Control and Cybernetics

vol. **31** (2002) No. 3

A stable homotopy approach to horizontal linear complementarity problems

by

D. Ralph

The Judge Institute of Management, University of Cambridge, Trumpington St, Cambridge CB2 1AG, UK

Abstract: We are interested in the solution of Horizontal Linear Complementarity Problems, HLCPs, that is complementarity problems with more variables than equations. Globally metrically regular HLCPs have nonempty solution sets that are stable with respect to "right-hand-side perturbations" of the data, hence are numerically attractive. The main purpose of the paper is to show how the stability or conditioning properties of globally metrically regular HLCPs are preserved by a homotopy framework for solving the HLCP that finds a "stable" direction at each iteration as a local minimizer of a strongly convex quadratic program with linear complementarity constraints, QPCC. Apart from intrinsic interest in numerical solution of HLCPs, this investigation has application in solving horizontal nonlinear complementarity problems and more broadly in the area of mathematical programs with complementarity constraints, MPCCs.

Keywords: horizontal linear complementarity problem, mathematical program with complementarity constraints, piecewise affine system, global metric regularity, pseudo-Lipschitz continuity, stable solution, homotopy method, path following, active set method, MPCC, MPEC, QPCC, MPCC-LICQ.

1. Introduction

We investigate solving the Horizontal Linear Complementarity Problem, HLCP:

$$\begin{aligned} &0 = L(x, y) = Mx + Ny + q \\ &0 \le x \perp y \ge 0 \end{aligned} \tag{1}$$

where x and y are vectors of variables in $\mathbb{R}^n$, M and N are given matrices in $\mathbb{R}^{m\times n}$, $q \in \mathbb{R}^m$ is also given, and $\perp$ denotes orthogonality ($x^T y = 0$ above).

We are interested in $n \ge m$, i.e. feasibility problems. The solution set of (1), denoted $\mathcal{F}$, will be polyhedral and generally nonconvex, i.e. the union of finitely many closed, convex polyhedra. Hence, we will approach its solution

The focus of this paper is *globally metrically regular* HLCPs, a concept which says that if the equation $L(x,y) = 0$ is perturbed to $L(x,y) = p$ for any $p \in \mathbb{R}^m$, then the solution sets of the original HLCP and the perturbed HLCP will be nonempty and separated (in terms of Hausdorff distance) by at most a constant factor of $\|p\|$. That is, the solution set of the HLCP is stable with respect to arbitrary right-hand-side perturbations. It turns out that stability with respect to right-hand-side perturbations implies stability with respect to much more general (functional) perturbations, see Dontchev and Hager (1994), Kummer (1999), a fact we will not use, however. Section 2 provides formal definitions.

While global metric regularity is our topic, the approach we consider can very well be applied to investigate *local* metric regularity since, for instance, local metric regularity about a solution $\bar{x}$ of a piecewise affine system $\Pi(x) = 0$ can be characterised by global regularity of the directional derivative $\Pi'(\bar{x};\cdot)$. (We use a kind of converse of this idea in Lemma 4.1 where it is shown that the global condition number is inherited locally.)

There are several motivations for this investigation. First, consideration of $\mathcal{F}$ is a step towards handling general nonconvex polyhedral sets. Using the concrete problem class of HLCPs makes the development a little more direct, however; for example we have the advantage that the set of complementary nonnegative pairs (x,y) is a piecewise affine manifold, Eaves (1976), Robinson (1993), in $\mathbb{R}^{2n}$, which has a convenient and explicit structure. Second, consider Horizontal Nonlinear Complementarity Problems, HNCPs, which have the same format as (1) except that $L(x,y)$ is replaced by a smooth nonlinear function $F : \mathbb{R}^{2n} \to \mathbb{R}^m$. The papers of Dontchev (1996), Kummer (1999) give convergence theory for a general type of Newton method that can be applied to solving HNCPs by solving a sequence of HLCPs in a stable way. The newly published book by Klatte and Kummer (2002) is recommended for this and other material on solvability and solution stability of nonsmooth mappings relating to complementarity problems. Third, apart from intrinsic interest, horizontal linear and nonlinear complementarity problems are important in Mathematical Programs with Equilibrium Constraints, MPECs, and in particular Mathematical Programs with Complementarity Constraints, MPCCs, where they appear as constraints in what would otherwise be standard nonlinear programs. See the monographs Luo, Pang and Ralph (1996), Outrata, Kočvara and Zowe (1998) for an introduction to this area, as well as the more recent publications, Fletcher, Leyffer, Ralph and Scholtes (2002), Fukushima and Pang (2000), Hu and Ralph (2002), Huang, Yang and Zhu (2001), Jiang and Ralph (1999), Luo, Pang and Ralph (1998), Ralph (2001), Scheel and Scholtes (2000), Scholtes (2001, 2002), Scholtes and Stöhr (2001), Fukushima and Tseng (2002), that will be referred to later.

With regard to nonlinear programming formulations, note that if we rewrite the orthogonality condition in (1) as a bilinear equation $x^T y = 0$ then the entire system, though smooth, violates classical nonlinear programming constraint bility. Therefore we prefer to treat the HLCP by taking explicit account of its piecewise affine structure. Nevertheless standard nonlinear programming methods applied to problems with such constraints can be very successful, as shown in Fletcher, Leyffer, Ralph and Scholtes (2002).

The "most stable" solution of the HLCP with respect to a given point (x^0, y^0) is simply the globally nearest solution, i.e. a global minimizer of $\|(x,y)-(x^0,y^0)\|$ subject to $(x,y) \in \mathcal{F}$. When using the Euclidean norm, this is equivalent to finding a global minimum of the problem with a strongly convex quadratic objective function:

$$\min \tfrac{1}{2}\|(x,y)-(x^0,y^0)\|^2 \quad \text{subject to} \quad (x,y) \in \mathcal{F}. \tag{2}$$

This is called quadratic program with complementarity constraints, QPCC. Given the polyhedral nature of $\mathcal{F}$, global optimization can be carried out by an enumerative procedure: for each of the finitely many convex polyhedra whose union is $\mathcal{F}$, find the nearest point of this set to (x^0, y^0) by solving a strongly convex quadratic program, QP. A more sophisticated global optimization method might attempt to use the problem structure, namely complementarity, to set up a branch and bound framework. However, the heavy hammer of global optimization is unnecessary as we explain next.

The main purpose of this paper is to show that the stability property of globally metrically regular HLCPs can be readily transferred to a numerical solution method, namely a PA homotopy method. The basic homotopy or path following or continuation idea is, of course, rather standard for "square" systems like linear complementarity problems, as demonstrated by the classic paper of Cottle and Dantzig (1974) that employs Lemke's method for this purpose. (For square systems, see Cottle, Pang and Stone, 1992, for a full treatment of linear complementarity problems, Eaves, 1976, for a general homotopy approach to more general PA systems and, for homotopy methods in the nonlinear case, Allgower and Georg, 1990.) Homotopy approaches to feasibility problems are rare, although some feasibility problems can be written as projection problems whose stationary conditions admit a homotopy approach.

The proposed homotopy approach is greedy in that it attempts to locally optimize progress at each iteration by finding a suitable direction along which to generate the path. The direction-finding subproblem at iteration k is a QPCC that is formed by a kind of "localisation" of (2) about the iterate z^k on the homotopy path, for which we seek a *local* minimizer. It turns out that the direction-finding subproblem is always feasible and, almost always, every feasible point satisfies a linear-independence condition called the MPCC-LICQ, see Scheel and Scholtes (2000). This means that finding a local minimizer is computationally practical; indeed we may use any of the growing family of methods, Fletcher, Leyffer, Ralph and Scholtes (2002), Fukushima and Pang (2000), Hu and Ralph (2002), Huang, Yang and Zhu (2001), Jiang and Ralph (1999), Luo, Pang and Ralph (1998), Ralph (2001), Scholtes (2001, 2002), Scholtes and Stöhr

MPCC-LICQ, and possibly other conditions. We will demonstrate that any local minimizer of the direction-finding subproblem satisfies two stability properties necessary to show that the homotopy path reaches the feasible set of the HLCP after a finite number of iterations, while preserving the numerical conditioning of the HLCP. For concreteness, we show how an active set method, Scholtes (2002), for QPCC can be applied to the direction-finding subproblem. In short, we propose a framework for stable solution of (1) that is readily implementable.

From the standpoint of practicality, there may be other methods for finding feasible points of the HLCP that are easier to describe or are attractive due to fast or robust implementations. For example, any of the methods mentioned can be applied directly to the QPCC (2), though the stability properties of a solution obtained in this way would have to be investigated.

The paper is laid out in the following way. Basic definitions relating to metric regularity are given in Section 2. A formal homotopy approach is presented in Section 3 including a finite convergence result, based on "face-stable" directions, that appears to be new. Section 4 gives the direction-finding QPCC and shows that any of its local minimizers satisfies the stability properties required. An active set method which is a specialisation of Scholtes (2002) is applied to this QPCC under an MPCC-LICQ. In Section 5, the main result is that for almost all starting points $z^0 = (x^0, y^0)$, and every iterate $z^k = (x^k, y^k)$ on the homotopy path, the direction-finding QPCC is feasible such that the MPCC-LICQ holds at all feasible points. We conclude by summarising the properties of the hybrid Homotopy-Active-Set method.

Before proceeding, we give a simple example as motivation.

EXAMPLE 1.1 Let ϵ be a small positive number, and consider the line in the (x_1, x_2)-plane, $x_2 = -\epsilon x_1$. We form a closely related HLCP by taking $m = 1$, $n = 2$, $M = [\epsilon\ 1] \in \mathbb{R}^{2\times 1}$, $N = -M$ and $q = 0$.

The reason for choosing such a trivial example, apart from simplicity, is that its stability properties are clear: Take $x^0 = (0,0), y^0 = (0,0)$. If the right-hand-side is perturbed away from zero to $p \geq 0$, then $x = (0,p), y = (0,0)$ is a solution of the perturbed HLCP: $Mx + Ny = p$, $0 \leq x \perp y \geq 0$. This solution is at distance p from the origin (x^0, y^0). If $p \leq 0$, we may take $x = (0,0)$ and $y = (0, -p)$. In fact, the Hausdorff distance between solution sets associated with different right-hand-sides p and p' is is exactly $|p - p'|/\sqrt{1+\epsilon^2}$, hence less than $|p - p'|$ no matter how small ϵ becomes. (In Section 2, the quantity $1/\sqrt{1+\epsilon^2}$ will be called the modulus of metric regularity of the HLCP.)

Now consider an intuitive pivotal approach to solve the perturbed HLCP where the right-hand-side is $p > 0$, given (x^0, y^0) at the origin as above. We identify a *complementary basis*, Cottle and Dantzig (1974), that is a set of m variables that would not violate complementarity if all were positive, for which the corresponding submatrix of $[M N]$ is invertible. Here $m = 1$ and we are free to choose any single variable. For instance, take the basis as x_1 with all other system $\epsilon x_1 = p$, hence produces the solution $(x, y) = ((p/\epsilon, 0), (0, 0))$ which is at a large distance p/ϵ from (x^0, y^0).

The difficulty is that we cannot be sure which basis is stable in the sense of producing a solution that is near to our starting point. We are only repeating what has been long known in linear algebra: the condition number, Golub and van Loan (1989), of a full rank rectangular matrix $A \in \mathbb{R}^{\ell \times n}$, $\ell < n$, cannot be approximated by the condition number of an arbitrary basis matrix.

The above idea of using invertible complementary bases was the subject of a previous investigation by the author, Ralph (2002). While this has a certain consistency in the history of pivotal methods for mathematical programming, the example warns it can lead to unnecessary numerical difficulties if applied naively. We modify this idea later to allow submatrices of complementary columns of $[MN]$ that have full rank but are not necessarily invertible (square).

2. The modulus of metric regularity

We assume the HLCP is globally metrically regular, as defined below. Let

$$\mathcal{P} = \{(x, y) \in \mathbb{R}^{2n} : 0 \le x \perp y \ge 0\}.$$

This is a polyhedral (nonconvex) set, i.e. the union of finitely many convex polyhedra. The perturbed feasible set, for $p \in \mathbb{R}^m$, is

$$\mathcal{F}(p) = \{z \in \mathcal{P} : L(z) = p\} = \mathcal{P} \cap L^{-1}(p)$$

where z denotes (x, y). The distance of any point z to $\mathcal{F}(p)$ is

$$\mathrm{dist}(z \mid \mathcal{F}(p)) = \inf\{\|z - z'\| : z' \in \mathcal{F}(p)\}$$

which is taken to be ∞ if $\mathcal{F}(p) = \emptyset$.

Let U and V be nonempty subsets of $\mathbb{R}^{2n}$ and $\mathbb{R}^m$, respectively. The modulus of (metric) regularity of (1) with respect to U, V, denoted $\gamma_L(U, V)$, is the infimum of $\gamma \ge 0$ such that

$$\mathrm{dist}(z \mid \mathcal{F}(p)) \le \gamma \|L(z) - p\|, \quad \forall z \in U \cap \mathcal{P}, \quad p \in V.$$

Note that we require $z \in \mathcal{P}$ in the above definition since we are only interested in complementary solutions of $L(z) = p$.

If $\gamma_L(\mathbb{R}^{2n}, \mathbb{R}^m) < \infty$ then we say that $\mathcal{F}$ is globally metrically regular. Also, we write γ_L for $\gamma_L(U, V)$ when U and V are clear from the context.

The study of metric regularity and related concepts is long and deep. Having $\gamma_L < \infty$ is variously referred to as pseudo-Lipschitz continuity, Lipschitz-like behaviour, Aubin continuity, or openness at linear rate of the multifunction $\mathcal{F} : \mathcal{P} \to \mathbb{R}^{2n}$. It is the same as metric or pseudo regularity of the system $z \in \mathcal{P}, L(z) = 0$. See Dontchev (1996), Dontchev and Hager (1994), Klatte and

Remarks

1. If the HLCP (1) is globally metrically regular, then $L(\mathcal{P}) = \mathbb{R}^m$, hence $[MN]$ is surjective and of course has full rank. A partial converse is that $\gamma_L(\mathbb{R}^n, \mathbb{R}^m) = \infty$ if $[MN]$ does not have full rank, because then $L(\mathcal{P}) \neq \mathbb{R}^m$, hence $\mathcal{F}(p) = \emptyset$ for some $p \in \mathbb{R}^m$.
2. Suppose $\mathcal{P} = \mathbb{R}^{2n}$ (drop the complementarity requirement), $U = \mathbb{R}^n$ and $V = \mathbb{R}^m$ (global metric regularity). Write $A = [MN] \in \mathbb{R}^{m\times 2n}$. Then
   (a) $\gamma_L = \|A^{-1}\|$ if A is an invertible square ($2n = m$) matrix. Note the classical condition number of an invertible matrix $\|A\|\|A^{-1}\|$, Golub and van Loan (1989).
   (b) $\gamma_L = \|A^T(AA^T)^{-1}\|$ if $\mathrm{rank}(A) = m$ ($\Rightarrow 2n \geq m$), and we are using the 2-norm.
   (c) $\gamma_L = \infty$ if $2n < m$ since $[MN]$ does not have full rank.
3. Consider a subclass of HLCPs for which $M = [AB]$ and $N = [-I\ \Theta]$, where $A, I \in \mathbb{R}^{m\times m}$, I is the identity, $B, \Theta \in \mathbb{R}^{(n-m)\times m}$ and Θ is the zero matrix. Partitioning the variable vectors $x = (t, u)$, $y = (v, w)$ where $t, v \in \mathbb{R}^m$ and $u, w \in \mathbb{R}^{n-m}$, we see that (1) takes the form of a parametric linear complementarity problem in t whose complementary vector is v,
   $$0 \leq t \perp v = At + Bu + q \geq 0$$
   $$0 \leq u$$
   where u can be thought of as a parameter, and w, which is supposed to be nonnegative and orthogonal to u, plays no role and can be fixed as the zero vector. Suppose there is a unique solution (t, v) for any fixed u and q, e.g. A is positive definite or, more generally, a P-matrix, Cottle, Pang and Stone (1992). In this case the solution (t, v) is a piecewise affine, hence globally Lipschitz function of (u, q) and global metric regularity follows easily, in the style of Luo, Pang and Ralph (1996, Section 4.4).
4. Mordukhovich's coderivative calculus is a general tool that is useful in characterising and investigating necessary and sufficient conditions for local (and, by extension, global) regularity of systems of equations, and even systems posed using set-valued mappings. See Mordukhovich (1997) for an introduction to coderivatives and applications, and Mordukhovich (1996) for the particular case of stability of solution maps to parametric variational inequalities, of which HLCPs are a special case.

3. A homotopy framework for globally metrically regular H-LCPs

Henceforth we use the Euclidean or 2-norm, $\|\cdot\| = \|\cdot\|_2$.

Assume (1) is globally metrically regular with modulus of regularity γ. We apply a homotopy framework for solving this HLCP given a starting point $z^0 =$

- Let $p^0 = L(z^0)$, the initial *residual*.
- Derive a *path* $z(t) = (x(t), y(t))$ such that

$$z(0) = z^0, \quad z(t) \in \mathcal{P}, \quad L(z(t)) = (1-t)p^0 \quad \text{for } t > 0. \tag{3}$$

- If the path extends to $t = 1$, then $(x, y) = z(1)$ solves (1).

The size of the initial residual, $\|p^0\|$, is a measure of the distance from z^0 to the solution set of the HLCP.

Our goal is to give a method that generates a path of points $z(t) \in \mathcal{F}((1-t)p^0)$ as above, such that the global "condition number" γ is preserved at each point on the path,

$$\operatorname{dist}[z^0 \mid \mathcal{F}((1-t)p^0)] \le \|z^0 - z(t)\| \le t\gamma\|p^0\|.$$

That is, computationally, we want to replace "dist" term by $\|z^0 - z(t)\|$. In this case, when we reach a solution of the problem at $t = 1$, we will have $\|z^0 - z(1)\| \le \gamma\|p^0\|$, which does indeed preserve the conditioning of the HLCP. Such a vector $z(1)$ might be called a stable solution of the HLCP relative to z^0.

In fact, we will show that a continuous stability property holds all along the path: for any $s, t \in [0, 1]$,

$$\|z(s) - z(t)\| \le |s - t|\gamma\|p^0\|. \tag{4}$$

We will do this by generating $z(\cdot)$ as a PA path in $\mathcal{P}$ with breakpoints $\{z^k = (x^k, y^k)\}_{k=0}^{K}$, where K is to be determined, and with corresponding scalars $0 = t_0 < t_1 < \ldots < t_K = 1$ such that for $k = 0, \ldots, K-1$,

$$L(z^{k+1}) = (1 - t_{k+1})p^0 \tag{5}$$

$$\|z^{k+1} - z^k\| \le \gamma(t_{k+1} - t_k)\|p^0\| \tag{6}$$

$$z^k + s(z^{k+1} - z^k) \in \mathcal{P} \text{ for } s \in [0, 1]. \tag{7}$$

Thus, the major work at each iteration of the homotopy method will be to find a suitable direction d that can be scaled by a stepsize $s > 0$ in order to generate the next iterate, $z^{k+1} = z^k + sd$.

3.1. A formal homotopy method

Given $z = (x, y) \in \mathcal{P}$, let $T(z|\mathcal{P})$ be the tangent (contingent) cone of $\mathcal{P}$ at z:

$$\begin{aligned} &T((x,y)|\mathcal{P}) \\ &\quad = \{(u, v) \in \mathbb{R}^n \times \mathbb{R}^n : \begin{array}{ll} u_i = 0 & \text{if } x_i = 0 < y_i, \\ v_i = 0 & \text{if } x_i > 0 = y_i, \\ 0 \le u_i \perp v_i \ge 0 & \text{if } x_i = 0 = y_i\}. \end{array} \end{aligned}$$

Since $\mathcal{P}$ is polyhedral, then $d \in T(z|\mathcal{P})$ if and only if $z + sd \in \mathcal{P}$ for all sufficiently small $s > 0$. (Actually, since $\mathcal{P}$ is conical, the characterisation holds for all

At iteration k, suppose we have $z^k \in \mathcal{P}$ and $t_k \in [0,1)$ with $L(z^k) = (1 - t_k)p^0$. Consider the subproblem of finding a direction $d = (u,v) \in \mathbb{R}^{2n}$ such that $L(z^k + sd) = L(z^k) - sL(z^k)$ and $z^k + sd \in \mathcal{P}$ for all small $s > 0$, or equivalently,

$$[MN]d = -L(z^k) \quad \text{where} \quad d \in T(z^k|\mathcal{P}). \tag{8}$$

Any solution d of (8) and scalar s satisfy

$$L(z^k + sd) = (1-s)L(z^k) = (1-s)(1-t_k)p^0.$$

Also, $z^k + sd \in \mathcal{P}$ for small $s > 0$. Observing that $s \in (0,1]$ implies $(1-s)(1-t_k) = 1 - t$ where $t = t_k + s(1-t_k) \in (t_k, 1]$, it follows that d extends "the" path for all small $s > 0$. So, let d^k be a solution of (8).

Next let s_k be the maximum value of $s \in (0,1]$ such that $z^k + sd^k \in \mathcal{P}$, the latter condition being equivalent to nonnegativity of $z^k + sd^k$ because $z^k \in \mathcal{P}$ and $d^k = (u,v) \in T(z^k|\mathcal{P})$ implies $x^k + su \perp y^k + sv$ for all scalars s. Let $t_{k+1} = t_k + s_k(1-t_k)$, hence $t_{k+1} > t_k$ since $s_k > 0$. If $s_k = 1$, that is, $z^{k+1} = z^k + d^k \geq 0$, then $t_{k+1} = 1$ and z^{k+1} solves the HLCP.

We are now in a position to state a formal homotopy method for HLCP.

Homotopy Method

0. **Initial conditions.** We are given $z^0 \in \mathcal{P}$. Let $t_0 = 0$, $k = 0$.
1. **Direction.** Find a solution $d = d^k$ of the subproblem (8).
2. **Stepsize.** Let $s_k = \max\{s \in (0,1] : z^k + sd^k \geq 0\}$.
3. **Update.** Let $z^{k+1} = z^k + s_k d^{k+1}$, $t_{k+1} = t_k + s_k(1-t_k)$, and $k = k+1$.
4. **Stopping test.** If $t_k = 1$ then STOP; z^k solves HLCP.
5. **Next iteration.** Go to step 1.

This algorithm is still formal rather than computational in that we have not discussed existence or calculation of directions required in step 1. Nevertheless, the above discussion shows that it is well defined, provided that a solution d of (8) can be found (in step 1) at each iteration, in which case the path relations (5) and (7) are immediate. Moreover if each d^k satisfies $\|d^k\| \leq \gamma\|L(z^k)\|$, i.e. d^k is "stable" solution of (8), then the promised stable path property (6) (and hence (4)) also follows. Existence and computation such stable directions are discussed in Section 4.

Our next step in this section is to show finite termination of the Homotopy Method, assuming it is well defined, at a solution of the HLCP. To achieve this we will need a further condition on the direction generated at each iteration that is based on a natural decomposition of the polyhedral nonconvex set $\mathcal{P}$ (and its tangent cones) into finitely many closed, convex, polyhedra (cones).

3.2. Branches of $\mathcal{P}$ and face-stable directions

By $\mathcal{I}(z^k)$ we denote the family of (possibly empty) subsets I of the index set satisfy

$$I \supset \{i : x_i^k > 0\}, \quad I^c \supset \{i : y_i^k > 0\}.$$

Define

$$P_I = \{(x, y) : \begin{array}{ll} x_i \geq 0 = y_i & \text{if } i \in I, \\ x_i = 0 \leq y_i & \text{if } i \in I^c\}. \end{array}$$

Each set P_I corresponding to $I \in \mathcal{I}(z^k)$ is called a *branch* of $\mathcal{P}$ at z^k. (The total number of branches of $\mathcal{P}$ at all feasible points can be as large as 2^n, i.e. exponential in the dimension of x and y, which helps to explain why optimization over such a feasible set is an NP-hard problem.) It can easily be seen, Luo, Pang and Ralph (1996), for small neighborhoods U of z^k, that $U \cap \mathcal{P}$ is contained in $\cup_{I \in \mathcal{I}(z^k)} P_I$, hence that

$$T(z^k \mid \mathcal{P}) = \bigcup_{I \in \mathcal{I}(z^k)} T(z^k \mid \mathcal{P}_I).$$

By a *face of* $\mathcal{P}$ we mean a face of one of its branches. By a *face of* $T(z^k \mid \mathcal{P})$ we mean a face of one of the convex polyhedral cones $T(z^k \mid \mathcal{P}_I)$ with $I \in \mathcal{I}(z^k)$.

It is well known, Rockafellar (1970), that the relative interiors of the faces of a polyhedral convex set C partition that set. In other words, every member of C lies in the relative interior of a unique face of C. Also, if C is convex polyhedral cone, then the faces of C are also convex polyhedral cones. We state a minor extension of this to piecewise affine manifolds such as $\mathcal{P}$.

LEMMA 3.1 *Let $z \in \mathcal{P}$ and $d \in T(z \mid \mathcal{P})$. There is a unique face $\mathcal{F}$ of $\mathcal{P}$ such that $z \in \operatorname{rint} \mathcal{F}$; moreover any face of $\mathcal{P}$ containing z also contains $\mathcal{F}$. Likewise there is a unique face $\mathcal{K}$ of $T(z \mid \mathcal{P})$ such that $d \in \operatorname{rint} \mathcal{K}$; moreover any face of $T(z \mid \mathcal{P})$ containing d also contains $\mathcal{K}$.*

Proof. It can be verified by inspection that the intersection of two or more faces of $\mathcal{P}$ is also a (possibly) empty face of $\mathcal{P}$, and that this property is inherited by it tangent cones. Hence the first statement follows the above property of closed convex polyhedral sets C, by taking C to be any branch P_I with $I \in \mathcal{I}(z)$. Likewise, the second statement follows by taking $C = T(z \mid \mathcal{P}_I)$ for any $I \in \mathcal{I}(z)$ such that $d \in T(z \mid \mathcal{P}_I)$. ∎

DEFINITION 3.2 *Let $z^k \in \mathcal{P}$ and d be a solution of (8). If there is a face $\mathcal{K}$ of $T(z^k \mid \mathcal{P})$ such that d solves*

$$\begin{array}{ll} \min & \frac{1}{2}\|d\|^2 \\ \text{subject to} & [MN]d = -L(z^k) \\ & d \in \mathcal{K} \end{array} \tag{9}$$

then we say z^k is a face-stable solution of (8), with respect to $\mathcal{K}$, *or simply*

There is no restriction placed on a direction $d \in T(z^k \mid \mathcal{P})$ by requiring it to lie in a face $T(z^k \mid \mathcal{P})$. However, for d to be face-stable it must be the shortest solution of (8) associated with some face of $T(z^k \mid \mathcal{P})$. In fact (9) is a strongly convex quadratic program, because we are using the 2-norm. The existence of a unique solution (depending on $\mathcal{K}$) therefore follows if the problem is feasible, as it is assumed to be by the above definition.

EXAMPLE 3.3 Returning to Example 1.1, recall the 2×1 matrices $M = [\epsilon\ 1]$, $N = -M$, where ϵ is small and positive, and let $q = 0 \in \mathbb{R}$. We have already seen that this HLCP is globally metrically regular with $\gamma = 1/\sqrt{1+\epsilon^2}$. Let $z^k = (x^k, y^k)$, with $x^k = (1,0)$, $y^k = (0,0)$, which has a residual of $L(z^k) = Mx^k + Ny^k + q = \epsilon$.

There are three faces of $\mathcal{P}$ containing z^k: F_1, given by $x_1, x_2 \geq 0 = y_1 = y_2$; F_2, given by $x_1, y_2 \geq 0 = x_2 = y_1$; and F_3, given by $x_1 \geq 0 = x_2 = y_1 = y_2$. F_1 and F_2 are two-dimensional faces that share F_3 as a common facet, where F_3 contains z^0 in its relative interior. The faces of the tangent cone $T(z^k \mid \mathcal{P})$ are tangent cones of these faces at z^k:

$$\begin{aligned}
\mathcal{K}_1 &= \{(u,v) \in \mathbb{R}^{2\times 2} : u_2 \geq 0, v = 0\} \\
\mathcal{K}_2 &= \{(u,v) \in \mathbb{R}^{2\times 2} : v_2 \geq 0, u_2 = v_1 = 0\} \\
\mathcal{K}_3 &= \{(u,v) \in \mathbb{R}^{2\times 2} : u_2 = v_1 = v_2 = 0\}.
\end{aligned}$$

We list the face-stable solutions of the system (8), i.e. minimizers $d = d^i$ of the quadratic program (9) with $\mathcal{K} = \mathcal{K}_i$. This QP is infeasible for $\mathcal{K}_1$ and yields solutions $d^2 = ((-\epsilon^2, 0), (0, \epsilon))/(1+\epsilon^2)$ for $\mathcal{K}_2$ and $d^3 = ((-1,0),(0,0))$ for $\mathcal{K}_3$. The solution d^2 has length $\epsilon\gamma$, whereas the solution d^3 is less stable, having length 1.

The following statement is an immediate consequence of Lemma 3.1.

COROLLARY 3.4 *Let $d \in T(z^k \mid \mathcal{P})$ and $\widehat{\mathcal{K}}$ be the face of $T(z^k \mid \mathcal{P})$ containing d in its relative interior. Then d is a face-stable solution of (9) if and only if it is face-stable with respect to $\widehat{\mathcal{K}}$.*

Another corollary is that if the constraints (8) are feasible, then any global minimizer of $\frac{1}{2}\|d\|^2$ subject to these constraints is face-stable, because it must be face-stable with respect to any face that contains it. This idea will be further developed in the next section.

The main result of this section is that if the homotopy method uses face-stable directions at each iteration then it finds a solution of the HLCP after finitely many iterations. The proof shows that certain kinds of cycles are impossible, namely that the "worst" face containing a direction d^k in its relative interior — where worst means d^k has the largest norm relative to the rightof the proof) — contains no other direction in its relative interior. For subsequent iterations, the same argument shows that the relative interior of the second-worst face can only visited once etc.

THEOREM 3.5 *Let $z^0 \in \mathcal{P}$. Suppose, for each iteration k, that there exists a face-stable solution d^k of (8). Then the Homotopy Method is well defined and it terminates after finitely many iterations with a solution of (1).*

Proof. The only claim that has not been verified in previous discussion is finite termination of the method (note that finite termination implies that the last iterate calculated solves the HLCP).

Suppose the method generates the sequence $\{(z^k, t_k)\}_{k=0}^{k^{\max}}$ where $k^{\max}$ is either the iteration number at termination, or ∞ if the sequence does not terminate. For each k, let $\mathcal{F}^k$ be the face of $\mathcal{P}$ with $z^k \in \operatorname{rint} \mathcal{F}^k$ and $\mathcal{K}^k$ be the face of $T(z^k \mid \mathcal{P})$ with $d^k \in \operatorname{rint} \mathcal{K}^k$. From the above, $\{t_k\}$ is strictly increasing. For each $k < k^{\max}$, define

$$\gamma_k = \frac{\|z^{k+1} - z^k\|}{(t_{k+1} - t_k)\|p^0\|} = \frac{\|d^k\|}{(1 - t_k)\|p^0\|}. \tag{10}$$

Now $\widehat{d^k} = (z^{k+1} - z^k)/s_k$ is face-stable which means, by Corollary 3.4, that it is the shortest vector in $\mathcal{K}^k$ to satisfy $[MN]d^k = -L(z^k) = -(1-t_k)p^0$. It follows by a scaling argument that $\gamma_k = \gamma_j$ if $\mathcal{K}^k = \mathcal{K}^j$, and therefore that γ_k takes on only finitely many values even if $k^{\max} = \infty$.

Suppose the maximum value of γ^k occurs in iteration $\widehat{k}$ (and possibly in other iterations), and let $\widehat{\mathcal{F}} = \mathcal{F}^{\widehat{k}}$, $\widehat{\mathcal{K}} = \mathcal{K}^{\widehat{k}}$ and $\widehat{\mathcal{G}} = \mathcal{F}^{\widehat{k}+1}$. We will show that no subsequent iteration z^k can have $(\mathcal{F}^k, \mathcal{K}^k, \mathcal{F}^{k+1}) = (\widehat{\mathcal{F}}, \widehat{\mathcal{K}}, \widehat{\mathcal{G}})$. The same argument can be applied recursively by considering the subsequent iterates $\{z^k\}_{k \geq \widehat{k}}$ in order to eliminate another triple $(\mathcal{F}, \mathcal{K}, \mathcal{G})$ from appearing more than once, and so on. Polyhedrality of $\mathcal{P}$ implies that there are only finitely many distinct triples $(\mathcal{F}, \mathcal{K}, \mathcal{G})$, where $\mathcal{F}$ and $\mathcal{G}$ are faces of P, and $\mathcal{K}$ is the face of the tangent cone to $\mathcal{P}$ at some point. Hence, the recursive argument implies that the algorithm must terminate after a finite number of iterations.

Assume, to get a contradiction, that at some iteration $K \in (\widehat{k}, k^{\max})$ we have $(\mathcal{F}^K, \mathcal{K}^K, \mathcal{F}^{K+1}) = (\widehat{\mathcal{F}}, \widehat{\mathcal{K}}, \widehat{\mathcal{G}})$. Since $\widehat{\mathcal{K}}$ is a face $\widehat{\mathcal{F}}$ of $T(z^{\widehat{k}} \mid \mathcal{P})$ then $(z^{\widehat{k}} + \widehat{\mathcal{K}}) \cap \mathcal{P}$ is a face of $\mathcal{P}$, denoted $\bar{\mathcal{F}}$. This face contains $z^{\widehat{k}+1}$, since $d^{\widehat{k}} \in \widehat{\mathcal{K}}$. Also $z^{\widehat{k}+1}$ is a relative interior point of the face $\widehat{\mathcal{G}}$; thus $\bar{\mathcal{F}} \supset \widehat{\mathcal{G}}$. As a result, our assumption $\mathcal{F}^{K+1} = \widehat{\mathcal{G}}$ yields that $z^{K+1} \in z^{\widehat{k}} + \widehat{\mathcal{K}}$.

Next, the triangle inequality gives

$$\|z^{K+1} - z^{\widehat{k}}\| \leq \sum^{K} \|z^{k+1} - z^k\|.$$

By squaring both sides and applying the Cauchy-Schwartz inequality, we observe that the above inequality is satisfied as an equality only if the directions $d^k/\|d^k\|$ are identical for $k = \widehat{k}, \ldots, K$; in fact for some positive scalars $\tau_{\widehat{k}+1}, \ldots, \tau_K$ we have $d^k = \tau_k d_{\widehat{k}}$. Recalling that $z^{k+1} = z^k + s_k d^k$ for each k, this leads to

$$z^{K+1} = z^{\widehat{k}} + \sum_{\widehat{k}}^{K} s_k d_k = z^{\widehat{k}} + \Big(s_{\widehat{k}} + \sum_{\widehat{k}+1}^{K} s_k \tau_k\Big) d_{\widehat{k}}$$

which is not possible since, from Step 2 of the Homotopy Method, if $s > s_{\widehat{k}}$ then $z^{\widehat{k}} + sd^{\widehat{k}} \not\geq 0$ hence $z^{\widehat{k}} + sd^{\widehat{k}} \notin \mathcal{P}$. We conclude that strict inequality holds, namely

$$\begin{aligned} \|z^{K+1} - z^{\widehat{k}}\| &< \sum_{k=\widehat{k}}^{K} \|z^{k+1} - z^k\| = \sum_{k=\widehat{k}}^{K} \gamma_k (t_{k+1} - t_k)\|p^0\| \\ &\leq \gamma_{\widehat{k}}(t_{K+1} - t_{\widehat{k}})\|p^0\|. \end{aligned} \tag{11}$$

Now

$$[MN](z^{K+1} - z^{\widehat{k}}) = L(z^{K+1}) - L(z^{\widehat{k}}) = (t_{\widehat{k}} - t_{K+1})p^0.$$

Thus for

$$\widetilde{d} = \frac{1 - t_{\widehat{k}}}{t_{K+1} - t_{\widehat{k}}}(z^{K+1} - z^{\widehat{k}})$$

we have

$$[MN]\widetilde{d} = -(1 - t_{\widehat{k}})p^0 = -L(z^{\widehat{k}}),$$

i.e. $\widetilde{d}$ solves (8). We also have $z^{K+1} - z^{\widehat{k}} \in \widehat{\mathcal{K}}$ from above, hence $\widetilde{d} \in \widehat{\mathcal{K}}$. Finally, using (11) and then (10) gives

$$\|\widetilde{d}\| < (1 - t_{\widehat{k}})\|p^0\| = \|d^{\widehat{k}}\|.$$

The desired contradiction arises because $d^{\widehat{k}}$ is face-stable which means, by Corollary 3.4, that it must be the shortest vector in $\widehat{\mathcal{K}}$ satisfying (8). ■

The usual convergence technique, Allgower and Georg (1990), Eaves (1976), for homotopy methods applied to square systems is quite different to the above proof, since the former relies on the path being locally uniquely defined at least generically (for infinitesimal perturbations of the right-hand-side vector $-L(z^0)$). This means that there is no need to require global metric regularity, which has the advantage that the method may be well defined even if regularity double back on itself) along a continuous piece of the path without jeopardising existence of the path.

We mention further that homotopy approaches to square piecewise affine systems, explored extensively in Eaves (1976), and used for solving one-parameter LCPs in Lemke's method, Cottle and Dantzig (1974), Cottle, Pang and Stone (1992), have the nice property in the (generic) nondegenerate case, that at each iteration there is a unique pivot that determines the next line segment on the path. Looking ahead to the next section, we see for the homotopy methods applied to general HLCP, that there are possibly many pivot choices, some of which are more stable than others.

4. Finding stable directions

By a γ-*stable* solution of (8) we mean a vector d that satisfies both this system and the bound $\|d\| \leq \gamma \|L(z^k)\|$. Here, γ is the the modulus of global metric regularity of the HLCP (1). In Example 3.3 there are two face-stable directions d^2 and d^3 identified at the point z^k, of which only d^2 is γ-stable.

If there exists a γ-stable solution then, obviously, any global minimizer of the following problem must also be γ-stable:

$$\begin{array}{ll} \min & \frac{1}{2}\|d\|^2 \\ \text{subject to} & [MN]d = -L(z^k) \\ & d \in T(z^k \mid \mathcal{P}). \end{array} \tag{12}$$

This is the *direction-finding subproblem*, which can be written as a QPCC by formulating the constraints as in (13) below.

Subsection 4.1 contains the main result of the section, Proposition 4.3, which says that each local minimizer d of the direction-finding subproblem is both γ-stable and face-stable. This yields a considerable reduction of computational effort compared to global optimization with regard to γ-stability. Nevertheless, even verifying stationarity of a feasible point may require examination of an exponential number of branches (of the tangent cone), another combinatorial optimization problem.

To make the local minimization idea more concrete, in Subsection 4.2 we apply an active set method, that can be derived from Scholtes (2002), to (12). (We could have instead applied the piecewise sequential quadratic programming, PSQP, method of Luo, Pang and Ralph, 1996, see also Ralph, 2001, as described for the case of QPCC in the Remark following Proposition 2 of Jiang and Ralph, 1999.) Then we discuss a regularity condition, called the MPCC-LICQ, which is sufficient for the Active Set Method to find a local minimizer of (12). Initialisation of the Active Set Method is the subject of Subsection 4.3.

Looking ahead to Section 5, we see that the MPCC-LICQ almost always holds as needed, hence the Active Set Method produces directions that are both generate a stable path that reaches the HLCP solution set after finitely many iterations. Note that we are not limited to the Active Set Method, since there are many other methods that also produce local minimizers under the MPCC-LICQ, and perhaps other assumptions, at limit points of the iteration sequence, see Fletcher, Leyffer, Ralph and Scholtes (2002), Fukushima and Pang (2000), Hu and Ralph (2002), Huang, Yang and Zhu (2001), Luo, Pang and Ralph (1998), Scholtes (2001), Scholtes and Stöhr (2001), Fukushima and Tseng (2002).

4.1. Local minima of the direction-finding subproblem

Recall our assumption, throughout, that the HLCP (1) is globally metrically regular of modulus γ. We will show stability of local solutions of (12) using two minor results. The first, Lemma 4.1, says that localised systems like (8) are at least as stable as the whole system (1). This is to be expected since the global modulus of regularity is necessarily more conservative than the local modulus. The second, Lemma 4.2, is a general result about projections.

LEMMA 4.1 *Write $A = [MN]$. For any $z \in \mathcal{P}$, $\widehat{d} \in \mathbb{R}^{2n}$ and $\widehat{q} \in \mathbb{R}^m$,*

$$\operatorname{dist}(\widehat{d} \mid A^{-1}(\widehat{q}) \cap T(z \mid \mathcal{P})) \le \gamma \|A\widehat{d} - \widehat{q}\|.$$

Proof. Write $p = L(z)$ and observe for any $d \in \mathbb{R}^{2n}$ and nonzero scalar t that

$$Ad = \widehat{q} \Leftrightarrow L(z + td) - L(z) = t\widehat{q} \Leftrightarrow L(z + td) = p + t\widehat{q}.$$

Also, $d \in T(z \mid \mathcal{P})$ if and only if $z + td \in \mathcal{P}$ for all small enough $t > 0$. These observations imply that there exists a scalar $t > 0$ for which

$$d \in A^{-1}(\widehat{q}) \cap T(z \mid \mathcal{P}) \Leftrightarrow z + td \in L^{-1}(p + t\widehat{q}) \cap \mathcal{P} = \mathcal{F}(p + t\widehat{q}).$$

Thus

$$\begin{aligned} &t \operatorname{dist}(\widehat{d} \mid A^{-1}(\widehat{q}) \cap T(z \mid \mathcal{P})) = \operatorname{dist}(z + t\widehat{d} \mid \mathcal{F}(p + t\widehat{q})) \\ &\le \gamma \|L(z + t\widehat{d}) - p - t\widehat{q}\| = \gamma \|L(z + t\widehat{d}) - L(z) - t\widehat{q}\| = t\gamma \|A\widehat{d} - \widehat{q}\| \end{aligned}$$

and we are done. ■

Next, we have a general fact about "local projections".

LEMMA 4.2 *Let $\widehat{x}$ be a local solution of the problem $\min_{x \in S} \|x\|$, where S is a subset of a normed space. Then for some $t \in [0, 1)$, $\widehat{x}$ is a global solution of $\min_{x \in S} \|x - t\widehat{x}\|$.*

Proof. For any $t \in [0, 1]$, the triangle inequality yields

$$\mathbb{B}(t\widehat{x}, (1 - t)\|\widehat{x}\|) \subset \mathbb{B}(0, \|\widehat{x}\|),$$

where $\mathbb{B}(x, r)$ is the closed ball of centre x and radius $r \ge 0$. By hypothesis

$t \in [0, 1)$ large enough such that $\mathbb{B}(t\widehat{x}, (1-t)\|\widehat{x}\|)$ is contained in U. On the one hand, if $x \in U \cap S$ then

$$\|x - t\widehat{x}\| \geq \|x\| - t\|\widehat{x}\| \geq (1-t)\|\widehat{x}\|.$$

On the other hand, if $x \in S \setminus U$ then $x \notin \mathbb{B}(t\widehat{x}, (1-t)\|\widehat{x}\|)$, hence $\|x - t\widehat{x}\| \geq (1-t)\|\widehat{x}\|$. Finally, $\widehat{x}$ lies in $S \cap \mathbb{B}(t\widehat{x}, (1-t)\|\widehat{x}\|)$, so it must be a global solution of $\min_{x \in S} \|x - t\widehat{x}\|$. ■

We present the main result of this section. A consequence is that if Step 1 of the Homotopy Method is defined by taking d as a local minimizer of (12), then the Homotopy Method is well defined, terminates finitely by Theorem 3.5, and the iterates (z^k, t_k) satisfy the desired properties (5)–(7).

PROPOSITION 4.3 *The direction-finding subproblem (12) is feasible and bounded below, such that every local minimizer is a γ-stable and face-stable solution of (8).*

Proof. Feasibility of (12) is implicit in the statement and proof of Lemma 4.1, and of course the objective function is bounded below by zero.

Let $\widehat{d}$ be a local minimizer of (12). The face-stable property is straightforward. For suppose that $\mathcal{K}$ is a face of $T(z^k \mid \mathcal{P})$ containing $\widehat{d}$, hence $\widehat{d}$ is a local minimizer of the convex QP (9). Then, $\widehat{d}$ is also a global minimizer of the same QP, and by definition is face-stable.

Showing γ-stability requires more effort. Write $A = [MN]$ and $p^k = -L(z^k)$. From Lemma 4.2, there exists $\widehat{t} \in [0, 1)$ such that $\widehat{d}$ is a global solution of the problem

$$\begin{array}{ll} \min & \frac{1}{2}\|d - \widehat{t}\widehat{d}\|^2 \\ \text{subject to} & Ad = p^k \\ & d \in T(z^k \mid \mathcal{P}). \end{array}$$

Thus

$$\begin{aligned} (1-\widehat{t})\|\widehat{d}\| &= \operatorname{dist}(\widehat{t}\widehat{d} \mid A^{-1}(p^k) \cap T(z^k \mid \mathcal{P})) \\ &\leq \gamma\|L(\widehat{t}\widehat{d}) - p^k\| \quad \text{by Lemma 4.1} \\ &= \gamma(1-\widehat{t})\|p^k\| = \gamma(1-\widehat{t})\|L(z^k)\|. \end{aligned}$$

Since $\widehat{t} < 1$, $\|\widehat{d}\| \leq \gamma\|L(z^k)\|$ as claimed. ■

4.2. An active set method

We leave it to the reader to check, by examining the tangent cone $T(z^k \mid \mathcal{P})$, that the feasible set of (12) can be written with $d = (u, v) \in \mathbb{R}^n \times \mathbb{R}^n$ as

$$\begin{array}{rl} [MN]d = -L(z^k) & \\ u_i = 0 & \text{for } i \in I_+(y^k) \\ 0 \leq u_i \perp v_i \geq 0 & \text{for } i \in B_0(z^k) \end{array} \qquad (13)$$

where for $x, y \in \mathbb{R}^n$,

$$I_+(x) = \{i : x_i > 0\}$$
$$I_0(x) = \{i : x_i = 0\}$$
$$B_0(x, y) = I_0(x) \cap I_0(y).$$

The letter B in $B_0(x, y)$ refers to the description of "bi-active" indices that form this set. If orthogonality was omitted from these constraints, we could apply a standard active set method, Fletcher (1987), to solve the resulting strictly convex QP. It will be a simple matter to adapt the standard approach to (12) by adding conditions (on the "entering candidates" in Step 5 below) that preserve complementarity of iterates. The resulting Active Set Method will retain the advantage of relatively cheap iterates, by rank-1 updates of matrix factorizations, and the possible disadvantages of cycling and of the number of iterations being exponential in the problem dimension.

The following algorithm is actually a specialisation of the active set method for quadratic programs with complementarity constraints given by Scholtes (2002).

Each iteration of the method requires a feasible point $\widehat{d} = (\widehat{u}, \widehat{v})$ and a corresponding pair I, J of index sets such that

$$\begin{aligned} &I \cup J \subset \{1, \ldots, n\}, \quad I \cap J = \emptyset, \\ &I \supset I_+(x^k) \cup I_+(\widehat{u}), \quad J \supset I_+(y^k) \cup I_+(\widehat{v}), \\ &A = [M_I N_J] \text{ has full rank } m, \end{aligned} \tag{14}$$

where M_I and N_J are the respective submatrices of M and N corresponding to columns indexed by $i \in I$ and $j \in J$. (The condition $I \cap J = \emptyset$ is required for complementarity.) By a slight abuse of terminology from complementary pivoting methods for linear complementarity problems, e.g. Cottle and Dantzig (1974), we refer to the triple d, I, J as a *complementary basis*, though the matrix A is not necessarily square, and $\widehat{d}$ itself as a *complementary basic feasible solution* of (13).

The iteration requires the solution of a strictly convex QP,

$$\min \tfrac{1}{2}\|w\|^2 \quad \text{subject to} \quad Aw = -L(z^k)$$

by solving the linear Lagrangean system

$$\begin{bmatrix} \mathrm{Id} & A^T \\ A & 0 \end{bmatrix} \begin{pmatrix} w \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ -L(z^k) \end{pmatrix} \tag{15}$$

where Id is the $n \times n$ identity matrix and $\mu \in \mathbb{R}^m$ is a standard Lagrange

Active Set Method

0. **Initial conditions.** We are given a complementary basic feasible point d of (13) and associated index sets I, J such that (14) holds. Our initial iterate $\widehat{d} = (\widehat{u}, \widehat{v})$ is d.
1. **Solve Lagrangean system.** Let (w, μ) solve (15).
2. **Stepsize.** Let $\widehat{w} = (\widehat{u}_I, \widehat{v}_J)$ and K be set of indices of w such that w_K does not correspond to any u_i, $i \in I_+(x^k)$, or v_j, $j \in I_+(y^k)$.
   Let $s = \max\{s \in [0,1] : (1-s)\widehat{w}_K + sw_K \geq 0\}$.
3. **Update iterate.** Let $(\widehat{u}_I, \widehat{v}_J) = (1-s)\widehat{w} + sw$ and $\widehat{d} = (\widehat{u}, \widehat{v})$.
4. **Leaving index.** If $s < 1$ choose a *leaving* variable, either u_i for some $i \in I \setminus I_+(x^k)$ such that $\widehat{u}_i = 0$, or v_j for some $j \in J \setminus I_+(y^k)$ with $\widehat{v}_j = 0$.

   Update basis: If u_i (or v_j) is leaving, let $I = I \setminus i$ (or $J = J \setminus j$, respectively) and delete the corresponding column from A.

   Go to Step 1.
5. **Entering candidate list.** Let the list of *entering candidates* be u_i for $i \in I^c \cap B_0(z^k) \cap B_0(\widehat{d})$ and v_j for $j \in J^c \cap B_0(z^k) \cap B_0(\widehat{d})$.
   Let E be the corresponding submatrix of columns of $[MN]$.
6. **Stopping test.** If the list of entering candidates is empty, or $\mu^T E \geq 0$ then STOP and return $\widehat{d} = (\widehat{u}, \widehat{v})$.
7. **Entering index.** Let c be a column of E with $\mu^T c < 0$, and the corresponding entering candidate be u_i or v_j.

   Update basis: If u_i (or v_j) is entering let $I = I \cup i$, $J = J \setminus i$ (or $J = J \cup j$, $I = I \setminus j$, resp.) and insert the column c into the basis matrix A.

   Go to Step 1.

The list of entering candidates in Step 5 has to allow for any variable u_i that is not already basic, and whose *complement* v_i neither has a positive current value $\widehat{v}_i$ nor is associated with the *free* index set $I_+(y^k)$. Entering candidates v_j satisfy similar conditions.

By a basic linear algebraic argument that is familiar in standard active set methods, a column can only be deleted (Step 4) from the basis matrix A if the new basis matrix still has full rank. Of course, adding a column (Step 7) to the basis matrix cannot affect the rank since the rank of A is already at its maximum value m. Therefore, like convergence proofs of active set methods for QPs, an inductive argument using the representation (13) easily establishes that the vector $\widehat{d} = (\widehat{u}, \widehat{v})$ and the index sets I, J satisfy (14) at every iteration. This means that Step 1, and hence the entire method, is well defined and moreover that each iterate $(\widehat{u}, \widehat{v})$ is a feasible point of (8). We have just sketched a proof

PROPOSITION 4.4 *The Active Set Method is well defined and generates a sequence of feasible points* $(\widehat{u}, \widehat{v})$ *of (8).*

Having a well defined method, however, does not preclude the possibility of cycling, i.e. taking a sequence of steps in which $(\widehat{u}, \widehat{v})$ remains constant and in which some complementary basis is eventually repeated. Cycling may only occur if there is more than one candidate for the leaving index at some iteration, as can be shown easily by induction. These comments apply equally to active set methods applied to QPs.

The meaning of the stopping condition is not clear at a first glance, given the definition of the entering candidate list. We next look at a linear independence (full rank) condition that provides some justification for the active set approach.

We have already identified the direction-finding subproblem (12) as a QPCC. We say that the *MPCC linear independence constraint qualification*, MPCC-LICQ, holds at a feasible point d of this problem if the active constraint gradients of the formulation (13), ignoring the orthogonality conditions, are linearly independent. See Luo, Pang and Ralph (1996, 1998), Scheel and Scholtes (2000) for details and more general discussion. To be explicit, let $(u, v) = d$ be a feasible point of (13), $I = I_+(x^k) \cup I_+(u)$, and $J = I_+(y^k) \cup I_+(v)$. Then the MPCC-LICQ at d is equivalent to requiring full rank of $[M_I N_J]$. Comparing this with the definition of a complementary basis (14), we see that the MPCC-LICQ requires that the first two lines of (14) imply the third line.

The beauty of the MPCC-LICQ is that it relieves the combinatorial difficulty of checking stationarity of a feasible point, see Luo, Pang and Ralph (1998), Scheel and Scholtes (2000). The next result is more or less well known in the study of MPCC; it is a corollary of the results of Luo, Pang and Ralph (1998), Scheel and Scholtes (2000) and is used in the general active set approach of Scholtes (2002). A proof will not be given for this reason.

PROPOSITION 4.5 *Let* $\widehat{d} = (\widehat{u}, \widehat{v})$ *be one of the iterates generated by the Active Set method. Suppose the MPCC-LICQ holds for (13) at* $\widehat{d}$. *Then either* $\widehat{d}$ *is a local minimizer of (12) and the algorithm stops (in Step 6), or* $\widehat{d}$ *is not a local minimizer of (12) and the algorithm takes a nonzero step (to strictly decrease the value of* $\|\widehat{d}\|$*) in the next iteration.*

An implication of this result is that cycling cannot occur at an iterate satisfying the MPCC-LICQ; the analogous result is well known in standard active set methods for QP. If, by luck, every iterate $(\widehat{u}, \widehat{v})$ generated by the Active Set Method satisfies the MPCC-LICQ, then the method is bound to stop, after a finite number of iterations, at a local minimizer of (12). Actually, it has been shown, Scholtes and Stöhr (2001), that the MPCC-LICQ is a generic property of MPCCs, a result that we will be able to use in Section 5 to show that the

4.3. Initialising the Active Set Method under another linear independence condition

Consider the homotopy path described by (3) which we restate here for convenience:

$$L(z) = (1-t)L(z^0), \quad 0 \leq x \perp y \geq 0.$$

As previously, let (z^k, t_k) be a point on the path with $z^k = (x^k, y^k)$. We now show how to find a starting complementary basic feasible solution of the feasible set (13), which is required in Step 0 of the Active Set Method.

We will see shortly in Proposition 5.1, for almost all z^0, that the MPCC-LICQ holds at all feasible points (z, t) of the path system (3) where t is considered to be a variable, i.e. $[M_I N_J L(z^0)]$ has full rank for $I = I_+(x^k)$ and $J = I_+(y^k)$. For now, we take this MPCC-LICQ condition at (z^k, t_k) for granted.

The easy case for defining an initial complementary basic feasible solution of (13) is when $[M_I N_J]$ has full rank, in which case we determine a vector $d = (u, v)$ with $[M_I N_J](u_I, v_J) = -L(z^k)$ and $u_{I^c} = 0$, $v_{J^c} = 0$. The hard case is when $[M_I N_J]$ is rank deficient, in which case full rank of $[M_I N_J L(z^0)]$ implies $\operatorname{rank}[M_I N_J] = m - 1$. The latter case can still be dealt with efficiently as we next show.

LEMMA 4.6 *Let (z^k, t_k) be a point on the homotopy path (3) at which MPCC-LICQ holds, $z^k = (x^k, y^k)$ and $I = I_+(x^k)$, $J = I_+(y^k)$. Either rank $[M_J\ N_J] = m$ or there exists a variable, either u_i or v_i for some $i \in B_0(z^k)$, and a corresponding column c, equal to M_i or N_j respectively, such that $[M_I N_J c]$ has full rank and for some $(u_I, v_J) \geq 0$ and scalar $\delta \geq 0$,*

$$M_I u_I + N_J v_J + \delta c = -L(z^k). \tag{16}$$

Proof. The result derives from the investigation of Ralph (2002). We sketch the proof.

We may assume without loss of generality that $\operatorname{rank}[M_I N_J] = m - 1$. Let $A = [M_I N_J]$ and $\mathcal{B}$ represent the family of matrices $B = [Ac]$ for which c is either M_i or N_i for some $i \in B_0(z^k)$. Assume, to contradict global metric regularity of (1), that none of the matrices $B \in \mathcal{B}$ has the required property. Consider the halfspace H of vectors (u_I, v_I, δ) where u_I and v_J are arbitrary, while $\delta \geq 0$. Observe that each $B \in \mathcal{B}$ maps H either to a half space or a hyperplane ($(m-1)$-dimensional subspace) in $\mathbb{R}^m$. If $B(H)$ is a hyperplane then c is necessarily in the range space (i.e. column space) of A, denoted G', hence $B(H) = G'$. If B maps H to a halfspace then the latter is defined by

$$H' = G' + \{\delta L(z^k) : \delta \geq 0\},$$

A separate argument shows that $[MN](T(z^k \mid \mathcal{P}))$ is the convex hull of the sets $B(H)$, for $B \in \mathcal{B}$. It follows that $[MN](T(z^k \mid \mathcal{P}))$ is contained in H', i.e. $[MN](T(z^k \mid \mathcal{P}))$ is not equal to $\mathbb{R}^m$. The equivalent systems (8) and (13) are therefore not metrically regular, contradicting Lemma 4.1. ■

Clearly, any solution u_J, v_J, δ, described above, corresponds to a complementary basic feasible solution of (13). Such a solution can be found by checking (16) for each of the columns c equal to M_i or N_i for $i \in B_0(z^k)$ in the following straightforward way. Given c, factorize $[MNc]$ (or its transpose, e.g. using the QR factorization, Golub and van Loan, 1989) in order to determine first a solution of $(\widehat{u}_J, \widehat{v}_J, \widehat{\delta})$ of (16), and, second, a basis for the nullspace or kernel of $[MNc]$. If $\widehat{\delta} < 0$ then existence of a solution of (16) with $\delta \geq 0$ is equivalent to the basis having at least one column whose last component is nonzero. If a particular column c does not provide satisfaction then rank-1 updates can be used to replace c by another valid column. The rank-1 updating procedure will also detect any column c for which $[MNc]$ has rank $m-1$; these columns are to be discarded. The first factorization could be carried out on $[MNL(z^k)]$, which has full rank because $L(z^k)$ is a nonzero multiple of $L(z^0) = p^0$, after which the rank-1 updating occurs as described.

5. Stability of the Homotopy-Active-Set Method for almost all starting points

The aim of this section is to complete the task of showing that the Homotopy Method is practical. We already know, by using a local minimizer of the direction-finding subproblem (12) at each iteration, that the Homotopy Method will terminate after finitely many steps at a stable solution of the HLCP. We know further that finding a local solution of (12) is practical, by applying a method designed for QPCCs or possibly MPCCs, Fletcher, Leyffer, Ralph and Scholtes (2002), Fukushima and Pang (2000), Hu and Ralph (2002), Huang, Yang and Zhu (2001), Luo, Pang and Ralph (1998), Scholtes (2001), (2002), Scholtes and Stöhr (2001), Fukushima and Tseng (2002), provided a suitable MPCC-LICQ holds at the iterates encountered by the method. It is left to Subsection 5.1 to prove, using Scholtes and Stöhr (2001), that the MPCC-LICQ holds as required throughout the Homotopy Method, for almost all initial points $z^0 = (x^0, y^0)$ of the method. Subsection 5.2 summarises the convergence properties of the hybrid method, which combines the Homotopy and Active Set methods, and also briefly considers the degenerate case in which some of the iterates (z^k, t_k) do not yield the required MPCC-LICQ properties.

5.1. MPCC-regular systems of complementarity constraints

Given $z = (x, y) \in \mathcal{P}$ and $p \in \mathbb{R}^m$, we have seen, see, (13), that the conditions

$$
\begin{aligned}
[MN]d &= p \\
u_i &= 0 \quad \text{for } i \in I_+(y) \\
0 \le u_i \perp v_i &\ge 0 \quad \text{for } i \in B_0(z) \\
v_j &= 0 \quad \text{for } j \in I_+(x).
\end{aligned} \tag{17}
$$

We say that this system is *MPCC-regular* if the MPCC-LICQ holds at every feasible point of (17). (The term "regular" was used in Scholtes and Stöhr (2001) but we prefer MPCC-regular to help distinguish it from metric regularity.)

By almost all or a.a. we mean all points but for those in a set of Lebesgue measure zero. By a.a. $z \in \mathcal{P}$ we mean all points of $\mathcal{P}$ except for those in a set whose intersection with each of the n-dimensional branches P_I has Lebesgue measure zero with respect to the affine hull (i.e. subspace $P_I - P_I$) of that branch. (An alternative definition could use the fact that $\mathcal{P}$ is homeomorphic to $\mathbb{R}^n$, Robinson, 1993, and define a.a. $z \in \mathcal{P}$ via zero measure sets in $\mathbb{R}^n$.)

PROPOSITION 5.1

1. *For a.a. $p \in \mathbb{R}^m$, and each $z \in \mathcal{P}$, the system (17) is MPCC-regular.*
2. *For a.a. $z^0 \in \mathcal{P}$, the following statements both hold:*
   (a) The homotopy path system (3) is MPCC-regular.
   (b) For every (z, t) with $t < 1$ on the path (3), the local system (17) with $p = -(1-t)L(z^0)$ is MPCC-regular.

Proof. 1. It is easy to show for fixed $z = (x, y)$ (i.e. fixed index sets $I_+(x)$, $I_+(y)$, $B_0(z)$) and a.a. p, that the solutions of (17) satisfy the MPCC-LICQ. The short proof of Corollary 2 from Scholtes and Stöhr (2001), using Sard's theorem, can be immediately adapted for this purpose.

Part 1 of the result now follows because there are only finitely many distinct triples $(I_+(y), B_0(z), I_+(x))$ for all possible z.

2(a). Similar to the first paragraph of the proof, it is an easy corollary of Sard's theorem that the system in (z, t) given by $L(z) = (1-t)p$, $0 \le x \perp y \ge 0$ is MPCC-regular for a.a. $p \in \mathbb{R}^m$.

For 2(b), the fact that L maps neighborhoods in $\mathcal{P}$ to neighborhoods in $\mathbb{R}^m$, which is a corollary of Lemma 4.1, yields MPCC-regularity of (17) with $p = -L(z^0)$ for a.a. z^0 and every z, by Part 1. Let z^0 be a point such that (17) is MPCC-regular for $p = -L(z^0)$ and every z. Consider the scaling $p = -sL(z^0)$ for any $s > 0$, and, given z, denote the corresponding feasible set (17) by $\mathcal{G}(z, z^0, s)$. Obviously $\mathcal{G}(z, z^0, s) = s\mathcal{G}(z, z^0, 1)$, hence MPCC-regularity of $\mathcal{G}(z, z^0, 1)$ implies MPCC-regularity of $\mathcal{G}(z, z^0, s)$. This holds for any $s > 0$ and z, and we are done. ∎

Metric regularity is not required for Part 1 of the Proposition (or for Sard's theorem). It is used in Part 2 because the statement involves the domain space

5.2. A summary of the Homotopy-Active-Set Method

As in Section 4, consider the Homotopy Method in which the direction-finding subproblem is solved by the Active Set Method. We call this hybrid the Homotopy-Active-Set or HAS Method.

Under the standing assumption that the HLCP (1) is globally metrically regular of modulus γ, the following statements are valid for a.a. $z^0 \in \mathbb{R}^m$:

1. At iteration k of the HAS method, where $t_k < 1$:
   (a) the MPCC-LICQ holds for the homotopy path (3) at the point (z^k, t_k), hence the Active Set Method for the direction-finding subproblem (12) can be initialised as described in Subsection 4.3.
   [Proposition 5.1, Part 2(a); Lemma 4.6]
   (b) the feasible set (13) of (12) is MPCC-regular;
   [Proposition 5.1, Part 2(b)]
   (c) The Active Set Method terminates after finitely many iterations at a local minimum of (12), hence a solution of (8) that is both γ-stable and face-stable. [Proposition 4.5, Proposition 4.3]
2. (a) The HAS Method is well defined and terminates after finitely many iterations, in iteration K, with a solution $z = z^K$ of the HLCP (1).
   [Theorem 3.5]
   (b) The iteration sequence $\{(z^k, t_k)\}_1^K$ satisfies the stable path property (6), in particular $\|z^K - z^0\| \le \gamma \|L(z^0)\|$. In addition $\{t_k\}_1^K$ is strictly increasing. [Subsection 3.1]
   (c) The piecewise affine path defined on $[0, 1]$ by
   $$z(t) = \frac{t_{k+1} - t}{t_{k+1} - t_k} z^k + \frac{t - t_k}{t_{k+1} - t_k} z^{k+1}$$
   for $t \in [t_k, t_{k+1}]$ and $k = 0, \ldots, K-1$, satisfies $z(t) \in \mathcal{P}$, $L(z(t)) = (1-t)L(z^0)$ and the continuous path stability property (4).
   [Subsection 3.1]

At this point it would be useful to mention that there is potential for an anti-cycling strategy that would allow any starting point $z^0 \in \mathcal{P}$.

Consider an iterate z^k such that the feasible set (13) is not MPCC-regular. We know that arbitrarily small perturbations of z^0 will mend this situation. Therefore we can in principle refine the Active Set Method to use infinitesimal perturbations so that it will not cycle, but will terminate in finitely many iterations. The terminal iterate $\widehat{d}$ will not necessarily be a local minimizer of (12), unless the MPCC-LICQ holds at $\widehat{d}$, but nevertheless it will still be a point satisfying (13) that is both γ-stable and face-stable, by continuity arguments. A similar infinitesimal anti-cycling strategy would be needed to initialise the Active Set Method.

Lexicographic ordering is a well known anti-cycling procedure for pivotal algorithms. For homotopy methods applied to square piecewise affine homotopy systems, lexicographic ordering is used, Eaves (1976), to solve the *degenerate* terms of infinitesimal perturbations, where the perturbations restore regularity similar to the above description.

For HLCP and every starting point $z^0 \in \mathcal{P}$, such anti-pivoting strategies would yield the same convergence results as listed above, apart from 1(a) and 1(b), i.e. omitting the statements involving MPCC-LICQ or MPCC-regularity.

Acknowledgements

The author gratefully acknowledges the comments and suggestions of Stefan Scholtes and two anonymous referees.

References

ALLGOWER, E.L. and GEORG, K. (1990) *Numerical Continuation Methods.* Springer-Verlag, Berlin.

CHEN, Y. AND FLORIAN, M. (1995) The nonlinear bilevel programming problem: formulations, regularity and optimality conditions. *Optimization*, **32**, 193–209.

COTTLE R.W. and DANTZIG, G.B. (1974) Complementary pivot theory of mathematical programming. In: G.B. Dantzig and B.C. Eaves, eds., *Studies in Optimization* 10, MAA Studies in Mathematics, 27–51.

COTTLE, R.W., PANG, J.S. and STONE, R.E. (1992) *The Linear Complementarity Problem.* Academic Press, Boston.

DONTCHEV, A.L. (1996A) Local convergence of the Newton method for generalized equations. *C.R. Acad. Sc. Paris* **332** Ser. I, 327–331.

DONTCHEV, A.L. (1996B) The Graves theorem revisited. *J. Convex Analysis* **3**, 45–53.

DONTCHEV, A.L. and HAGER, W.W. (1994) Implicit functions, Lipschitz maps and stability in optimization. *Mathematics of Operations Research* **19**, 753–768.

EAVES, B.C. (1976) A short course in solving equations with PL homotopies. In: R.W. Cottle and C.E. Lemke, eds., *Nonlinear Programming* (SIAM-AMS Proceedings 9, 73–143), American Mathematical Society, Providence, RI.

FLETCHER, R. (1987) *Practical Methods of Optimization.* 2nd ed., John Wiley and Sons, New York.

FLETCHER, R., LEYFFER, S., RALPH, D. and SCHOLTES, S. (2002) Local convergence of SQP methods for Mathematical Programs with Equilibrium Constraints. Numerical Analysis Report NA/209, Department of Mathematics, University of Dundee.

FUKUSHIMA, M. and PANG, J.S. (2000) Convergence of a smoothing continuation method for mathematical programs with complementarity constraints. In: *Ill-posed Variational Problems and Regularization Techniques.*

FUKUSHIMA, M. and TSENG, P. (2002) An implementable active-set algorithm for computing a B-stationary point of the mathematical program with linear complementarity constraints. *SIAM J Optimization* **12**, 724–739.

GOLUB, G. and VAN LOAN, C. (1989) *Matrix Computations.* 2nd edition, The Johns Hopkins University Press, Baltimore.

HU, X. and RALPH, D. (2002) Convergence of a penalty method for mathematical programming with complementarity constraints. To appear in *Journal of Optimization Theory and Applications.*

HUANG, X.X., YANG, X.Q. and ZHU, D.L. (2001) A sequential smooth penalization approach to mathematical programs with complementarity constraints. Manuscript, Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China.

JIANG, H. and RALPH, D. (1999) QPECgen, a MATLAB generator for mathematical programs with quadratic objectives and affine variational inequality constraints. *Computational Optimization and Applications.* **13**, 25–59.

KLATTE, D. and KUMMER, B. (2002) *Nonsmooth Equations in Optimization.* Kluwer Academic Publishers, Boston.

KUMMER, B. (1999) Metric regularity: characterizations, nonsmooth variations and successive approximation. *Optimization*, **46**, 247–281.

LUO, Z.-Q., PANG, J.-S. and RALPH, D. (1996) *Mathematical Programs with Equilibrium Constraints.* Cambridge University Press.

LUO, Z.-Q., PANG, J.-S. and RALPH, D. (1998) Piecewise sequential quadratic programming for mathematical programs with nonlinear complementarity constraints. In: A. Migdalas, P.M. Pardalos and P. Värbrand, eds., *Multilevel Optimization: Algorithms and Applications.* Kluwer Academic Publishers, Dordrecht, 209–229.

MORDUKHOVICH, B.S. (1996) Generalized differentiation and stability of solutions maps to parametric variational inequalities. In: V. Lakshmikantham, ed., *World Congress of Nonlinear Analysts '92.* Proceedings of the First World Congress of Nonlinear Analysts, Tampa, Florida (August 19–26, 1992), W. De Gruyter, 2311–2323.

MORDUKHOVICH, B.S. (1997) Coderivatives of set-valued mappings: calculus and applications. *Nonlinear Analysis: Theory, Methods & Applications* **30**, 3059–3070.

OUTRATA, J., KOČVARA, M. and ZOWE, J. (1998) *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints.* Kluwer Academic Publishers, Boston.

RALPH, D. (2001) Optimization with equilibrium constraints: a piecewise SQP approach. In: C.A. Floudas and P.M. Pardalos, eds., *Encyclopedia of Optimization.* Kluwer Academic Publishers, Dordrecht.

RALPH, D. (2002) Notes on metrically regular systems. Manuscript, Judge Institute of Management, Cambridge University, Trumpington St, Cambridge, CB2 1AG, UK.

ROBINSON, S.M. (1993) Normal maps induced by linear transformations. *Mathematics of Operations Research*, **17**, 691–714.

ROCKAFELLAR, R.T. (1970) *Convex Analysis*. Princeton, New Jersey.

SCHEEL, H. and SCHOLTES, S. (2000) Mathematical Programs with Complementarity Constraints: Stationarity, Optimality, and Sensitivity. *Mathematics of Operations Research* **25**, 1–22.

SCHOLTES, S. (2001) Convergence Properties of a Regularization Scheme for Mathematical Programs with Complementarity Constraints. *SIAM J. Optim.*, **11** (4), 918–936.

SCHOLTES, S. (1999) Active set methods for inverse linear complementarity problems. Manuscript, Judge Institute of Management, Cambridge University, Trumpington St, Cambridge, CB2 1AG, UK.

SCHOLTES, S. and STÖHR, M. (2001) How stringent is the linear independence assumption for mathematical programs with complementarity constraints. *Mathematics of Operations Research*, **26**, 851–863.