SIAM J. NUMER. ANAL.
Vol. 38, No. 4, pp. 1200–1216

SMOOTHING METHODS AND SEMISMOOTH METHODS FOR NONDIFFERENTIABLE OPERATOR EQUATIONS*

XIAOJUN CHEN†, ZUHAIR NASHED‡, AND LIQUN QI§

Abstract. We consider superlinearly convergent analogues of Newton methods for nondifferentiable operator equations in function spaces. The superlinear convergence analysis of semismooth methods for nondifferentiable equations described by a locally Lipschitzian operator in R^n is based on Rademacher's theorem which does not hold in function spaces. We introduce a concept of slant differentiability and use it to study superlinear convergence of smoothing methods and semismooth methods in a unified framework. We show that a function is slantly differentiable at a point if and only if it is Lipschitz continuous at that point. An application to the Dirichlet problems for a simple class of nonsmooth elliptic partial differential equations is discussed.

Key words. smoothing methods, semismooth methods, superlinear convergence, nondifferentiable operator equation, nonsmooth elliptic partial differential equations

AMS subject classifications. 65J15, 65H10, 65J20

PII. S0036142999356719

1. Introduction.

This paper considers the nonlinear operator equation

$$F(x) = 0, \tag{1.1}$$

where $F : D \subset X \to Y$ is a continuous mapping, X and Y are Banach spaces, and D is an open domain in X. In a number of problems, the operator F is nondifferentiable. For example, a class of such problems arises in optimal control problems for parabolic partial differential equations with bound constraints on the control [16, 17] and leads to the operator equation

$$F(x) = x - P(K(x)) = 0,$$

where K is a completely continuous map from $L^\infty(\Omega)$ to $C(\Omega)$ for some bounded region $\Omega \subseteq \Re^m$, and P is the map on $C(\Omega)$ given by

$$P(K(x))(t) = \begin{cases} l(t), & K(x)(t) \le l(t), \\ K(x)(t), & l(t) \le K(x)(t) \le u(t), \\ u(t), & K(x)(t) \ge u(t) \end{cases}$$

for given l and u in $C(\Omega)$. A paradigm for such problems is the Urysohn integral equation of the second kind. Another class of nonsmooth equations related to magnetohydrodynamics (MHD) equilibria [18, 37] will be discussed in section 4. A paradigm

*Received by the editors May 24, 1999; accepted for publication April 26, 2000; published electronically October 6, 2000.
http://www.siam.org/journals/sinum/38-4/35671.html

†Department of Mathematics and Computer Science, Shimane University, Matsue 690-8504, Japan (chen@math.shimane-u.ac.jp). The work of this author was supported in part by the Japan Society of the Promotion of Science grant C11640119.

‡Department of Mathematical Sciences, University of Delaware, Newark, DE 19716 (nashed@math.udel.edu). The research of this author was partially supported by a grant from the National Science Foundation.

§Department of Applied Mathematics, Hong Kong Polytechnic University, Kowloon, Hong Kong (maqilq@polyu.edu.hk). The work of this author was partially supported by the Australian Research Council.

for this class is the Dirichlet problem for nonsmooth elliptic partial differential equations as discussed in section 4.

The nonsmoothness poses serious difficulties and challenges for devising for nonsmooth problems analogues of existing iterative methods, which use smoothness. For example, the Newton method assumes that F is Fréchet differentiable and is defined by

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \tag{1.2}$$

where F' is the Fréchet derivative of F. What are suitable analogues of Newton's method when F is not smooth?

Iterative methods for nondifferentiable equations have been studied for decades [7, 10, 15, 16, 31, 34, 38, 40]. Among these methods, smoothing methods and semismooth methods for nondifferentiable equations arising from variational inequalities and complementarity problems in R^n have been studied extensively in the last few years [2, 3, 4, 5, 9, 11, 14, 34].

Superlinear convergence analysis of semismooth Newton methods for equations defining a locally Lipschitian operator in R^n uses the notions of generalized Jacobian [12] and semismoothness [32, 35], which are based on the Rademacher theorem. The Rademacher theorem states that if $F : R^n \to R^m$ is a locally Lipschitzian function, then F is differentiable almost everywhere. For a locally Lipschitzian function $F : R^n \to R^m$, Clarke defined the generalized Jacobian [12] by

$$\partial F(x) = \mathrm{co}\{ \lim_{\substack{x_i \to x \\ x_i \in D_F}} F'(x_i)\},$$

where D_F is the set of points where F is differentiable. For nonsmooth equations described by a locally Lipschitzian operator F from R^n into R^n, the generalized Newton method is defined by

$$x_{k+1} = x_k - V_k^{-1}F(x_k), \qquad V_k \in \partial F(x_k). \tag{1.3}$$

Qi [32] and Qi and Sun [35] established superlinear convergence of (1.3) by using a concept of semismoothness. The concept of semismoothness was introduced by Mifflin for real-valued functions [20]. In [35], F is said to be *semismooth* at x if the limit

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

exists for any $h \in R^n$. Local behavior of the generalized Newton method is analyzed in [13, 31, 32, 35].

The Rademacher theorem does not hold in function spaces. Hence the above definitions of generalized Jacobian and semismoothness cannot be used in function spaces.

In this paper, we introduce notions of slanting functions and slant differentiability of operators in Banach spaces, and use these notions to formulate a concept of semismoothness in infinite dimensional spaces, which coincides with the above notion of semismoothness in the case of a locally Lipschitzian mapping on R^n. These notions will play a pivotal role in the formulation and convergence analysis of analogues of Newton's method (smoothing and semismooth methods) for nondifferentiable operator equations in function spaces.

The main feature of smoothing Newton methods in R^n is to approximate F by a parametric function $f(x,\epsilon) : R^n \times R_{++} \to R^n$ which is continuously differentiable with respect to x and then to use the partial derivative $f_x(x_k,\epsilon_k)$ at each step of the Newton-like iteration. The error of $f(\cdot,\epsilon_k)$ to F is bounded by $\|F(x) - f(x,\epsilon_k)\| \le O(\epsilon_k)$ and $\epsilon_k \to 0$ as $k \to \infty$. For complementarity problems, many smoothing functions have the following properties [9]:

$$f^o(x) \equiv \lim_{\epsilon\downarrow 0} f_x(x,\epsilon) \quad \text{exists for every } x \in R^n \tag{1.4}$$

and

$$\lim_{h\to 0} \frac{\|F(x+h) - F(x) - f^o(x+h)h\|}{\|h\|} = 0. \tag{1.5}$$

The properties (1.4) and (1.5) suggest a superlinearly convergent Newton method [6, 11]:

$$x_{k+1} = x_k - f^o(x_k)^{-1}F(x_k). \tag{1.6}$$

Note that $f^o : R^n \to R^{n\times n}$ is a single valued function, and the superlinear convergence of (1.6) is not based on the Rademacher theorem. This opens a way to study Newton methods for nonsmooth problems in function spaces.

The organization of the paper is as follows. In section 2 we introduce the notion of a slanting function f^o and slant differentiability for a general nonsmooth function F in Banach spaces and study some of their interesting properties. Using slanting functions, we extend in section 3 the semismooth Newton method and the smoothing Newton method to Banach spaces. An application to a class of nonsmooth Dirichlet problems is studied in section 4.

We use $\alpha, \beta, \gamma, \ldots,$ to denote scalars. The set of all positive real numbers is denoted by R_{++}. Let $L(X,Y)$ denote the set of all bounded linear operators on X into Y.

2. Slant differentiability.

A function $F : D \subset X \to Y$ is said to be *(one-sided) directionally differentiable* at x if the limit

$$\delta^+ F(x;h) := \lim_{t\to 0^+} \frac{F(x+th) - F(x)}{t}$$

exists, in which case $\delta^+ F(x;h)$ is called the (one-sided) directional derivative of F at x with respect to the direction h. For brevity we will drop "one-sided" in what follows since this is the only notion of directional derivative that occurs in this paper.

A function $F : D \subset X \to Y$ is said to be *B-differentiable* at a point x if it is directionally differentiable at x, and

$$\lim_{h\to 0} \frac{F(x+h) - F(x) - \delta^+ F(x;h)}{\|h\|} = 0. \tag{2.1}$$

In this case, we call $\delta^+ F(x;\cdot)$ the B-derivative of F at x. See [39] for the B-differentiability (B for Bouligand).

In finite dimensional Euclidean spaces, Shapiro [41] showed that a locally Lipschitzian function F is B-differentiable at x if and only if it is directionally differentiable at x. Moreover, Qi and Sun [35] showed that F is semismooth at x if and only if F is B-differentiable (hence directionally differentiable) at x and for each $V \in \partial F(x+h)$

$$\delta^+ F(x;h) - Vh = o(\|h\|).$$

However, these results do not hold in function spaces since the generalized Jacobian is defined only in finite dimensional spaces. To circumvent this difficulty in infinite dimensional spaces we introduce the following notion of slant differentiability.

DEFINITION 2.1. *A function $F : D \subset X \to Y$ is said to be* slantly differentiable *at $x \in D$ if there exists a mapping $f^o : D \to L(X,Y)$ such that the family $\{f^o(x+h)\}$ of bounded linear operators is uniformly bounded in the operator norm for h sufficiently small and*

$$\lim_{h\to 0} \frac{F(x+h) - F(x) - f^o(x+h)h}{\|h\|} = 0.$$

The function f^o is called a slanting function *for F* **at** *x.*

DEFINITION 2.2. *A function $F : D \subset X \to Y$ is said to be slantly differentiable in an open domain $D_0 \subset D$ if there exists a mapping $f^o : D \to L(X,Y)$ such that f^o is a slanting function for F at every point $x \in D_0$. In this case, f^o is called a* slanting function *for F* **in** *D_0.*

DEFINITION 2.3. *Suppose that $f^o : D \to L(X,Y)$ is a slanting function for F at $x \in D$. We call the set*

$$\partial_S F(x) := \{\lim_{x_k \to x} f^o(x_k)\}$$

the slant derivative *of F associated with f^o at $x \in D$. Here, $\lim_{x_k\to x} f^o(x_k)$ means the limit of $f^o(x_k)$ for any sequence $\{x_k\} \subset D$ such that $x_k \to x$ and $\lim_{x_k\to x} f^o(x_k)$ exists, and $\partial_S F(x)$ is the set of all such limits. (Note that $f^o(x) \in \partial_S F(x)$, so $\partial_S F(x)$ is always nonempty.)*

Slant differentiability captures a property that appears implicitly in some convergence proofs of Newton-type methods for solving nonsmooth equations as well as ill-posed smooth equations. For example, consider the parameterized Newton method for solving ill-posed smooth equations. To overcome ill-posedness and singularity, we use

$$x_{k+1} = x_k - (F'(x_k) + \lambda_k I)^{-1} F(x_k),$$

where λ_k is chosen such that $F'(x_k) + \lambda_k I$ is nonsingular. Let $f^o(x) = F'(x) + \lambda_x I$ and assume $\lambda_x \to 0$ as $x \to x^*$. Then f^o is a slanting function for F at x^* if $F'(x)$ is uniformly bounded in a neighborhood of x^*.

We now make a few comments on some unusual properties of slanting functions which also explain the choice of the terms "slanting function" and "slant derivative."

Remarks.

(1) Unlike other notions of derivatives, the term "$f^o(x)h$" does not appear in Definition 2.1, so for a slanting function f^o for F at x, $f^o(x)$ itself is not characterized *in general* by a limit of a quotient or a sequence.

(2) A function F may be slantly differentiable at every point of D, but there is no common slanting function of F at all points of D. For example, if F is Fréchet differentiable at x, we take $f^o(u) := F'(x)$ for all $u \in D$, then f^o is a slanting function for F at x. But f^o in general is not a slanting function of F at other points of D. If

F is continuously differentiable in D and we take $f^o(u) := F'(u)$ for all $u \in D$, then f^o is a slanting function for F at every point of D.

(3) A slanting function f^o for F at x is a single valued function. A slantly differentiable function F at x can have infinitely many slanting functions at x. Even if F is continuously differentiable in D, F still can have infinitely many slanting functions for all points of D. For example, we may let f^o take the same values of F' except at a finite number of points of D, and take arbitrary values at these finite number of points. Then such f^o is still a slanting function of F for all points of D. One may conjecture that if F is continuously differentiable in D and f^o is a slanting function for F in D, then f^o coincides with F' except possibly on a set of measure zero.

(4) If f^o and p^o are both slanting functions for F at x (in D), then

$$h^o := \lambda f^o + (1-\lambda)p^o$$

is also a slanting function for F at x (in D), where $\lambda \in [0,1]$. Moreover,

$$\lim_{h\to 0} \|f^o(x+h)h - p^o(x+h)h\| = 0.$$

On the other hand, if f^o and g^o are slanting functions for F and G at x (in D), respectively, then $h^o := \alpha f^o + \beta g^o$ is a slanting function for $\alpha F + \beta G$ at x (in D) where α and β are constants. Note that such a result for linear combination does not hold for the generalized Jacobian [12].

(5) f^o is not continuous in general. For example, let $X = Y = R$ and $F(x) = \max(0, x)$. Let δ be a real number. Then the function

$$f^o(x) = \begin{cases} 1, & x > 0, \\ \delta, & x = 0, \\ 0, & x < 0 \end{cases}$$

is a slanting function for F in X. The slant derivative of F for $x \in X$ is

$$\partial_S F(x) = \begin{cases} 1, & x > 0, \\ \{0, \delta, 1\}, & x = 0, \\ 0, & x < 0. \end{cases}$$

In fact it is easy to see that if f^o is continuous at x, then F is differentiable at x and $F'(x) = f^o(x)$. The slant derivative of F associated with f^o at x reduces to a singleton $\partial_S F(x) = \{f^o(x)\}$.

(6) For a locally Lipschitzian function $F : R^n \to R^m$, if F is semismooth at x, then any single valued selection of the Clarke–Jacobian or the B-subdifferential is a slanting function of F at x. This may not be true if F is not semismooth at x. For example, let $X = Y = R$,

$$F(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

The derivative $F'(x)$ is discontinuous at 0. The function F is slantly differentiable at 0. Indeed let f^o be any function for which $\lim_{h\to 0} f^o(h) = 0$. Then

$$\lim_{h\to 0} \frac{h^2 \sin \frac{1}{h} - f^o(h)h}{h} = 0.$$

Hence F is slantly differentiable at 0 with infinitely many slanting functions for F at 0. Note that such f^o is not a slanting function for F at every point $x \in X \setminus \{0\}$. If we let $f^o \equiv F'$, then f^0 is a slanting function for F at every point $x \in X \setminus \{0\}$, but not a slanting function for F at 0.

(7) For a slantly differentiable function F at x, the set $\partial_S F(x)$ is dependent on the choice of a slanting function for F at x. Associated with any slanting function, the set $\partial_S F(x)$ is bounded, since $f^o(x+h)$ is uniformly bounded for h sufficiently small. For example, let $X = Y = R$,

$$F(x) = \begin{cases} x \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Let

$$f^o(x) = \begin{cases} \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Then f^o is a slanting function for F at 0 and $\partial_S F(0) = [-1, 1]$. We note that F is neither directionally differentiable at 0 nor Lipschitzian in any neighborhood of 0. Note that the function f^o in this example is not slantly differentiable at 0.

(8) A continuous function is not necessarily slantly differentiable. For example, let $X = Y = R$,

$$F(x) = \begin{cases} \sqrt{x} & \text{if } x \geq 0, \\ -\sqrt{-x} & \text{if } x < 0. \end{cases}$$

Since $F(h) - F(0) = h/\sqrt{|h|}$, and $1/\sqrt{|h|} \to \infty$ as $h \to 0$, there is no uniformly bounded function f^o such that $F(h) - F(0) - f^o(h)h = o(h)$.

Definition 2.4. *A function $F : X \to Y$ is said to be* Lipschitz continuous *at x if there is a positive constant L such that for all sufficiently small h,*

$$\|F(x+h) - F(x)\| \leq L\|h\|.$$

We now present a necessary and sufficient condition for slant differentiability. For the proof of sufficiency we need the following lemma, which is a corollary of the Hahn–Banach theorem.

Lemma 2.5. *Let X be a normed space and h be a fixed element of X, $h \neq 0$. Then there exists an element $g \in X^*$, where X^* is the (norm) dual of X, such that*

$$g(h) = \|h\| \quad \text{and} \quad \|g\| = 1.$$

(Note by definition of X^*, g is a continuous linear functional on X, so it is bounded.)

Theorem 2.6. *An operator $F : X \to Y$ is slantly differentiable at x if and only if F is Lipschitz continuous at x.*

Proof. Suppose that F is slantly differentiable at x. By the definition of slant differentiability, there are $C > 0$ and $\delta > 0$ such that for all $\|h\| \leq \delta$, $\|f^o(x+h)\| \leq C$ and

$$\frac{\|F(x+h) - F(x) - f^o(x+h)h\|}{\|h\|} \leq C.$$

Hence, for all $\|h\| \leq \delta$,

$$\|F(x+h) - F(x)\| \leq C\|h\| + C\|h\| \leq 2C\|h\|.$$

Conversely, suppose that F is Lipschitz continuous at x. We shall show that F is slantly differentiable at x by constructing a slanting function for F at x. For each fixed $h \neq 0$, by Lemma 2.5, there exists a continuous linear functional $g_h \in X^*$ such that $g_h(h) = \|h\|$ and $\|g_h\| = 1$. With x fixed as above, define the following function on an open domain containing x:

$$f^o(x+h) := \frac{1}{\|h\|}(F(x+h) - F(x))g_h \tag{2.2}$$

for $h \neq 0$, and define $f^o(x)$ to be any bounded linear operator on X into Y. Then f^o maps D into $L(X, Y)$ since each g_h is in X^*. For any $z \in X$,

$$f^o(x+h)z = \frac{F(x+h) - F(x)}{\|h\|} g_h(z).$$

Thus

$$\begin{aligned} \|f^o(x+h)z\| &= \frac{\|F(x+h) - F(x)\|}{\|h\|} |g_h(z)| \\ &\leq \frac{\|F(x+h) - F(x)\|}{\|h\|} \|z\|. \end{aligned}$$

Therefore

$$\sup_{z \neq 0} \frac{\|f^o(x+h)z\|}{\|z\|} \leq \frac{\|F(x+h) - F(x)\|}{\|h\|},$$

that is,

$$\|f^o(x+h)\| \leq \frac{\|F(x+h) - F(x)\|}{\|h\|}.$$

Thus for sufficiently small h,

$$\|f^o(x+h)\| \leq L.$$

Now using (2.2), and $g_h(h) = \|h\|$, we obtain

$$\begin{aligned} f^o(x+h)h &= \frac{F(x+h) - F(x)}{\|h\|} g_h(h) \\ &= F(x+h) - F(x). \qquad \square \end{aligned}$$

If X is a Hilbert space, then by the Riesz representation theorem every continuous linear functional on X can be represented by an inner product. Thus the formula (2.2) can be written in the form

$$f^o(x+h) = \frac{1}{\|h\|}(F(x+h) - F(x))\langle \cdot, h/\|h\| \rangle.$$

COROLLARY 2.7 (mean value theorem for slantly differentiable functions). *Suppose that $F : D \subset X \to Y$ is slantly differentiable at x. Then for any $h \neq 0$ such that $x + h$ is in D, there exists a slanting function for F at x such that*

$$F(x+h) - F(x) = f^o(x+h)h.$$

Proof. This follows from the first part of Theorem 2.6 and the proof of the second part of the same theorem. ∎

Note that the above form of the mean value theorem is in equality form. It is a selection theorem from the set of slanting functions for F at x. Mean value theorems for smooth operators whose range is an infinite dimensional space are usually given in the form of inequalities involving norms or majorants or an inclusion form involving the closed convex hull of the set of values of the derivative. For a comprehensive overview of various types of mean value theorems for smooth operators, see pp. 171–186 of [23].

PROPOSITION 2.8. *Suppose that F is slantly differentiable at x, and let f^o be a slanting function for F at x.*

(a) *F is directionally differentiable at x if and only if*

$$\lim_{t\to 0^+} f^o(x+th)h$$

exists. If F is directionally differentiable at x, then

$$\delta^+ F(x;h) = \lim_{t\to 0^+} f^o(x+th)h.$$

(b) *F has a B-derivative at x if and only if*

$$\lim_{t\to 0^+} f^o(x+th)h$$

exists uniformly *with respect to h on each bounded set (say, on $\|h\| = 1$).*

Proof. (a) Let $h \in X$ with $\|h\| = 1$, and $t > 0$. Then

$$\lim_{t\to 0^+} \frac{\|F(x+th) - F(x) - f^o(x+th)(th)\|}{t} = 0$$

is equivalent to

$$\lim_{t\to 0^+} \|\frac{F(x+th) - F(x)}{t} - f^o(x+th)(h)\| = 0.$$

Hence if F is directionally differentiable, then

$$\delta^+ F(x;h) = \lim_{t\to 0^+} \frac{F(x+th) - F(x)}{t} = \lim_{t\to 0^+} f^o(x+th)h.$$

The converse is also true.

(b) This follows from part (a) and the known (and easy to prove) fact that F has a B-derivative at x if and only if

$$\lim_{t\to 0^+} \frac{F(x+th) - F(x)}{t}$$

exists uniformly with respect to h on each bounded set (see, for example, Nashed [23], where a hierarchy of notions of differentiability is characterized by convergence of "remainder" quotients $R(th)/t$ as t approaches zero, uniformly with respect to h in various classes of subsets). ∎

THEOREM 2.9. *Suppose F is slantly differentiable at x and let f^o be a slanting function for F at x. Then the following statements are equivalent.*

(a) *For some function $g : X \to Y$ which is $o(\|h\|)$, $f^o(x+h)h + g(h)$ is positively homogeneous of degree 1 in h.*

(b) $\lim_{t\to 0^+} f^o(x+th)h$ *exists for every $h \in X$ and*

$$\lim_{\|h\|\to 0} \frac{\lim_{t\to 0^+} f^o(x+th) - f^o(x+h)h}{\|h\|} = 0.$$

(c) *F is B-differentiable at x, and*

$$\delta^+ F(x;h) - f^o(x+h)h = o(\|h\|).$$

Proof. (a)$\Rightarrow$ (b): If $f^o(x+h)h + g(h)$ is positively homogeneous of degree 1 in h, then for any fixed $t > 0$,

$$f^o(x+th)(th) + g(th) = t(f^o(x+h)h + g(h)),$$

so

$$f^o(x+th)h - f^o(x+h)h = g(h) - g(th)/t. \tag{2.3}$$

Note that $g(h) = o(\|h\|)$ if and only if $g(th) = o(t)$ for each fixed h, uniformly in h on each bounded set. Hence for any $h \in X$

$$\lim_{t\to 0^+} f^o(x+th)h = f^o(x+h)h + g(h)$$

uniformly with respect to h on each bounded set. Moreover,

$$\lim_{\|h\|\to 0} \frac{\lim_{t\to 0^+} f^o(x+th) - f^o(x+h)h}{\|h\|} = \lim_{\|h\|\to 0} \frac{g(h)}{\|h\|} = 0.$$

(b) $\Rightarrow$ (c): By part (b) of Proposition 2.8, statement (b) implies that F is B-differentiable and $\lim_{t\to 0^+} f(x+th)h = \delta^+ F(x;h)$. Hence, statement (c) holds.

(c) $\Rightarrow$ (a): Since $\delta^+ F(x;h) = f^o(x+h)h + o(\|h\|)$, and $\delta^+ F(x;h)$ is positively homogeneous of degree 1 in h, we have (a). ∎

PROPOSITION 2.10. *Suppose that F is slantly differentiable in a neighborhood of x and let f^o be a slanting function for F in the neighborhood of x. Then the following two statements are equivalent.*

(a) *There are a neighborhood $\mathcal{N}_x$ of x and a positive constant C such that for any $u \in \mathcal{N}_x$, $f^o(u)$ is nonsingular and $\|f^o(u)^{-1}\| \le C$.*

(b) *There are a neighborhood $\hat{\mathcal{N}}_x$ and a positive constant $\hat{C}$ such that for any $u \in \hat{\mathcal{N}}_x$, every $V \in \partial_S F(u)$ is nonsingular and $\|V^{-1}\| \le \hat{C}$.*

Proof. Part (a) $\Longrightarrow$ part (b): It is straightforward from the definition of $\partial_S F(u)$.

Part (b) $\Longrightarrow$ part (a): It is due to the fact that $f^o(u) \in \partial_S F(u)$. ∎

PROPOSITION 2.11. *Suppose that F is slantly differentiable at x and let f^o be a slanting function for F at x. If there are a neighborhood $\mathcal{N}_x$ of x and a positive constant C such that for any $u \in \mathcal{N}_x$, $f^o(u)$ is nonsingular and $\|f^o(u)^{-1}\| \le C$, then there is a positive constant $\hat{C}$ such that every $V \in \partial_S F(x)$ is nonsingular and $\|V^{-1}\| \le \hat{C}$. Moreover, if Y is a finite dimensional space, the converse holds.*

Proof. The first part follows the definition of $\partial_S F(x)$. The second part is due to the fact that every bounded sequence has a convergent subsequence in finite dimensional spaces. Indeed, if the second part is not true, then there is a sequence $\{u_k\}$ such that $u_k \to x$ and either all $f^o(u_k)$ are singular or $\|f^o(u_k)^{-1}\| \to \infty$. By the definition of $\partial_S F(x)$, there is a subsequence $\{u_{k_j}\} \subset \{u_k\}$ such that $f^o(u_{k_j}) \to V$, and $V \in \partial_S F(x)$ is singular. This contradicts the assumption. ∎

3. Smoothing functions and semismooth functions. We generalize the definition of smoothing functions for a nonsmooth function and the concept of semismmothness of a nonsmooth function in finite dimensional spaces to infinite dimensional spaces.

DEFINITION 3.1. *We say that $f : D \times R_{++} \to Y$ is a* smoothing function *of F if f is continuously differentiable with respect to x and for any $x \in D$ and any $\epsilon > 0$,*

$$\|F(x) - f(x, \epsilon)\| \leq \mu\epsilon, \tag{3.1}$$

where μ is a positive constant.

The smoothing function f is said to satisfy the slant derivative consistency property *at $\hat{x}$ (in D) if*

$$\lim_{\epsilon \to 0^+} f_x(x, \epsilon) = f^o(x) \in L(X, Y) \tag{3.2}$$

exists for x in a neighborhood of $\hat{x}$ (in D) and f^o serves as a slanting function for F at $\hat{x}$ (in D).

Note that the limit in (3.2) is in the topology of the operator norm, so the pointwise convergence of $f_x(x, \epsilon)h$ to $f^o(x)h$ for each fixed h is uniform on the set $\|h\| = 1$.

DEFINITION 3.2. *We say that F is* semismooth *at x if there is a slanting function f^o for F in a neighborhood $\mathcal{N}_x$ of x, such that f^o and the associated slant derivative satisfy the following two conditions.*

(a) $\lim_{t \to 0^+} f^o(x + th)h$ *exists for every $h \in X$ and*

$$\lim_{\|h\| \to 0} \frac{\lim_{t \to 0^+} f^o(x + th)h - f^o(x + h)h}{\|h\|} = 0.$$

(b)

$$f^o(x + h)h - Vh = o(\|h\|) \qquad \textit{for all} \quad V \in \partial_S F(x + h).$$

THEOREM 3.3. *Suppose that F is slantly differentiable in a neighborhood $\mathcal{N}_x$ of x, and let f^o be a slanting function for F in $\mathcal{N}_x$. Then F is semismooth at x if and only if F is B-differentiable at x and*

$$\delta^+ F(x; h) - Vh = o(\|h\|) \qquad \textit{for all} \quad V \in \partial_S F(x + h), \tag{3.3}$$

where $\partial_S F$ is the slant derivative associated with f^o in $\mathcal{N}_x$.

Proof. Suppose that F is semismooth at x. Then from Theorem 2.9, part (a) of Definition 3.2 implies that F is B-differentiable and

$$\delta^+ F(x; h) - f^o(x + h)h = o(\|h\|).$$

Thus part (b) of Definition 3.2 implies (3.3).

Now we suppose that F is B-differentiable at x and (3.3) holds. Then for all $V \in \partial_S F(x + h)$,

$$\begin{aligned} & f^o(x + h)h - Vh \\ &= F(x + h) - F(x) - \delta^+ F(x; h) + (\delta^+ F(x; h) - Vh) \\ &\qquad -(F(x + h) - F(x) - f^o(x + h)h) \\ &= o(\|h\|). \end{aligned}$$

Hence part (b) of Definition 3.2 holds. Moreover, we have

$$\delta^+ F(x;h) - f^o(x+h)h = o(\|h\|).$$

By Theorem 2.9, part (a) of Definition 3.2 holds and so F is semismooth at x. ∎

Theorem 3.3 implies that the definition of semismoothness used here coincides with the definition of Qi and Sun [35] in finite dimensional spaces if we take a single valued selection of the Clarke–Jacobian or the B-subdifferential as the slanting function.

To illustrate Theorem 3.3, we consider the system of "min" nonsmooth equations in R^n

$$F(x) := \min(p(x), q(x)) = 0,$$

where p and q are continuously differentiable functions from R^n into itself. This system is equivalent to the complementarity problem

$$p(x) \geq 0, \qquad q(x) \geq 0, \qquad p(x)^T q(x) = 0.$$

Chen, Qi, and Sun [9] showed that every smoothing function $f(x,\epsilon)$ in the Chen–Mangasarian smoothing function family [5] for the nonsmooth function F satisfies (3.2). In particular, for $i = 1, 2, \ldots, n$,

$$f_i^o(x) = \begin{cases} p_i'(x), & p_i(x) < q_i(x), \\ q_i'(x), & q_i(x) < p_i(x), \\ \alpha p_i'(x) + (1-\alpha) q_i'(x) & \text{otherwise}, \end{cases}$$

where $\alpha \in [0,1]$ is dependent on the choice of a smoothing function. Such $f^o(x)$ belongs to the set

$$\partial_C F(x) = \partial F_1(x) \times \partial F_2(x) \times \cdots \times \partial F_n(x)$$

at every point $x \in R^n$. (See [33] for $\partial_C F(x)$.) Hence, every smoothing function in the Chen–Mangasarian smoothing function family satisfies the slant derivative consistency property in R^n. Moreover, the associated slant derivative is, for $i = 1, 2, \ldots, n$,

$$\partial_S F_i(x) = \begin{cases} p_i'(x), & p_i(x) < q_i(x), \\ q_i'(x), & q_i(x) < p_i(x), \\ \{p_i'(x),\, q_i'(x),\, \alpha p_i'(x) + (1-\alpha) q_i'(x)\} & \text{otherwise}, \end{cases}$$

which is bounded, nonempty, and satisfies

$$\partial_S F(x) \subseteq \partial_C F(x).$$

Furthermore, the following fact is known [35]:

$$\delta^+ F(x;h) - Vh = o(\|h\|) \qquad \text{for every} \qquad V \in \partial_C F(x+h).$$

On the other hand, by Proposition 2.8, we know

$$\delta^+ F(x;h) = \lim_{t \to 0^+} f^o(x+th)h.$$

Hence, the nonsmooth function F is semismooth in the sense of Definition 3.2.

Now we consider superlinearly convergent Newton-type methods for nonsmooth equations with slanting differentiable operators.

THEOREM 3.4. *Suppose that F is slantly differentiable at a solution x^* of* (1.1). *Let f^o be a slanting function for F at x^* and $\|f^o(x)^{-1}\| \le M$ in a neighborhood $\mathcal{N}$ of x^*, where M is a positive constant. Then the iterative sequence $\{x_k\}$ generated by the Newton-type method*

$$x_{k+1} = x_k - A(x_k)^{-1}F(x_k) \tag{3.4}$$

superlinearly converges to x^* in a neighborhood $\mathcal{N}_0$ of x^*. Here $A(x) \in L(X,Y)$ and

$$\|A(x^*+h) - f^o(x^*+h)\| \to 0 \qquad \text{as} \quad \|h\| \to 0. \tag{3.5}$$

Proof. By Definition 2.1 and the Banach lemma [29], there is a neighborhood $\mathcal{N}_0$ of x^*, $\mathcal{N}_0 \subset \mathcal{N}$, and positive constants $M_0 \ge M$, $\rho_1, \rho_2 \in (0,1), \rho_1 + \rho_2 < 1$ such that for any $x \in \mathcal{N}_0$, $A(x)$ is nonsingular and $\|A(x)^{-1}\| \le M_0$,

$$\|F(x) - F(x^*) - f^o(x)(x-x^*)\| \le \frac{\rho_1}{M_0}\|x-x^*\|$$

and

$$\|A(x) - f^o(x)\| \le \frac{\rho_2}{M_0}.$$

Therefore starting from any $x_0 \in \mathcal{N}_0$ the Newton method (3.4) is well defined and the successive iterates satisfy the following inequalities:

$$\begin{aligned}
&\|x_{k+1} - x^*\| \\
&= \|x_k - x^* - A(x_k)^{-1}F(x_k) + A(x_k)^{-1}F(x^*)\| \\
&\le \|A(x_k)^{-1}\|(\|F(x_k) - F(x^*) - f^o(x_k)(x_k - x^*)\| + \|(A(x_k) - f^o(x_k))(x_k - x^*)\|) \\
&\le M_0\frac{\rho_1+\rho_2}{M_0}\|x_k - x^*\| \\
&\le (\rho_1+\rho_2)\|x_k - x^*\|.
\end{aligned}$$

Hence the sequence $\{x_k\}$ converges to x^*. Moreover, using Definition 2.1 and (3.5), the inequalities above imply

$$\|x_{k+1} - x^*\| = o(\|x_k - x^*\|). \qquad \square$$

Using Theorem 3.4 and Proposition 2.10, we can immediately obtain the following theorem.

THEOREM 3.5. *Suppose that F is slantly differentiable at a solution x^* of* (1.1). *Let f^o be a slanting function for F at x^* and $\|f^o(x)^{-1}\| \le M$ in a neighborhood $\mathcal{N}$ of x^*, where M is a positive constant. Then the following statements hold.*

(a) *The Newton-type method* (1.6) *superlinearly converges to x^* in a neighborhood $\mathcal{N}_0$ of x^*.*

(b) *If $f : D \times R_{++} \to Y$ is a smoothing function of F which satisfies the slant derivative consistency property* (3.2) *in $\mathcal{N}$, then the smoothing Newton method*

$$x_{k+1} = x_k - f'(x_k, \epsilon_k)^{-1}F(x_k) \tag{3.6}$$

superlinearly converges to x^* in a neighborhood $\mathcal{N}_0$ of x^*.

(c) *If F is semismooth at x^*, then the semismooth Newton method*

$$x_{k+1} = x_k - V_k^{-1}F(x_k), \qquad V_k \in \partial_S F(x_k) \tag{3.7}$$

superlinearly converges to x^* in a neighborhood $\mathcal{N}_0$ of x^*.

4. An application to a class of nonsmooth elliptic partial differential equations. Let $\Omega \subset R^2$ be a bounded region with piecewise smooth boundary $\partial\Omega$ and let W be the class of functions from Ω to R satisfying

$$\int_\Omega |u(x)|dx < \infty.$$

Let $X := L^1(\Omega, R)$ be the space of functions in W endowed with the norm

$$\|u\| = \int_\Omega |u(x)|dx.$$

We consider the following nonsmooth Dirichlet problem:

$$\begin{cases} -\triangle u &= P(u) \quad \text{in } \Omega, \\ \quad u &= 0 \qquad \text{on } \partial\Omega, \end{cases} \tag{4.1}$$

where $P : X \to R$ is a continuous function.

Let

$$H(u) = \int_\Omega G(x, y)P(u(y))dy,$$

where G is the Green function for the boundary value problem (e.g., see [42])

$$\begin{cases} -(\triangle u)(x) &= \delta(x - y),\ x, y \quad \text{in } \Omega, \\ \quad u &= 0 \qquad\qquad\qquad \text{on } \partial\Omega, \end{cases} \tag{4.2}$$

where $\delta(\cdot - y)$ is the Dirac "generalized function" at y in Ω.

The nonsmooth integral equation

$$F(u) := u - H(u) = 0$$

is equivalent to the nonsmooth Dirichlet problem (4.1).

THEOREM 4.1. *Suppose that $p : X \times R_{++} \to X$ is a smoothing function of P satisfying*

$$\|P(u) - p(u, \epsilon)\| \le \kappa\epsilon,$$

where κ is a positive constant. Then

$$f(u, \epsilon) := u - \int_\Omega G(x, y)p(u(y), \epsilon)dy$$

is a smoothing function of F and

$$\|F(u) - f(u, \epsilon)\| \le \kappa\mu\epsilon,$$

where

$$\mu = \sup_{x\in\Omega} \int_\Omega \|G(x, y)\|dy.$$

Proof. It is easy to see that f is continuously differentiable with respect to u and

$$f'(u, \epsilon) = I - \int_\Omega G(x, y)p'(u(y), \epsilon)dy.$$

Moreover,

$$\begin{aligned}\|F(u) - f(u,\epsilon)\| &= \| \int_\Omega G(x,y)P(u(y)) - G(x,y)p(u(y),\epsilon)dy\| \\ &\le \int_\Omega \|G(x,y)\|\|P(u(y)) - p(u(y),\epsilon)\|dy \\ &\le \kappa\epsilon \int_\Omega \|G(x,y)\|dy \\ &\le \kappa\mu\epsilon. \qquad \square\end{aligned}$$

THEOREM 4.2. *Suppose that P is slantly differentiable at u, and let p^o be a slanting function for P at u. Then F is slantly differentiable at u, and*

$$f^o(u) = I - \int_\Omega G(x,y)p^o(u(y))dy$$

is a slanting function for F at u

Proof. Using the definition of f^o as given in the statement of the theorem, we have

$$\begin{aligned}&F(u+h) - F(u) - f^o(u+h)h \\ &= -\left(\int_\Omega G(x,y)P(u+h)dy - \int_\Omega G(x,y)P(u)dy - \int_\Omega G(x,y)p^o(u+h)dyh\right) \\ &= -\left(\int_\Omega G(x,y)(P(u+h) - P(u) - p^o(u+h)h)dy\right).\end{aligned}$$

Since p^o is a slanting function for P at u, the above equality implies that f^o is a slanting function for F at u. $\square$

By properties of integrals, we have the following proposition.

PROPOSITION 4.3. (1) *If $p : X \times R_{++} \to X$ is a smoothing function of P which satisfies the slant derivative consistency property* (3.2) *at u, then $f : X \times R_{++} \to X$ is a smoothing function of F which satisfies the slant derivative consistency property* (3.2) *at u.*

(2) *If P is semismooth at u, then F is semismooth at u.*

The above results demonstrate that superlinearly convergent smoothing methods or semismooth methods can be developed for the nonsmooth Dirichlet problem (4.1).

For example, we consider

$$P(u) = \lambda \max(0, u - \alpha),$$

that is,

$$P(u)(x) = \begin{cases} 0, & u(x) \le \alpha, \\ \lambda(u(x) - \alpha) & \text{otherwise,} \end{cases}$$

where λ and α are constants. Let

$$p(u,\epsilon) = \frac{\lambda}{2}(u - \alpha + \sqrt{(u-\alpha)^2 + 4\epsilon^2}).$$

Then p is a smoothing function of P in X, and satisfies the slant derivative consistency property in X since

$$\lim_{\epsilon\to 0} p'(u,\epsilon) = \lim_{\epsilon\to 0} \frac{\lambda}{2}\left(1 + \frac{u-\alpha}{\sqrt{(u-\alpha)^2 + 4\epsilon^2}}\right) = p^o(u) = \begin{cases} 0, & u(x) < \alpha, \\ \frac{\lambda}{2}, & u(x) = \alpha, \\ \lambda & \text{otherwise.} \end{cases}$$

We next show that p^o is a slant function for P in X. In fact, for any $u, h(\neq 0) \in X$, we have

$$\begin{aligned}
&\|P(u+h) - P(u) - p^o(u+h)h\| \\
&= \int_\Omega |P(u+h)(x) - P(u)(x) - p^o(u+h)h(x)|dx \\
&\leq \lambda \left\{ \int_{(u(x)+h(x)-\alpha)(u(x)-\alpha)<0} |u(x)-\alpha|dx + \int_{u(x)+h(x)=\alpha} |u(x)-\alpha| + \frac{1}{2}|h(x)|dx \right\}.
\end{aligned}$$

Note that

$$|u(x)-\alpha| < |h(x)| \qquad \text{if} \quad (u(x)+h(x)-\alpha)(u(x)-\alpha) < 0.$$

Letting $\|h\| \to 0$, we have

$$\begin{aligned}
&\frac{\lambda}{\|h\|} \left\{ \int_{(u(x)+h(x)-\alpha)(u(x)-\alpha)<0} |u(x)-\alpha|dx + \int_{u(x)+h(x)=\alpha} |u(x)-\alpha| + \frac{1}{2}|h(x)|dx \right\} \\
&\leq \frac{3\lambda}{2} \int_{0<|u(x)-\alpha|\leq|h(x)|} dx \\
&\to 0.
\end{aligned}$$

Moreover, we can show that P is semismooth in X. Hence by using

$$f(u, \epsilon) := u - \frac{1}{2} \int_\Omega G(x, y)(u + \sqrt{u^2 + 4\epsilon^2})dy$$

we can obtain superlinearly convergent smoothing methods and semismooth methods for the nonsmooth Dirichlet problem

$$\left\{ \begin{array}{lll} -\triangle u & = \lambda \max(0, u-\alpha) & \text{in } \Omega, \\ u & = \phi(x, y) & \text{on } \partial\Omega. \end{array} \right. \tag{4.3}$$

This nonsmooth Dirichlet problem is related to MHD equilibria [18].

We report numerical results for the following example in Table 1.

Example 4.1.

$$\left\{ \begin{array}{lll} -\triangle u & = \frac{\pi^2}{2} \max(0, u-1) & \text{in } \Omega = (0,1) \times (0,1), \\ u & = \phi(x, y) & \text{on } \partial\Omega, \end{array} \right.$$

where $\phi(0, \xi) = \phi(\xi, 0) = 1 + 2\cos(\frac{\pi}{2}\xi)$ and $\phi(1, \xi) = \phi(\xi, 1) = 1 - \pi\xi$. This problem has an exact solution

$$u(x, y) = \left\{ \begin{array}{ll} 1 + 2\cos(\frac{\pi}{2}(x+y)), & x + y \leq 1, \\ 1 + \pi(1 - x - y) & \text{otherwise.} \end{array} \right.$$

We use method (1.6) with the five-point finite difference method. The stopping criterion is $\|F_n(x)\|_\infty \leq 10^{-12}$. Here F_n is the finite difference approximation function with grids n. We report the value of $\|F_n(x)\|_\infty$ at the last five iterations.

Nonsmooth optimization and operator equations involving nonsmooth operators are becoming crucial in various areas of computational and applied mathematics, for example, in nonsmooth mechanics [21, 22, 30], optimal design of electromagnetic

TABLE 1
Numerical result of Example 4.1: $\|F_n(x^k)\|_\infty$.

n	$k=7$	$k=8$	$k=9$	$k=10$	$k=11$
625	1.3×10^{-8}	1.0×10^{-9}	8.3×10^{-11}	6.6×10^{-12}	5.4×10^{-13}
1225	6.8×10^{-9}	5.5×10^{-10}	4.4×10^{-11}	3.6×10^{-12}	2.9×10^{-13}
2025	4.1×10^{-9}	3.4×10^{-10}	2.7×10^{-11}	2.2×10^{-12}	1.8×10^{-13}

devices [21], ill-posed problems involving nonsmooth operators and variational inequalities [19, 24], bounded variation regularization and nondifferentiable optimization problems in nonreflexive spaces [26, 27], inverse source problems [25], free boundary problems [28], multibody system identification [1], and nonlinear complementary problems (see [8] and references cited therein). Various classes of these problems can be reformulated as nonsmooth equations with locally Lipschitzian operators. Hence the smoothing methods and semismooth methods studied in this paper can be applied to these problems.

Acknowledgment. This research was initiated while the first author worked in the School of Applied Mathematics at the University of New South Wales where the second author visited in 1993 and 1998. We would like to express our appreciation for the support.

REFERENCES

[1] M. ANITESCU, D. I. COROIAN, M. Z. NASHED, AND F. A. POTRA, *Outer inverses and multi-body system identification*, Numer. Funct. Anal. Optim., 17 (1996), pp. 661–678.

[2] J. BURKE AND S. XU, *The global linear convergence of a non-interior path-following algorithm for linear complementarity problem*, Math. Oper. Res., 23 (1998), pp. 719–734.

[3] B. CHEN AND P. T. HARKER, *A non-interior-point continuation method for linear complementarity problems*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 1168–1190.

[4] B. CHEN AND P. T. HARKER, *Smooth approximations to nonlinear complementarity problems*, SIAM J. Optim., 7 (1997), pp. 403–420.

[5] C. CHEN AND O. L. MANGASARIAN, *A class of smoothing functions for nonlinear and mixed complementarity problems*, Comput. Optim. Appl., 5 (1996), pp. 97–138.

[6] X. CHEN, *Superlinear convergence of smoothing quasi-Newton methods for nonsmooth equations*, J. Comput. Appl. Math., 80 (1997), pp. 105–126.

[7] X. CHEN, *Global and superlinear convergence of inexact Uzawa methods for saddle point problems with nondifferentiable mappings*, SIAM J. Numer. Anal., 35 (1998), pp. 1130–1148.

[8] X. CHEN, Z. NASHED, AND L. QI, *Convergence of Newton's method for singular smooth and nonsmooth equations using adaptive outer inverses*, SIAM J. Optim., 7 (1997), pp. 445–462.

[9] 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*, Math. Comp., 67 (1998), pp. 519–540.

[10] X. CHEN AND T. YAMAMOTO, *Convergence domains of certain iterative methods for solving nonlinear equations*, Numer. Funct. Anal. Optim., 10 (1989), pp. 37–48.

[11] X. CHEN AND Y. YE, *On homotopy-smoothing methods for box-constrained variational inequalities*, SIAM J. Control Optim., 37 (1999), pp. 589–616.

[12] F. H. CLARKE, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[13] A. FISHER, *Solution of monotone complementarity problems with locally Lipschitzian functions*, Math. Programming, 76 (1997), pp. 513–532.

[14] 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, 1997, pp. 105–116.

[15] M. HEINKENSCHLOSS, C. T. KELLEY, AND H. T. TRAN, *Fast algorithms for nonsmooth compact fixed-point problems*, SIAM J. Numer. Anal., 29 (1992), pp. 1769–1792.

[16] C. T. KELLEY, *Identification of the support of nonsmoothness,* in Large Scale Optimization:

State of the Art, W. W. Hagar, D. W. Hearn, and P. M. Pardalos, eds., Kluwer Academic Publishers, Boston, 1993, pp. 192–205.

[17] C. T. Kelley and E. W. Sachs, *Multilevel algorithms for constrained compact fixed point problems*, SIAM J. Sci. Comput., 15 (1994), pp. 645–667.

[18] F. Kikuchi, *Finite element analysis of a nondifferentiable nonlinear problem related to MHD equilibria,* J. Fac. Sci. Univ. Tokyo Sect. IA Math., 35 (1988), pp. 77–101.

[19] F. Liu and M. Z. Nashed, *Regularization of nonlinear ill-posed variational inequalities and convergence rates*, Set-Valued Anal., 6 (1998), pp. 313–344.

[20] R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optim., 15 (1977), pp. 959–972.

[21] J. J. Moreau, P. D. Panagiotopoulos, and G. Strang, eds., *Topics in Nonsmooth Mechanics*, Birkhäuser, Basel, 1988.

[22] Z. Naniewicz and P. D. Panagiotopoulos, *Mathematical Theory of Hemivariational Inequalities and Applications*, Marcel Dekker, Inc., New York, 1995.

[23] M. Z. Nashed, *Differentiability and related properties of nonlinear operators: Some aspects of the role of differentials in nonlinear function analysis*, in Nonlinear Functional Analysis and Applications, L. B. Rall, ed., Academic Press, New York, 1971, pp. 103–309.

[24] M. Z. Nashed and F. Liu, *On nonlinear ill-posed problems* II: *Monotone operator equations and monotone variational inequalities*, in Theory and Applications of Nonlinear Operators of Accretive and Monotone Type, A. G. Kartsatos, ed., Marcel Dekker, Inc., New York, 1996, pp. 223–240.

[25] M. Z. Nashed and O. Scherzer, *Stable approximation of nondifferentiable optimization problems with variational inequalities*, in Recent Developments in Optimization Theory and Nonlinear Analysis, Y. Censor and S. Reich, eds., Contemp. Math. 204, AMS, Providence, RI, 1997, pp. 155–170.

[26] M. Z. Nashed and O. Scherzer, *Stable approximation of a minimal surface problem with variational inequalities*, Abstr. Appl. Anal., 2 (1997), pp. 137–161.

[27] M. Z. Nashed and O. Scherzer, *Least squares and bounded variation regularization with nondifferentiable functionals*, Numer. Funct. Anal. Optim., 19 (1998), pp. 873–901.

[28] P. Neittaanmäki, ed., *Numerical Methods for Free Boundary Problems*, Internat. Ser. Numer. Math. 99, Birkhäuser, Basel, 1991.

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

[30] P. D. Panagiotopoulos, *Inequality Problems in Mechanics and Applications*, Birkhäuser, Basel, 1985.

[31] J.-S. Pang and L. Qi, *Nonsmooth equations: Motivation and algorithms*, SIAM J. Optim., 3 (1993), pp. 443–465.

[32] L. Qi, *Convergence analysis of some algorithms for solving nonsmooth equations,* Math. Oper. Res., 18 (1993), pp. 227–244.

[33] L. Qi, *C-differentiability, C-differential operators and generalized Newton methods*, AMR 96/5, Applied Mathematics Report, University of New South Wales, Sydney, 1996.

[34] L. Qi and X. Chen, *A globally convergent successive approximation method for severely nonsmooth equations*, SIAM J. Control Optim., 33 (1995), pp. 402–418.

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

[36] L. B. Rall, *Computational Solution of Nonlinear Operator Equations*, Wiley, New York, 1969.

[37] J. Rappaz, *Approximation of a nondifferentiable nonlinear problem related to MHD equilibria*, Numer. Math., 45 (1984), pp. 117–133.

[38] W. C. Rheinboldt, *A unified convergence theory for a class of iterative process*, SIAM J. Numer. Anal., 5 (1968), pp. 42–63.

[39] S. M. Robinson, *Local structure of feasible sets in nonlinear programming. Part* III. *Stability and sensitivity*, Math. Programming Stud., 30 (1987), pp. 45–66.

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

[41] A. Shapiro, *On concepts of directional differentiability*, J. Optim. Theory Appl., 66 (1990), pp. 477–487.

[42] I. Stakgold, *Green's Functions and Boundary Value Problems*, 2nd ed., Wiley, New York, 1998.