A Survey of Some Nonsmooth Equations and Smoothing Newton Methods *

Liqun Qi and Defeng Sun

School of Mathematics
The University of New South Wales
Sydney 2052, Australia

(April 20, 1998; Revised August 13, 1998)

Abstract

In this article we review and summarize recent developments on nonsmooth equations and smoothing Newton methods. Several new suggestions are presented.

1 Introduction

Suppose that $H : \Re^n \to \Re^n$ is locally Lipschitz but not necessarily continuously differentiable. To solve

$$H(x) = 0 \tag{1.1}$$

has become one of most active research directions in mathematical programming. The early study of nonsmooth equations can be traced back to [Eav71, Man75, Man76]. The system of nonsmooth equations arises from many applications. Pang and Qi [PaQ93] reviewed eight problems in the study of complementarity problems, variational inequality problems and optimization problems, which can be reformulated as systems of nonsmooth equations. In this paper, we review recent developments of algorithms for solving nonsmooth equations. Section 2 is devoted to semismooth Newton methods and Section 3 discusses smoothing Newton methods. We make several final remarks in Section 4.

2 Semismooth Newton methods

2.1 Local Semismooth Newton methods

Some extensions of Newton and quasi-Newton methods have been developed; Josephy [Jos79a, Jos79b] for strongly regular generalized equations which were extensively discussed by Robinson [Rob80, Rob82, Rob83], Pang and Chan [PaC82] for variational inequalities and complementarity problems. Some early studies of solving systems of nonsmooth equations by Newton's methods include Kojima and Shindo [KoS86] for piece-wise smooth

*This work is supported by the Australian Research Council. E-mail: L.Qi@unsw.edu.au (Liqun Qi), sun@maths.unsw.edu.au (Defeng Sun)

equations, Pang [Pan90] and Robinson [Rob94] for B-differentiable equations and Kummer [Kum88] for locally Lipschitz functions. Subsequently, a number of generalized Newton methods were proposed for nonsmooth equations and related problems, see, for example [DiF95, Fis92, HaX90, Kum88, OuZ95, Pan90, Pan91, Qi93, QiS93, PaQ93, PaG93, Ral94, Mor94, XiH90a, XiH90b]. Some methods solve nonlinear subproblems at each step, see [Jos79a, Rob94, Pan90, Pan91, Ral94, DiF95, XiH90a, XiH90b]. In some special cases, these nonlinear subproblems are solvable. Superlinear convergence results were established based on this. Qi and Sun [QiS93] analysed a generalized Newton method for solving (1.1), in which a system of linear equations is solved at each step. Superlinear convergence of their generalized Newton method is based on a key concept of so-called semismoothness. Independently, Kummer [Kum92] presented a general analysis of superlinear convergence for generalized Newton methods of nonsmooth equations under similar conditions used in Qi and Sun [QiS93].

Suppose that H of (1.1) is locally Lipschitz. By Rademacher's Theorem, H is almost everywhere differentiable. Let

$$D_H = \{x \mid H \text{ is differentiable at } x\}.$$

Then the generalized Jacobian of H at x can be defined by

$$\partial H(x) = \text{conv}\partial_B H(x),$$

where

$$\partial_B H(x) = \{ \lim_{\substack{x^j \to x \\ x^j \in D_H}} H'(x^j)\}.$$

The generalized Newton method for solving (1.1) can be defined as follows: Having the vector x^k, compute x^{k+1} by

$$x^{k+1} = x^k - V_k^{-1} H(x^k), \tag{2.1}$$

where $V_k \in \partial H(x^k)$.

The generalized Newton method (2.1) reduces to the classical Newton method for a system of equations if H is continuously differentiable. The classical Newton method has a favorable feature that the sequence $\{x^k\}$ generated by (2.1) is locally superlinearly (quadratically) convergent to a solution x^* if $H'(x^*)$ is nonsingular (and H' is Lipschitz continuous) [OrR70, DeS83]. However, in general the iterative method (2.1) is not convergent for nonsmooth equations (1.1). See [Kum88] for a counterexample.

In order to establish superlinear convergence of the generalized Newton method (2.1), we state the concept of semismoothness. Let H be directionally differentiable at x. H is said to be semismooth at x if

$$Vd - H'(x; d) = o(\|d\|), \quad d \to 0$$

and H is called strongly semismooth at x if

$$Vd - H'(x; d) = O(\|d\|^2), \quad d \to 0,$$

where $V \in \partial H(x + d)$. Semismoothness was originally introduced by Mifflin [Mif77] for functionals. Semismooth functionals play an important role in the global convergence theory of nonsmooth optimization, see Polak [Pol97]. Qi and Sun [QiS93] extended the concept of semismoothness to vector-valued functions.

Using semismoothness, Qi and Sun [QiS93] presented the following convergence theorem for the generalized Newton method (2.1).

Theorem 2.1 *Suppose that $H(x^*) = 0$ and that all $V \in \partial H(x^*)$ are nonsingular. Then the generalized Newton method (2.1) is Q-superlinearly convergent in a neighborhood of x^* if H is semismooth at x^*, and quadratically convergent if H is strongly semismooth at x^*.*

In [Kum92], Kummer independently discussed sufficient and necessary conditions for the convergence of Newton method based on generalized derivatives. One of the conditions for guaranteeing convergence (see Theorem 2 of [Kum92]) is that for any $V \in \partial H(x + d),\ d \to 0$,

$$H(x + d) - H(x) - Vd = o(\|d\|). \tag{2.2}$$

Since H is locally Lipschitz continuous, from [Sha90] we know that if $H'(x; d)$ exists, then $H'(x; d)$ coincides with the B-derivative of H at x; i.e.,

$$\lim_{d \to 0} \frac{H(x + d) - H(x) - H'(x; d)}{\|d\|} = 0.$$

So, if $H'(x; d)$ exists, then (2.2) implies that for any $V \in \partial H(x + d),\ d \to 0$,

$$Vd - H'(x; d) = o(\|d\|).$$

Hence (2.2) implies the semismoothness of H at x if $H'(x; d)$ exists. On the other hand, the semismoothness of H at x implies (2.2) since $H'(x; d)$ exists in this case [Page 465, SuH97].

Note that the nonsingularity of $\partial H(x^*)$ in the above theorem is somewhat restrictive in some cases. Qi [Qi93] presented a modified version of (2.1) which may be stated as follows

$$x^{k+1} = x^k - V_k^{-1} H(x^k), \tag{2.3}$$

where $V_k \in \partial_B H(x^k)$. The difference of this version from (2.1) is that V_k is chosen from $\partial_B H(x^k)$ rather than the convex hull of $\partial_B H(x^k)$. Analogous to Theorem 2.1, Qi [Qi93] established the following result.

Theorem 2.2 *Suppose that $H(x^*) = 0$ and that all $V \in \partial_B H(x^*)$ are nonsingular. Then the generalized Newton method (2.3) is Q-superlinearly convergent in a neighborhood of x^* if H is semismooth at x^*, and quadratically convergent at x^* if H is strongly semismooth at x^*.*

For inexact versions of (2.1) and (2.3) and their superlinear convergence theorems, see [MaQ95, FaK97a].

Pang and Qi [PaQ93] generalized the superlinear convergence results of Dennis-Moré [DeM77] for quasi-Newton methods of smooth equations.

Theorem 2.3 *Assume that H is semismooth at x^* and that all elements in $\partial_B H(x^*)$ are nonsingular. Let $\{x^k\} \subseteq D$ be any sequence that converges to x^* with $x^k \neq x^*$ for all k. Then $\{x^k\}$ converges Q-superlinearly to x^* and $H(x^*) = 0$ if and only if*

$$\lim_{k\to\infty} \frac{\|H(x^k) + V_k d^k\|}{\|d^k\|} = 0, \tag{2.4}$$

where $V_k \in \partial_B H(x^k)$ and $d^k = x^{k+1} - x^k$.

Theorems 2.1, 2.2 and 2.3 generalize the convergence results of classical Newton method for smooth equations without assuming differentiability of H. On the other hand, those convergence results can be applied to some important mathematical programming problems such as nonlinear complementarity problems, variational inequalities, the Karush-Kuhn-Tucker (KKT) conditions of optimization problems and other related problems. Contrary to Newton's methods for solving nonsmooth equations, the direct application of quasi-Newton methods to nonsmooth equations is not very successful. Nevertheless, several quasi-Newton methods for solving structured nonsmooth equations have been successfully obtained. See [JQCS96] for a review about this topic.

2.2 Globalization of Semismooth Newton Methods

It is noted that (2.1) is only convergent locally under the semismoothness assumption. A natural question is that if (2.1) can be globalized similar to classical Newton's method for solving smooth equations. In general, the answer is negative because θ is not continuously differentiable, where for any $x \in \Re^n$, $\theta(x)$ is defined by

$$\theta(x) = \frac{1}{2}\|H(x)\|^2. \tag{2.5}$$

Fortunately, in some special but important cases, θ can be continuously differentiable though H itself is not smooth. For example, if $H(x) = \max(0, x), x \in \Re$, then H is not differentiable at $x = 0$, but θ is continuously differentiable on $\Re$. For the sake of globalization, probably the most useful function is the Fischer-Burmeister function $\phi : \Re^2 \to \Re^2$ [FiB92]

$$\phi(a, b) = \sqrt{a^2 + b^2} - (a + b). \tag{2.6}$$

It can be verified that ϕ is strongly semismooth everywhere [JiQ97, FaS97, Fis97]. Let $\psi : \Re^2 \to \Re$ be defined by

$$\psi(a, b) = \phi(a, b)^2.$$

Then ψ is continuously differentiable on $\Re^2$, which was first discovered by Kanzow [Kan94a]. This property is very nice and, yet, is surprising at the first glance. The Fischer-Burmeister function has attracted a number of authors' attention and has been used extensively to study the nonlinear complementarity problem NCP(F): Finding $x \in \Re^n$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad x^T F(x) = 0,$$

where $F : D \subseteq \Re^n \to \Re^n$. Since $\phi(a, b) = 0$ if and only if $a, b \geq 0$ and $ab = 0$. To solve NCP(F) is then equivalent to find a solution of the nonsmooth equation $H(x) = 0$ with H give by

$$H_i(x) = \phi(x_i, F_i(x)), \quad i = 1, ..., n. \tag{2.7}$$

Suppose that $D = \Re^n$ and F is continuously differentiable on $\Re^n$. Then by the continuous differentiability of ψ [Kan94a], θ is continuously differentiable on $\Re^n$. See [Bil95, CCK97, DFK96, FaK97a, FaS97, FFK96, FFK97, FFK98 Fis95, Fis97, GeK96, Jia96, Jia97a, Kan94b, KaK98, KaQ97, Qi97, QiJ97, QiS98, SuW98, Tse96, YaF97] for more discussions about Fischer-Burmeister function and its generalizations. Among those papers, De Luca et al. provided a damped semismooth Newton method for solving NCP(F). Although, De Luca et al.'s algorithm was designed for solving $H(x) = 0$ with H given by (2.7), it is apparent that it can be applied to solve any nonsmooth equations $H(x) = 0$ by requiring θ to be continuously differentiable.

It has long been known [Eav71, Man76, Pan90] that to solve NCP(F) (suppose that $D = \Re^n$) is equivalent to solve another equation $H(x) = 0$ with H given by

$$H_i(x) = \min(x_i, F_i(x)), \quad i = 1, ..., n. \tag{2.8}$$

It is also known (e.g., [KoS86, Rob92]) that to solve NCP(F) is equivalent to solve the normal equation

$$H(y) := F(y_+) + y - y_+ = 0 \tag{2.9}$$

in the sense that if x is a solution of NCP(F), then $y := x - F(x)$ is a solution of (2.9); conversely, if y is a solution of (2.9), then $x := y_+$ is a solution of NCP(F). Contrary to the case based on (2.7), the function θ defined by (2.5) is not continuously differentiable on $\Re^n$ if H is defined by (2.8) or (2.9). Even so, globally convergent methods based on (2.9) have been successfully designed [DiF95,Ral94,Rob94].

Next, we state a damped semismooth Newton method due to De Luca et el. [DFK96] by assuming that θ is continuously differentiable. We leave the case that θ is not continuously differentiable to the next section, where smoothing methods will be introduced to address the nonsmoothness issue.

Algorithm 2.1 (Damped Semismooth Newton Method [DFK96])

Step 0. *Given* $x^0 \in \Re^n$, $\beta > 0$, $p > 2$, $\rho, \sigma \in (0, 1/2)$. $k := 0$.

Step 1. *Select an element* $V_k \in \partial_B H(x^k)$ *and solve*

$$H(x^k) + V_k d = 0. \tag{2.10}$$

Let d^k *be the solution of (2.10) if it is solvable. If (2.10) is unsolvable or if the condition*

$$\nabla\theta(x^k)^T d^k \le -\beta\|d^k\|^p$$

is not satisfied, let $d^k = -\nabla\theta(x^k)$.

Step 2. *Let* m_k *be the smallest nonnegative integer* m *such that*

$$\theta(x^k + \rho^m d^k) - \theta(x^k) \le 2\sigma\rho^m \nabla\theta(x^k)^T d^k.$$

Set $t_k = \rho^{m_k}$ *and* $x^{k+1} = x^k + t_k d^k$.

Step 3. *Replace* k *by* $k + 1$ *and go to Step 1.*

De Luca et al. proved the following two theorems.

Theorem 2.4 *Suppose that H is semismooth and θ is continuously differentiable. Then any accumulation point x^* of $\{x^k\}$ generated by Algorithm 2.1 is a stationary point of θ, i.e., $\nabla\theta(x^*) = 0$. If x^* is a solution of $H(x) = 0$ and all matrices in $\partial_B H(x^*)$ are nonsingular, then the whole sequence $\{x^k\}$ converges to x^* superlinearly and quadratically if H is strongly semismooth at x^*.*

Theorem 2.5 *Suppose that H is defined by (2.7) and F is continuously differentiable on $\Re^n$. Then any accumulation point x^* of $\{x^k\}$ generated by Algorithm 2.1 is a stationary point of θ, i.e., $\nabla\theta(x^*) = 0$. If $F'(x^*)$ is a P_0 matrix, in particular, a positive semidefinite matrix, then x^* is a solution of $H(x) = 0$. Moreover, if NCP(F) is R-regular [Rob83] at x^*, then the whole sequence $\{x^k\}$ converges to x^* superlinearly, and quadratically if F' is Lipschitz continuous around x^*.*

Related algorithms to Algorithm 2.1 have been designed in [FaS97, JiQ97, Jia97a, JiR97] etc. A recent study of Fischer-Burmeister function with various hybrid techniques is included in [DFK97].

3 Smoothing Methods

3.1 Introduction

Recently there has been an increasing interest in smoothing methods for solving nonsmooth equations arising from complementarity problems and variational inequalities, e.g., see [Sma86, ChH93, ChH95, ChH97, ChM95, ChM96, GaM97, Kan96, KaJ98, QiC95]. The feature of smoothing methods is to construct a smoothing approximation function $G : \Re^n \times \Re_{++} \to \Re^n$ of H such that for any $\varepsilon > 0$ and $x \in \Re^n$, $G(\cdot, \varepsilon)$ is continuously differentiable on $\Re^n$ and satisfies

$$\|H(x) - G(x, \varepsilon)\| \to 0 \quad \text{as } \varepsilon \downarrow 0 \tag{3.1}$$

and then to find a solution of (1.1) by (inexactly) solving the following problems for a given positive sequence $\{\varepsilon^k\}$, $k = 0, 1, 2...$,

$$G(x, \varepsilon^k) = 0. \tag{3.2}$$

The well-known smoothing function is the so called Chen-Harker-Kanzow-Smale (CHKS) function for approximating $\max(0, w), w \in \Re$. The CHKS function is defined by

$$\xi(w, \varepsilon) = \frac{\sqrt{w^2 + 4\varepsilon^2} + w}{2}, \quad (w, \varepsilon) \in \Re^2. \tag{3.3}$$

See [Sma86, ChH93, Kan96]. Suppose that ξ is defined by (3.3) and define $P : \Re^n \times \Re \to \Re$ by

$$P_i(y, \varepsilon) = \xi(y_i, \varepsilon), \ (y, \varepsilon) \in \Re^n \times \Re, \ i = 1, ..., n. \tag{3.4}$$

Then the smoothing function G of H, defined by (2.8) and (2.9), can be described by

$$G(x, \varepsilon) := x - P(x - F(x), \varepsilon), \quad (x, \varepsilon) \in \Re^n \times \Re_{++} \tag{3.5}$$

and

$$G(y,\varepsilon) := F(P(y,\varepsilon)) + y - P(y,\varepsilon), \quad (y,\varepsilon) \in \Re^n \times \Re_{++}, \tag{3.6}$$

respectively. For any $a, b, \varepsilon \in \Re$, define

$$\phi(a,b,\varepsilon) = \sqrt{a^2+b^2+\varepsilon^2} - (a+b). \tag{3.7}$$

This function is a smoothed form of Fischer-Burmeister function and was first defined by Kanzow [Kan96]. Jiang [Jia97b] proves that $\psi(\cdot,\cdot,\cdot) := \phi(\cdot,\cdot,\cdot)^2$ is continuously differentiable on $\Re^3$. Define $G : \Re^{n+1} \to \Re^n$ by

$$G_i(x,\varepsilon) := \phi(x_i, F_i(x), \varepsilon), \quad i = 1,2,...,n. \tag{3.8}$$

Suppose that F is continuously differentiable on $\Re^n$. The function G defined by (3.5), (3.6) and (3.7), respectively, satisfies (3.1) and for any $\varepsilon \in \Re_{++}$, $G(\cdot,\varepsilon)$ is continuously differentiable on $\Re^n$.

Smoothing techniques have also been used to solve mathematical programming problems with equilibrium constraints [FJQ98, FLP98, JiR98].

3.2 Jacobian Smoothing Newton method

The smoothing Newton method is also called as splitting Newton method or non-interior point Newton method or homotopy Newton method. It uses the derivative of G with respect to the first variable in the Newton method, namely

$$x^{k+1} = x^k - t_k G'_x(x^k,\varepsilon^k)^{-1} H(x^k), \tag{3.9}$$

where $\varepsilon^k > 0$, $G'_x(x^k,\varepsilon^k)$ denotes the derivative of G with respect to the first variable at (x^k,ε^k) and $t_k > 0$ is the stepsize. Smoothing Newton method (3.9) for solving nonsmooth equation (1.1) has been studied for decades in different areas (see [CQS98] for references). In some previous papers, method (3.9) is called a splitting method because $H(\cdot)$ is split into a smooth part $G(\cdot,\varepsilon)$ and a nonsmooth part $H(\cdot) - G(\cdot,\varepsilon)$. The global and linear convergence of (3.9) has been discussed in [QiC95, Qi95]. In [CQS98], the authors define a Jacobian consistency property and show that the smoothing approximation functions in [ChM96, GaM97] have this property. Under suitable conditions, they proved that the sequence $\{x^k\}$ generated by the smoothing Newton method is bounded and each accumulation point is a solution of (1.1). Let $\partial_C H$ be defined by

$$\partial_C H(x) = \partial H_1(x) \times \partial H_2(x) \times \ldots \times \partial H_n(x).$$

This definition can be seen as a special case of the C-differential operator discussed in [Qi96].

Definition 3.1 *Let H be a Lipschitz continuous function in $\Re^n$. We call $G : \Re^n \times \Re_{++} \to \Re^n$ a smoothing approximation function of H if G is continuously differentiable with respect to the first variable and there is a constant $\mu > 0$ such that for any $x \in \Re^n$ and $\varepsilon \in \Re_{++}$,*

$$\|G(x,\varepsilon) - H(x)\| \le \mu\varepsilon. \tag{3.10}$$

Furthermore, if for any $x \in \Re^n$,

$$\lim_{\varepsilon \downarrow 0} \operatorname{dist}(G'_x(x,\varepsilon), \partial_C H(x)) = 0, \tag{3.11}$$

then we say G satisfies the Jacobian consistency property.

In this subsection we assume that G has the Jacobian consistency property, and present a smoothing Newton method with a line search based on G. This smoothing Newton method was called Jacobian smoothing Newton method in [KaP97] in order to differentiate it from other smoothing methods.

We denote

$$\theta_k(x) = \frac{1}{2}\|G(x, \varepsilon_k)\|^2.$$

Algorithm 3.2 (Jacobian Smoothing Newton Method [CQS98])

Step 0. *Given* $\rho, \alpha, \eta \in (0, 1)$, $\gamma \in (0, +\infty)$ *and a starting point* $x^0 \in \Re^n$. *Choose* $\sigma \in (0, \frac{1}{2}(1-\alpha))$, *and* $\mu > 0$ *satisfying (3.10). Let* $\beta_0 = \|H(x^0)\|$ *and* $\varepsilon^0 = \frac{\alpha}{2\mu}\beta_0$. $k := 0$.

Step 1. *Solve*

$$H(x^k) + G'_x(x^k, \varepsilon^k)d^k = 0. \tag{3.12}$$

Let d^k *be the solution of (3.12).*

Step 2. *Let* m_k *be the smallest nonnegative integer* m *such that*

$$\theta_k(x^k + \rho^m d^k) - \theta_k(x^k) \le -2\sigma\rho^m\theta(x^k)$$

Set $t_k = \rho^{m_k}$ *and* $x^{k+1} = x^k + t_k d^k$.

Step 3. 3.1 *If* $\|H(x^{k+1})\| = 0$, *terminate.*

3.2 *If* $\|H(x^{k+1})\| > 0$ *and*

$$\|H(x^{k+1})\| \le \max\{\eta\beta_k, \alpha^{-1}\|H(x^{k+1}) - G(x^{k+1}, \varepsilon^k)\|\}, \tag{3.13}$$

we let

$$\beta_{k+1} = \|H(x^{k+1})\|$$

and choose an ε^{k+1} *satisfying*

$$0 < \varepsilon^{k+1} \le \min\{\frac{\alpha}{2\mu}\beta_{k+1}, \frac{\varepsilon^k}{2}\}$$

and

$$\mathrm{dist}(G'_x(x^{k+1}, \varepsilon^{k+1}), \partial_C H(x^{k+1})) \le \gamma\beta_{k+1}.$$

3.3 *If* $\|H(x^{k+1})\| > 0$ *but (3.13) does not hold, we let* $\beta_{k+1} = \beta_k$ *and* $\varepsilon^{k+1} = \varepsilon^k$.

Step 4. *Replace* k *by* $k+1$ *and go to Step 1.*

Assumption 3.1 (i) *The level set*

$$D_0 = \{x \in \Re^n : \theta(x) \le (1+\alpha)^2\theta(x^0)\}$$

is bounded.

(ii) *For any* $\varepsilon \in \Re_{++}$ *and* $x \in D_0$, $G'_x(x, \varepsilon)$ *is nonsingular.*

In [CQS98], Chen et al. proved

Theorem 3.1 *Suppose that Assumption 3.1 holds. Then Algorithm 3.2 is well defined and the generated sequence* $\{x^k\}$ *remains in* D_0 *and satisfies*

$$\lim_{k\to 0} H(x^k) = 0.$$

Theorem 3.2 *Suppose that Assumption 3.1 holds. Suppose that for an accumulation point* x^* *of the sequence* $\{x^k\}$*, all* $V \in \partial_C H(x^*)$ *are nonsingular and that* H *is semismooth at* x^**. Then* x^* *is a solution of* $H(x) = 0$ *and the sequence* $\{x^k\}$ *generated by Algorithm 3.2 converges to* x^* *superlinearly. Moreover, if* F *is strongly semismooth at* x^**, then* $\{x^k\}$ *converges to* x^* *quadratically.*

There are several modifications of Algorithm 3.2. Chen and Ye [ChY98] combined a finite termination strategy with Algorithm 3.2 to get finitely convergent results for box constrained linear variational inequality problems. Kanzow and Pieper [KaP97] introduced a gradient step to handle the singular case of the iteration matrices and provided numerical results on computing complementarity problems.

3.3 Squared Smoothing Newton Methods

Algorithm 3.2 provided a general convergence result on the variable x. Its convergence analysis strongly depends on the Jacobian consistency property. It has been verified in [CQS98] that many smoothing functions satisfy it. However, on the other hand, the smoothing functions based on normal maps [Rob92], which only require the mapping to be defined on the feasible region instead of on $\Re^n$, do not satisfy this property. See [QSZ97, ZSQ98, SuQ97] for the smoothing forms of normal maps. Moreover, even if G satisfies the Jacobian consistency property, its regularized forms, e.g., $G(x, \varepsilon) := G(x, \varepsilon) + \varepsilon x$, which are useful to get stronger results, do not satisfy the Jacobian consistency property. In addition, more smoothing functions which do not satisfy the Jacobian consistency property arise. See Section 3.4 for a new example. In order to circumvent one or several of these difficulties, a class of squared smoothing Newton methods were introduced in [QSZ97]. For the sake of convenience, we suppose that for any $\varepsilon < 0$ and $x \in \Re^n$, $G(x, \varepsilon) = G(x, -\varepsilon)$ and $G(x, 0) = H(x)$. Define $\Phi : \Re^{n+1} \to \Re^{n+1}$ by

$$\Phi(x, \varepsilon) = \begin{pmatrix} G(x, \varepsilon) \\ \varepsilon \end{pmatrix}, \tag{3.14}$$

where G is continuously differentiable at any $z := (x, \varepsilon) \in \Re^n \times \Re_{++}$ and satisfies (3.1).

Choose $\bar{\varepsilon} \in \Re_{++}$ and $\gamma \in (0, 1)$ such that $\gamma\bar{\varepsilon} < 1$. Let $\bar{z} := (0, \bar{\varepsilon}) \in \Re^n \times \Re$. Define the merit function $\eta : \Re^{n+1} \to \Re_+$ by

$$\eta(z) := \|\Phi(z)\|^2 \tag{3.15}$$

and define $\beta : \Re^{n+1} \to \Re_+$ by

$$\beta(z) := \gamma \min\{1, \eta(z)\}.$$

Let

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

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

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

Algorithm 3.3 (Squared Smoothing Newton Method [QSZ97])

Step 0. *Choose constants* $\delta \in (0,1)$, $\sigma \in (0,1/2)$. *Let* $\varepsilon^0 := \bar{\varepsilon}$, $x^0 \in \Re^n$ *be an arbitrary point and* $k := 0$.

Step 1. *If* $\Phi(z^k) = 0$ *then stop. Otherwise, let* $\beta_k := \beta(z^k)$.

Step 2. *Compute* $d^k := (\Delta x^k, \Delta \varepsilon^k) \in \Re^n \times \Re^n$ *by*

$$\Phi(z^k) + \Phi'(z^k) d^k = \beta_k \bar{z}. \tag{3.16}$$

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

$$\eta(z^k + \delta^l d^k) \leq [1 - 2\sigma(1 - \gamma\bar{\varepsilon})\delta^l]\eta(z^k). \tag{3.17}$$

Define $z^{k+1} := z^k + \delta^{l_k} d^k$.

Step 4. *Replace* k *by* $k+1$ *and go to Step 1.*

Remark. Algorithm 3.3 was first proposed in [QSZ97] and was then used in [Sun98, Qih97, ZSQ98] for solving regularized reformulations of complementarity problems and variational inequality problems [FaK97b, RaG97, SzG98] and in [QiL97] for solving extended order linear complementarity problems.

Assumption 3.2 *G is continuously differentiable at any* $(x, \varepsilon) \in \Re^n \times \Re_{++}$ *and* $G'_x(x,\varepsilon)$ *is nonsingular.*

Qi et al. [QSZ97] proved the following two theorems.

Theorem 3.3 *Suppose that Assumption 3.2 is satisfied. Then an infinite sequence* $\{z^k\}$ *is generated by Algorithm 3.3 and each accumulation point* $\tilde{z}$ *of* $\{z^k\}$ *is a solution of* $\Phi(z) = 0$.

Stronger global convergent results have been obtained in [Sun98, Qih97, ZSQ98] by considering various regularized forms of Φ.

Theorem 3.4 *Suppose that Assumption 3.2 is satisfied and* z^* *is an accumulation point of the infinite sequence* $\{z^k\}$ *generated by Algorithm 3.3. Suppose that* Φ *is semismooth at* z^* *and that all* $V \in \partial\Phi(z^*)$ *are nonsingular. Then the whole sequence* $\{z^k\}$ *converges to* z^*,

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

and

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

Furthermore, if Φ *is strongly semismooth at* z^*, *then*

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

and

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

In Theorem 3.4 we obtained the superlinear (quadratic) convergence of ε^k and $\|z^k - z^*\|$ instead of $\|x^k - x^*\|$ while in Theorem 3.2 we obtain the superlinear (quadratic) of $\|x^k - x^*\|$, but neither ε^k nor $\|z^k - z^*\|$.

Suppose that G is defined by (3.8) and that F is continuously differentiable. Jiang [Jia97b] provided a different form of Φ, which was defined by

$$\Phi(x, \varepsilon) := \begin{pmatrix} G(x, \varepsilon) \\ e^{\varepsilon} - 1 \end{pmatrix}. \tag{3.18}$$

An interesting property of such defined Φ is that for any $\varepsilon > 0$ and any $\lambda \in (0, 1]$,

$$\varepsilon + \lambda \Delta \varepsilon > 0$$

and

$$\varepsilon + \lambda \Delta \varepsilon < \varepsilon,$$

where $d := (\Delta x, \Delta \varepsilon) \in \Re^n \times \Re$ is a solution of

$$\Phi(x, \varepsilon) + \Phi'(x, \varepsilon) d = 0.$$

Based on this observation, Jiang [Jia97b] designed a smoothing Newton method for solving NCP(F) with the assumption that F is a P_0 function. By using the continuous differentiability of $\eta := \|\Phi\|^2$ and the assumption that the search directions are bounded, which can be satisfied by assuming that F is a uniform P-function, Jiang [Jia97b] proved global and local superlinearly (quadratically) convergent results of his method. It is noted that Jiang's approach can be used to any Φ with η being continuously differentiable. Moreover, the assumption that the search directions are bounded might be removed off if similar techniques used in Algorithm 3.3 are applied. Another approach to remove off the boundedness assumption of the search directions made in [Jia97b] is to use gradient directions combined with Newton directions [QiL98].

In Algorithm 3.3, we have assumed that the iteration matrices are nonsingular. This may not be satisfied if for every $\varepsilon > 0$, $G(\cdot, \varepsilon)$ is not a P_0-function. See [GoT97] for a discussion on the P_0 property of $G(\cdot, \varepsilon)$. In this case, we suggest to use a modified method, which we will introduce below, if η is continuously differentiable on $\Re^{n+1}$. Let $Z := \Re^n \times \Re_+$ and Π_Z be the orthogonal projection operator onto Z.

Algorithm 3.4 (Modified Squared Smoothing Newton Method)

Step 0. *Choose constants* $\delta \in (0, 1)$, $\sigma \in (0, 1/2)$, $\alpha > 0$, *and* $p > 2$. *Let* $\varepsilon^0 := \bar{\varepsilon}$, $x^0 \in \Re^n$ *be an arbitrary point and* $k := 0$.

Step 1. *If* $\Phi(z^k) = 0$ *then stop. Otherwise, let* $\beta_k := \beta(z^k)$.

Step 2. *Compute* $d^k := (\Delta x^k, \Delta \varepsilon^k) \in \Re^n \times \Re$ *by*

$$\Phi(z^k) + \Phi'(z^k) d^k = \beta_k \bar{z}. \tag{3.19}$$

If (3.19) is not solvable or if the condition

$$\nabla \eta(z^k)^T d^k \le -\alpha \|d^k\|^p \tag{3.20}$$

does not hold, let

$$d^k := (\Delta x^k, \Delta \varepsilon^k) = \Pi_Z[z^k - \nabla \eta(z^k)] - z^k.$$

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

$$\eta(z^k + \delta^l d^k) \le \eta(z^k) + \sigma\delta^l \nabla\eta(x^k)^T d^k \tag{3.21}$$

and

$$\varepsilon^k + \delta^l \Delta\varepsilon^k > 0.$$

Define $z^{k+1} := z^k + \delta^{l_k} d^k$.

Step 4. *Replace k by $k+1$ and go to Step 1.*

Remark. If d^k is a solution of (3.19), then

$$\nabla\eta(z^k)^T d^k \le -2(1-\gamma\bar{\varepsilon})\eta(z^k)$$

and

$$\varepsilon^k + \mu\Delta\varepsilon^k > 0 \quad \forall \mu \in [0,1].$$

Theorem 3.5 *Suppose that η is continuously differentiable on $\Re^{n+1}$. Then Algorithm 3.4 generates an infinite sequence $\{z^k\}$ with $z^k \in \Re^n \times \Re_{++}$ and any accumulation point $\tilde{z}$ of $\{z^k\}$ satisfies*

$$\tilde{z} - \Pi_Z[\tilde{z} - \nabla\eta(\tilde{z})] = 0. \tag{3.22}$$

Proof: First, it is easy to see that Algorithm 3.4 is well defined and generates an infinite sequence $\{z^k\}$ with $z^k \in \Re^n \times \Re_{++}$. Suppose that $\tilde{z}$ is an accumulation point of $\{z^k\}$ but does not satisfy (3.22). Then from $\tilde{z} - \Pi_Z[\tilde{z} - \nabla\eta(\tilde{z})] \neq 0$ and $\tilde{\varepsilon} \ge 0$, we obtain that $\nabla\eta(\tilde{z}) \neq 0$. By taking a subsequence if necessary, assume that $z^k \to \tilde{z}$. We consider two cases:

(i) There are infinitely many k such that condition (3.20) does not hold. Define

$$K = \{k |\ (3.20) \text{ does not hold at the } kth \text{ step}\}.$$

Suppose that $K = \{k_1, k_2, ...\}$. Then $d^{k_j} \to \tilde{d} := \Pi_Z[\tilde{z} - \nabla\eta(\tilde{z})] - \tilde{z} \neq 0$ as $j \to \infty$.

Therefore, there exist a neighborhood $\mathcal{N}(\tilde{z})$ and a positive number $\bar{\mu} \in (0,1]$ such that for all $z := (x, \varepsilon) \in \mathcal{N}(\tilde{z})$, $d := (\Delta x, \Delta\varepsilon) = \Pi_Z[z - \nabla\eta(z)] - z$, and $\mu \in (0, \bar{\mu})$, we have

$$\eta(z + \mu d) \le \eta(z) + \sigma\mu\nabla\eta(z)^T d$$

and if $\varepsilon > 0$, then

$$\varepsilon + \mu\Delta\varepsilon = (1-\mu)\varepsilon + \mu\Pi_{\Re_+}[\varepsilon - \nabla_\varepsilon\eta(z)] > 0 \quad \forall \mu \in (0,1).$$

Thus we can conclude that

$$\nabla\eta(z^{k_j})^T d^{k_j} \to 0,$$

which is a contradiction because $-\nabla\eta(z^{k_j})^T d^{k_j} \ge \|d^{k_j}\|^2$ and $d^{k_j} \to \tilde{d} \neq 0$.

(ii) There are infinitely many k such that condition (3.20) holds. Define

$$K = \{k |\ (3.20) \text{ holds at the } kth \text{ step}\}.$$

Suppose that $K = \{k_1, k_2, ...\}$. Then $\{d^{k_j}\}$ is bounded. By taking a subsequence if necessary, assume that $d^{k_j} \to \tilde{d}$ and $\Phi'(z^{k_j}) \to V \in \partial\Phi(\tilde{z})$ as $j \to \infty$. Then $\tilde{d}$ satisfies

$$\Phi(\tilde{z}) + V\tilde{d} = \beta(\tilde{z})\bar{z} \tag{3.23}$$

and

$$\nabla\eta(\tilde{z})^T\tilde{d} \leq -\alpha\|\tilde{d}\|^p.$$

Since η is continuously differentiable on $\Re^{n+1}$, we obtain from (3.23) that

$$\nabla\eta(\tilde{z})^T\tilde{d} \leq -2(1-\gamma\bar{\varepsilon})\eta(\tilde{z}). \tag{3.24}$$

Apparently, $\eta(\tilde{z}) \neq 0$, and so, $\nabla\eta(\tilde{z})^T\tilde{d} < 0$. Then there exist a neighborhood $\mathcal{N}(\tilde{z})$ of $\tilde{z}$ and a positive number $\bar{\mu} \in (0,1]$ such that for all $z := (x,\varepsilon) \in \mathcal{N}(\tilde{z})$ with $\varepsilon > 0$, $\mu \in (0,\bar{\mu})$ and all $d := (\Delta x, \Delta\varepsilon) \in \Re^n \times \Re$ sufficiently close to $\tilde{d}$ satisfying

$$\Phi(z) + \Phi'(z)d = \beta(z)\bar{z}$$

and

$$\nabla\eta(z)^T d \leq -\alpha\|d\|^p,$$

we have

$$\eta(z+\mu d) \leq \eta(z) + \sigma\mu\nabla\eta(z)^T d$$

and

$$\varepsilon + \mu\Delta\varepsilon > 0.$$

This implies that

$$\nabla\eta(\tilde{z})^T\tilde{d} = \nabla\eta(z^{k_j})^T d^{k_j} \to 0,$$

which is a contradiction because by (3.24), $\nabla\eta(\tilde{z})^T\tilde{d} = 0$ implies $\eta(\tilde{z}) = 0$.

We have completed the proof of the theorem. ∎

In Theorem 3.5 we prove that any accumulation point $\tilde{z}$ of $\{z^k\}$ is a solution of (3.22). This point $\tilde{z}$ may be not a solution of $\Phi(z) = 0$. However, under some suitable conditions any solution of (3.22) is a solution of $\Phi(z) = 0$ as well. For example, let us consider NCP(F). Suppose that F is a continuously differentiable P_0 function on $\Re^n$ and G is defined by (3.8). From (3.22), $V_x^T G(\tilde{z}) = 0$, $\tilde{\varepsilon} \geq 0$, $V_\varepsilon^T G(\tilde{z}) + \tilde{\varepsilon} \geq 0$ and $\tilde{\varepsilon}(V_\varepsilon^T G(\tilde{z}) + \tilde{\varepsilon}) = 0$, where $V \in \partial G(\tilde{z})$ and $V = (V_x\ V_\varepsilon)$ with $V_x \in \Re^n \times \Re^n$ and $V_\varepsilon \in \Re^n$. If $\tilde{\varepsilon} > 0$, then $V_x = G'_x(\tilde{z})$ is a P-matrix [Kan96]. So, $G(\tilde{z}) = 0$, $\tilde{\varepsilon} \geq 0$ and $\tilde{\varepsilon}^2 = 0$. This contradicts our assumption that $\tilde{\varepsilon} > 0$. Then $\tilde{\varepsilon}$ must be 0. From [FaS97], we know that $V_x^T G(\tilde{z}) = 0$ implies that $G(\tilde{z}) = 0$. Then $\tilde{z}$ is a solution of $\Phi(z) = 0$. Comparing to Algorithm 3.3, the advantage of Algorithm 3.4 is that during the process of iteration the iterate matrices are not required to be nonsingular. On the other hand, if for every $\varepsilon > 0$, $G(\cdot,\varepsilon)$ is known to be a P_0 function, by considering $G(x,\varepsilon) := G(x,\varepsilon) + \varepsilon x$ if necessary, then from Theorem 3.4 we know that any accumulation point of $\{z^k\}$ generated by Algorithm 3.3 is a solution of $\Phi(z) = 0$. In general, the same result may not hold for Algorithm 3.4. For an example, see $G(x,\varepsilon) = F(x) + \varepsilon x$, where F is a continuously differentiable P_0-function. See [GoT97] for more discussions about the P_0 property of various G.

Theorem 3.6 *Suppose that η is continuously differentiable on $\Re^{n+1}$ and z^* is an accumulation point of the infinite sequence $\{z^k\}$ generated by Algorithm 3.4. Suppose that z^* is a solution of $\Phi(z) = 0$ and that Φ is semismooth at z^*. If all $V \in \partial\Phi(z^*)$ are nonsingular, then the whole sequence $\{z^k\}$ converges to z^*,*

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

and

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

Furthermore, if Φ is strongly semismooth at z^*, then

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

and

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

Proof: The proof of this theorem is similar to that of Theorem 3.4 by noting that when $z \to z^*$,

$$\|\beta(z)\bar{z}\| = O(\|\Phi(z)\|^2),$$

which, together with the nonsingularity of all $V \in \partial\Phi(z^*)$, guarantees that (3.20) holds for all z^k sufficiently close to z^*. ∎

3.4 Smoothing Newton Methods with Neighborhood Constraints

The paper of Burke and Xu [BuX98a] initialized the study of smoothing Newton methods with various neighborhood constraints. A typical class of neighborhoods can be described as follows. Let

$$\mathcal{N}(\beta) := \{x \in \Re^n : \frac{\|G(x,\varepsilon)\|}{\varepsilon} \leq \beta, \text{ with } 0 < \varepsilon\}, \tag{3.25}$$

for $\beta > 0$. Given an $\varepsilon > 0$ and a starting point in $\mathcal{N}(\beta)$. Then apply Newton's methods to $G(x, \varepsilon) = 0$ with ε fixed to find a new $x \in \mathcal{N}(\beta)$ and a decreased ε. Along this line, see [BuX98b, ChX98, ChC97, PeL97, Tse98, Xu96, Xu97]. Since $\mathcal{N}(\beta)$ is not a very restricted set, it is usually easy to find a staring point in this neighborhood. In order to make $\mathcal{N}(\beta)$ bounded, we need a restrictive condition, e.g., for NCP(F) F is assumed to be a P_0 +R_0 function [ChC97]. So, several refined neighborhoods are introduced. For example

$$\mathcal{N}_1(\beta) := \{x \in \Re^n : G(x,\varepsilon) \in \Omega, \frac{\|G(x,\varepsilon)\|}{\varepsilon} \leq \beta, \text{ with } 0 < \varepsilon\}, \tag{3.26}$$

where $\beta > 0$ and Ω is a convex set containing the origin. Along this line, see [HoY96, ChC98, QiS98, BuX97, BuX98c].

Except [Tse98], where NCP(F) was considered and an active set strategy was introduced, the superlinear convergence of this class of smoothing Newton methods requires the assumption of the differentiability of H at a solution point. There is still no smoothing Newton method to have superlinear convergence without this restrictive assumption if the refined neighborhoods are applied.

For NCP(F), these smoothing methods require F to be well defined on $\Re^n$, which may not hold for a few problems. Moreover, the P_0 property or monotonicity of F may only hold on $\Re^n_+$ instead of on $\Re^n$. These require us to define new smoothing functions to use the definition of F on $\Re^n_+$ only. Apparently, the normal map and its smoothing forms are a choice. In the following, we give a new approach.

It is easy to verify that to solve NCP(F) is also equivalent to solve

$$H(x) := \min(x, F(x_+)) = x - [x - F(x_+)]_+ = 0. \tag{3.27}$$

Suppose that ξ and P are defined by (3.3) and (3.4), respectively. Define $G : \Re^n \times \Re \to \Re^n$ by

$$G(x, \varepsilon) := x - P(x - F(P(x, \varepsilon)), \varepsilon), \ (x, \varepsilon) \in \Re^n \times \Re.$$

Then G is a smoothing approximation of H and F is only required to have definition on $\Re^n_+$. This defined smoothing function G and related forms may be useful to improve the properties of the above smoothing Newton methods with neighborhood constraints.

4 Final Remarks

In this paper we reviewed some recent developments of Newton's methods for solving nonsmooth equations: semismooth Newton methods and smoothing Newton methods. These two classes of methods are different but related to each other, in particular in the local high order convergence analysis. Some related topics include

- continuously differentiable optimization approach (see [Fuk96] and references therein).
- interior point methods (see [Wri97] and references therein).

Finally, we would like to mention that some standard test libraries like `GAMSLIB` and `MCPLIB` have been constructed by Ferris and his colleagues [DFPR94, DiF95b, FeR96]. These test problems within the MATLAB programming environment are publically available and can be downloaded by anonymous ftp from `ftp.cs.wisc.edu:math-prog/matlab/`. In fact, these problems should be seriously tested for any numerical methods designed for solving nonsmooth equations developed in the future.

Acknowledgement

The authors would like to thank Dr. Houduo Qi for his comments on an earlier version of this paper.

References

[Bil95] S.C. Billups, *Algorithms for complementarity problems and generalized equations*, Ph.D thesis, Department of Computer Sciences, University of Wisconsin–Madison, USA, 1995.

[BiF97] S.C. Billups and M.C. Ferris, "QPCOMP: a quadratic programming based solver for mixed complementarity problems", *Mathematical Programming, 76* (1997), 533–562.

[BuX97] J. Burke and S. Xu, "A non-interior predictor-corrector path following algorithm for the monotone linear complementarity problem", Preprint, Department of Mathematics, University of Washington, Seattle, WA 98195, September 1997.

[BuX98a] J. Burke and S. Xu, "A polynomial time interior–point path–following algorithm for LCP based on Chen–Harker–Kanzow smoothing techniques", to appear in *Mathematical Programming.*

[BuX98b] J. Burke and S. Xu, "The global linear convergence of a non-interior path-following algorithm for linear complementarity problem", to appear in *Mathematics of Operations Research.*

[BuX98c] J. Burke and S. Xu, "A non-interior predictor-corrector path following method for LCP", *Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, M. Fukushima and L. Qi, eds., Kluwer Academic Publisher, Nowell, Maryland, 1998, pp. 45–64.

[ChC97] B. Chen and X. Chen, "A global linear and local quadratic continuation smoothing method for variational inequalities with box constraints", Preprint, Department of Management and Systems, Washington State University, Pullman, March 1997.

[ChC98] B. Chen and X. Chen, "A global and local superlinear continuation-smoothing method for $P_0 + R_0$ and monotone NCP", to appear in *SIAM Journal on Optimization.*

[CCK97] B. Chen, X. Chen and C. Kanzow, "A penalized Fischer-Burmeister NCP-function: theoretical investigation and numerical results", Applied Mathematics Report 97/28, School of Mathematics, the University of New South Wales, Sydney 2052, Australia, September 1997.

[ChH93] B. Chen and P.T. Harker, "A non-interior-point continuation method for linear complementarity problems", *SIAM Journal on Matrix Analysis and Applications, 14* (1993), 1168–1190.

[ChH95] B. Chen and P.T. Harker, "A continuation method for monotone variational inequalities", *Mathematical Programming, 69* (1995), 237–253.

[ChH97] B. Chen and P.T. Harker, "Smooth approximations to nonlinear complementarity problems", *SIAM Journal on Optimization, 7* (1997), 403–420.

[ChX98] B. Chen and N. Xiu, "A global linear and local quadratic non-interior continuation method for nonlinear complementarity problems based on Chen-Mangasarian smoothing function", to appear in *SIAM Journal on Optimization.*

[ChM95] C. Chen and O.L. Mangasarian, "Smoothing methods for convex inequalities and linear complementarity problems", *Mathematical Programming, 71* (1995), 51–69.

[ChM96] C. Chen and O.L. Mangasarian, "A class of smoothing functions for nonlinear and mixed complementarity problems", *Computational Optimization and Applications, 5* (1996), 97–138.

[CQS98] X. Chen, L. Qi and D. Sun, "Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities", *Mathematics of Computation, 67* (1998), 519-540.

[ChY98] X. Chen and Y. Ye, "On homotopy-smoothing methods for variational inequalities", to appear in *SIAM Journal on Control and Optimization.*

[Cla83] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[DFK96] 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.

[DFK97] T. De Luca, F. Facchinei and C. Kanzow, "A theoretical and numerical comparison of some semismooth algorithms for complementarity problems", Mathematical Programming Technical Report 97-15, Computer Sciences Department, University of Wisconsin-Madison, Madison, WI, December 1997.

[DeM77] J.E. Dennis, Jr. and J.J. Moré, "Quasi-Newton methods: Motivation and theory", *SIAM Review, 19* (1977), 46-89.

[DeS83] J.E. Dennis, Jr. and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, New Jersey, 1983.

[DiF95] S.P. Dirkse and M.C. Ferris, "The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems", *Optimization Methods and Software, 5* (1995), 123–156,

[DiF95b] S.P. Dirkse and M.C. Ferris, "MCPLIB: A collection of nonlinear mixed complementarity problems", *Optimization Methods and Software, 5* (1995), 319–345.

[DFPR94] S.P. Dirkse, M.C. Ferris, P.V. Preckel, and T. Rutherford, "The GAMS callable program library for variational and complementarity solvers", Mathematical Programming Technical Report 94-07, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, 1994.

[Eav71] B.C. Eaves, "On the basic theorem of complementarity", *Mathematical Programming, 1* 1971, 68–75.

[FFK96] F. Facchinei, A. Fischer and C. Kanzow, "Inexact Newton methods for semismooth equations with applications to variational inequality problems", in *Nonlinear Optimization and Applications*, G. Di Pillo and F. Giannessi, eds., Plenum Press, New York, 1996, pp. 125-139.

[FFK97] F. Facchinei, A. Fischer and C. Kanzow, A semismooth Newton method for variational inequalities: The case of box constraints, in *Complementarity and Variational Problems: State of the Art*, M.C. Ferris and J.S. Pang, eds., SIAM, Philadelphia, Pennsylvania, 1997, pp. 76—90.

[FFK98] 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.

[FJQ98] F. Facchinei, H. Jiang, and L. Qi, "A smoothing method for mathematical programming with equilibrium constraints", to appear in *Mathematical Programming.*

[FaK97a] F. Facchinei and C. Kanzow, "A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems", *Mathematical Programming, 76* (1997), 493–512.

[FaK97b] F. Facchinei and C. Kanzow, "Beyond monotonicity in regularization methods for nonlinear complementarity problems", to appear in *SIAM Journal on Control and Optimization.*

[FaS97] 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.

[FeR96] M. C. Ferris and T. F. Rutherford, "Accessing realistic mixed complementarity problems within MATLAB", in Nonlinear Optimization and Applications, G. Di Pillo and F. Giannessi, eds., Plenum Press, New York, 1996, pp. 141–153.

[Fis92] A. Fischer, "A special Newton-type optimization method", *Optimization, 24* (1992), 269–284.

[Fis95] 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 Publishers, New Jersy, 1995, pp. 88–105.

[Fis97] A. Fischer, "Solution of monotone complementarity problems with locally Lipschitzian functions", *Mathematical Programming, 76* (1997), 513–532.

[Fuk96] M. Fukushima, "Merit functions for variational inequality and complementarity problems", in *Nonlinear Optimization and Applications*, G. Di Pillo and F. Giannessi eds., Plenum Publishing Corporation, New York, 1996, pp. 155–170.

[FLP98] M. Fukushima, Z.-Q. Luo, and J.-S. Pang, "A globally convergent sequential quadratic programming algorithm for mathematical programming problems with linear complementarity constraints", *Computational Optimization and Applications, 10* (1998), 5–34.

[GaM97] S.A. Gabriel and J.J. Moré, "Smoothing of mixed complementarity problems", in *Complementarity and Variational Problems: State of the Art*, M.C. Ferris and J.S. Pang, eds., SIAM, Philadelphia, Pennsylvania, 1997, pp. 105–116.

[GeK96] C. Geiger and C. Kanzow, "On the resolution of monotone complementarity problems", *Computational Optimization and Applications, 5* (1996), 155–173.

[GoT97] M.S. Gowda and M.A. Tawhid, "Existence and limiting behavior of associated with P_0-equations", to appear in *Computational Optimization and Applications*.

[HaP90] 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.

[HaX90] P.T. Harker and B. Xiao, "Newton method for the nonlinear complementarity problem: A B-differentiable equation approach", *Mathematical Programming, 48* (1990), 339–357.

[HoY96] 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, Japan, December 1996.

[Jia96] H. Jiang, "Unconstrained minimization approaches to nonlinear complementarity problems", *Journal of Global Optimization*, 9 (1996), 169–181.

[Jia97a] H. Jiang, "Global convergence analysis of the generalized Newton and Gauss-Newton methods of the Fischer-Burmeister equation for the complementarity problem", to appear in *Mathematics of Operations Research*.

[Jia97b] H. Jiang, "Smoothed Fischer-Burmeister equation methods for the nonlinear complementarity problem", Preprint, Department of Mathematics & Statistics, the University of Melbourne, Victoria 3052, Australia, June 1997.

[JiQ97] H. Jiang and L. Qi, "A new nonsmooth equation approach to nonlinear complementarity problems", *SIAM Journal on Control and Optimization, 35* (1997), 178–193.

[JQCS] H. Jiang, L. Qi, X. Chen and D. Sun, "Semismoothness and superlinear convergence in nonsmooth optimization and nonsmooth equations", in *Nonlinear Optimization and Applications*, G. Di Pillo and F. Giannessi eds., Plenum Publishing Corporation, New York, 1996, pp.197–212.

[JiR97] H. Jiang and D. Ralph, "Smooth SQP methods for mathematical programs with nonlinear complementarity constraints", Manuscript, Department of Mathematics & Statistics, University of Melbourne, December 1997.

[JiR98] H. Jiang and D. Ralph, "Global and local superlinear convergence analysis of Newton-type methods for semismooth equations with smooth least squares", *Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, M. Fukushima and L. Qi, eds. Kluwer Academic Publisher, Nowell, Maryland, 1998, pp. 181–210.

[Jos79a] N.H. Josephy, "Newton's method for generalized equations", Technical Report 1965, Mathematics Research Center, University of Wisconsin, Madison, 1979.

[Jos79b] N.H. Josephy, "Quasi-Newton methods for generalized equations", Technical Report 1966, Mathematics Research Center, University of Wisconsin, Madison, 1979.

[Kan94a] C. Kanzow, "An unconstrained optimization technique for large-scale linearly constrained convex minimization problems", *Computing, 53* (1994), 101–117.

[Kan94b] C. Kanzow, "Some equation-based methods for the nonlinear complementarity problem", *Optimization Methods and Software, 3* (1994), 327–340.

[Kan96] C. Kanzow, "Some noninterior continuation methods for linear complementarity problems", *SIAM Journal on Matrix Analysis and Applications, 17* (1996), 851–868.

[KaJ98] C. Kanzow and H. Jiang, "A continuation method for (strongly) monotone variational inequalities", *Mathematical Programming, 81* (1998), 140–157.

[KaK98] C. Kanzow and H. Kleinmichel, "A new class of semismooth Newton-type methods for nonlinear complementarity problems", to appear in *Computational Optimization and Applications.*

[KaP97] C. Kanzow and H. Pieper, "Jacobian smoothing methods for general complementarity problems", to appear in *SIAM Journal on Optimization.*

[KaQ97] C. Kanzow and H.-D. Qi, "A QP-free constrained Newton-type method for variational inequality problems", to appear in *Mathematical Programming.*

[KoS86] M. Kojima and S. Shindo, "Extensions of Newton and quasi-Newton methods to systems of PC^1 equations", *Journal of Operations Research Society of Japan, 29* (1986), 352-374.

[Kum88] B. Kummer, "Newton's method for non-differentiable functions", in *Advances in Mathematical Optimization*, J. Guddat, B. Bank, H. Hollatz, P. Kall, D. Klatte, B. Kummer, K. Lommatzsch, L. Tammer, M. Vlach and K. Zimmerman, eds., Akademi-Verlag, Berlin, pp. 114-125, 1988.

[Kum92] B. Kummer, "Newton's method based on generalized derivatives for nonsmooth functions: Convergence analysis", in: W. Oettli and D. Pallaschke, eds., "Advances in Optimization", Springer-Verlag, Berlin, pp. 171-194, 1992.

[Man75] O.L. Mangasarian, "Unconstrained Lagrangians in nonlinear programming", *SIAM Journal of Control, 13* (1975), 772–791.

[Man76] O.L. Mangasarian, "Unconstrained methods in nonlinear programming", in; *Nonlinear Programming*, SIAM-AMS Proceedings, Volume IX, American Math. Soc., Providence, Rhode Island, pp. 169-184.

[MaQ95] J.M. Martínez and L. Qi, "Inexact Newton methods for solving nonsmooth equations", *Journal of Computational and Applied Mathematics,* 60 (1995), 127-145.

[Mif77] R. Mifflin, "Semismooth and semiconvex functions in constrained optimization", *SIAM Journal on Control and Optimization, 15* (1977), 957–972.

[Mor74] J.J. Moré, "Coercivity conditions in nonlinear complementarity problems", *SIAM Review, 16* (1974), 1–16.

[Mor96] J.J. Moré, "Global methods for nonlinear complementarity problems", *Mathematics of Operations Research, 21* (1996), 589–614.

[OrR70] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

OuZ95 J.V. Outrata and J. Zowe, "A Newton method for a class of quasi-variational inequalities", *Computational Optimization and Applications,* 4 (1995), 5–21.

[Pan90] J.-S. Pang, "Newton's method for B-differentiable equations", *Mathematics of Operations Research, 15* (1990), 311–341.

[Pan91] 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.

[PaC82] J.S. Pang and D. Chan, "Iterative methods for variational and complementarity problems", *Mathematical Programming, 24* (1982), 284–313.

[PaG93] J.S. Pang and S.A. Gabriel, "NE/SQP: A robust algorithm for the nonlinear complementarity problem", *Mathematical Programming, 60* (1993), 295–337.

[PaQ93] J.-S. Pang and L. Qi, "Nonsmooth equations: Motivation and algorithms", *SIAM Journal on Optimization, 3* (1993), 443–465.

[PeL97] J. Peng and Z. Lin "A non-interior continuation method for generalized linear complementarity problems", Preprint, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, August 1997.

[Pol97] E. Polak, *Optimization: Algorithms and Consistent Approximations*, Springer, New York, 1997.

[Qih97] H.-D. Qi, "A regularized smoothing Newton method for box constrained variational inequality problems with P_0 functions", Preprint, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, August 1997. Revised March 1998.

[QiL97] H.-D. Qi and L.Z. Liao, "A smoothing Newton method for extended vertical linear complementarity problems", Preprint, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, August 1997. Revised April 1998.

[QiL98] H.-D. Qi and L.Z. Liao, "A smoothing Newton method for general nonlinear complementarity problems", Preprint, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, March 1998.

[Qi93] L. Qi, "Convergence analysis of some algorithms for solving nonsmooth equations", *Mathematics of Operations Research, 18* (1993), 227–244.

[Qi95] L. Qi, "Trust region algorithms for solving nonsmooth equations", *SIAM Journal on Optimization, 5* (1995), 219–230.

[Qi96] L. Qi, "C-differentiability, C-Differential operators and generalized Newton methods", AMR 96/5, Applied Mathematics Report, the University of New South Wales, Sydney 2052, 1996.

[Qi97] 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, the University of New South Wales, Sydney 2052, Australia, June 1997. Revised July 1998.

[QiC95] 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.

[QiJ97] 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.

[QiS98] L. Qi and D. Sun, "Improving the convergence of non-interior point algorithms for nonlinear complementarity problems", to appear in *Mathematics of Computation.*

[QSZ97] L. Qi, D. Sun and G. Zhou, "A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities", AMR 97/13, Applied Mathematics Report, School of Mathematics, the University of New South Wales, Sydney 2052, Australia, June 1997. Revised August 1998.

[QiS93] L. Qi and J. Sun, "A nonsmooth version of Newton's method", *Mathematical Programming, 58* (1993), 353–367.

[Ral94] D. Ralph, "Global convergence of damped Newton's method for nonsmooth equations, via the path search", *Mathematics of Operations Research, 19* (1994), 352–389.

[RaG97] G. Ravindran and M.S. Gowda, "Regularization of P_0-functions in box variational inequality problems", Department of Mathematics and Statistics, University of Maryland Baltimore Country, Baltimore, August 1997. (Revised October 1997).

[Rob80] S.M. Robinson, "Strongly regular generalized equations", *Mathematics of Operations Research, 5* (1980), 43-62.

[Rob82] S.M. Robinson, "Generalized equations and their solutions, part II: Applications to nonlinear programming", *Mathematical Programming Study, 19* (1982), 200-221.

[Rob83] S.M. Robinson, "Generalized equations", in *Mathematical Programming: The State of the Art*, A. Bachem, M. Grötschel and B. Korte, eds., Springer-Verlag, Berlin, 1983, pp. 346–347.

[Rob91] S.M. Robinson, "An implicit-function theorem for a class of nonsmooth equations", *Mathematics of Operations Research, 16* (1991), 292–309.

[Rob92] S. M. Robinson, "Normal maps induced by linear transformation", *Mathematics of Operations Research, 17* (1992), 691–714.

[Rob94] S.M. Robinson, "Newton's method for a class of nonsmooth functions", *Set-Valued Analysis, 2* (1994), 291–305.

[Sha90] A. Shapiro, "On concepts of directional differentiability", *Journal of Optimization Theory and Applications, 66* (1990), 477–487.

[Sma86] S. Smale, "Algorithms for solving equations", in *Proceedings of the International Congress of Mathematicians*, Berkeley, California, pp. 172–195, 1986.

[Sun98] D. Sun, "A regularization Newton method for solving nonlinear complementarity problems", to appear in *Applied Mathematics and Optimization*.

[SuH97] 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.

[SuQ97] D. Sun and L. Qi, "On NCP-functions", Applied Mathematics Report AMR 97/32, School of Mathematics, the University of New South Wales, Sydney 2052, Australia, October 1997. (Revised April 1998).

[SuW98] D. Sun and R. S. Womersley, "A new unconstrained differentiable merit function for box constrained variational inequality problems and a damped Gauss-Newton method", to appear in *SIAM Journal on Optimization*.

[SzG98] R. Sznajder and M.S. Gowda, "On the limiting behavior of trajectory of regularized solutions of a P_0-complementarity problem", *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, M. Fukushima and L. Qi, eds., Kluwer Academic Publisher, Nowell, Maryland, 1998, pp. 371–380.

[Tse96] 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.

[Tse98] P. Tseng, “Analysis of a non-interior continuation method based on Chen-Mangasarian functions for complementarity problems”, *Reformulation - Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, M. Fukushima and L. Qi, eds., Kluwer Academic Publisher, Nowell, Maryland, 1998, pp. 381–404.

[Wri97] S. Wright, *Primal-Dual Interior Point Methods*, SIAM, Philadelphia, PA, 1997.

[XiH94a] B. Xiao and P.T. Harker, “A nonsmooth Newton method for variational inequalities, I: Theory”, *Mathematical Programming, 65* (1994), 151–194.

[XiH94b] B. Xiao and P.T. Harker, “A nonsmooth Newton method for variational inequalities, II: Numerical results”, *Mathematical Programming, 65* (1994), 195-216.

[YaF97] N. Yamashita and M. Fukushima, “Modified Newton methods for solving semismooth reformulations of monotone complementarity problems”, *Mathematical Programming, 76* (1997), 469–491.

[Xu96] S. Xu, “The global linear convergence of an infeasible non-interior path-following algorithm for complementarity problems with uniform P-functions”, Preprint, Department of Mathematics, University of Washington, Seattle, WA 98195, December 1996.

[Xu97] S. Xu, “The global linear convergence and complexity of a non-interior path-following algorithm for monotone LCP based on Chen-Harker-Kanzow-Smale smooth functions”, Preprint, Department of Mathematics, University of Washington, Seattle, WA 98195, February 1997.

[ZSQ98] G. Zhou, D. Sun and L. Qi, “Numerical experiments for a class of squared smoothing Newton methods for complementarity and variational inequality problems”, to appear in *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, M. Fukushima and L. Qi eds., Kluwer Academic Publisher, Nowell, Maryland, 1998, pp. 421–441.