Journal of Computational Mathematics, Vol.18, No.3, 2000, 289–304.

# BROYDEN'S METHOD FOR SOLVING VARIATIONAL INEQUALITIES WITH GLOBAL AND SUPERLINEAR CONVERGENCE *

Yu-fei Yang Dong-hui Li
(*Department of Applied Mathematics, Hunan University, Changsha 410082, China*)

**Abstract**

In this paper, we establish a quasi-Newton method for solving the KKT system arising from variational inequalities. The subproblems of the proposed method are lower-dimensional mixed linear complementarity problems. A suitable line search is introduced. We show that under suitable conditions, the proposed method converges globally and superlinearly.

*Key words*: Variational inequality, quasi-Newton method, global convergence, superlinear convergence

## 1. Introduction

We are concerned with the following variational inequality problem of finding an $x \in X$ such that

$$(y-x)^T f(x) \geq 0, \quad \forall y \in X, \tag{1.1}$$

where $f : R^n \to R^n$ is assumed to be a continuously differentiable function, and $X \subseteq R^n$ is specified by

$$X = \{x \in R^n : g_i(x) \leq 0, i = 1, 2, \ldots, m;\ \ h_j(x) = 0, j = 1, 2, \ldots, l\}, \tag{2.2}$$

where $g_i : R^n \to R$ and $h_j : R^n \to R$ are twice continuously differentiable functions. The variational inequality (1.1) is denoted by $\mathrm{VI}(X, f)$. An important special case of $\mathrm{VI}(X, f)$ is the so-called nonlinear complementarity problem ($\mathrm{NCP}(f)$) with $X = R^n_+ = \{x \in R^n \,|\, x \geq 0\}$. Variational inequality and nonlinear complementarity problems have attracted many people's attention because of their wide practice background. We refer to [5] for a review on it.

Originated by Josephy [7, 8], quasi-Newton methods are now an important class of iterative methods for solving $\mathrm{VI}(X, f)$ and $\mathrm{NCP}(f)$, and have attracted of many veseachers (see [2, 5, 6, 7, 8, 10, 13, 16, 18] etc.).

One kind of quasi-Newton methods is the linearized quasi-Newton methods where the $\mathrm{VI}(X, f)$ and $\mathrm{NCP}(f)$ are approximated successively by a sequence of linear problems $\mathrm{VI}(X, f^k)$ and $\mathrm{NCP}(f^k)$ (e.g., see [8] and [13]), where

$$f^k(x) = f(x_k) + B_k(x - x_k),$$

* Received September 16, 1996.

and $B_k$ is an $n \times n$ matrix. Another kind of quasi-Newton methods is established based on equivalent nonsmooth equations of $\mathrm{VI}(X, f)$ and $\mathrm{NCP}(f)$ (e.g., see [2, 6, 16]). The local convergence properties of these quasi-Newton methods are very similar to those of quasi-Newton methods for smooth equations and optimizations. In particular, they converge superlinearly. However, so far the studies for quasi-Newton methods focused on their local behavior. This paper devotes to develop a globally convergent quasi-Newton method. To do this, we simply review the related globally convergent Newton methods for solving $\mathrm{VI}(X, f)$.

The globally convergent Newton methods have developed vary fast in the past two decades. Using appropriate reformulations, there have been established various kinds of damped Newton methods (e.g., see [11, 12, 15, 17]). The ideas of Newton methods for solving smooth equations and mathematical programming were applied to solve variational inequalities successfully. Basically, in these methods, Armijo-type line search was used. Under suitable conditions, these methods converge globally and superlinearly/quadratically. For quasi-Newton methods, however, Armijo-type line search seems inappropriate because its implementation relies on the calculation of $\nabla f(x)$. Therefore, to establish globally convergent quasi-Newton methods, it seems necessary to adopt new line search technique. In [18], on the basis of the derivative-free line search given by Griewank [4] on quasi-Newton methods for smooth equations, a derivative-free line search was proposed, and a quasi-Newton method for solving $\mathrm{NCP}(f)$ was proposed. Under certain conditions, the method converges globally and superlinearly. This line search was further studied in [10]. In this paper, we adopt the same line search to establish a Broyden-like method for solving $\mathrm{VI}(X, f)$. Under suitable conditions, we prove its global and superlinear convergence. The method in this paper is quite different from the ones in [10] and [18]. The advantages of method in this paper are: first, the method is suitable to solving general variational inequality problems. Second, the subproblem is a lower-dimensional mixed linear complementarity problem. Moreover, as we will show in section 4, the method converges superlinearly without strict complementarity assumption.

The organization of the paper is as follows. In the next section, we do some preliminaries, and describe the method. In section 3 and 4, we prove global and superlinear convergence of the proposed method.

## 2. Preliminaries

In this section, we do some preliminaries. We first introduce the Karush-Kuhn-Tucker system for $\mathrm{VI}(X, f)$. It is well-known ( e.g., see [5] or [12]) that if $x$ is a solution of $VI(X, f)$ and a suitable constraint qualification is satisfied at $x$, then there exist multiplier vectors $u \in R^m$ and $v \in R^l$ such that the following mixed complementarity conditions are satisfied:

$$\begin{cases} f(x) + \nabla g(x)u + \nabla h(x)v = 0, \\ u \geq 0, \quad g(x) \leq 0 \quad \text{and} \quad u^T g(x) = 0, \\ h(x) = 0, \end{cases} \tag{2.1}$$

where $g(x) = (g_1(x), g_2(x), \ldots, g_m(x))^T$ and $h(x) = (h_1(x), h_2(x), \ldots, h_l(x))^T$. A triple $z = (x, u, v) \in R^{n+m+l}$ satisfying system (2.1) is called a Karush-Kuhn-Tucker vector of $VI(X, f)$. If $f$ is the gradient mapping of some function $F : R^n \to R$, then (2.1) is just the KKT system of the constrained optimization problem

$$\min_{x \in X} F(x).$$

Define function $H : R^{n+m+l} \to R^{n+m+l}$ by

$$H(z) = \begin{pmatrix} f(x) + \nabla g(x)u + \nabla h(x)v \\ \min(u, -g(x)) \\ -h(x) \end{pmatrix}, \tag{2.2}$$

where $z = (x, u, v) \in R^{n+m+l}$. Then the KKT system (2.1) can be represented compactly as the following system of nonlinear equations

$$H(z) = 0. \tag{2.3}$$

The function $H$ is generally not differentiable in the sense of Frechét derivative but directionally differentiable. The directional derivative of $H$ at $z = (x, u, v)$ along direction $d = (y, \mu, \nu)$ is given by (see [11])

$$H'(z, d) = \begin{pmatrix} \nabla L(z)^T d \\ (-\nabla g_i(x)^T y)_{i \in \alpha(z)} \\ (\min(\mu_i, -\nabla g_i(x)^T y))_{i \in \beta(z)} \\ (\mu_i)_{i \in \gamma(z)} \\ -\nabla h(x)^T y \end{pmatrix},$$

where $L : R^{n+m+l} \to R^n$ is given by

$$L(z) = f(x) + \nabla g(x)u + \nabla h(x)v$$

and $\alpha(z)$, $\beta(z)$ and $\gamma(z)$ are index sets defined by

$$\alpha(z) = \{i : u_i > -g_i(x)\}, \quad \beta(z) = \{i : u_i = -g_i(x)\}, \quad \gamma(x) = \{i : u_i < -g_i(x)\}.$$

Based on the equivalent nonsmooth equation (2.3), there have developed many globally convergent Newton methods. We will establish a quasi-Newton method on the basis of the damped Newton methods in [11] and [12]. To do so, let us simply review the related damped Newton method in [11] and [12].

In the method proposed in [11], the subproblem is the following equation

$$H(z^k) + H'(z^k, d^k) = 0. \tag{2.4}$$

Though equation (2.4) is a nonlinear, it is equivalent to a lower-dimensional mixed linear complementarity problem (see [11]). Moreover, the direction $d_k$ generated by (2.4) is a descent directional of the norm function

$$\theta(z) = \frac{1}{2} H(z)^T H(z) = \frac{1}{2}(\|L(z)\|^2 + \|\min(u, -g(x))\|^2 + \|h(x)\|^2),$$

here and throughout the paper, $\|\cdot\|$ denotes the Euclidean norm of vectors. With Armijo-type line search, the damped Newton method converges globally if there is an accumulation point $z^*$ of $\{z^k\}$ at which $H$ is strongly F-differentiable (see [11]). In [12], a modified method was proposed. The modified method retains global convergence without the strong F-differentiability assumption. In the modified method, the subproblem is still a lower-dimensional mixed linear complementarity problem but a little different from the one in [11].

A natural idea to establish quasi-Newton methods is to use the following linear equation as an approximation of (2.4)

$$H(z^k) + B_k d^k = 0.$$

The related work can be seen in [2, 6, 16] etc.. Local superlinear convergence of these related methods was proved. Our consideration is different from above work. We establish a quasi-Newton method where the subproblem is an affine variational inequality problem. The KKT system of the subproblem is still a lower-dimensional mixed linear complementarity problem. To do so, define index sets

$$\alpha_-(z) = \{i : 0 > u_i > -g_i(x)\}, \quad \gamma_-(z) = \{i : u_i < -g_i(x) < 0\},$$
$$\bar{\alpha}(z) = \alpha(z) \backslash \alpha_-(z), \quad \bar{\beta}(z) = \beta(z) \cup \alpha_-(z) \cup \gamma_-(z), \quad \bar{\gamma}(z) = \gamma(z) \backslash \gamma_-(z).$$

Let

$$f^k(x) = f(x^k) + B_k(x - x^k) \tag{2.5}$$

and $X^k \subseteq R^n$ be the polyhedral set defined by the linear system

$$\begin{aligned} g_{\bar{\alpha}}(x^k) + \nabla g_{\bar{\alpha}}(x^k)^T(x - x^k) &= 0, \\ g_{\bar{\beta}}(x^k) + \nabla g_{\bar{\beta}}(x^k)^T(x - x^k) &\le 0, \\ h(x^k) + \nabla h(x^k)^T(x - x^k) &= 0, \end{aligned}$$

here $\bar{\alpha}, \bar{\beta}$ and $\bar{\gamma}$ denote $\bar{\alpha}(z^k), \bar{\beta}(z^k)$ and $\bar{\gamma}(z^k)$, respectively.

Let $(\bar{x}^k, \bar{u}^k_{\bar{\alpha}}, \bar{u}^k_{\bar{\beta}}, \bar{v}^k)$ be a KKT point for $VI(X^k, f^k)$ and $\bar{u}^k_{\bar{\gamma}} = 0$. We determine the direction $d^k = (y^k, \mu^k, \nu^k)$ by

$$y^k = \bar{x}^k - x^k, \quad \mu^k = \bar{u}^k - u^k, \quad \nu^k = \bar{v}^k - v^k \tag{2.6}$$

It is not difficult to see that $d^k$ satisfies the following lower-dimensional mixed linear complementarity problem

$$\begin{cases} L(z^k) + B_k y^k + \nabla g(x^k)\mu^k + \nabla h(x^k)\nu^k = 0, \\ g_{\bar{\alpha}}(x^k) + \nabla g_{\bar{\alpha}}(x^k)^T y^k = 0, \\ \min(u^k_{\bar{\beta}} + \mu^k_{\bar{\beta}}, -g_{\bar{\beta}}(x^k) - \nabla g_{\bar{\beta}}(x^k)^T y^k) = 0, \\ u^k_{\bar{\gamma}} + \mu^k_{\bar{\gamma}} = 0, \\ h(x^k) + \nabla h(x^k)^T y^k = 0. \end{cases} \tag{2.7}$$

If in (2.7), $B_k$ is replaced by $\nabla_x L(z^k)^T$, then (2.7) reduces the subproblem of the damped Newton method in [12].

We have established the subproblem of the quasi-Newton method. An important task is then to allocate a suitable line search. Since in quasi-Newton methods, the calculation of $\nabla_x L(z)$ should be avoided, it is necessary to adopt a derivative-free line search. In this paper, we use the derivative-free line search proposed in [10], which is motivated from [4]. Given a positive sequence $\{\sigma_k\}$ satisfying

$$\sum_{k=0}^{\infty} \sigma_k \leq \sigma < \infty. \tag{2.8}$$

Define

$$q_k(\lambda) = \frac{H(z^k)^T(H(z^k) - H(z^k + \lambda d^k)) + \frac{1}{2}\sigma_k \|H(z^k)\|^2}{\max\{\|H(z^k) - H(z^k + \lambda d^k)\|^2, \min(\|\lambda d^k\|^2, \frac{2}{3}\|H(z^k)\|^2)\}}. \tag{2.9}$$

Then it is clear that $\lim_{\lambda \to 0^+} q_k(\lambda) = +\infty$, because as $\lambda \to 0^+$, the numerator of $q_k(\lambda)$ goes to a positive constant $\frac{1}{2}\sigma_k \|H(z^k)\|$ but the denominator goes to 0. Consequently for a constant $\epsilon > 0$, the following inequality is satisfied for all small positive $\lambda_k$:

$$q_k(\lambda_k) \geq \frac{1}{2} + \epsilon. \tag{2.10}$$

Now, we state the steps of our quasi-Newton method.

**Algorithm 1**

Choose $\rho \in (0,1), \epsilon \in (0, \frac{1}{6})$, initial vector $z^0 = (x^0, u^0, v^0) \in R^{n+m+l}$ and initial matrix $B_0 \in R^{n \times n}$. Set $k := 0$.

**Step 1.** If $\theta(z^k) = 0$, stop. Otherwise, solve (2.7) and let $d^k = (y^k, \mu^k, \nu^k)$.

**Step 2.** Determine $\lambda_k = \rho^{i_k}$, where $i_k$ is the smallest nonnegative integer $i$ such that (2.10) holds for $\lambda_k = \rho^i$.

**Step 3.** Set $z^{k+1} = z^k + \lambda_k d^k$.

**Step 4.** Update $B_k$ to get $B_{k+1}$ by $B_{k+1} = B_k + \Delta B_k$, where

$$\Delta B_k = \begin{cases} 0, & \text{if } s^k = 0, \\ \dfrac{(r_k - B_k s^k)(s^k)^T}{\|s^k\|^2}, & \text{if } s^k \neq 0, \end{cases}$$

where $r_k = L(x^{k+1}, u^k, v^k) - L(x^k, u^k, v^k)$, $s^k = x^{k+1} - x^k$.

**Step 5.** Replace $k$ by $k+1$ and go to Step 1.

## 3. Global Convergence

In this section, we prove global convergence of Algorithm 1. Similar to [12], throughout this section and next section, we make the following blanket assumption:

**Assumption A0.** each subproblem $VI(X^k, f^k)$ has at least one solution.

**Lemma 3.1**. *For all $k$, we have either*

$$(1 + \sigma_k)\theta_k - \theta_{k+1} \geq \epsilon \|z^{k+1} - z^k\|^2 \tag{3.1}$$

*or*

$$\theta_{k+1} \leq (1 + \sigma_k - \frac{4}{3}\epsilon)\theta_k \tag{3.2}$$

*where $\theta_k = \theta(z^k)$. Moreover, we have $\theta_k \le e^{\sigma}\theta_0$ and $\{\theta_k\}$ converges.*

*Proof.* The inequalities (3.1) and (3.2) are straightforward from the definition of $q_k(\lambda)$ and the fact that $2q_k(\lambda_k) - 1 \ge 2\epsilon$. Moreover, (3.1) and (3.2) imply that $(1+\sigma_k)\theta_k - \theta_{k+1} \ge 0$ holds for all $k$. Since $\sigma_k$ satisfies (2.8), we get the convergence of $\theta_k$ from Lemma 3.3 of [3]. □

Define

$$N_1 = \{k | (1+\sigma_k)\theta_k - \theta_{k+1} \ge \epsilon \|z^{k+1} - z^k\|^2\}$$

and

$$N_2 = \{k | \theta_{k+1} \le (1 + \sigma_k - \frac{4}{3}\epsilon)\theta_k\}.$$

Then, it is not difficult to show from Lemma 3.1 that

**Lemma 3.2**. *Let $\{z^k\}$ be generated by Algorithm* 1.

(1) *If $N_2$ is finite, then*

$$\sum_{k=0}^{\infty} \|z^{k+1} - z^k\|^2 < \infty. \tag{3.3}$$

*(2) If $N_2$ is infinite, then $\{\theta_k\}$ converges to 0. Moreover, any accumulation point of $\{z^k\}$ is a KKT point for $VI(X, f)$.*

Define index sets

$$\alpha_+(z) = \{i : u_i > -g_i(x) \text{ and } u_i > 0\},$$
$$\alpha_0(z) = \{i : u_i > -g_i(x) \text{ and } u_i = 0\},$$
$$\gamma_+(z) = \{i : u_i < -g_i(x) \text{ and } g_i(x) < 0\},$$
$$\gamma_0(z) = \{i : u_i < -g_i(x) \text{ and } g_i(x) = 0\},$$

and matrix $A(z)$ and vector $C(z)$

$$A(z) = \begin{pmatrix} \nabla_x L(z)^T & \nabla g_{\alpha_+(z)}(x) & \nabla h(x) \\ -\nabla g_{\alpha_+(z)}(x)^T & 0 & 0 \\ -\nabla h(x)^T & 0 & 0 \end{pmatrix}, \quad C(z) = \begin{pmatrix} \nabla g_{\tilde{\beta}(z)}(x) \\ 0 \\ 0 \end{pmatrix},$$

where $\tilde{\beta}(z) = \bar{\beta}(z) \cup \alpha_0(z) \cup \gamma_0(z)$

The following regular definition comes from [12].

**Definition 3.1**. *Let $z = (x, u, v)$ be an arbitrary vector. We say that $H$ defined by* (2.2) *is regular at $z$ if*

(1) *the matrix $A(z)$ is nonsingular, and*

(2) *the matrix*

$$B(z) = C(z)^T A(z)^{-1} C(z)$$

*is a P-matrix.*

*If $H$ is regular at $z$, we say that $z$ is a regular point of $H$.*

We introduce two useful lemmas (see [12]).

**Lemma 3.3**. *Suppose that $z^k$ is a regular point of $H$. Then there exists a unique KKT point for $VI(X^k, \tilde{f}^k)$, where $\tilde{f}^k(x) = f(x^k) + \nabla_x L(z^k)^T(x - x^k)$.*

**Lemma 3.4**. *Let $z = (x, u, v)$ be a regular point of $H$. Then, there exists a neighbourhood $U(z)$ of $z$ such that every vector in $U(z)$ is also regular. Moreover, when $z^k$ runs over $U(z)$, the solution set of Newton subproblem (i.e., (2.7) with $B_k = \nabla_x L(z^k)^T$) is bounded.*

To analyze the global convergence of Algorithm 1, by Lemma 3.2, we only need to give the global convergence when $N_2$ is finite. To do so, we first take the following assumption:

**Assumption A.** (1) The level set

$$\Omega = \{z \in R^{n+m+l} \mid \theta(z) \leq e^{\sigma}\theta_0\}$$

is bounded.

(2) f is continuously differentiable, $g_i$, $i = 1, 2, \ldots, m$ and $h_j$, $j = 1, 2, \ldots, l$ are twice continuously differentiable.

(3) $\nabla_x L(z)$ is Lipschtiz continuous, i.e., there is a constant $L > 0$ such that

$$\|\nabla_x L(z) - \nabla_x L(z')\| \leq L\|z - z'\|, \quad \forall z, z'.$$

In a similar way to the proof of Theorem 3.1 in [9], we get the following useful lemma.

**Lemma 3.5**. *Let Assumption A and (3.3) hold, then*

$$\lim_{l\to\infty} \frac{1}{l} \sum_{k=0}^{l-1} \eta_k^2 = 0 \tag{3.4}$$

*where*

$$\eta_k = \begin{cases} 0, & \text{if } s^k = 0, \\ \dfrac{\|r_k - B_k s^k\|}{\|s^k\|}, & \text{if } s^k \neq 0. \end{cases}$$

*In particular, there exists a subsequence $\{\eta_k\}_{k\in K}$ converging to zero.*

Since $\{z^k\}_{k\in K} \subseteq \Omega$ is bounded, there exists at least one accumulation point. We now state the global convergence theorem of Algorithm 1 as follows.

**Theorem 3.1**. *Let Assumption A hold and $N_2$ be finite. Suppose that there exists an accumulation point, $z^*$, of $\{z^k\}_{k\in K}$ that is a regular point of $H$. Then $z^*$ is a KKT point of $VI(X, f)$. Moreover, any accumulation point of $\{z^k\}$ is a KKT point of $VI(X, f)$.*

Before proving Theorem 3.1, we first prove some lemmas.

**Lemma 3.6**. *Let the conditions of Theorem* 3.1 *hold. Then there exists a subsequence* $\{z^k\}_{k\in K_1}$ *of* $\{z^k\}_{k\in K}$ *converging to* $z^*$ *such that the sequences* $\{d^k\}_{k\in K_1}$ *and* $\{\frac{d^k}{\|H(z^k)\|}\}_{k\in K_1}$ *are bounded.*

*Proof.* Assume that the subsequence $\{z^k\}_{k\in K_1}$ of $\{z^k\}_{k\in K}$ converges to $z^*$. Since $z^*$ is a regular point, $z^k$ is also a regular point when $k\in K_1$ is sufficiently large, and hence the KKT system of $VI(X^k,\tilde{f}^k)$ has a unique solution. Let $d^k=(y^k,\mu^k,\nu^k)$ be a solution of (2.7) and $\tilde{d}^k=(\tilde{y}^k,\tilde{\mu}^k,\tilde{\nu}^k)$ be the related Newton direction, i.e., the solution of (2.7) where $B_k$ is replaced by $\nabla_x L(z^k)^T$.

Since $z^*$ is a regular point, by Lemma 3.4, $\{\tilde{d}^k\}_{k\in K_1}$ is bounded. Moreover, in a similar way to the proof of Lemma 3 in [12], we can deduce that $\{\tilde{d}^k/\|H(z^k)\|\}_{k\in K_1}$ is bounded.

Let $\bar{z}^k=z^k+d^k=(\bar{x}^k,\bar{u}^k,\bar{v}^k)$ and $\tilde{z}^k=z^k+\tilde{d}^k=(\tilde{x}^k,\tilde{u}^k,\tilde{v}^k)$, where $\bar{u}^k_{\bar{\gamma}}=\tilde{u}^k_{\bar{\gamma}}=0$. Denote

$$\bar{A}(z^k)=\begin{pmatrix}\nabla_x L(z^k)^T & \nabla g_{\bar{\alpha}}(x^k) & \nabla h(x^k)\\ -\nabla g_{\bar{\alpha}}(x^k)^T & 0 & 0\\ -\nabla h(x^k)^T & 0 & 0\end{pmatrix},\quad \bar{C}(z^k)=\begin{pmatrix}\nabla g_{\bar{\beta}}(x^k)\\ 0\\ 0\end{pmatrix},$$

$$p(z^k)=\begin{pmatrix}f(x^k)\\ -g_{\bar{\alpha}}(x^k)\\ -h(x^k)\end{pmatrix},\quad \tilde{w}^k=\begin{pmatrix}\tilde{y}^k\\ \tilde{u}^k_{\bar{\alpha}}\\ \tilde{v}^k\end{pmatrix},\quad \bar{w}^k=\begin{pmatrix}y^k\\ \bar{u}^k_{\bar{\alpha}}\\ \bar{v}^k\end{pmatrix}$$

and

$$p_1(z^k)=p(z^k)+\begin{pmatrix}(B_k-\nabla_x L(z^k)^T)y^k\\ 0\\ 0\end{pmatrix}$$

Then by (2.7), $(\tilde{u}^k_{\bar{\beta}},\tilde{w}^k)$ and $(\bar{u}^k_{\bar{\beta}},\bar{w}^k)$ satisfy

$$\begin{cases}p(z^k)+\bar{A}(z^k)\tilde{w}^k+\bar{C}(z^k)\tilde{u}^k_{\bar{\beta}}=0,\\ \min\{\tilde{u}^k_{\bar{\beta}},-g_{\bar{\beta}}(x^k)-\bar{C}(z^k)^T\tilde{w}^k\}=0\end{cases}\quad\text{and}\quad\begin{cases}p_1(z^k)+\bar{A}(z^k)\bar{w}^k+\bar{C}(z^k)\bar{u}^k_{\bar{\beta}}=0,\\ \min\{\bar{u}^k_{\bar{\beta}},-g_{\bar{\beta}}(x^k)-\bar{C}(z^k)^T\bar{w}^k\}=0\end{cases}$$

So, by Proposition 2 in [11], there exists a constant $L_1>0$ such that $\forall k\in K_1$ sufficiently large,

$$\begin{aligned}\|\tilde{d}^k-d^k\| &= \|\tilde{z}^k-\bar{z}^k\|=\|(\tilde{u}^k_{\bar{\beta}},\tilde{w}^k)-(\bar{u}^k_{\bar{\beta}},\bar{w}^k)\|\\ &\le L_1\|p_1(z^k)-p(z^k)\|=L_1\|(\nabla_x L(z^k)^T-B_k)y^k\|.\end{aligned}\tag{3.5}$$

Let $K_1=K_2\cup K_3$, where $K_2=\{k\in K_1\,|\,s^k=0\}$ and $K_3=\{k\in K_1\,|\,s^k\neq 0\}$. Notice that $s^k=\lambda_k y^k$, it follows from (3.5) that $d^k=\tilde{d}^k$ for each $k\in K_2$. Consequently, $\{d^k\}_{k\in K_2}$ and $\{d^k/\|H(z^k)\|\}_{k\in K_2}$ are bounded.

Consider $k\in K_3$. Denote

$$A_{k+1}=\int_0^1\nabla_x L(x^k+ts^k,u^k,v^k)^T dt.$$

Then by the mean value theorem, $r_k = A_{k+1}s^k$. Moreover, by Lipschitz continuity of $\nabla_x L(z)$, we have

$$\|A_{k+1} - \nabla_x L(z^k)^T\| \le L\|x^{k+1} - x^k\|. \tag{3.6}$$

It then follows that for any $k \in K_3$

$$\begin{aligned} \|(\nabla_x L(z^k)^T - B_k)y^k\|/\|y^k\| &\le \|\nabla_x L(z^k)^T - A_{k+1}\| + \|(A_{k+1} - B_k)y^k\|/\|y^k\| \\ &\le L\|x^{k+1} - x^k\| + \|r_k - B_k s^k\|/\|s^k\| \\ &= L\|x^{k+1} - x^k\| + \eta_k, \end{aligned} \tag{3.7}$$

where the last inequality follows from (3.6) and the fact that $s^k = \lambda_k y^k$. Taking limits in both sides as $k \to \infty$ with $k \in K_3$, by means of (3.3) and Lemma 3.5, we get

$$\lim_{k\in K_3, k\to\infty} \|(\nabla_x L(z^k)^T - B_k)y^k\|/\|y^k\| = \lim_{k\in K_3, k\to\infty} \eta_k = 0. \tag{3.8}$$

Thus by (3.5) and (3.8), we get that $\{d^k\}_{k\in K_3}$ and $\{d^k/\|H(z^k)\|\}_{k\in K_3}$ are bounded. The proof is then completed. □

**Lemma 3.7**. *Let the conditions of Lemma 3.6 hold. Suppose that $\{\delta_k\}$ is a positive sequence satisfying that $\{\delta_k\}_{k\in K_1}$ converge to zero, where $K_1$ is the set defined by Lemma 3.6. Denote $\hat{z}^k = z^k + \delta_k d^k = (\hat{x}^k, \hat{u}^k, \hat{v}^k)$. Then we have*

$$\liminf_{\substack{k \to \infty \\ k \in K_1}} H(z^k)^T(H(z^k) - H(\hat{z}^k))/\delta_k \ge 2\theta(z^*) \tag{3.9}$$

*Proof.* The assumption $\{\delta_k\}_{k\in K_1} \to 0$ implies that $\{\hat{z}^k\}_{k\in K_1} \to z^*$. It is then easy to see that for all $k \in K_1$ sufficiently large

$$\begin{cases} \alpha(z^*) \subset \alpha(z^k) \cap \alpha(\hat{z}^k), \quad \gamma(z^*) \subset \gamma(z^k) \cap \gamma(\hat{z}^k), \\ \alpha_-(z^*) \subset \alpha_-(z^k) \cap \alpha_-(\hat{z}^k), \quad \gamma_-(z^*) \subset \gamma_-(z^k) \cap \gamma_-(\hat{z}^k), \\ \alpha_+(z^*) \subset \alpha_+(z^k) \cap \alpha_+(\hat{z}^k), \quad \gamma_+(z^*) \subset \gamma_+(z^k) \cap \gamma_+(\hat{z}^k). \end{cases} \tag{3.10}$$

Let

$$H(z^k)^T(H(z^k) - H(\hat{z}^k)) = \sum_{r=1}^{5} T_r,$$

where

$$\begin{aligned} T_1 &= L(z^k)^T(L(z^k) - L(\hat{z}^k)), \\ T_2 &= \sum_{i\in\alpha(z^*)} g_i(x^k)(g_i(x^k) - g_i(\hat{x}^k)), \\ T_3 &= \sum_{i\in\beta(z^*)} \min(u_i^k, -g_i(x^k))(\min(u_i^k, -g_i(x^k)) - \min(\hat{u}_i^k, -g_i(\hat{x}^k))), \\ T_4 &= \sum_{i\in\gamma(z^*)} u_i^k(u_i^k - \hat{u}_i^k), \\ T_5 &= \sum_{j=1}^{l} h_j(x^k)(h_j(x^k) - h_j(\hat{x}^k)). \end{aligned}$$

By means of (2.7), (3.10) and Lemma 2 in [12], we estimate $T_i$, $i = 1, 2, 4, 5$ as follows.

$$\begin{aligned}
T_1 &= -L(z^k)^T \nabla L(z^k)(\hat{z}^k - z^k) + o(\|\hat{z}^k - z^k\|) \\
&= \delta_k L(z^k)^T (L(z^k) - (\nabla_x L(z^k)^T - B_k) y^k) + o(\|\hat{z}^k - z^k\|) \\
&= \delta_k L(z^k)^T L(z^k) + o(\|\hat{z}^k - z^k\|), \\
T_2 &= -\sum_{i \in \alpha(z^*)} g_i(x^k) \nabla g_i(x^k)^T (\hat{x}^k - x^k) + o(\|\hat{x}^k - x^k\|) \\
&= -\delta_k \sum_{i \in \alpha(z^*)} g_i(x^k) \nabla g_i(x^k)^T y^k + o(\|\hat{x}^k - x^k\|) \\
&\geq \delta_k \sum_{i \in \alpha(z^*)} (g_i(x^k))^2 + o(\|\hat{x}^k - x^k\|), \\
T_4 &= -\delta_k \sum_{i \in \gamma(z^*)} u_i^k \mu_i^k \geq \delta_k \sum_{i \in \gamma(z^*)} (u_i^k)^2, \\
T_5 &= -\sum_{j=1}^{l} h_j(x^k) \nabla h_j(x^k)(\hat{x}^k - x^k) + o(\|\hat{x}^k - x^k\|) \\
&= \delta_k \sum_{j=1}^{l} (h_j(x^k))^2 + o(\|\hat{x}^k - x^k\|),
\end{aligned}$$

To estimate $T_3$, let $T_3 = T_6 + T_7 + T_8$, where

$$\begin{aligned}
T_6 &= \sum_{i \in \beta_+(z^*)} \min(u_i^k, -g_i(x^k))(\min(u_i^k, -g_i(x^k)) - \min(\hat{u}_i^k, -g_i(\hat{x}^k))), \\
T_7 &= \sum_{i \in \beta_0(z^*)} \min(u_i^k, -g_i(x^k))(\min(u_i^k, -g_i(x^k)) - \min(\hat{u}_i^k, -g_i(\hat{x}^k))), \\
T_8 &= \sum_{i \in \beta_-(z^*)} \min(u_i^k, -g_i(x^k))(\min(u_i^k, -g_i(x^k)) - \min(\hat{u}_i^k, -g_i(\hat{x}^k))),
\end{aligned}$$

with

$$\begin{aligned}
\beta_+(z^*) &= \{i : u_i^* = -g_i(x^*) > 0\}, \\
\beta_0(z^*) &= \{i : u_i^* = -g_i(x^*) = 0\}, \\
\beta_-(z^*) &= \{i : u_i^* = -g_i(x^*) < 0\}.
\end{aligned}$$

For $i \in \beta_+(z^*)$, it is clear that $i \in \bar{\alpha}(z^k) \cup \beta(z^k) \cup \bar{\gamma}(z^k)$ for all $k \in K_1$ sufficiently large. So, we get

$$T_6 \geq \delta_k \sum_{i \in \beta_+(z^*)} (\min(u_i^k, -g_i(x^k)))^2 + o(\|\hat{x}^k - x^k\|).$$

For $i \in \beta_-(z^*)$, we have $i \in \alpha_-(z^k) \cup \beta_-(z^k) \cup \gamma_-(z^k)$ for all $k \in K_1$ sufficiently large. So, we get

$$T_8 \geq \delta_k \sum_{i \in \beta_-(z^*)} u_i^k(-g_i(x^k)) + o(\|\hat{x}^k - x^k\|).$$

Finally, consider the term $T_7$. Noting that the function $\min(u, -g(x))$ is Lipschitzian, and that $\min(u_i^k, -g_i(x^k))$ approaches zero for $i \in \beta_0(z^*)$, it is not difficult to deduce

$$\lim_{\substack{k \to \infty \\ k \in K_1}} T_7/\delta_k = 0$$

Summarizing the above expressions for the various $T_r$ terms, we deduce

$$\liminf_{\substack{k \to \infty \\ k \in K_1}} H(z^k)^T (H(z^k) - H(\hat{z}^k))/\delta_k \geq 2\theta(z^*)$$

This completes the proof. □

**Lemma 3.8**. *Let the conditions of Lemma* 3.7 *hold. Then*

$$\lim_{\substack{k \to \infty \\ k \in K_1}} \|H(z^k) - H(\hat{z}^k)\|^2/\delta_k = 0.$$

*Proof.* Let

$$\|H(z^k) - H(\hat{z}^k)\|^2 = \sum_{r=1}^{5} E_r,$$

where

$$\begin{aligned}
E_1 &= (L(z^k) - L(\hat{z}^k))^T (L(z^k) - L(\hat{z}^k)), \\
E_2 &= \sum_{i \in \alpha(z^*)} (g_i(x^k) - g_i(\hat{x}^k))^2, \\
E_3 &= \sum_{i \in \beta(z^*)} (\min(u_i^k, -g_i(x^k)) - \min(\hat{u}_i^k, -g_i(\hat{x}^k)))^2, \\
E_4 &= \sum_{i \in \gamma(z^*)} (u_i^k - \hat{u}_i^k)^2, \\
E_5 &= \sum_{j=1}^{l} (h_j(x^k) - h_j(\hat{x}^k))^2.
\end{aligned}$$

In a similar way to the proof of Lemma 3.7, we can deduce that for all $k \in K_1$ sufficiently large

$$\begin{aligned}
E_1 &= \delta_k^2 L(z^k)^T L(z^k) + o(\|\hat{x}^k - x^k\|), \\
E_2 &= \delta_k^2 \sum_{i \in \alpha(z^*)} (\nabla g_i(x^k)^T y^k)^2 + o(\|\hat{x}^k - x^k\|), \\
E_4 &= \delta_k^2 \sum_{i \in \gamma(z^*)} (\mu_i^k)^2 \leq \delta_k^2 \|d^k\|^2, \\
E_5 &= \delta_k^2 \sum_{j=1}^{l} (h_j(x^k))^2 + o(\|\hat{x}^k - x^k\|).
\end{aligned}$$

For the term $E_3$, noting that the function $\min(u, -g(x))$ is Lipschitzian, thus there exists a constant $L_2 > 0$ such that for all $k \in K_1$ sufficiently large

$$E_3 \leq L_2^2 \|d^k\|^2 \delta_k^2.$$

The above analysis and Assumption A show that

$$\lim_{k \to \infty, k \in K_1} E_i/\delta_k = 0, \qquad i = 1, 2, \ldots, 5.$$

Thus, this completes the proof. □

Now, we prove the global convergence of Algorithm 1.

*Proof of Theorem 3.1.* Set

$$G(z^k, d^k) = \begin{pmatrix} L(z^k) + \nabla_x L(z^k)^T y^k + \nabla g(x^k)\mu^k + \nabla h(x^k)\nu^k \\ -g_{\bar{\alpha}}(x^k) - \nabla g_{\bar{\alpha}}(x^k)^T y^k \\ \min(u^k_{\bar{\beta}} + \mu^k_{\bar{\beta}}, -g_{\bar{\beta}}(x^k) - \nabla g_{\bar{\beta}}(x^k)^T y^k) \\ u^k_{\bar{\gamma}} + \mu^k_{\bar{\gamma}} \\ -h(x^k) - \nabla h(x^k)^T y^k \end{pmatrix}. \tag{3.11}$$

Then by (2.7) we get

$$\|G(z^k, d^k)\| = \|(\nabla_x L(z^k)^T - B_k)y^k\|. \tag{3.12}$$

By Lemma 3.6, $\{d^k\}_{k\in K_1}$ is bounded, without loss of generality, we may assume that $d^k \to d^*(k \in K_1,\ k \to \infty)$. It follows from (3.12) and (3.8) that

$$\lim_{k\in K_1, k\to\infty} G(z^k, d^k) = 0. \tag{3.13}$$

Let $\lambda^* = \liminf_{k\in K_1} \lambda_k$. Then $\lambda^* \geq 0$. If $\lambda^* > 0$, then by (3.3), $d^* = 0$. Therefore, from (3.11) and (3.13), we have $H(z^*) = 0$. If $\lambda^* = 0$, then $\{\lambda_k\}_{k\in K_1}$ has a subsequence converging to zero. Without loss of generality, we assume that $\{\lambda_k\}_{k\in K_1} \to 0$. Set $\delta_k = \lambda_k/\rho$ for $k \in K_1$. Then by Step 2 of Algorithm 1, when $k \in K_1$ is sufficiently large, $\delta_k$ does not satisfy (2.10), that is, we have $q_k(\delta_k) < \frac{1}{2} + \epsilon$, or equivalently,

$$\begin{aligned} & H(z^k)^T(H(z^k) - H(z^k + \delta_k d^k))/\delta_k \\ < \ & (\frac{1}{2} + \epsilon) \max\{\delta_k^{-1}\|H(z^k) - H(z^k + \delta_k d^k)\|^2, \delta_k\|d^k\|^2\}. \end{aligned} \tag{3.14}$$

Therefore, it follows from Lemma 3.7 and Lemma 3.8 that $\theta(z^*) = 0$. In other words, $z^*$ is a KKT point of VI$(X, f)$. Since $\theta(z^k)$ converges by Lemma 3.1, we assert that every accumulation point of $z^k$ is a zero point of $\theta$, and hence a KKT point of VI$(X, f)$. □

As a corollary of Lemma 3.6 and Theorem 3.1, we have

**Corollary 3.1**. *Let the conditions of Lemma 3.6 hold. Then* $\{d^k\}_{k\in K_1}$ *converges to zero.*

## 4. Superlinear Convergence

In this section, we prove the superlinear convergence of Algorithm 1. To do so, we need the following assumption:

**Assumption B.** (1) The sequence $\{z^k\}$ generated by Algorithm 1 converges to $z^*$ and $H(z^*) = 0$.

(2) $H$ is regular at $z^*$.

**Lemma 4.9**. *Let Assumption A and B hold. Then there exists a neighborhood* $U(z^*)$ *of* $z^*$ *and positive constants* $C_1 < C_2$ *such that for any* $z \in U(z^*)$,

$$C_1\|z - z^*\| \leq \|H(z)\| \leq C_2\|z - z^*\|. \tag{4.1}$$

*Proof.* Since $z^*$ is a regular point of the function $H(z)$, by Lemma 3.4, there exists a neighborhood $U(z^*)$ of $z^*$ such that for any $z \in U(z^*)$, $z$ is also a regular point of function $H(z)$ and $\alpha(z^*) \subset \alpha(z), \gamma(z^*) \subset \gamma(z)$. By the mean-value theorem, there exists a point $\tilde{z} = z^* + t(z - z^*)$, $0 \leq t \leq 1$ such that

$$\begin{aligned} H(z) &= H(z) - H(z^*) = \begin{pmatrix} L(z) - L(z^*) \\ (-g_i(x) + g_i(x^*))_{i \in \alpha(z)} \\ (u_i - u_i^*)_{i \in \beta(z) \cup \gamma(z)} \\ -h(x) + h(x^*) \end{pmatrix} \\ &= \begin{pmatrix} \nabla_x L(\tilde{z})^T (x - x^*) + \nabla g(\tilde{x})(u - u^*) + \nabla h(\tilde{x})(v - v^*) \\ -\nabla g_{\alpha(z)}(\tilde{x})^T (u - u^*)_{\alpha(z)} \\ (u - u^*)_{\beta(z) \cup \gamma(z)} \\ -\nabla h(\tilde{x})^T (v - v^*) \end{pmatrix} + o(\|z - z^*\|) e \\ &\triangleq P(z)\omega(z) + o(\|z - z^*\|) e, \end{aligned}$$

where $e \in R^{n+m+l}$ stands for the vector whose components are 1 and

$$\begin{aligned} P(z) &= \begin{pmatrix} \nabla_x L(\tilde{z})^T & \nabla g_{\alpha(z)}(\tilde{x}) & \nabla h(\tilde{x}) & \nabla g_{\beta(z) \cup \gamma(z)}(\tilde{x}) \\ -\nabla g_{\alpha(z)}(\tilde{x})^T & 0 & 0 & 0 \\ -\nabla h(\tilde{x})^T & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{\beta(z) \cup \gamma(z)} \end{pmatrix}, \\ \omega(z) &= \begin{pmatrix} x - x^* \\ (u - u^*)_{\alpha(z)} \\ v - v^* \\ (u - u^*)_{\beta(z) \cup \gamma(z)} \end{pmatrix}. \end{aligned}$$

It is not difficult to see from the regularity that $P(z)$ is uniformly nonsingular as $z$ is close to $z^*$ and hence $P(z)^{-1}$ is bounded in a neighbourhood of $z^*$. (4.1) then follows from the last equality. □

**Lemma 4.10**. *Let Assumption A and B hold. Then*

(1) *we have*

$$\sum_{k=0}^{\infty} \|z^{k+1} - z^k\|^2 < \infty.$$

(2) *For all $k$ sufficiently large, we have*

$$\|z^k + d^k - z^*\| \leq M(\|z^k - z^*\| + \|s^k\| + \eta_k)\|d^k\|, \tag{4.2}$$

*where $M$ is a suitable positive constant.*

*Proof.* The conclusion (1) can be obtained by Lemma 4.9 and the fact $q_k(\lambda_k) \geq \frac{1}{2} + \epsilon$. We only need to prove the conclusion (2).

Denote $S^k = \{i : g_i(x^k) + \nabla g_i(x^k)^T y^k = 0, i \in \bar{\beta}(z^k)\}$ and $\bar{S}^k = \bar{\beta}(z^k)\backslash S^k$. Then from (2.7), we have

$$\begin{cases} L(z^k) + B_k y^k + \nabla g(x^k)\mu^k + \nabla h(x^k)\nu^k = 0, & \\ g_i(x^k) + \nabla g_i(x^k)^T y^k = 0, & i \in \bar{\alpha}(z^k) \cup S^k. \\ h(x^k) + \nabla h(x^k)^T y^k = 0, & \\ u_i^k + \mu_i^k = 0, & i \in \bar{\gamma}(z^k) \cup \bar{S}^k. \end{cases} \tag{4.3}$$

Define matrices

$$P_k = \int_0^1 \nabla L(z^* + \tau(z^k - z^*))^T d\tau, \;\; Q_k = \int_0^1 \nabla g_{\bar{\alpha}(z^k)\cup S^k}(x^* + \tau(x^k - x^*))^T d\tau$$

and

$$R_k = \int_0^1 \nabla h(x^* + \tau(x^k - x^*))^T d\tau.$$

Then by the mean value theorem, we have

$$L(z^k) = L(z^k) - L(z^*) = P_k(z^k - z^*), \;\; g_{\bar{\alpha}(x^k)\cup S^k}(x^k) = Q_k(x^k - x^*), \;\; h(x^k) = R_k(x^k - x^*).$$

It then follows from (4.3) that

$$\begin{cases} P_k(z^k - z^*) + \nabla L(z^k)^T d^k - \nabla_x L(z^k)^T y^k + B_k y^k &=& 0, \\ Q_k(x^k - x^*) + \nabla g_{\bar{\alpha}\cup S^k}(x^k)^T y^k &=& 0, \\ R_k(x^k - x^*) + \nabla h(x^k)^T y^k &=& 0, \\ (u^k + \mu^k - u^*)_{\bar{\gamma}\cup\bar{S}^k} &=& 0, \end{cases} \tag{4.4}$$

where $\bar{\alpha}^k = \bar{\alpha}(z^k)$ and $\bar{\gamma}^k = \bar{\gamma}(z^k)$. (4.4) can be rewritten as

$$\begin{cases} \nabla L(z^k)^T(z^k + d^k - z^*) &=& (\nabla_x L(z^k)^T - B_k)y^k + (\nabla L(z^k)^T - P_k)(z^k - z^*), \\ \nabla g_{\bar{\alpha}^k\cup S^k}(x^k)^T(x^k + y^k - x^*) &=& (\nabla g_{\bar{\alpha}^k\cup S^k}(x^k)^T - Q_k)(x^k - x^*), \\ \nabla h(x^k)^T(x^k + y^k - x^*) &=& (\nabla h(x^k) - R_k)(x^k - x^*) \\ (u^k + \mu^k - u^*)_{\bar{\gamma}^k\cup\bar{S}^k} &=& 0. \end{cases} \tag{4.5}$$

Denote

$$U_k = \begin{pmatrix} \nabla_x L(z^k)^T & \nabla g_{\bar{\alpha}^k\cup S^k}(x^k) & \nabla h(x^k) & \nabla g_{\bar{\gamma}^k\cup\bar{S}^k}(x^k) \\ \nabla g_{\bar{\alpha}\cup S^k}(x^k)^T & 0 & 0 & 0 \\ \nabla h(x^k)^T & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{\bar{\gamma}^k\cup\bar{S}^k} \end{pmatrix}$$

and

$$t_k = \begin{pmatrix} (\nabla_x L(z^k)^T - B_k)y^k + (\nabla L(z^k)^T - P_k)(z^k - z^*), \\ (\nabla g_{\bar{\alpha}^k\cup S^k}(x^k)^T - Q_k)(x^k - x^*), \\ (\nabla h(x^k) - R_k)(x^k - x^*) \\ 0_{\bar{\gamma}^k\cup\bar{S}^k}. \end{pmatrix}$$

Then (4.5) is represented as a compact form

$$U_k \begin{pmatrix} x^k + y^k - x^* \\ (u^k + \mu^k - u^*)_{\bar{\alpha}^k \cup S^k} \\ v^k + \nu^k - v^* \\ (u^k + \mu^k - u^*)_{\bar{\gamma}^k \cup \bar{S}^k} \end{pmatrix} = t_k. \tag{4.6}$$

Moreover, by means of the Lipschitz continuity of $\nabla_x L(z)$, the twice differentiability of $g$ and $h$, it is not difficult to deduce that there is a positive constant $M_1$ such that

$$\|t_k\| \leq M_1(\|z^k - z^*\| + \|s^k\| + \eta_k)\|d^k\|. \tag{4.7}$$

By the regularity and the assumption that $z^k \to z^*$, we claim that when $k$ is sufficiently large, $U_k^{-1}$ exits and is bounded. Therefore, we get (4.2) from (4.6) and (4.7). □

**Lemma 4.11**. *Let Assumption A and B hold. Then there exists a positive constant $\tilde{\eta} < 1$ such that when $\eta_k \leq \tilde{\eta}$ and $k$ is sufficiently large, $\lambda_k \equiv 1$.*

*Proof.* From Lemma 3.6, Lemma 4.9 and Lemma 4.10, there exists a constant $C > 0$ such that for all $k$ sufficiently large

$$\begin{aligned} \|H(z^k + d^k)\| \leq C_2\|z^k + d^k - z^*\| &\leq C_2 M(\|z^k - z^*\| + \|s^k\| + \eta_k)\|d^k\| \\ &\leq CM(\|z^k - z^*\| + \|s^k\| + \eta_k)\|H(z^k)\|, \end{aligned}$$

where the last inequality follows from Lemma 3.6. Since $z^k \to z^*$ and $s^k \to 0$ as $k \to \infty$, the last inequality shows that there exists a positive constant $\tilde{\eta} < 1$ such that when $\eta_k \leq \tilde{\eta}$ and $k$ is sufficiently large

$$\|H(z^k + d^k)\| \leq \frac{1}{2}\|H(z^k)\|$$

It is not difficult to verify that for these k, $q_k(1) \geq \frac{1}{2} + \epsilon$, This means $\lambda_k \equiv 1$. □

In a similar way to Theorem 4.4 in [10], we can prove the superlinear convergence of Algorithm 1.

**Theorem 4.2**. *Let Assumption A and B hold. Then we have*

$$\sum_{k=0}^{\infty} \|z^{k+1} - z^k\| < \infty.$$

*Moreover, $\{z^k\}$ converges superlinearly.*

**Acknowledgment.** The authors would like to thank Professor Shu-zi Zhou for his guidance about the paper. We also thank the two anonymous referees for their helpful comments which improved the presentation of this paper.

## References

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

[2] X. Chen and L. Qi, A parameterized Newton method and a quasi-Newton method for solving nonsmooth equations, *Computational Optimization and Applications*, 3(1994), 157-179.

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

[4] A. Griewank, The 'global' convergence of Broyden-like methods with a suitable line search, *Journal of Australia Mathematical Society, Ser. B* 28(1986) 75-92.

[5] P. T. Harker and J. S. Pang, Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications, *Mathematical Programming* 48(1990) 161-220.

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

[7] N.H. Josephy, Newton's method for generalized equations, Technical Report 1965, Mathematical Research Center, University of Wisconsin, Madison, 1979.

[8] N.H. Josephy, Quasi-Newton methods for generalized equations, Technical Report 1966, Mathematical Research Center, University of Wisconsin, Madison, 1979.

[9] D. H. Li, Global convergence of nonsingular Broyden's methods for unconstrained optimizations, *Chinese Journal of Numerical Mathematics and Applications*, 17, 1995 No.4, 85-94.

[10] D. H. Li, J. P. Zeng and S. Z. Zhou, A line search technique for nonlinear complementarity problems, to appear in "Proceeding of ISOR'96 Conference". Guilin, China, 1996.

[11] J. S. Pang, Newton's method for B-differentiable equations, *Mathematics of Operations Research*, 15(1990), 311-341.

[12] J. S. Pang, A B-differentiable equation-based, globally and locally quadratically convergent algorithm for nonlinear programs, complementarity and variational inequality problems, *Mathematical Programming*, 51(1991), 101-131.

[13] J.S. Pang and D. Chan, Iterative methods for variational and complementarity problems, Mathematical Programming 24 (1982), 284-313.

[14] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, *Mathematics of Operations Research* 18(1993) 217-224.

[15] L. Qi and X. Chen, A globally convergent successive approximation method for severely nonsmooth equations, *SIAM Journal on Control and Optimization* 33(1995) 402-418.

[16] 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), 463-480.

[17] K. Taji, F. Fukushima and T. Ibaraki, A globally convergent Newton method for solving strongly variational inequalities, *Mathematical Programming* 58(1993) 369-383.

[18] S. Z. Zhou, D. H. Li and J. P. Zeng, A successive approximation quasi-Newton process for nonlinear complementarity problems, in D-Z Du, L. Qi and R. Womersley eds. "Recent Advances in Nonsmooth Optimizations", World Scientific Publishers, New York, 1995, 459-472.