorithm (1965) for the special case $K \equiv \mathbb{R}^n_+$, the nonnegative orthant of $\mathbb{R}^n$.

In §2 we describe the theoretical algorithm and apply several results of Eaves and Robinson to establish its finite termination for coherently oriented normal maps. In §3 we carefully describe an implementation of such a method, under the assumption that C is given by

$$C := \{z | Bz \geq b, Hz = h\}.$$

In §4 we extend several well known results for linear complementarity problems to the affine variational inequality. In particular, we generalize the notions of copositive, copositive-plus and L-matrices from the complementarity literature and prove that our algorithm processes variational inequalities associated with such matrices. That is, when the algorithm is applied to such a problem, either a solution is found, or the problem is infeasible in a well-specified sense. Our definition of L-matrices is new and enables the treatment of both coherently oriented normal maps and copositive-plus matrices within the same framework. Furthermore, this result (Theorem 4.4) includes many of the standard existence results for complementarity problems and variational inequalities as special cases.

A word about our notation. For any vectors x and y in $\mathbb{R}^n$, $\langle x, y\rangle$ or $x^T y$ denotes the inner product of x and y, and in this paper, these two notations are freely interchangeable. Each $m \times n$ matrix A represents a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$, the symbol A refers to either the matrix or the linear map as determined by the context. Given a linear map A from $\mathbb{R}^n$ to $\mathbb{R}^m$, for any $X \subset \mathbb{R}^n$, the set $A(X) := \{y \in \mathbb{R}^m | y = Ax, \text{ for some } x \in X\}$ is called the image of X under A; for any set $Y \subset \mathbb{R}^m$, the set $A^{-1}(Y) := \{x \in \mathbb{R}^n | Ax \in Y\}$ is referred to as the inverse image of Y under A. In particular, the set $\ker A := A^{-1}(\{0\})$ is called the kernel of A and the set $\operatorname{im} A := A(\mathbb{R}^n)$ is called the image of A. Given a nonempty, closed, convex set C in $\mathbb{R}^n$, $\operatorname{rec} C := \{d \in \mathbb{R}^n | x + \lambda d \in C, \forall x \in C, \forall \lambda \geq 0\}$ is called the recession cone of C and $\operatorname{lin} C = \operatorname{rec} C \cap -\operatorname{rec} C$ is the lineality of C. If F is a function from $\mathbb{R}^n$ to $\mathbb{R}^n$, then F_C represents the normal map defined above. If C is a polyhedral convex convex set, a subset G is called a face of C if there exists a vector $c \in \mathbb{R}^n$ such that $G = \arg\max_{x \in C} c^T x$.

2. Theoretical algorithm. We describe briefly a theoretical algorithm that is guaranteed to find a solution in finitely many steps when the homeomorphism condition developed in Robinson (1992) holds. This method is a realization of the general path-following algorithm described and justified in Eaves (1976). In what follows we use various terms and concepts that are explained in Eaves (1976). A more detailed description of an implementation of the method is given in §3; here we deal with theoretical considerations underpinning the method. Other related work can be found in Burke and Moré (1994).

In order to formulate the algorithm, it is important to understand the underlying geometric structure of the problem. Our approach relies heavily on the normal manifold of the set C, (Robinson 1992), which we will now describe. Note that the normal cone to a convex set C at a point $x \in C$ is given by

$$N_C(x) = \{n | \langle n, c - x\rangle \leq 0, \forall c \in C\}.$$

It is well known (Burke and Moré 1994, and Robinson 1992) that the normal cone is constant on the relative interior of a face, that is $N_C(y) = N_C(x)$, whenever $x, y \in \operatorname{ri} C$. The normal manifold is generated by the faces and these normal cones as follows:

THEOREM 2.1. *Let C be a nonempty polyhedral convex set in $\mathbb{R}^n$ and let $\{F_i | i \in \mathcal{I}\}$ be the nonempty faces of C. For $i \in \mathcal{I}$, define N_{F_i} to be the common value of $N_C(\cdot)$ on $\operatorname{ri} F_i$ and let $\sigma_i := F_i + N_{F_i}$. The normal manifold $\mathcal{N}_C$ of C consists of the pair $(\mathbb{R}^n, \mathcal{S})$, where $\mathcal{S} := \{\sigma_i | i \in \mathcal{I}\}$, and $\mathbb{R}^n = \cup_{i \in \mathcal{I}} \sigma_i$. The faces of the σ_i having dimension $k \geq 0$ are called the k-cells of $\mathcal{N}_C$. $\mathcal{N}_C$ is a subdivided piecewise linear manifold of dimension n.*

It can be seen that the normal map A_C will agree in each n-cell of this manifold with an affine map, and therefore, with each such cell we can associate the determinant of the corresponding linear transformation. If each of these determinants has the same sign, we say that A_C is coherently oriented. For example, if A is the matrix representing the linear map A with respect to the standard coordinate system in $\mathbb{R}^n$ and $C = \mathbb{R}^n_+$, the nonnegative orthant in $\mathbb{R}^n$, then A_C is coherently oriented if and only if A is a P-matrix. The following is the central result from Robinson 1992.

THEOREM 2.2. *The normal map A_C is a Lipschitzian homeomorphism of $\mathbb{R}^n$ into $\mathbb{R}^n$ if and only if A_C is coherently oriented.*

We will assume first of all that A_C is a homeomorphism of $\mathbb{R}^n$ onto $\mathbb{R}^n$, so that the same-sign condition holds and describe the algorithm within this framework. Later in the paper, this condition will be weakened. The first step of the algorithm is to determine if C contains any lines. If it does, take orthonormal bases for lin C and its orthogonal complement according to the scheme explained in (Robinson 1992, Proposition 4.1). The factoring procedure explained there shows how to reduce the problem to one (which we shall also write $A_C(x) = a$) in a possibly smaller space, in which the set C appearing in this problem contains no lines. In that case, as shown in Robinson (1992), the determinants associated with A_C in the various cells of $\mathcal{N}_C$ must all have positive sign. Further, C will have an extreme point, say x_e, and as pointed out in Robinson (1992, §5) the normal cone $N_C(x_e)$ must have an interior. Let e be any element of int $N_C(x_e)$. An implementation of the factoring procedure is given as stage one of the method described in §3. The construction of an extreme point and element in the interior of the normal cone corresponds to stage two of that method.

Now construct a piecewise-linear manifold $\mathcal{M}$ from $\mathcal{N}_C$ by forming the Cartesian product of each cell of $\mathcal{N}_C$ with $\mathbb{R}_+$, the nonnegative half-line in $\mathbb{R}$. This $\mathcal{M}$ will be a PL$(n + 1)$-manifold in $\mathbb{R}^{n+1}$, as can easily be verified (see Eaves 1976, Example 4.3). Define a PL function F: $\mathcal{M} \to \mathbb{R}^n$ (where $\mathbb{R}^n$ is regarded as a PL manifold of one cell) by:

$$F(x, \mu) = A_C(x) - (\mu e + a).$$

We shall consider solutions $x(\mu)$ of $F(x, \mu) = 0$; it is clear from (NE) that $x(0)$ will solve our problem. Note that since we have assumed A_C to be a homeomorphism, the function $x(\cdot)$ is single-valued and defined on all of $\mathbb{R}_+$, though this property is not essential to our argument.

Now define $w(\mu) = x_e + (a - Ax_e) + \mu e$. It is clear that since

$$w(\mu) = x_e + \mu\left[e + \mu^{-1}(a - Ax_e)\right] \tag{1}$$

for large positive μ, $w(\mu)$ lies interior to the cell $x_e + N_C(x_e)$ of $\mathcal{N}_C$. Therefore $(w(\mu), \mu)$ lies interior to the cell $[x_e + N_C(x_e)] \times R_+$ of $\mathcal{M}$, and so it is a regular point of $\mathcal{M}$. (Given F, $\mathcal{M}$ and $\mathbb{R}^n$ as above, $(w(\mu), \mu)$ is a regular point if it is not contained in any cell σ with $\dim F(\sigma) < n$.) Further, for such μ we have $\pi_C(w(\mu)) = x_e$, so that

$$F(w(\mu), \mu) = Ax_e + (a - Ax_e) + \mu e - (\mu e + a) = 0,$$

and therefore for some $\mu_0 \geq 0$, $F^{-1}(0)$ contains the ray $\{(w(\mu), \mu) | \mu \geq \mu_0\}$.

Now we apply the algorithm of Eaves (1976) to the PL equation $F(x, \mu) = 0$, using a ray start at $(w(\mu_1), \mu_1)$ for some $\mu_1 > \mu_0$ and proceeding in the direction $(-e, -1)$. As the manifold $\mathcal{M}$ is finite, according to Eaves (1976, Theorem 15.13) the algorithm generates, in finitely many steps, either a point $(x_*, \mu_*) = 0$, or a ray in $F^{-1}(0)$ different from the starting ray. As the boundary of $\mathcal{M}$ is $\mathcal{N}_C \times \{0\}$, we see that in the first case $\mu_* = 0$ and, by our earlier remarks, x_* then satisfies $A_C(x_*) = a$. Therefore in order to justify the algorithm we need only show that it cannot produce a ray different from the starting ray.

The algorithm in question permits solving the perturbed system $F(x_\epsilon, \mu_\epsilon) = p(\epsilon)$, where $p(\epsilon)$ is of the form

$$p(\epsilon) = \sum_{i=1}^{n} p_i \epsilon^i$$

for appropriately chosen vectors p_i. It is shown in Eaves (1976) that $p(\epsilon)$ is a regular value of F for each small positive ϵ, and it then follows by Eaves (1976, Theorem 9.1) that for such ϵ, $F^{-1}(p(\epsilon))$ is a connected 1-manifold $Y(\epsilon)$, whose boundary is equal to its intersection with the boundary of $\mathscr{M}$, and which is subdivided by the chords formed by its intersections with the cells of $\mathscr{M}$ that it meets. Finally, for an easily computed function

$$b(\epsilon) = \sum_{i=1}^{n} b_i \epsilon^i,$$

we have $(w(\mu_1), \mu_1) + b(\epsilon) \in Y(\epsilon)$, and for small positive ϵ this point evidently lies on a ray in $F^{-1}(p(\epsilon))$. Because we start on this ray, $Y(\epsilon)$ cannot be homeomorphic to a circle, and therefore it is homeomorphic to an interval.

A simple computation at the starting point shows that the curve index Eaves (1976, §12) at that point is -1. By Eaves (1976, Lemma 12.1) this index will be constant along $Y(\epsilon)$. However, a computation similar to that in Eaves (1976, Lemma 12.3) shows that in each cell of $\mathscr{M}$, if the direction of $Y(\epsilon)$ in that cell is (r, ρ) then

$$(\operatorname{sgn} \rho)(\operatorname{sgn} \det T) = -1,$$

where T is the linear transformation associated with A_C in the corresponding cell of $\mathscr{N}_C$. Under our hypotheses, $\det T$ must be positive, and therefore ρ is negative everywhere along $Y(\epsilon)$. But this means that the parameter μ decreases strictly in each cell of linearity that $Y(\epsilon)$ enters, and it follows from the structure of $\mathscr{M}$ that after finitely many steps we must have $\mu = 0$, and therefore we have a point x_ϵ with $A_C(x_\epsilon) = a + p(\epsilon)$.

Now in practice the algorithm does not actually use a positive ϵ, but only maintains the information necessary to compute $Y(\epsilon)$ for all small positive ϵ, employing the lexicographic ordering to resolve possible ambiguities when $\epsilon = 0$. Therefore after finitely many steps it will actually have computed x_0 with $A_C(x_0) = a$.

Note that for linear complementarity problems, the above algorithm corresponds to Lemke's method (1965). It is well known that for linear complementarity problems associated with **P**-matrices, Lemke's method terminates at a solution. For variational inequalities, we have a similar result due to the analysis above.

THEOREM 2.3. *Given the problem* (NE), *assume that A_C is coherently oriented*; *then the path following method given in this section terminates at a solution of* (NE). *Furthermore, the parameter μ decreases monotonically to zero.*

3. Algorithm implementation. The previous section described a method for solving the Affine Variational Inequality over a general polyhedral set and showed (under a lexicographical ordering) that a coherently oriented normal equation (NE) can be solved in a finite number of iterations by a path-following method. In this section, we describe the numerical implementation of such a method, giving emphasis to the numerical linear algebra required to perform the steps of the algorithm.

We shall specialize to the case where C is given as

$$C := \{z \mid Bz \geq b, Hz = h\}, \tag{2}$$

and we shall assume that the linear transformation A is represented by the matrix A in our current coordinate system. We can describe our method to solve the normal equation in three stages. Note that by "solving," we mean producing a pair $(x, \pi(x))$, where x is a solution of (NE) and $\pi(x)$ is the projection of x onto the underlying set C.

In the first stage we remove lines from the set C, to form a reduced problem (over $\tilde{C}$) as outlined in the theory above. The lineality space of C as defined by (2) is

$$\operatorname{lin} C = \ker \begin{bmatrix} B \\ H \end{bmatrix}.$$

We calculate bases for the lineality space and its orthogonal complement by performing a QR factorization (with column pivoting) of $[B^T \quad H^T]$. If $[W \quad V]$ represents these bases, the reduced problem is to solve the normal equation

$$\tilde{A}_{\tilde{C}} y = \tilde{a}, \quad \text{where} \tag{3}$$

$$\tilde{C} = \{z \mid \tilde{B}z \geq b, \tilde{H}z = h\}, \qquad \tilde{B} = BV, \qquad \tilde{H} = HV. \tag{4}$$

Here

$$\tilde{A} = U^T A U, \tilde{a} = V^T(I - AZ)a, \quad \text{with} \tag{5}$$

$$Z = W(W^T A W)^{-1} W^T, \qquad U = (I - ZA)V, \tag{6}$$

and Z satisfies $Z^T A Z = Z^T$. In practice, $\tilde{A}$ and $\tilde{a}$ are calculated using one LU factorization of $W^T A W$. Furthermore, the solution pair $(x, \pi(x))$ of the original normal equation (NE) can be recovered from the solution pair $(y, \pi(y))$ of (3) using the identities

$$x_l = Z(a - AV\pi(y)),$$

$$x = x_l + Vy,$$

$$\pi(x) = x_l + V\pi(y).$$

Therefore, we can assume that the problem has the form (3), with $\tilde{C}$ given by (4) and that the matrix $\begin{bmatrix} \tilde{B} \\ \tilde{H} \end{bmatrix}$ has full column rank. We note that a similar construction is needed in Ralph (1992) and Robinson (1993).

In the second stage, we determine an extreme point of the set $\tilde{C}$, and using this information reduce the problem further by forcing the iterates to lie in the affine space generated by the equality constraints. More precisely, we have the following result:

LEMMA 3.1. *Suppose $y_e \in \tilde{C}$ and Y is a basis for the kernel of $\tilde{H}$. Then $\bar{y}$ solves* (3) *if and only if $\bar{y} = y_e + Y\bar{x}$ where $\bar{x}$ solves*

$$\bar{A}_{\bar{C}} x = \bar{a}. \tag{7}$$

Here $\bar{A} = Y^T\tilde{A}Y$, $\bar{a} = Y^T(\tilde{a} - \tilde{A}y_e)$ *and* $\bar{C} = \{z \mid \tilde{B}Yz \geq b - \tilde{B}y_e\}$. *Furthermore,* $\tilde{B}Y$ *has full column rank if and only if* $\begin{bmatrix} \tilde{B} \\ \tilde{H} \end{bmatrix}$ *has full column rank.*

Thus, to reduce our problem to one over an inequality constrained polyhedral set, it remains to show how we generate the point $y_e \in \tilde{C}$. In fact we show how to generate y_e as an extreme point of $\tilde{C}$ and further, how to project this extreme point into an extreme point of $\bar{C}$. The following result is a well known characterization of extreme points of polyhedral sets (Murty 1976, §3.4).

LEMMA 3.2. *Let u be partitioned into free and constrained variables* $(u_{\mathscr{F}}, u_{\mathscr{C}})$. *u is an extreme point of* $\mathscr{D} = \{u = (u_{\mathscr{F}}, u_{\mathscr{C}}) \mid Du = d, u_{\mathscr{C}} \geq 0\}$ *if and only if* $u \in \mathscr{D}$ *and* $\{d_i \mid i \in \mathscr{B}\}$ *are linearly independent, where* $\mathscr{B} := \mathscr{F} \cup \{j \in \mathscr{C} \mid u_j > 0\}$.

If we adopt the terminology of linear programming, then the variables corresponding to $\mathscr{B}$ are called basic variables; similarly, the columns of D corresponding to $\mathscr{B}$ are called basic columns; extreme points are called basic feasible solutions.

The extreme points of systems of inequalities and equalities are defined in an analogous manner. Note that extreme points of $\tilde{C}$ are (by definition) precisely the extreme points of

$$\begin{bmatrix} \tilde{B} & -I \\ \tilde{H} & 0 \end{bmatrix} \begin{bmatrix} z \\ s \end{bmatrix} = \begin{bmatrix} b \\ h \end{bmatrix}, \qquad s \geq 0. \tag{8}$$

The slack variables s are implicitly defined by z, so without ambiguity we will refer to the above extreme point as z. For other systems of inequalities and equations a similar convention will be used. The following lemma outlines our method for constructing the relevant extreme points.

LEMMA 3.3. *Suppose* $\begin{bmatrix} \tilde{B} \\ \tilde{H} \end{bmatrix}$ *has linearly independent columns, Y is a basis of the kernel of* $\tilde{H}$ *and* $\bar{B} = \tilde{B}Y$. *Then* y_e *is an extreme point of* (8) *if and only if* $y_e = y_* + Yz_*$, *for some* y_*, z_* *where* $\tilde{H}y_* = h$ *and* z_* *is an extreme point of*

$$\begin{bmatrix} \bar{B} & -I \end{bmatrix} \begin{bmatrix} z \\ s \end{bmatrix} = b - \tilde{B}y_*, \qquad s \geq 0. \tag{9}$$

In our method we produce an extreme point of (8) as follows. Find orthonormal bases U and Y for im $\tilde{H}$ and ker $\tilde{H}$ respectively. This can be carried out by a singular value decomposition of $\tilde{H}$ or by QR factorizations of $\tilde{H}$ and $\tilde{H}^T$ (in fact, Y could be calculated as a by-product of stage 1 of the algorithm). In particular, if

$$\tilde{H}^T = \begin{bmatrix} Z & Y \end{bmatrix} \begin{bmatrix} R \\ 0 \end{bmatrix},$$

then Y is the orthonormal basis of ker $\tilde{H}$ and we can let $y_* = ZR^{-T}h$, using this value of y_* in (9). If $b \notin \text{im } \bar{B}$, then find an extreme point of (9) by solving the following auxiliary problem with the revised simplex method:

$$\begin{aligned} &\text{minimize} \quad z_{\text{aux}} \\ &\text{subject to} \quad \begin{bmatrix} B & b - \tilde{B}y_* \end{bmatrix} \begin{bmatrix} z \\ z_{\text{aux}} \end{bmatrix} \geq b - \tilde{B}y_*, \qquad z_{\text{aux}} \geq 0. \end{aligned}$$

Note that $z = 0, z_{\text{aux}} = 1$ is an initial feasible point for this problem, with basic variables (z, z_{aux}). In contrast to the usual square basis matrix (with corresponding LU factors), we use a QR factorization of the nonsquare basis matrix. The calculations of dual variables and incoming columns are performed in a least squares sense using the currently available QR factorization. This factorization is updated at each pivot step either by using a rank-one update to the factorization or by adding a column to the factorization (see Golub and Van Loan 1983). In order to invoke Lemma 3.1, we let $y_e = y_* + Yz_*$ be the feasible point needed to define (7).

Note that in the well known method of Lemke, stages one and two are trivial since $C = \mathbb{R}^n_+$ has no lines and a single extreme point at 0. Furthermore, stage one is an exact implementation of the theory outlined in the previous section and stage two corresponds to determining an extreme point and treating the defining equalities of C in an effective computational manner.

It remains to describe stage three of our method. We are able to assume that our problem is given as

$$\bar{A}_{\bar{C}} x = \bar{a}, \tag{10}$$

with $\bar{C} = \{z | \bar{B}z \geq \bar{b}\}$, where $\bar{B}$ has full column rank and x_e is an extreme point of $\bar{C}$ (easily determined from z_*). We also have available a basis matrix corresponding to this extreme point along with a QR factorization, courtesy of stage two.

The method that we use to solve this problem is precisely a realization of the general scheme for piecewise linear equations developed by Eaves (1976). The general method of Eaves (assuming a ray start and regular value v) moves along the curve $F^{-1}(v)$ in the direction d_1 from x_1. Note that a direction $d \neq 0$ points into σ at x if $x \in \sigma$ and $x + \theta d \in \sigma$ for all sufficiently small θ. The complete algorithm is given as Algorithm 1.

Algorithm 1

Initialize:. Let L_{σ_k} denote the linear map representing F on the cell σ_k. Determine (x_1, σ_1, d_1) satisfying

$$L_{\sigma_1} d_1 = 0, d_1 \text{ points into } \sigma_1 \text{ at } x_1, \tag{11}$$

$$F(x_1) = v,$$

$$x_1 \in \sigma_1 \in \mathscr{M}, x_1 \in \operatorname{int}\{x - \theta d_1 | \theta \geq 0\} \subset F^{-1}v. \tag{12}$$

Iteration:. Given (x_k, σ_k, d_k) let

$$\theta_k := \sup\{\theta | x_k + \theta d_k \in \sigma_k\}. \tag{13}$$

If $\theta_k = +\infty$ then ray termination.
If $x_{k+1} := x_k + \theta_k d_k \in \partial\mathscr{M}$ then boundary termination.
Otherwise determine $(x_{k+1}, \sigma_{k+1}, d_{k+1}), d_{k+1} \neq 0$, satisfying

$$L_{\sigma_{k+1}} d_{k+1} = 0, \text{ and } d_{k+1} \text{ points into } \sigma_{k+1} \text{ from } x_{k+1}, \tag{14}$$

$$\sigma_{k+1} \in \mathscr{M} \setminus \{\sigma_k\} \text{ with } x_{k+1} \in \sigma_{k+1}.$$

Set $k = k + 1$ and repeat iteration.

How does this relate to the description we gave in the previous section? The manifold we consider is $\mathscr{M} = \mathscr{N}_{\bar{C}} \times \mathbb{R}_+$, and the corresponding cells $\sigma_{\mathscr{A}}$ are given by $(F_{\mathscr{A}} + N_{F_{\mathscr{A}}}) \times \mathbb{R}_+$, where $F_{\mathscr{A}}$ are the faces of $\bar{C}$.

A face of $\overline{C}$ is described by the set of constraints from the system $\overline{B}z \geq \overline{b}$ which are active. Let $\mathscr{A}$ represent such a set so that

$$F_{\mathscr{A}} = \left\{z | \overline{B}_{\mathscr{A}} z = \overline{b}_{\mathscr{A}}, \overline{B}_{\mathscr{I}} z \geq \overline{b}_{\mathscr{I}}\right\},$$

where $\mathscr{I}$ is the complement of the set $\mathscr{A}$. The normal cone to the face (the normal cone to $\overline{C}$ at some point in the relative interior of $F_{\mathscr{A}}$) is given by

$$\left\{\overline{B}^T u | u_{\mathscr{A}} \leq 0, u_{\mathscr{I}} = 0\right\}.$$

It now follows that an algebraic description of $(x, \mu) \in \sigma_{\mathscr{A}}$ is that there exist $(x, z, u_{\mathscr{A}}, s_{\mathscr{I}}, \mu)$ which satisfy

$$\begin{aligned} \overline{B}_{\mathscr{A}} z &= \overline{b}_{\mathscr{A}}, \\ \overline{B}_{\mathscr{I}} z - s_{\mathscr{I}} &= \overline{b}_{\mathscr{I}}, \; s_{\mathscr{I}} \geq 0, \\ x &= z + \overline{B}_{\mathscr{A}}^T u_{\mathscr{A}}, \; u_{\mathscr{A}} \leq 0, \\ \mu &\geq 0. \end{aligned} \tag{15}$$

In particular, if x_e is the given extreme point, the corresponding face of the set $\overline{C}$ is used to define the initial cell σ_1. The piecewise linear system we solve is

$$F(x, \mu) := \overline{A}_{\overline{C}}(x) - (\mu e + \overline{a}) = 0,$$

where e is a point in the interior of $N_{\overline{C}}(x_e)$. An equivalent description of $N_{\overline{C}}(x_e)$ is given by $\{\overline{B}_{\mathscr{A}}^T u | u \leq 0\}$, from which it is clear that the interior of this set is nonempty if and only if $\overline{B}_{\mathscr{A}}$ has full column rank.

LEMMA 3.4. *If x_e is an extreme point of $\{z | \overline{B}z \geq \overline{b}\}$ with active constraints $\mathscr{A}$, then $\overline{B}_{\mathscr{A}}$ has full column rank.*

PROOF. By definition,

$$G := \begin{bmatrix} \overline{B}_{\mathscr{A}} & 0 \\ \overline{B}_{\mathscr{I}} & -I \end{bmatrix}$$

has linearly independent columns. If $\overline{B}_{\mathscr{A}}$ does not have linearly independent columns, then $\overline{B}_{\mathscr{A}} w = 0$, for some $w \neq 0$, so that

$$G \begin{bmatrix} w \\ \overline{B}_{\mathscr{I}} w \end{bmatrix} = 0,$$

with $(w, \overline{B}_{\mathscr{I}} w) \neq 0$, a contradiction to the linear independence of the columns of G. □

This is a simple proof (in this particular instance) of the comment from the previous section that the normal cone has interior at an extreme point. For consistency, we shall let e be any point in this interior $\{\overline{B}_{\mathscr{A}}^T u | u \leq 0\}$, and for concreteness we could take

$$e = -\overline{B}_{\mathscr{A}}^T \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Hence F is specified, $v = 0$ and the cells of $\sigma_{\mathscr{A}}$ are defined. By solving the perturbed system $F(x_\epsilon, \mu_\epsilon) = p(\epsilon)$ (as outlined in §2), we know that $F^{-1}(p(\epsilon))$ is a connected 1-manifold whose boundary is equal to its intersection with the boundary of $\mathscr{M}$ and which is subdivided by the chords formed by its intersections with the cells of $\mathscr{M}$ that it meets. In practice, this means that (under the lexicographical ordering induced by $p(\epsilon)$) we may assume nondegeneracy. Thus, if ties ever occur in the description that follows, we will always choose the lexicographical minimum from those which achieve the tie.

Note that if $(x, \mu) \in \sigma_{\mathscr{A}}$ as defined in (15) then

$$F(x, \mu) = \bar{A}z + x - z - \mu e - \bar{a}.$$

It follows that if $(x, \mu) \in \sigma_{\mathscr{A}} \cap F^{-1}(0)$ (i.e. (x, μ) is in one of the chords mentioned in the previous paragraph), then there exist $(x, z, u_{\mathscr{A}}, s_{\mathscr{I}}, \mu)$ satisfying

$$\begin{aligned} x - z &= -\bar{A}z + \mu e + \bar{a}, \\ \bar{B}_{\mathscr{A}}z &= \bar{b}_{\mathscr{A}}, \\ \bar{B}_{\mathscr{I}}z - s_{\mathscr{I}} &= \bar{b}_{\mathscr{I}},\ s_{\mathscr{I}} \geq 0, \\ x - z &= \bar{B}_{\mathscr{A}}^T u_{\mathscr{A}},\ u_{\mathscr{A}} \leq 0, \\ \mu &\geq 0. \end{aligned} \tag{16}$$

Furthermore, these equations determine the chord on the current cell of the manifold, or in the notation used to describe the algorithm of Eaves, the map $L_{\sigma_{\mathscr{A}}}$. The direction is determined from (11) by solving $L_{\sigma_{\mathscr{A}}} d = 0$, which can be calculated by solving

$$\begin{aligned} \Delta x - \Delta z &= -\bar{A}\Delta z + e\Delta\mu, \\ \bar{B}_{\mathscr{A}}\Delta z &= 0, \\ \bar{B}_{\mathscr{I}}\Delta z - \Delta s_{\mathscr{I}} &= 0, \\ \Delta x - \Delta z &= \bar{B}_{\mathscr{A}}^T \Delta u_{\mathscr{A}}. \end{aligned} \tag{17}$$

At the first iteration, $\bar{B}_{\mathscr{A}}$ has full column rank, so that $\Delta z = 0$, which also implies that $\Delta s_{\mathscr{I}} = 0$. The remaining system of equations is

$$\begin{aligned} \Delta x &= e\Delta\mu, \\ \Delta x &= \bar{B}_{\mathscr{A}}^T \Delta u_{\mathscr{A}}. \end{aligned}$$

We choose $\Delta\mu = -1$ in order to force the direction to move into σ_1 (as required by (11)), and then it follows that $\Delta x = -e$ for the choice of e outlined above $\Delta u_{\mathscr{A}} = (1, \dots, 1)^T$. The actual choice $x_1 = (w(\mu), \mu)$ given in the previous section ensures that (12) is satisfied.

We can now describe the general iteration and the resultant linear algebra that it entails. We are given a current point $(x, z, u_{\mathscr{A}}, s_{\mathscr{I}}, \mu)$ satisfying (16) for some cell $\sigma_{\mathscr{A}}$ and a direction $(\Delta x, \Delta z, \Delta u_{\mathscr{A}}, \Delta s_{\mathscr{I}}, \Delta\mu)$ satisfying (17). The value of θ_k to satisfy (13) can be calculated by the following ratio test; that is to find the largest θ such that

$$\text{(18)} \qquad \begin{aligned} u_{\mathscr{A}} + \theta \Delta u_{\mathscr{A}} &\leq 0, \\ s_{\mathscr{I}} + \theta \Delta s_{\mathscr{I}} &\geq 0, \\ \mu + \theta \Delta \mu &\geq 0. \end{aligned}$$

Ray termination occurs if $\Delta u_{\mathscr{A}} \leq 0$, $\Delta s_{\mathscr{I}} \geq 0$ and $\Delta \mu \geq 0$. Obviously, if $\mu + \theta \Delta \mu = 0$, then we have a solution. Otherwise, at least one of the $\{u_i | i \in \mathscr{A}\}$ or $\{s_i | i \in \mathscr{I}\}$ hits a bound in (18). By the lexicographical ordering we can determine the "leaving" variable from these uniquely. The set $\mathscr{A}$ is updated (corresponding to moving onto a new cell of the manifold) and a new direction is calculated as follows: if u_i, $i \in \mathscr{A}$ is the leaving variable, then $\mathscr{A} := \mathscr{A} \setminus \{i\}$, $\Delta s_i = 1$ and the new direction is found by solving (17); if s_i, $i \in \mathscr{I}$ is the leaving variable, then $\mathscr{A} := \mathscr{A} \cup \{i\}$, $\Delta u_i = -1$ and the new direction is found by solving (17). Note that in both cases, the choice of one component of the direction ensures movement into the new (uniquely specified) cell $\sigma_{\mathscr{A}}$ and forces a unique solution of (17).

The linear algebra needed for an implementation of the method is now clear. The actual steps used to carry out stage 3 are now described. First of all, x is eliminated from (16) to give

$$\begin{aligned} -\overline{A}z + \mu e + \bar{a} &= \overline{B}_{\mathscr{A}}^T u_{\mathscr{A}} + \overline{B}_{\mathscr{I}}^T u_{\mathscr{I}}, \\ \overline{B}_{\mathscr{A}} z - s_{\mathscr{A}} &= \bar{b}_{\mathscr{A}}, \\ \overline{B}_{\mathscr{I}} z - s_{\mathscr{I}} &= \bar{b}_{\mathscr{I}}, \\ \mu \geq 0, u_{\mathscr{A}} \leq 0, u_{\mathscr{I}} = 0, &\qquad s_{\mathscr{I}} \geq 0, s_{\mathscr{A}} = 0. \end{aligned}$$

Note that we have added in the variables which are set to zero for completeness. The QR factorization corresponding to the given extreme point is used to eliminate the variables z. In fact, we take as our initial active set $\mathscr{A}$, the variables corresponding to $Q\hat{R}$, where $\hat{R}$ is the invertible submatrix of R. Thus $z = \overline{B}_{\mathscr{A}}^{-1}(s_{\mathscr{A}} + \bar{b}_{\mathscr{A}})$, and substituting this into the above gives

$$\begin{aligned} -\overline{A}\overline{B}_{\mathscr{A}}^{-1}\left(s_{\mathscr{A}} + \bar{b}_{\mathscr{A}}\right) + \mu e + \bar{a} &= \overline{B}_{\mathscr{A}}^T u_{\mathscr{A}} + \overline{B}_{\mathscr{I}}^T u_{\mathscr{I}}, \\ \overline{B}_{\mathscr{I}} \overline{B}_{\mathscr{A}}^{-1}\left(s_{\mathscr{A}} + \bar{b}_{\mathscr{A}}\right) - s_{\mathscr{I}} &= \bar{b}_{\mathscr{I}}, \\ \mu \geq 0, u_{\mathscr{A}} \leq 0, u_{\mathscr{I}} = 0, s_{\mathscr{I}} \geq 0, &\qquad s_{\mathscr{A}} = 0. \end{aligned}$$

Essentially we treat this system as in the method of Lemke. An initial basis is given by $(u_{\mathscr{A}}, s_{\mathscr{I}})$ and complementary pivots can then be executed (using the variables u and s as the complementary pair). Any basis updating technique or anti-cycling rule can be incorporated from the literature on linear programming and complementarity. In fact we have an initial QR factorization of the basis available from the given factorization if needed.

We showed in the previous section that if A_C was coherently oriented then following the above path gives a monotonic decrease in μ. However, the proof of the finite termination of the method (possibly ray termination) goes through without this assumption, and in the following section we will look at other conditions which guarantee that the method terminates either with a solution or a proof that no solution exists. The coherent orientation results are direct analogues of the P-matrix results for the linear complementarity problem—the results we shall give now generalize the notions of coposition-plus and L-matrices.

4. Existence results. The following definitions are generalizations of those found in the literature.

DEFINITION 4.1. Let K be a given closed convex cone. A matrix A is said to be copositive with respect to the cone K if

$$\langle x, Ax \rangle \geq 0, \quad \forall x \in K.$$

A matrix A is said to be copositive-plus with respect to the cone K if it is copositive with respect to K and

$$\langle x, Ax \rangle = 0, \qquad x \in K \quad \Rightarrow \quad (A + A^T)x = 0.$$

DEFINITION 4.2. Let K be a given closed convex cone. A matrix A is said to be L-matrix with respect to K if both

(a) For every $q \in \text{ri}(K^D)$, the solution set of the generalized complementarity problem

$$z \in K, \qquad Az + q \in K^D, \qquad z^T(Az + q) = 0 \tag{19}$$

is contained in lin K.

(b) For any $z \neq 0$ such that

$$z \in K, \qquad Az \in K^D, \qquad z^T Az = 0,$$

there exists $z' \neq 0$, such that z' is contained in every face of K containing z and $-A^T z'$ is contained in every face of K^D containing Az.

To see how these definitions relate to the standard ones given in the literature on linear complementarity problems (e.g. Murty 1988 and Cottle, Pang and Stone 1992), consider the case that $C = \mathbb{R}^n_+$ and $K = \text{rec}\, C = \mathbb{R}^n_+$. Condition a) says that LCP(q, A) has a unique solution 0 for all $q > 0$. Condition b) states that, if $z \neq 0$ is a solution of LCP$(0, A)$, then there exists $z \neq 0$ such that z' is contained in every face of $\mathbb{R}^n_+$ containing z and $-A^T z'$ is contained in every face of $\mathbb{R}^n_+$ containing Az. In particular, $z' \in \{x \in \mathbb{R}^n | x_i = 0\}$, for all $i \in \{i | z_i = 0\}$. Hence $z'_i = 0$ for each i such that $z_i = 0$. That is, supp $z' \subset$ supp z. In another words, there exists a diagonal matrix $D \geq 0$ such that $z' = Dz$. Similarly, there exists a diagonal matrix $E \geq 0$ such that $-A^T z' = EAz$. Hence $(EA + A^T D)z = 0$. where, $D, E \geq 0$ and $Dz \neq 0$. Thus the notion of L-matrix defined here is a natural extension of that presented in Murty (1988). The following lemma shows that the class L-matrices the class of copositive-plus matrices.

LEMMA 4.3. *If a matrix A is copositive-plus with respect to a closed convex cone K, then it is an L-matrix with respect to K.*

PROOF. Suppose that $q \in \text{ri}(K^D)$ and $z \in K \setminus \text{lin}\, K$, then $\pi_{(\text{lin}\, K)^\perp}(z) \neq 0$. Furthermore, there exists an $\epsilon > 0$, such that $q - \epsilon \pi_{(\text{lin}\, K)^\perp}(z) \in K^D$, since aff$(K^D) = (\text{lin}\, K)^\perp$ (cf. Rockafellar 1970, Theorem 14.6). It follows that

$$0 \leq \left\langle z, q - \epsilon \pi_{(\text{lin}\, K)^\perp}(z) \right\rangle = \langle z, q \rangle - \epsilon \left\langle z, \pi_{(\text{lin}\, K)^\perp}(z) \right\rangle = \langle z, q \rangle - \epsilon \|\pi_{(\text{lin}\, K)^\perp}(z)\|_2^2.$$

That is $\langle z, q\rangle \geq \epsilon\|\pi_{(\text{lin } K)^\perp}(z)\|_2^2 > 0$. Also $z^TAz \geq 0$ since A is copositive with respect to K. Thus $z^T(Az + q) = z^TAz + z^Tq \geq z^Tq > 0$. This shows that the set $K \setminus \text{lin } K$ does not contain any solution of (19). Therefore the solution set of the problem (19) is contained in lin K.

To complete the proof, note that for any $z \in K$, such that $Az \in K^D$ and $z^TAz = 0$, we have $Az + A^Tz = 0$, or $-A^Tz = Az$, since A is copositive-plus. So the condition b) of Definition 4.2 is satisfied with $z' = z$. □

We now come to the main result of this section.

THEOREM 4.4. *Suppose C is a polyhedral convex set and A is an L-matrix with respect to* rec C *which is invertible on the lineality space of C. Then exactly one of the following occurs*:

- *The method given above solves (AVI)*;
- *the following system has no solution*:

$$Ax - a \in (\text{rec } C)^D, \qquad x \in C. \tag{20}$$

PROOF. Suppose that $C = \{z | Bz \geq b, Hz = h\}$. We may assume that (AVI) is in the form (10) due to Lemma A.4 and Lemma A.5 and our assumption regarding the lineality space of C. The pivotal method fails to solve (AVI) only if, at some iterate x_k, it reaches an unbounded direction d_{k+1} in σ_{k+1}. We know that x_k satisfies (16), and the direction d_{k+1} which satisfies $L_{\sigma_{k+1}} d_{k+1} = 0$ can be found by solving (17). Suppose $(\Delta x, \Delta z, \delta u_{\mathscr{A}}, \Delta s_{\mathscr{I}}, \Delta \mu)$ is a solution of (17), then

$$\Delta u_{\mathscr{A}} \leq 0, \qquad \Delta s_{\mathscr{I}} \geq 0, \qquad \Delta \mu \geq 0, \tag{21}$$

provided that $x_k + \theta d_{k+1}$ is an unbounded ray. By reference to (17), we have

$$\begin{aligned} \bar{B}_{\mathscr{A}}^T \Delta u_{\mathscr{A}} + \bar{A}\Delta z &= e\Delta\mu, \\ \bar{B}_{\mathscr{A}}\Delta z &= 0, \\ \bar{B}_{\mathscr{I}}\Delta z &= \Delta s_{\mathscr{I}} \geq 0. \end{aligned} \tag{22}$$

That is, Δz satisfies

$$\Delta z \in \text{rec } \bar{C},$$

$$\bar{A}\Delta z - e\Delta\mu = \bar{B}_{\mathscr{A}}^T(-\Delta u_{\mathscr{A}}) \in (\text{rec } \bar{C})^D,$$

$$\Delta z^T(\bar{A}\Delta z - e\Delta\mu) = \Delta z^T \bar{B}_{\mathscr{A}}^T(-\Delta u_{\mathscr{A}}) = -(\bar{B}_{\mathscr{A}}\Delta z)^T \Delta u_{\mathscr{A}} = 0.$$

If $\Delta\mu > 0$, then $e\Delta\mu \in \text{int } N_{\bar{C}}(x_e)$, hence $-e\Delta\mu \in \text{int}(\text{rec } \bar{C})^D$. The above system has a unique solution $\Delta z = 0$ by the fact that $\bar{A}$ is an L-matrix with respect to rec $\bar{C}$ and lin $\bar{C} = \{0\}$. Therefore the terminating ray is the starting ray, a contradiction. Thus $\Delta\mu = 0$. It follows that $\Delta z \in \text{rec } \bar{C}$, $\bar{A}\Delta z \in (\text{rec } \bar{C})^D$, and $\Delta z^T \bar{A}\Delta z = 0$. Therefore there exist $\tilde{z} \neq 0$, such that $\tilde{z}$ is contained in every face of rec $\bar{C}$ containing Δz, and that $-\bar{A}^T\tilde{z}$ is contained in every face of $(\text{rec } \bar{C})^D$ containing $\bar{A}\Delta z$. We observe that, since $x_k \in \sigma_k \cap \sigma_{k+1} \cap F^{-1}(0)$, there exist z_k, u_k, s_k, and μ_k such that (16) is satisfied. It is easy to verify that Δz is in the face

$$G_1 = \left\{z \in \text{rec } \bar{C} | z^T(\bar{B}^T u_k) = 0\right\}$$

of rec $\overline{C}$, and $\overline{A}\Delta z$ is in the face

$$G_2 = \left\{z \in (\operatorname{rec} \overline{C})^D | z = \overline{B}^T u, u = (u_{\mathscr{A}}, 0) \geq 0\right\}$$

of $(\operatorname{rec} \overline{C})^D$, and thus

$$-\overline{A}^T \tilde{z} = \overline{B}^T \tilde{u} \in G_2, \quad \text{for some } \tilde{u} = (\tilde{u}_{\mathscr{A}}, 0) \geq 0. \tag{23}$$

Consequently, by (16) we have

$$\bar{a} = x_k - z_k + \overline{A} z_k - e\mu,$$

$$\tilde{u}^T\left(\overline{B} z_k - \bar{b}\right) = (\tilde{u}_{\mathscr{A}}^T, 0)\begin{pmatrix} 0 \\ s_{\mathscr{I}} \end{pmatrix} = 0, \quad \text{and}$$

$$\tilde{z}^T(x_k - z_k) = \tilde{z}^T \overline{B}^T u_k = 0,$$

since $\tilde{z} \in G_1$. Therefore

$$\begin{aligned}
\tilde{u}^T \bar{b} + \tilde{z}^T \bar{a} &= \tilde{u}^T\left(\bar{b} - \overline{B} z_k\right) + \tilde{u}^T \overline{B} z_k + \tilde{z}^T\left(x_k - z_k + \overline{A} z_k - e\mu\right) \\
&= \left(\overline{B}^T \tilde{u} + \overline{A}^T \tilde{z}\right)^T z_k - \mu e^T \tilde{z} \\
&= -\mu e^T \tilde{z} > 0,
\end{aligned}$$

in which the last inequality is due to $\tilde{z} \in \operatorname{rec} \overline{C}$ and $e \in \operatorname{int} N_{\overline{C}}(x_e) \subset -\operatorname{int}(\operatorname{rec} \overline{C})^D$. We now claim that the system

$$\overline{A}x - \bar{a} \in (\operatorname{rec} \overline{C})^D, \qquad x \in \overline{C} \tag{24}$$

has no solution. To see this, let $x \in \overline{C}$, then $\tilde{u}^T \overline{B} x + \tilde{z}^T \overline{A} x = 0$, as a result of (23). Subtract from this the inequality $\tilde{u}^T \bar{b} + \tilde{z}^T \bar{a} > 0$ which we have just proven. Then $\tilde{u}^T(\overline{B}x - \bar{b}) + \tilde{z}^T(\overline{A}x - \bar{a}) < 0$. It is obvious that $\tilde{u}^T(\overline{B}x - \bar{b}) \geq 0$, hence $\tilde{z}^T(\overline{A}x - \bar{a}) < 0$. But $\tilde{z} \in \operatorname{rec} \overline{C}$. Thus $\overline{A}x - \bar{a} \notin (\operatorname{rec} \overline{C})^D$.

The proof is complete by noting that (24) has a solution if and only if (20) has a solution. □

As a special case of this theorem, we have the following result for copositive-plus matrices.

COROLLARY 4.5. *Suppose C is a polyhedral convex set, A is copositive-plus with respect to* rec C *and invertible on the lineality space of C. Then exactly one of the following occurs*:

- *The method given above solves (AVI)*;
- *the following system has no solution*:

$$Ax - a \in (\operatorname{rec} C)^D, \qquad x \in C. \tag{25}$$

PROOF. Obvious, in view of Lemma 4.3. □

We can also prove Theorem 2.3 (without the proof of monotonic decrease in μ) as a special case of Theorem 4.4 by noting the following lemma.

LEMMA 4.6. *Suppose A_C is coherently oriented. Then*
(a) $A_{\mathrm{rec}\, C}$ *is coherently oriented;*
(b) A *is an L-matrix with respect to* rec C.

PROOF. a) This follows from the proof of Theorem 4.3 of Robinson (1992).
b) By the first part, $A_{\mathrm{rec}\, C}$ is coherently oriented, so by Robinson (1992, Theorem 4.3) it is a Lipschitzian homeomorphism, and hence $A_{\mathrm{rec}\, C}(x) = q$ has a unique solution for all q. Therefore, parts (i) and (ii) of the definition of L-matrices are trivially satisfied by the unique solution $\{0\}$. □

Note also that if C is compact, then any matrix A is an L-matrix with respect to rec C. Thus Theorem 4.4 also recovers the standard existence theory for variational inequalities over compact sets.

5. Computational results. The algorithm described in this paper has been implemented in MATLAB (The Math Works 1992). Copies of the code and the testing script files are available from the second author.

The algorithm NEPOLY is implemented as 3 function files in MATLAB. The development of the code is exactly as outlined in §3. The first function removes the lineality of the set C, then calls the second routine which proceeds to determine an extreme point and factor out the equality constraints. Having accomplished this, the third routine then executes the pivot steps. We note in particular, that Lemke's original pivot algorithm can be carried out just using the third routine, since the defining set $C = \mathbb{R}^n_+$ has no lines, no equality constraints and a single extreme point 0.

We now present two tables of our results of applying this algorithm to some small quadratic programs. In Table 1 we present a comparison of NEPOLY to the standard QP solver that is available as part of the optimization toolbox of MATLAB. This QP solver is an active set method, similar to that described in Gill, Murray and Wright (1981). Further details available in The Math Works (1992).

The problems that we generate are of the form

$$\text{(26)} \qquad \begin{aligned} &\text{minimize} && \tfrac{1}{2}x^TQx + c^Tx + \tfrac{1}{2}y^Ty \\ &\text{subject to} && Ax + By = b, \qquad x \geq 0, \end{aligned}$$

TABLE 1
NEPOLY and MATLAB QP

m	n	p	NEPOLY time	MATLAB QP time
10	10	10	0.3	0.8
20	10	10	0.2	0.2
30	20	10	0.3	0.3
10	40	10	3.4	10.5
10	10	50	0.6	5.7
20	20	30	1.2	4.6
10	60	20	5.8	45.1
70	10	30	0.8	0.9
40	40	40	4.6	14.3
100	10	10	0.5	0.6
10	10	100	3.1	9.8
10	100	10	28.0	121.1
50	30	40	7.9	6.8
40	100	60	32.4	208.5
80	40	100	10.2	37.4
60	60	100	13.3	114.5

where $Q \in \mathbb{R}^{n \times n}$, $A \in \mathbb{R}^{p \times n}$ and $B \in \mathbb{R}^{p \times m}$. The minimum principle generates an affine variational inequality which under convexity is equivalent to (26). In general, the variational inequality represents necessary optimality conditions for (26).

We generate Q as a random sparse symmetric matrix. Unfortunately, the MATLAB QP solver did not solve (26) unless Q was positive semidefinite, so in Table 1, Q was generated positive semidefinite. All other matrices were generated using the MATLAB random generator, although the feasible region was guaranteed to be nonempty.

MATLAB 4.0 was used with dedicated access to a Hewlett Packard 9000/705 workstation. The times reported are elapsed times in seconds using the built in stopwatch timer of MATLAB. The ordering of entries in the table is by total problem size. Since the problems are convex, both codes always found the solution of (26). The constraint error was always less than 10^{-14}. All MATLAB codes reported here do not use the sparse matrix facility of MATLAB.

Notice that NEPOLY solves all but one of these instances quicker than the MATLAB code. On the bigger problems, NEPOLY is much quicker that QP. These results are averaged over 10 randomly generated problems of the given size. The times vary slightly for different random problems of the same dimension, but the main conclusion is that NEPOLY outperforms MATLAB QP.

In Table 2, we present similar results comparing NEPOLY with a standard Lemke code. As outlined above, NEPOLY is easily adapted to generate the Lemke path as a special case. In order to carry out this comparison, we reformulate (26) as the following quadratic program:

$$\begin{aligned} \text{minimize} \quad & \tfrac{1}{2}x^TQx + c^Tx + \tfrac{1}{2}(z - e\xi)^T(z - e\xi) \\ & Ax + B(z - e\xi) \geq b, \\ \text{subject to} \quad & e^T(Ax + B(z - e\xi)) \leq e^Tb, \\ & x, z, \xi \geq 0. \end{aligned}$$

The necessary optimality conditions for this problem give rise to a standard form LCP to which Lemke's method can then be applied. Table 2 reports the iteration count

TABLE 2
NEPOLY and MATLAB QP

			NEPOLY		Lemke	
m	n	p	iter	time	iter	time
10	10	10	8	0.3	46	2.6
10	10	10	9	0.3	69	3.5
20	10	5	0	0.1	64	4.0
10	14	24	9	0.5	75	7.7
13	26	10	37	2.4	80	11.3
13	26	10	29	2.3	114	16.5
13	26	10	18	2.1	F	
20	40	20	32	4.6	126	46.5
20	40	20	23	2.8	173	62.6
10	50	30	F		F	
30	30	30	20	2.1	168	60.8
50	30	40	10	8.1	196	109.6
10	50	70	F		F	
40	70	50	40	13.5	298	471.2
40	100	60	55	33.8	323	1199.5
80	40	100	29	21.9	349	860.7

and elapsed time for problems of various sizes. In all cases, the problems were solved to high accuracy (constraint errors less than 10^{-14}).

Notice on some of the problems, one or other of the codes failed (denoted by F in the table). This is because for these experiments, Q was generated sparse and symmetric but not positive definite. The convergence theory does not guarantee finding a solution in these case, but note that the number of failures are small for NEPOLY. The number of failures can be made large by testing problems with large n since the failures are entirely due to the indefiniteness of Q. However, it is easy to infer that NEPOLY is significantly quicker than the standard Lemke code.

6. Conclusions. We have presented a method for solving affine variational inequalities and demonstrated its implementability. Further, the algorithm has been used to generate a new class of matrices (L-matrices) for which the corresponding affine variational inequality is solvable, or provably infeasible. Our theory is shown to unify several existing results in the literature. An implementation shows the method performs well in comparison with a standard active set method and Lemke's complementary pivot algorithm when applied to quadratic programs. Further testing is needed to ascertain whether the technique is effective for large scale problems.

Appendix A. Invariance properties of L-matrices. In this appendix we show that the property of L-matrix with respect to a polyhedral convex cone is invariant under the two reductions presented in §3. We begin with the following technical lemmas.

LEMMA A.1. *Let C, $\tilde{C}$, and $\bar{C}$ be as in* (*AVI*), (3) *and* (10); *V and Y be as in* (6) *and Lemma* 3.1. *Then*

$$\operatorname{rec} C = V(\operatorname{rec} \tilde{C}), \tag{27}$$

$$\operatorname{rec} \tilde{C} = Y(\operatorname{rec} \bar{C}), \quad \textit{and} \tag{28}$$

$$V^T\big((\operatorname{rec} C)^D\big) = (\operatorname{rec} \tilde{C})^D, \tag{29}$$

$$Y^T\big((\operatorname{rec} \tilde{C})^D\big) = (\operatorname{rec} \bar{C})^D. \tag{30}$$

Furthermore

$$V^T\big(\operatorname{ri}((\operatorname{rec} C))^D\big) = \operatorname{ri}(\operatorname{rec} \tilde{C})^D, \tag{31}$$

$$Y^T\big(\operatorname{ri}(\operatorname{rec} \tilde{C})^D\big) = \operatorname{ri}(\operatorname{rec} \bar{C})^D. \tag{32}$$

PROOF. (27) and (28) are obvious from definition.

Based on these two equations and Rockafellar (1970, Corollary 16.3.2), we have

$$\begin{aligned}(\operatorname{rec} C)^D &= -(\operatorname{rec} C)^o = -(V \operatorname{rec} \tilde{C})^o \\ &= -(V^T)^{-1}(\operatorname{rec} \tilde{C})^o = (V^T)^{-1}(\operatorname{rec} \tilde{C})^D,\end{aligned}$$

where $K^o = -K^D$ is the polar cone of K and $(V^T)^{-1}$ is the inverse image of the linear map V^T (also see Rockafellar 1970). Similarly

$$(\operatorname{rec} \tilde{C})^D = (Y \operatorname{rec} \bar{C})^D = (Y^T)^{-1}(\operatorname{rec} \bar{C})^D.$$

So we have proven (29) and (30).

(31) and (32) can be obtained from (29) and (30) by applying Rockafellar (1970, Theorem 6.6). □

LEMMA A.2. *For* $z \in \operatorname{rec} C$, $\tilde{z} \in \operatorname{rec} \tilde{C}$, *and* $\bar{z} \in \operatorname{rec} \bar{C}$, *define*

$$D(z) := \left\{d \in (\operatorname{rec} C)^D | \langle d, z\rangle = 0\right\},$$

$$\tilde{D}(\tilde{z}) := \left\{\tilde{d} \in (\operatorname{rec} \tilde{C})^D | \langle \tilde{d}, \tilde{z}\rangle = 0\right\},$$

$$\bar{D}(\bar{z}) := \left\{\bar{d} \in (\operatorname{rec} \bar{C})^D | \langle \bar{d}, \bar{z}\rangle = 0\right\}.$$

Then

$$\tilde{D}(\tilde{z}) = V^T D(V\tilde{z}\ ), \tag{33}$$

$$\bar{D}(\bar{z}) = Y^T \tilde{D}(Y\bar{z}), \tag{34}$$

where V *and* Y *are as in* (6) *and Lemma* 3.1.

PROOF.

$$\tilde{D}(\tilde{z}) = \left\{\tilde{d} \in (\operatorname{rec} \tilde{C})^D | \langle \tilde{d}, \tilde{z}\rangle = 0\right\} = \left\{\tilde{d} \in V^T(\operatorname{rec} C)^D | \langle \tilde{d}, \tilde{z}\rangle = 0\right\}$$

$$= V^T\left\{d \in (\operatorname{rec} C)^D | \langle d^T, V\tilde{z}\rangle = 0\right\} = V^T D(V\tilde{z}).$$

The other equation can be proven similarly. □

Actually, for $z \in \operatorname{rec} C$, $D(z)$ is the set of vectors defining faces of $\operatorname{rec} C$ containing z, a vector z' is in every face of $\operatorname{rec} C$ containing z if and only if $\langle d, z'\rangle = 0$ and all $d \in D(z)$. Similar observation can also be made for the set $\tilde{C}$ and $\bar{C}$.

LEMMA A.3. *For* $w \in (\operatorname{rec} C)^D$, $\tilde{w} \in (\operatorname{rec} \tilde{C})^D$, *and* $\bar{w} \in (\operatorname{rec} \bar{C})^D$, *define*

$$R(w) := \{r \in \operatorname{rec} C | \langle r, w\rangle = 0\},$$

$$\tilde{R}(\tilde{w}) := \left\{\tilde{r} \in \operatorname{rec} \tilde{C} | \langle \tilde{r}, \tilde{w}\rangle = 0\right\},$$

$$\bar{R}(\bar{w}) := \left\{\bar{r} \in \operatorname{rec} \bar{C} | \langle \bar{r}, \bar{w}\rangle = 0\right\}.$$

Then

$$V\tilde{R}(V^T w) = R(w), \tag{35}$$

$$Y\bar{R}(Y^T \tilde{w}) = \tilde{R}(\tilde{w}), \tag{36}$$

where V *and* Y *are as in* (6) *and Lemma* 3.1.

PROOF.

$$R(w) = \{r \in \operatorname{rec} C | \langle r, w\rangle = 0\} = \left\{r \in V(\operatorname{rec} \tilde{C}) | \langle r, w\rangle = 0\right\}$$

$$= V\left\{\tilde{r} \in \operatorname{rec} \tilde{C} | \langle \tilde{r}, V^T w\rangle = 0\right\} = V\tilde{R}(V^T w).$$

The other equation can be proven similarly. □

Similar to the case of Lemma A.2, for $w \in (\operatorname{rec} C)^D$, $R(w)$ is the set of vectors defining faces of $(\operatorname{rec} C)^D$ containing w, a vector w' is in every face of $(\operatorname{rec} C)^D$ containing w if and only if $\langle r, w'\rangle = 0$ for all $r \in R(z)$. The situation is similar for the set $\tilde{C}$ and $\bar{C}$.

Now, we come to the invariance of the L-matrix property.

LEMMA A.4. *Given the problems* (3) *and* (10). *Suppose* $\tilde{A}$ *is an L-matrix with respect to* $\operatorname{rec} \tilde{C}$, *then* $\bar{A}$ *is an L-matrix with respect to* $\operatorname{rec} \bar{C}$.

PROOF. For $\bar{z} \in \operatorname{rec} \bar{C}, Y\bar{z} \in \operatorname{rec} \tilde{C}$. For any $\bar{q} \in \operatorname{ri}(\operatorname{rec} \bar{C})^D$, there exists $\tilde{q} \in \operatorname{re}(\operatorname{rec} \tilde{C})^D$ such that $\bar{q} = Y^T\tilde{q}$ due to (32). If $\bar{A}\bar{z} + \bar{q} \in (\operatorname{rec} \bar{C})^D$ then

$$Y^T\tilde{A}Y\bar{z} + Y^T\tilde{q} \in (\operatorname{rec} \bar{C})^D,$$

by definition of $\bar{A}$. Hence

$$\langle \tilde{A}Y\bar{z} + \tilde{q}, Y\bar{z}\rangle = \langle Y^T\tilde{A}Y\bar{z} + Y^T\tilde{q}, \bar{z}\rangle \geq 0, \quad \forall \bar{z} \in \operatorname{rec} \bar{C}.$$

It follows from (28) that

$$\langle \tilde{A}Y\bar{z} + \tilde{q}, \tilde{z}\rangle \geq 0, \quad \forall \tilde{z} \in \operatorname{rec} \tilde{C}.$$

Thus $\tilde{A}Y\bar{z} + \tilde{q} \in (\operatorname{rec} \tilde{C})^D$. Therefore $\bar{z}$ satisfies

$$\bar{z} \in \operatorname{rec} \bar{C}, \quad \bar{A}\bar{z} + \bar{q} \in (\operatorname{rec} \bar{C})^D, \quad \text{and} \quad \bar{z}^T\left(\bar{A}\bar{z} + \bar{q}\right) = 0, \tag{37}$$

with $\bar{q} \in \operatorname{ri}(\operatorname{rec} \bar{C})^D$, implying that $Y\bar{z}$ satisfies

$$Y\bar{z} \in \operatorname{rec} \tilde{C}, \quad \tilde{A}Y\bar{z} + \tilde{q} \in (\operatorname{rec} \tilde{C})^D, \quad \text{and} \quad (Y\bar{z})^T\left[\tilde{A}(Y\bar{z}) + \tilde{q}\right] = 0, \tag{38}$$

with $\tilde{q} \in \operatorname{ri}(\operatorname{rec} \tilde{C})^D$. Thus, the solution $Y\bar{z}$ of (38) is contained in $\operatorname{lin} \tilde{C} = \{0\}$, which implies that $\bar{z} = 0$. Thus the solution of (37) is $\{0\} \subset \operatorname{lin} \bar{C}$.

For any $0 \neq \bar{z} \in \operatorname{rec} \bar{C}$ such that

$$\bar{A}\bar{z} \in (\operatorname{rec} \bar{C})^D \quad \text{and} \quad \bar{z}^T\bar{A}\bar{z} = 0,$$

we have, $0 \neq Y\bar{z} \in \operatorname{rec} \tilde{C}$, and

$$\tilde{A}Y\bar{z} \in (\operatorname{rec} \tilde{C})^D \quad \text{and} \quad (Y\bar{z})^T\tilde{A}(Y\bar{z}) = 0.$$

So, there exists $0 \neq \tilde{z} \in \operatorname{rec} \tilde{C}$ such that $\tilde{z}$ is contained in every face of $\operatorname{rec} \tilde{C}$ containing $Y\bar{z}$, and $-\tilde{A}^T\tilde{z}$ is contained in every face of $(\operatorname{rec} \tilde{C})^D$ containing $\tilde{A}Y\bar{z}$. That is

$$\langle \tilde{d}, \tilde{z}\rangle = 0. \quad \forall \tilde{d} \in \tilde{D}(Y\bar{z}),$$

$$\langle \tilde{r}, -\tilde{A}^T\tilde{z}\rangle = 0, \quad \forall \tilde{r} \in \tilde{R}(\tilde{A}Y\bar{z}).$$

Consequently, there exists $0 \neq \bar{z}' \in \operatorname{rec} \bar{C}$ such that $\tilde{z} = Y\bar{z}'$. For any $\bar{d} \in \bar{D}(\bar{z}), \bar{d} = Y^T\tilde{d}$ for some $\tilde{d} \in \tilde{D}(Y\bar{z})$. Hence

$$\langle \bar{d}, \bar{z}'\rangle = \langle Y^T\tilde{d}, \bar{z}'\rangle = \langle \tilde{d}, Y\bar{z}'\rangle = 0.$$

So, $\bar{z}'$ is contained every face of $\operatorname{rec} \bar{C}$ containing $\bar{z}$. Moreover, for any $\bar{r} \in \bar{R}(\bar{A}\bar{z})$

$$\langle \bar{r}, -\bar{A}^T\bar{z}'\rangle = \langle Y\bar{r}, -\tilde{A}^TY\bar{z}'\rangle = \langle Y\bar{r}, -\tilde{A}^T\tilde{z}\rangle = 0,$$

since $Y\bar{z} \in \tilde{R}(\tilde{A}Y\bar{z})$. We see that $-\bar{A}^T\bar{z}'$ is contained in every face of $(\operatorname{rec} \bar{C})^D$ containing $\bar{A}\bar{z}$. Thus $\bar{A}$ is an L-matrix with respect to $\bar{C}$. □

Lemma A.5. *Given the problems (NE) and (3). Suppose A is an L-matrix with respect to* rec C, *then $\tilde{A}$ is an L-matrix with respect to* rec $\tilde{C}$.

Proof. For any $\tilde{z} \in \operatorname{rec} \tilde{C}$, $V\tilde{z} \in \operatorname{rec} C$ and

$$U\tilde{z} = \left(V - W(W^TAW)^{-1}W^TAV\right)\tilde{z} = V\tilde{z} - W(W^TAW)^{-1}W^TAV\tilde{z} \in \operatorname{rec} C,$$

since $W(W^TAW)^{-1}W^TAV\tilde{z} \in \operatorname{lin} C$. For any $\tilde{q} \in \operatorname{ri}(\operatorname{rec} \tilde{C})^D$, there exists $q \in \operatorname{ri}(\operatorname{rec} C)^D$ such that $\tilde{q} = V^Tq$. If $\tilde{A}\tilde{z} + \tilde{q} \in (\operatorname{rec} \tilde{C})^D$ then

$$U^TAU\tilde{z} + V^Tq \in (\operatorname{rec} \tilde{C})^D, \qquad q \in (\operatorname{rec} C)^D,$$

by definition of $\tilde{A}$. But

$$U^TAU = V^TAU - V^TA^TW(W^TAW)^{-T}W^TAU = V^TAU,$$

since $W^TAU = 0$, as can be directly verified. Thus

$$V^T(AU\tilde{z} + q) = V^TAU\tilde{z} + V^Tq \in (\operatorname{rec} \tilde{C})^D, \qquad q \in (\operatorname{rec} C)^D,$$

which implies

$$\langle AU\tilde{z} + q, V\tilde{z} \rangle = \langle V^T(AU\tilde{z} + q), \tilde{z} \rangle \geq 0, \quad \forall \tilde{z} \in \operatorname{rec} \tilde{C}.$$

It follows from (27) that

$$\langle AU\tilde{z} + q, z \rangle \geq 0, \quad \forall z \in \operatorname{rec} C.$$

Thus $AU\tilde{z} + q \in (\operatorname{rec} C)^D$. Also, $(U\tilde{z})^T[A(U\tilde{z}) + q] = \tilde{z}^T\tilde{A}\tilde{z} = 0$. Therefore $\tilde{z}$ satisfies

$$\tilde{z} \in \operatorname{rec} \tilde{C}, \quad \tilde{A}\tilde{z} + \tilde{q} \in (\operatorname{rec} \tilde{C})^D, \quad \text{and} \quad \tilde{z}^T\left(\tilde{A}\tilde{z} + \tilde{q}\right) = 0, \tag{39}$$

with $\tilde{q} \in \operatorname{ri}(\operatorname{rec} \tilde{C})^D$. This implies $U\tilde{z}$ satisfies

$$U\tilde{z} \in \operatorname{rec} C, \quad AU\tilde{z} + q \in (\operatorname{rec} C)^D, \quad \text{and} \quad (U\tilde{z})^T[A(U\tilde{z}) + q] = 0, \tag{40}$$

with $q \in \operatorname{ri}(\operatorname{rec} C)^D$. Hence the solution $U\tilde{z} \in \operatorname{lin} \operatorname{rec} C = \operatorname{lin} C$. But then

$$V\tilde{z} \in W(W^TAW)^{-1}A^TV\tilde{z} + \operatorname{lin} C \subset \operatorname{lin} C,$$

which, by the definition of V, implies $\tilde{z} = 0$. This shows that the solution of (39) is contained in $\operatorname{lin} \tilde{C} = \{0\}$.

For any $0 \neq \tilde{z} \in \operatorname{rec} \tilde{C}$ such that

$$\tilde{A}\tilde{z} \in (\operatorname{rec} \tilde{C})^D \quad \text{and} \quad \tilde{z}^T\tilde{A}\tilde{z} = 0,$$

we have $0 \neq U\tilde{z} \in \operatorname{rec} C$, and

$$V^TAU\tilde{z} = U^TAU\tilde{z} = \tilde{A}\tilde{z} \in (\operatorname{rec} \tilde{C})^D,$$

which implies $A(U\tilde{z}) \in (\operatorname{rec} C)^D$. We also have

$$(U\tilde{z})^TA(U\tilde{z}) = \tilde{z}^T\tilde{A}\tilde{z} = 0.$$

So, there exists $0 \neq z' \in \operatorname{rec} C$ such that z' is contained in every face of $\operatorname{rec} C$ containing $U\tilde{z}$, and that $-A^Tz'$ is contained in every face of $(\operatorname{rec} C)^D$ containing $A(U\tilde{z})$. That is

$$\langle d, z' \rangle = 0, \quad \forall d \in D(U\tilde{z}),$$

$$\langle r, -Az' \rangle = 0, \quad \forall r \in R(AU\tilde{z}).$$

Consequently, there exists $0 \neq \tilde{z}' \in \operatorname{rec} \tilde{C}$, such that $z' = V\tilde{z}'$, and for any $\tilde{d} \in \tilde{D}(\tilde{z})$, we have $\tilde{d} = V^Td$, for some $d \in D(V\tilde{z})$. Since $d \in (\operatorname{rec} C)^D, W^Td = 0$, therefore $\langle d, V\tilde{z} \rangle = \langle d, U\tilde{z} \rangle$, so $d \in D(V\tilde{z})$ implies $d \in D(U\tilde{z})$, hence

$$\langle \tilde{d}, \tilde{z}' \rangle = \langle V^Td, \tilde{z}' \rangle = \langle d, V\tilde{z}' \rangle = \langle d, z' \rangle = 0.$$

So, $\tilde{z}'$ is contained in every face of $\operatorname{rec} \tilde{C}$ containing $\tilde{z}$. For any $\tilde{r} \in \tilde{R}(\tilde{A}\tilde{z})$,

$$\begin{aligned}\langle \tilde{r}, -\tilde{A}^T\tilde{z}' \rangle &= \langle \tilde{r}, -U^TA^TU\tilde{z}' \rangle = \langle \tilde{r}, -U^TA^TV\tilde{z}' \rangle = \langle \tilde{r}, -U^TA^Tz' \rangle \\ &= \langle \tilde{r}, -V^TA^Tz' \rangle = \langle V\tilde{r}, -A^Tz' \rangle = \langle r, -A^Tz' \rangle = 0,\end{aligned}$$

since $r = V\tilde{r} \in R(AU\tilde{z})$ as a result of (36). This proves that $-\tilde{A}^T\tilde{z}'$ is contained in every face of $(\operatorname{rec} \tilde{C})^D$ containing $\tilde{A}\tilde{z}$. □

Acknowledgement. We are grateful to Professor Stephen Robinson for several enlightening discussion on the material contained in this paper and for his insights into the practicality of the normal equation formulation of the affine variational inequality problem. This material is based on research supported by the National Science Foundation Grant CCR-9157632 and the Air Force Office of Scientific Research Grant AFOSR-89-0410.

References

Burke, J. V. and J. J. Moré (1994). Exposing constraints. *SIAM J. Optimization* **4** 573–595.

Cottle, R. W., J. S. Pang and R. E. Stone (1992). *The Linear Complementarity Problem*. Academic Press, New York.

Dai, Y. and D. Talman (1993). Linear stationary point problems on unbounded polyhedra. *Math. Oper. Res.* **18** 635–644.

____, G. van der Laan, A. J. J. Talman and Y. Yamamoto (1991). A simplicial algorithm for the nonlinear stationary problem on an unbounded polyhedron. *SIAM J. Optimization* **1** 151–165.

Eaves, B. C. (1971). On the basic theorem of complementarity. *Math. Programming* **1** 68–87.

____ (1976). A short course in solving equations with pl homotopies. In R. W. Cottle and C. E. Lemke, Eds., *Nonlinear Programming*, pages 73–143, Providence, RI, 1976. American Mathematical Society, SIAM-AMS Proceedings.

____ (1978a). Computing stationary points. *Mathematical Programming Study* **7** 1–14.

____ (1978b). Computing stationary points, again. In O. L. Mangasarian, R. R. Meyer, and S. M. Robinson, Eds., *Nonlinear Programming 3*, pages 391–405. Academic Press, New York.

Gill, P. E., W. Murray and M. H. Wright (1981). *Practical Optimization*. Academic Press, London.

Golub, G. H. and C. F. Van Loan (1983). *Matrix Computations*. The John Hopkins University Press, Baltimore, MD.

Karamardian, S. (1976). Complementarity problems over cones with monotone and pseudomonotone maps. *J. of Optimization Theory and Applications* **18** 445–454.

Lemke, C. E. (1965). Bimatrix equilibrium points and mathematical programming. *Management Sci.* **11** 681–689.

The Math Works (1992). *MATLAB User's Guide*. The Math Works, Inc. Natick, MA.

Murty, K. G. (1976). *Linear and Combinatorial Programming*. John Wiley & Sons, NY.

____ (1988). *Linear Complementarity, Linear and Nonlinear Programming*. Helderman-Verlag, Berlin.

Ralph, D. (1992). On branching numbers of normal manifolds. Technical Report 92-1283, Computer Science Department, Cornell University, Ithaca, NY.

Robinson, S. M. (1992). Normal maps induced by linear transformations. *Math. Oper. Res.* **17** 691–714.

____ (1993). Nonsingularity and symmetry for linear normal maps. *Math. Prog.* **62** 415–425.

Rockafellar, R. T. (1970). *Convex Analysis*. Princeton University Press, Princeton, NJ.

Talman, A. J. J. and Y. Yamamoto (1989). A simplicial algorithm for stationary point problems. *Math. Oper. Res.* **14** 383–399.

Yamamoto, Y. (1987). A path-following algorithm for stationary point problems. *J. Operations Society of Japan* **30** 181–198.

M. C. Ferris: Computer Sciences Department, University of Wisconsin-Madison, 1210 West Dayton Street, Madison, Wisconsin 53706; e-mail: ferris@cs.wisc.edu