On the Resolution of Monotone Complementarity Problems

Carl Geiger and Christian Kanzow *

Institute of Applied Mathematics

University of Hamburg

Bundesstrasse 55

D–20146 Hamburg

Germany

April, 1994

Abstract. A reformulation of the nonlinear complementarity problem (NCP) as an unconstrained minimization problem is considered. It is shown that any stationary point of the unconstrained objective function is already a solution of NCP if the mapping F involved in NCP is continuously differentiable and monotone. A descent algorithm is described which uses only function values of F. Some numerical results are given.

Key words. nonlinear complementarity problems, unconstrained minimization, stationary points, global minima, descent methods.

AMS (MOS) subject classification. 90C33, 90C30, 65K05.

Abbreviated title. Resolution of Monotone Complementarity Problems.

1 Introduction

Consider the complementarity problem NCP(F)

$$x \geq 0, F(x) \geq 0, x^T F(x) = 0, \tag{1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a given function. In a number of recent papers, this problem has been reformulated as a minimization problem in order to apply well–developed optimization methods to problem (1). This might be of particular interest

Preprint 82, Institute of Applied Mathematics, University of Hamburg, April 1994.

*correspondence to: Christian Kanzow, e-mail: kanzow@math.uni-hamburg.de

in the large–scale case. For example, Mangasarian and Solodov [13] introduce an unconstrained minimization problem with the property that any global minimizer of their objective function is a solution of (1) (see Section 5 for a more detailed description). Yamashita and Fukushima [21] prove that each stationary point of Mangasarian and Solodov's function is already a global minimum and thus a solution of (1) if the function F is continuously differentiable and $F'(x)$ is positive definite for all $x \in \mathbb{R}^n$. This has also been shown in [8] for a more general class of functions.

In case $F'(x)$ is only assumed to be positive semidefinite for all $x \in \mathbb{R}^n$, Friedlander, Martnez and Santos [4] have shown that problem (1) can be formulated as a bound constrained optimization problem in such a way that each Karush–Kuhn–Tucker point of this constrained optimization problem leads to a solution of (1). As a specialization of a more general result for variational inequality problems, Fukushima [5] also obtains a bound constrained optimization formulation of (1), for which he proves equivalence to problem (1) for monotone functions F, see also Taji, Fukushima and Ibaraki [18].

In this paper, we make use of a tool introduced in [8] in order to rewrite problem (1) as an unconstrained optimization problem. In Section 2, we show that each stationary point of the unconstrained objective function is a solution of (1) if F is a continuously differentiable and monotone function. Some global and local convergence properties are proved in Section 3. A descent method for our unconstrained objective function is proposed in Section 4 which does not use any derivative information of F. It is shown that any stationary point is already a solution of NCP(F) for this method. Section 5 contains a short review of Mangasarian and Solodov's approach. Some numerical results are given in Section 6. The results are compared with the ones obtained using Mangasarian and Solodov's function. We conclude this paper with some final remarks in Section 7.

2 The Equivalence Theorem

Let $\varphi_F : \mathbb{R}^2 \to \mathbb{R}$ be the function defined by

$$\varphi_F(a, b) := \sqrt{a^2 + b^2} - a - b. \tag{2}$$

This function has recently been introduced by Fischer in order to characterize the Karush–Kuhn–Tucker conditions of a nonlinear program (see [1]) and the linear complementarity problem (see [2]) as a (nondifferentiable) system of equations. Here, we are interested in the square of Fischer's function, namely

$$\varphi(a, b) := \frac{1}{2} \left(\sqrt{a^2 + b^2} - a - b \right)^2 . \tag{3}$$

Some easily established properties of this function are summarized in the following lemma, see also [10].

2.1 Lemma. (i) $\varphi(a,b)=0 \Longleftrightarrow a\geq 0, b\geq 0, ab=0$.
(ii) $\varphi(a,b)\geq 0\,\forall\,(a,b)^T\in\mathbb{R}^2$.
(iii) φ *is continuously differentiable for all* $(a,b)^T\in\mathbb{R}^2$*, in particular* $\nabla\varphi(0,0)=(0,0)^T$.
(iv) $\frac{\partial\varphi}{\partial a}(a,b)\frac{\partial\varphi}{\partial b}(a,b)\geq 0\ \forall\ (a,b)^T\in\mathbb{R}^2$.
(v) $\frac{\partial\varphi}{\partial a}(a,b)\frac{\partial\varphi}{\partial b}(a,b)=0 \Longrightarrow \varphi(a,b)=0$.

Now, consider the nonlinear complementarity problem (1) and the related unconstrained optimization problem

$$\min_{x\in\mathbb{R}^n}\Psi(x), \tag{4}$$

where $\Psi:\mathbb{R}^n\to\mathbb{R}$ is defined by

$$\Psi(x):=\sum_{i=1}^n\varphi(x_i,F_i(x)), \tag{5}$$

$F_i:\mathbb{R}^n\to\mathbb{R}$ being the ith component function of F $(i\in I:=\{1,\ldots,n\})$. Due to Lemma 2.1, properties (i) and (ii), we have the following result:

2.2 Lemma. *Assume that the complementarity problem (1) has at least one solution. Then* $x^*\in\mathbb{R}^n$ *solves the complementarity problem if and only if* x^* *is a global minimum of the unconstrained minimization problem (4).*

The equivalence stated in Lemma 2.2 is not true if the complementarity problem (1) is not solvable. This is shown in the next

2.3 Example. Let $n=1$ and $F(x):=-x^2-1$. Then it is not difficult to see that the corresponding function $\Psi(x)=1/2\left(\sqrt{x^2+(x^2+1)^2}-x+x^2+1\right)^2$ has compact level sets and therefore must have a global minimum. On the other hand, the complementarity problem itself has obviously no solutions.

The problem of finding a global minimum is in general quite difficult. It is therefore of interest under what assumptions on the function F stationary points of Ψ are already global minima. The following result has been shown in [8].

2.4 Theorem. *Let* $F\in C^1(\mathbb{R}^n)$ *have a positive definite Jacobian* $F'(x)$ *for all* $x\in\mathbb{R}^n$*. Then* x^* *is a global minimum of* Ψ *if and only if* x^* *is a stationary point of* Ψ.

In fact, a more general theorem has been proved in [8], since it was a main purpose of that paper to provide general conditions on the functions φ and F such that Theorem 2.4 is true for an entire class of functions Ψ. For the particular function Ψ defined in (5)/(3), however, we can prove the following stronger result. Note that this result holds although Ψ is in general a nonconvex function. Moreover, the result is independent of whether or not the complementarity problem is solvable.

2.5 Theorem. *Let $F \in C^1(\mathbb{R}^n)$ be a monotone function, i.e. $(x-y)^T(F(x) - F(y)) \geq 0$ for all $x, y \in \mathbb{R}^n$. Then $x^* \in \mathbb{R}^n$ is a global minimum of the unconstrained optimization problem (4) if and only if x^* is a stationary point of Ψ.*

Proof. First, let x^* be a global minimum of Ψ. Since F is continuously differentiable, our function Ψ is also continuously differentiable because of Lemma 2.1 (iii). Thus, the gradient of Ψ exists and vanishes in x^*. Next, assume that x^* is a stationary point of Ψ, i.e., let

$$0 = \nabla\Psi(x^*) = \sum_{i=1}^{n} \left(\frac{\partial\varphi}{\partial a}(x_i^*, F_i(x^*))e_i + \frac{\partial\varphi}{\partial b}(x_i^*, F_i(x^*))\nabla F_i(x^*) \right), \tag{6}$$

where e_i denotes the ith column vector of the identity matrix I_n. Let us abbreviate the vectors $\left(\ldots, \frac{\partial\varphi}{\partial a}(x_i^*, F_i(x^*)), \ldots\right)^T$ and $\left(\ldots, \frac{\partial\varphi}{\partial b}(x_i^*, F_i(x^*)), \ldots\right)^T$ by $\frac{\partial\varphi}{\partial a}(x^*, F(x^*))$ and $\frac{\partial\varphi}{\partial b}(x^*, F(x^*))$, respectively. Then, the stationary conditions (6) can be rewritten as

$$0 = \frac{\partial\varphi}{\partial a}(x^*, F(x^*)) + F'(x^*)^T \frac{\partial\varphi}{\partial b}(x^*, F(x^*)). \tag{7}$$

Premultiplying (7) by $\left(\frac{\partial\varphi}{\partial b}(x^*, F(x^*))\right)^T$ yields

$$0 = \sum_{i=1}^{n} \left(\frac{\partial\varphi}{\partial a}(x_i^*, F_i(x^*)) \frac{\partial\varphi}{\partial b}(x_i^*, F_i(x^*)) \right) + \frac{\partial\varphi}{\partial b}(x^*, F(x^*))^T F'(x^*)^T \frac{\partial\varphi}{\partial b}(x^*, F(x^*)). \tag{8}$$

Since F is monotone, the Jacobian $F'(x^*)$ is positive semidefinite (see, e.g., Ortega and Rheinboldt [17], p. 142). Using property (iv) of Lemma 2.1, we therefore obtain from (8):

$$\frac{\partial\varphi}{\partial a}(x_i^*, F_i(x^*)) \frac{\partial\varphi}{\partial b}(x_i^*, F_i(x^*)) = 0 \quad (i = 1, \ldots, n).$$

This, however, yields

$$\varphi(x_i^*, F_i(x^*)) = 0 \quad (i = 1, \ldots, n)$$

because of Lemma 2.1 (v). Consequently, we have $\Psi(x^*) = 0$, i.e., x^* is a global minimizer of Ψ. $\square$

From Lemma 2.2 and Theorem 2.5 we directly obtain the following result:

2.6 Corollary. *Let $F \in C^1(\mathbb{R}^n)$ be a monotone function. If the complementarity problem (1) is solvable, then x^* is a solution of (1) if and only if x^* is a stationary point of Ψ.*

3 Convergence Properties

We first prove that the level sets of our unconstrained objective function (5) are bounded for strongly monotone functions F. Recall that $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to be strongly monotone (with modulus $\mu > 0$) if

$$(x-y)^T(F(x)-F(y)) \geq \mu\|x-y\|^2 \ \forall\ x, y \in \mathbb{R}^n. \tag{9}$$

It is well–known that for $F \in C^1(\mathbb{R}^n)$, condition (9) is equivalent to

$$d^T F'(x) d \geq \mu\|d\|^2 \ \ \forall x \in \mathbb{R}^n \ \ \forall d \in \mathbb{R}^n, \tag{10}$$

see Ortega and Rheinboldt [17], p. 142. It turns out that the following result is of great help.

3.1 Lemma. *Let* $\{(a^k, b^k)\}_{k\in\mathbb{N}} \subseteq \mathbb{R}^2$ *be any sequence such that* $|a^k|, |b^k| \to +\infty \ \ (k \in \mathbb{N})$. *Then* $\varphi(a^k, b^k) \to \infty \ \ (k \in \mathbb{N})$.

Proof. This follows immediately from Lemma 2.8 in [9]. □

We are now ready to state the main result of this section.

3.2 Theorem. *Suppose that* F *is continuous and strongly monotone. Let* $x^0 \in \mathbb{R}^n$ *be any given vector, and let* $L(x^0) := \{x \in \mathbb{R}^n | \Psi(x) \leq \Psi(x^0)\}$ *be the corresponding level set. Then* $L(x^0)$ *is compact.*

Proof. Assume that there is a sequence $\{x^k\}_{k\in\mathbb{N}} \subseteq L(x^0)$ such that $\lim_{k\to\infty} \|x^k\| = \infty$. Define the index set

$$J := \{i \in I | \{x_i^k\}_{k\in\mathbb{N}} \text{ is unbounded}\}.$$

By our assumption $J \neq \emptyset$. Let $\{y^k\}_{k\in\mathbb{N}}$ denote the sequence defined by

$$y_i^k := \begin{cases} 0 & \text{if } i \in J, \\ x_i^k & \text{if } i \notin J. \end{cases}$$

From the definition of $\{y^k\}_{k\in\mathbb{N}}$ and the strong monotonicity of F, we get

$$\left.\begin{array}{rcl} \mu \sum_{i\in J}(x_i^k)^2 & = & \mu\|x^k - y^k\|^2 \\ & \leq & \sum_{i=1}^n (x_i^k - y_i^k)(F_i(x^k) - F_i(y^k)) \\ & = & \sum_{i\in J} x_i^k (F_i(x^k) - F_i(y^k)) \\ & \leq & \sqrt{\sum_{i\in J}(x_i^k)^2} \sum_{i\in J} |F_i(x^k) - F_i(y^k)|. \end{array}\right\} \tag{11}$$

Since $\sum_{i\in J}(x_i^k)^2 \neq 0$ for all $k \in K_1$, K_1 being an infinite subset of $\mathbb{N}$, we obtain from (11):

$$\mu\sqrt{\sum_{i\in J}(x_i^k)^2} \leq \sum_{i\in J} |F_i(x^k) - F_i(y^k)| \ \ (k \in K_1). \tag{12}$$

Due to the boundedness of the sequence $\{y^k\}_{k\in K_1}$ and the continuity of the functions F_i $(i \in J)$, the sequences $\{F_i(y^k)\}_{k\in K_1}$ $(i \in J)$ remain also bounded. Because of (12), we therefore have

$$|F_{i_0}(x^k)| \to \infty$$

for an index $i_0 \in J$. From the definition of the index set J, it follows that

$$|x^k_{i_0}| \to \infty \;\; (k \in K_2 \subseteq K_1).$$

Consequently, Lemma 3.1 yields

$$\varphi(x^k_{i_0}, F_{i_0}(x^k)) \to \infty \;\; (k \in K_2).$$

This, however, contradicts the fact that

$$\varphi(x^k_{i_0}, F_{i_0}(x^k)) \le \Psi(x^k) \le \Psi(x^0) \;\; \forall \;\; k \in \mathbb{N}.$$

□

We emphasize that Theorem 3.2 is true for any function φ satisfying the condition of Lemma 3.1. Furthermore, note that this result is independent of any differentiability assumptions.

Theorem 3.2 implies that if we apply a line search descent method to minimize the objective function Ψ such that the search directions satisfy, e.g., an angle condition and the steplength procedure is, say, efficient in the sense defined by Warth and Werner [19] and Werner [20], then any accumulation point of this sequence is a stationary point of Ψ and thus a solution of NCP(F) because of Corollary 2.6. Moreover, since NCP(F) has a unique solution for strongly monotone F, the entire sequence converges to this solution.

The following result shows that the Hessian matrix of $\Psi(x)$ is positive definite at a solution x^* under certain assumptions. This result is a special case of a more general theorem proved in [8].

3.3 Theorem. *Let $x^* \in \mathbb{R}^n$ be a nondegenerate solution of NCP(F), i.e., $x^*_i + F_i(x^*) > 0$ $(i \in I)$. Let F be twice continuously differentiable. Assume that the gradients $\nabla F_i(x^*)$ $(i \notin I^* := \{i \in I | x^*_i = 0\})$ and e_i $(i \in I^*)$ are linearly independent. Then the Hessian matrix $\nabla^2\Psi(x^*)$ exists and is positive definite.*

As a consequence of Theorem 3.3, any descent method for solving problem (4) finally achieves its known local rate of convergence.

4 A Descent Method

We present a descent method for minimizing our unconstrained objective function Ψ which does not need any explicit derivatives of the function F involved in the nonlinear complementarity problem. Moreover, we prove a global convergence result for this descent method. Given an iterate $x^k \in \mathbb{R}^n$, let $\frac{\partial\varphi}{\partial a}(x^k, F(x^k))$ and $\frac{\partial\varphi}{\partial b}(x^k, F(x^k))$ denote the n-vectors having as ith components $\frac{\partial\varphi}{\partial a}(x_i^k, F_i(x^k))$ and $\frac{\partial\varphi}{\partial b}(x_i^k, F_i(x^k))$, respectively. Let

$$d^k := -\frac{\partial\varphi}{\partial b}(x^k, F(x^k)) \tag{13}$$

be a search direction. By the following lemma, d^k is a descent direction of Ψ at x^k under monotonicity assumptions.

4.1 Lemma. *Let $x^k \in \mathbb{R}^n$ and let $F \in C^1(\mathbb{R}^n)$ be a monotone function. Then the search direction d^k defined in (13) satisfies the descent condition $\nabla\Psi(x^k)^T d^k < 0$ as long as x^k is not a solution of NCP(F). Moreover, if F is strongly monotone with modulus $\mu > 0$, then*

$$\nabla\Psi(x^k)^T d^k \le -\mu\|d^k\|^2.$$

Proof. Using the representations (6)/(7) of the gradient $\nabla\Psi(x^k)$ and the definition (13) of d^k, we obtain

$$\nabla\Psi(x^k)^T d^k = -\sum_{i=1}^n \frac{\partial\varphi}{\partial a}(x_i^k, F_i(x^k))\frac{\partial\varphi}{\partial b}(x_i^k, F_i(x^k)) - (d^k)^T F'(x^k)^T d^k. \tag{14}$$

By our assumptions, the Jacobian matrix $F'(x^k)$ is positive semidefinite. Consequently, we obtain from (14) and Lemma 2.1 (iv):

$$\nabla\Psi(x^k)^T d^k \le 0.$$

Assume that $\nabla\Psi(x^k)^T d^k = 0$. Then $\frac{\partial\varphi}{\partial a}(x_i^k, F_i(x^k))\frac{\partial\varphi}{\partial b}(x_i^k, F_i(x^k)) = 0$ for all $i \in I$. Lemma 2.1 (v) therefore yields $\varphi(x_i^k, F_i(x^k)) = 0$ $(i \in I)$, i.e., $x^k \in \mathbb{R}^n$ solves NCP(F) in contrast to our assumption.

If F is strongly monotone with modulus $\mu > 0$, we obtain from (14), Lemma 2.1 (iv) and (10)

$$\nabla\Psi(x^k)^T d^k \le -(d^k)^T F'(x^k)^T d^k \le -\mu\|d^k\|^2.$$

□Lemma 4.1 motivates the following algorithm:

4.2 Algorithm.

(S.0): Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a strongly monotone function. Define $\Psi : \mathbb{R}^n \to \mathbb{R}$ as in (5). Let $x^0 \in \mathbb{R}^n, \epsilon > 0, \sigma \in (0,1)$ and $\beta \in (0,1)$. Set $k := 0$.

(S.1): If $\Psi(x^k) < \epsilon$, stop: x^k is an approximate solution of NCP(F).

(S.2): Let $d^k := -\frac{\partial\varphi}{\partial b}(x^k, F(x^k))$.

(S.3): Compute a steplength $t_k = \beta^{m_k}$, where m_k is the smallest nonnegative integer m satisfying the Armijo–type condition

$$\Psi(x^k + \beta^m d^k) \le \Psi(x^k) - \beta^m\sigma\|d^k\|^2.$$

(S.4): Set $x^{k+1} := x^k + t_k d^k, k := k+1$ and go to (S.1).

The next theorem is a global convergence result for Algorithm 4.2.

4.3 Theorem. *Let $F \in C^1(\mathbb{R}^n)$ be a strongly monotone function with modulus $\mu > 0$. Let $x^0 \in \mathbb{R}^n$ be any given starting point, and let $L(x^0)$ denote its level set. Assume that the Jacobian F' is Lipschitz–continuous in $L(x^0)$. If $\sigma < \mu$ then the sequence $\{x^k\}_{k\in\mathbb{N}}$ generated by Algorithm 4.2 is well–defined and converges to the unique solution x^* of NCP(F).*

Proof. It follows from the assumptions that $\nabla\Psi$ is a Lipschitz–continuous function in $L(x^0)$, i.e.

$$\|\nabla\Psi(x) - \nabla\Psi(y)\| \le L\|x-y\| \text{ for all } x, y \in L(x^0)$$

and some constant $L > 0$. Therefore, using Lemma 4.1 and the Mean Value Theorem, we obtain for $x^{k+1} = x^k + td^k, 0 \le t \le 1$, and $\theta^k = x^k + \vartheta_k(x^{k+1} - x^k), \vartheta_k \in (0,1)$:

$$\begin{aligned}
\Psi(x^{k+1}) - \Psi(x^k) &= \nabla\Psi(\theta^k)^T(x^{k+1} - x^k) \\
&= t\nabla\Psi(x^k)^T d^k + t\left(\nabla\Psi(\theta^k) - \nabla\Psi(x^k)\right)^T d^k \\
&\le -t\mu\|d^k\|^2 + tL\|\theta^k - x^k\|\,\|d^k\| \\
&= \left(-t\mu + t^2 L\right)\|d^k\|^2.
\end{aligned}$$

Therefore, the inequality

$$\Psi(x^k + td^k) \le \Psi(x^k) - \sigma t\|d^k\|^2$$

holds for all $0 \le t \le \min\{1, (\mu - \sigma)/L\}$. Consequently, the steplength t_k computed in step (S.3) of Algorithm 4.2 is bounded from below by

$$t_k \ge \min\{\beta, \beta(\mu - \sigma)/L\}. \tag{15}$$

In particular, a steplength $t_k > 0$ satisfying the Armijo–type condition in step (S.3) can always be found, i.e., Algorithm 4.2 is well–defined. Since the sequence $\{\Psi(x^k)\}_{k\in\mathbb{N}}$ is monotonically decreasing and nonnegative, it follows from

$$\Psi(x^{k+1}) \le \Psi(x^k) - t_k\sigma\|d^k\|^2$$

and (15) that

$$\lim_{k\to\infty} \|d^k\|^2 = 0.$$

This implies that

$$\lim_{k\to\infty} \frac{\partial\varphi}{\partial b}(x^k, F(x^k)) = 0.$$

Therefore and because of Lemma 2.1 (v), any accumulation point of the sequence $\{x^k\}_{k\in\mathbb{N}}$ is a solution of problem NCP(F). Since the sequence $\{x^k\}_{k\in\mathbb{N}}$ remains in

$L(x^0)$ and since $L(x^0)$ is compact by Theorem 3.2, there exists at least one accumulation point x^*. Due to the strong monotonicity of F, problem NCP(F) has a unique solution, so the entire sequence $\{x^k\}_{k\in\mathbb{N}}$ must converge to x^*. □

Note that the proof of Theorem 4.3 in particular guarantees the existence of a solution of the nonlinear complementarity problem associated with strongly monotone functions F.

5 The Approach of Mangasarian and Solodov

We give a short review of the method recently proposed by Mangasarian and Solodov [13] and further analysed in Yamashita and Fukushima [21], which is closely related to our approach. We note, however, that the presentation of their method given here differs from the one in [13] and [21].

Mangasarian and Solodov introduce the function

$$\varphi_{MS}(a,b;\alpha) := ab + \frac{1}{2\alpha}\left(\max{}^2\{0, a-\alpha b\} - a^2 + \max{}^2\{0, b-\alpha a\} - b^2\right) \qquad (16)$$

and prove the following result:

5.1 Lemma. *For any parameter $\alpha > 1$, the following holds:*
(i) $\varphi_{MS}(a,b;\alpha) \geq 0 \ \forall \ (a,b)^T \in \mathbb{R}^2$.
(ii) $\varphi_{MS}(a,b;\alpha) = 0 \Longleftrightarrow a \geq 0, b \geq 0, ab = 0$.

Based on the function (16), the unconstrained optimization problem

$$\min_{x\in\mathbb{R}^n} M(x;\alpha) := \sum_{i=1}^{n} \varphi_{MS}(x_i, F_i(x);\alpha) \qquad (17)$$

is considered in [13]. Due to Lemma 5.1, there is a one–to–one correspondence between global minimizers of problem (17) and solutions of the complementarity problem NCP(F). Yamashita and Fukushima [21] prove that stationary points of $M(x;\alpha)$ are already solutions of NCP(F) if F is differentiable and has a positive definite Jacobian $F'(x)$ for all $x \in \mathbb{R}^n$ (see also [8]). This is a stronger assumption than the one used in our Theorem 2.5, in particular, since Yamashita and Fukushima were able to show by a counterexample that their result is incorrect even for strictly monotone functions F. Moreover, Yamashita and Fukushima [21] prove a result analogous to our Lemma 4.1, but once again they need the positive definiteness of $F'(x)$ to prove the descent condition $\nabla M(x;\alpha)^T d < 0$, where the search direction d is given by

$$d := -\alpha \frac{\partial \varphi_{MS}}{\partial b}(x, F(x);\alpha).$$

6 Numerical Results

In this section, we compare Mangasarian and Solodov's reformulation (17) of the nonlinear complementarity problem with our approach (4). We first present some results for Algorithm 4.2 using $\varepsilon = 10^{-5}, \sigma = 10^{-4}, \beta = 0.5$ and the starting vector $x^0 = (0, \ldots, 0)^T \in \mathbb{R}^n$. The algorithm has been applied to two linear complementarity problems, i.e., $F(x) = Mx + q$ is an affine–linear function, $M \in \mathbb{R}^{n \times n}, q \in \mathbb{R}^n$. The first example is given by

$$M = \begin{pmatrix} 4 & -1 & 0 & \ldots & 0 \\ -1 & 4 & -1 & \ldots & 0 \\ 0 & -1 & 4 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & 0 & 0 & -1 & 4 \end{pmatrix}, q = (-1, \ldots, -1)^T, \qquad (18)$$

the second example has the data

$$M = diag(1/n, 2/n, \ldots, 1), q = (-1, \ldots, -1)^T. \qquad (19)$$

The numerical results are given in Tables 1 and 2 for different dimensions n. The rows denoted by Ψ and M contain the number of iterations needed by our approach (4) and by Mangasarian and Solodov's reformulation (17), respectively. The first example has been solved without any problems, and the number of iterations remains almost constant. On the other hand, both methods have substantial difficulties in solving the second example, and the number of iterations increases linearly with the dimension n.

Table 1: Number of iterations for example (18) using Algorithm 4.2

n	8	16	32	64	128	256
Ψ	35	42	43	43	43	43
M	10	11	12	13	13	14

Table 2: Number of iterations for example (19) using Algorithm 4.2

n	8	16	32	64	128	256
Ψ	36	79	165	337	682	1374
M	173	347	696	1395	2791	5584

Algorithm 4.2 is in general only well–defined for strongly monotone functions. However, due to Corollary 2.6, we are also interested in monotone functions. We

therefore present a standard line search method. Since the objective functions of both minimization problems ((4) and (17)) are only once continuously differentiable, it is in general not advisable to use a second order method. We have therefore decided to use a limited memory BFGS method (see Nocedal [15]), which is only based on gradient information and which has recently been shown to be one of the most successful methods for (large–scale) unconstrained optimization, see Gilbert and Lemarchal [6], Liu and Nocedal [11], Nocedal [16], Nash and Nocedal [14] and Zou et al. [23]. Below we give a formal description of the limited memory BFGS method. It makes use of the abbreviations $g^k := \nabla\Psi(x^k)$, $s^k := x^{k+1} - x^k$ and $y^k := g^{k+1} - g^k$.

6.1 Limited memory BFGS method.

(S.0) (Initial data):
Choose $x^0 \in \mathbb{R}^n, m \in \mathbb{N}, \varepsilon > 0, \sigma \in (0, \frac{1}{2}), \rho \in (\sigma, 1)$ and a symmetric and positive definite starting matrix $H_0 \in \mathbb{R}^{n \times n}$. Set $k := 0$.

(S.1) (Stopping criterion):
If $\Psi(x^k) < \varepsilon$, stop: x^k is an approximate solution of problem (1).

(S.2) (Computation of a search direction):
Compute $d^k := -H_k g^k$.

(S.3) (Computation of a steplength):
Compute a steplength $t_k > 0$ satisfying the strong Wolfe conditions

$$\begin{aligned} \Psi(x^k + t_k d^k) &\leq \Psi(x^k) + \sigma t_k (g^k)^T d^k, \\ |\nabla\Psi(x^k + t_k d^k)^T d^k| &\leq -\rho (g^k)^T d^k. \end{aligned}$$

(S.4) (Update):
Set $x^{k+1} := x^k + t_k d^k$. Define $\rho_k := 1/(y^k)^T s^k$ and $V_k := I_n - \rho_k y^k (s^k)^T$. Let $\hat{m} := \min\{k, m-1\}$. Update H_0 $\hat{m}+1$ times using the pairs $\{(s^j, y^j)\}_{j=k-\hat{m}}^k$, i.e., let

$$\begin{aligned} H_{k+1} = & \left(V_k^T \dots V_{k-\hat{m}}^T\right) H_0 (V_{k-\hat{m}} \dots V_k) \\ & + \rho_{k-\hat{m}} \left(V_k^T \dots V_{k-\hat{m}+1}^T\right) s^{k-\hat{m}} (s^{k-\hat{m}})^T (V_{k-\hat{m}+1} \dots V_k) \\ & + \rho_{k-\hat{m}+1} \left(V_k^T \dots V_{k-\hat{m}+2}^T\right) s^{k-\hat{m}+1} (s^{k-\hat{m}+1})^T (V_{k-\hat{m}+2} \dots V_k) \\ & \vdots \\ & + \rho_k s^k (s^k)^T. \end{aligned}$$

(S.5) (Loop):
Set $k := k+1$ and go to (S.1).

6.2 Remark. a.) In our numerical experiments, we have chosen the following values for the parameters in step (S0): $\varepsilon = 10^{-5}, \sigma = 10^{-4}, \rho = 0.9$ and $H_0 = I$.

b.) The steplength $t_k > 0$ satisfying the strong Wolfe conditions has been computed via the algorithm described in Fletcher [3].

c.) It is computationally advantageous to replace the matrix H_0 in step (S4) by a matrix $H_k^{(0)}$, where $H_k^{(0)} = \gamma_k H_0$ and γ_k is a scaling parameter. Here, we follow Liu and Nocedal [11], who recommend the choice $\gamma_k = (y^k)^T s^k / \|y^k\|^2$.

d.) The matrices H_k do not have to be formed explicitly. Instead, the last $\hat{m}+1$ vectors s^j and y^j are stored, and the search direction can be computed with this data using the two–loop recursion described in Nocedal [15].

As test problems, we have chosen some convex constrained optimization problems, namely problems 34, 35, 66 and 76 from the book of Hock and Schittkowski [7]: Their Karush–Kuhn–Tucker (KKT) optimality conditions lead to complementarity problems of dimensions 8, 4, 8 and 7, respectively. It is important to note that these complementarity problems are monotone, but not strictly monotone. Consequently, it is not guaranteed for these problems that any stationary point of Mangasarian and Solodov's objective function (17) is already a solution of the complementarity problem. Furthermore, we note that problems 35 and 76 are quadratic programming problems, so the corresponding complementarity problems are linear, whereas problems 34 and 66 lead to nonlinear complementarity problems.

The results obtained with algorithm 6.1 being applied to the four test problems are summarized in Tables 3–6. We report the number of iterations for $m = 5$ and $m = 7$ (recall that m denotes the number of vector pairs (s^j, y^j) stored in the limited memory BFGS method). The stopping criterion of algorithm 6.1 has been changed as follows:

$$\text{If } \|\nabla\Psi(x^k)\| < \varepsilon, \text{ then terminate the iteration.}$$

The parameter α of Mangasarian and Solodov's function has been taken as $\alpha = 1.1$. We stress that Mangasarian and Solodov's function is the zero function in the limiting case $\alpha = 1$. This can easily be verified, see also Lemma 2.1 in Luo et al. [12]. Therefore, since this function is continuous in $\alpha > 0$, it is an "almost" linear function for $\alpha \approx 1$, and good numerical performance of the corresponding method is expected in this case.

TABLE 3: Number of iterations for example HS 34

	m=5		m=7	
starting vector	Ψ	M	Ψ	M
(1,1,1,1,1,1,1,1)	114	101	112	99
(2,2,2,2,2,2,2,2)	107	106	105	94
(1,1,1,0,0,0,0,0)	101	100	102	99
(-1,-1,-1,1,1,1,1,1)	110	98	104	89
(1,1,1,-10,-10,-10,-10,-10)	114	112	113	95

TABLE 4: Number of iterations for example HS 35

	m=5		m=7	
starting vector	Ψ	M	Ψ	M
(0.5,0.5,0.5,1)	30	30	27	29
(10,10,10,10)	43	9*	34	9*
(100,100,100,100)	44	10*	42	10*
(100,10,1,0)	43	14*	40	14*
(-1,-10,-100,-1000)	53	17*	48	19*

TABLE 5: Number of iterations for example HS 66

	m=5		m=7	
starting vector	Ψ	M	Ψ	M
(0,1.05,2.9,0,0,0,0,0)	39	35	29	32
(0,0,0,1,1,1,1,1)	64	44	45	29
(-1,-1,-1,1,1,1,1,1)	43	45	46	32
(1,1,1,-1,-1,-1,-1,-1)	61	62	45	41
(-1,-1,-1,0,1,2,3,4)	62	41	52	37

TABLE 6: Number of iterations for example HS 76

	m=5		m=7	
starting vector	Ψ	M	Ψ	M
(0.5,0.5,0.5,0.5,0,0,0)	47	42	33	33
(0,0,0,0,0,0,0)	48	33	36	32
(10,10,10,10,10,10,10)	102	27*	73	36
(0,1,0,1,0,1,0)	41	40	34	31
(1,2,3,4,3,2,1)	50	42	39	30

The results in Tables 3–6 indicate the following: If both methods converge to a solution of the underlying problem, then Mangasarian and Solodov's method is usually slightly superior to our method. However, in several instances, their method converge to a stationary point which is not a solution of the corresponding complementarity problem (this is indicated by an asterisk in the tables), whereas our method solves these problems as guaranteed by our theory, see Corollary 2.6. To us, it is surprising how often Mangasarian and Solodov's method converge only to a stationary point for example HS 35. We have therefore tested with some randomly generated starting values, and again, in most cases convergence was observed to stationary points only. However, this behaviour of their method has been observed for the linear complementarity problems only (mostly for example HS 35), whereas the nonlinear complementarity problems have been solved by both methods without any difficulties. We have not an explanation for this yet.

Nevertheless, we can summarize the results as follows: Mangasarian and Solodov's reformulation behaves slightly better than ours, but their method should only be used for complementarity problems whose associated function F has a positive definite Jacobian everywhere. If the Jacobian is only positive semidefinite, i.e., if F is a monotone function, their approach is not a reliable technique, and our method is preferable.

7 Final Remarks

In this paper, we have presented a reformulation of the nonlinear complementarity problem as an unconstrained optimization problem. It has been shown that this reformulation is equivalent to the complementarity problem for monotone functions F. Since several complementarity problems are just monotone and in general have not a positive definite Jacobian, we feel that our approach is an important extension of Mangasarian and Solodov's method.

In a very recent paper, Yamashita and Fukushima [22] have extended Mangasarian and Solodov's approach to the generalized complementarity problem. We believe that a similar extension is possible for our method. Again, however, it should be possible to prove similar results as in [22] under weaker assumptions.

References

[1] A. Fischer (1992): *A special Newton–type optimization method.* Optimization 24, pp. 269–284.

[2] A. Fischer (1994): *A globally and Q–quadratically convergent Newton–type method for positive semidefinite linear complementarity problems.* Journal of Optimization Theory and Applications, to appear.

[3] R. Fletcher (1987): *Practical Methods of Optimization* (second edition). John Wiley and Sons, Chichester.

[4] A. Friedlander, J. M. Martnez and S. A. Santos (1993): *Resolution of linear complementarity problems using minimization with simple bounds.* Technical Report, Department of Applied Mathematics, University of Campinas, Campinas, Brazil.

[5] M. Fukushima (1992): *Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems.* Mathematical Programming (Series A) 53, pp. 99–110.

[6] J. Ch. Gilbert and C. Lemarchal (1989): *Some numerical experiments with variable–storage quasi–Newton algorithms.* Mathematical Programming (Series A) 45, pp. 407–435.

[7] W. Hock and K. Schittkowski (1981): *Test Examples for Nonlinear Programming Codes.* Lecture Notes in Economics and Mathematical Systems 187, Springer–Verlag, Berlin, Germany.

[8] C. Kanzow (1993): *Nonlinear complementarity as unconstrained optimization.* Preprint 67, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany (revised February, 1994).

[9] C. Kanzow (1993): *Global convergence properties of some iterative methods for linear complementarity problems.* Preprint 72, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany (revised January, 1994).

[10] C. Kanzow (1993): *An unconstrained optimization technique for large–scale linearly constrained convex minimization problems.* Preprint 74, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany.

[11] D. C. Liu and J. Nocedal (1989): *On the limited memory BFGS method for large scale optimization.* Mathematical Programming (Series A) 45, pp. 503–528.

[12] Z.–Q. Luo, O. L. Mangasarian, J. Ren and M. V. Solodov (1994): *New error bounds for the linear complementarity problem.* Mathematics of Operations Research, to appear.

[13] O. L. Mangasarian and M. V. Solodov (1993): *Nonlinear complementarity as unconstrained and constrained minimization.* Mathematical Programming (Series B) 62, pp. 277–297.

[14] St. G. Nash and J. Nocedal (1991): *A numerical study of the limited memory BFGS method and the truncated–Newton method for large scale optimization.* SIAM Journal on Optimization 1, pp. 358–372.

[15] J. Nocedal (1980): *Updating quasi–Newton matrices with limited storage.* Mathematics of Computation 35, pp. 773–782.

[16] J. Nocedal (1990): *The performance of several algorithms for large scale unconstrained optimization.* In: Th. F. Coleman and Y. Li (eds.): Large–Scale Numerical Optimization. SIAM, Philadelphia, pp. 138–151.

[17] J. M. Ortega and W. C. Rheinboldt (1970): *Iterative Solution of Nonlinear Equations in Several Variables.* Academic Press, New York – San Francisco – London.

[18] K. Taji, M. Fukushima and T. Ibaraki (1993): *A globally convergent Newton method for solving strongly monotone variational inequalities.* Mathematical Programming (Series A) 58, pp. 369–383.

[19] W. Warth and J. Werner (1977): *Effiziente Schrittweitenfunktionen bei unrestringierten Optimierungsaufgaben.* Computing 19, pp. 59–72.

[20] J. Werner (1978): *ber die globale Konvergenz von Variable–Metrik–Verfahren mit nicht–exakter Schrittweitenbestimmung.* Numerische Mathematik 31, pp. 321–334.

[21] N. Yamashita and M. Fukushima (1994): *On stationary points of the implicit Lagrangian for nonlinear complementarity problems.* Journal of Optimization Theory and Applications, to appear.

[22] N. Yamashita and M. Fukushima (1993): *Implicit Lagrangian for generalized complementarity problems.* Technical Report 93010, Nara Institute of Science and Technology, Ikoma, Nara 630–01, Japan.

[23] X. Zou, I. M Navon, M. Berger, K. H. Phua, T. Schlick and F. X. Le Dimet (1993): *Numerical experience with limited–memory quasi–Newton and truncated Newton methods.* SIAM Journal on Optimization 3, pp. 582–608.