Newton–Type Methods for Variational Inequalities *

Jiye Han and Defeng Sun †
Institute of Applied Mathematics, Chinese Academy of Sciences
Beijing 100080, P. R. China.
Email: jyhan%AMath3@amath3.amt.ac.cn

February 1996; Revised August 1996

Abstract. Josephy's Newton and quasi–Newton methods are the basic methods for solving variational inequalities (VI). The subproblem needed to solve is a linear variational inequality problem with the same constraint to the original problem. In this paper, we provide such Newton and quasi–Newton methods for solving VI that the subproblem needed to solve is a linear system. This generalizes the previous results of the authors for nonlinear complementarity problem (NCP) and variational inequalities with polyhedral set constraints. Moreover, we provide globally and superlinearly (quadratically) convergent hybrid Newton methods for solving VI.

Key Words. Variational inequalities, Newton method, quasi–Newton method, Q–superlinear convergence.

Abbreviated Title: Newton–type Methods for VI

*This work is supported by the National Natural Science Foundation of China.

†The present address: School of Mathematics, The University of New South Wales, Sydney 2052, Australia. Email: sun@maths.unsw.edu.au

1. Introduction

Let $F : D(\supseteq S) \subseteq \Re^n \to \Re^n$ be a continuously differentiable mapping and S be a nonempty closed convex set in $\Re^n$. Variational inequalities, denoted by $VI(S, F)$, is to find a vector $x \in S$ such that

$$F(x)^T(y - x) \geq 0 \quad \text{for all } \ y \in S.$$

In the special case where $S = \Re^n_+$, VI reduces to complementarity problem. A comprehensive survey of VI is given in [9].

Basic methods for solving $VI(S, F)$ are Josephy's Newton and quasi–Newton methods [11, 12]: Let $x^0 \in S$. In general, given $x^k \in S$, we let x^{k+1} be a solution of $VI(S, F^k)$, which is the nearest one to x^k if it is not unique, where

$$F^k(x) := F(x^k) + A(x^k)(x - x^k).$$

If $A(x^k) = F'(x^k)$, this is Josephy's Newton method. On the other hand, if $A(x^k)$ is approximated by some quasi–Newton update, this is Josephy's quasi–Newton method. Under some assumptions, the above methods have high–order convergence rate. However, they also suffer from some drawbacks: First, the subproblem needed to solve is not a linear system; Second, when the subproblem has more than one solution, to find a solution as required will cause numerical difficulties. When S is a polyhedral set, some modifications aimed at eliminating such drawbacks have been discussed in [22, 8]. In this paper we will generalize the ideas in [22, 8] to the case:

$$S = \{y \in \Re^n \mid h_i(y) \leq 0, i = 1, \cdots, m\}, \tag{1.1}$$

where each h_i is twice continuously differentiable and convex. As well as Josephy's Newton and quasi–Newton methods such methods converge only locally. As a remedy for this, in this paper we will propose some hybrid methods.

In order to get globally convergent Newton–type methods, we need a differentiable merit function. Early merit functions such as the regularized gap function [7] are intended to reformulate $VI(S, F)$ as a constrained differentiable optimization problem. Recently, Peng [16] showed that the difference of two regularized gap functions constitutes an unconstrained differentiable optimization problem equivalent to the $VI(S, F)$. Later, Yamashita, Taji and Fukushima [24] extended the idea of Peng [16] and investigated some important properties related to this merit function. Specifically, the latter authors considered the function $g_{\alpha\beta} : \Re^n \to \Re$ defined by

$$g_{\alpha\beta}(x) = f_\alpha(x) - f_\beta(x), \tag{1.2}$$

where α and β are arbitrary positive parameters such that $\alpha < \beta$ and f_α is the regularized gap function

$$f_\alpha(x) = \max_{y \in S} \left\{ \langle F(x), x - y \rangle - \frac{\alpha}{2} \|x - y\|^2 \right\}. \tag{1.3}$$

(The function f_β is defined similarly with α replaced by β.) In the special case $\beta = 1/\alpha$ and $\alpha < 1$ in (1.2), the function $g_{\alpha\beta}$ reduces to the merit function studied by Peng [16]. This function $g_{\alpha\beta}$ is called D–gap function, where D stands for the word "difference". Although $g_{\alpha\beta}$ is continuously differentiable, its gradient function may be not Lipschitz continuous. So we can not directly use the standard optimization methods (Newton–type methods) to solve $g_{\alpha\beta}$ to find a solution of $VI(S, F)$. However, globally and superlinearly convergent Newton–type methods still exist [23]. Since the superlinear convergence condition used in [23] is stronger than that used in Section 2, we will propose a hybrid method in section 3 by considering the results in [23] and Section 2. In practice, F may be not defined on the whole space of $\Re^n$, in Section 4 we discuss a safeguarded Newton method to avoid the possible difficulty caused by the fact that D may not equal to $\Re^n$.

2. Local Newton–type methods

In this section and the next one we assume that $D = \Re^n$, i.e., F is defined on the whole space of $\Re^n$. It is easy to see (for example, see [9]) that to find a solution of $VI(S, F)$ is equivalent to find a solution of the following equation:

$$E(x) := x - \Pi_S(x - F(x) = 0,$$

where Π_S is the orthogonal projection operator on S. In [23], a computable generalized Jacobian $\partial_C \Pi_S(x)$ for the projection operator Π_S at x is proposed under the so called constant rank constraint qualification (CRCQ) at $\Pi_S(x)$. The CRCQ holds at $\Pi_S(x)$ if the linear independent constraint qualification holds at $\Pi_S(x)$ and holds automatically everywhere if S is a polyhedral set. For any matrix $P \in \partial_C \Pi_S(x)$, we have $P^T = P, P^2 = P$. For details, see [8, 23].

Denote

$$\mathcal{W}(x) = \{W \in \Re^{n\times n} | \; W = I - P(I - F'(x)), \; P \in \partial_C \Pi_S(x - F(x))\}.$$

Newton's method for solving $VI(S, F)$

Given $x^0 \in \Re^n$.
Do for $k = 0, 1, ...$:
Choose $P_k \in \partial_C \Pi_S(x^k - F(x^k))$ and compute

$$W_k := I - P_k(I - F'(x^k)).$$

Solve

$$W_k d + E(x^k) = 0 \tag{2.1}$$

for d^k.

$$x^{k+1} = x^k + d^k. \tag{2.2}$$

Theorem 2.1 *Suppose that $F : \Re^n \to \Re^n$ is continuously differentiable and x^* is a solution of $VI(S, F)$. Suppose that the CRCQ holds at $\Pi_S(x^* - F(x^*))$ and all $W_* \in \mathcal{W}(x^*)$ are nonsingular, then there exists a neighborhood N of x^* such that when the initial vector x^0 is chosen in N, the entire sequence $\{x^k\}$ generated by (2.2) is well defined and converges to x^* Q–superlinearly. Furthermore, if F' is Lipschitz continuous around x^* and all ∇h_i^2, $i = 1, ..., m$ are Lipschitz continuous around $\Pi_S(x^* - F(x^*))$, then the convergence is Q–quadratic.*

Proof. By Lemma 2.3 of [23], $\partial_C \Pi_S(\cdot)$ is upper semicontinuous at $(x^* - F(x^*))$ and there exists a neighborhood U of x^* such that for any $P \in \partial_C \Pi_S(x - F(x))$ and $x \in U$

$$\Pi_S(x - F(x)) - \Pi_S(x^* - F(x^*)) - P[x - F(x) - (x^* - F(x^*))] = o(\|x - F(x) - (x^* - F(x^*))\|)$$

(or $= O(\|x - F(x) - (x^* - F(x^*))\|^2)$ if all $\nabla^2 h_i$ are Lipschitz continuous around $\Pi_S(x^* - F(x^*))$).

The upper semicontinuity of $\partial_C \Pi_S(\cdot)$ at $(x^* - F(x^*))$ implies that all $W \in \mathcal{W}(x)$ are nonsingular when x is sufficiently close to x^*. Therefore the algorithm is well defined for $k = 0$ and

$$
\begin{aligned}
\|x^{k+1} - x^*\| &= \|x^k - W_k^{-1} E(x^k) - x^*\| \\
&\leq \|W_k^{-1}\| \|E(x^k) - E(x^*) - W_k(x^k - x^*)\| \\
&= \|W_k^{-1}\| \|\Pi_S(x^k - F(x^k)) - \Pi_S(x^* - F(x^*)) - P_k(I - F'(x^k))(x^k - x^*)\| \\
&\leq \|W_k^{-1}\| [\|\Pi_S(x^k - F(x^k)) - \Pi_S(x^* - F(x^*)) - P_k[x^k - F(x^k) - (x^* - F(x^*))] \\
&\quad + \|P_k[x^k - F(x^k) - (x^* - F(x^*)) - (I - F'(x^k))(x^k - x^*)]\|] \\
&= o(\|x^k - F(x^k) - (x^* - F(x^*))\|) + O(\|F(x^k) - F(x^*) - F'(x^k)(x^k - x^*)\|) \\
&= o(\|x^k - x^*\|).
\end{aligned}
$$

Thus we obtain the Q–superlinear convergence of $\{x^k\}$. Finally when F' and $\nabla^2 h_i$ are Lipschitz continuous, we can easily modify the above arguments to get the Q–quadratic convergence. □

Remark 2.1. For the assumption on the nonsingularity of all $W_* \in \mathcal{W}(x^*)$, we just point out that if $F'(x^*)$ is positive definite on $\Re^n$, such an assumption is satisfied. For a weaker condition, the reader may refer to Proposition 3.1 of [8] for a discussion.

Quasi–Newton method (Broyden's case [3])

Given $x^0 \in \Re^n$, $D_0 \in \Re^{n \times n}$ (an approximation of $F'(x^0)$).
Do for $k = 0, 1, \ldots$:
Choose $P_k \in \partial_C \Pi_S(x^k - F(x^k))$ and compute

$$V_k := I - P_k(I - D_k).$$

Solve

$$V_k d + E(x^k) = 0 \tag{2.3}$$

for d^k.

$$x^{k+1} = x^k + d^k, \tag{2.4}$$

$$y^k = F(x^{k+1}) - F(x^k),$$

$$D_{k+1} = D_k + \frac{(y^k - D_k d^k)(d^k)^T}{(d^k)^T d^k}.$$

Theorem 2.2 *Suppose that $F : \Re^n \to \Re^n$ is continuously differentiable, x^* is a solution of $VI(S, F)$, F' is Lipschitz continuous in a neighborhood of x^* and the Lipschitz constant is γ. Suppose that the CRCQ holds at $\Pi_S(x^* - F(x^*))$ and all $W_* \in \mathcal{W}(x^*)$ are nonsingular. There exist positive constants ε, δ such that if $\|x^0 - x^*\| \leq \varepsilon$ and $\|D_0 - F'(x^0)\| \leq \delta$, then the sequence $\{x^k\}$ generated by (2.4) is well defined and converges Q–superlinearly to x^*.*

Proof. By considering Theorem 2.1, we may verify the results of Theorem 2.2 by a similar argument to the proof of Theorem 2.2 of [8]. Here we omit the detail. □

3. A hybrid Newton method based on D–gap function

In [23], an approximation of the generalized Hessian of f_α at x is defined as

$$\tilde{H}_C f_\alpha(x) = \{V \in \Re^{n\times n} \mid \quad V = F'(x) + (F'(x)^T - \alpha I)(I - P_\alpha(I - \alpha^{-1}F'(x)^T)^T), \\ P_\alpha \in \partial_C \Pi_S(x - \alpha^{-1}F(x))\}$$

and $\tilde{H}_C f_\beta(x)$ is defined similarly. Then the computable generalized Hessian of $g_{\alpha\beta}$ at x is defined as

$$H_C g_{\alpha\beta}(x) = \tilde{H}_C f_\alpha(x) - \tilde{H}_C f_\beta(x). \tag{3.1}$$

Let

$$\theta(x) = \frac{1}{2}E(x)^T E(x).$$

Using these sets and the Newton's method presented in Section 2, we may give the following hybrid method:

Hybrid Newton method for $VI(S,F)$

Step 0. Given $x^0 \in \Re^n$, $\tau, \eta, \eta_1 \in (0,1)$, $\gamma, \rho, \varepsilon \in (0,\infty)$. $k := 0$.
Step 1. Choose $P_k \in \partial_C \Pi_S(x^k - F(x^k))$ and compute

$$W_k := I - P_k(I - F'(x^k)).$$

Step 2. Solve

$$W_k d + E(x^k) = 0 \tag{3.2}$$

for d^k.

Step 3. If $\theta(x^k + d^k) \le \eta\theta(x^k)$, let

$$x^{k+1} = x^k + d^k.$$

$k := k+1$. Go to step 1. Otherwise, go to Step 4.

Step 4. Let $y^{k,0} = x^k$. $j := 0$.

Step 5. Choose $V_{k,j} \in H_C g_{\alpha\beta}(y^{k,j})$. If $V_{k,j}$ is nonsingular and

$$(V_{k,j}^{-1}\nabla g_{\alpha\beta}(y^{k,j}))^T \nabla g_{\alpha\beta}(y^{k,j}) \ge \rho\|\nabla g_{\alpha\beta}(y^{k,j})\|^{2+\varepsilon}, \tag{3.3}$$

let

$$s^{k,j} = -V_{k,j}^{-1}\nabla g_{\alpha\beta}(y^{k,j});$$

otherwise, let

$$s^{k,j} = -\gamma\nabla g_{\alpha\beta}(y^{k,j}).$$

Step 6. Let m_j be the smallest nonnegative integer m such that

$$g_{\alpha\beta}(y^{k,j} + \tau^m s^{k,j}) - g_{\alpha\beta}(y^{k,j}) \le \eta_1\tau^m (s^{k,j})^T \nabla g_{\alpha\beta}(y^{k,j})$$

holds. Let

$$y^{k,j+1} = y^{k,j} + \tau^{m_j} s^{k,j}.$$

Step 7. If $\theta(y^{k,j+1}) \le \eta\theta(x^k)$, let

$$x^{k+1} = y^{k,j+1}.$$

$k := k+1$. Go to Step 1. Otherwise, $j := j+1$. Go to Step 5.

Theorem 3.1 *Suppose that F is strongly monotone on $\Re^n$ and the CRCQ holds at $\Pi_S(y)$ for any $y \in \Re^n$. Then the above hybrid Newton method is well–defined and any accumulation point of the infinite sequence $\{x^k\}$ is a solution of $VI(S,F)$. Furthermore, if the set $G := \{x \in \Re^n | \theta(x) \leq \theta(x^0)\}$ is bounded then the sequence $\{x^k\}$ will converge to the unique solution $\bar{x}$ of $VI(S,F)$ Q–superlinearly. Moreover, the convergence is Q–quadratic if $F'(y)$ is Lipschitz continuous around $\bar{x}$ and ∇h_i, $i = 1, ..., m$ are Lipschitz continuous around $\Pi_S(\bar{x} - F(\bar{x}))$.*

Proof. Since F is strongly monotone on $\Re^n$, any stationary point of $g_{\alpha\beta}$ is a solution of $VI(S,F)$ [16, 24]. This means that in the above hybrid Newton method Steps 4–7 cannot loop and an infinite sequence $\{x^k\}$ will be generated. From the algorithm we have

$$\begin{aligned} \theta(x^{k+1}) &\leq \eta\theta(x^k) \\ &\leq (\eta)^{k+1}\theta(x^0). \end{aligned}$$

Therefore,

$$\lim_{k\to\infty} \theta(x^k) = 0,$$

which means that any accumulation point of $\{x^k\}$ is a solution of $VI(S,F)$.

The Q–superlinear (–quadratic) convergence of $\{x^k\}$ may be obtained by Theorem 2.1 and Remark 2.1 easily. □.

4. A safeguarded Newton method under pseudomonotone condition

In many problems, the mapping F is only defined on the set S or outside S some monotonicity condition on F, which is essential in designing effective algorithms, will lose [6], it is desirable to consider algorithms similar to those discussed in Sections 2–3 while keeping the iteration sequence in S. Such approaches have already been discussed in [8] for solving $VI(S,F)$ with S given by a polyhedral set. The results in [8] may be easily generalized to the case that S is given by (1.1). In this section, however, under the assumption that F is pseudomonotone at a solution, we will provide a globally and superlinearly convergent hybrid method while keeping the feasibility. The tools used here are the so–called normal maps [19] and the projection and contraction (PC) method for solving $VI(S,F)$ [20]. Normal maps for $VI(S,F)$ are defined by

$$H(z) := F(\Pi_S(z)) + z - \Pi_S(z). \tag{4.1}$$

It is easy to verify (for example, see [19]) that if $H(z) = 0$, then the point $x := \Pi_S(z)$ solves $VI(S,F)$; conversely if x solves $VI(S,F)$, then with $z := x - F(x)$ one has $H(z) = 0$. Therefore the equation $H(z) = 0$ is an equivalent way of formulating the variational inequality problem $VI(S,F)$. It is deserved to point out that although H is defined on the whole space of $\Re^n$, F is only required to be defined on S. Similar to the discussion in Section 2 we may give the Newton–type methods for solving $H(z) = 0$. Again, these are only locally convergent methods. In order to obtain a globally and locally superlinear convergent method, we first describe a globally convergent method recently obtained by Sun [20].

Let S^* denote the solution set of $VI(S,F)$ and

$$E(x,\beta) = x - \Pi_S[x - \beta F(x)]. \tag{4.2}$$

When $\beta = 1$, $E(x, 1) = E(x)$.

Choose an arbitrary constant $\eta \in (0, 1)$. When $x \notin S^*$, define

$$\eta(x) = \begin{cases} \max\{\eta, 1 - \dfrac{t(x)}{\|E(x,1)\|^2}\}, & \text{if } t(x) > 0 \\ \\ 1, & \text{otherwise} \end{cases} \tag{4.3}$$

and

$$s(x) = \begin{cases} (1 - \eta(x))\dfrac{\|E(x,1)\|^2}{t(x)}, & \text{if } t(x) > 0 \\ \\ 1, & \text{otherwise} \end{cases}, \tag{4.4}$$

where $t(x) = \{F(x) - F(\Pi_C[x - F(x)])\}^T E(x, 1)$.

Define

$$b(x, \beta) = F(\Pi_S[x - \beta F(x)]) - F(x) + E(x, \beta)/\beta. \tag{4.5}$$

Then we can describe a globally convergent method appeared in [20].

Projection and Contraction (PC) Method

Step 0. Choose an arbitrary vector $x^0 \in S$. Choose positive constants $\eta,\ \alpha \in (0,\ 1),\ 0 < \Delta_1 \leq \Delta_2 < 2$. $k := 0$, go to step 1.

Step 1. Calculate $\eta(x^k)$ and $s(x^k)$. If $s(x^k) = 1$, let $\beta_k = 1$; otherwise determine $\beta_k = s(x^k)\alpha^{m_k}$, where m_k is the smallest nonnegative integer m such that

$$\begin{aligned} &\{F(x^k) - F(\Pi_S[x^k - s(x^k)\alpha^m F(x^k)])\}^T E(x^k, s(x^k)\alpha^m) \\ &\qquad \leq (1 - \eta(x^k))\|E(x^k, s(x^k)\alpha^m)\|^2/(s(x^k)\alpha^m) \end{aligned} \tag{4.6}$$

holds.

Step 2. Calculate $b(x^k, \beta_k)$.

Step 3. Calculate

$$\rho_k = E(x^k, \beta_k)^T b(x^k, \beta_k)/\|b(x^k, \beta_k)\|^2. \tag{4.7}$$

Step 4. Take $\gamma_k \in [\Delta_1,\ \Delta_2]$ and set

$$x^{k+1} = \Pi_S[x^k - \gamma_k \rho_k b(x^k, \beta_k)]. \tag{4.8}$$

$k := k + 1$, go to step 1.

The mapping F is said to be pseudomonotone at a solution $x^* \in S^*$ over S if

$$F(x)^T(x - x^*) \geq 0 \quad \text{for all } x \in S.$$

Theorem 4.1 *[20]. Suppose that F is continuous over S and pseudomonotone at a solution point x^* over S. Then the infinite sequence $\{x^k\}$ generated by the above PC method is bounded and there exists a subsequence of $\{x^k\}$ converging to a solution of $VI(S, F)$.*

When S is of the following form

$$S = \{x \in \Re^n | \ l \leq x \leq u\}, \tag{4.9}$$

where l and u are two vectors of $\{R \cup \{\infty\}\}^n$, we can give an improved form of the PC method. For any $x \in S$ and $\beta > 0$, denote

$$N(x,\beta) = \{i| \ (x_i \leq l_i \text{ and } (b(x,\beta))_i \geq 0) \text{ or } (x_i \geq u_i \text{ and } (b(x,\beta))_i \leq 0)\},$$

$$B(x,\beta) = \{1,...,n\} \backslash N(x,\beta). \tag{4.10}$$

Denote $b_N(x,\beta)$ and $b_B(x,\beta)$ as follows

$$(b_N(x,\beta))_i = \begin{cases} 0, & \text{if } i \in B(x,\beta) \\ (b(x,\beta))_i, & \text{otherwise} \end{cases},$$

$$(b_B(x,\beta))_i = (b(x,\beta))_i - (b_N(x,\beta))_i, \quad i = 1,...,n. \tag{4.11}$$

Then for any $x^* \in S^*$ and $x \in S$,

$$(x - x^*)^T b_B(x,\beta) \geq (x - x^*)^T b(x,\beta). \tag{4.12}$$

So if in the PC method we set

$$x^{k+1} = \Pi_S[x^k - \gamma_k \bar{\rho}_k b_B(x^k, \beta_k)] \tag{4.13}$$

where

$$\bar{\rho}_k = E(x^k,\beta_k)^T b(x^k,\beta_k)/\|b_B(x^k,\beta_k)\|^2,$$

then the convergence Theorem 4.1 holds for the modified PC method. In practice, we will use the iterative form (4.13) when S is of the form (4.9).

Define

$$\mathcal{N}(z) = \{W \in \Re^{n\times n} | \ N = F'(\Pi_S(z))P + I - P, \ P \in \partial_C \Pi_S(z)\}$$

and

$$r(z) = \frac{1}{2}\|H(z)\|^2.$$

Safeguarded Newton Method

Step 0. Choose an arbitrary vector $z^0 \in \Re^n$. Choose scalars $\eta, \ \alpha, \gamma, \varepsilon_0 \in (0,1)$, $\sigma \in (0, 1/2)$, and $0 < \Delta_1 \leq \Delta_2 < 2$. $k := 0$, go to step 1.

Step 1. Choose $W_k \in \mathcal{N}(z^k)$.

Step 2. If W_k is singular, go to step 6; otherwise solve

$$W_k s + H(z^k) = 0$$

for s^k. If

$$r(z^k + s^k) \leq (1-\sigma) r(z^k), \tag{4.14}$$

let $z^{k+1} = z^k + s^k, \ k := k+1$, go to step 1; otherwise, go to step 3.

Step 3. If $r'(z^k; s^k) < -\varepsilon_0 r(z^k)$, let $d^k = s^k$ and go to step 5; otherwise go to step 4.

Step 4. If $r'(z^k; -s^k) < -\varepsilon_0 r(z^k)$, let $d^k = -s^k$ and go to step 5; otherwise, go to step 6.

Step 5. (safeguarding step) Let $\beta^k = \alpha^{m_k}$, where m_k is the first nonnegative integer m such that

$$r(z^k + \alpha^m d^k) \le r(z^k) + \sigma \alpha^m r'(z^k; d^k)$$

or

$$\alpha^m \le \gamma$$

holds.
If $\beta^k \ge \gamma$, let $z^{k+1} = z^k + \beta_k d^k$, $k := k+1$, and go to step 1; otherwise, go to step 6.

Step 6. Set $x^{k,0} = \Pi_S(z^k)$ and $i := 0$. Take $x^{k,0}$ as the initial vector and use PC method till to get a sequence $\{x^{k,0}, x^{k,1}, ..., x^{k,i(k)}\}$ such that $i(k)$ is the first nonnegative integer i such that

$$r(x^{k,i} - F(x^{k,i})) \le (1-\sigma) r(z^k).$$

Set $z^{k+1} = x^{k,i(k)} - F(x^{k,i(k)})$ and $k := k+1$. Go to step 1.

Before giving the convergence theorem, we make several remarks.

Remark 4.1. We use the safeguarding step because H is not continuously differentiable.
Remark 4.2. The pseudomonotonicity assumption of F is used only when the Newton step fails.
Remark 4.3. The finite termination of Step 6 at a non–solution point z^k is guaranteed by Theorem 4.1 and the continuity of F and Π_S.

Theorem 4.2 *Let F be continuously differentiable over S. Suppose that F is pseudomonotone at a solution of $VI(S,F)$ over S, and CRCQ holds. Then the sequence $\{z^k\}$ generated by the above safeguarded Newton method is well defined and $\lim_{k\to\infty} r(z^k) = 0$. Furthermore, if all $W \in \mathcal{N}(\bar{z})$ are nonsingular at an accumulation point $\bar{z}$, then $\{z^k\}$ converges to $\overline{z}$ Q–superlinearly and Q–quadratically if F' is Lipschitz continuous in a neighborhood of $\bar{z}$ and all $\nabla^2 h_i,\ i = 1, ..., m$ are Lipschitz continuous around $\Pi_S(z)$.*

Proof. According to the safeguarded Newton method, we have

$$\begin{aligned} r(z^{k+1}) &\le (1-\sigma\gamma\varepsilon_0) r(z^k) \\ &\le (1-\sigma\gamma\varepsilon_0)^{k+1} r(z^0). \end{aligned}$$

Therefore,

$$\lim_{k\to\infty} r(z^k) = 0.$$

Furthermore, if $\overline{z}$ is an accumulation point of $\{z^k\}$ and all $W \in \mathcal{N}(\bar{z})$ are nonsingular, then similar to the proof of Theorem 2.1 we may prove that when z^k is close enough to $\overline{z}$, we have

$$\|z^k + s^k - \bar{z}\| = o(\|z^k - \bar{z}\|).$$

This, and the fact that $\partial_B H(\bar{z}) \subseteq \mathcal{N}(\bar{z})$, implies that (4.14) is satisfied if z^k is close enough to $\bar{z}$ and the full Newton step will be taken. So by modifying the proof of Theorem 2.1, we may obtain the Q–superlinear (–quadratic) convergence of $\{z^k\}$ under the assumptions. □

There are several papers on the topic of this section under the assumption that F is pseudomonotone at a solution [1, 2, 21]. In [1, 2] the subproblem needed to solve is not a linear system of equations (a quadratic programming subproblem for $VI(S, F)$ with S given by (4.9)) while in [21] (Chapter 6) F is required to be defined on the whole space of $\Re^n$, which was not made here. For a generalization of the above safeguarded Newton method we can ask F to be only semismooth instead of continuously differentiable as in [21]. For the discussion on the applications of the concept of semismoothness to nonsmooth equations see [18, 17, 15].

References

[1] S.C. Billups, *Algorithms for Complementarity Problems and Generalized Equations*, Ph.D. Thesis, Computer Sciences Department, University of Wisconsin, Madison, WI, 1995.

[2] S.C. Billups and M.C. Ferris, "QPCOMP: A quadratic programming based solver for mixed complementarity problems", *Mathematical Programming*, to appear.

[3] C.G. Broyden, "A class of methods for solving nonlinear simultaneous equations", *Mathematics of Computation* 19 (1965) 577–593.

[4] J.E. Dennis and J.J. Moré, "A characterization of superlinear convergence and its application to quasi–Newton methods", *Mathematics of Computation* 28 (1974) 549–560.

[5] J.E. Dennis and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice–Hall, Englewood Cliffs, N.J., 1983).

[6] S.P. Dirkse, M.C. Ferris, P.V. Preckel and T. Rutherfud, The GAMS callable program library for variational and complementarity solvers, Technical Report 94–07, Computer Sciences Department, University of Wisconsin, Madison, WI, 1994.

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

[8] J. Han and D. Sun, "Newton and quasi–Newton methods for normal maps with polyhedral sets", *Journal of Optimization Theory and Applications* 94 (1997), to appear.

[9] 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.

[10] C.–M. Ip and T. Kyparisis, "Local convergence of quasi–Newton methods for B–differentiable equations", *Mathematical Programming* 56 (1992) 71–89.

[11] N.H. Josephy, "Newton's method for generalized equations", Technical Summary Report No. 1965, Mathematical Research Center, University of Wisconsin (Madison, WI, 1979).

[12] N.H. Josephy, "Quasi–Newton methods for generalized equations", Technical Summary Report No. 1966, Mathematical Research Center, University of Wisconsin (Madison, WI, 1979).

[13] M. Kojima and S. Shindo, "Extensions of Newton and quasi–Newton methods to systems of PC^1 equations", *Journal of the Operations Research Society of Japan* 29 (1986) 352–374.

[14] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).

[15] J.–S. Pang and L. Qi, "Nonsmooth equations: Motivation and algorithms", *SIAM Journal on Optimization* 3 (1993) 443–465.

[16] J.M. Peng, Equivalence of variational inequality problems to unconstrained optimization, Technical Report, State Key Laboratory of Scientific and Engineering Computing, Academia Sinica, Beijing, China (April 1995), to appear in *Mathematical Programming*.

[17] L. Qi, "Convergence analysis of some algorithms for solving nonsmooth equations", *Mathematics of Operations Research* 18 (1993) 227–244.

[18] L. Qi and J. Sun, "A nonsmooth version of Newton's method", *Mathematical Programming* 58 (1993) 353–367.

[19] S.M. Robinson, "Normal maps induced by linear transformation", *Mathematics of Operations Research* 17 (1992) 691–714.

[20] D. Sun, "A class of iterative methods for solving nonlinear projection equations", *Journal of Optimization Theory and Applications* 91 (1996) 123–140. to appear.

[21] D. Sun, *Algorithms and Convergence Analysis for Nonsmooth Optimization and Nonsmooth Equations*, Ph.D. Thesis, Institute of Applied Mathematics, Chinese Academy of Sciences, 1994.

[22] D. Sun and J. Han, "Newton and quasi–Newton methods for a class of nonsmooth equations and related problems", *SIAM Journal on Optimization* 7 (1997), to appear.

[23] D. Sun, M. Fukushima and L. Qi, "A computable generalized Hessian of the D–gap function and Newton–type methods for variational inequality problems", in *Complementarity and Variational Problems: State of the Art*, M.C. Ferris and J.S. Pang, eds., SIAM, Philadelphia, Pennsylvania, 1997.

[24] N. Yamashita, K. Taji and M. Fukushima, Unconstrained optimization reformulations of variational inequality problems, *Journal of Optimization Theory and Applications*, to appear.