MATHEMATICS OF OPERATIONS RESEARCH
Vol. 16, No. 2, May 1991
Printed in U.S.A.

AN IMPLICIT-FUNCTION THEOREM FOR A CLASS OF NONSMOOTH FUNCTIONS*†

STEPHEN M. ROBINSON

In this paper we introduce the concept of strong approximation of functions, and a related concept of strong Bouligand (B-) derivative, and we employ these ideas to prove an implicit-function theorem for nonsmooth functions. This theorem provides the same kinds of information as does the classical implicit-function theorem, but with the classical hypothesis of strong Fréchet differentiability replaced by strong approximation, and with Lipschitz continuity replacing Fréchet differentiability of the implicit function. Therefore it is applicable to a considerably wider class of functions than is the classical theorem.

In the last part of the paper we apply this implicit function result to analyze local solvability and stability of perturbed generalized equations.

1. Introduction. This paper develops an implicit-function theorem for functions which, although not differentiable in the conventional sense, can be approximated in a certain strong sense by functions having Lipschitzian inverses. This theorem resembles the usual implicit-function theorem, in that when the "partial derivative" of a function with respect to a certain set of variables is invertible, then the requirement that the function be zero defines the values of those variables as functions of the other variables appearing in it. However, the words "partial derivative" appear in quotation marks because in this case the derivative involved is not the usual partial derivative, but rather the approximating function mentioned above.

In addition to proving this theorem, we prove an analogue for locally Lipschitzian functions of the Banach lemma for linear operators, which may have other applications. We also establish some techniques for approximating the implicit function whose existence we prove, and we apply the results to sensitivity analysis for nonlinear generalized equations on Hilbert spaces.

Many authors have recently contributed to the study of implicit-function and inverse function theorems for functions that are not differentiable in the conventional sense. For example, Aubin [2] developed numerous results dealing with the approximation of multifunctions in various spaces, and the use of these approximations to provide different concepts of derivative. Further, in [1] he studied Kuratowski limits of epigraphs of functions, and showed how to derive properties of the limit multifunctions from these.

*Received August 15, 1988; revised December 1989.
AMS 1980 subject classification. Primary: 47H15, 49A27, 90C30.
IAOR 1973 subject classification. Main: Programming: Nonlinear.
OR/MS Index 1978 subject classification. Primary: 434 Mathematics/Functions. Secondary: 657 Programming/Nonlinear/Theory.
Key words. Implicit function, Bouligand derivative, B-derivative, strong B-derivative, implicit differentiation, variational inequality, generalized equation, strong approximation.
†The research reported here was sponsored by the National Science Foundation under Grants CCR-8502202 and CCR-8801489, and by the Air Force Systems Command, USAF, under Grants AFOSR-88-0090 and AFOSR-89-0058. The U.S. Government has certain rights to this material, and is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation thereon. A preliminary version of this paper appeared in July 1988 as Working Paper WP-88-67 of the International Institute for Applied Systems Analysis, Laxenburg, Austria.

0364-765X/91/1602/0292/$01.25

Recently, Frankowska studied inverse-function theorems using hypotheses on higher-order variations of the function in question, rather than the usual assumptions on the first derivative. These hypotheses enabled her to obtain results for various kinds of set-valued maps, as well as for maps defined on spaces more general than the usual normed linear space. These results are reported in [7], [8], [9]; see also further references therein.

Cornet and Laroque [5] applied the Clarke inverse-function theorem [4, §7.1] to the problem of sensitivity in nonlinear programming, and showed that when a certain generalized Jacobian is nonsingular one has existence of a Lipschitzian trajectory of optimizers. Along much the same lines, Jongen et al. [16] recently proved an implicit-function theorem for nonsmooth functions in $\mathbf{R}^n$ under the basic assumption that the generalized Jacobian was nonsingular. They the applied this result to the problem of stability in nonlinear programming. Their inverse-function theorem strongly resembles that of Clarke, but is stated in function spaces rather than in $\mathbf{R}^n$.

In [31], Pshenichny proved implicit-function theorems for multifunctions with convex graphs, and for other functions whose graphs could be locally approximated in a certain sense. Related work was given in [42] by Ursescu and in [33] by the present author.

In [14] Jittorntrum proved under very weak assumptions a theorem establishing the existence of a continuous implicit function. The basic tool used in his proof was the Brouwer theorem on invariance of domain.

Several papers of Ioffe have extended the classical inverse-function and implicit-function theorems to more general classes of functions. For example, [11] gave such theorems in terms of "predifferentials," and [12, Theorems 11.14, 11.16] presented such results in terms of "fans," each of these being a type of approximating function or multifunctions extending the classical notion of derivative. In [13, Theorem 2] a global inverse-function theorem under a hypothesis on the modulus of surjection of the function involved is presented.

Recent work of Rockafellar [39], [40] and of King and Rockafellar [18] has dealt with stability questions in parametric optimization and parametric generalized equations. Much of this work relates closely to the applications we discuss in §4.

Finally, in a very remarkable paper [21] Kummer has given a *characterization* of Lipschitzian invertibility for functions on $\mathbf{R}^n$, using directional derivatives. He also gives an example to show that his hypotheses are essentially weaker than those of the Clarke inverse-function theorem [4], and that therefore his technique can establish invertibility in some cases in which the Clarke theorem cannot be applied. Related work, and applications to parametric optimization, are in [22] and [23].

A considerable amount of work in this area has been focused on more specific classes of problems, particularly those involving generalized equations and this application to the solution of optimization problems. An implicit-function type theorem for generalized equations was established in [34], and was employed there to obtain sensitivity results about nonlinear programming problems. Kojima [19] obtained similar sensitivity information using very different techniques. Later, Jittorntrum [15] showed that optimal solutions of nonlinear programming problems were directionally differentiable under suitable assumptions. Then, in [36] and [37], Robinson established B-differentiability properties of solutions of finite-dimensional variational inequalities over polyhedral sets, and nonlinear programming problems, respectively.

In a recent series of papers [24], [25], [26] Kyparisis has extended the above work. In [25] he showed how to extend the type of result proved in [15] and [34] to variational problems over sets defined by systems of inequalities and equations, while in [24] he developed differentiability properties of a specific type of generalized equation (the nonlinear complementarity problem). Finally, in [26] he dealt with the case of variational problems over polyhedral convex sets, and showed how the results of [33] and [34] could be extended and sharpened, in particular giving conditions for continuous differentiability (as opposed to B-differentiability).

Dafermos [6] studied variational problems in which the underlying set may vary, and particularly in which it may be given by a system of nonlinear equations and inequalities. For her results, she assumed strong monotonicity of the function involved. Another approach to this general problem area was explored by Qiu and Magnanti [32], who considered the case in which the solution could be multivalued, and introduced ideas of directional differentiability for such functions.

Work very close to that of this paper was reported in two recent papers of Pang [27], [28]. In [27], Pang developed a Newton iterative method for solving equations involving B-differentiable functions. He analyzed this method under the assumption that the key derivative involved was a "strong B-derivative," as defined in [27], and he applied the method to some problems in mathematical programming. Although we introduce a strong B-derivative in this paper also, our definition is not equivalent to that of [27].

In [28], Pang considered variational problems over polyhedral convex sets, and he determined when the solutions of such problems were Fréchet differentiable. He suggested a continuation-type method for computing such problems. A similar continuation method was proposed and investigated by Park [29].

The rest of this paper is organized in three sections: in §2 we briefly review B-derivatives, and introduce new concepts of strong approximation and strong B-derivative. We show that with strong B-derivatives one can establish a formula for the B-derivative of a function of several variables as a linear expression in partial B-derivatives. Then in §3 we give the main results of the paper: an extension of the classical Banach lemma of functional analysis to locally Lipschitzian functions, an implicit-function theorem, and a theorem on approximation of the implicit function by simpler functions. Finally, in §4 we sketch an application of this theory to parametric solutions of generalized equations.

2. Approximation and strong approximation. This section introduces two kinds of functional approximation that we shall need in what follows. These extend the principal ideas inherent in the concept of ordinary or strong Fréchet derivative to situations in which the approximating functions may not be linear.

For the rest of this section we let X, Y, W, and Z be normed linear spaces, and we consider neighborhoods Ξ of x_0 in X, H of y_0 in Y, and Ω of w_0 in W. F and G will be functions from $\Xi \times \mathrm{H}$ to Z, h_1 and h_2 functions from Ω to Z, f a function from Ξ to Z, and g a function from H to Z.

DEFINITION 2.1. f approximates F in x at (x_0, y_0) (written $f \sim_x F$ at (x_0, y_0)) if

$$F(x, y_0) - f(x) = o(x - x_0).$$

Similarly, $g \sim_y F$ at (x_0, y_0) if

$$F(x_0, y) - g(y) = o(y - y_0).$$

This approximation is the kind that one is used to seeing in partial derivatives: for example, to say that F has a partial derivative in x at (x_0, y_0) amounts to saying that $F(x_0, y_0) + F_x(x_0, y_0)(x - x_0) \sim_x F$ at (x_0, y_0). However, for the implicit-function analysis of §3 we shall need a stronger notion, given in the next definition.

DEFINITION 2.2. h_1 strongly approximates h_2 at w_0 (written $h_1 \approx h_2$ at w_0) if for each $\epsilon > 0$ there exists a neighborhood U of w_0 such that whenever w and w' belong to U,

$$\|[h_1(w) - h_2(w)] - [h_1(w') - h_2(w')]\| \leqslant \epsilon(\|w - w'\|).$$

This definition amounts to requiring $h_1 - h_2$ to have a strong Fréchet derivative of zero at w_0, though neither h_1 nor h_2 is assumed to be differentiable in any sense. Evidently, for fixed w_0, $\approx$ is an equivalence relation. It is also clear that we could define strong approximation for functions of several groups of variables (for example, $G \approx_{(x,y)} F$ at (x_0, y_0)) by replacing W by $X \times Y$ and making the obvious substitutions.

The following proposition shows that $\approx$ is preserved under composition from the right with any Lipschitzian function. We introduce temporarily another normed linear space S and a function q from S to W with $q(s_0) = w_0$.

PROPOSITION 2.3. *Suppose that q is locally Lipschitzian at s_0. If $h_1 \approx h_2$ at w_0, then $h_1 \circ q \approx h_2 \circ q$ at s_0.*

PROOF. Let the local Lipschitz modulus of q at s_0 be λ. If s and s' are near s_0, then $w = q(s)$ and $w' = q(s')$ will be near w_0. For any positive ϵ, if s and s' are close enough to s_0, we have

$$\begin{aligned}
&\|[(h_1 \circ q)(s) - (h_2 \circ q)(s)] - [(h_1 \circ q)(s') - (h_2 \circ q)(s')]\| \\
&\quad = \|[h_1(w) - h_2(w)] - [h_1(w') - h_2(w')]\| \\
&\quad \leqslant \epsilon\|w - w'\| = \epsilon\|q(s) - q(s')\| \leqslant \epsilon\lambda\|s - s'\|.
\end{aligned}$$

As ϵ was arbitrary, we have $(h_1 \circ q) \approx (h_2 \circ q)$ at s_0. ■

We shall also use the following definition of strong approximation involving functions such as f and g, defined on only one of the spaces involved.

DEFINITION 2.4. f strongly approximates F in x at (x_0, y_0) (written $f \approx_x F$ at (x_0, y_0)) if for each $\epsilon > 0$ there exist neighborhoods U of x_0 and V of y_0 such that whenever x and x' belong to U and y belongs to V we have

$$\|[F(x, y) - f(x)] - [F(x', y) - f(x')]\| \leqslant \epsilon\|x - x'\|.$$

We make a similar definition for strong approximation in y. Note that one has both $f \approx_x F$ and $g \approx_y F$ at (x_0, y_0) exactly when $f(x) + g(y) \approx_{(x,y)} F$ at (x_0, y_0) in the sense of Definition 2.2.

We saw in Proposition 2.3 that strong approximation was preserved under composition with a locally Lipschitzian function. Exactly the same phenomenon appears in the case of the "single-variable" strong approximation we have just defined. We show this in the following result.

PROPOSITION 2.5. *Let h be a function from Ω into X, locally Lipschitzian at w_0, with $h(w_0) = x_0$. Define r and R by $r = f \circ h$ and $R(w, y) = F(h(w), y)$. If $f \approx_x F$ at (x_0, y_0), then $r \approx_w R$ at (w_0, y_0).*

PROOF. Let $\epsilon > 0$ and assume that $f \approx_x F$ at (x_0, y_0). Find neighborhoods U of x_0 and V of y_0 so that if x and x' belong to U and $y \in V$ then

$$\|[F(x, y) - f(x)] - [F(x', y) - f(x')]\| \leqslant \epsilon\|x - x'\|.$$

Let P be a neighborhood of w_0 small enough so that h is Lipschitzian on P with modulus λ, and so that if $w \in P$ then $h(w) \in U$. For any w and w' in P and any $y \in V$, we have

$$\begin{aligned}
&\|[R(w, y) - r(w)] - [R(w', y) - r(w')]\| \\
&\quad = \|[F(h(w), y) - f(h(w))] - [F(h(w'), y) - f(h(w'))]\| \\
&\quad \leqslant \epsilon \|h(w) - h(w')\| \leqslant \epsilon\lambda\|w - w'\|,
\end{aligned}$$

which shows that $r \approx_w R$ at (w_0, y_0). ∎

We shall use these definitions and relationships later in the paper, but we apply them here to obtain some new results about so-called Bouligand derivatives, or B-derivatives, of functions. In order to make clear what is involved, we provide here a definition and a brief summary of some properties of B-derivatives.

Given a function f from an open subset Ξ of a normed linear space X into another normed linear space Z, we say that f is *B-differentiable* at a point $x_0 \in \Xi$ if there is a positively homogeneous function $Df(x_0)$: $X \to Z$ such that

$$f(x_0 + h) = f(x_0) + Df(x_0)h + o(h). \tag{2.1}$$

In terms of the notation we have just introduced, (2.1) means that $f(x_0) + Df(x_0)(x - x_0) \sim_x f$ at x_0. The function $Df(x_0)$ is necessarily unique if it exists.

The B-derivative terminology was introduced in [37], and the results cited were established in the finite-dimensional case. Subsequently, Shapiro [41] showed that in a great many situations this definition and several others introduced in the literature are actually equivalent. This equivalence is especially useful in the case of finite-dimensional spaces.

The definition of the B-derivative somewhat resembles that of the classical Fréchet (F-) derivative, but there are some differences. One of these is that, although we can define a partial B-derivative in the usual way for a function $F(x, y)$ from a product $X \times Y$ of normed linear spaces to Z (by defining, for example, $D_xF(x_0, y_0)$ to be the B-derivative at x_0 of $F(\cdot, y_0)$), we do not obtain the addition formula: in general,

$$DF(x_0, y_0)(h, k) \neq D_xF(x_0, y_0)h + D_yF(x_0, y_0)k. \tag{2.2}$$

To see that inequality holds in general, we can consider the Euclidean norm function on $\mathbf{R}^2$: its B-derivative at the origin is itself, yet the partial B-derivatives are the absolute value functions of the coordinates, and the sum of these is not the Euclidean norm.

In order to recover the addition formula, we need to strengthen the requirements placed on a B-derivative. For that purpose we introduce the following definition.

DEFINITION 2.6. Let F: $X \times Y \to Z$, and suppose F has a partial B-derivative $D_xF(x_0, y_0)$ with respect to x at (x_0, y_0). We say $D_xF(x_0, y_0)$ is *strong* if $F(x_0, y_0) + D_xF(x_0, y_0)(x - x_0) \approx_x F$ at (x_0, y_0).

In [27], Pang also introduced a definition of strong B-derivative, but his definition differs from the above in the case in which $D_xF(x_0, y_0)$ is not a linear operator. He showed that a strong B-derivative under his definition must in fact be a strong F-derivative; thus his requirements are considerably stronger than those of Definition 2.6, since that definition does not imply that the function in question is F-differentiable. For a simple example of such a case, consider the function F defined on $\mathbf{R}$ by

$$F(x) = \begin{cases} x + x^2, & \text{if } x \geqslant 0, \\ -x - x^2, & \text{if } x < 0. \end{cases}$$

The following result shows that if one of the B-derivatives involved in the sum formula is strong, then equality holds in (2.2). Further, if both are strong, then the B-derivative obtained from the sum formula is strong too. We use as norm on $X \times Y$ the sum of the norms on X and Y.

PROPOSITION 2.7. *Let F: $X \times Y \to Z$, and let $(x_0, y_0) \in X \times Y$. Assume that F has partial B-derivatives with respect to x and to y at (x_0, y_0).*

(a) *If $D_x F(x_0, y_0)$ is strong, then F is B-differentiable at (x_0, y_0), and*

$$DF(x_0, y_0)(h, k) = D_x F(x_0, y_0)h + D_y F(x_0, y_0)k. \tag{2.3}$$

(b) *If both $D_x F(x_0, y_0)$ and $D_y F(x_0, y_0)$ are strong, then $DF(x_0, y_0)$ is strong.*

PROOF. For simplicity we suppress (x_0, y_0) and write D, D_x, and D_y. Assume first that the latter two exist and that D_x is strong. Then

$$\begin{aligned} F(x_0 + h, y_0 + k) &- F(x_0, y_0) - D_x h - D_y k \\ &= \left[F(x_0 + h, y_0 + k) - F(x_0, y_0 + k) - D_x h\right] \\ &\quad + \left[F(x_0, y_0 + k) - F(x_0, y_0) - D_y k\right]. \end{aligned}$$

For any $\epsilon > 0$ and for small enough $\|h\|$ and $\|k\|$, the first group of terms on the right is bounded in norm by $\epsilon\|h\|$ and the second by $\epsilon\|k\|$. This is enough to prove (a).

Now if both partial derivatives are strong, then we have

$$\begin{aligned} &\left[F(x_0 + h_1, y_0 + k_1) - F(x_0, y_0) - D_x h_1 - D_y k_1\right] \\ &\quad - \left[F(x_0 + h_2, y_0 + k_2) - F(x_0, y_0) - D_x h_2 - D_y k_2\right] \\ &= \left[F(x_0 + h_1, y_0 + k_1) - F(x_0, y_0) - D_x h_1\right] \\ &\quad - \left[F(x_0 + h_2, y_0 + k_1) - F(x_0, y_0) - D_x h_2\right] \\ &\quad + \left[F(x_0 + h_2, y_0 + k_1) - F(x_0, y_0) - D_y k_1\right] \\ &\quad - \left[F(x_0 + h_2, y_0 + k_2) - F(x_0, y_0) - D_y k_2\right]. \end{aligned}$$

For any small positive ϵ and for h_1, h_2, k_1, and k_2 close to zero, the difference of the first two terms is bounded in norm by $\epsilon\|h_1 - h_2\|$, and that of the second by $\epsilon\|k_1 - k_2\|$. Therefore the entire expression is bounded by $\epsilon(\|h_1 - h_2\| + \|k_1 - k_2\|)$, and so the total B-derivative $DF(x_0, y_0)$ is strong. ■

3. Inversion and implicit-function results. This section contains the principal results of the paper. We first establish an analogue of the Banach lemma on invertibility of perturbed linear operators, then prove the main implicit-function theorem, and finally establish an approximation result for the implicit function.

The following lemma can be regarded as an extension of the well known Banach perturbation lemma [17, Th. 4(2.V)] from linear operators to locally Lipschitzian functions. The classical lemma says that if an invertible linear operator A is perturbed by adding another linear operator of norm less than $\|A^{-1}\|^{-1}$, then the sum is still invertible and the norm of its inverse is bounded by a simple formula. The present lemma gives a similar statement but with linear operators replaced by locally Lipschitzian functions, and with the norms replaced by the Lipschitz moduli of the functions.

In stating the lemma we use the symbol $B(x, \rho)$ to indicate the closed ball of radius ρ about a point x. Also, for a function f from a metric space (X, d) to another metric space (Y, e) we let

$$\delta(f, X) = \inf\{e[f(x_1), f(x_2)]/d(x_1, x_2) | x_1 \neq x_2, x_1, x_2 \in X\}.$$

This quantity is used in [30], with the notation d instead of δ. Clearly $\delta(f, X) \neq 0$ only if f is one-to-one on X; however, δ actually provides more information (it is actually the reciprocal of a Lipschitz constant for the inverse of f).

LEMMA 3.1.† *Let (X, d) be a complete metric space, Ω an open subset of X, and Y a normed linear space. Let f and h be functions from Ω into Y, h being Lipschitzian with modulus η. Let $x_0 \in \Omega$ with $f(x_0) = y_0$. Assume that*:
1. $f(\Omega) \supset B(y_0, \alpha)$,
2. $0 \leqslant \eta < \delta := \delta(f, \Omega)$,
3. $\Omega \supset B(x_0, \delta^{-1}\alpha)$, *and*
4. $\theta := (1 - \eta\delta^{-1})\alpha - \|h(x_0)\| \geqslant 0$.

Then

$$(f + h)\big(B(x_0, \delta^{-1}\alpha)\big) \supset B(y_0, \theta),$$

and

$$\delta(f + h, \Omega) \geqslant \delta - \eta > 0.$$

PROOF. For $y \in B(y_0, \theta)$ and $x \in B(x_0, \delta^{-1}\alpha)$ let $T_y(x) = f^{-1}(y - h(x))$. Note that

$$\|y - h(x) - y_0\| \leqslant \|y - y_0\| + \|h(x) - h(x_0)\| + \|h(x_0)\|$$

$$\leqslant \theta + \eta(\delta^{-1}\alpha) + \|h(x_0)\| = \alpha,$$

so that the set $T_y(x)$ is nonempty; in fact it consists of a single point because $\delta > 0$. Therefore T_y is a function on $B(x_0, \delta^{-1}\alpha)$. This function is actually a self-map of $B(x_0, \delta^{-1}\alpha)$ because for each x in that ball,

$$d\big(T_y(x), x_0\big) = d\big(f^{-1}(y - h(x)), f^{-1}(y_0)\big) \leq \delta^{-1}\alpha.$$

†After this paper had been accepted for publication, Alexander Ioffe kindly informed the author of related work appearing in the paper of A. V. Dimitruk, A. A. Milyutin, and N. P. Osmolovskii, "Lyusternik's theorem and the theory of extrema," Russian Mathematical Surveys 1980, No. 6, pp. 11–51; cf. Theorems 1.2, 1.3.

Finally, if x_1 and x_2 are elements of $B(x_0, \delta^{-1}\alpha)$, then

$$d\big(T_y(x_1), T_y(x_2)\big) = d\big(f^{-1}(y - h(x_1)), f^{-1}(y - h(x_2))\big)$$
$$\leq \delta^{-1}\eta d(x_1, x_2),$$

and since $\delta^{-1}\eta < 1$ we see that T_y is a strong contraction. Now by the contraction mapping principle [17, Th. 1(1.XVI)] T_y has a fixed point $x(y)$ in $B(x_0, \delta^{-1}\alpha)$. Clearly $(f + h)(x(y)) = y$, and $x(y)$ is the only point in Ω satisfying this equation since

$$\begin{aligned}\delta(f + h, \Omega) &= \inf\{\|[f(u) - f(v)] + [h(u) - h(v)]\|/d(u, v) | u \neq v, u, v \in \Omega\} \\ &\geq \delta(f, \Omega) - \sup\{\|h(u) - h(v)\|/d(u, v) | u \neq v, u, v \in \Omega\} \\ &\geq \delta - \eta > 0,\end{aligned}$$

and therefore in particular $f + h$ is one-to-one on Ω. ■

Here is the main implicit-function theorem.

THEOREM 3.2. *Let X be a Banach space and Y and Z be normed linear spaces. Let x_0 and y_0 be points of X and Y respectively, and let Ξ be a neighborhood of x_0 and H a neighborhood of y_0. Suppose F is a function from $\Xi \times \mathrm{H}$ to Z with $F(x_0, y_0) = 0$, and f is a function from Ξ to Z with $f(x_0) = 0$. Assume further that*

a. $f \approx_x F$ *at* (x_0, y_0).

b. *For each $x \in \Xi$, $F(x, \cdot)$ is Lipschitzian on H with modulus ϕ.*

c. *$f(\Xi)$ is a neighborhood of the origin in Z.*

d. $\delta(f, \Xi) =: d_0 > 0$.

Then for each $\lambda > d_0^{-1}\phi$ there exist neighborhoods $U(x_0)$ and $V(y_0)$, and a function x from V to U, such that:

i. $x(y_0) = x_0$.

ii. *$x(\cdot)$ is Lipschitzian on V with modulus λ.*

iii. *For each $y \in V$, $x(y)$ is the unique solution in U of $F(x, y) = 0$.*

PROOF. Fix $\lambda > d_0^{-1}\phi$. Choose positive numbers α, κ, and ϵ, and a neighborhood V of y_0 such that $V \subset \mathrm{H}$ and the following hold:

(1) $\epsilon < d_0$ and $\phi(d_0 - \epsilon)^{-1} \leq \lambda$.

(2) For each x, x' in $B(x_0, d_0^{-1}\alpha)$ and each y in V, $\|[F(x', y) - f(x')] - [F(x, y) - f(x)]\| \leq \epsilon\|x' - x\|$.

(3) $f(\Xi) \supset B(0, \alpha)$.

(4) For each y in V, $\kappa + \|F(x_0, y)\| \leq (1 - \epsilon d_0^{-1})\alpha$.

Now write $\Omega = B(x_0, d_0^{-1}\alpha)$. Fix $y \in V$ and define a function h_y from Ω into Z by $h_y(x) = F(x, y) - f(x)$. Note that by point (2) above, h_y is Lipschitzian on Ω with modulus ϵ. We now verify in turn each of the hypotheses of Lemma 3.1. The quantities η, h, and d appearing in the lemma correspond to ϵ, h_y, and d_0 here.

For hypothesis 1, let $z \in B(0, \alpha)$. From point (3) above, we know there is some $x \in \Xi$ with $f(x) = z$. Further,

$$\|z - 0\| = \|f(x) - f(x_0)\| \geq d_0\|x - x_0\|,$$

so that $\|x - x_0\| \leq d_0^{-1}\|z\| \leq d_0^{-1}\alpha$. Therefore $f(\Omega) \supset B(0, \alpha)$.

To verify hypothesis 2, note that by point (1), $0 < \epsilon < d_0$, where d_0 was chosen to be $\delta(f, \Xi)$; this is not greater than $\delta(f, \Omega)$ because $\Omega \subset \Xi$.

For hypothesis 3, note that $\Omega = B(x_0, d_0^{-1}\alpha)$ by definition.

Finally, for hypothesis 4 we use point (4) to verify that

$$\theta(y) := (1 - \epsilon d_0^{-1})\alpha - \|h_y(x_0)\| = (1 - \epsilon d_0^{-1})\alpha - \|F(x_0, y)\| \geqslant \kappa > 0.$$

Now we apply the lemma to conclude that (for this fixed y)

$$\delta(F(\cdot, y), \Omega) \geqslant d_0 - \epsilon, \tag{3.1}$$

and

$$F(\cdot, y)(\Omega) \supset B(0, \theta(y)) \supset B(0, \kappa). \tag{3.2}$$

For each $z \in B(0, \kappa)$ and each $y \in V$, let $j_y(z)$ be the unique $x \in \Omega$ with $F(x, y) = z$; existence and uniqueness are guaranteed by (3.2) and (3.1) respectively. Define x from V to Ω by $x(y) = j_y(0)$. Take U to be the ball Ω; then conclusion (iii) of the theorem will hold. Conclusion (i) also holds because $F(x_0, y_0) = 0$.

Now for any y_1 and y_2 in V, (3.1) yields

$$\begin{aligned}(d_0 - \epsilon)\|x(y_2) - x(y_1)\| &\leqslant \|F(x(y_2), y_2) - F(x(y_1), y_2)\| \\ &= \|F(x(y_1), y_1) - F(x(y_1), y_2)\| \\ &\leqslant \phi\|y_1 - y_2\|,\end{aligned}$$

so that $x(\cdot)$ is Lipschitzian on V with modulus $\phi(d_0 - \epsilon)^{-1} \leqslant \lambda$, and therefore conclusion (ii) holds. ∎

The next theorem gives information about how to approximate the implicit function $x(y)$ whose existence was established by Theorem 3.2. We approximate the function $F(x, y)$ by the separable function $f(x) + g(y)$, then define $x_f(y)$ by $f(x_f(y)) + g(y) = 0$; under the assumptions of Theorem 3.2, x_f will be well defined because f^{-1} is defined on a neighborhood of the origin. We then ask how we can use approximation properties of f and g to infer approximation properties of x_f.

For some applications (such as those discussed in §4) it may be relatively easy to compute $x_f(y)$. In such cases this method can provide good numerical approximations to values of x.

THEOREM 3.3. *Let x_0, y_0, F, and f be as in Theorem 3.2. Let g be a continuous function from* H *to Z with $g(y_0) = 0$, and define $x_f(y)$ for y near y_0 by*

$$f(x_f(y)) + g(y) = 0.$$

(a) *If $g \sim_y F$ at (x_0, y_0), then $x_f \sim x$ at y_0.*

(b) *If $g \approx_y F$ at (x_0, y_0) and f is a continuous linear transformation, then $x_f \approx x$ at y_0.*

Note that the hypothesis in (b) means that f is the strong partial Fréchet derivative of F in x at (x_0, y_0), since the hypotheses of Theorem 3.2 include the assumption that $f \approx_x F$ at (x_0, y_0).

PROOF. Fix $\epsilon > 0$. First, using the assumptions in (a), choose a neighborhood $Q \subset V$ of y_0 such that whenever $y \in Q$, one has $x(y)$ and $x_f(y)$ in U, with

$$\|[F(x(y), y) - f(x(y))] - [F(x_0, y) - f(x_0)]\| \leqslant \epsilon\|x(y) - x_0\|, \tag{3.3}$$

and

$$\|F(x_0, y) - g(y)\| \leqslant \epsilon \|y - y_0\|. \tag{3.4}$$

Now we have

$$\begin{aligned} 0 &= F(x(y), y) - f(x_f(y)) - g(y) \\ &= [F(x(y), y) - f(x(y))] - [F(x_0, y) - f(x_0)] \\ &\quad + [f(x(y)) - f(x_f(y))] + [F(x_0, y) - g(y)]. \end{aligned}$$

Using (3.3) and (3.4) together with the fact that f^{-1} has a local Lipschitz constant of d_0^{-1} at the origin, we find that

$$\begin{aligned} \|x(y) - x_f(y)\| &\leqslant d_0^{-1} \|f(x(y)) - f(x_f(y))\| \\ &\leqslant d_0^{-1}(\epsilon \|x(y) - x_0\| + \epsilon \|y - y_0\|) \leqslant \epsilon d_0^{-1}(\lambda + 1)\|y - y_0\|, \end{aligned}$$

and therefore $x_f \sim x$ at y_0.

Next, assume that $g \approx_y F$ at (x_0, y_0) and that f is linear. Again fix $\epsilon > 0$. Choose neighborhoods $P \subset U$ of x_0 and $Q \subset V$ of y_0 so that if $y \in Q$ then $x(y)$ and $x_f(y)$ belong to P and further, if (x, y) and (x', y') are any elements of $P \times Q$ then

$$\|[F(x', y') - f(x')] - [F(x, y') - f(x)]\| \leqslant \epsilon \|x' - x\|, \tag{3.5}$$

and

$$\|[F(x, y') - g(y')] - [F(x, y) - g(y)]\| \leqslant \epsilon \|y' - y\|. \tag{3.6}$$

If y and y' are any elements of Q then

$$\begin{aligned} &[f(x(y)) - f(x_f(y))] - [f(x(y')) - f(x_f(y'))] \\ &\quad = \{[F(x(y'), y') - f(x(y'))] - [F(x(y), y') - f(x(y))]\} \\ &\quad\quad + \{[f(x_f(y')) + g(y')] - [f(x_f(y)) + g(y)]\} \\ &\quad\quad + \{[F(x(y), y') - g(y')] - [F(x(y), y) - g(y)]\}; \end{aligned} \tag{3.7}$$

recall that $F(x(y), y) = 0 = F(x(y'), y')$. The first group of terms on the right-hand side of (3.7) is bounded in norm by $\epsilon\|x(y') - x(y)\|$ because of (3.5); the second group is zero; the third group is bounded in norm by $\epsilon\|y' - y\|$ because of (3.6). Applying f^{-1} to (3.7) and using the estimates just given, we find that

$$\begin{aligned} \|[x(y) - x_f(y)] - [x(y') - x_f(y')]\| &\leqslant \|f^{-1}\|(\epsilon\|x(y') - x(y)\| + \epsilon\|y' - y\|) \\ &\leqslant \epsilon\|f^{-1}\|(\lambda + 1)\|y' - y\|, \end{aligned}$$

which proves that $x_f \approx x$ at y_0. ∎

One might ask whether the assumption in part (b) of Theorem 3.3 that f is linear could be removed, while still retaining the conclusion that $x_f \approx x$ at y_0. The following counterexample shows that this cannot be done.

Define a piecewise linear homeomorphism of $\mathbf{R}^2$ onto itself by

$$\gamma(y) = \begin{cases} (y_1, y_2), & \text{if } y_2 \geqslant 0, \\ (y_1, 5y_2), & \text{if } y_2 < 0, \end{cases}$$

and let $F: \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}^2$ be defined by

$$F(x, y) = \gamma^{-1}(x) - \left(y_1, y_2 + y_1^2\right).$$

If we let $f(x) = \gamma^{-1}$, then $f \approx_x F$ at $(x_0, y_0) = (0,0)$. We take $g(y) = -y$ (the linearization obtained using the continuous partial F-derivative of F in y at the origin). If we solve for the implicit function $x(y)$, we find that

$$x(y) = \gamma\left(y_1, y_2 + y_1^2\right) = \gamma(y + r(y)),$$

where $r(y) = (0, y_1^2)$. The function x_f is just γ. Now let

$$\Delta(y) = x(y) - x_f(y) = \gamma(y + r(y)) - \gamma(y).$$

To say that $x_f \approx x$ at y_0 is to say that Δ has a strong F-derivative of zero at y_0. We shall show that this is not so.

For large n, define

$$y^1(n) = (n^{-1}, 0), \qquad y^2(n) = (n^{-1}, -2n^{-2}).$$

Clearly both $y^1(n)$ and $y^2(n)$ approach the origin as $n \to \infty$. However,

$$\Delta(y^1(n)) = \gamma((n^{-1}, n^{-2})) - \gamma((n^{-1}, 0)) = (0, n^{-2}),$$

and

$$\begin{aligned}\Delta(y^2(n)) &= \gamma((n^{-1}, -n^{-2})) - \gamma((n^{-1}, -2n^{-2})) \\ &= (n^{-1}, -5n^{-2}) - (n^{-1}, -10n^{-2}) = (0, 5n^{-2}).\end{aligned}$$

Hence

$$\|\Delta(y^2(n)) - \Delta(y^1(n))\| = \|(0, 4n^{-2})\| = 2\|y^2(n) - y^1(n)\|,$$

and therefore $x_f \not\approx x$ at y_0.

We can apply the approximation results of Theorem 3.3 to the particular case in which the approximation functions are B-derivatives, and thereby obtain an analogue of the classical implicit-differentiation formula.

COROLLARY 3.4. *Let X, Y, Z, Ξ, H, and F be as in Theorem* 3.2; *in particular, for some ϕ and each $x \in \Xi$, $F(x, \cdot)$ is assumed Lipschitzian on* H *with modulus ϕ. Assume that F has partial B-derivatives with respect to x and to y at (x_0, y_0), and that*:

(a) *$D_x F(x_0, y_0)$ is strong.*

(b) *$D_x F(x_0, y_0)(\Xi - x_0)$ is a neighborhood of the origin in Z.*

(c) $\delta(D_x F(x_0, y_0), \Xi - x_0) =: d_0 > 0$.

The for each $\lambda > d_0^{-1}\phi$ there are neighborhoods U of x_0 and V of y_0, and a function $x: V \to U$ satisfying conclusions (i), (ii), *and* (iii) *of Theorem* 3.2. *Further, the function*

x is B-differentiable at y_0 with

$$Dx(y_0)(k) = D_xF(x_0, y_0)^{-1}[-D_yF(x_0, y_0)(k)]. \tag{3.8}$$

PROOF. We set $f(x) = D_xF(x_0, y_0)(x - x_0)$. The assumption that the partial B-derivative in x is strong means that $f \approx_x F$ at (x_0, y_0). Therefore conclusions (i), (ii), and (iii) of Theorem 3.2 follow immediately. Next, if we set $g = D_yF(x_0, y_0)$ we see that the definition of B-derivative implies that $g \sim_y F$ at (x_0, y_0). By conclusion (a) of Theorem 3.3, we have $x_f \sim x$ at y_0, where x_f is defined by

$$D_xF(x_0, y_0)[x_f(y) - x_0] + D_y(x_0, y_0)(y - y_0) = 0. \tag{3.9}$$

Note that (3.9) defines $x_f(y)$ for all y, since $D_xF(x_0, y_0)$ is assumed one-to-one from a neighborhood of the origin in X onto a neighborhood of the origin in Z, and it is positively homogeneous. Therefore we have

$$x_f(y) - x_0 = D_xF(x_0, y_0)^{-1}[-D_yF(x_0, y_0)(y - y_0)]. \tag{3.10}$$

If we write $y = y_0 + k$ and define $z(k)$ to be the right-hand side of (3.8), we find from (3.10) that

$$x_f(y_0 + k) - x_0 = z(k). \tag{3.11}$$

The function z is positively homogeneous. Further, since $x_f \sim x$ at y_0, (3.11) implies that $x(y_0) + z(y - y_0) \sim x$ at y_0. Therefore x is B-differentiable at y_0 with

$$Dx(y_0)(k) = z(k) = D_xF(x_0, y_0)^{-1}[-D_yF(x_0, y_0)(k)],$$

as required. ■

4. Application to generalized equations. In this section we show how to apply the theory developed in §3 to parametric generalized equations. For this purpose, we suppose throughout this section that X is a Hilbert space containing a closed convex set C, and that Ω is an open subset of X meeting C. We denote by H a neighborhood of a point y_0 in a normed linear space Y, and we let F be a function from $\Omega \times \mathrm{H}$ to X.

Now for fixed $y \in \mathrm{H}$ we pose the problem of finding $u_0 \in C \cap \Omega$ such that

$$\langle c - u_0, F(u_0, y)\rangle \geqslant 0 \quad \text{for each } c \in C, \tag{4.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of X. If we introduce the *normal cone* [38],

$$N_C(u_0) = \begin{cases} \{x \in X | \langle c - u_0, x\rangle \leqslant 0 \text{ for each } c \in C\}, & \text{if } u_0 \in C, \\ \varnothing, & \text{if } u_0 \notin C, \end{cases}$$

then we can rewrite (4.1) as

$$0 \in F(u_0, y) + N_C(u_0); \tag{4.2}$$

because of the form of (4.2) these problems are called generalized equations. Since a solution of (4.1) or (4.2) will in general be a function of y, one might ask how that function would behave under various assumptions on f and C. Indeed, an implicit-function theorem for such problems was established in [34] as a forerunner of the more general formulation in §3 of this paper. Here we show how to reformulate (4.1) so that the theory of §3 can be applied to it, and in the process we illustrate how the composition formula of Proposition 2.3 can be applied.

Although the development in [34] proceeded in terms of multifunctions, it seems desirable if possible to analyze (4.1) using ordinary functional methods, since single-valued functions are more familiar and easier to handle than are multifunctions. There is a well-known way to do this; to illustrate it we introduce the *projector* on C, namely the function Π_C that takes any point $x \in X$ to the point $\Pi_C(x)$ that is closest to x in C. It is easy to show that such a closest point exists and is unique, and that Π_C is a contraction: that is, it is Lipschitzian with modulus 1. For this and numerous other results about projectors, see [43].

Now let $W = X \times X$, and let $g\colon X \to W$ be defined by

$$g(x) = (\Pi_C(x), x - \Pi_C(x)). \tag{4.3}$$

This g is sometimes called the "Minty map"; it is a Lipschitzian homeomorphism of X onto the graph of N_C (see [3]). If we define G; $\Omega \times X \times \mathrm{H} \to X$ by

$$G(u, v, y) = F(u, y) + v, \tag{4.4}$$

then solving (4.2) is equivalent to finding a solution x of

$$G(g(x), y) = 0, \tag{4.5}$$

since for any such x the point $\Pi_C(x)$ solves (4.2), whereas if u_0 satisfies (4.2) then with $x = u_0 - F(u_0, y)$ we have $g(x) = (u_0, -F(u_0, y))$, and so this x satisfies (4.5). In this way we have reduced the generalized equation (4.2) to the problem of finding a zero of the single-valued, continuous function given in (4.5).

The theory of §3 now provides a tool to establish the existence and investigate the behavior of parametric solutions of (4.2). If we wish to apply that theory, we need to verify the principal assumption of Theorem 3.2. Supposing that u_0 is a solution of (4.2) for $y = y_0$ and that we wish to investigate parametric solutions $u(y)$ for y near y_0, then we see that we need to produce a function strongly approximating $G(g(x), y)$ at (x_0, y_0), where $x_0 = u_0 - F(u_0, y_0)$. It turns out that if $F(u, y)$ has a strong partial B-derivative in u at $(\Pi_C(x_0), y_0)$, then the composition of that derivative with g gives us the required approximating function. We prove this in the following proposition.

PROPOSITION 4.1. *Let x_0 be a point of X such that $g(x_0) = (u_0, x_0 - u_0)$ with $u_0 \in \Omega$. Let y_0 be a point of Y, and suppose that F has a strong partial B-derivative $D_uF(u_0, y_0)$ with respect to u at (u_0, y_0). Then*

$$\begin{aligned} D_uF(u_0, y_0)[\Pi_C(x) - u_0] + [(x - x_0) - (\Pi_C(x) - u_0)]& \\ \approx_x G(g(x), y) \quad \text{at } (x_0, y_0),& \end{aligned} \tag{4.6}$$

where G is defined by (4.4).

PROOF. For any $v' \in X$,

$$D_uG(u_0, v', y_0) = D_uF(u_0, y_0),$$

and each of these partial B-derivatives is strong (that on the right by hypothesis, and that on the left because $G(\cdot, v', y_0)$ is a translate of $F(\cdot, y_0)$). This, together with the definition of strong approximation and the fact that v enters linearly in the definition of G, yields

$$D_u F(u_0, y_0)(u - u_0) + (v - v_0) \approx_{(u,v)} G \quad \text{at } (u_0, v_0, y_0).$$

Now applying Proposition 2.5, we conclude that (4.6) holds. ■

It is worth pointing out that the left side of (4.6) can be regarded as a linearization of the generalized equation (4.2). Specifically, let us define

$$(4.7) \quad f(x) = D_u F(u_0, y_0)\left[\Pi_C(x) - u_0\right] + \left[(x - x_0) - (\Pi_C(x) - u_0)\right],$$

and suppose that $G(u_0, v_0, y_0) = 0$: that is,

$$F(\Pi_C(x_0), y_0) + (x_0 - \Pi_C(x_0)) = 0.$$

Then $f(x) = 0$ if and only if, for each $c \in C$,

$$(4.8) \qquad \langle c - \Pi_C(x), F(u_0, y_0) + D_u F(u_0, y_0)(\Pi_C(x) - u_0)\rangle \geqslant 0.$$

Therefore setting the left side of (4.6) to zero just yields the nonsmooth equation equivalent to the (B-)linearization of (4.1) in u at (u_0, y_0).

Having (4.6), we can now apply the main implicit-function theorem to obtain the following theorem:

THEOREM 4.2. *Assume the notation and hypotheses of Proposition* 4.1. *Suppose also that F is continuous on $\Omega \times \mathrm{H}$, and that Ξ is a neighborhood of x_0 with $\Pi_C(\Xi) \subset \Omega$. Assume the following*:

(a) $F(\Pi_C(x_0), y_0) + (x_0 - \Pi_C(x_0)) = 0$.

(b) *For each $x \in \Xi$, $F(\Pi_C(x), \cdot\,)$ is Lipschitzian on* H *with modulus ϕ.*

(c) *$f(\Xi)$ is a neighborhood of the origin in X, where f is the function on X defined by* (4.7).

(d) $\delta(f, \Xi) =: d_0 > 0$.

Then for any $\lambda > d_0^{-1}\phi$ there exist neighborhoods N of x_0, U of $\Pi_C(x_0)$, *and Y of y_0, and a function $x\colon Y \to N$, such that*:

(1) $x(y_0) = x_0$.

(2) *$x(\cdot)$ is Lipschitzian with modulus λ.*

(3) *For each $y \in Y$, $\Pi_C(x(y))$ is the unique point u in $U \cap C$ such that for each $c \in C$,*

$$\langle c - u, F(u, y)\rangle \geqslant 0.$$

PROOF. We shall apply Theorem 3.2 to the composed function $G(g(x), y)$. The first of the four main hypotheses of Theorem 3.2 is satisfied here because of (4.6). The second also holds because hypothesis (b) of the present theorem guarantees that for each $x \in \Xi$, $G(g(x), y)$ will be Lipschitzian in y on H with modulus ϕ. The third and fourth hypotheses of Theorem 3.2 appear as hypotheses (c) and (d) here. We conclude that there exist neighborhoods Y_0 of y_0 and N of x_0, and a function x: $y_0 \to N$ satisfying conclusions (1) and (2), and such that for each $y \in Y_0$, $x(y)$ is the unique point of N such that $G(g(x), y) = 0$. In particular, this tells us that for each $c \in C$,

$$\langle c - \Pi_C(x(y)), F(\Pi_C(x(y)), y)\rangle \geqslant 0.$$

Now observe that g carries X onto the graph of N_C, and its inverse is the restriction of the addition operator to that graph: that is, the operator that takes (u, v) into $u + v$. We know that $u_0 = \Pi_C(x_0)$ and $v_0 = x_0 - u_0$; then because g is a homeomorphism, $g(N)$ is a neighborhood of (u_0, v_0) in the graph of N_C. It follows that there are neighborhoods U_0 of u_0 in X and V of v_0 in X, such that if $u \in U_0$, $v \in V$, and (u, v) belongs to the graph of N_C, then $u + v \in N$.

Further, since we assumed F continuous, and since $v_0 = -F(u_0, y_0)$ by hypothesis (a), we find that there are neighborhoods $U \subset U_0$ and $Y \subset Y_0$ such that whenever $(u, y) \in U \times Y$ then $-F(u, y) \in V$.

Now suppose that $y \in Y$, $u \in U \cap C$, and for each $c \in C$,

$$\langle c - u, F(u, y)\rangle \geqslant 0.$$

It follows that the point $v := -F(u, y)$ belongs to $V \cap N_C(u)$, and therefore the point $x := u + v$ belongs to N. It is also clear that $g(x) = (u, v)$, and the definition of v implies that $G(g(x), y) = 0$. The uniqueness conclusion of Theorem 3.2 then tells us that $x = x(y)$. But then $u = \Pi_C(x) = \Pi_C(x(y))$, and the theorem is proved. ∎

In general, (4.8) may be simpler to solve than (4.1); in particular, when $D_u F(u_0, y_0)$ is actually linear (i.e., a Fréchet derivative), then (4.8) is a *linear* variational inequality over C, contrasted to the *nonlinear* variational problem (4.1). In that case, we recover the type of result proved in [34, Theorem 2.1] under the more restrictive assumption that the derivative with respect to x was a continuous Fréchet derivative. (*Note*: In [34, Theorem 2.1] it is not explicitly stated that the space X must be complete; however, this assumption is necessary, as it was in Theorem 3.2 here, since the contraction mapping principle is used.)

In fact, this situation also provides an example of a case in which the Lipschitzian composition formulation is needed. The reason for this is that, even in finite-dimensional spaces X, the projector Π_C may not be B-differentiable. For a clever example constructed in $\mathbf{R}^3$, see [20]. Therefore, for general closed convex sets C, we cannot expect to use the B-derivative in x of the composed function $G(g(x), y)$ appearing in the proof of Theorem 4.2. The use of the composition therefore enables us to obtain a result that we could not have found directly.

There are, however, important special cases in which a B-derivative of Π_C exists. For example, Haraux [10] studied special convex sets in Hilbert space, which he called "polyhedric" sets, and showed how to compute the directional derivative in that case. When attention is further restricted to polyhedral convex sets in $\mathbf{R}^n$, then the directional derivative has a particularly simple form, as pointed out for example by Pang in [27, Lemma 5(i)]. Pang refers for the proof to a paper [36] of the present author. Since [36] was never published, we give a proof of the result here for the sake of completeness.

To state the result, we begin with a polyhedral convex set $C \subset \mathbf{R}^n$ and a point $z \in \mathbf{R}^n$, and we let $x = \Pi_C(z)$. We are interested in determining the form of $\Pi_C(z + h)$ for small h. Note that by the definition of x we have $(x, z - x) \in N_C$: that is, $z - x \in N_C(x)$. Therefore the set

$$F = \{c \in C | \langle z - x, c - x\rangle = 0\}$$

is a face of C containing x. It is the set of points at which the maximum on C of the linear functional $\langle z - x, \cdot\rangle$ is achieved, and we call it the *critical face* of C corresponding to $z - x$. Likewise, the tangent cone to F at x, which we denote by K, is called the *critical cone* of C corresponding to z and $z - x$; it is easy to show that K is also the intersection of the tangent cone to C at x with the subspace orthogonal to

$z - x$. The following fundamental lemma describes the properties of F that we shall need here.

LEMMA 4.3 [35, Lemma 3.5]. *Let C be a polyhedral convex set in $\mathbf{R}^n$, let $y_0 \in \mathbf{R}^n$, and let F be the critical face of C corresponding to y_0. There is a neighborhood U of y_0 such that if $y \in U$ then the critical face of C corresponding to y is the critical face of F corresponding to y.*

We use this result to prove the following proposition, from which in turn we obtain a corollary containing the result we are seeking.

PROPOSITION 4.4. *Let C be a polyhedral convex set in $\mathbf{R}^n$, and let $(x, y) \in N_C$. Let K be the critical cone of C corresponding to (x, y). There is a neighborhood V of the origin in $\mathbf{R}^n \times \mathbf{R}^n$ such that*

$$[(x, y) + V] \cap N_C = (x, y) + [V \cap N_K].$$

PROOF. Let F be the critical face of C corresponding to y. By Lemma 4.3, there is a neighborhood Q of the origin in $\mathbf{R}^n$ such that for $k \in Q$ the critical face of C corresponding to $y + k$ is the critical face of F corresponding to $y + k$. Further, since F is polyhedral and K is its tangent cone at x, there is an open neighborhood P of the origin in $\mathbf{R}^n$ such that $(x + P) \cap F = x + (P \cap K)$. Note that since the normal cone is a local construct, if $p \in P$ then $N_F(x + p) = N_K(p)$. Now let $V = P \times Q$ and let $(h, k) \in V$.

To say $(x + h, y + k)$ belongs to N_C is to say that $x + h$ belongs to the critical face of C corresponding to $y + k$. By Lemma 4.3, this is the same as saying that $x + h$ belongs to the critical face of F corresponding to $y + k$, or in other words that $(x + h, y + k) \in N_F$. In turn, since $h \in P$, this is equivalent to $(h, y + k) \in N_K$. However, since y is orthogonal to the affine hull of K, it lies in the lineality space of $N_K(h)$, and therefore $(h, y + k) \in N_K$ is equivalent to $(h, k) \in N_K$. ■

Here is the result we set out to prove.

COROLLARY 4.5 [27, Lemma 5(i)]. *Let C be a polyhedral convex set in $\mathbf{R}^n$, let $z \in \mathbf{R}^n$, and let $x = \Pi_C(z)$. Let K be the critical cone of C corresponding to x and $z - x$. There is a neighborhood Q of the origin in $\mathbf{R}^n$ such that for each $q \in Q$,*

$$\Pi_C(z + q) = x + \Pi_K(q).$$

PROOF. Write $y = z - x$; for any q define h by $x + h = \Pi_C(z + q)$. We need to prove that for small q, $h = \Pi_K(q)$.

Observe that since for any convex set S the operator Π_S is $(I + N_S)^{-1}$, the statement $x + h = \Pi_C(z + q)$ is equivalent to $(x + h, y + (q - h)) \in N_C$. Further, since the projector is a contraction, there is a neighborhood Q of the origin in $\mathbf{R}^n$ such that for each $q \in Q$ we have $(h, q - h) \in V$, where V is the neighborhood in $\mathbf{R}^n \times \mathbf{R}^n$ provided by Proposition 4.4. Assume that q has been chosen in Q.

By Proposition 4.4, the statement $(x + h, y + (q - h)) \in N_C$ is equivalent to $(h, q - h) \in N_K$. In turn, as noted above this is equivalent to $(q, h) \in \Pi_K$; that is, $h = \Pi_K(q)$. ■

As Pang [27] noted, Corollary 4.5 immediately implies the weaker result

$$D\Pi_C(z) = \Pi_K,$$

so that not only is the projector (strongly) B-differentiable, but its B-derivative also has a particularly simple and convenient form.

References

[1] Aubin, J. P. (1987). Graphical Convergence of Set-Valued Maps. Working Paper WP-87-83, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[2] ______. (1987). Differential Calculus of Set-Valued Maps. An Update. Working Paper WP-87-93, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[3] Brézis, H. (1973). *Opérateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*. North Holland Mathematics Studies No. 5, North Holland, Amsterdam, The Netherlands.

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

[5] Cornet, B. and Laroque, G. (1987). Lipschitz Properties of Solutions in Mathematical Programming. *J. Math. Anal. Appl.* **53**, 407–427.

[6] Dafermos, S. (1988). Sensitivity Analysis in Variational Inequalities. *Math. Oper. Res.* **13**, 421–434.

[7] Frankowska, H. (1987). An Open Mapping Principle for Set-Valued Maps. *J. Math. Anal. Appl.* **127**, 172–180.

[8] ______. (1988). High Order Inverse Function Theorems, Working Paper WP-88-018, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[9] ______. (1988). Some Inverse Mapping Theorems. Preprint S.I.S.S.A. 113M, Scuola Internazionale di Studi Avanzati, Trieste, Italy.

[10] Haraux, A. (1977). How to Differentiate the Projection on a Convex Set in Hilbert Space. Some Applications to Variational Inequalities. *J. Math. Soc. Japan* **29**, 615–631.

[11] Ioffe, A. D. (1979). Différentielles Généralisées d'Applications Localement Lipschitziennes d'un Espace de Banach dans un Autre. *C. R. Acad. Sci. Paris Sér. A* **289**, 637–640.

[12] ______. (1981). Nonsmooth Analysis: Differential Calculus of Nondifferentiable Mappings. *Trans. Amer. Math. Soc.* **266**, 1–56.

[13] ______. (1987). Global Surjection and Global Inverse Mapping Theorems in Banach Spaces. *Ann. New York Acad. Sci.* **491**, 181–189.

[14] Jittorntrum, K. (1978), An Implicit Function Theorem. *J. Optim. Theory Appl.* **25**, 575–577.

[15] ______. (1984). Solution Point Differentiability Without Strict Complementarity in Nonlinear Programming. *Math. Programming Stud.* **21**, 127–138.

[16] Jongen, H. Th., Klatte, D., and Tammer, K. (1990). Implicit Functions and Sensitivity of Stationary Points. *Math. Programming* **49**, 123–138.

[17] Kantorovich, L. V. and Akilov, G. P. (1964). *Functional Analysis in Normed Spaces*. Macmillan, New York.

[18] King, A. J. and Rockafellar, R. T. (1989). Sensitivity Analysis for Nonsmooth Generalized Equations. Research Report RC 14639 (#65641), IBM Research Division, Yorktown Heights, NY.

[19] Kojima, M. (1980). Strongly Stable Solutions in Nonlinear Programs. In S. M. Robinson (Ed.), *Analysis and Computation of Fixed Points*, 93-138. Academic Press, New York.

[20] Kruskal, J. B. (1969). Two Convex Counterexamples: A Discontinuous Envelope Function and a Nondifferentiable Nearest-Point Mapping. *Proc. Amer. Math. Soc.* **23**, 697–703.

[21] Kummer, B. (1988). The Inverse of a Lipschitz Function in $\mathbf{R}^n$: Complete Characterization by Directional Derivatives. Preprint Nr. 195 (Neue Folge), Sektion Mathematik, Humboldt-Universität zu Berlin, Berlin, DDR.

[22] ______. (1989). An Implicit-Function Theorem for $C^{0,1}$-Equations and Parametric $C^{1,1}$-Optimization. Preprint, Sektion Mathematik, Humboldt-Universität zu Berlin, Berlin, DDR.

[23] ______. (1989). Some Pathological Lipschitz Functions. Preprint, Sektion Mathematik, Humboldt-Universität zu Berlin, Berlin, DDR.

[24] Kyparisis, J. (1986). Uniqueness and Differentiability of Solutions of Parametric Nonlinear Complementarity Problems. *Math. Programming* **36**, 105–113.

[25] ______. (1987). Sensitivity Analysis Framework for Variational Inequalities. *Math. Programming* **38**, 203–213.

[26] ______. (1988). Perturbed Solutions of Variational Inequality Problems Over Polyhedral Sets. *J. Optim. Theory Appl.* **57**, 295–305.

[27] Pang, J. S. (1990). Newton's Method for B-Differentiable Equations. *Math. Oper. Res.* **15**, 311–341.

[28] ______. (1990). Solution Differentiability and Continuation of Newton's Method for Variational Inequality Problems Over Polyhedral Sets. *J. Optim. Theory Appl.* **66**, 121–135.

[29] Park, K. (1989). Continuation Methods for Nonlinear Programming. Ph.D. Dissertation, Department of Industrial Engineering, University of Wisconsin-Madison.

[30] Potra, F. A. and Pták, V. (1984). *Nondiscrete Induction and Iterative Processes*, Pitman, Boston.

[31] Pshenichny, B. N. (1986). Implicit Function Theorems for Multi-Valued Mappings. Working Paper WP-86-45, International Institute for Applied Systems Analysis, Laxenburg, Austria.

[32] Qiu, Y. and Magnanti, T. L. (1987). Sensitivity Analysis for Variational Inequalities. Working Paper OR 163-87, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA.

[33] Robinson, S. M. (1976). Regularity and Stability for Convex Multivalued Functions. *Math. Oper. Res.* **1**, 130–143.
[34] ______. (1980). Strongly Regular Generalized Equations. *Math. Oper. Res.* **5**, 43–62.
[35] ______. (1984). Local Structure of Feasible Sets in Nonlinear Programming, Part II: Nondegeneracy. *Math. Programming Stud.* **22**, 217–230.
[36] ______. (1985). Implicit B-Differentiability in Generalized Equations. Technical Summary Report No. 2854, Mathematics Research Center, University of Wisconsin-Madison, Madison, WI.
[37] ______. (1987). Local Structure of Feasible Sets in Nonlinear Programming, Part III: Stability and Sensitivity. *Math. Programming Stud.* **30**, 45–66. Corrigenda, *Math. Programming* **49** (1990), 143.
[38] Rockafellar, R. T. (1970). *Convex Analysis*, Princeton Univ. Press, Princeton, NJ.
[39] ______. (1989). Perturbation of Generalized Kuhn-Tucker Points in Finite-Dimensional Optimization. In F. H. Clarke et al. (Eds.), *Nonsmooth Analysis and Related Topics*, Plenum Press, 393–402.
[40] ______. (1989). Nonsmooth Analysis and Parametric Optimization. Preprint, Department of Mathematics, University of Washington, Seattle, WA.
[41] Shapiro, A. (1990). On Concepts of Directional Differentiability. *J. Math. Anal. Appl.* **66**, 477–487.
[42] Ursescu, C. (1975). Multifunctions with Closed Convex Graph. *Czechoslovak Math. J.* **25**, 438–441.
[43] Zarantonello, E. H. (1971). Projections on Convex Sets in Hibert Space and Spectral Theory. In E. H. Zarantonello (Ed.), *Contributions to Nonlinear Functional Analysis*, Academic Press, New York, 237–424.

DEPARTMENT OF INDUSTRIAL ENGINEERING, UNIVERSITY OF WISCONSIN, MADISON, WISCONSIN 53706-1572