International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

An Implicit-Function Theorem for B-Differentiable Functions

Robinson, S.M.

IIASA Working Paper

WP-88-067

July 1988

WORKING PAPER

AN IMPLICIT-FUNCTION THEOREM FOR B-DIFFERENTIABLE FUNCTIONS

Stephen M. Robinson

July 1988
WP-88-67

IIASA
International Institute
for Applied Systems Analysis

AN IMPLICIT-FUNCTION THEOREM FOR B-DIFFERENTIABLE FUNCTIONS

Stephen M. Robinson

July 1988
WP-88-67

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

A function from one normed linear space to another is said to be Bouligand differentiable (B-differentiable) at a point if it is directionally differentiable there in every direction, and if the directional derivative has a certain uniformity property. This is a weakening of the classical idea of Fréchet (F-) differentiability, and it is useful in dealing with optimization problems and in other situations in which F-differentiability may be too strong.

In this paper a concept of strong B-derivative is introduced and this idea is employed to prove an implicit-function theorem for B-differentiable functions. This theorem provides the same kinds of information as does the classical implicit-function theorem, but with B-differentiability in place of F-differentiability. Therefore it is applicable to a considerably wider class of functions than is the classical theorem.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

CONTENTS

An Implicit-Function Theorem for B-Differentiable Functions

Stephen M. Robinson
University of Wisconsin-Madison †

Abstract

A function from one normed linear space to another is said to be Bouligand differentiable (B-differentiable) at a point if it is directionally differentiable there in every direction, and if the directional derivative has a certain uniformity property. This is a weakening of the classical idea of Fréchet (F-) differentiability, and it is useful in dealing with optimization problems and in other situations in which F-differentiability may be too strong.

In this paper we introduce a concept of strong B-derivative, and we employ this idea to prove an implicit-function theorem for B-differentiable functions. This theorem provides the same kinds of information as does the classical implicit-function theorem, but with B-differentiability in place of F-differentiability. Therefore it is applicable to a considerably wider class of functions than is the classical theorem.

AMS(MOS) Subject Classifications: 47H15, 49A27, 90C30

IAOR Subject Classification: Programming: Nonlinear

OR/MS Subject Classifications: 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

---

† Address: Department of Industrial Engineering, University of Wisconsin–Madison, 1513 University Avenue, Madison, WI 53706. 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 Grant No. AFOSR-88-0090. The US Government has certain rights in this material, and is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

1. Introduction

This paper develops an implicit-function theorem for certain functions which, although not differentiable in the conventional sense, still obey a weak kind of differentiability called Bouligand differentiability (B-differentiability). 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 above because in this case the derivative involved is not the usual partial derivative, but rather a partial B-derivative.

Numerous 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.

Recently, Frankowska studied inverse-function theorems using hypotheses on higher-order variations of the function in questions, 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]; 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.* [11] recently proved an implicit-function theorem for nonsmooth functions in $\mathbf{R}^n$ under the basic assumption that the generalized Jacobian was nonsingular, They then applied this result to the problem of stability in nonlinear programming. The basic device in the proof is again the inverse function theorem of Clarke.

In [21], 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 [28] by Ursescu and in [23] by Robinson.

A considerable amount of work in this area has been focused on more specific classes of problems, particularly those involving variational inequalities (generalized equations) and their application to solution of optimization problems. An implicit- function type theorem for generalized equations was established in [24], and was employed there to obtain sensitivity results about nonlinear programming problems. Kojima [13] obtained similar sensitivity information using very different techniques. Later, Jittorntrum [10] showed that optimal solutions of nonlinear programming problems were directionally differentiable under suitable assumptions. Then, in [25] and [26], 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 [15,16,17] Kyparisis has extended the above work. In [16] he showed how to extend the type of result proved in [10] and [24] to variational problems over sets defined by systems of inequalities and equations, while in [15] he developed differentiability properties of a specific type of variational problem (the nonlinear complementarity problem). Finally, in [17] he dealt with the case of variational problems over polyhedral convex sets, and showed how the results of [23] and [24] 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 [22], 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 very recent papers of Pang [18,19]. In [18], 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 [18], 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 [18].

In [19], 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 solutions. A similar continuation method has been proposed, and is currently being investigated, by Park [20].

The rest of this paper is organized in three sections: in §2 we briefly review B-derivatives, and introduce the new concept of 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, the implicit-function theorem, and a theorem on B-differentiability of the implicit function. We also show that the hypotheses of the latter theorem cannot, in general, be improved. Finally, in §4 we sketch an application of this theory to parametric solutions of variational inequalities (generalized equations).

2. B-derivatives.

We provide here a definition and a brief summary of some properties of B-derivatives, and we also introduce some new material that will be used in what follows.

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

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

This function $Df(x_0)$ is necessarily unique if it exists. In the case in which f is locally Lipschitzian at x_0, which applies to all the results of this paper, we have easily that $Df(x_0)$ inherits the Lipschitz modulus of f, that the chain rule works for B-derivatives, and that B-differentiation is distributive with respect to addition and scalar multiplication.

The B-derivative terminology was introduced in [26], and the results cited were established in the finite-dimensional case. Subsequently, Shapiro [27] 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. For example, it can be used to show easily that if a locally Lipschitz function on a finite-dimensional space is regular in the sense of Clarke [4, p. 39], then it is B-differentiable.

In all of the above properties the B-derivative resembles 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, $D_x f(x_0, y_0)$ to be the B-derivative at x_0 of $f(\cdot, y_0)$), we do not obtain the addition formula: in general,

$$Df(x_0, y_0)(h, k) \neq D_x f(x_0, y_0)h + D_y f(x_0, y_0)k. \tag{2.2}$$

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. It uses a special class of Lipschitzian functions; in general, we shall denote by $Lip(\lambda)$ the functions (on whatever spaces are under discussion) that are Lipschitzian with modulus λ, and by $Flip(x_0)$ (Flat Lipschitzian) the linear space of functions ϕ having the following property: $\phi(x_0) = 0$ and for each $\epsilon > 0$ there is a neighborhood U of x_0 such that ϕ is $Lip(\epsilon)$ on U.

Definition 2.1. *Let $f : X \times Y \to Z$, and suppose f has a partial B-derivative $D_x f(x_0, y_0)$ with respect to x at (x_0, y_0). We say $D_x f(x_0, y_0)$ is* **strong** *if for each $\epsilon > 0$ there are neighborhoods U of the origin in X and V of y_0 such that for each $y \in V$ the function (of h) $f(x_0 + h, y) - f(x_0, y) - D_x f(x_0, y_0)h$ is $Lip(\epsilon)$ on U.*

Roughly speaking, this definition says that the partial B-derivative is strong if the "remainder" function belongs to $Flip(x_0)$, uniformly for values of y near y_0. Of course, the definition applies also in the case in which the space Y is vacuous, in which case we refer to a "strong B-derivative," without using the word "partial."

In [18], Pang also introduced a definition of strong B-derivative, but his definition differs from the above in the case in which $D_x f(x_0, y_0)$ is not a linear operator. He showed that a strong B-derivative under his definition must in fact be a strong F-derivative, so his requirements are considerably stronger than those of Definition 2.1, 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

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

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 2.2. *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

$$\begin{aligned} f(x_0 + h, y_0 + k) - f(x_0, y_0) - D_x h - D_y k & \\ &= [f(x_0 + h, y_0 + k) - f(x_0, y_0 + k) - D_x h] \\ &+ [f(x_0, y_0 + k) - f(x_0, y_0) - D_y k] \\ &= o(h) + o(k) = o(\|h\| + \|k\|). \end{aligned}$$

This is enough to prove (a). Now if both partial derivatives are strong, then we have

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

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. The Implicit-Function Theorem.

This section contains the main results of the paper. We begin with a lemma that contains the essential perturbation information necessary to establish the implicit-function theorem, and we follow that with the theorem itself.

The following lemma can be regarded as an extension of the well known Banach perturbation lemma [12, 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 the lemma, we use the notation $B(x,\xi)$ for the closed ball about x of radius ξ, either in a normed space or in a more general metric space. We also use some shorthand notation pertaining to functions and multifunctions: $f|S$ denotes the restriction of f to S, while $f \cap T$ denotes the function or multifunction whose value at x is $f(x) \cap T$.

Lemma 3.1. *Let (X,ρ) be a complete metric space and Y be a normed linear space. Let F be a multifunction from X to Y with $y_0 \in F(x_0)$, and let h be a function from X into Y. Assume that:*

a. $F^{-1}|B(y_0,\alpha) \in Lip(\delta)$.

b. $h \in Lip(\eta)$ *on* $B(x_0,\alpha\delta)$.

c. $\theta := \alpha(1-\delta\eta) - \|h(x_0)\| > 0$.

Then $[(F+h)^{-1}|B(y_0,\theta)] \cap B(x_0,\alpha\delta)$ is a Lipschitzian function with modulus $\delta/(1-\delta\eta)$.

Proof: For $y \in B(y_0,\theta)$ and $x \in B(x_0,\alpha\delta)$ define $\Phi_y(x) = F^{-1}[y-h(x)]$. Note that for such x and y,

$$\begin{aligned}\|y-h(x)-y_0\| &\le \|h(x_0)\| + \|y-y_0\| + \|h(x)-h(x_0)\|\\ &\le \alpha(1-\delta\eta) + \eta(\alpha\delta) = \alpha,\end{aligned}$$

so that Φ_y is single-valued on $B(x_0,\alpha\delta)$. Also,

$$\rho[\Phi_y(x_1),\Phi_y(x_2)] \le \delta\|h(x_1)-h(x_2)\| \le \delta\eta\rho[x_1,x_2],$$

so that Φ_y is a strong contraction with modulus $\delta\eta < 1$. Further, we have

$$\begin{aligned}\rho[x_0,\Phi_y(x_0)] &= \rho[F^{-1}(y_0),F^{-1}(y-h(x_0))]\\ &\le \delta\|y_0-y-h(x_0)\| \le \alpha\delta(1-\delta\eta),\end{aligned}$$

so that Φ_y carries $B(x_0,\alpha\delta)$ into itself.

By the contraction mapping theorem [12, Th. 1(1.XVI)], Φ_y has a fixed point $x(y)$ that is unique in the ball $B(x_0,\alpha\delta)$. Because of the way in which Φ_y was defined, this $x(y)$ must satisfy $(F+h)x(y) = y$. Now let y_1 and y_2 be two points in $B(y_0,\theta)$, and let $x_i = x(y_i)$ for $i = 1,2$. Then

$$\begin{aligned}\rho[x_1,x_2] &= \rho[F^{-1}(y_1-h(x_1)),F^{-1}(y_2-h(x_2))]\\ &\le \delta\|y_1-y_2\| + \delta\eta\rho[x_1,x_2].\end{aligned}$$

It follows that

$$\rho[x_1, x_2] \leq \delta(1 - \delta\eta)^{-1}\|y_1 - y_2\|,$$

and therefore $x(y)$ is single-valued and Lipschitzian with modulus $\delta/(1 - \delta\eta)$. ∎

The next result is our main implicit-function theorem. It uses the perturbation lemma together with the notion of strong B-differentiability discussed in §2. For simplicity we let the space Y be a normed linear space, although a result could be derived also for the case in which Y is just a topological space. However, in that case we cannot discuss the notions of Lipschitz modulus or of B-differentiability for the implicit function, and the theorem loses some strength. As before, we use the sum of the norms on X and Y as the norm on $X \times Y$.

The theorem may look somewhat strange because it involves the composition of the function $f(\cdot, y)$ with another Lipschitzian function g, whereas no such composition is involved in the classical implicit-function theorem. Further, there are no statements about B-differentiability of the implicit function, whereas one would expect to be able to make such statements. There are good reasons for these differences: we shall illustrate in §4 a substantial application area in which the composition formulation is essential. Further, if we set the inner function in the composition equal to the identity, then we recover a result much like the conventional implicit-function theorem, and we give later in this section a theorem on B-differentiability of the implicit function in that case.

Finally, although we do not give results about B-differentiability in the present case, the theorem does contain an approximation result showing that a simpler function (denoted by x_L) approximates the implicit function to within the small order of $y - y_0$. We show in §4 how this approximation may be useful even when statements about B-differentiability cannot be made.

Theorem 3.2. *Let X be a Banach space and let W, Y and Z be normed linear spaces. Let g be a Lipschitzian function with modulus γ from a neighborhood Γ of $x_0 \in X$ to W, with $g(x_0) = w_0$, and let f be a Lipschitzian function with modulus ϕ from a neighborhood Ω of $(w_0, y_0) \in W \times Y$ to Z, satisfying:*

a. *$f(g(x_0), y_0) = 0$.*

b. *f has a partial B-derivative $D_w f(w_0, y_0)(\cdot)$ with respect to w that is strong at (w_0, y_0).*

c. *The function $L(x) = D_w(w_0, y_0)[g(x) - g(x_0)]$ has an inverse that is locally Lipschitzian at the origin in Z with modulus δ.*

Then for each $\xi > \phi\delta$ there exist neighborhoods V of x_0 and U of y_0, and a Lipschitzian function $x : U \to V$, such that:

a. *$x(y_0) = x_0$;*

b. *For each $y \in U$, $x(y)$ is the unique solution in V of $f(g(x), y) = 0$;*

c. *$x \in Lip(\xi)$;*

d. *For each $y \in U$, the equation*

$$f(w_0, y) + L(x) = 0$$

has a solution $x_L(y)$ that is unique in V, and $x(y) = x_L(y) + o(y - y_0)$.

Proof: For simplicity, throughout the proof we write D_w in place of $D_w(w_0, y_0)$.

Let $\xi > \phi\delta$, and let ϵ satisfy $0 < \epsilon < \phi\gamma^{-1}[(\phi\delta)^{-1} - \xi^{-1}]$. Note that we then have

$$\xi > \phi\delta/(1 - \gamma\delta\epsilon); \tag{3.1}$$

we shall need this estimate later.

Now choose neighborhoods Q of w_0 and U of y_0, and a positive scalar α, so that the following three requirements are satisfied:

a. Whenever $y \in U$ and w_1 and w_2 belong to Q, one has

$$\begin{aligned} \|[f(w_1, y) - f(w_0, y) - D_w(w_1 - w_0)] \\ - [f(w_2, y) - f(w_0, y) - D_w(w_2 - w_0)] \leq \epsilon\|w_1 - w_2\|. \end{aligned} \tag{3.2}$$

(This can be done because D_w is a strong B-derivative.)

b. For each $y \in U$, $B(0, \alpha) - f(w_0, y) \subset P$, where P is the neighborhood of the origin in Z on which L^{-1} is Lipschitzian.

c. If $x \in V$, then $x \in \Gamma$ and $g(x) \in Q$, where we define V to be $B(x_0, \alpha\delta)$.

For any given $y \in U$, define $F_y(x)$ for $x \in V$ by

$$F_y(x) = f(w_0, y) + L(x).$$

Note that $F_y^{-1}(z) = L^{-1}[z - f(w_0, y)]$, and this function is Lipschitzian for $z \in B(0, \alpha)$, with modulus δ. Next, for $x \in V$ and $y \in U$ let

$$h_y(x) = f(g(x), y) - F_y(x);$$

then for $y \in U$ and x_1 and x_2 in V we have by using (3.2),

$$\|h_y(x_1) - h_y(x_2)\| \leq \epsilon\gamma\|x_1 - x_2\|.$$

Therefore h_y is Lipschitzian on V with modulus $\epsilon\gamma$.

Keeping y arbitrary but fixed in U, apply Lemma 3.1 to F_y and h_y, with $\eta = \epsilon\gamma$. Since the sum of these two functions is $f(g(x), y)$, we find that there is an inverse function j_y defined on the ball about the origin in Z of radius $\alpha(1 - \delta\epsilon\gamma)$, that is Lipschitzian with modulus ξ/ϕ (from (3.1)), such that for any z in that ball we have $f(g(j_y(z)), y) = z$, and $x = j_y(z)$ is the only point in V that satisfies $f(g(x), y) = z$. Note that the domain of definition and the Lipschitz modulus of the inverse function j_y are independent of y.

Now for each $y \in U$, let $x(y) = j_y(0)$. This $x(y)$ belongs to V, and it is the only point in V that satisfies $f(g(x), y) = 0$. It is clear that $x(y_0) = x_0$, so we have proved (a) and (b). For (c), select two points of U, say y_1 and y_2, and let $x_i = x(y_i)$ for $i = 1, 2$. Then by definition

$$x_1 = j_{y_1}(0).$$

Also,

$$x_2 = j_{y_1}[f(g(x_2), y_1)] = j_{y_1}[f(g(x_2), y_1) - f(g(x_2), y_2)],$$

since we know $f(g(x_2), y_2) = 0$. Therefore

$$\|x_1 - x_2\| \leq (\xi/\phi)\phi\|y_1 - y_2\| = \xi\|y_1 - y_2\|,$$

and so $x(\cdot)$ is Lipschitzian with modulus ξ. Therefore (c) is proved, and we only have to show that the estimate in (d) holds.

First, note that $x_L(y) = F_y^{-1}(0)$, while $x(y) = F_y^{-1}(-h_y[x(y)])$. Therefore by the Lipschitzian property of F_y^{-1} we have $\|x(y) - x_L(y)\| \leq \delta\|h_y(x(y))\| = \delta\|h_y(x(y)) - h_y(x_0)\|$, where we used the fact that $h_y(x_0) = 0$. Now recall that (by (3.2)) the function h_y is $Flip(x_0)$, and $x(y)$ is Lipschitzian in y. If we choose any $\beta > 0$, by taking y close enough to y_0 we shall have

$$\|h_y(x(y)) - h_y(x_0)\| \leq \beta\|x(y) - x_0\| \leq \beta\xi\|y - y_0\|.$$

But then

$$\|x(y) - x_L(y)\| \leq \delta\beta\xi\|y - y_0\|,$$

which shows that $x(y) - x_L(y) = o(y - y_0)$, as required. ∎

Theorem 3.2 shows how to approximate the implicit function $x(y)$, but it does not discuss differentiability of that function, because no assumptions are made about differentiability of g, or of f with respect to y. The next result shows that when g is the identity and the partial B-derivative of f with respect to y exists at (x_0, y_0), then we obtain the same kind of differentiation formula as is found in the classical implicit-function theorem.

Theorem 3.3. *Assume the notation and hypotheses of Theorem 3.2, with $X = W$ and $g = I$. If $D_y f(x_0, y_0)$ exists, then x is B-differentiable at y_0 with*

$$Dx(y_0)k = D_x f(x_0, y_0)^{-1}[-D_y f(x_0, y_0)k]. \tag{3.3}$$

Further, if $D_y(x_0, y_0)$ is strong and $D_x(x_0, y_0)$ is linear (that is, a strong Fréchet derivative), then $Dx(y_0)$ is a strong B-derivative.

Proof: We have for y near y_0,

$$\begin{aligned} 0 &= f(x(y), y) - f(x_0, y_0) \\ &= [f(x(y), y) - f(x_0, y)] + [f(x_0, y) - f(x_0, y_0)]. \end{aligned}$$

Define functions $r(y)$ and $s(y)$ by

$$f(x_0, y) - f(x_0, y_0) = D_y f(x_0, y_0)(y - y_0) + s(y),$$

and

$$f(x(y), y) - f(x_0, y) = D_x f(x_0, y_0)[x(y) - x_0] + r(y).$$

Then

$$\begin{aligned} x(y) - x_0 &= D_x f(x_0, y_0)^{-1}[f(x(y), y) - f(x_0, y) - r(y)] \\ &= D_x f(x_0, y_0)^{-1}[-D_y f(x_0, y_0)(y - y_0) - r(y) - s(y)], \end{aligned} \tag{3.4}$$

and we obtain

$$\|x(y) - x_0 - D_x f(x_0, y_0)^{-1}[-D_y f(x_0, y_0)(y - y_0)]\| \leq \delta\|r(y) + s(y)\|. \tag{3.5}$$

The strong derivative property and the Lipschitz continuity of x imply that $r(y) \in Flip(y_0)$. In particular, $r(y) = o(y - y_0)$. But $s(y) = o(y - y_0)$ by definition of the B-derivative, and this together with (3.5) shows that the formula in (3.3) holds.

Now if $D_y f(x_0, y_0)$ is strong, then $s(y) \in Flip(y_0)$. If $D_x(x_0, y_0)$ is linear, then (3.4) yields

$$\begin{aligned} x(y) - x_0 - D_x f(x_0, y_0)^{-1}[-D_y f(x_0, y_0)(y - y_0)] \\ = -D_x f(x_0, y_0)^{-1}[r(y) + s(y)]. \end{aligned} \tag{3.6}$$

Since the right-hand side of (3.6) is evidently $Flip(y_0)$, we see that in this case $Dx(y_0)$ is strong. ∎

One might ask whether the assumption in Theorem 3.3 that $D_x f(x_0, y_0)$ is linear could be replaced by the assumption that it is a strong B-derivative, while still retaining the conclusion that $Dx(y_0)$ is strong. The following counterexample shows that this cannot be done.

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

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

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

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

We have $D_x f(0, 0) = \gamma^{-1}$, and this is a strong B-derivative. Further, f has a continuous partial F-derivative in y. If we solve for the implicit function $x(y)$, we find that

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

so that $x(0) = 0$ and (by the chain rule) $Dx(0) = \gamma$.

Now let

$$\begin{aligned} \Delta(y) &= x(y) - x(0) - D_x(0)(y - 0) \\ &= \gamma(y + r(y)) - \gamma(y), \end{aligned}$$

where $r(y) = (0, y_1^2)$. To say $Dx(0)$ is strong is to say that $\Delta \in Flip(0)$. 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,

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

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 $\Delta \notin Flip(0)$.

4. Application to Variational Problems.

In this section we sketch an application of the theory developed in §3 to parametric variational inequalities. Such problems involve a closed convex set C in a Hilbert space X, and a function F from the product of X and a normed linear space Y to X. For fixed $y \in Y$, one wishes to find a point $c_0 \in C$ such that for each $c \in C$,

$$\langle c - c_0, F(c_0, y) \rangle \geq 0, \tag{4.1}$$

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

$$N_C(c_0) = \{ x \in X \mid \text{for each } c \in C, \langle c - c_0, x \rangle \leq 0 \},$$

then we can rewrite (4.1) as

$$0 \in F(c_0, y) + N_C(c_0), \tag{4.2}$$

and because of the form of (4.2) these problems are also 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 [24] 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 why the composition formulation adopted in that section can be important.

Although the development in [24] 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 belongs to $Lip(1)$. For this and numerous other results about projectors, see [29].

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

$$g(x) = (\Pi_C(x), x - \Pi_C(x)).$$

This g is sometimes called the "Minty map"; it is a Lipschitzian homeomorphism of X onto the graph of N_C (see [3]). Define $f : W \times Y \to X$ by

$$f((u, v), y) = F(u, y) + v.$$

Then it is simple to show that solving (4.1) is equivalent to finding a solution x of

$$f(g(x), y) = 0, \tag{4.3}$$

since for any such x the point $\Pi_C(x)$ solves (4.1), whereas any point satisfying (4.1) also satisfies (4.3). In this way we have reduced the variational problem to the problem of finding a zero of the single-valued, continuous function given in (4.3).

The theory of §3, particularly Theorem 3.2, now provides a tool to establish the existence and investigate the behavior of parametric solutions of (4.1) in the case in which F is B-differentiable. If we wish to apply that theory, we need to verify the principal assumption of Theorem 3.2. If we suppose that x_0 is a solution of (4.3) for $y = y_0$ and that we wish to investigate parametric solutions $x(y)$ for y near y_0, then we see that the critical fact to verify is that $D_w f(g(x_0), y_0) \circ g$ has a single-valued, Lipschitzian inverse defined from a neighborhood M of the origin in X to a neighborhood N of x_0. If we write D for $D_w F(g(x_0), y_0)$, then the definitions of f and g mean that we need to show that the function $D(\Pi_C(x)) + (x - \Pi_C(x))$ has an inverse of the kind just described. However, by our earlier remarks this is completely equivalent to verifying that for each z in some neighborhood M of the origin in X, there is some $x(z) \in C$ that is the only solution in a neighborhood N of x_0 of the variational inequality

$$\langle c - x, D(x) - z \rangle \leq 0 \text{ for each } c \in C. \tag{4.4}$$

In general, (4.4) may be simpler than (4.1); in particular, when D is actually linear (i.e., a Fréchet derivative), then (4.4) is a *linear* variational inequality over C, as contrasted to the *nonlinear* variational problem (4.1). In that case, we recover the type of result proved in [24, Th. 2.1] under the more restrictive assumption that the derivative with respect to x was a continuous Fréchet derivative. (*Note:* In [24, Th. 2.1] it is not explicitly stated that the space X must be complete; however, this assumption is necessary (as it was in Th. 3.2 here) since the contraction theorem is used.)

In fact, this situation also provides the example, promised in §3, 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 [14]. Therefore, for general closed convex sets C, we cannot expect to use the B-derivative of the function g appearing in (4.3), and thus the use of the composition is necessary.

There are, however, important special cases in which a B-derivative derivative of Π_C exists. For example, Haraux [9] 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, and the results of §3 can then be applied with $g = I$. We intend to deal with that case in a separate paper.

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] Aubin, J. P. (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 (North-Holland Mathematics Studies No. 5), Amsterdam, The Netherlands.

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

[5] Cornet, B. and Laroque, G. (1983). Lipschitz Properties of Solutions in Mathematical Programming. *Discussion Paper 8321*, Center for Operations Research and Econometrics, Louvain-la-Neuve, Belgium.

[6] Dafermos, S. (1986). Sensitivity Analysis in Variational Inequalities. Manuscript, to appear in *Math. Operations Res.*

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

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

[9] 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.

[10] Jittorntrum, K. (1984). Solution Point Differentiability Without Strict Complementarity in Nonlinear Programming. *Math. Programming Study* **21**, 127-138.

[11] Jongen, H. Th., Klatte, D., and Tammer, K. (1988). Implicit Functions and Sensitivity of Stationary Points. *Preprint No. 1*, Lehrstuhl C für Mathematik, Rheinisch-Westfälische Technische Hochschule, Aachen, West Germany.

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

[13] 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.

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

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

[16] Kyparisis, J. (1987). Sensitivity Analysis Framework for Variational Inequalities. *Mathematical Programming* **38**, 203-213.

[17] Kyparisis, J. (1988). Perturbed Solutions of Variational Inequality Problems Over Polyhedral Sets. *J. Optimization Theory and Appl.* **57**, 295-305.

[18] Pang, J. S. (1988). Newton's Method for B-Differentiable Equations. Manuscript, Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD.

[19] Pang, J. S. (1988). Solution Differentiability and Continuation of Newton's Method for Variational Inequality Problems Over Polyhedral Sets. Manuscript, Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD.

[20] Park, K. (1988). A Continuation Method for Nonlinear Programming by Generalized Equations. Ph.D. Dissertation, Department of Industrial Engineering, University of Wisconsin-Madison (in progress).

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

[22] 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.
[23] Robinson, S. M. (1976). Regularity and Stability for Convex Multivalued Mappings. *Math. Operations Res.* **1**, 130-143.
[24] Robinson, S. M. (1980). Strongly Regular Generalized Equations. *Math. Operations Res.* **5**, 43-62.
[25] Robinson, S. M. (1985). Implicit B-Differentiability in Generalized Equations. *Technical Summary Report No. 2854*, Mathematics Research Center, University of Wisconsin-Madison, Madison, WI.
[26] Robinson, S. M. (1987). Local Structure of Feasible Sets in Nonlinear Programming, III: Stability and Sensitivity. *Math. Prog. Study* **30**, 45-66.
[27] Shapiro, A. (1986). On Concepts of Directional Differentiability. Manuscript, Department of Mathematics and Applied Mathematics, University of South Africa, Pretoria, South Africa.
[28] Ursescu, C. (1975). Multifunctions with Closed Convex Graph. *Czechoslovak Math. J.* **25**, 438-441.
[29] Zarantonello, E. H. (1971). Projections on Convex Sets in Hilbert Space and Spectral Theory. In: E. H. Zarantonello, ed., *Contributions to Nonlinear Functional Analysis*, 237-424. Academic Press, New York.