MATHEMATICS OF OPERATIONS RESEARCH
Vol. 22, No. 2, May 1997
Printed in U.S.A.

ON THE CONVERGENCE OF THE ITERATION SEQUENCE OF INFEASIBLE PATH FOLLOWING ALGORITHMS FOR LINEAR COMPLEMENTARITY PROBLEMS

J. FRÉDÉRIC BONNANS AND FLORIAN A. POTRA

A generalized class of infeasible-interior-point methods for solving horizontal linear complementarity problems is analyzed and sufficient conditions are given for the convergence of the sequence of iterates produced by methods in this class. In particular it is shown that the largest step path following algorithms generates convergent iterates even when starting from infeasible points. The computational complexity of the latter method is discussed in detail and its local convergent rate is analyzed. The primal-dual gap of the iterates produced by this method is superlinearly convergent to zero. A variant of the method has quadratic convergence.

1. Introduction. In this paper we study the convergence of a class of infeasible interior point methods for solving the horizontal linear complementarity problem (HLCP):

$$xs = 0,$$
$$Qx + Rs = b, \tag{1.1}$$
$$x, s \geq 0,$$

where $b \in \mathbb{R}^n$, and $Q, R \in \mathbb{R}^{n \times n}$ are such that for any $u, v \in \mathbb{R}^n$,

$$Qu + Rv = 0 \quad \text{implies } u^T v \geq 0. \tag{1.2}$$

If Q, R is a pair of $n \times n$-matrices satisfying (1.2) we say that Q, R is a *positive semidefinite pair*. It is known (see, e.g., Bonnans and Gonzaga 1993) that this problem trivially includes the linear programming problem (LP) and the convex quadratic programming problem (QP) in their usual formulations, and thus provides a quite general framework for the study of algorithms. Of course any LP and QP can also be written as a standard linear complementarity problem (SLCP) which is a HLCP where R is the identity matrix and $-Q$ is positive semidefinite. As mentioned by Wright (1993, 1996) little loss of efficiency is involved in solving LPs and QPs by embedding them in an algorithm for SLCP, provided the linear algebra takes into account the specific structure of the individual problem. However, as will be seen below, in the analysis of the algorithms it is very convenient to permute the components of x and s such that x denotes "the large variables" and s denotes "the small variables". Then the SLCP becomes an HLCP which gives one more reason (besides the symmetry of the formulation) to consider HLCP. For recent theoretical work on the relationship between different formulations of linear complementarity problems we refer to Güler (1993), Tütüncü and Todd (1992), Anitescu, Lesaja and Potra (1994).

Received December 29, 1994; revised November 1, 1995 and July 20, 1996.
AMS 1991 subject classification. Primary: 90C33; Secondary: 49M15, 65K05.
OR/MS Index 1978 subject classification. Primary: Programming/complementarity.
Key words. Linear complementarity problem, infeasible central path, interior-point algorithm, predictor-corrector algorithm, largest step algorithm, shifted analytic center.

0364-765X/97/2202/0378/$05.00

The present paper is organized as follows. In §2 we present a generic infeasible path following algorithm and study its basic properties. In §3 we propose an extension of the largest step path following algorithm for arbitrary positive infeasible starting points and prove its global convergence. The computational complexity of this infeasible path following algorithm depends on the quality of the starting point. If the starting points are large enough then the algorithm has $O(nL)$ iteration complexity. If a certain measure of feasibility at the starting point is small enough then the algorithm has $O(\sqrt{n}L)$ iteration complexity. In particular, for feasible starting points we recover the result of McShane (1994). In §4 we give sufficient conditions for convergence of the sequences generated by the generic infeasible path following algorithm. Our results generalize the corresponding results obtained by Bonnans and Gonzaga (1993) in the feasible case. In the last section we use these general results to prove that for problems having a strictly complementary solution the sequence (x^k, s^k) of iterates generated by the largest step infeasible path following algorithm is convergent and that the sequences $(x^k)^T s^k$ and $r^k := b - Qx^k - Rs^k$ measuring "optimality" and "feasibility" are superlinearly convergent to zero. Moreover we show that the largest step infeasible path following algorithm can be modified à la Gonzaga (1994) so that the latter sequences are quadratically convergent while preserving the same global convergence and polynomial complexity as the original algorithm (actually, our analysis allows us to extend to infeasible algorithms the asymptotic results of Gonzaga (1994) and Gonzaga, Bonnans (1994)). For proving this, we generalize the analysis of the largest step path following algorithm done by Gonzaga (1994) in the feasible case. To our knowledge this algorithm is the first infeasible interior point algorithm for LCP with quadratic convergence and $O(nL)$ iteration complexity that uses only one matrix factorization and two backsolves per iteration. We mention that the infeasible interior point algorithms for LCP of Potra (1994) and Potra and Sheng (1994) also have quadratic convergence and $O(nL)$ iteration complexity but the former requires two matrix factorizations and three backsolves while the latter requires two matrix factorizations and two backsolves. Also, we note that the algorithms from Wright (1993, 1994, 1996), and Zhang (1994) have only $O(n^2L)$-iteration complexity. Kojima, Mizuno and Todd (1995), mention that the $O(nL)$ infeasible-interior-point algorithms for linear programming considered in that paper can be generalized for linear complementarity problems, but the superlinear convergence of the resulting algorithms has not yet been established. Very recently, Ye (1994) has proposed a homogeneous self-dual reformulation of a standard form linear complementarity problem, for which a feasible starting point is always available, and has proposed an $O(\sqrt{n}L)$ iteration algorithm for this problem, but has not proved any superlinear convergence results.

Conventions. By $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$ we denote the set of real, nonnegative real, and positive real numbers. Given a vector x, the corresponding upper case symbol denotes as usual the diagonal matrix X defined by the vector. The symbol e represents the vector of all ones, with dimension given by the context.

We denote component-wise operations on vectors by the usual notations for real numbers. Thus, given two vectors u, v of the same dimension, uv, u/v, etc. will denote the vectors with components $u_i v_i$, u_i/v_i, etc. This notation is consistent as long as component-wise operations always have precedence in relation to matrix operations. Note that $uv \equiv Uv$ and if A is a matrix, then $Auv \equiv AUv$, but in general $Auv \neq (Au)v$. We denote the nullspace and range space of a matrix A by $\mathcal{N}(A)$ and $\mathcal{R}(A)$ respectively.

We frequently use the $O(\cdot)$ and $\Omega(\cdot)$ notation to express the relationship between functions. Our most common usage will be associated with a sequence $\{x^k\}$ of vectors and a sequence $\{\tau_k\}$ of positive real numbers. In this case $x^k = O(\tau_k)$ means that there is a constant K (dependent on problem data) such that for every $k \in \mathbb{N}$, $\|x^k\| \le K\tau_k$. Similarly, if $x^k > 0$, $x^k = \Omega(\tau_k)$ means that $(x^k)^{-1} = O(1/\tau_k)$. Finally, $x^k \approx \tau_k$ means that $x^k = O(\tau_k)$ and $x^k = \Omega(\tau_k)$.

We use the same notations for a point x in a set parameterized by τ, say $\mathscr{E}_\tau$. We say that $x = O(\tau)$ (resp. $x = \Omega(\tau)$, $x \approx \tau$) whenever there is a constant K such that $\|x\| \le K\tau$ (resp. $x^{-1} = O(1/\tau)$, $x \approx \tau$) for all $x \in \mathscr{E}_\tau$, and all small enough τ. In particular, $x \approx 1$ in $\mathscr{E}_\tau$ means that there are constants $K_2 > K_1 > 0$, such that any $x \in \mathscr{E}_\tau$ satisfies $K_1 \le x_i \le K_2$, $i = 1, \ldots, n$. Given two vector functions x and y, $x \approx y$ means that $x_i \approx y_i$ for $i = 1, \ldots, n$.

2. A generic infeasible path-following algorithm. We define the measure of optimality of $(x, s) \in \mathbb{R}_+^{2n}$ as

$$\mu = \frac{1}{n} x^T s, \tag{2.3}$$

and its measure of feasibility as $\|r\|$ where r is the residual in the linear part of (1.1),

$$r = b - Qx - Rs. \tag{2.4}$$

It is easily seen that the measure of optimality of (x, s) is in fact the normalized 1-norm of the residual of the nonlinear part of (1.1), $n\mu = \|xs\|_1$, so that finding a solution of HLCP (1.1) means finding a pair $(x, s) \in \mathbb{R}_+^{2n}$ with $\mu = 0$ and $r = 0$.

We consider algorithms for solving the HLCP (1.1) that follow approximately the *infeasible central path pinned on a*, defined as the set of triplets (x, s, τ) that satisfy

$$\begin{aligned} xs &= \tau e, \\ Qx + Rs &= b - \tau a, \end{aligned} \tag{2.5}$$

where a is a constant vector related to the starting point

$$w^0 = \left(x^0, s^0, \tau_0\right) \in \mathbb{R}_{++}^{2n+1} \tag{2.6}$$

by

$$a = r^0/\tau_0, \tag{2.7}$$

with r^0 being the residual in the linear part of (1.1) at the starting point,

$$r^0 = b - Qx^0 - Rs^0. \tag{2.8}$$

It is easily seen that if a is defined as above then the second equation of (2.5) is satisfied at the starting point. If we choose arbitrary $s^0 > 0$, $\tau_0 > 0$ and take $x^0 = \tau_0/s^0$ then the first equation of (2.5) is also satisfied, i.e., the starting point chosen in this way belongs to the infeasible central path pinned on a. It is easily seen that for any triplet (x, s, τ) belonging to the infeasible central path we have $\mu = \tau$. Even if the starting point is on an infeasible central path, the subsequent points (x, s, τ) produced by the algorithm will not be on this path and therefore μ and τ will be different in general. Nevertheless, because the second equation in (2.5) is linear, and we consider algorithms based on Newton's method, it follows that if the second equation in (2.5) is satisfied by the starting point, then it will be satisfied by all subsequent points. We will assume that the algorithms under consideration produce

points (x, s, τ) with decreased values of τ (i.e., $\tau \le \tau_0$) and which belong to a certain neighborhood of the infeasible central path. More precisely we assume that the points belong to a "large" neighborhood of the form

$$\mathscr{V}_\nu := \left\{w = (x, s, \tau) \in \mathbb{R}_+^n \times \mathbb{R}_+^n \times \mathbb{R}_+;\, Qx + Rs = b - \tau a,\, \nu\tau e \le xs \le \nu^{-1}\tau e,\, \tau \le \tau_0\right\}$$

where $0 < \nu < 1$ is a given constant, or to a "small" neighborhood

$$\mathscr{N}_\alpha := \left\{w = (x, s, \tau) \in \mathbb{R}_+^n \times \mathbb{R}_+^n \times \mathbb{R}_+;\, Qx + Rs = b - \tau a,\, \left\|\frac{xs}{\tau} - e\right\| \le \alpha,\, \tau \le \tau_0\right\},$$

where $\alpha > 0$ is another given constant. It is easily seen that

$$\mathscr{N}_\alpha \subset \mathscr{V}_\nu \quad \text{for all } 0 < \nu \le 1 - \alpha, \tag{2.9}$$

and

$$\mathscr{V}_\nu \subset \mathscr{N}_\alpha \quad \text{for all } \alpha \ge \sqrt{n}\left(\frac{1}{\nu} - 1\right). \tag{2.10}$$

We also note that

$$\nu\tau \le \mu \le \nu^{-1}\tau \quad \text{for all } (x, s, \tau) \in \mathscr{V}_\nu, \tag{2.11}$$

so that in $\mathscr{V}_\nu$ the parameter τ and the optimality measure μ have the same size, while, as remarked before, on the infeasible central path these two quantities coincide. Moreover

$$Qx + Rs = b - \tau a \quad \text{implies } r = \tau a = (\tau/\tau_0) r^0, \tag{2.12}$$

so that on $\mathscr{V}_\nu$ the feasibility measure $\|r\|$ also has the same size as τ.

At a typical point of an algorithm belonging to the class to be studied in this paper we have already computed a point $(x, s, \tau) \in \mathscr{V}_\nu$, and we want to compute a new point in $\mathscr{V}_\nu$ with a smaller value of τ. Such a point can be obtained as an approximate solution of (2.5) with τ replaced by $\gamma\tau$ where $\gamma \in [0, 1]$ is the desired reduction factor. The associated Newton direction (u, v) is given by the following linear system:

$$\begin{aligned} su + xv &= -xs + \gamma\tau e, \\ Qu + Rv &= \tau(1 - \gamma)a, \end{aligned} \tag{2.13}$$

which under the monotonicity assumption (1.2) has a unique solution. The new point is obtained by taking a steplength $\theta \in (0, 1]$ along this direction, i.e.,

$$x^\# = x + \theta u, \qquad s^\# = x + \theta v. \tag{2.14}$$

It is easily seen that

$$Qx^\# + Rs^\# = b - (1 - \theta + \theta\gamma)\tau. \tag{2.15}$$

If we take by definition

$$\tau_\# := (1 - \theta + \theta\gamma)\tau, \tag{2.16}$$

then $\tau_\# \le \tau \le \tau_0$, so that the new point $(x^\#, s^\#, \tau_\#)$ belongs to $\mathscr{V}_\nu$, provided the "centering parameter" $\gamma \in [0,1]$ and the steplength $\theta \in (0,1]$ are chosen such that

$$\nu\tau_\# e \le x^\# s^\# \le \nu^{-1}\tau_\# e. \tag{2.17}$$

On the other hand if we want the new point $(x^\#, s^\#)$ to belong to $\mathscr{N}_\alpha$ then the steplength $\theta \in (0,1]$ has to be chosen such that

$$\|x^\# s^\# - \tau_\# e\|^2 \le \alpha^2\tau_\#^2. \tag{2.18}$$

We stress the fact that the algorithms we intend to study do not necessarily choose first a parameter γ, then solve the Newton equations (2.13) to obtain direction (u, v), and determine steplength θ such that (2.17) is satisfied. For example, in the largest step path following algorithm to be considered in the next section we solve (2.13) with $\gamma = 0$ to obtain the affine scaling direction (u^a, v^a), and with $\gamma = 1$ to obtain the affine centering direction (u^c, v^c), and then we set $(u, v) = \gamma(u^c, v^c) + (1-\gamma)(u^a, v^a)$ which clearly verifies (2.13) for any γ. The new point is obtained by taking $\theta = 1$ and determining γ such that (2.18) is satisfied. We will see that this involves the solution of a quartic equation in γ. According to (2.9) this ensures that the new point belongs to $\mathscr{V}_\nu$ for all $\nu \le 1 - \alpha$. With this observation we are ready to state the Generic Infeasible Path-Following Algorithm (GIPFA):

ALGORITHM (GIPFA). Data: $\nu, (x^0, s^0, \tau_0) \in \mathscr{V}_\nu$
$k := 0$

REPEAT
$x := x_k,\ s := s_k,\ \tau := \tau_k$;
Compute by some procedure scalars $\gamma \equiv \gamma_k \in [0,1]$, $\theta \equiv \theta_k \in \mathbb{R}_{++}$ and vectors u, v such that (2.13) is satisfied and the point $w^\# = (x^\#, s^\#, \tau_\#)$ defined by (2.14), (2.16) belongs to $\mathscr{V}_\nu$;
$x^{k+1} := x^\#,\ s^{k+1} := s^\#,\ \tau_{k+1} = \tau_\#$;
$k := k + 1$.

This general algorithm includes many infeasible interior point algorithms, such as the predictor-corrector algorithm of Potra and Sheng (1994), where centering steps alternate with affine-scaling steps, and the largest step infeasible path following method that was briefly described above and will be studied in detail in the next section. We will assume that GIPFA produces a sequence $\{\tau_k\}$ that converges R-linearly to zero, in the sense that

$$\limsup_{k\to\infty} (\tau_k)^{1/k} < 1. \tag{2.19}$$

All infeasible interior point algorithms we know of have this property. It is easily seen that if (2.19) is satisfied then any accumulation point of the sequence $\{(x^k, s^k)\}$ produced by GIPFA converges to a solution of (1.1). We will prove in the next theorem that if (1.1) has a solution then the set $\mathscr{V}_\nu$ is bounded and therefore the sequence $\{(x^k, s^k)\}$ has accumulation points. The problem is then under what conditions is there a unique accumulation point; that is, under what conditions does $\{(x^k, s^k)\}$ converge to a solution of our problem? In §4 we will prove that if our problem has a strictly complementary solution then there exists $\bar\lambda > 0$ such that if for all sufficiently large k we have either $w^k \in \mathscr{N}_\alpha$ or $\theta_k\gamma_k \le \bar\lambda$ then $\{(x^k, s^k)\}$ converges to a strictly complementary solution (x^*, s^*) of (1.1). Moreover if $\Sigma_{k=0}^\infty \theta_k\gamma_k = +\infty$, then (x^*, s^*) is the shifted analytic center of the optimal facet. Before ending this section let us introduce some useful notation and prove a theorem showing that the existence of a solution to the HLCP (1.1) implies the boundedness of $\mathscr{V}_\nu$.

We denote the feasible set, and the set of solutions of (1.1) as

$$\mathscr{F} := \{(x, s) \in \mathbb{R}_+^n \times \mathbb{R}_+^n;\ Qx + Rs = b\}, \tag{2.20}$$

and

$$\mathscr{S} := \{(x, s) \in \mathscr{F};\ xs = 0\}. \tag{2.21}$$

Also, we denote the set of strictly complementary solutions by

$$\mathscr{S}^0 := \{(x, s) \in \mathscr{S};\ x + s > 0\}. \tag{2.22}$$

Finally, the set of ϵ-approximate solutions of (1.1) is defined by

$$\mathscr{S}_\epsilon := \{(x, s) \in \mathbb{R}_+^n \times \mathbb{R}_+^n;\ x^Ts \le \epsilon,\ \|Qx + Rs - b\| \le \epsilon\}. \tag{2.23}$$

It is well known that if the HLCP (1.1) represents a linear programming problem, then if $\mathscr{S}$ is nonempty so is $\mathscr{S}^0$. This is not true in general, since it is easy to construct an HLCP with nonempty $\mathscr{S}$ and empty $\mathscr{S}^0$. The existence of a strictly complementary solution will be an essential assumption in §4. However in the following theorem and in the next section we do not make this assumption.

THEOREM 2.1. *If $\mathscr{S}$ is nonempty then for any $(x, s, \tau) \in \mathscr{V}_\nu$ we have*

$$(x^Ts^0 + s^Tx^0) \le (2\nu^{-1} + \zeta)n\tau_0, \tag{2.24}$$

where

$$\zeta = \inf\{((x^0)^T s^* + (s^0)^T x^*)/(n\tau_0) : (x^*, s^*) \in \mathscr{S}\}. \tag{2.25}$$

PROOF. For any $(x^*, s^*) \in \mathscr{S}$ and any $(x, s, \tau) \in \mathscr{V}_\nu$ we can write

$$Q(\tau_0 x - \tau x^0 - (\tau_0 - \tau)x^*) + R(\tau_0 s - \tau s^0 - (\tau_0 - \tau)s^*) = 0,$$

and since Q, R is a positive semidefinite pair (see (1.2)) we have

$$(\tau_0 x - \tau x^0 - (\tau_0 - \tau)x^*)^T(\tau_0 s - \tau s^0 - (\tau_0 - \tau)s^*) \ge 0,$$

wherefrom, by using the fact that $(x, s, \tau) \in \mathscr{V}_\nu$ implies $x^Ts \le \nu^{-1}n\tau$ and $\tau \le \tau_0$, we deduce

$$\begin{aligned}\tau\tau_0(x^Ts^0 + s^Tx^0) &\le \tau\tau_0(x^Ts^0 + s^Tx^0) + \tau_0(\tau_0 - \tau)(x^{*T}s + s^{*T}x)\\ &\le \tau_0^2 x^Ts + \tau^2 x^{0T}s^0 + \tau(\tau_0 - \tau)(x^{*T}s^0 + s^{*T}x^0)\\ &\le n\tau_0^2\nu^{-1}\tau + n\tau^2\tau_0\nu^{-1} + \tau\tau_0(x^{*T}s^0 + s^{*T}x^0).\end{aligned}$$

Dividing by $\tau\tau_0$, we obtain the desired inequality. □

3. The largest step infeasible path following algorithm. As mentioned in §2 the Largest Step Infeasible Path Following Algorithm (LSIPFA) can be obtained from GIPFA by assuming the starting point to belong to $\mathscr{N}_\alpha$, taking $\theta = 1$, and γ the smallest value such that the new point is still in $\mathcal{N}_\alpha$. This smallest value is computed in the algorithm as the root of a quartic. In order to define the quartic, one has to find the solution of (2.13) by first solving the following two linear systems:

$$su^a + xv^a = -xs, \quad Qu^a + Rv^a = r \tag{3.1}$$

$$su^c + xv^c = -xs + \tau e, \quad Qu^c + Rv^c = 0. \tag{3.2}$$

The parameter γ has to be determined such that

$$\|x(\gamma)s(\gamma) - \gamma\tau e\| \le \alpha\gamma\tau, \tag{3.3}$$

where

$$x(\gamma) = x + u(\gamma), \qquad s(\gamma) = s + v(\gamma), \tag{3.4}$$

and

$$u(\gamma) = \gamma u^c + (1-\gamma)u^a, \quad v(\gamma) = \gamma v^c + (1-\gamma)v^a. \tag{3.5}$$

It is easily seen that

$$x(\gamma)s(\gamma) = \gamma\tau e + u(\gamma)v(\gamma). \tag{3.6}$$

Since

$$u(\gamma)v(\gamma) = \gamma^2 u^c v^c + \gamma(1-\gamma)(u^a v^c + u^c v^a) + (1-\gamma)^2 u^a v^a, \tag{3.7}$$

the function

$$\phi(\gamma) := \|\tau^{-1}u(\gamma)v(\gamma)\|^2 - \alpha^2\gamma^2 \tag{3.8}$$

is a quartic. Condition (3.3) is equivalent to $\phi(\gamma) \le 0$. We will show that $\phi(0) \ge 0$ and $\phi(1) < 0$ so that the quartic $\phi(\gamma)$ has either one or three zeros in the half-open interval $[0, 1[$. We denote by $\bar{\gamma}$ the largest one. It follows that $\bar{\gamma}$ is uniquely defined by

$$\phi(\bar{\gamma}) = 0, \qquad \phi(\gamma) < 0 \quad \text{for all } \gamma \in]\bar{\gamma}, 1]. \tag{3.9}$$

In what follows we assume that $\bar{\gamma}$ is computed exactly. However, the whole theory can be rewritten in terms in appropriately chosen upper bounds of $\bar{\gamma}$ that are computable in $O(1)$ arithmetic operations (see also Gonzaga 1994).

By a continuity argument (3.9) ensures that $(x(\gamma), s(\gamma)) \in \mathbf{R}^{2n}_{++}$, $\forall\gamma \in]\bar{\gamma}, 1]$, so that $(x(\gamma), s(\gamma), \gamma\tau) \in \mathcal{N}_\alpha$, $\forall\gamma \in]\bar{\gamma}, 1]$. Therefore by defining

$$x^\# := x + u(\bar{\gamma}), \qquad s^\# := s + v(\bar{\gamma}), \qquad \tau_\# = := \bar{\gamma}\tau, \tag{3.10}$$

we have

$$(x^\#, s^\#, \tau_\#) \in \mathcal{N}_\alpha. \tag{3.11}$$

The complete algorithm can be written as follows:

ALGORITHM 3.1. Data: $0 < \alpha \le .25\epsilon > 0$, $(x^0, s^0, \tau_0) \in \mathcal{N}_\alpha$, where

$$|\mu_0/\tau_0 - 1| \le \beta - 1, \qquad \mu_0 = \frac{1}{n}(x^0)^T s^0, \qquad \beta = 1 + \alpha/\sqrt{n}\,. \tag{3.12}$$

$k := 0$
REPEAT
  $x := x^k$, $s := s^k$, $\tau := \tau_k$;
  Solve the linear systems (3.1) and (3.2);
  Compute $\bar{\gamma} \equiv \gamma_k$ the largest root of the quartic (3.8) from the interval $[0, 1[$;
  $x^{k+1} := x^\#$, $s^{k+1} := s^\#$, $\tau_{k+1} := \tau_\#$, according to (3.10) and (3.5);
  $k := k + 1$

First let us note that for the standard choice of starting points $x^0 = \rho_1 e$, $s^0 = \rho_2 e$ we have $(x^0, s^0, \rho_1\rho_2) \in \mathcal{N}_0 \subset \mathcal{N}_\alpha$. More generally we consider a permutation matrix

$$L = [L_1 L_2], \qquad L_i \in \mathbb{R}^{n\times n_i}, \quad i = 1, 2, n_1 + n_2 = n,$$

and starting points of the form

$$x^0 = \rho_1 L_1 e^1 + \rho_2 L_2 e^2, \qquad s^0 = \rho_2 L_1 e^1 + \rho_1 L_2 e^2, \tag{3.13}$$

where $e^i \in \mathbb{R}^{n_i}$, $i = 1, 2$ are the vectors of all ones of corresponding dimension. Obviously

$$x^0 s^0 = \rho_1\rho_2 e, \qquad \mu_0 = \rho_1\rho_2,$$

so that we have again $(x^0, s^0, \rho_1\rho_2) \in \mathcal{N}_0 \subset \mathcal{N}_\alpha$.

In what follows we will prove that the above algorithm is globally convergent in the sense that the sequence $\{\tau_k\}$ converges to zero at a global Q-linear rate. This implies that the sequence $\{\mu_k\}$ converges to zero at a global R-linear rate. Under some assumptions on the starting points we will derive polynomial complexity bounds. In our proofs we will often use the following two lemmas. The first one contains a simple but fundamental result that can easily be proved, while the second one is more technical and has been proved in Monteiro and Tsuchiya (1992) (see also Bonnans and Gonzaga 1993 and Güler 1993):

PROPOSITION 3.2. *If Q, R is a positive semidefinite pair then the matrix $Q - R$ is invertible.*

PROPOSITION 3.3. *If Q, R is a positive semidefinite pair then $Qu + Rv = 0$ implies $u \in \mathscr{R}(R^T)$ and $v \in \mathscr{R}(Q^T)$.*

In the next proposition we show how to reduce a HLCP to a SLCP. A more general result is given by Güler (1993). Our reduction scheme is inspired from Bonnans and Gonzaga (1993).

PROPOSITION 3.4. *Let Q, R be a positive semidefinite pair and denote*

$$\hat{n}_1 = \operatorname{rank}(Q), \qquad \check{n}_1 = \operatorname{rank}(R), \qquad \hat{n}_2 = n - \hat{n}_1, \qquad \check{n}_2 = n - \check{n}_1.$$

Let

$$\hat{L} = \left[\hat{L}_1 \hat{L}_2\right], \quad \check{L} = \left[\check{L}_1 \check{L}_2\right], \quad \hat{L}_i \in \mathbb{R}^{n \times \hat{n}_i}, \quad \check{L}_i \in \mathbb{R}^{n \times \check{n}_i}, \quad i = 1,2$$

be two permutation matrices, such that $\operatorname{rank}(Q\hat{L}_1) = \hat{n}_1$ *and* $\operatorname{rank}(R\check{L}_1) = \check{n}_1$. *Consider the matrices*

$$\hat{Q}_i = Q\hat{L}_i, \quad \hat{R}_i = R\hat{L}_i, \quad \check{Q}_i = Q\check{L}_i, \quad \check{R}_i = R\check{L}_i, \quad i = 1,2$$

and for any vector $y \in \mathbb{R}^n$ *denote*

$$\hat{y}_i = \hat{L}_i^T y, \quad \check{y}_i = \check{L}_i^T y, \quad i = 1,2.$$

Then the matrices

$$\hat{W} = \left[\hat{Q}_1 \hat{R}_2\right], \quad \check{W} = \left[\check{R}_1 \check{Q}_2\right]$$

are invertible and the matrices

$$\hat{M} = -\hat{W}^{-1}\left[\hat{R}_1 \hat{Q}_2\right], \quad \check{M} = -\check{W}\left[\check{Q}_1 \check{R}_2\right]$$

are positive semidefinite. Moreover the following three equations are equivalent:

$$Qu + Rv = g, \tag{3.14}$$

$$\begin{bmatrix} \hat{u}_1 \\ \hat{v}_2 \end{bmatrix} = \hat{M}\begin{bmatrix} \hat{v}_1 \\ \hat{u}_2 \end{bmatrix} + \hat{W}^{-1}g, \tag{3.15}$$

$$\begin{bmatrix} \check{v}_1 \\ \check{u}_2 \end{bmatrix} = \check{M}\begin{bmatrix} \check{u}_1 \\ \check{v}_2 \end{bmatrix} + \check{W}^{-1}g. \tag{3.16}$$

PROOF. Assume that $\hat{W}$ is not invertible. Then there is a vector $z \in \mathbb{R}^n - \{0\}$ such that

$$\hat{Q}_1 \hat{z}_1 + \hat{R}_2 \hat{z}_2 = 0. \tag{3.17}$$

Since $\hat{Q}_1$ is full rank we must have $\hat{z}_2 \neq 0$. The matrices $\hat{Q} = Q\hat{L}$, $\hat{R} = R\hat{L}$ clearly form a positive semidefinite pair and equation (3.17) can be written as

$$\hat{Q}\begin{bmatrix} \hat{z}_1 \\ 0 \end{bmatrix} + \hat{R}\begin{bmatrix} 0 \\ \hat{z}_2 \end{bmatrix} = 0.$$

According to Proposition 3.3 there is a vector $y \in \mathbf{R}^n$ such that

$$0 = \hat{Q}_1^T y, \quad \hat{z}_2 = \hat{Q}_2^T y.$$

Since $\operatorname{rank}(\hat{Q}_1) = \operatorname{rank}(Q)$ there is a matrix M_{21} such that $\hat{Q}_2 = \hat{Q}_1 M_{21}$ and therefore we obtain the contradiction

$$0 \neq \hat{z}_2 = \hat{Q}_2^T y = M_{21}^T \hat{Q}_1^T y = 0.$$

Hence $\hat{W}$ is invertible. Then the equivalence between (3.14) and (3.15) is easily verified. Taking $g = 0$ in (3.14) and (3.15) we deduce that for any $h \in \mathbb{R}^n$ there are vectors $u, v \in \mathbb{R}^n$ such that

$$h = \begin{bmatrix} \hat{v}_1 \\ \hat{u}_2 \end{bmatrix}, \quad \hat{M}h = \begin{bmatrix} \hat{u}_1 \\ \hat{v}_2 \end{bmatrix}, \quad Qu + Rv = 0.$$

It follows that

$$h^T \hat{M} h = u^T v \geq 0.$$

Hence $\hat{M}$ is positive semidefinite. Thus all the claims related to permutation $\hat{L}$ are proved. The corresponding claims for $\check{L}$ are proved similarly. □

LEMMA 3.5. *If Q, R is a positive semidefinite pair of $n \times n$-matrices and x, s, f, g are four n-dimensional vectors with $x > 0$ and $s > 0$, then the linear system*

$$su + xv = f, \tag{3.18}$$

$$Qu + Rv = g, \tag{3.19}$$

has a unique solution (u, v) and the following relations are satisfied

$$\|Yu\|^2 + \|Y^{-1}v\|^2 \leq \xi^2 + 2\vartheta(\vartheta + \xi), \tag{3.20}$$

$$\|Uv\|^2 \leq \tfrac{1}{8}\xi^4 + \tfrac{1}{2}\vartheta(\xi + \vartheta)(\xi^2 + 2\vartheta(\xi + \vartheta)), \tag{3.21}$$

where

$$Y = X^{-1/2} S^{1/2}, \quad \xi = \left\|(XS)^{-1/2} f\right\|, \quad \vartheta = \min\left\{\left\|\hat{y}(\hat{W}^{-1}g)\right\|, \left\|\check{y}(\check{W}^{-1}g)\right\|\right\},$$

the matrices $\hat{W}, \check{W}$ are defined in Proposition 3.4 and

$$\hat{y} = \begin{bmatrix} \hat{x}_1^{-1/2} \hat{s}_1^{1/2} \\ \hat{x}_2^{1/2} \hat{s}_2^{-1/2} \end{bmatrix}, \quad \check{y} = \begin{bmatrix} \check{x}_1^{-1/2} \check{s}_1^{1/2} \\ \check{x}_2^{1/2} \check{s}_2^{-1/2} \end{bmatrix}.$$

PROOF. Apply Proposition 3.4 and Corollary 3.8 of Potra (1994). □

We are ready to prove now that the point $\bar{\gamma}$ from (3.9) exists and is unique, so that Algorithm 3.1 is well defined.

LEMMA 3.6. *The quartic (3.8) has at least one zero in the interval $[0, 1[$, and its largest zero $\bar{\gamma}$ in $[0, 1]$ satisfies the inequality*

$$0 \leq \bar{\gamma} \leq \hat{\gamma} < 1, \tag{3.22}$$

where $\hat{\gamma}$ is the unique zero in $]0, 1[$ of the quadratic

$$\psi(\gamma) := \gamma^2 \alpha_{cc} + \gamma(1 - \gamma)\alpha_{ac} + (1 - \gamma)^2 \alpha_{aa} - \gamma\alpha, \tag{3.23}$$

with coefficients

$$\alpha_{cc} = \|u^c v^c\|/\tau, \quad \alpha_{ac} = \|u^c v^a + v^a v^c\|/\tau, \quad \alpha_{aa} = \|u^a v^a\|/\tau. \tag{3.24}$$

PROOF. First, let us prove that $\phi(0) \geq 0$ and $\phi(1) < 0$ so that (3.9) indeed defines a unique number $\bar{\gamma} \in [0,1]$. From $(x, s) \in \mathcal{N}_\alpha$ and Lemma 3.5 it follows that:

$$\|u^c v^c\| \leq \frac{1}{\sqrt{8}} \|(xs)^{-1}\|_\infty \|xs - \tau e\|^2$$

$$\leq \frac{\alpha^2 \tau^2}{\sqrt{8}(1-\alpha)\tau} = \frac{\alpha^2 \tau}{\sqrt{8}(1-\alpha)}.$$

Using the fact that $\alpha \leq .25$ implies $\alpha/(1-\alpha) \leq 1/3$ we obtain

$$\alpha_{cc} \leq \frac{\alpha}{3\sqrt{8}}. \tag{3.25}$$

Therefore

$$\phi(1) \leq \frac{\alpha^2}{72} - \alpha^2 < 0,$$

and

$$\phi(0) = \|u^a v^a\|^2/\tau^2 \geq 0.$$

The quartic $\phi(\gamma)$ has either one or three zeros in $[0,1[$. Therefore the largest root $\bar{\gamma}$ of $\phi(\gamma) = 0$ in $[0,1]$ exists and is characterized by (3.9). By the triangle inequality and (3.7), we have that $\phi(\gamma) \leq 0$ whenever $\psi(\gamma) \leq 0$.

Finally, it is easily seen that the quadratic function (3.23) satisfies $\psi(0) \geq 0$, $\psi(1) < 0$ so that the equation $\psi(\gamma) = 0$ has a unique root $\hat{\gamma} \in]0,1[$. Also $\psi(\gamma) \leq 0 \Rightarrow \phi(\gamma) \leq 0$ which shows that (3.22) is true, as required. □

We can now prove the following general result about the largest step infeasible path following algorithm.

THEOREM 3.7. *Algorithm* 3.1 *is well defined and for any integer* $k \geq 0$ *we have*:

$$(x^k, s^k, \tau_k) \in \mathcal{N}_\alpha, \tag{3.26}$$

$$\left(1 - \frac{\alpha}{\sqrt{n}}\right)\tau_k \leq \mu_k \leq \left(1 + \frac{\alpha}{\sqrt{n}}\right)\tau_k, \tag{3.27}$$

$$\tau_k = \psi_k \tau_0, \qquad r^k = \psi_k r^0, \tag{3.28}$$

where

$$\psi_{k+1} = \prod_{j=0}^{k} \gamma_j. \tag{3.29}$$

PROOF. The fact that the algorithm is well defined follows from Lemma 3.6. The relation $(x^k, s^k, \tau_k) \in \mathcal{N}_\alpha$ follows from the choice of γ_k. We have by definition $\tau_{k+1} = \gamma_k \tau_k$, and it is easily verified that $r^{k+1} = \gamma_k r^k$. This shows that (3.28) is satisfied. From the definition of the algorithm (see (3.12)) we have $|\mu_0/\tau_0 - 1| \leq \beta - 1$ so that (3.27) is satisfied for $k = 0$. Using the notation of Algorithm 3.1 and equation (3.6) we can write

$$\mu_\# = \tau_\# + u^T v/n.$$

By using (3.8) and the fact that the orthogonal projection of uv onto span(e) is $n^{-1}u^T v e$ we deduce that

$$\alpha \tau_\# \geq \|uv\| \geq \|n^{-1}(u^T v)e\| = \frac{|u^T v|}{\sqrt{n}}.$$

From the last two equations it follows that (3.27) is satisfied for all $k \geq 0$. □

Let us note that in the above theorem we have not supposed that HLCP (1.1) has a solution. In what follows we will prove that if a solution exists then the sequence $\{\tau_k\}$ converges to zero at a global Q-linear rate by showing that the quantity $\hat{\gamma}$ from Lemma 3.6 has an upper bound strictly less than one. We first have to find bounds for the coefficients (3.24).

LEMMA 3.8. *The coefficients* α_{cc}, α_{ac} *and* α_{aa} *defined by* (3.24) *satisfy*:

$$\alpha_{cc} \leq \frac{\alpha}{3\sqrt{8}}, \qquad \alpha_{ac} \leq \sqrt{\frac{\alpha\delta^*\sqrt{8}}{3}} \leq \sqrt{\alpha\delta^*}, \qquad \alpha_{aa} \leq \delta^*, \tag{3.30}$$

where

$$\delta^* = n\beta\sqrt{.125 + \eta(1+\eta)(.5 + \eta(1+\eta))}, \tag{3.31}$$

$$\eta = \sqrt{n}\left(2(1-\alpha)^{-1} + \zeta\right)\sqrt{1/(\beta(1-\alpha))}\,\chi_0, \tag{3.32}$$

with ζ *defined by* (2.25) *and*

$$\chi_0 = \min\left\{ \left\| \begin{bmatrix} \tilde{x}_1^0 \\ \tilde{s}_2^0 \end{bmatrix}^{-1} (\hat{W}^{-1} r^0) \right\|_\infty, \left\| \begin{bmatrix} \tilde{s}_1^0 \\ \tilde{x}_2^0 \end{bmatrix}^{-1} (\check{W}^{-1} r^0) \right\|_\infty \right\}. \tag{3.33}$$

PROOF. The first inequality in (3.30) has already been verified in the proof of Lemma 3.6 (see (3.25)). In order to prove the last inequality in (3.30) we apply Lemma 3.5 with $f = -Xs$ and $g = r = b - Qx - Rs$. Using (3.27) we deduce immediately that

$$\xi = \|(XS)^{1/2} e\| = \sqrt{n\mu} \leq \sqrt{n\beta\tau}. \tag{3.34}$$

In order to obtain a bound for ϑ we first note that with the notation introduced in Proposition 3.4 we can write:

$$\begin{aligned} \|\hat{y}(\hat{W}^{-1} r)\| &= \frac{\tau}{\tau_0}\|\hat{y}(\hat{W}^{-1} r^0)\| \leq \frac{\tau}{\tau_0}\|(xs)^{-1/2}\|_\infty \left\| \begin{bmatrix} \hat{s}_1 \\ \hat{x}_2 \end{bmatrix} (\hat{W}^{-1} r^0) \right\| \\ &\leq \frac{\sqrt{\tau}}{\tau_0\sqrt{1-\alpha}} \left\| \begin{bmatrix} \hat{s}_1 \\ \hat{x}_2 \end{bmatrix} (\hat{W}^{-1} r^0) \right\|. \end{aligned} \tag{3.35}$$

Let us denote

$$\hat{W}^{-1} r^0 = w^0 = \begin{bmatrix} w_1^0 \\ w_2^0 \end{bmatrix}.$$

We have

$$\|\hat{s}_1 w_1^0\| \le \|\hat{s}_1 w_1^0\|_1 \le \left\|(\hat{x}_1^0)^{-1} w_1^0\right\|_\infty \|\hat{s}_1 \hat{x}_1^0\|_1 \le s^T x^0 \left\|(\hat{x}_1^0)^{-1} w_1^0\right\|_\infty.$$

In a similar manner we get

$$\|\hat{x}_2 w_2^0\| \le x^T s^0 \left\|(\hat{s}_2^0)^{-1} w_2^0\right\|_\infty,$$

and consequently we obtain

$$\left\|\begin{bmatrix}\hat{s}_1\\ \hat{x}_2\end{bmatrix}(\hat{W}^{-1}r^0)\right\| \le \|\hat{s}_1 w_1^0\| + \|\hat{x}_2 w_2^0\| \le (s^T x^0 + x^T s^0)\left\|\begin{bmatrix}\hat{x}_1^0\\ \hat{s}_2^0\end{bmatrix}^{-1}(\hat{W}^{-1}r^0)\right\|_\infty.$$

From (3.35) and Theorem 2.1 we deduce that

$$\|\hat{y}(\hat{W}^{-1}r)\| \le \left(2(1-\alpha)^{-1} + \zeta\right)\sqrt{1/(1-\alpha)}\left\|\begin{bmatrix}\hat{x}_1^0\\ \hat{s}_2^0\end{bmatrix}^{-1}(\hat{W}^{-1}r^0)\right\|_\infty n\sqrt{\tau}.$$

We can prove similarly that

$$\|\check{y}(\check{W}^{-1}r)\| \le \left(2(1-\alpha)^{-1} + \zeta\right)\sqrt{1/(1-\alpha)}\left\|\begin{bmatrix}\check{s}_1^0\\ \check{x}_2^0\end{bmatrix}^{-1}(\check{W}^{-1}r^0)\right\|_\infty n\sqrt{\tau},$$

and therefore we get

$$\vartheta \le \eta\sqrt{n}\beta\tau. \tag{3.36}$$

Finally, by substituting (3.34), (3.36) in (3.21) we get $\alpha_{aa} \le \delta^*$ which proves the last equation in (3.30). Let us note that according to Lemma 3.5 the above inequalities imply that

$$\|Yu^a\|^2 + \|Y^{-1}v^a\|^2 \le \sqrt{8}\,\delta^*\tau. \tag{3.37}$$

Similarly we have (see (3.25)):

$$\|Yu^c\|^2 + \|Y^{-1}v^c\|^2 \le \frac{\alpha\tau}{3}. \tag{3.38}$$

Using the above inequalities we can write

$$\begin{aligned}\|u^c v^a + u^a v^c\| &\le \|Yu^c\|\,\|Y^{-1}v^a\| + \|Yu^a\|\,\|Y^{-1}v^a\|\\ &\le \left(\|Yu^c\|^2 + \|Y^{-1}v^c\|^2\right)^{1/2}\left(\|Yu^a\|^2 + \|Y^{-1}v^a\|^2\right)^{1/2}\\ &\le \sqrt{\frac{\alpha\delta^*\sqrt{8}}{3}}\,\tau \le \sqrt{\alpha\delta^*}\,\tau,\end{aligned}$$

wherefrom we deduce the required bound for α_{ac}. □

We can now give an upper bound for the parameters γ_k generated by Algorithm 3.1.

LEMMA 3.9. *If $\mathcal{S}$ is nonempty then, with the notation introduced in Lemma 3.6, we have*

$$0 < \bar{\gamma} \le \hat{\gamma} < \tilde{\gamma} := 1 - \frac{\sqrt{\alpha}}{8\sqrt{\delta^*}}.$$

PROOF. Using (3.30) we have for any $0 < \gamma < 1$,

$$\begin{aligned}\psi(\gamma) &\le \frac{\alpha\gamma^2}{3\sqrt{8}} + \gamma(1-\gamma)\sqrt{\alpha\delta^*} + (1-\gamma)^2\delta^* - \gamma\alpha\\ &= \left(\frac{\gamma}{3\sqrt{8}} - 1\right)\alpha\gamma + \gamma(1-\gamma)\sqrt{\alpha\delta^*} + (1-\gamma)^2\delta^*.\end{aligned}$$

Since $\delta^* \ge n\beta\sqrt{.125} \ge \sqrt{.125} > .35$ we infer

$$\tilde{\gamma} = 1 - \frac{\sqrt{\alpha}}{8\sqrt{\delta^*}} \ge 1 - \frac{\sqrt{.25}}{8\sqrt{.35}} \ge .8.$$

Therefore

$$\psi(\tilde{\gamma}) \le \left(\frac{1}{3\sqrt{8}} - 1\right)\alpha\tilde{\gamma} + \frac{\tilde{\gamma}\alpha}{8} + \frac{\alpha}{64} \le \left(-\frac{3}{5} + \frac{1}{64}\right)\alpha < 0.$$

Because ψ is decreasing on $[0,1]$ and $\psi(\hat{\gamma}) = 0$ we deduce that $\hat{\gamma} < \tilde{\gamma}$, which completes the proof of our lemma. □

From the above lemma it follows that if the problem has a solution, then the sequence $\{\tau_k\}$ converges to zero at a global Q-linear rate less than or equal to $1 - \sqrt{\alpha}/8\sqrt{\delta^*}$, i.e.,

$$\tau_{k+1} \le \left(1 - \frac{\sqrt{\alpha}}{8\sqrt{\delta^*}}\right)\tau_k. \tag{3.39}$$

Using Theorem 3.7, this implies that our algorithm will produce an ϵ approximate solution in a finite number of steps. More precisely we have the following result.

THEOREM 3.10. *If $\mathcal{S}$ is nonempty then Algorithm 3.1 finds a $(x^*, s^*) \in \mathcal{S}_\epsilon$ in at most*

$$K_\epsilon = \left\lceil \frac{|\ln(\epsilon/\epsilon_0)|}{\left|\ln\left(1 - \sqrt{\alpha}/8\sqrt{\delta^*}\right)\right|}\right\rceil,$$

steps, where $\epsilon_0 = \max\{\beta(x^0)^T s^0, \|r^0\|\}$.

We can obtain $O(\sqrt{n}L)$-iteration complexity in case the starting point is almost feasible.

COROLLARY 3.11. *If $\mathcal{S}$ is nonempty and if there is a constant κ_0 independent of n such that*

$$(1+\zeta)\chi_0 \le n^{-1/2}\kappa_0,$$

then

$$K_\epsilon = O\left(\sqrt{n}\,\ln(\epsilon_0/\epsilon)\right).$$

If we take starting points of the form (3.13) with $L = \hat{L}$ or $L = \check{L}$ we obtain polynomiality if the points are "large enough."

COROLLARY 3.12. *If $\mathscr{S}$ is nonempty and the starting points are chosen of the form*

$$(3.40)\quad x^0 = \hat{\rho}_1 \hat{L}_1 e^1 + \hat{\rho}_2 \hat{L}_2 e^2, \qquad s^0 = \hat{\rho}_2 \hat{L}_1 e^1 + \hat{\rho}_1 \hat{L}_2 e^2, \qquad \tau_0 = \hat{\rho}_1 \hat{\rho}_2,$$

with

$$\hat{\rho}_2 \geq n^{-1} \|\hat{x}_2^* + \hat{s}_1^*\|_1,$$

$$\hat{\rho}_1 \geq \max\left\{ \hat{\rho}_2 \|\hat{M}e\|_\infty, \|\hat{W}^{-1} b\|_\infty, n^{-1} \|\hat{x}_1^* + \hat{s}_2^*\|_1 \right\},$$

for some $(x^*, s^*) \in \mathscr{S}$, or of the form

$$(3.41)\quad x^0 = \check{\rho}_2 \check{L}_1 e^1 + \check{\rho}_1 \check{L}_2 e^2, \qquad s^0 = \check{\rho}_1 \check{L}_1 e^1 + \check{\rho}_2 \check{L}_2 e^2, \qquad \tau_0 = \check{\rho}_1 \check{\rho}_2,$$

with

$$\check{\rho}_2 \geq n^{-1} \|\check{x}_1^* + \check{s}_2^*\|_1,$$

$$\check{\rho}_1 \geq \max\left\{ \check{\rho}_2 \|\check{M}e\|_\infty, \|\check{W}^{-1} b\|_\infty, n^{-1} \|\check{x}_2^* + \check{s}_1^*\|_1 \right\},$$

for some $(x^*, s^*) \in \mathscr{S}$, then

$$K_\epsilon = O(n \ln(\epsilon_0 / \epsilon)).$$

PROOF. We give a proof only for the choice (3.40). For this choice we have

$$\zeta \leq n^{-1} \left(\|\hat{x}_1^* + \hat{s}_2^*\|_1 / \hat{\rho}_1 + \|\hat{x}_2^* + \hat{s}_1^*\|_1 / \hat{\rho}_2 \right) \leq 2$$

and

$$\chi_0 \leq \left\| \begin{bmatrix} \hat{x}_1^0 \\ \hat{s}_2^0 \end{bmatrix}^{-1} (\hat{W}^{-1} r^0) \right\|_\infty = \left\| \hat{\rho}_1^{-1} \left(\hat{W}^{-1} b + \hat{\rho}_2 \hat{M} e \right) - e \right\|_\infty \leq 3,$$

and the claim follows from Lemma 3.8 and Theorem 3.10. □

4. General sufficient conditions for convergence of the iterates. This section is rather technical and the reader who is interested only in the main result should go directly to the statement of Theorem 4.13. The proof of the main result is based on a series of technical lemmas, some of which may be of independent interest to specialists.

Large variables and small variables. For the remainder of this paper we will assume that our HLCP (1.1) has a strictly complementary solution. It is well known that in this case there is a unique partition

$$\mathscr{B} \cup \mathscr{N} = \{1, \ldots, n\}, \qquad \mathscr{B} \cap \mathscr{N} = \varnothing,$$

such that for any $(x, s) \in \mathscr{S}^0$ we have $([x]_i > 0,\ [s]_i = 0,\ \forall i \in \mathscr{B})$ and $([x]_i = 0,\ [s]_i > 0,\ \forall i \in \mathscr{N})$. This means that the "small" and "large" variables are invariant for any strictly complementary solution. Let us denote the corresponding partition of Q and R by

$$Q = [Q_B Q_N], \qquad R = [R_B R_N].$$

Also, for any vector $y \in \mathbb{R}^n$ we denote by y_B the vector of components $[y]_i$, $i \in \mathscr{B}$, and by y_N the vector of components $[y]_i$, $i \in \mathscr{N}$. With this notation the constraint $Qx + Rs = b$ can be rewritten as

$$[Q_B R_N] \begin{bmatrix} x_B \\ s_N \end{bmatrix} + [R_B Q_N] \begin{bmatrix} s_B \\ x_N \end{bmatrix} = b.$$

Following Bonnans and Gonzaga (1993) we rename the variables in the following sequence:

$$Q \leftarrow [Q_B R_N], \qquad R \leftarrow [R_B Q_N], \qquad x \leftarrow \begin{bmatrix} x_B \\ s_N \end{bmatrix},$$

$$s \leftarrow \begin{bmatrix} s_B \\ x_N \end{bmatrix}, \qquad N \leftarrow \varnothing, \qquad B \leftarrow \{1, \ldots, n\}.$$

With this reordering, the solution set (also called optimal face) of (1.1) is written simply as

$$(4.1)\qquad \mathscr{S} = \left\{ (x, s) \in \mathbb{R}^{2n} | s = 0,\ Qx = b,\ x \geq 0 \right\}.$$

The relative interior of the optimal face is composed by the set of all strictly complementary solutions,

$$(4.2)\qquad \mathscr{S}^0 = \left\{ (x, s) \in \mathbb{R}^{2n} | s = 0,\ Qx = b,\ x > 0 \right\}.$$

We note that after reordering, Q, R remains a positive semidefinite pair. In addition, the Newton directions as well as the neighborhoods $\mathscr{V}_\nu$ and $\mathscr{N}_\alpha$ of the infeasible central path as defined in §2 are invariant with respect to this transformation. This means that all algorithms based on the Newton step and those neighborhoods of the infeasible central path (in particular GIPFA) are invariant with respect to permutation of variables. Of course, the algorithms never use the knowledge of the optimal partition, which is unknown: the algorithms are always defined in terms of the original problem and we assume that the optimal face is characterized by (4.1) only when analyzing the algorithms. Therefore in the analysis we will always refer to x as the vector of large variables and to s as the vector of small variables. At a solution of our problem the small variables vanish. In the proof we need a pair $(\tilde{x}, \tilde{s})$ such that

$$(4.3)\qquad Q\tilde{x} + R\tilde{s} = a.$$

We mention that such a pair always exists since we can take for example $\tilde{x} = (Q - R)^{-1} a$ and $\tilde{s} = -\tilde{x}$.

LEMMA 4.1. *If $w = (x, s, \tau) \in \mathscr{V}_\nu$, then $x \approx 1$ and $s \approx \tau$.*

PROOF. According to Theorem 2.1 we have $x = O(1)$, and $s = O(1)$ and from the definition of $\mathscr{V}_\nu$ it follows that $xs \approx \tau$. Let $(\bar{x}, \bar{s})$ be a strictly complementary solution, i.e., $\bar{x} > 0$ and $\bar{s} = 0$. Using

$$Qx + Rs + \tau a = b = Q\bar{x},$$

and (4.3) we get

$$Q(x - \bar{x} + \tau\tilde{x}) + R(s + \tau\tilde{s}) = 0.$$

It follows that $0 \le (x - \bar{x} + \tau\tilde{x})^T(s + \tau\tilde{s})$, i.e.,

$$(\bar{x} - \tau\tilde{x})^T s \le x^T(s + \tau\tilde{s}) - \tau(\bar{x} - \tau\tilde{x})^T\tilde{s}.$$

Set $\bar{\alpha} := \min_{1\le i\le n} \bar{x}_i$. Whenever τ is small enough, we have $\bar{x} - \tau\tilde{x} \ge 0.5\bar{\alpha}e$, so that $s \le (2/\bar{\alpha})x^T(s + \tau\tilde{s})$. By using $xs \approx \tau$ and $x = O(1)$ we deduce that $s = O(\tau)$. The conclusion follows by again using $x = O(1)$ and $xs \approx \tau$. □

We want to study the limit-points of sequences $\{(x^k, s^k, \tau_k)\}$ of elements of $\mathscr{V}_\nu$ when $\tau_k \to 0$. They are related to the shifted analytic barrier function of the optimal face which is defined as follows.

$$\pi_{\tilde{s}}(x) := -\sum_{i=1}^{n} \log x_i - \tilde{s}^T x. \tag{4.4}$$

This is to be compared to the standard barrier function

$$\pi(x) := -\sum_{i=1}^{n} \log x_i. \tag{4.5}$$

We note that the shifted analytic barrier function depends on $\tilde{s}$, not only on a. However, for two different choices of $\tilde{s}$, the two corresponding barrier functions differ only by a fixed constant on the set of all x having the same image through Q. This is a consequence of Proposition 3.3. We define the shifted analytic center of $\mathscr{S}$ as the argument of the minimum of $\pi_{\tilde{s}}$ over the relative interior $\mathscr{S}^0$ of $\mathscr{S}$. The existence of such a minimum may be checked using the same arguments as for the analytic center. For future reference, we note that $\pi_{\tilde{s}}$ is an analytical function with gradient

$$\nabla\pi_{\tilde{s}}(x) = -x^{-1} - \tilde{s}. \tag{4.6}$$

In addition, $\pi_{\tilde{s}}$ is strictly convex so that the shifted analytic center of $\mathscr{S}$ is the unique point x^* that satisfies the optimality system

$$\nabla\pi_{\tilde{s}}(x^*) \in \mathscr{R}(Q^T); \qquad Qx^* = b,\ x^* > 0.$$

Let us denote the projected gradient of the shifted analytic barrier function by

$$g(x) = -P_Q(x^{-1} + \tilde{s}), \tag{4.7}$$

where P_Q denotes the orthogonal projection on $\mathscr{N}(Q)$. It follows that the shifted analytic center of $\mathscr{S}$ is the unique point x_a^* such that

$$g(x_a^*) = 0; \qquad Qx_a^* = b,\ x_a^* > 0. \tag{4.8}$$

We note that from Proposition 3.3 it follows that x_a^* depends only on a and not on the particular choice of $\tilde{s}$.

Now we can state and prove a result concerning the asymptotic behavior of sequences of elements of $\mathscr{V}_\nu$.

LEMMA 4.2. *Let $w^k = (x^k, s^k, \tau_k) \in \mathscr{V}_\nu$ for $k = 0, 1, \dots,$ and denote*

$$q^k := \frac{x^k s^k}{\tau_k} - e.$$

If $\tau_k \to 0$ then there is a subsequence $K \subset \mathbb{N}$ such that, for some $(x, q) \in \mathbb{R}^n \times \mathbb{R}^n$,

$$\lim_{k\in K}(x^k, s^k, q^k) = (x, 0, q) \quad \textit{and} \quad -\nabla\pi_{\tilde{s}}(x) + \frac{q}{x} \in \mathscr{R}(Q^T).$$

PROOF. By Theorem 2.1 $\{w^k\}$ is bounded. According to (2.10), the amount q^k is also bounded. Extracting a convergent subsequence if necessary, we may assume that $(x^k, s^k, q^k) \to (x, s, q)$. By virtue of Lemma 4.1, $x > 0$ and $s = 0$; hence $Qx = b$. Using

$$Qx^k + Rs^k + \tau_k a = b,$$

and substituting $b = Qx$ and $a = Q\tilde{x} + R\tilde{s}$, we obtain

$$Q\left(x^k - x + \tau_k\tilde{x}\right) + R\left(s^k + \tau_k\tilde{s}\right) = 0.$$

By Proposition 3.3 we have $s^k + \tau_k\tilde{s} \in \mathscr{R}(Q^T)$, or equivalently,

$$\mathscr{R}(Q^T) \ni \frac{s^k}{\tau_k} + \tilde{s} = (x^k)^{-1} + \frac{q^k}{x^k} + \tilde{s}.$$

Passing to the limit, we get

$$x^{-1} + \tilde{s} + \frac{q}{x} \in \mathscr{R}(Q^T),$$

as desired. □

COROLLARY 4.3. *If $\{(x^k, s^k, \tau_k)\} \subset \mathscr{V}_\nu$ with $\lim_{k\to\infty}\tau_k = 0$ and $\lim_{k\to\infty}(x^k s^k/\tau_k) = e$, then x^k converges toward the shifted analytic center of $\mathscr{S}$.*

Asymptotic study of the Newton step. Now we come back to the study of the Newton step with a generic right-hand side, namely

$$\begin{aligned} su + xv &= f, \\ Qu + Rv &= r. \end{aligned} \tag{4.9}$$

We use the scaling vectors

$$d = \sqrt{\frac{\tau x}{s}}, \qquad \phi = \sqrt{\frac{xs}{\tau}}.$$

According to Lemma 4.1, $d \approx 1$ and $\phi \approx 1$ whenever $w = (x, s, \tau) \in \mathscr{V}_\nu$. We often use the relation $\phi d = x$. The scaled equations are deduced from (4.9) by multiplying the first equation by $d/\tau x = 1/\tau\phi$, which results in

$$d\frac{s}{\tau x}u + \frac{d}{\tau}v = \frac{df}{\tau x} = \frac{f}{\tau}\phi^{-1},$$

and then using the definition of d to obtain,

$$d^{-1}u + d\frac{v}{\tau} = \frac{f}{\tau}\phi^{-1}. \tag{4.10}$$

Defining now the scaled variables and operators

$$\bar{u} := d^{-1}u, \qquad \bar{v} := d\frac{v}{\tau}, \tag{4.11}$$

$$\overline{Q} := QD, \qquad \overline{R} := \tau RD^{-1}, \tag{4.12}$$

we obtain the scaled Newton equations

$$\begin{aligned} \bar{u} + \bar{v} &= \frac{f}{\tau}\phi^{-1}, \\ \overline{Q}\bar{u} + \overline{R}\bar{v} &= r. \end{aligned} \tag{4.13}$$

The scaling used in the present section is advantageous when a strictly complementary solution exists and variables are permuted as in the previous section so that x and s are the big and small variables respectively. This will allow us to represent the solution of the original Newton equation (2.5) as $O(\tau)$ perturbations of quantities that are easily analyzed. In order to do so we represent the solution of the scaled equations (4.13) in terms of orthogonal projections. First we summarize some facts about orthogonal projections (for proofs see Bonnans and Gonzaga 1993).

Given $A \in \mathbb{R}^{m\times n}$, $q \in \mathbb{R}^m$, we define the projection operators $P_{A,q}$ and P_A by

$$x \mapsto P_{A,q}x = \operatorname{argmin}\{\|w - x\| \mid Aw = q\},$$

and $P_A = P_{A,0}$. Since P_A is the orthogonal projection on $\mathscr{N}(A)$, and therefore a linear operator in $\mathbb{R}^n$, the same notation will be used for its matrix representation. Similarly, we denote by $\tilde{P}_A = I - P_A$ the orthogonal projection on $\mathscr{R}(A^T)$ and the matrix representing it.

LEMMA 4.4. *For any $x \in \mathbb{R}^n$ and $q \in \mathbb{R}^m$, $P_{A,q}x = P_A x + P_{A,q}0$.*

LEMMA 4.5. *Let $\mathscr{D} \subset \mathbb{R}^n$ be such that $d \approx 1$ whenever $d \in \mathscr{D}$. Then for $y \in \mathbb{R}^n$, $q \in \mathscr{R}(A)$, $d \in \mathscr{D}$,*

$$P_{AD,q}0 = O(\|q\|),$$

$$P_{AD,q}y = O(\|q\|) + O(\|y\|),$$

$$\|DP_{AD}Dy\| \approx \|P_A y\|.$$

LEMMA 4.6 (GONZAGA AND TAPIA 1992, LEMMA 3.3). *Let $g \in \mathbb{R}^n$ be such that $\|g - e\|_\infty \le \hat{\alpha}$, where $\hat{\alpha} \in (0,1)$. Let $p \in \mathbb{R}^n$, set $G = \operatorname{diag}(g)$, and consider the projections $\hat{h} = P_A p$, $h = gP_{AG}gp$. Then*

$$\|h - \hat{h}\| \le \hat{\alpha}(1 + \hat{\alpha})\frac{2 - \hat{\alpha}}{1 - \hat{\alpha}}\|\hat{h}\|.$$

Now we can prove the following result.

LEMMA 4.7. *Consider the system (4.13). Let $\breve{u}, \breve{v}$ be such that $Q\breve{u} + R\breve{v} = r$. Then*

$$\bar{u} = P_{\overline{Q}}\left(\frac{f}{\tau}\phi^{-1} - d\frac{\breve{v}}{\tau}\right) + P_{\overline{Q},r-Rv}0. \tag{4.14}$$

PROOF. Let us write (4.13) as

$$\bar{u} - \left(\frac{f}{\tau}\phi^{-1} - d\frac{\breve{v}}{\tau}\right) = d\frac{\breve{v}}{\tau} - \bar{v},$$

$$\overline{Q}\bar{u} = r - Rv.$$

From $Q(u - \breve{u}) + R(v - \breve{v}) = 0$ we deduce with Proposition 3.3 that $\breve{v} - v \in \mathscr{R}(Q^T)$; hence $(d/\tau)(v - \breve{v}) = \bar{v} - d(\breve{v}/\tau) \in \mathscr{R}(\overline{Q}^T)$, i.e.,

$$\bar{u} - \left(\frac{f}{\tau}\phi^{-1} - d\frac{\breve{v}}{\tau}\right) \in \mathscr{R}(\overline{Q}^T),$$

$$\overline{Q}\bar{u} = r - Rv,$$

from which the result follows. □

Let us consider now the Newton equation (2.13). The scaled system associated with (2.13) is

$$\begin{aligned} \bar{u} + \bar{v} &= \gamma\phi^{-1} - \phi, \\ \overline{Q}\bar{u} + \overline{R}\bar{v} &= \tau(1 - \gamma)a. \end{aligned} \tag{4.15}$$

We consider again a pair $(\bar{x}, \bar{s})$ satisfying (4.3). Applying the above results, we obtain the following important lemma:

LEMMA 4.8. *If $(x, s, \tau) \in \mathscr{V}_\nu$, then the solutions of the linear systems (2.13) and (4.15) satisfy*

(i) $\bar{v} = \bar{v}(\gamma) = O(1)$, $v = O(\tau)$;
(ii) $u = u(\gamma) = \gamma\, dP_{\overline{Q}}(\phi^{-1} + d\bar{s}) + O(\tau)$;
(iii) $v/\tau = d^{-1}[\gamma\tilde{P}_{\overline{Q}}(\phi^{-1} + d\bar{s}) - \phi - \gamma\, d\bar{s}] + O(\tau)$.

PROOF. Define

$$\hat{u} := \bar{u} - \tau(1 - \gamma)d^{-1}\bar{x},$$

$$\hat{v} := \bar{v} - \tau(1 - \gamma)d\frac{\bar{s}}{\tau}.$$

Then, by using (4.15) we get

$$\hat{u} + \hat{v} = \gamma\phi - \phi^{-1} - \tau(1-\gamma)\Big(d^{-1}\tilde{x} + d\frac{\tilde{s}}{\tau}\Big) = O(1),$$

$$\overline{Q}\hat{u} + \overline{R}\hat{v} = 0.$$

By applying Lemma 3.5 with $X = S = I$, $f = O(1)$ and $r = 0$ it follows that $\hat{u} \approx 1$ and $\hat{v} \approx 1$. Coming back to the definition of $\hat{u}$ and $\hat{v}$, we deduce that $\bar{v} = O(1)$. Also, $v = \tau d^{-1}\hat{v} = O(\tau)$, which proves (i).

Now let us prove that

$$u = \gamma\, dP_{\overline{Q}}(\phi^{-1} + d\tilde{s}) + \tau\, dP_{\overline{Q},a}0 - dP_{\overline{Q},Rv}0. \tag{4.16}$$

We apply Lemma 4.7 with

$$\check{v} = (1-\gamma)\tau\tilde{s}, \qquad f = -xs + \gamma\tau e = \tau\phi(-\phi + \gamma\phi^{-1}).$$

Using the linearity of projections (see Lemma 4.4), and the fact that $r = \tau a$, we get

$$u = dP_{\overline{Q}}(-\phi + \gamma\phi^{-1} - (1-\gamma)\,d\tilde{s}) + dP_{\overline{Q},\tau a - Rv}0$$

$$= dP_{\overline{Q}}(-\phi + \gamma\phi^{-1} - (1-\gamma)\,d\tilde{s}) + \tau\, dP_{\overline{Q},a}0 - dP_{\overline{Q},Rv}0.$$

We obtain (4.16) if we can prove that $P_{\overline{Q}}(\phi + d\tilde{s}) = 0$, i.e.,

$$\phi + d\tilde{s} \in \mathcal{N}(\overline{Q}^T)^{\perp} = \mathcal{R}(\overline{Q}^T). \tag{4.17}$$

As $\phi/d = s/\tau$, this boils down to $s + \tau\tilde{s} \in \mathcal{R}(Q^T)$, which is a consequence of

$$Q(x + \tau\tilde{x}) + R(s + \tau\tilde{s}) = b \in \mathcal{R}(Q)$$

and Proposition 3.3. Hence (4.16) is satisfied.

By Lemma 4.5, $dP_{\overline{Q},a}0 = O(\|a\|) = O(1)$ and $dP_{\overline{Q},Rv}0 = O(\|v\|) = O(\tau)$, so that (ii) holds.

Now, according to (4.15) we have

$$\bar{u} + \bar{v} = \gamma\phi^{-1} - \phi = \gamma(\phi^{-1} + d\tilde{s}) - (\phi + \gamma\, d\tilde{s}),$$

and by using (ii) we get

$$\bar{v} = \gamma\bar{P}_{\overline{Q}}(\phi^{-1} + d\tilde{s}) - (\phi + \gamma\, d\tilde{s}) + O(\tau).$$

As $v/\tau = d^{-1}\bar{v}$, we obtain (iii). □

Let us now prove a very useful result about the new point produced by a step of GIPFA.

Lemma 4.9. *Given $(x, s, \tau) \in \mathscr{V}_v$, let (u, v) be the solution of the Newton equation (2.13) and define the new point $(x^{\#}, s^{\#}, \tau_{\#})$ as in (2.14) and (2.16). Then*

(i) $x^{\#}s^{\#} = (1-\theta)xs + \theta\gamma\tau e + \theta^2 uv$;

(ii) $uv/\gamma\tau = [\gamma\bar{P}_{\overline{Q}}(\phi^{-1} + d\tilde{s}) - \phi - \gamma\, d\tilde{s}]P_{\overline{Q}}(\phi^{-1} + d\tilde{s}) + O(\tau/\gamma)$;

(iii) $u^T v/\gamma\tau = (1-\gamma)\tilde{s}^T dP_{\overline{Q}}(\phi^{-1} + d\tilde{s}) + O(\tau/\gamma)$;

(iv) $$\frac{x^{\#}s^{\#}}{\tau_{\#}} - e = \frac{1-\theta}{1-\theta+\theta\gamma}\Big(\frac{xs}{\tau} - e\Big) + \frac{\theta^2}{1-\theta+\theta\gamma}\frac{uv}{\tau}.$$

Proof. Using (2.13), we obtain

$$(x + \theta u)(s + \theta v) = xs + \theta(xv + su) + \theta^2 uv,$$

$$= (1-\theta)xs + \theta\gamma\tau e + \theta^2 uv.$$

This proves (i). Relation (ii) follows from Lemma 4.8. We deduce that

$$\frac{u^T v}{\gamma\tau} = -(\phi + \gamma\, d\tilde{s})^T P_{\overline{Q}}(\phi^{-1} + d\tilde{s}) + O\Big(\frac{\tau}{\gamma}\Big).$$

Using (4.17), we get (iii). Relation (iv) is an easy consequence of (i). □

The shifted analytic barrier function associated with large variables. The direction of displacement for large variables is, up to $O(\tau)$, equal to γl, with

$$l := dP_{\overline{Q}}(\phi^{-1} + d\tilde{s}). \tag{4.18}$$

This vector l is, up to $O(\tau)$, equal to the centering step u^c (see (3.2)). If the variables are well centered, then u^c is close to 0. It is easily seen that $l = 0$ if and only if

$$\phi^{-1} + d\tilde{s} \in \mathcal{R}(\overline{Q}^T) = \mathcal{R}(DQ^T).$$

Since $d\phi = x$, this is equivalent to

$$\mathcal{R}(Q^T) \ni d^{-1}(\phi^{-1} + d\tilde{s}) = x^{-1} + \tilde{s}.$$

The last expression is opposite to the gradient of the shifted barrier function $\pi_{\tilde{s}}$ (see (4.6)). Therefore $l = 0$ if and only if $g(x) = 0$, where g is the projected gradient of the shifted analytic barrier function given by (4.7).

We now interpret l as a perturbed Newton step for minimizing the shifted barrier function over some affine space. Consider the nonlinear problem parameterized by x:

$$\min_{y \in \mathbb{R}^n} \pi_{\tilde{s}}(y);\ Qy = Qx. \tag{4.19}$$

As the Hessian of $\pi_{\tilde{s}}$ at x is X^{-2}, the Newton step associated with this minimization problem is defined by

$$\min_{\hat{l} \in \mathbb{R}^n} \frac{1}{2}\hat{l}^T X^{-2}\hat{l} - \hat{l}^T(x^{-1} + \tilde{s}); \qquad Q\hat{l} = 0, \tag{4.20}$$

whose solution is characterized by

$$X^{-2}\hat{l} \in x^{-1} + \tilde{s} + \mathcal{R}(Q^T); \qquad Q\hat{l} = 0.$$

Writing the equivalent system

$$X^{-1}\hat{l} \in x(x^{-1} + \tilde{s}) + \mathcal{R}((QX)^T); \qquad QX(X^{-1}\hat{l}) = 0,$$

we obtain the following expression for the Newton step,

$$\hat{l} = xP_{QX}(x(x^{-1} + \tilde{s})) = xP_{QX}(e + x\tilde{s}). \tag{4.21}$$

If we change the Hessian of problem (4.20) to D^{-2}, we obtain the modified problem

$$\min_{l \in \mathbb{R}^n} \frac{1}{2} l^T D^{-2} l - l^T(x^{-1} + \tilde{s}); \qquad Ql = 0,$$

whose optimality system is equivalent to

$$D^{-1} l \in d(x^{-1} + \tilde{s}) + \mathscr{R}\big((QD)^T\big); \qquad QD(D^{-1}l) = 0.$$

The unique primal solution of the latter system is given by (4.18). In short, l is the displacement obtained by applying a Newton-like step for minimizing $\pi_{\tilde{s}}(x)$ with the Hessian being approximated by D^{-2}.

The distance between the Newton-like step l and the exact Newton step $\hat{l}$ may be estimated using Lemma 4.6. Let us define the proximity of the large variables $x > 0$ as the norm of $x^{-1}\hat{l}$ (the scaled Newton centering step for solving (4.19)):

$$\delta(x) := \| P_{QX}(e + x\tilde{s}) \|.$$

The next lemma shows that the proximity is never larger than the usual measure of centering.

LEMMA 4.10. *If* $(x, s) \in \mathscr{V}_\nu$, *then* $\delta(x) \le \|xs/\tau - e\|$.

PROOF. Because P_{QX} is the orthogonal projection onto $\mathscr{N}(QX)$ we have

$$\| P_{QX}(e + x\tilde{s}) \| = \min\{\|e + x\tilde{s} - z\| \,|\, z \in \mathscr{R}(XQ^T)\}.$$

By (4.17), $\mathscr{R}(XQ^T) \ni \phi(\phi + d\tilde{s}) = sx/\tau + x\tilde{s}$. Hence, from the relation above, $\|P_{QX}(e + x\tilde{s})\| \le \|e - xs/\tau\|$. □

Now we can prove the following two useful lemmas.

LEMMA 4.11. *Let l and $\hat{l}$ be given by* (4.18) *and* (4.21), *respectively. If* $(x, s) \in \mathscr{N}_\alpha$, *with* $\alpha \le 0.25$, *then*

$$\pi_{\tilde{s}}(x + l) \le \pi_{\tilde{s}}(x) - 0.06\|\hat{l}\|^2.$$

PROOF. The method of proof consists in comparing l and the Newton direction $\hat{l}$. We have

$$\pi_{\tilde{s}}(x + l) - \pi_{\tilde{s}}(x) = \pi(x + l) - \pi(x) - \tilde{s}^T l = \pi(e + x^{-1}l) - \tilde{s}^T l.$$

If we set $h := x^{-1}l$, $\hat{h} := x^{-1}\hat{l}$, then

$$\pi(x + l) - \pi(x) - \tilde{s}^T l = \pi(e + h) - \tilde{s}^T l,$$

where π is the (standard) barrier function (4.5). The quadratic approximation of the barrier function gives us the following property (shown for instance in Gonzaga 1989): if $\|h\| < 1$, then

$$\pi(e + h) \le -e^T h + \frac{\|h\|^2}{2} + \frac{1}{3}\frac{\|h\|^3}{1 - \|h\|},$$

and consequently,

$$\pi_{\tilde{s}}(x + l) - \pi_{\tilde{s}}(x) \le -(e + x\tilde{s})^T h + \frac{\|h\|^2}{2} + \frac{1}{3}\frac{\|h\|^3}{1 - \|h\|}. \tag{4.22}$$

We want to apply Lemma 4.6 to $h := x^{-1}l$, $\hat{h} := x^{-1}\hat{l}$, $g := dx^{-1} = \phi^{-1}$, $p = e + x\tilde{s}$, and $A := QX$. As $(x, s) \in \mathscr{N}_\alpha$, $\|\phi^2 - e\| \le 0.25$ and consequently $1.25e \ge \phi^2 \ge 0.75e$. It follows that $-.106e < \phi^{-1} - e < .155e$. Finally, by taking $\hat{\alpha} = .155$ in Lemma 4.6 we obtain $\|h - \hat{h}\| \le 0.4\|\hat{h}\|$. By Lemma 4.10, $\|\hat{h}\| = \delta(x) \le 0.25$, so that

$$\|h\| \le \|\hat{h}\| + \|q\| \le 1.4\|\hat{h}\| \le 0.35,$$

where we have denoted $q := h - \hat{h}$. Since $\|h\| < 1$ we can use (4.22). We have $(e + x\tilde{s})^T \hat{h} = (e + x\tilde{s})^T P_{QX}(e + x\tilde{s}) = \|\hat{h}\|^2$, so that we obtain

$$\pi_{\tilde{s}}(x + l) - \pi_{\tilde{s}}(x) \le -\|\hat{h}\|^2 - (e + x\tilde{s})^T q + \frac{\|\hat{h}\|^2 + \|q\|^2 + 2\hat{h}^T q}{2} + \frac{1}{3}\frac{\|h\|^3}{1 - \|h\|}.$$

Let us prove that $q \in \mathscr{N}(QX)$. As $q = h - \hat{h}$ and $\hat{h} \in \mathscr{N}(QX)$, we have to show that $h \in \mathscr{N}(QX)$. This follows immediately by writing

$$QXh = QX\phi^{-1}P_{\overline{Q}}(\phi^{-1} + d\tilde{s}) = \overline{Q}P_{\overline{Q}}(\phi^{-1} + d\tilde{s}) = 0. \tag{4.23}$$

Since $\hat{h} = P_{QX}(e + x\tilde{s})$, we have $\hat{h}^T q = (e + x\tilde{s})^T q$, and by using $\|q\| \le 0.4\|\hat{h}\|$ we deduce

$$\pi_{\tilde{s}}(x + l) - \pi_{\tilde{s}}(x) \le -\frac{\|\hat{h}\|^2}{2} + \frac{\|q\|^2}{2} + \frac{1}{3}\frac{\|h\|^3}{1 - \|h\|},$$

$$\le \left(-\frac{1}{2} + \frac{0.16}{2} + \frac{1}{3}\frac{\|h\|}{1 - \|h\|}\frac{\|h\|^2}{\|\hat{h}\|^2}\right)\|\hat{h}\|^2.$$

Using again $\|\hat{h}\| \le 0.25$ and $\|h\| \le 1.4\|\hat{h}\| \le 0.35$, we obtain the conclusion. □

LEMMA 4.12. *If* $w = (x, s, \tau) \in \mathscr{V}_\nu$, *then there exists* $\bar{\rho} > 0$, $\Delta > 0$ *such that*

$$\pi_{\tilde{s}}(x + \rho l) \le \pi_{\tilde{s}}(x) - \rho\Delta\|g(x)\|^2, \quad \textit{for all } 0 \le \rho < \bar{\rho}. \tag{4.24}$$

PROOF. In $\mathscr{V}_\nu$, we have $x \approx 1$. Since the Hessian of $\pi_{\tilde{s}}$ is equal to X^{-2} it follows that there are constants $M > 0$ and $\rho_1 > 0$ such that

$$\pi_{\tilde{s}}(x + \rho l) - \pi_{\tilde{s}}(x) \le -\rho l^T(x^{-1} + \tilde{s}) + \frac{\rho^2}{2}M\|l\|^2, \quad \text{for all } 0 < \rho \le \rho_1.$$

Also,

$$l = DP_{QD}D(x^{-1} + \tilde{s}), \qquad \hat{l} = XP_{QX}X(x^{-1} + \tilde{s}), \qquad g(x) = P_Q(x^{-1} + \tilde{s}),$$

so that we can apply Lemma 4.5 to deduce

$$l \approx \hat{l} \approx g(x). \tag{4.25}$$

From the definition of an orthogonal projection we have $d^{-1}l - d(x^{-1} + \tilde{s}) \in \mathscr{R}(DQ^T)$. Hence $d^{-2}l - (x^{-1} + \tilde{s}) \in \mathscr{R}(Q^T)$. From (4.23) it follows that $l = xh \in \mathscr{N}(Q)$, and therefore we can write

$$l^T(x^{-1} + \tilde{s}) = l^T D^{-2} l.$$

On the other hand, since $d \approx 1$ and $l \approx g(x)$, there exists $\Delta_1 > 0$ and $M_1 > 0$ such that

$$-l^T D^{-2} l \le -\Delta_1 \| g(x) \|^2, \qquad M\|l\|^2 \le M_1 \| g(x) \|^2.$$

If we define

$$\bar{\rho} := \frac{\Delta_1}{M_1}, \qquad \Delta := \frac{\Delta_1}{2M_1},$$

then we can write

$$\pi_{\tilde{s}}(x + \rho l) - \pi_{\tilde{s}}(x) \le -\rho\Delta_1 \| g(x) \|^2 + \frac{\rho^2}{2} M_1 \| g(x) \|^2$$

$$= -\rho\left(\Delta_1 - \rho\frac{M_1}{2}\right) \| g(x) \|^2$$

$$\le -\rho\Delta \| g(x) \|^2,$$

for any $0 \le \rho \le \bar{\rho}$, as was to be proved. □

We end this section by stating and proving our main result.

THEOREM 4.13. *Suppose that the linear complementarity problem* (1.1) *has a strictly complementary solution* (*i.e.*, $\mathscr{S}^0$ *is not empty*) *and that the sequence* $\{\tau_k\}$ *produced by Algorithm GIPFA converges R-linearly to zero. If one of the following two conditions are satisfied*:

(i) $w^k \in \mathscr{N}_\alpha$ *for all sufficiently large* k,

(ii) $\lim_{k\to\infty} \theta_k \gamma_k = 0$,

then $\{(x^k, s^k)\}$ *converges to a strictly complementary solution* $(x^*, s^*) \in \mathscr{S}^0$ *of* (1.1). *Moreover if* $\sum_{k=0}^{\infty} \theta_k \gamma_k = +\infty$ *then* (x^*, s^*) *is the shifted analytic center* (x_a^*, s_a^*) *of the optimal face.*

PROOF. Set

$$l^k := d^k P_{QD^k}\left((\phi^k)^{-1} + d^k \tilde{s}\right).$$

By Lemma 4.8, we have

$$u^k = \theta_k \gamma_k l^k + O(\tau_k).$$

As l^k is bounded in $\mathscr{V}_\nu$, and τ_k converges linearly to 0, we deduce that x^k converges whenever $\sum_{k=1}^{\infty} \theta_k \gamma_k < +\infty$. It remains to prove that if $\sum_{k=1}^{\infty} \theta_k \gamma_k = +\infty$, then x^k converges to the shifted analytic center of the optimal face. By Lemmas 4.11 and 4.12, the fact that $l^k \approx g(x^k)$, and the hypothesis of our theorem we deduce that there is a constant $\tilde{\Delta}$ such that

$$\pi_{\tilde{s}}(x^k + \theta_k \gamma_k l^k) \le \pi_{\tilde{s}}(x^k) - \theta_k \gamma_k \tilde{\Delta} \| g(x^k) \|^2$$

for all sufficiently large k. The shifted analytic barrier $\pi_{\tilde{s}}$ is Lipschitz continuous in $\mathscr{V}_\nu$, so that there exists a constant $\tilde{M}$ such that

$$\pi_{\tilde{s}}(x^{k+1}) \le \pi_{\tilde{s}}(x^k) - \theta_k \gamma_k \tilde{\Delta} \| g(x^k) \|^2 + \tilde{M}\tau_k. \tag{4.26}$$

Since τ_k converges R-linearly, and the shifted barrier function is bounded from below on the compact set $\mathscr{V}_\nu$, we obtain

$$\sum_{k=1}^{\infty} \theta_k \gamma_k \| g(x^k) \|^2 < \infty.$$

Since $\sum_{k=1}^{\infty} \theta_k \gamma_k = \infty$ there is a subsequence $\{x^k\}_{k\in K}$ such that $\lim_{k\in K} \|g(x^k)\|^2 = 0$. The set $\mathscr{V}_\nu$ is compact and therefore we may assume without loss of generality that $\{x^k\}_{k\in K}$ is convergent. Its limit clearly verifies (4.8), so that we can write

$$\lim_{k\in K} x^k = x_a^*.$$

Using (2.19) and $\lim_{k\in K} \pi_{\tilde{s}}(x^k) = \pi_{\tilde{s}}(x_a^*)$ we deduce that for any $\epsilon > 0$, there exists $k' \in \mathbb{N}$ such that

$$\pi_{\tilde{s}}(x^{k'}) < \pi_{\tilde{s}}(x_a^*) + \epsilon \quad \text{and} \quad \sum_{k=k'}^{\infty} \tau_k < \epsilon.$$

By virtue of (4.26) we get

$$\pi_{\tilde{s}}(x^{k'+j}) \le \pi_{\tilde{s}}(x^{k'}) + \tilde{M} \sum_{k=k'}^{\infty} \tau_k \le \pi_{\tilde{s}}(x_a^*) + (\tilde{M} + 1)\epsilon, \qquad \forall j \in \mathbb{N},$$

i.e., $\limsup\{\pi_{\tilde{s}}(x^k)\} \le \pi_{\tilde{s}}(x_a^*)$. On the other hand, any limit-point x^* of $\{x^k\}$ satisfies $x^* > 0$ and $\pi_{\tilde{s}}(x^* \ge \pi_{\tilde{s}}(x_a^*)$. It follows that $\pi_{\tilde{s}}(x^k) \to \pi_{\tilde{s}}(x_a^*)$, which by the compactness of $\mathscr{V}_\nu$ implies the convergence of x^k towards x_a^*. □

Let us note that condition (ii) of the above theorem could be replaced with the formally more general condition

$$\theta_k \gamma_k \le \bar{\rho} \quad \text{for all sufficiently large } k,$$

where $\bar{\rho}$ is the quantity considered in Lemma 4.12. However, the latter condition would be very difficult to verify since there is no simple way to find bounds for $\bar{\rho}$. The largest step infeasible path following algorithm studied in the previous section clearly satisfies condition (i), and in the next section we will show that it also satisfies (ii) since $\theta_k = 1$ and $\lim_{k\to\infty} \gamma_k = 0$. We note that if the largest step infeasible path following algorithm is considered in connection to larger neighborhoods of the infeasible central path then (ii) may still be satisfied while (i) is not. We also note that the algorithms of Wright (1993, 1994, 1996) as well as the algorithms of Potra and Sheng (1995) satisfy (ii) since $\gamma_k = 0$ for k sufficiently large, but they do not satisfy (i).

5. Asymptotic convergence analysis for the largest step infeasible path following algorithm. In what follows we first use Theorem 4.13 to deduce that the sequence $\{w^k\}$ produced by the largest step infeasible path following algorithm converges. Then it is easy to prove that the sequences $\{\|r^k\|\}$ and $\{(x^k)^T s^k\}$ measuring feasibility and optimality are superlinearly convergent to zero. We shall end the paper by showing that the largest step infeasible path following algorithm can be modified in such a way that the latter sequences are quadratically convergent.

THEOREM 5.1. *Assume that the set of strictly complementary solutions $\mathcal{S}^0$ is nonempty and that $\alpha \le 0.25$. Then the sequence $\{(x^k, s^k, \tau_k)\}$ produced by Algorithm 3.1 converges. Moreover the sequence $\{\tau_k\}$ converges to zero Q-superlinearly, and consequently the sequences $\{\|r^k\|\}$ and $\{(x^k)^T s^k\}$ also converge Q-superlinearly to zero.*

PROOF. We assume that the variables have been permuted such that x contains the large variables and s contains the small variables. From Lemma 4.1 we have $s^k \approx \tau_k$ and therefore according to (3.39) it follows that the small variables converge to zero, $s^k \to 0$. We know by Theorem 4.13 that $\{x^k\}$ also converges, $x^k \to x^*$. Hence $u^k \to 0$. It follows that

$$\alpha = \left\| \frac{x^{k+1}s^{k+1}}{\gamma_k \tau_k} - e \right\| = \left\| \frac{u^k v^k}{\gamma_k \tau_k} \right\| = o\left(\frac{v^k}{\gamma_k \tau_k} \right).$$

As $v^k \approx \tau_k$, we deduce that $\alpha = o(1/\gamma_k)$, which implies $\gamma_k \to 0$. Since $\gamma_k = \tau_{k+1}/\tau_k$, this means that $\tau_k \to 0$ Q-superlinearly. From (3.28) it follows that r^k converges Q-superlinearly to zero as well (even component-wise!). Finally, from (3.27) we deduce that

$$\frac{\mu_{k+1}}{\mu_k} \le \left[\left(1 + \frac{\alpha}{\sqrt{n}}\right) \Big/ \left(1 - \frac{\alpha}{\sqrt{n}}\right) \right] \frac{\tau_{k+1}}{\tau_k},$$

which shows that $\{(x^k)^T s^k\}$ also converges to zero Q-superlinearly. □

We now consider a variant of this algorithm in which a safeguard is added so as to obtain quadratic convergence. Our basic tool is Lemma 5.4, due (in the framework of feasible algorithms) to Gonzaga (1994). This lemma gives a precise estimate of $\|xs(1)/\tau - e\|$, which is itself strongly related to the centering of the large variable.

We define $\delta^\#(\gamma)$ as the proximity obtained at the point $(x(\gamma), s(\gamma), \gamma\tau)$, i.e.,

$$\delta^\#(\gamma) := \left\| \frac{x(\gamma)s(\gamma)}{\gamma\tau} - e \right\|.$$

ALGORITHM 5.2. Data: $0 < \alpha \le .25$, $\epsilon > 0$, $\epsilon' \in (0, 1 - 1/2\sqrt{n})$, $(x^0, s^0, \tau_0) \in \mathcal{N}_\alpha$, where

$$\tau_0 = \mu_0/\beta = (x^0)^T s^0/(n\beta), \qquad \beta = 1 + \alpha/\sqrt{n}. \tag{5.27}$$

$k := 0$
REPEAT
$x := x^k$, $s := s^k$, $\tau := \tau_k$;
Solve the linear systems (3.1) and (3.2);
Compute $\bar{\gamma}$ the largest root of the quartic (3.8) from the interval [0, 1[;
Safeguard. If $\delta^\#(0.1) \le 0.2$, then set $\gamma_k := \bar{\gamma}$, else $\gamma_k := \max(\epsilon', \bar{\gamma})$;
$x^{k+1} := x^\#$, $s^{k+1} := s^\#$, $\tau_{k+1} := \tau_\#$, according to (3.10) and (3.5);
$k := k + 1$.

THEOREM 5.3. *Assume that the set of strictly complementary solutions $\mathcal{S}^0$ is nonempty and that $\alpha \le 0.25$. Then the sequence $\{(x^k, s^k, \tau_k)\}$ produced by Algorithm 5.2 converges, and the safeguard is activated only a finite number of times. The complexity results stated in Corollary 3.11 and Corollary 3.12 hold also for Algorithm 5.2. Moreover the sequence $\{\tau_k\}$ converges to zero quadratically, and consequently the sequences $\{\|r^k\|\}$ and $\{(x^k)^T s^k\}$ also converge quadratically to zero.*

Here is a general idea of the proof. We have to prove that $\gamma = O(\tau)$. We first prove the following lemma, that establishes a strong connection between the proximity at a new point and the amount $\|xs(1)/\tau - e\|$. In the proof of Theorem, we show that if the large variables are sufficiently well centered, then $\|xs(1)/\tau - e\|$ is small enough to obtain $\gamma = O(\tau)$ using (5.28). Now if the safeguard was activated an infinite number of times, the sequences of points would converge to the analytical center by Theorem 4.13, and we would get a contradiction using (5.28). Knowing that the safeguard is not active after a finite number of steps, we deduce with (5.29) an estimate of $\|xs(1)/\tau - e\|$ that allows us to check with (5.28) that $\gamma = O(\tau)$, as desired.

The lemma below was proved by Gonzaga (1994) in the framework of feasible algorithms. Although the proof needs no modification, we give it for the self-containedness of the paper.

LEMMA 5.4. *Consider $w \in \mathcal{N}_\alpha$, $\gamma \in (0, 1]$. Then*

$$\delta^\#(\gamma) \le (1 - \gamma)\left\| \frac{xs(1)}{\tau} - e \right\| + \gamma\delta^\#(1) + O\left(\frac{\tau}{\gamma}\right), \tag{5.28}$$

$$\delta^\#(\gamma) \ge (1 - \gamma)\left\| \frac{xs(1)}{\tau} - e \right\| - \gamma\delta^\#(1) + O\left(\frac{\tau}{\gamma}\right). \tag{5.29}$$

PROOF. By Lemma 4.8, we have $x(0) = x + u(0) = x + O(\tau)$ and $s(0) = s + v(0) = s + \tau d^{-1}\phi + O(\tau^2) = O(\tau^2)$. It follows that

$$x(\gamma) = \gamma x(1) + (1 - \gamma)x(0) = \gamma x(1) + (1 - \gamma)x + O(\tau),$$

$$s(\gamma) = \gamma s(1) + (1 - \gamma)s(0) = \gamma s(1) + O(\tau^2).$$

Multiplying these expressions, we obtain

$$\frac{x(\gamma)s(\gamma)}{\gamma\tau} = \gamma\frac{x(1)s(1)}{\tau} + (1 - \gamma)\frac{xs(1)}{\tau} + O\left(\frac{\tau}{\gamma}\right). \tag{5.30}$$

The result follows by subtracting $e = \gamma e + (1 - \gamma)e$ and taking norms. □

PROOF OF THE THEOREM. We first prove that the safeguard cannot be activated an infinite number of times. Indeed, if the safeguard was activated an infinite number of times, then $\sum_{k=1}^{\infty} \theta_k \gamma_k = 1$, as $\theta_k = 1$ at all iterations and $\gamma_k \ge \epsilon'$ at the restoration steps. From Theorem 4.13 we deduce that $\{x^k\}$ converges to the shifted analytic center. It follows then that $g(x^k) \to 0$, i.e., $g(x) \approx o(1)$, where g is the projected gradient of the shifted analytic barrier function given by (4.7). Using Lemma 4.5, (4.21) and the fact that $x \approx 1$ and $d \approx 1$, we deduce $P_{\bar{Q}}(\phi^{-1} + d\bar{s}) \approx o(1)$ (see (4.25)). From the equation of the centering step (3.2), Lemma 4.8, and $s = O(\tau)$ we obtain

$$\frac{xs(1)}{\tau} - e = -\frac{su^c}{\tau} = O\left(dP_{\bar{Q}}(\phi^{-1} + d\bar{s})\right) + O(\tau).$$

Hence $\|xs(1)/\tau - e\| \approx o(1)$. From (3.6) and (3.30) we have

$$\delta^{\#}(0.1) \le \alpha_{cc} \le \frac{\alpha}{3\sqrt{8}} \le 0.03.$$

Finally, according to (5.28) we deduce that

$$\delta^{\#}(0.1) \le 0.9 \times o(1) + 0.1 \times 0.03 + O(\tau) \le 0.003 + o(1).$$

Consequently the safeguard will not be activated after a certain iteration, which gives the desired contradiction.

Having established that the safeguard cannot be activated an infinite number of times, we deduce with (5.29) that after a certain iteration,

$$0.9\left\|\frac{xs(1)}{\tau} - e\right\| \le 0.2 + 0.1 \times 0.03 + O(\tau),$$

whence after a finite number of iterations $\|xs(1)/\tau - e\| \le 0.23$. With (5.28), it follows that

$$0.25 \le 0.23 + O\left(\frac{\tau}{\gamma}\right),$$

which implies $\gamma = O(\tau)$, as was to be proved. The fact that the complexity results stated in Corollaries 3.11 and 3.12 hold for Algorithm 5.2 follows from the inequality $\epsilon' < 1 - 1/(2\sqrt{n})$. □

Acknowledgment. The work of the second author was supported in part by NSF, Grant DMS 9305760. The main results of this paper were obtained during a visit of the second author to INRIA in June 1994, supported by the projects PROMATH and MENUSIN.

References

Anitescu, M., G. Lesaja, F. A. Potra (1994). *On the equivalence between different formulations and algorithms for the $P_*(\kappa)$ – LCP.* Reports on Computational Mathematics, Department of Mathematics, The University of Iowa, Iowa City, IA 52242, USA.

Bonnans, J. F., C. C. Gonzaga (1993). Convergence of interior point algorithms for the monotone linear complementarity problem. *Math. Oper. Res.* **21** 1–25.

Gonzaga, C. C. (1989). Conical projection algorithms for linear programming. *Math. Programming* **43** 151–173.

—— (1997). The largest step path following algorithm for monotone linear complementary problems. *Math. Programming* **76** 309–332.

——, J. F. Bonnans (1997). Fast convergence of the simplified largest step path following algorithm. *Math. Programming* **76** 95–115.

——, R. A. Tapia (1992). On the convergence of the Mizuno-Todd-Ye algorithm to the analytic center of the solution set. *SIAM J. Optim.* (to appear).

Güler, O. (1995). Generalized linear complementarity problems. *Math. Oper. Res.* **20** 441–448.

McShane, K. A. (1994). Superlinearly convergent $O(\sqrt{n}L)$-iteration interior-point algorithms for linear programming and the monotone linear complementarity problem. *SIAM J. Optim.* **4** 247–261.

Mizuno, S., M. Kojima, M. J. Todd (1995). Infeasible-interior-point primal-dual potential-reduction algorithms for linear programming. *SIAM J. Optim.* **5** 52–67.

Monteiro, R. D. C., T. Tsuchiya. (1996). Limiting behavior of the derivatives of certain trajectories associated with a monotone horizontal linear complementarity problem, *Math. Oper. Res.* **21** 793–814.

Potra, F. A. (1996). An $O(nL)$ infeasible-interior-point algorithm for LCP with quadratic convergence. *Annals Oper. Res.* **62** 81–102.

——, R. Sheng (1994). *A modified $O(nL)$ infeasible-interior-point algorithm for LCP with quadratic convergence.* Reports on Computational Mathematics 54, Department of Mathematics, The University of Iowa, Iowa City, IA 52242, USA. *SIAM J. Optimization*, to appear.

——, —— (1995a). *A path following method for LCP with superlinearly convergent iteration sequence.* Reports on Computational Mathematics 69, Department of Mathematics, The University of Iowa, Iowa City, IA 52242, USA. *Annals Oper. Res.*, to appear.

——, —— (1995b). *A superlinearly convergent infeasible-interior-point algorithm for degenerate LCP.* Reports on Computational Mathematics 66, Department of Mathematics, The University of Iowa, Iowa City, IA 52242, USA.

Tütüncü, R. H., M. J. Todd (1995). Reducing horizontal linear complementarity problems. *Linear Algebra Appl.* **223/224** 717–729.

Wright, S. J. (1993). A path-following infeasible-interior-point algorithm for linear complementarity problems. *Optim. Methods and Software* **2** 79–106.

—— (1994). An infeasible interior point algorithm for linear complementarity problems. *Math. Programming* **67** 29–52.

—— (1996). A path-following interior-point algorithm for linear and quadratic problems. *Ann. Oper. Res.* **62** 103–130.

Ye, Y. (1997). On homogeneous and self-dual algorithm for lcp. *Math. Programming* **76** 211–222.

Zhang, Y. (1994). On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem. *Math. Oper. Res.* (to appear).

J. F. Bonnans: INRIA, B.P. 105, 78153 Rocquencourt, France; e-mail: Frederic.Bonnans@inria.fr.

F. A. Potra: Department of Mathematics, The University of Iowa, Iowa City, Iowa 52242; e-mail: potra@math.uiowa.edu.