Appl Math Optim 40:315–339 (1999)

Applied Mathematics & Optimization

A Regularization Newton Method for Solving Nonlinear Complementarity Problems*

D. Sun

School of Mathematics, University of New South Wales,
Sydney 2052, NSW, Australia
sun@maths.unsw.edu.au

Communicated by J. Stoer

Abstract. In this paper we construct a regularization Newton method for solving the nonlinear complementarity problem (NCP(F)) and analyze its convergence properties under the assumption that F is a P_0-function. We prove that every accumulation point of the sequence of iterates is a solution of NCP(F) and that the sequence of iterates is bounded if the solution set of NCP(F) is nonempty and bounded. Moreover, if F is a monotone and Lipschitz continuous function, we prove that the sequence of iterates is bounded if and only if the solution set of NCP(F) is nonempty by setting $t = \frac{1}{2}$, where $t \in [\frac{1}{2}, 1]$ is a parameter. If NCP(F) has a locally unique solution and satisfies a nonsingularity condition, then the convergence rate is superlinear (quadratic) without strict complementarity conditions. At each step, we only solve a linear system of equations. Numerical results are provided and further applications to other problems are discussed.

Key Words. Nonlinear complementarity problem, Nonsmooth equations, Regularization, Generalized Newton method, Convergence.

AMS Classification. 90C33, 90C30, 65H10.

* This work was supported by the Australian Research Council.

1. Introduction

In this paper we consider numerical methods for solving the nonlinear complementarity problem (NCP(F) for abbreviation):

$$x \geq 0, \qquad F(x) \geq 0 \quad \text{and} \quad x^T F(x) = 0, \tag{1.1}$$

where F: $\Re^n \to \Re^n$ is any given function which we assume to be continuously differentiable throughout this paper.

This problem has attracted much attention due to its various applications. We refer the reader to [17], [25], [15], and [11] for a review. The methods considered here are intended to handle singular nonlinear complementarity problems, in which the derivative of the mapping F may be seriously ill-conditioned. The singularity problem will prevent most of the currently available algorithms from converging to a solution of NCP(F). In the literature there are two classes of methods that can be used to deal with the singular nonlinear complementarity problems: regularization methods [5], [27] and proximal point methods [35], [36]. Proximal point methods have been investigated by many researchers. See the recent report of Eckstein and Ferris [6] and references therein for a review. The methods discussed in this paper are in the class of regularization methods. This class of methods try to circumvent the singularity problem by considering a sequence of perturbed problems, which possibly have better conditions. For nonlinear complementarity problems, the simplest regularization technique is to use the so-called *Tikhonov-regularization*, which consists in solving a sequence of complementarity problems NCP(F_ε):

$$x \geq 0, \qquad F_\varepsilon(x) \geq 0 \quad \text{and} \quad x^T F_\varepsilon(x) = 0, \tag{1.2}$$

where $F_\varepsilon(x) := F(x) + \varepsilon x$ and ε is a positive parameter converging to zero. In this paper, for the convenience of discussion, we allow ε to take a nonpositive value also.

Regularization methods, closely related to proximal point methods, for solving monotone complementarity problems have been studied in the literature by several authors, see, e.g., [37] or Theorem 5.6.2 of [3]. Recently, Facchinei and Kanzow [9] generalized most of the classic results for the monotone complementarity problems to the larger class of P_0 complementarity problems. In particular, they established a result on the behavior of the inexact solutions of the perturbed problems. This result states that if F is a P_0-function and the solution set $\mathcal{S}$ of NCP(F) is nonempty and bounded then the (inexact) solutions of NCP(F_ε) are uniformly bounded as $\varepsilon \to 0_+$. For a precise description, see Theorem 5.4 of [9]. Facchinei and Kanzow's result generalized a result in [41] for linear complementarity problems when the exact solutions of perturbed problems are considered. In this paper we consider a regularization Newton method by using Facchinei and Kanzow's result, a result developed in Section 2 for monotone complementarity problems and the techniques established recently on smoothing methods [32]. For this regularization Newton method we, under the assumption that F is a P_0-function, prove that every accumulation point of the sequence of iterates is a solution of NCP(F) and that the iteration sequence is bounded if the solution set of NCP(F) is nonempty and bounded. Moreover, if F is monotone and Lipschitz continuous, we prove that the sequence of iterates is bounded if and only if the solution set of NCP(F) is nonempty by setting $t = \frac{1}{2}$, where $t \in [\frac{1}{2}, 1]$ is a parameter. If NCP(F) has a locally unique solution and satisfies a nonsingularity condition, then the convergence rate is superlinear (quadratic) without assuming the strict complementarity conditions. To our knowledge, this is the first method, not only in the scope of regularization methods, to have all these properties. It is necessary to point out that in order to make the sequence of iterates bounded for P_0 or monotone nonlinear complementarity problems, people usually use various neighborhoods in both interior point methods [40], [43], [44], and noninterior point methods [18], [1], [31] such that the sequence of iterates stays in these neighborhoods. In general, to keep the sequence of iterates staying in these neighborhoods one needs additional work, probably time consuming, and it may prevent the sequence of iterates from converging to a solution superlinearly if strict complementarity conditions are not satisfied. A neighborhood is also introduced in this paper, however, this neighborhood, which does not appear in our algorithm and is only used to analyze the convergence properties of the algorithm, allows us to achieve superlinear convergence no matter whether strict complementarity conditions hold or not. These features can be seen clearly in the following discussion.

The organization of this paper is as follows. In the next section we study some preliminary properties of the reformulated nonsmooth equations and their solutions. In Section 3 we state the algorithm and prove several propositions related to the algorithm. In Section 4 we establish the global convergence of the algorithm. We analyze the superlinear and quadratic convergence properties of the algorithm in Section 5 and give preliminary numerical results in Section 6. Final conclusions are given in Section 7.

A word about our notation is in order. Let $\|\cdot\|$ and $\|\cdot\|_\infty$ denote the l_2 norm and the l_∞ norm of $\Re^m$, respectively. For a continuously differentiable function Φ: $\Re^m \to \Re^m$, we denote the Jacobian of Φ at $x \in \Re^m$ by $\Phi'(x)$, whereas the transposed Jacobian is denoted as $\nabla\Phi(x)$. If W is an $m \times m$ matrix with entries W_{jk}, $j, k = 1, \ldots, m$, and $\mathcal{J}$ and $\mathcal{K}$ are index sets such that $\mathcal{J}, \mathcal{K} \subseteq \{1, \ldots, m\}$, we denote by $W_{\mathcal{J}\mathcal{K}}$ the $|\mathcal{J}| \times |\mathcal{K}|$ submatrix of W consisting of entries W_{jk}, $j \in \mathcal{J}, k \in \mathcal{K}$. If $W_{\mathcal{J}\mathcal{J}}$ is nonsingular, we denote by $W/W_{\mathcal{J}\mathcal{J}}$ the Schur-complement of $W_{\mathcal{J}\mathcal{J}}$ in W, i.e., $W/W_{\mathcal{J}\mathcal{J}} := W_{\mathcal{K}\mathcal{K}} - W_{\mathcal{K}\mathcal{J}} W_{\mathcal{J}\mathcal{J}}^{-1} W_{\mathcal{J}\mathcal{K}}$, where $\mathcal{K} = \{1, \ldots, m\} \backslash \mathcal{J}$.

2. Some Preliminaries

We need the following definitions concerning matrices and functions.

Definition 2.1. A matrix $W \in \Re^{n \times n}$ is called a

- P_0-matrix if every one of its principal minors is nonnegative;
- P-matrix if every one of its principal minors is positive.

Obviously, a positive semidefinite matrix is a P_0-matrix and a positive definite matrix is a P-matrix.

Definition 2.2. A function $F\colon \Re^n \to \Re^n$ is called a

- P_0-function if, for every x and y in $\Re^n$ with $x \neq y$, there is an index i such that

  $$x_i \neq y_i, \qquad (x_i - y_i)(F_i(x) - F_i(y)) \geq 0;$$

- P-function if, for every x and y in $\Re^n$ with $x \neq y$, there is an index i such that

  $$x_i \neq y_i, \qquad (x_i - y_i)(F_i(x) - F_i(y)) > 0;$$

- uniform P-function if there exists a positive constant μ such that, for every x and y in $\Re^n$, there is an index i such that

  $$(x_i - y_i)(F_i(x) - F_i(y)) \geq \mu \|x - y\|^2;$$

- monotone function if, for every x and y in $\Re^n$,

  $$(x - y)^T(F(x) - F(y)) \geq 0;$$

- strongly monotone function if there exists a positive constant μ such that, for every x and y in $\Re^n$,

  $$(x - y)^T(F(x) - F(y)) \geq \mu \|x - y\|^2.$$

It is known that every strongly monotone function is a uniform P-function and every monotone function is a P_0-function. Furthermore, the Jacobian of a continuously differentiable P_0-function (uniform P-function) is a P_0-matrix (P-matrix).

The Fischer–Burmeister function $\psi\colon \Re^2 \to \Re$ introduced by Fischer in [12] is defined by

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

The function $\psi(\cdot)$ has the following important property:

$$\psi(a, b) = 0 \quad \Longleftrightarrow \quad a, b \geq 0, \qquad ab = 0.$$

Another important property of $\psi(\cdot)$ is that $\psi(\cdot)^2$ is continuously differentiable on the whole space $\Re^2$ [21]. The advent of the Fischer–Burmeister function, has attracted a lot of attention in the area of nonlinear complementarity and variational inequality problems, see, e.g., [13], [20], [10], [4], [8], [46] and references therein. It was first used by Facchinei and Kanzow in [9] to study regularization methods. Here we choose to use it for the sake of simplicity. However, there are many other functions that can be chosen from [29]. In particular, we can use the following function which was originally discussed in [38]:

$$\psi(a, b) := \begin{cases} \sqrt{a^2 + b^2} - (a + b) & \text{if} \quad a \geq 0, \quad b \geq 0, \\ b & \text{if} \quad a \geq 0, \quad b < 0, \\ a & \text{if} \quad a < 0, \quad b \geq 0, \\ -\sqrt{a^2 + b^2} & \text{if} \quad a < 0, \quad b < 0. \end{cases}$$

The above function is linear in the second and fourth quadrants, but its square is still continuously differentiable [38].

Let $z := (\varepsilon, x) \in \Re \times \Re^n$ and let $F_{\varepsilon,i}$ be the ith component of F_ε, $i \in N := \{1, 2, \ldots, n\}$. Define $G\colon \Re^{n+1} \to \Re^n$ by

$$G_i(\varepsilon, x) := \psi(x_i, F_{\varepsilon,i}(x)), \qquad i \in N, \tag{2.1}$$

where $\psi(\cdot)$ is the Fischer–Burmeister function. Denote $H\colon \Re^{n+1} \to \Re^{n+1}$ by

$$H(z) := \begin{pmatrix} \varepsilon \\ G(z) \end{pmatrix}, \qquad z = (\varepsilon, x) \in \Re \times \Re^n. \tag{2.2}$$

Such a defined function H is locally Lipschitz continuous because ψ is locally Lipschitz continuous [12]. It is easy to see that

$$H(\varepsilon, x) = 0 \iff \varepsilon = 0, \qquad x \text{ is a solution of NCP}(F),$$
$$(\varepsilon, x) \in \Re \times \Re^n.$$

Let $\varphi\colon \Re^{n+1} \to \Re_+$ be defined by

$$\varphi(z) := \|G(z)\|^2.$$

Then, because $\psi(\cdot)^2$ is continuously differentiable on $\Re^2$, $\varphi(\cdot)$ is continuously differentiable on the whole space $\Re^{n+1}$. Define the merit function $f\colon \Re^{n+1} \to \Re_+$ by

$$f(z) := \|H(z)\|^2 = \varepsilon^2 + \varphi(z).$$

Then f is also continuously differentiable on $\Re^{n+1}$.

Lemma 2.1 [22]. *Let $\{a^k\}, \{b^k\} \subseteq \Re$ be any two sequences such that either $a^k, b^k \to \infty$ or $a^k \to -\infty$ or $b^k \to -\infty$. Then $|\psi(a^k, b^k)| \to \infty$.*

The following lemma was provided in the proof of Proposition 3.4 in [9].

Lemma 2.2. *Suppose that F is a P_0-function. Then for any sequence $\{x^k\}$ such that $\|x^k\| \to \infty$, there exist an index $i \in N$ and a subsequence $\{x^{k_j}\}$ such that either $x_i^{k_j} \to \infty$ and $F_i(x^{k_j})$ does not tend to $-\infty$ or $x_i^{k_j} \to -\infty$.*

The next result generalizes a result in [9] and is more suitable to our discussion. However, its proof should be credited to Facchinei and Kanzow.

Proposition 2.1. *Suppose that F is a P_0-function and that $\tilde{\varepsilon}, \bar{\varepsilon}$ are two given positive numbers such that $\bar{\varepsilon} \geq \tilde{\varepsilon}$. Then for any sequence $\{z^k = (\varepsilon^k, x^k)\}$ such that $\varepsilon^k \in [\tilde{\varepsilon}, \bar{\varepsilon}]$ and $\|x^k\| \to \infty$ we have*

$$\lim_{k\to\infty} f(z^k) = \infty. \tag{2.3}$$

Proof. For the sake of contradiction, suppose that there exists a sequence $\{z^k = (\varepsilon^k, x^k)\}$ such that $\varepsilon^k \in [\tilde{\varepsilon}, \bar{\varepsilon}]$ and $\|x^k\| \to \infty$ and $f(z^k)$ is bounded. Then from Lemma 2.2, by taking a subsequence if necessary, there exists an index $i \in N$ such that either $x_i^k \to \infty$ and $F_i(x^k)$ does not tend to $-\infty$ or $x_i^k \to -\infty$. Since $\varepsilon^k \in [\tilde{\varepsilon}, \bar{\varepsilon}]$, for the above i either $x_i^k \to \infty$ and $F_i(x^k) + \varepsilon^k x_i^k \to \infty$ or $x_i^k \to -\infty$. Hence, from Lemma 2.1,

$$|\psi(x_i^k, F_i(x^k) + \varepsilon^k x_i^k)| \to \infty,$$

which means $f(x^k) \to \infty$. This is a contradiction. So, we complete our proof. □

The following result is extracted from Theorem 5.4 of [9].

Lemma 2.3. *Suppose that F is a P_0-function and that the solution set of* NCP(F) *is nonempty and bounded. Suppose that $\{\varepsilon^k\}$ and $\{\eta^k\}$ are two infinite sequences such that, for each $k \geq 0$, $\varepsilon^k > 0$, $\eta^k \geq 0$ satisfying* $\lim_{k\to\infty} \varepsilon^k = 0$, $\lim_{k\to\infty} \eta^k = 0$. *For each $k \geq 0$, let $x^k \in \Re^n$ satisfy $\|G(\varepsilon^k, x^k)\| \leq \eta^k$. Then $\{x^k\}$ remains bounded and every accumulation point of $\{x^k\}$ is a solution of* NCP(F).

Lemma 2.4. *Suppose that F is a monotone function. Let $X \in \Re^n$ be any nonempty closed convex subset of $\Re^n$ and for any given $\varepsilon > 0$ define $W_\varepsilon\colon \Re^n \to \Re^n$ by*

$$W_\varepsilon(u) := u - \Pi_X[u - (F(u) + \varepsilon u)], \qquad u \in \Re^n,$$

where for any $v \in \Re^n$, $\Pi_X(v)$ is the orthogonal projection of v onto X. Let x be a solution of $W_\varepsilon(u) = 0$ and let y be a solution of $W_\varepsilon(u) - \delta = 0$, where $\delta \in \Re^n$ is any given vector. Then if $x \neq y$,

$$\|y - x\| \leq \left[(1+\varepsilon) + \frac{\|F(y) - F(x)\|}{\|y - x\|}\right] \frac{\|\delta\|}{\varepsilon}. \tag{2.4}$$

Proof. First, for the projection operator $\Pi_X(\cdot)$ for any $u, v \in \Re^n$ we have from [47] that

$$(u - v)^T(\Pi_X(u) - \Pi_X(v)) \geq \|\Pi_X(u) - \Pi_X(v)\|^2.$$

By letting $u := y - (F(y) + \varepsilon y)$ and $v := x - (F(x) + \varepsilon x)$ in the above relation we obtain

$$\begin{aligned}
&\{[y - (F(y) + \varepsilon y)] - [x - (F(x) + \varepsilon x)]\}^T \\
&\qquad \times\{\Pi_X[y - (F(y) + \varepsilon y)] - \Pi_X[x - (F(x) + \varepsilon x)]\} \\
&\quad \geq \|\Pi_X[y - (F(y) + \varepsilon y)] - \Pi_X[x - (F(x) + \varepsilon x)]\|^2,
\end{aligned}$$

which, combining with the assumptions that $x = \Pi_X[x - (F(x) + \varepsilon x)]$ and $y = \Pi_X[y - (F(y) + \varepsilon y)] + \delta$, gives

$$\{[y - (F(y) + \varepsilon y)] - [x - (F(x) + \varepsilon x)]\}^T\{y - \delta - x\} \geq \|y - x - \delta\|^2.$$

By rearrangements,

$$\varepsilon\|y-x\|^2+\|\delta\|^2 \le (1+\varepsilon)\delta^T(y-x)-(y-x)^T(F(y)-F(x)) \\ +\delta^T(F(y)-F(x)).$$

Hence, because F is monotone, $(y-x)^T(F(y)-F(x)) \ge 0$, we have

$$\begin{aligned}\varepsilon\|y-x\|^2+\|\delta\|^2 &\le (1+\varepsilon)\delta^T(y-x)+\delta^T(F(y)-F(x)) \\ &\le (1+\varepsilon)\|\delta\|\|y-x\|+\|\delta\|\|F(y)-F(x)\|.\end{aligned}$$

Since $\|\delta\|^2 \ge 0$, the above relation proves (2.4). □

Proposition 2.2. *Suppose that F is a monotone function and that the solution set $\mathcal{S}$ of* NCP(F) *is nonempty. Suppose that $\{\varepsilon^k\}$ and $\{\eta^k\}$ are two infinite sequences such that for each $k \ge 0$, $\varepsilon^k > 0$, $\eta^k \ge 0$, $\eta_k \le C\varepsilon^k$, and $\lim_{k\to\infty}\varepsilon^k = 0$, where $C > 0$ is a constant. For each $k \ge 0$, let $x^k \in \Re^n$ satisfy $\|G(\varepsilon^k, x^k)\| \le \eta^k$. Suppose that $x^* = \arg\min_{x\in\mathcal{S}}\|x\|$ and that F is Lipschitz continuous, i.e., there exists a constant $L > 0$ such that, for any $u, v \in \Re^n$,*

$$\|F(u)-F(v)\| \le L\|u-v\|. \tag{2.5}$$

Then $\{x^k\}$ remains bounded and every accumulation point of $\{x^k\}$ is a solution of NCP(F).

Proof. Suppose that, for any $\varepsilon > 0$, $x(\varepsilon)$ is a solution of NCP(F_ε). For any $\varepsilon > 0$, let $W_\varepsilon\colon \Re^n \to \Re^n$ be defined by

$$W_\varepsilon(u) := u - \Pi_{\Re^n_+}[u-(F(u)+\varepsilon u)] = \min\{u, F(u)+\varepsilon u\}, \qquad u \in \Re^n.$$

Then, from Lemma 2.4,

$$\|x^k - x(\varepsilon^k)\| \le \left[(1+\varepsilon^k)+\frac{\|F(x^k)-F(x(\varepsilon^k))\|}{\|x^k-x(\varepsilon^k)\|}\right]\frac{\|W_{\varepsilon^k}(x^k)\|}{\varepsilon^k}. \tag{2.6}$$

By Tseng [39], for any two numbers $a, b \in \Re$, we have

$$\frac{1}{2+\sqrt{2}}|\min\{a,b\}| \le |\psi(a,b)| \le (2+\sqrt{2})|\min\{a,b\}|.$$

Hence,

$$\|W_{\varepsilon^k}(x^k)\| \le (2+\sqrt{2})\|G(\varepsilon^k, x^k)\|. \tag{2.7}$$

Then by using (2.5)–(2.7) and the assumption that $\|G(\varepsilon^k, x^k)\| \le C\varepsilon^k$ we have

$$\|x^k - x(\varepsilon^k)\| \le \left[(1+\varepsilon^k)+L\right](2+\sqrt{2})C.$$

This implies that $\{x^k\}$ remains bounded because $x(\varepsilon^k) \to x^*$ as $k \to \infty$ [5]. By the continuity of f every accumulation point of $\{x^k\}$ is a solution of NCP(F). □

In order to design high-order convergent Newton methods we need the concept of semismoothness. Semismoothness was originally introduced by Mifflin [24] for functionals. Convex functions, smooth functions, and piecewise linear functions are examples of semismooth functions. The composition of semismooth functions is still a semismooth function [24]. In [33] Qi and Sun extended the definition of semismooth functions to $\Phi\colon \Re^{m_1} \to \Re^{m_2}$. A locally Lipschitz continuous vector valued function $\Phi\colon \Re^{m_1} \to \Re^{m_2}$ has a generalized Jacobian $\partial\Phi(x)$ as in [2]. Φ is said to be *semismooth* at $x \in \Re^{m_1}$, if

$$\lim_{\substack{V\in\partial\Phi(x+th')\\ h'\to h,\ t\downarrow 0}} \{Vh'\}$$

exists for any $h \in \Re^{m_1}$. It has been proved in [33] that Φ is semismooth at x if and only if all its component functions are. Also, $\Phi'(x; h)$, the directional derivative of Φ at x in the direction h, exists for any $h \in \Re^{m_1}$ if Φ is semismooth at x.

Lemma 2.5 [33]. *Suppose that $\Phi\colon \Re^{m_1} \to \Re^{m_2}$ is a locally Lipschitzian function and semismooth at x. Then*

(i) *for any $V \in \partial\Phi(x+h),\ h \to 0,$*

$$Vh - \Phi'(x; h) = o(\|h\|);$$

(ii) *for any $h \to 0,$*

$$\Phi(x+h) - \Phi(x) - \Phi'(x; h) = o(\|h\|).$$

A stronger notion than semismoothness is strong semismoothness. $\Phi(\cdot)$ is said to be *strongly semismooth* at x if Φ is semismooth at x and, for any $V \in \partial\Phi(x+h)$, $h \to 0$,

$$Vh - \Phi'(x; h) = O(\|h\|^2).$$

(Note that in [33] and [28] different names for strong semismoothness are used.) A function Φ is said to be a (strongly) semismooth function if it is (strongly) semismooth everywhere.

Lemma 2.6. *Suppose that F is continuously differentiable. Then the function H defined by* (2.2) *is a semismooth function. If $F'(\cdot)$ is Lipschitz continuous in a neighborhood of a point $x \in \Re^n$, then H is strongly semismooth at $(\varepsilon, x) \in \Re^{n+1}$.*

Proof. Since a function is (strongly) semismooth at a point if and only if its component functions are, to prove that H is a semismooth function we only need to prove that H_i, $i = 1, \ldots, n+1$, are semismooth functions and to prove H is strongly semismooth at (ε, x) we only need to prove that H_i, $i = 1, \ldots, n+1$, are strongly semismooth at (ε, x). Apparently, H_1 is a strongly semismooth function because, for any $z = (\varepsilon, y) \in \Re \times \Re^n$, $H_1(z) = \varepsilon$. Next, we consider H_i, $i = 2, \ldots, n+1$. By noting that $\psi\colon \Re^2 \to \Re$ is a (strongly) semismooth function [30, Lemma 3.1], and that the composition of two semismooth functions is a semismooth function we conclude that H_i, $i = 2, \ldots, n+1$, are semismooth functions. If $F'(\cdot)$ is Lipschitz continuous in a neighborhood of a point $x \in \Re^n$, then $\tilde{F}$ is strongly semismooth at $(\varepsilon, x) \in \Re^{n+1}$, where for any $z = (\varepsilon, y) \in$

$\Re \times \Re^n$, $\tilde{F}(z) := F_\varepsilon(y)$. Hence by Theorem 19 of [14] we can conclude that H_i, $i = 2, \ldots, n+1$, are strongly semismooth at $(\varepsilon, x) \in \Re^{n+1}$. So, we complete our proof. □

3. A Regularization Newton Method

Choose $\bar{\varepsilon} \in (0, \infty)$ and $\gamma \in (0, 1)$ such that $\gamma\bar{\varepsilon} < 1$. Let $t \in [\frac{1}{2}, 1]$ and let $\bar{z} := (\bar{\varepsilon}, 0) \in \Re \times \Re^n$. Define β: $\Re^{n+1} \to \Re_+$ by

$$\beta(z) := \gamma \min\{1, f(z)^t\}.$$

Let

$$\Omega := \{z = (\varepsilon, x) \in \Re \times \Re^n \mid \varepsilon \geq \beta(z)\bar{\varepsilon}\}.$$

Then, because, for any $z \in \Re^{n+1}$, $\beta(z) \leq \gamma$, it follows that, for any $x \in \Re^n$,

$$(\bar{\varepsilon}, x) \in \Omega.$$

We prove in what follows that if we choose the initial point $z^0 = (\bar{\varepsilon}, x^0)$, where $x^0 \in \Re^n$ is an arbitrary point, then the sequence of iterates generated by Algorithm 3.1, which is introduced later, will remain in Ω. This is an important feature because through this neighborhood we can prevent ε from approaching 0 too fast and thus avoid the iterates converging to a nonsolution point.

Proposition 3.1. *The following relations hold*:

$$H(z) = 0 \iff \beta(z) = 0 \iff H(z) = \beta(z)\bar{z}.$$

Proof. It follows from the definitions of $H(\cdot)$ and $\beta(\cdot)$ that

$$H(z) = 0 \iff \beta(z) = 0 \quad \text{and} \quad \beta(z) = 0 \implies H(z) = \beta(z)\bar{z}.$$

Then we only need to prove

$$H(z) = \beta(z)\bar{z} \implies \beta(z) = 0.$$

From $H(z) = \beta(z)\bar{z}$ we have

$$\varepsilon = \beta(z)\bar{\varepsilon} \quad \text{and} \quad G(z) = 0.$$

Hence, from the definitions of $f(\cdot)$ and $\beta(\cdot)$ and the fact that $\gamma\bar{\varepsilon} < 1$, we get

$$f(z) = \varepsilon^2 + \|G(z)\|^2 = \varepsilon^2 = \beta(z)^2\bar{\varepsilon}^2 \leq \gamma^2\bar{\varepsilon}^2 < 1.$$

Hence,

$$\beta(z) = \gamma f(z)^t = \gamma\beta(z)^{2t}\bar{\varepsilon}^{2t}. \tag{3.1}$$

If $\beta(z) \neq 0$, it follows from (3.1) and the fact $\beta(z) \leq \gamma$ that

$$1 = \gamma\beta(z)^{2t-1}\bar{\varepsilon}^{2t} \leq \gamma^{2t}\bar{\varepsilon}^{2t},$$

which contradicts the fact that $\gamma\bar{\varepsilon} < 1$. This contradiction completes our proof. □

Algorithm 3.1.

Step 0. Choose constants $\delta \in (0, 1)$, $t \in [\frac{1}{2}, 1]$, and $\sigma \in (0, \frac{1}{2})$. Let $\varepsilon^0 := \bar{\varepsilon}$, $x^0 \in \Re^n$ be an arbitrary point and $k := 0$.
Step 1. If $H(z^k) = 0$, then stop. Otherwise, let $\beta_k := \beta(z^k) = \gamma \min\{1, f(z^k)^t\}$.
Step 2. Choose $V_k \in \partial H(z^k)$ and compute $\Delta z^k = (\Delta\varepsilon^k, \Delta x^k) \in \Re \times \Re^n$ by

$$H(z^k) + V_k \Delta z^k = \beta_k \bar{z}. \tag{3.2}$$

Step 3. Let l_k be the smallest nonnegative integer l satisfying

$$f(z^k + \delta^l \Delta z^k) \leq [1 - 2\sigma(1 - \gamma\bar{\varepsilon})\delta^l] f(z^k). \tag{3.3}$$

Define $z^{k+1} := z^k + \delta^{l_k} \Delta z^k$.
Step 4. Replace k by $k + 1$ and go to Step 1.

Remark. (i) The direction Δz^k computed in (3.2) is an approximated generalized Newton direction of H at z^k because $\beta_k \bar{z}$ is introduced in the right-hand side of (3.2). To introduce $\beta_k \bar{z}$ to (3.2) has two benefits: (a) it ensures that all ε^k are positive, and (b) it keeps the whole sequence of iterates $\{z^k\}$ in Ω, see Propositions 3.2–3.4. As we claimed in the Introduction, (b) plays an important role in proving the global convergence of Algorithm 3.1 under the assumption that F is a P_0-function only. It is noted that Ω does no appear in our algorithm and thus it does not need additional work to guarantee (b).

(ii) The idea of introducing one or several parameters in the reformulated systems as free variables is not new, see, for example, [18], [31], [32], and [19]. Here we adopted the approach developed in [32] on smoothing methods mainly because it guarantees both features (a) and (b) in (i).

(iii) For any $V \in \partial H(z)$, $z = (\varepsilon, x) \in \Re \times \Re^n$ there exists a $W = (W_\varepsilon \; W_x) \in \partial G(z)$ with $W_\varepsilon \in \Re^n$ and $W_x \in \Re^{n \times n}$ such that

$$V = \begin{pmatrix} 1 & 0 \\ W_\varepsilon & W_x \end{pmatrix}.$$

Suppose that F is a P_0-function. Then, for any $x \in \Re^n$ and $\varepsilon > 0$, $F'(x)$ is a P_0-matrix and $F'_\varepsilon(x)$ is a P-matrix. Hence, for any $x \in \Re^n$ and $\varepsilon > 0$, W_x is nonsingular, refer to Proposition 3.2 of [20] for a proof. It thus follows that all $V \in \partial H(z)$ are nonsingular, $z \in \Re_{++} \times \Re^n$. Hence, from (3.2), for any $k \geq 0$ and $\varepsilon^k > 0$, there exists a $W_k \in \partial G(z^k)$ such that

$$(\nabla\varphi(z^k))^T \Delta z^k = 2G(z^k)^T W_k \Delta z^k = -2G(z^k)^T G(z^k) = -2\varphi(z^k). \tag{3.4}$$

Proposition 3.2. *Suppose that F is a P_0-function. Then, for any $k \geq 0$, if $z^k = (\varepsilon^k, x^k) \in \Re_{++} \times \Re^n$, then Algorithm* 3.1 *is well defined at k*th *iteration and $z^{k+1} \in \Re_{++} \times \Re^n$.*

Proof. First, from Proposition 3.1 and that $\varepsilon^k \in \Re_{++}$, we have

$$\beta_k = \beta(z^k) > 0.$$

By using (3.2) in Algorithm 3.1, we get

$$\Delta\varepsilon^k = -\varepsilon^k + \beta_k \bar{\varepsilon}. \tag{3.5}$$

Hence, for any $\alpha \in [0, 1]$, we have

$$\varepsilon^k + \alpha\Delta\varepsilon^k = (1-\alpha)\varepsilon^k + \alpha\beta_k\bar{\varepsilon} \in \Re_{++}. \tag{3.6}$$

Then, from (3.6), (3.2), and the fact that $\beta(z) = \gamma \min\{1,\ f(z)^t\} \leq \gamma f(z)^{1/2}$ (note that $t \in [\frac{1}{2}, 1]$), for any $\alpha \in [0, 1]$ we have

$$\begin{aligned}(\varepsilon^k + \alpha\Delta\varepsilon^k)^2 &= [(1-\alpha)\varepsilon^k + \alpha\beta_k\bar{\varepsilon}]^2 \\ &= (1-\alpha)^2(\varepsilon^k)^2 + 2(1-\alpha)\alpha\beta_k\varepsilon^k\bar{\varepsilon} + \alpha^2\beta_k^2\bar{\varepsilon}^2 \\ &\leq (1-\alpha)^2(\varepsilon^k)^2 + 2\alpha\beta_k\varepsilon^k\bar{\varepsilon} + O(\alpha^2) \\ &\leq (1-\alpha)^2(\varepsilon^k)^2 + 2\alpha\gamma f(z^k)^{1/2}\|H(z^k)\|\bar{\varepsilon} + O(\alpha^2) \\ &= (1-2\alpha)(\varepsilon^k)^2 + 2\alpha\gamma\bar{\varepsilon} f(z^k) + O(\alpha^2).\end{aligned} \tag{3.7}$$

Define

$$g(\alpha) := \varphi(z^k + \alpha\Delta z^k) - \varphi(z^k) - \alpha(\nabla\varphi(z^k))^T\Delta z^k.$$

Since $\varphi(\cdot)$ is continuously differentiable at any $z^k \in \Re^{n+1}$,

$$g(\alpha) = o(\alpha).$$

Then, from (3.4) and (3.2), for any $\alpha \in [0, 1]$ we have

$$\begin{aligned}\|G(z^k + \alpha\Delta z^k)\|^2 &= \varphi(z^k + \alpha\Delta z^k) \\ &= \varphi(z^k) + \alpha(\nabla\varphi(z^k))^T\Delta z^k + g(\alpha) \\ &= \varphi(z^k) - 2\alpha\varphi(z^k) + o(\alpha) \\ &= (1-2\alpha)\varphi(z^k) + o(\alpha).\end{aligned} \tag{3.8}$$

It then follows from (3.7) and (3.8) that, for all $\alpha \in [0, 1]$, we have

$$\begin{aligned}f(z^k + \alpha\Delta z^k) &= \|H(z^k + \alpha\Delta z^k)\|^2 \\ &= (\varepsilon^k + \alpha\Delta\varepsilon^k)^2 + \|G(z^k + \alpha\Delta z^k)\|^2 \\ &\leq (1-2\alpha)(\varepsilon^k)^2 + 2\alpha\gamma\bar{\varepsilon} f(z^k) + (1-2\alpha)\|G(z^k)\|^2 + o(\alpha) \\ &= (1-2\alpha) f(z^k) + 2\alpha\gamma\bar{\varepsilon} f(z^k) + o(\alpha) \\ &= [1 - 2(1-\gamma\bar{\varepsilon})\alpha] f(z^k) + o(\alpha).\end{aligned} \tag{3.9}$$

The inequality (3.9) shows that there exists a positive number $\bar{\alpha} \in (0, 1]$ such that, for all $\alpha \in [0, \bar{\alpha}]$,

$$f(z^k + \alpha \Delta z^k) \leq [1 - 2\sigma(1 - \gamma\bar{\varepsilon})\alpha] f(z^k),$$

which completes our proof. □

Proposition 3.3. *Suppose that F is a P_0-function. For each $k \geq 0$, if $\varepsilon^k \in \Re_{++}$, $z^k \in \Omega$, then for any $\alpha \in [0, 1]$ such that*

$$f(z^k + \alpha \Delta z^k) \leq [1 - 2\sigma(1 - \gamma\bar{\varepsilon})\alpha] f(z^k), \tag{3.10}$$

it holds that $z^k + \alpha \Delta z^k \in \Omega$.

Proof. We prove this proposition by considering the following two cases:

(i) $f(z^k) > 1$. Then $\beta_k = \gamma$. It therefore follows from $z^k \in \Omega$ and $\beta(z) = \gamma \min\{1, f(z)^t\} \leq \gamma$ for any $z \in \Re^{n+1}$ that, for all $\alpha \in [0, 1]$, we have

$$\begin{aligned} \varepsilon^k + \alpha\Delta\varepsilon^k - \beta(z^k + \alpha\Delta z^k)\bar{\varepsilon} &\geq (1-\alpha)\varepsilon^k + \alpha\beta_k\bar{\varepsilon} - \gamma\bar{\varepsilon} \\ &\geq (1-\alpha)\beta_k\bar{\varepsilon} + \alpha\beta_k\bar{\varepsilon} - \gamma\bar{\varepsilon} \\ &= (1-\alpha)\gamma\bar{\varepsilon} + \alpha\gamma\bar{\varepsilon} - \gamma\bar{\varepsilon} \\ &= 0. \end{aligned} \tag{3.11}$$

(ii) $f(z^k) \leq 1$. Then, for any $\alpha \in [0, 1]$ satisfying (3.10), we have

$$f(z^k + \alpha \Delta z^k) \leq [1 - 2\sigma(1 - \gamma\bar{\varepsilon})\alpha] f(z^k) \leq 1. \tag{3.12}$$

So, for any $\alpha \in [0, 1]$ satisfying (3.10),

$$\beta(z^k + \alpha\Delta z^k) = \gamma f(z^k + \alpha\Delta z^k)^t.$$

Hence, again because $z^k \in \Omega$, by using the first inequality in (3.12), for any $\alpha \in [0, 1]$ satisfying (3.10) we have

$$\begin{aligned} \varepsilon^k + \alpha\Delta\varepsilon^k - \beta(z^k + \alpha\Delta z^k)\bar{\varepsilon} &= (1-\alpha)\varepsilon^k + \alpha\beta_k\bar{\varepsilon} - \gamma f(z^k + \alpha\Delta z^k)^t\bar{\varepsilon} \\ &\geq (1-\alpha)\beta_k\bar{\varepsilon} + \alpha\beta_k\bar{\varepsilon} - \gamma[1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha]^t f(z^k)^t\bar{\varepsilon} \\ &= \beta_k\bar{\varepsilon} - \gamma[1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha]^t f(z^k)^t\bar{\varepsilon} \\ &= \gamma f(z^k)^t\bar{\varepsilon} - \gamma[1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha]^t f(z^k)^t\bar{\varepsilon} \\ &= \gamma\{1 - [1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha]^t\} f(z^k)^t\bar{\varepsilon} \\ &\geq 0. \end{aligned} \tag{3.13}$$

Thus, by combining (3.11) and (3.13), we have proved that, for all $\alpha \in [0, 1]$ satisfying (3.10),

$$z^k + \alpha\Delta z^k \in \Omega.$$

This completes our proof. □

Proposition 3.4. *Suppose that F is a P_0-function. Then Algorithm* 3.1 *is not finite and generates a sequence $\{z^k\}$ with $\varepsilon^k \in \Re_{++}$, $\{z^k\} \in \Omega$ for all k.*

Proof. First, because $z^0 = (\bar{\varepsilon}, x^0) \in \Omega$, we have from Propositions 3.2 and 3.3 that z^1 is well defined, $\varepsilon^1 \in \Re_{++}$, and $z^1 \in \Omega$. Then, by repeatedly resorting to Propositions 3.2 and 3.3, we can prove that an infinite sequence $\{z^k\}$ is generated, $\varepsilon^k \in \Re^n_{++}$, and $z^k \in \Omega$. □

4. Global Convergence

Lemma 4.1. *Suppose that F is a P_0-function and that $\{z^k = (\varepsilon^k, x^k) \in \Re \times \Re^n\}$ is an infinite sequence generated by Algorithm* 3.1. *Then, for any $k \geq 0$,*

$$0 < \varepsilon^{k+1} \leq \varepsilon^k \leq \bar{\varepsilon}. \tag{4.1}$$

Proof. First, $\varepsilon^0 = \bar{\varepsilon} > 0$. Then, from the design of Algorithm 3.1 and that, for any $z \in \Re^{n+1}$, $\beta(z) = \gamma \min\{1,\ f(z)^t\} \leq \gamma$, we have

$$\varepsilon^1 = (1 - \delta^{l_0})\varepsilon^0 + \delta^{l_0}\beta(z^0)\bar{\varepsilon} \leq (1 - \delta^{l_0})\bar{\varepsilon} + \delta^{l_0}\gamma\bar{\varepsilon} \leq \bar{\varepsilon}.$$

Hence (4.1) holds for $k = 0$. Suppose that (4.1) holds for $k = i - 1$. We now prove that (4.1) holds for $k = i$. From the design of Algorithm 3.1 we have

$$\varepsilon^{i+1} = (1 - \delta^{l_i})\varepsilon^i + \delta^{l_i}\beta(z^i)\bar{\varepsilon}.$$

Since, from Proposition 3.4, $z^i \in \Omega$, we have $\varepsilon^i \geq \beta(z^i)\bar{\varepsilon}$. Thus,

$$\varepsilon^{i+1} \leq (1 - \delta^{l_i})\varepsilon^i + \delta^{l_i}\varepsilon^i = \varepsilon^i$$

and

$$\varepsilon^{i+1} \geq (1 - \delta^{l_i})\beta(z^i)\bar{\varepsilon} + \delta^{l_i}\beta(z^i)\bar{\varepsilon} = \beta(z^i)\bar{\varepsilon} > 0.$$

So, (4.1) holds for $k = i$. We complete our proof. □

Theorem 4.1. *Suppose that F is a P_0-function. Then Algorithm* 3.1 *is not finite and generates a sequence $\{z^k\}$ with*

$$\lim_{k \to \infty} f(z^k) = 0. \tag{4.2}$$

In particular, any accumulation point $\tilde{z}$ of $\{z^k\}$ is a solution of $H(z) = 0$.

Proof. It follows from Proposition 3.4 that an infinite sequence $\{z^k\}$ is generated such that $\{z^k\} \in \Omega$. From the design of Algorithm 3.1, $f(z^{k+1}) < f(z^k)$ for all $k \geq 0$. Hence the two sequences $\{f(z^k)\}$ and $\{\beta(z^k)\}$ are monotonically decreasing. Since

$f(z^k), \beta(z^k) \geq 0$ $(k \geq 0)$, there exist $\tilde{f}, \tilde{\beta} \geq 0$ such that $f(z^k) \to \tilde{f}$ and $\beta(z^k) \to \tilde{\beta}$ as $k \to \infty$. If $\tilde{f} = 0$, then we obtain the desired result. Suppose that $\tilde{f} > 0$. Then $\tilde{\beta} > 0$. Since $\{z^k\} \in \Omega$, $\varepsilon^k \geq \beta(z^k)\bar{\varepsilon}$. It then follows from Lemma 4.1 that $\bar{\varepsilon} \geq \varepsilon^k \geq \tilde{\beta}\bar{\varepsilon}$. Hence, by Proposition 2.1, the infinite sequence $\{z^k\}$ must be bounded because otherwise $\{f(z^k)\}$ must be unbounded, which is impossible by the monotonicity decreasing property of $\{f(z^k)\}$. Then there exists at least one accumulation point $\tilde{z} = (\tilde{\varepsilon}, \tilde{x}) \in \Re \times \Re^n$ of $\{z^k\}$ such that $\tilde{\varepsilon} \in [\tilde{\beta}\bar{\varepsilon}, \bar{\varepsilon}]$. By taking a subsequence if necessary, we may assume that $\{z^k\}$ converges to $\tilde{z}$. It is easy to see that $\tilde{f} = f(\tilde{z})$, $\beta(\tilde{z}) = \tilde{\beta}$, and $\tilde{z} \in \Omega$. Then any $V \in \partial H(\tilde{z})$ are nonsingular because $\tilde{\varepsilon} > 0$ and F is a P_0-function. Hence, there exists a closed neighbourhood $\mathcal{N}(\tilde{z})$ of $\tilde{z}$ such that for any $z = (\varepsilon, x) \in \mathcal{N}(\tilde{z})$ we have $\varepsilon \in \Re_{++}$ and that all $V \in \partial H(z)$ are nonsingular. For any $z \in \mathcal{N}(\tilde{z})$ and $V \in \partial H(z)$, let $\Delta z = (\Delta\varepsilon, \Delta x) \in \Re \times \Re^n$ be the unique solution of the following equation:

$$H(z) + V\Delta z = \beta(z)\bar{z}, \tag{4.3}$$

and, for any $\alpha \in [0, 1]$, define

$$g_z(\alpha) = \varphi(z + \alpha\Delta z) - \varphi(z) - \alpha(\nabla\varphi(z))^T \Delta z.$$

From (4.3), for any $z \in \mathcal{N}(\tilde{z})$,

$$\Delta\varepsilon = -\varepsilon + \beta(z)\bar{\varepsilon}.$$

Then, for all $\alpha \in [0, 1]$ and all $z \in \mathcal{N}(\tilde{z})$,

$$\varepsilon + \alpha\Delta\varepsilon = (1 - \alpha)\varepsilon + \alpha\beta(z)\bar{\varepsilon} \in \Re_{++}. \tag{4.4}$$

It follows from the Mean Value Theorem that

$$g_z(\alpha) = \alpha \int_0^1 [\varphi'(z + \theta\alpha\Delta z) - \varphi'(z)]\Delta z \, d\theta.$$

Since $\varphi'(\cdot)$ is uniformly continuous on $\mathcal{N}(\tilde{z})$ and $\{\Delta z\}$ is bounded for all $z \in \mathcal{N}(\tilde{z})$ (refer to Proposition 3.1 of [33] for a proof),

$$\lim_{\alpha \downarrow 0} \|g_z(\alpha)\|/\alpha = 0.$$

Then, from (4.4), (4.3), and the fact that $\beta(z) \leq \gamma f(z)^{1/2}$, for all $\alpha \in [0, 1]$ and all $z \in \mathcal{N}(\tilde{z})$, we have

$$\begin{aligned}
(\varepsilon + \alpha\Delta\varepsilon)^2 &= [(1 - \alpha)\varepsilon + \alpha\beta(z)\bar{\varepsilon}]^2 \\
&= 1 - \alpha)^2\varepsilon^2 + 2(1 - \alpha)\alpha\beta(z)\varepsilon\bar{\varepsilon} + \alpha^2\beta(z)^2\bar{\varepsilon}^2 \\
&\leq (1 - \alpha)^2\varepsilon^2 + 2\alpha\beta(z)\varepsilon\bar{\varepsilon} + O(\alpha^2) \\
&\leq (1 - \alpha)^2\varepsilon^2 + 2\alpha\gamma f(z)^{1/2}\|H(z)\|\bar{\varepsilon} + O(\alpha^2) \\
&= (1 - \alpha)^2\varepsilon^2 + 2\alpha\gamma\bar{\varepsilon} f(z) + O(\alpha^2)
\end{aligned} \tag{4.5}$$

and

$$\|G(z + \alpha\Delta z)\|^2 = \varphi(z + \alpha\Delta z)$$

$$
\begin{aligned}
&= \varphi(z) + \alpha(\nabla\varphi(z))^T \Delta z + g_z(\alpha) \\
&= \varphi(z) - 2\alpha\varphi(z) + o(\alpha) \\
&= (1-2\alpha)\varphi(z) + o(\alpha). \qquad (4.6)
\end{aligned}
$$

It then follows from (4.5) and (4.6) that, for all $\alpha \in [0, 1]$ and all $z \in \mathcal{N}(\tilde{z})$, we have

$$
\begin{aligned}
f(z+\alpha\Delta z) &= \|H(z+\alpha\Delta z)\|^2 \\
&= (\varepsilon + \alpha\Delta\varepsilon)^2 + \varphi(z+\alpha\Delta z) \\
&\leq (1-\alpha)^2\varepsilon^2 + 2\alpha\gamma\bar{\varepsilon} f(z) + (1-\alpha)^2\varphi(z) + o(\alpha) + O(\alpha^2) \\
&= (1-\alpha)^2 f(z) + 2\alpha\gamma\bar{\varepsilon} f(z) + o(\alpha) \\
&= (1-2\alpha) f(z) + 2\alpha\gamma\bar{\varepsilon} f(z) + o(\alpha) \\
&= [1 - 2(1-\gamma\bar{\varepsilon})\alpha] f(z) + o(\alpha). \qquad (4.7)
\end{aligned}
$$

Then from inequality (4.7) we can find a positive number $\bar{\alpha} \in (0, 1]$ such that, for all $\alpha \in [0, \bar{\alpha}]$ and all $z \in \mathcal{N}(\tilde{z})$,

$$f(z+\alpha\Delta z) \leq [1 - 2\sigma(1-\gamma\bar{\varepsilon})\alpha] f(z).$$

Therefore, for a nonnegative integer l such that $\delta^l \in (0, \bar{\alpha}]$, we have

$$f(z^k + \delta^l \Delta z^k) \leq [1 - 2\sigma(1-\gamma\bar{\varepsilon})\delta^l] f(z)$$

for all sufficiently large k. Then, for every sufficiently large k, we see that $l^k \leq l$ and hence $\delta^{l_k} \geq \delta^l$. Therefore,

$$f(z^{k+1}) \leq [1 - 2\sigma(1-\gamma\bar{\varepsilon})\delta^{l_k}] f(z^k) \leq [1 - 2\sigma(1-\gamma\bar{\varepsilon})\delta^l] f(z^k)$$

for all sufficiently large k. This contradicts the fact that the sequence $\{f(z^k)\}$ converges to $\tilde{f} > 0$. This contradiction shows that (4.2) must hold. In particular, if there exists an accumulation point $\tilde{z}$ of $\{z^k\}$, then, by the continuity of f, $f(\tilde{z}) = 0$, and so $H(\tilde{z}) = 0$. So, we complete our proof. □

Theorem 4.1 shows that (4.2) holds if F is a P_0-function and any accumulation point $\tilde{z}$ of $\{z^k\}$ is a solution of NCP(F). This does not mean that $\{z^k\}$ has an accumulation point. Apparently, if NCP(F) has no solution, $\{z^k\}$ cannot have an accumulation point. In order to make $\{z^k\}$ have an accumulation point we at least need to assume that the solution set of NCP(F) is nonempty.

Theorem 4.2. *Suppose that F is a P_0-function and that the solution set $\mathcal{S}$ of* NCP(F) *is nonempty and bounded. Then the infinite sequence $\{z^k\}$ generated by Algorithm* 3.1 *is bounded and any accumulation point of $\{z^k\}$ is a solution of $H(z) = 0$.*

Proof. From Theorem 4.1, $\lim_{k\to\infty} f(z^k) = 0$. Then $\varepsilon^k \to 0$ and $\varphi(\varepsilon^k, x^k) \to 0$ as $k \to \infty$. By Lemma 4.1, for any $k \geq 0$, $\varepsilon^k > 0$. Then by applying Lemma 2.3 we obtain that $\{x^k\}$ is bounded and any accumulation point of $\{x^k\}$ is a solution of NCP(F). Since

$\lim_{k\to\infty} \varepsilon^k = 0$, we have in fact proved that $\{z^k\}$ is bounded and any accumulation point of $\{z^k\}$ is a solution of $H(z) = 0$. □

Corollary 4.1. *Suppose that F is a monotone function and that there exists a strictly feasible point $\hat{x}$, i.e., $\hat{x} > 0$, $F(\hat{x}) > 0$. Then the infinite sequence $\{z^k\}$ generated by Algorithm* 3.1 *is bounded and any accumulation point of $\{z^k\}$ is a solution of $H(z) = 0$.*

Proof. Since NCP(F) has a strictly feasible point, its solution set $\mathcal{S}$ is nonempty and bounded [17, Theorem 3.4]. By noting that a monotone function is always a P_0–function, we obtain the results of this theorem by Theorem 4.2. □

In Theorem 4.2 we did not state whether the sequence $\{z^k\}$ is bounded or not if the solution set of NCP(F) is nonempty but unbounded. However, if F is monotone, we can have such a result.

Theorem 4.3. *Suppose that F is a monotone function and in Algorithm* 3.1 *the parameter t is set to be $\frac{1}{2}$, i.e, $\beta(z) = \gamma \min\{1,\ f(z)^{1/2}\}$. Then, if the iteration sequence $\{z^k\}$ is bounded, the solution set $\mathcal{S}$ of* NCP(F) *is nonempty*; *conversely, if the solution set $\mathcal{S}$ is nonempty and* (2.5) *in Proposition* 2.2 *holds, the iteration sequence $\{z^k\}$ is bounded and any accumulation point of $\{z^k\}$ is a solution of $H(z) = 0$.*

Proof. First, we suppose that $\{z^k\}$ is bounded. Then there exists at least one accumulation point, say $z^* = (\varepsilon^*, x^*) \in \Re \times \Re^n$. By Theorem 4.1, z^* is a solution of $H(z) = 0$. Then $\varepsilon^* = 0$ and x^* is a solution of NCP(F). So, $\mathcal{S}$ is nonempty.

Conversely, suppose that the solution set $\mathcal{S}$ is nonempty. For each $k \geq 0$, let $x(\varepsilon^k)$ be a solution of NCP(F_{ε^k}), then $x(\varepsilon^k)$ converges to $x^* = \arg\min_{x\in\mathcal{S}} \|x\|$. In Theorem 4.1, we have proved that $\lim_{k\to\infty} \|H(\varepsilon^k, x^k)\| = 0$. Then there exists a $\bar{k} \geq 0$ such that, for all $k \geq \bar{k}$, we have

$$\varepsilon^k \geq \beta(z^k)\bar{\varepsilon} = \gamma f(z^k)^{1/2}\bar{\varepsilon} = \gamma[(\varepsilon^k)^2 + \|G(\varepsilon^k, x^k)\|^2]^{1/2}\bar{\varepsilon}.$$

Hence,

$$\|G(\varepsilon^k, x^k)\| \leq \frac{\sqrt{1-\gamma^2\bar{\varepsilon}^2}}{\gamma\bar{\varepsilon}}\varepsilon^k = C\varepsilon^k$$

with

$$C := \frac{\sqrt{1-\gamma^2\bar{\varepsilon}^2}}{\gamma\bar{\varepsilon}}.$$

Thus by Proposition 2.2 the sequence $\{x^k\}$ must be bounded. Then, because $\lim_{k\to\infty} f(z^k) = 0$ by Theorem 4.1, $\varepsilon^k \to 0$ as $k \to \infty$, we conclude that the sequence $\{z^k\}$ is bounded and from Theorem 4.1 any accumulation point of $\{z^k\}$ is a solution of $H(z) = 0$. □

5. Superlinear and Quadratic Convergence

In this section we discuss the superlinear and quadratic convergence of Algorithm 3.1 by assuming that there is a locally unique solution. By a recent result [7], [16] this, under the assumption that F is a P_0-function, is equivalent to saying that NCP(F) has only one solution, and, thus, $\mathcal{S}$ is bounded. We then in this section assume the parameter $t = 1$, i.e., $\beta(z) = \gamma \min\{1, f(z)\}$.

Theorem 5.1. *Suppose that F is a P_0-function and that the solution set $\mathcal{S}$ of* NCP(F) *is nonempty and bounded. Suppose that $z^* := (\varepsilon^*, x^*) \in \Re \times \Re^n$ is an accumulation point of the infinite sequence $\{z^k\}$ generated by Algorithm* 3.1 *and that all $V \in \partial H(z^*)$ are nonsingular. Then the whole sequence $\{z^k\}$ converges to z^*,*

$$\|z^{k+1} - z^*\| = o(\|z^k - z^*\|) \tag{5.1}$$

and

$$\varepsilon^{k+1} = o(\varepsilon^k). \tag{5.2}$$

Furthermore, if F' is locally Lipschitz around x^*, then

$$\|z^{k+1} - z^*\| = O(\|z^k - z^*\|^2) \tag{5.3}$$

and

$$\varepsilon^{k+1} = O(\varepsilon^k)^2. \tag{5.4}$$

Proof. First, from Theorem 4.1, z^* is a solution of $H(z) = 0$. Then, from Proposition 3.1 of [33], for all z sufficiently close to z^* and for all $V \in \partial H(z)$,

$$\|V^{-1}\| = O(1).$$

Under the assumptions, from Lemma 2.6 we know that H is semismooth (strongly semismooth, respectively) at z^*. Then, from Lemma 2.5 for z^k sufficiently close to z^*, we have

$$\begin{aligned} \|z^k + \Delta z^k - z^*\| &= \|z^k + V_k^{-1}[-H(z^k) + \beta_k \bar{z}] - z^*\| \\ &= O(\|H(z^k) - H(z^*) - V_k(z^k - z^*)\| + \beta_k \bar{\varepsilon}) \\ &= o(\|z^k - z^*\|) + O(f(z^k)) \quad (= O(\|z^k - z^*\|^2) + O(f(z^k))). \end{aligned} \tag{5.5}$$

Then, because H is locally Lipschitz continuous around z^*, for all z^k close to z^*,

$$f(z^k) = \|H(z^k)\|^2 = O(\|z^k - z^*\|^2). \tag{5.6}$$

Therefore, from (5.5) and (5.6), because H is semismooth (strongly semismooth, respectively) at z^*, for all z^k sufficiently close to z^*,

$$\|z^k + \Delta z^k - z^*\| = o(\|z^k - z^*\|) \qquad (= O(\|z^k - z^*\|^2)). \tag{5.7}$$

By following the proof of Theorem 3.1 of [28], for all z^k sufficiently close to z^*, we have

$$\|z^k - z^*\| = O(\|H(z^k) - H(z^*)\|). \tag{5.8}$$

Hence, because H is semismooth (strongly semismooth, respectively) at z^*, for all z^k sufficiently close to z^*, we have

$$\begin{aligned} f(z^k + \Delta z^k) &= \|H(z^k + \Delta z^k)\|^2 \\ &= O(\|z^k + \Delta z^k - z^*\|^2) \\ &= o(\|z^k - z^*\|^2) \qquad (= O(\|z^k - z^*\|^4)) \\ &= o(\|H(z^k) - H(z^*)\|^2) \qquad (= O(\|H(z^k) - H(z^*)\|^4)) \\ &= o(f(z^k)) \qquad (= O(f(z^k)^2)). \end{aligned} \tag{5.9}$$

Therefore, for all z^k sufficiently close to z^* we have

$$z^{k+1} = z^k + \Delta z^k,$$

which, together with (5.7), proves (5.1), and if F' is locally Lipschitz around x^*, proves (5.3).

Next, from the definition of β_k and the fact that $z^k \to z^*$ as $k \to \infty$, for all k sufficiently large,

$$\beta_k = \gamma f(z^k) = \gamma \|H(z^k)\|^2.$$

Also, because, for all k sufficiently large, $z^{k+1} = z^k + \Delta z^k$, we have for all k sufficiently large that

$$\varepsilon^{k+1} = \varepsilon^k + \Delta\varepsilon^k = \beta_k \bar{\varepsilon}.$$

Hence, for all k sufficiently large,

$$\varepsilon^{k+1} = \gamma \|H(z^k)\|^2 \bar{\varepsilon},$$

which, together with (5.1), (5.6), and (5.8), gives

$$\lim_{k\to\infty} \frac{\varepsilon^{k+1}}{\varepsilon^k} = \lim_{k\to\infty} \frac{\|H(z^k)\|^2}{\|H(z^{k-1})\|^2} = \lim_{k\to\infty} \frac{\|H(z^k) - H(z^*)\|^2}{\|H(z^{k-1}) - H(z^*)\|^2} = 0.$$

This proves (5.2). If F' is locally Lipschitz around x^*, then from the above argument we can easily get (5.4). So, we complete our proof. □

In Theorem 5.1 we assumed that all $V \in \partial H(z^*)$ are nonsingular at a solution point z^* of $H(z) = 0$. Next, we give a sufficient condition such that this assumption is satisfied. Let $z^* = (\varepsilon^*, x^*) \in \Re \times \Re^n$ be a solution point of $H(z) = 0$. Then, apparently, $\varepsilon^* = 0$ and x^* is a solution of NCP(F). For convenience of handling notation we denote

$$\mathcal{I} := \{i \,|\, 0 < x_i^* \ \&\ F_i(x^*) = 0,\ i \in N\},$$

$$\mathcal{J} := \{i \mid x_i^* = 0 \ \&\ F_i(x^*) = 0,\ i \in N\},$$

and

$$\mathcal{K} := \{i \mid x_i^* = 0 \ \&\ F_i(x^*) > 0,\ i \in N\}.$$

Then

$$\mathcal{I} \cup \mathcal{J} \cup \mathcal{K} = N.$$

By rearrangement we assume that $F'(x^*)$ can be rewritten as

$$F'(x^*) = \begin{pmatrix} F'(x^*)_{\mathcal{I}\mathcal{I}} & F'(x^*)_{\mathcal{I}\mathcal{J}} & F'(x^*)_{\mathcal{I}\mathcal{K}} \\ F'(x^*)_{\mathcal{J}\mathcal{I}} & F'(x^*)_{\mathcal{J}\mathcal{J}} & F'(x^*)_{\mathcal{J}\mathcal{K}} \\ F'(x^*)_{\mathcal{K}\mathcal{I}} & F'(x^*)_{\mathcal{K}\mathcal{J}} & F'(x^*)_{\mathcal{K}\mathcal{K}} \end{pmatrix}.$$

The nonlinear complementarity problem is said to be R-regular at x^* if $F'(x^*)_{\mathcal{I}\mathcal{I}}$ is nonsingular and its Schur-complement in the matrix

$$\begin{pmatrix} F'(x^*)_{\mathcal{I}\mathcal{I}} & F'(x^*)_{\mathcal{I}\mathcal{J}} \\ F'(x^*)_{\mathcal{J}\mathcal{I}} & F'(x^*)_{\mathcal{J}\mathcal{J}} \end{pmatrix}$$

is a P-matrix, see [34].

Proposition 5.1. *Suppose that $z^* = (\varepsilon^*, x^*) \in \Re \times \Re^n$ is a solution of $H(z) = 0$. If the nonlinear complementarity problem is R-regular at x^*, then all $V \in \partial H(z^*)$ are nonsingular.*

Proof. It is easy to see that for any $V \in \partial H(z^*)$ there exists a $W = (W_\varepsilon \quad W_x) \in \partial G(z^*)$ with $W_\varepsilon \in \Re^n$ and $W_x \in \Re^{n\times n}$ such that

$$V = \begin{pmatrix} 1 & 0 \\ W_\varepsilon & W_x \end{pmatrix}.$$

Hence, proving V is nonsingular is equivalent to proving W_x is nonsingular. Since $\varepsilon^* = 0$, by using standard analysis (see, e.g., Proposition 3.2 of [10]), we can prove that all such generated matrices W_x are nonsingular. Then we complete our proof. ☐

The above proposition shows that all the conclusions of Theorem 5.1 hold if the assumption that all $V \in \partial H(z^*)$ are nonsingular is replaced by that the nonlinear complementarity problem is R-regular at x^*.

6. Numerical Results

In this section we present some numerical experiments for the nonmonotone line search version of Algorithm 3.1: Step 3 is replaced by

Step 3′. Let l_k be the smallest nonnegative integer l satisfying

$$z^k + \delta^l \Delta z^k \in \Omega \tag{6.1}$$

and

$$f(z^k + \delta^l \Delta z^k) \le \mathcal{W} - 2\sigma(1 - \gamma\bar{\varepsilon})\delta^l f(z^k), \tag{6.2}$$

where $\mathcal{W}$ is any value satisfying

$$f(z^k) \le \mathcal{W} \le \max_{j=0,1,\ldots,M^k} f(z^{k-j})$$

and M^k are nonnegative integers bounded above for all k such that $M^k \le k$. Define $z^{k+1} := z^k + \delta^{l_k} \Delta z^k$.

Remark. (i) We choose a nonmonotone line search here because in most cases it increases the stability of algorithms.

(ii) The requirement (6.1) is for guaranteeing the global convergence of the algorithm. This requirement automatically holds for our algorithm with a monotone line search, see Proposition 3.3. The consistency between (6.1) and (6.2) can be seen clearly from Propositions 3.2 and 3.3.

In the implementation we choose $\mathcal{W}$ as follows:

(1) Set $\mathcal{W} = f(z^0)$ at the beginning of the algorithm.
(2) Keep the value of $\mathcal{W}$ fixed as long as

$$f(z^k) \le \min_{j=0,1,\ldots,5} f(z^{k-j}). \tag{6.3}$$

(3) If (6.3) is not satisfied at the kth iteration, set $\mathcal{W} = f(z^k)$.

For a detailed description of the above nonmonotone line search technique and its motivation, see [4].

The above algorithm was implemented in Matlab and run on a DEC Alpha Server 8200. Throughout the computational experiments, the parameters used in the algorithm were $\delta = 0.5$, $\sigma = 0.5 \times 10^{-4}$, $t = 1$, $\bar{\varepsilon} = 1$, and $\gamma = 0.2$. We used $f(z) \le 10^{-12}$ as the stopping rule. The numerical results are summarized in Table 1 for different problems tested. In Table 1, Dim denotes the number of the variables in the problem, Start. point denotes the starting point, Iter denotes the number of iterations, which is also equal to the number of Jacobian evaluations for the function F, NF denotes the number of function evaluations for the function F, and FF denotes the value of f at the final iterate. In the following, we give a brief description of the tested problems. The source reported for the problem is not necessarily the original one.

Problem 1. This is the Kojima-Shindo problem, see [26]. F is not a P_0-function. This problem has two solutions: $x^1 = (\sqrt{6}/2, 0, 0, 0.5)$ and $x^2 = (1, 0, 3, 0)$.
Starting points: (a) $(1, 1, 1, 1)$, (b) $(-1, -1, -1, -1)$, (c) $(0, 0, 0, 0)$.

Table 1. Numerical results for the algorithm.

Problem	Dim.	Start. point	Iter	NF	FF
Problem 1	4	a	8	13	2.2×10^{-23}
	4	b	10	15	2.2×10^{-13}
	4	c	fail		
Problem 2	10,000	a	9	11	4.3×10^{-16}
	10,000	b	9	12	1.3×10^{-13}
Problem 3	5	a	22	23	2.3×10^{-18}
	5	b	19	23	1.5×10^{-25}
Problem 4	4	a	5	6	1.5×10^{-16}
	4	b	5	6	1.5×10^{-16}
Problem 5	10	a	9	10	3.2×10^{-13}
Problem 6a	4	a	6	9	8.2×10^{-13}
	4	b	5	6	4.3×10^{-13}
	4	c	6	9	8.2×10^{-13}
Problem 6b	4	a	8	12	2.5×10^{-15}
	4	b	6	7	4.6×10^{-19}
	4	c	8	12	2.5×10^{-15}
Problem 7	4	a	20	117	3.4×10^{-24}
	4	b	7	12	4.2×10^{-16}
Problem 8	42	a	13	15	2.7×10^{-17}
	42	b	12	16	1.5×10^{-14}
Problem 9	50	a	27	68	6.6×10^{-13}
	50	b	29	67	6.6×10^{-13}
Problem 10	1,000	a	8	14	6.5×10^{-18}
Problem 11	106	a	31	61	5.6×10^{-13}

Problem 2. This is a linear complementarity problem. See the first example of [20] for the data.
Starting points: (a) $(0, 0, 0, 0)$, (b) $(1, 1, 1, 1)$.

Problem 3. This is the fourth example of [42]. This problem represents the *KKT* conditions for a convex programming problem involving exponentials. The resulting F is monotone on the positive orthant but not even P_0 on R^n.
Starting points: (a) $(0, 0, \ldots, 0)$, (b) $(1, 1, \ldots, 1)$.

Problem 4. This is a modification of the Mathiesen example of a Walrasian equilibrium model as suggested in [21]. F is not defined everywhere and does not belong to any known class of functions.
Starting points: (a) $(0, 0, 0, 0)$, (b) $(1, 1, 1, 1)$.

Problem 5. This is the Nash–Cournot production problem [26]. F is not twice continuously differentiable. F is a P-function on the strictly positive orthant.
Starting point: (a) $(1, 1, 1, 1)$.

Problem 6. This is a Mathiesen equilibrium problem [23], [26], in which F is not defined everywhere. Two set of constants were used: $(\alpha, b_2, b_3) = (0.75, 1, 0.5)$ and

$(\alpha, b_2, b_3) = (0.9, 5, 3)$. We use Problem 6a and 6b to represent this problem with these two set of constants, respectively.
Starting points: (a) $(1, 1, 1, 1)$, (b) $(0.5, 0.5, 0.5, 0.5)$, (c) $(0, 0, 0, 0)$.

Problem 7. This is the Kojima–Josephy problem, see [4]. F is not a P_0-function. The problem has a unique solution which is not R-regular.
Starting points: (a) $(0, 0, 0, 0)$, (b) $(1, 1, 1, 1)$.

Problem 8. This is a problem arising from a spatial equilibrium model, see [26]. F is a P-function and the unique solution is R-regular.
Starting points: (a) $(0, 0, \dots, 0)$, (b) $(1, 1, \dots, 1)$.

Problem 9. This is a traffic equilibrium problem with elastic demand, see [26].
Starting points: (a) All the components are 0 except x_1, x_2, x_3, x_{10}, x_{11}, x_{20}, x_{21}, x_{22}, x_{29}, x_{30}, x_{40}, x_{45} which are 1, x_{39}, x_{42}, x_{43}, x_{46} which are 7, x_{41}, x_{47}, x_{48}, x_{50} which are 6, and x_{44} and x_{49} which are 10, (b) $(0, 0, \dots, 0)$.

Problem 10. This is the third problem of [42], which is a linear complementarity problem with $F(x) = Mx + q$. M is not even semimonotone and none of the standard algebraic techniques can solve it. Let q be the vector with -1 in the eighth coordinate and zeros elsewhere. The continuation method of [42] fails on this problem.
Starting point: (a) $(1, 1, \dots, 1)$.

Problem 11. This is the 106-variable Von Thünen problem [26], [45]. This problem is a challenge to the algorithms designed in the literature for solving nonlinear complementarity problems. The data of this problem was down-loaded from Paul Tseng's home page http://www.math.washington.edu/~tseng/, where the data was originally obtained from Jong-Shi Pang.
Starting point: (a) $(100, 100, \dots, 100)$,

The numerical results reported in Table 1 showed that the algorithm proposed in this paper works well for Problems 1–10. For the challenging Problem 11, things seem complicated because it was observed during the process of computation that some elements of the iteration sequence become negative such that the function may take complex values. In such cases we still allowed our algorithm to continue to see what would happen. Surprisingly, our algorithm stopped in a relatively small number of iterations with a very small residue. We then checked the approximate solution xs obtained from our algorithm and this time, not surprisingly, we found that it is not a real number but one with a relatively small imaginary part. Let $\mathrm{R}(xs)$ denote the real part of xs and denote a new point xs' by

$$xs'_i = \max\{10^{-20}, \mathrm{R}(xs)_i\}, \qquad i = 1, 2, \dots, 106. \tag{6.4}$$

This new point is a very good approximation to the solution with

$$\|\min\{xs', F(xs')\}\|_\infty \le 1.3 \times 10^{-9}.$$

This is the first time that such a good approximation point is obtained for the 106-variable Von Thünen problem. All the components of xs' are around $10^{-8} \sim 10^{-20}$ except $xs'_{101} = 0.0032834$, $xs'_{102} = 0.0231250$, $xs'_{103} = 0.0133340$, $xs'_{104} = 0.0061591$, $xs'_{105} = 0.0033426$, and $xs'_{106} = 0.0019377$. If in (6.4) 10^{-20} is replaced by 10^{-15}, then we can only get a point xs' with

$$\|\min\{xs', F(xs')\}\|_\infty \leq 6.8 \times 10^{-5}.$$

This shows that the 106-variable Von Thünen problem is very sensitive to the change of variable and its implementation for various algorithms must be preprocessed considerably.

7. Conclusions

In this paper we constructed a regularization Newton method for solving nonlinear complementarity problems under the assumption that F is a P_0-function by using the recent developments on regularization methods and smoothing methods. The convergence results discussed in this paper are very favorable. Even stronger results have been obtained for monotone complementarity problems. The numerical results showed that our algorithm works well for the problems tested. With regard to the nice theoretical results of our algorithm, the computational results reported are very encouraging. We expect our algorithm can also solve large-scale problems well.

By utilizing some box constrained variational inequality problem (BVIP) functions (see [29] for several interesting BVIP functions), the approach developed in this paper can also be used to solve the BVIP: find $x^* \in X$ such that

$$(x - x^*)^T F(x^*) \geq 0 \qquad \text{for all} \quad x \in X, \tag{7.1}$$

where $X := \{x \in \Re^n |\ l \leq x \leq u\}$, $l \in \{\Re \cup \{-\infty\}\}^n$, $u \in \{\Re \cup \{\infty\}\}^n$, and $l < u$. See [25] and [11] for various applications of BVIPs along with nonlinear complementarity problems.

There are a few nonlinear complementarity and variational inequality problems in which the mapping F is not well defined outside its feasible region [11]. Then our algorithm for these problems is not well defined on the whole space $\Re^n$. However, the approach developed in [32] on smoothing methods provided a way to circumvent this difficulty. We leave these as future research topics.

Acknowledgments

The author would like to thank Guanglu Zhou for his helpful comments on an earlier version of this paper and the referee for his careful reading and detailed comments.

References

1. B. Chen and X. Chen, A global and local superlinear continuation-smoothing method for $P_0 + R_0$ and monotone NCP, Preprint, Department of Management and Systems, Washington State University, Pullman, WA, May 1997. Revised November 1997. SIAM Journal on Optmization, to appear.

2. F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.
3. R.W. Cottle, J.-S. Pang, and R.E. Stone, The Linear Complementarity Problem, Academic Press, Boston, 1992.
4. T. De Luca, F. Facchinei, and C. Kanzow, A semismooth equation approach to the solution of nonlinear complementarity problems, Mathematical Programming, 75 (1996), 407–439.
5. A.L. Dontchev and T. Zolezzi, Well-Posed Optimization Problems, Lecture Notes in Mathematics 1543, Springer-Verlag, Berlin, 1993.
6. J. Eckstein and M.C. Ferris, Smooth methods of multipliers for complementarity problems, RUCTOR Research Report 27-96, Faculty of Management and RUTCOR, Rutgers University, New Brunswick, NJ 08903, February 1997.
7. F. Facchinei, Structural and stability properties of P_0 nonlinear complementarity problems, Mathematics of Operations Research, 23 (1998), 735–745.
8. F. Facchinei, A. Fischer, and C. Kanzow, Regularity properties of a semismooth equation reformulation of variational inequalities, SIAM Journal on Optimization, 8 (1998), 850–869.
9. F. Facchinei and C. Kanzow, Beyond monotonicity in regularization methods for nonlinear complementarity problems, DIS Working Paper 11-97, Università di Roman "La Sapienza," Rome, May 1997. SIAM Journal on Control and Optimization, to appear.
10. F. Facchinei and J. Soares, A new merit function for nonlinear complementarity problems and a related algorithm, SIAM Journal on Optimization, 7 (1997), 225–247.
11. M.C. Ferris and J.-S. Pang, Engineering and economic applications of complementarity problems, SIAM Review, 39 (1997), 669–713.
12. A. Fischer, A special Newton-type optimization method, Optimization, 24 (1992), 269–284.
13. A. Fischer, An NCP-function and its use for the solution of complementarity problems, in Recent Advances in Nonsmooth Optimization, D. Du, L. Qi, and R. Womersley, eds., World Scientific, NJ, pp. 88–105, 1995.
14. A. Fischer, Solution of monotone complementarity problems with locally Lipschitzian functions, Mathematical Programming, 76 (1997), 513–532.
15. M. Fukushima, Merit functions for variational inequality and complementarity problems, in Nonlinear Optimization and Applications, G. Di Pillo and F. Giannessi, eds., Plenum, New York, pp. 155–170, 1996.
16. M.S. Gowda and R. Sznajder, "Weak univalence and connectedness of inverse images of continuous functions", Preprint, Department of Mathematics & Statistics, University of Maryland Baltimore County, Baltimore, MD 21250, April 1997.
17. P.T. Harker and J.-S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problem: a survey of theory, algorithms and applications, Mathematical Programming, 48 (1990), 161–220.
18. K. Hotta and A. Yoshise, Global convergence of a class of non-interior-point algorithms using Chen–Harker–Kanzow functions for nonlinear complementarity problems, Discussion Paper Series No. 708, Institute of Policy and Planning Sciences, University of Tsukuba, Tsukuba, Ibaraki 305, December 1996.
19. H. Jiang, Smoothed Fischer-Burmeister equation methods for the nonlinear complementarity problem, Preprint, Department of Mathematics, University of Melbourne, Victoria 3052, June 1997.
20. H. Jiang and L. Qi, A new nonsmooth equation approach to nonlinear complementarity problems, SIAM Journal on Control and Optimization, 35 (1997), 178–193.
21. C. Kanzow, An unconstrained optimization technique for large-scale linearly constrained convex minimization problems, Computing, 53 (1994), 101–117.
22. C. Kanzow, Global convergence properties of some iteration methods for linear complementarity problems, SIAM Journal on Optimization, 6 (1996), 326–341.
23. L. Mathiesen, An algorithm based on a sequence of linear complementarity problems applied to a Walrasian equilibrium model: an example, Mathematical Programming, 37 (1987), 1–18.
24. R. Mifflin, Semismooth and semiconvex functions in constrained optimization, SIAM Journal on Control and Optimization, 15 (1977), 957–972.
25. J.–S. Pang, Complementarity problems, in Handbook of Global Optimization, R. Horst and P. Pardalos, eds., Kluwer, Boston, pp. 271–338, 1995.
26. J.-S. Pang and S.A. Gabriel, NE/SQP: a robust algorithm for the nonlinear complementarity problem, Mathematical Programming, 60 (1993), 295–337.
27. B.T. Polyak, Introduction to Optimization, Optimization Software Inc., New York, 1987.

28. L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, Mathematics of Operations Research, 18 (1993), 227–244.
29. L. Qi, Regular pseudo-smooth NCP and BVIP functions and globally and quadratically convergent generalized Newton methods for complementarity and variational inequality problems, Applied Mathematics Report AMR 97/14, School of Mathematics, University of New South Wales, Sydney 2052, NSW, June 1997. To appear in Mathematics of Operations Research.
30. L. Qi and H. Jiang, Semismooth Karush–Kuhn–Tucker equations and convergence analysis of Newton and quasi-Newton methods for solving these equations, Mathematics of Operations Research, 22 (1997), 301–325.
31. L. Qi and D. Sun, Improving the convergence of non-interior point algorithms for nonlinear complementarity problems, Applied Mathematics Report AMR 97/11, School of Mathematics, University of New South Wales, Sydney 2052, NSW, May 1997. Revised February 1998. Mathematics of Computation, to appear.
32. L. Qi, D. Sun and G. Zhou, A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities, Applied Mathematics Report AMR 97/13, School of Mathematics, University of New South Wales, Sydney 2052, NSW, June 1997.
33. L. Qi and J. Sun, A nonsmooth version of Newton's method, Mathematical Programming, 58 (1993), 353–367.
34. S.M. Robinson, Generalized equations, in Mathematical Programming: The State of the Art, A. Bachem, M. Grötschel and B. Korte, eds., Springer-Verlag, Berlin, pp. 346–347, 1983.
35. R.T. Rockafellar, Monotone operators and the proximal point algorithms, SIAM Journal on Control and Optimization, 14 (1976), 877–898.
36. R.T. Rockafellar, Monotone operators and augmented Lagrangian methods in nonlinear programming, in Nonlinear Programming 3, O. L. Mangasarian, R. R. Meyer, and S. M. Robinson, eds., Academic Press, New York, pp. 1–25, 1978.
37. P.K. Subramanian, A note on least two norm solutions of monotone complementarity problems, Technical Report 2844, Mathematical Research Center, University of Wisconsin, Madison, WI, July 1985.
38. D. Sun and R.S. Womersley, A new unconstrained differentiable merit function for box constrained variational inequality problems and a damped Gauss–Newton method, Applied Mathematics Report AMR 96/37, School of Mathematics, University of New South Wales, Sydney 2052, NSW, December 1996. Revised September 1997. SIAM Journal on Optimization, to appear.
39. P. Tseng, Global behaviour of a class of merit functions for the nonlinear complementarity problem, Journal of Optimization Theory and Applications, 89 (1996), 17–37.
40. P. Tseng, An infeasible path-following method for the monotone complementarity problem", SIAM Journal on Optimization, 7 (1997), 386–402.
41. V. Venkateswaran, An algorithm for the linear complementarity problem with a P_0-matrix, SIAM Journal on Matrix Analysis and Applications, 14 (1993), 967–977.
42. L.T. Watson, Solving the nonlinear complementarity problem by a homotopy method, SIAM Journal on Control and Optimization, 17 (1979), 36–46.
43. S. Wright, Primal–Dual Interior Point Methods, SIAM, Philadelphia, PA, 1997.
44. S. Wright and D. Ralph, A superlinear infeasible-interior-point algorithm for monotone complementarity problems, Mathematics of Operations Research, 21 (1996), 815–838.
45. B. Xiao and P.T. Harker, A nonsmooth Newton method for variational inequalities, II: Numerical results, Mathematical Programming, 65 (1994), 195-216.
46. N. Yamashita and M. Fukushima, Modified Newton methods for solving semismooth reformulations of monotone complementarity problems, Mathematical Programming, 76 (1997), 469–491.
47. E.H. Zarantonello, Projections on convex sets in Hilbert space and spectral theory, in Contributions to Nonlinear Functional Analysis, E. H. Zarantonello, ed., Academic Press, New York, pp. 237-424, 1971.

Accepted 25 *March* 1998