ANNALES DE L'I. H. P., SECTION C

R. T. ROCKAFELLAR

Proto-differentiability of set-valued mappings and its applications in optimization

Annales de l'I. H. P., section C, tome S6 (1989), p. 449-482
<http://www.numdam.org/item?id=AIHPC_1989__S6__449_0>

PROTO-DIFFERENTIABILITY OF SET-VALUED MAPPINGS AND ITS APPLICATIONS IN OPTIMIZATION

R.T. ROCKAFELLAR
Department of Mathematics, University of Washington, Seattle, USA

Abstract. Proto-differentiability of a set-valued mapping (multifunction) G from one Euclidean space to another is defined in terms of graphical convergence of associated difference quotient multifunctions. The nature and consequences of the property are investigated in considerable detail to provide a basis for applications. Applications are demonstrated for the theory of optimization by verifying the proto-differentiability of some of the most important multifunctions in that theory, specifically multifunctions giving the set of solutions to a parameterized system of constraints or to a parameterized variational inequality or collection of optimality conditions. The fact that such multifunctions are actually differentiable in this a generalized sense has not previously been detected.

Keywords: Multifunctions, proto-differentiability, semi-differentiability, Lipschitz properties, generalized equations, variational inequalities, perturbations of constraints, set convergence, nonsmooth analysis, optimization.

* This research was supported in part by grants from the National Science Foundation and the Air Force Office of Scientific Research at the University of Washington, Seattle.

1. Introduction

Set-valued mappings, or *multifunctions* as we shall call them here, arise in several ways in connection with problems of optimization. In a typical situation one has a problem $(P(u))$ in $\mathbb{R}^n$ that depends on a parameter vector $u \in \mathbb{R}^d$. A multifunction $G : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ can be defined by letting $G(u)$ denote the set of all points x satisfying the constraints of $(P(u))$, which could be equations and inequalities involving certain functions with both x and u as arguments. Alternatively $G(u)$ could be the set of optimal solutions to $(P(u))$ or the set of points x satisfying a collection of necessary conditions for optimality, and so forth. Multiplier vectors could also be involved: $G(u)$ could consist of the pairs $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ satisfying something like the Kuhn-Tucker conditions for problem $(P(u))$, for instance. In all these cases the multivaluedness of G is an inherent feature, or at least a strong possibility. Even when $G(u)$ is the optimal solution set, it may contain more than one element for certain choices of u that cannot realistically be left out of consideration.

Such circumstances raise serious difficulties for the study of how $G(u)$ can change relative to changes in u. Classical notions of continuity and differentiability, developed for functions rather than multifunctions, obviously do not apply.

The goal of this paper is to demonstrate that a wide class of multifunctions important in optimization nonetheless enjoys a property that we call *proto-differentiability*. To help with understanding the consequences of this fact, it is essential that the nature of proto-differentiability be illuminated at the same time. This we do from several angles, exploring in particular the relationship of proto-differentiability to true differentiability and an attractive concept of *semi-differentiability*.

From a geometric point of view, proto-differentiation of a multifunction corresponds to looking at certain tangent cones to the graph of the multifunction. Such a pattern of analysis has been pioneered by Aubin [**1**], [**2**], [**3**], [**4**]. It has already been shown to yield much information of use in connection with optimization and allied subjects. Aubin has focused chiefly on the contingent cone and the Clarke tangent cone, while Frankowska [**5**], [**7**], [**8**] has made potent use of an intermediate type of tangent cone (first treated by Ursescu [**20**]), which in this paper is called the derivative cone. The distinguishing feature of proto-differentiability is its requirement that the contingent cone and derivative cone coincide. In this it may be compared, as far as the geometry of graphs of multifunctions is concerned, with the study of tangential *regularity* in the sense of Clarke [**8**], [**9**]. Tangential regularity requires the contingent cone to coincide with the Clarke tangent cone rather than merely with the derivative cone. This is a stronger property which corresponds, in the case of multifunctions, to a concept we refer to as *strict* proto-differentiability.

All this tangent cone terminology could be bewildering to someone not accustomed to it, so the reader may be glad to know that proto-differentiability can be defined in a relatively simple and natural way without it. The fact that this property is commonly present for the multifunctions of interest in optimization has not previously been applied or even recognized. Roughly speaking, multifunctions expressing feasibility turn out to be proto-differentiable when an appropriate constraint qualification is satisfied, whereas those expressing optimality or the like, such as subgradient multifunctions, are proto-differentiable when the parameterization is sufficiently rich. The first fact, although not previously demonstrated in the generality furnished here, is not very surprising in view of the studies in the framework of nonsmooth analysis that have already been made of tangential regularity of sets defined by constraints [9, pp. 55–57], [10, Prop. 4.4]. The second fact is less expected.

Multifunctions of the second type have indeed been investigated previously for certain differential properties connected with tangent cones to their graphs, but the results, in concentrating on the Clarke tangent cone, have primarily been somewhat negative. The case of subgradient multifunctions illustrates this well. Results in Rockafellar [11] establish that for multifunctions ∂f associated with convex functions and saddle functions, or more generally for any maximal monotone multifunction, the Clarke tangent cone to the graph at any point is always a *subspace* and thus is incapable of reflecting anything other than "two-sided" aspects of differentiation. To the extent that a "corner" of the graph of ∂f may be involved, the Clarke tangent cone has to be degenerate. For such multifunctions, therefore, *strict* proto-differentiability is a property that has very powerful consequences when it is present—and theorems can be stated about it being present *almost everywhere* on the graph (cf. Rockafellar [11])—but which is unusable in characterizing local one-sided behavior. It cannot be invoked at *every* point of the graph of ∂f unless the graph happens to be smooth and f itself is correspondingly a generalized sort of C^2 function.

Sensitivity analysis of the kind carried out by Aubin [2] in convex programming, which in effect assumes strict proto-differentiability at the point under scrutiny, as pointed out in Rockafellar [11], suffers a serious limitation therefore in its applicability. This limitation is removed if strict proto-differentiability can be replaced by proto-differentiability.

Although we leave to another paper [12] the full study of proto-differentiability in the case of subgradient multifunctions in convex analysis, we do cover in §6 of the present paper a related case with equal claim to importance in sensitivity analysis. This concerns the multifunction which gives the Kuhn-Tucker points in a smooth (not necessarily convex) programming problem.

2. Proto-differentiability.

The convergence of sets in $\mathbf{R}^n$ will be a key ingredient in our definitions of generalized differentiability of multifunctions. A family of sets $S_t \subset \mathbf{R}^n$ parameterized by $t > 0$ is said to converge to a set $S \subset \mathbf{R}^n$ as $t \downarrow 0$, written

$$(2.1) \qquad S = \lim_{t \downarrow 0} S_t,$$

if S is closed and

$$(2.2) \qquad \lim_{t \downarrow 0} \operatorname{dist}(w, S_t) = \operatorname{dist}(w, S) \text{ for all } w \in \mathbf{R}^n,$$

where "dist" denotes Euclidean distance. It is often convenient to view this property as the equation

$$(2.3) \qquad S = \liminf_{t \downarrow 0} S_t = \limsup_{t \downarrow 0} S_t,$$

where

$$(2.4) \qquad \liminf_{t \downarrow 0} S_t := \{w \mid \limsup_{t \downarrow 0} \operatorname{dist}(w, S_t) = 0\},$$

$$(2.5) \qquad \limsup_{t \downarrow 0} S_t := \{w \mid \liminf_{t \downarrow 0} \operatorname{dist}(w, S_t) = 0\}.$$

Note that the points w belonging to the "lim inf" in (2.4) are the ones expressible as the limit of a family of elements $w_t \in S_t$ defined for all t in some interval $(0, \tau)$, whereas the ones belonging to the "lim sup" in (2.5) need only be expressible as the limit of some sequence $w^\nu \in S_{t^\nu}$ corresponding to a sequence $t^\nu \downarrow 0$. (We use superscript ν in this paper as the universal index for sequences: $\nu = 1, 2, \ldots$). Both of the sets (2.4) and (2.5) are necessarily closed.

Yet another way of characterizing the concept of set convergence is the following: (2.1) holds if and only if S is a closed set such that for arbitrarily large $\rho > 0$ and arbitrarily small $\varepsilon > 0$, there exists $\tau > 0$ for which

$$(2.6) \qquad S_t \cap \rho B \subset S + \varepsilon B \text{ and } S \cap \rho B \subset S_t + \varepsilon B \text{ when } t \in (0, \tau).$$

Here B denotes the closed unit ball in the Euclidean norm but could be replaced by any bounded neighborhood of the origin. This characterization results from the fact that the distance functions in (2.2) are uniformly Lipschitzian in w (with modulus 1). Therefore.

if they converge pointwise on $\mathbb{R}^n$ as asserted in (2.2), they actually converge uniformly on all bounded subsets of $\mathbb{R}^n$.

The sets to which we shall want to apply such convergence in order to define proto-differentiability are the graphs of multifunctions. Recall that the *graph* of a multifunction $G : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ is the set

$$\operatorname{gph} G = \{(u, x) \mid x \in G(u)\}. \tag{2.7}$$

The *effective domain* of G, on the other hand, is

$$\operatorname{dom} G = \{u \mid G(u) \neq \emptyset\}, \tag{2.8}$$

while the *effective range* of G is

$$\operatorname{rge} G = \{x \mid \exists u \text{ with } x \in G(u)\}. \tag{2.9}$$

In terms of the *inverse* G^{-1} of G, defined by

$$u \in G^{-1}(x) \iff x \in G(u), \tag{2.10}$$

one obviously has

$$\operatorname{rge} G = \operatorname{dom} G^{-1} \text{ and } \operatorname{dom} G = \operatorname{rge} G^{-1}. \tag{2.11}$$

One says that G has *closed graph* if the $\operatorname{gph} G$ is closed as a subset of $\mathbb{R}^d \times \mathbb{R}^n$. Clearly G is of closed graph if and only if G^{-1} is of closed graph.

Definition 2.1. Let $G : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ be any multifunction, not necessarily of closed graph, and let $u \in \operatorname{dom} G$ and $x \in G(u)$. Let $D_t : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ be the difference quotient multifunction at u relative to x, defined by

$$D_t(\omega) = [G(u + t\omega) - x]/t \text{ for } t > 0. \tag{2.12}$$

We shall say that G is *proto-differentiable* at u relative to x if there is a multifunction $D : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ such that D_t converges in graph to D, i.e. the set $S_t = \operatorname{gph} D_t$ converges in $\mathbb{R}^d \times \mathbb{R}^n$ to the set $S = \operatorname{gph} D$ as $t \downarrow 0$. In this event we shall call D the *proto-derivative* of G at u relative to x and employ the notation $D = G'_{u,x}$.

The relationship between proto-differentiability and other ideas of differentiability will be examined in §3. For now we develop the concept along its own natural lines.

beginning with the geometry of epigraphs and the connection between Definition 2.1 and the definitions of generalized differentiability proposed by Aubin [1], [2].

It will help us if we return temporarily to the study of a set $C \subset \mathbf{R}^n$ and a point $x \in C$. The set

$$\limsup_{t \downarrow 0} t^{-1}(C - x) \tag{2.13}$$

is known as the *contingent cone* to C at x, having first been given that name by Bouligand [13] in 1932. It is generated by all the directions from which x can be approached by a sequence in C. Specifically, a vector ξ belongs to the "lim sup" in (2.13) if and only if there exists a sequence of points $x^\nu \in C$ and scalars $t^\nu \downarrow 0$ such that $(x^\nu - x)/t^\nu \to \xi$. The word *cone* in this context refers to a set that can be expressed as a union of rays emanating from the origin, i.e. a set that is closed under the operation of nonnegative scalar multiplication. The contingent cone (2.13) is characterized as the smallest cone containing 0 and all the (direction) vectors ξ with $|\xi| = 1$ expressible as in the form

$$\xi = \lim_{\nu \to \infty} (x^\nu - x)/|x^\nu - x|$$

where $x^\nu \in C$, $x^\nu \neq x$, and the limit in question exists. (The sequence $\{x^\nu\}_{\nu=1}^{\infty}$ is said to converge to x *in the direction of* ξ in this case.)

The set

$$\liminf_{t \downarrow 0} t^{-1}(C - x) \tag{2.14}$$

will be called here the *derivative cone* of C at x. It is less well known but has been employed by a number of authors, expecially by Frankowska [5], [6], [7], who refers to it as the "intermediate cone" because it lies between the contingent cone and the Clarke tangent cone to C at x. Our calling it the derivative cone is suggested by the following characterization, which ties in with a long tradition in mathematical programming.

Let us say that $y : [0, \tau) \to \mathbf{R}^n$ is an *emanating arc* in C at x if $y(t) \in C$ for all $t \in [0, \tau)$, $y(0) = x$, $y(t) \to x$ as $t \downarrow 0$, and the limit

$$y'_+(0) := \lim_{t \downarrow 0} [y(t) - y(0)]/t \tag{2.15}$$

exists. Then $y'_+(0)$ is the (right) *derivative* of y at x. The set (2.14) turns out to consist of all the vectors $\xi \in \mathbf{R}^n$ expressible in $\xi = y'_+(0)$ for the various emanating arcs y in C at x. This is apparent from the description given earlier to the "lim inf" of S_t as $t \downarrow 0$

when applied to $S_t = t^{-1}[C - x]$. The idea of forming a cone that consists of derivative vectors like ξ has been followed in mathematical programming since the early days in the development of optimality conditions, except that the arcs y have usually been considered to be differentiable on an interval $[0, \tau)$ rather than just "right differentiable" at $t = 0$.

Both the contingent cone (2.13) and the derivative cone (2.14) are always closed cones containing the origin. The second is obviously contained within the first. We shall say that C is *approximable* at x if the two cones coincide, i.e. if the limit set

$$\lim_{t \downarrow 0} t^{-1}[C - x] \tag{2.16}$$

exists. This property dictates that the functions

$$\begin{aligned} d_t(\xi) &= \operatorname{dist}(\xi, t^{-1}[C - x]) \\ &= [\operatorname{dist}(x + t\xi, C) - \operatorname{dist}(x, C)]/t \end{aligned} \tag{2.17}$$

converge as $t \downarrow 0$ to a function

$$d(\xi) = \operatorname{dist}(\xi, K), \tag{2.18}$$

where K is a certain closed set, necessarily containing the origin. Shapiro [14] has called the set K, when it exists, the *approximating* cone to C at x. He did not connect it up with the theory of set convergence, however.

Proposition 2.2. *The multifunction $G : \mathbf{R}^d \rightrightarrows \mathbf{R}^n$ is proto-differentiable at u relative to a point $x \in G(u)$ if and only if the set* gph G *is approximable at (u, x). The graph of the proto-derivative multifunction $G'_{u,x}$ then equals the approximating cone to* gph G *at (u, x), which is simultaneously the derivative cone and the contingent cone.*

Proof. To establish this on the basis of what has just been explained, all one needs is the observation that for $C = \operatorname{gph} G$ the set $t^{-1}[C - (u, x)]$ is just $\operatorname{gph} D_t$. $\square$

Some elementary consequences of the definition of proto-differentiability will be recorded next.

Proposition 2.3. *Suppose that the proto-derivative $G'_{u,x}$ exists. Then for every $\omega \in \mathbf{R}^d$ one has*

$$G'_{u,x}(\omega) = \limsup_{\substack{\omega' \to \omega \\ t \downarrow 0}} [G(u + t\omega') - x]/t \tag{2.19}$$

and at the same time

$$G'_{u,x}(\omega) = \{\xi \mid \textit{for some arc } v : [0,\tau) \to \mathbf{R}^d \textit{ with } v(0) = u, \\ v'_+(0) = \omega, \textit{ one can select } y(t) \in G(v(t)) \\ \textit{for all } t \in [0,\tau) \textit{ so that } y(0) = x,\ y'_+(0) = \xi\}. \tag{2.20}$$

Conversely, if for every $\omega \in \mathbf{R}^d$ the set defined by the right side of (2.19) coincides with the set defined by the right side of (2.20), then the proto-derivative $G'_{u,x}$ exists.

Proof. The right side of (2.19) defines the set $D^+(\omega)$, where D^+ is the multifunction whose graph is the contingent cone to $\operatorname{gph} G$ at (x, u),

$$\operatorname{gph} D^+ = \limsup_{t \downarrow 0} \operatorname{gph} D_t. \tag{2.21}$$

The right side of (2.20), on the other hand, defines the st $D^-(\omega)$, where D^- is the multifunction whose graph is the derivative cone to $\operatorname{gph} G$ at (u, x),

$$\operatorname{gph} D^- = \liminf_{t \downarrow 0} \operatorname{gph} D_t. \tag{2.22}$$

The proposition comes down then again to the definition of proto-differentiability: $G'_{x,u}$ exists if and only if $D^+ = D^-$, in which event $G'_{x,u} = D^+ = D^-$. □

Proposition 2.4. *Let $G : \mathbf{R}^d \rightrightarrows \mathbf{R}^n$ be proto-differentiable at u relative to x, where $x \in G(u)$. Then the derivative multifunction $G'_{x,u} : \mathbf{R}^d \rightrightarrows \mathbf{R}^n$ has closed graph and satisfies*

$$0 \in G'_{u,x}(0), \textit{ and } G'_{u,x}(\lambda\omega) = \lambda G'_{u,x}(\omega) \textit{ for all } \omega \in \mathbb{R}^d \textit{ and } \lambda > 0. \tag{2.23}$$

Moreover $G'_{u,x}(0)$ is a closed cone which includes the contingent cone to $G(u)$ at x and therefore contains more than just 0 when x is not an isolated point of $G(u)$.

Proof. From Proposition 2.2 we know that the graph of $G'_{x,u}$ is a certain closed cone containing $(0,0)$. In particular it equals the contingent cone to $\operatorname{gph} G$ at (u, x). Everything follows at once from this. □.

Next we obtain a simple characterization of proto-differentiability by elaborating the meaning of graphical convergence for the difference quotient multifunctions D_t. We use the notation that the image of a set U under a multifunction G is the set

$$G(U) := \bigcup_{u \in U} G(u). \tag{2.24}$$

Proposition 2.5. *Let $G : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ be any multifunction and let $u \in \operatorname{dom} G$, $x \in G(u)$. In order that G be proto-differentiable at u relative to x, it is necessary and sufficient that there exist a closed-graph multifunction $D : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ (which will be $G'_{u,x}$) for which the following holds. For every $\varepsilon > 0$ (no matter how small) and $\rho > 0$ (no matter how large) one can find $\tau > 0$ such that*

$$D_t(\omega) \cap \rho B \subset D(\omega + \varepsilon B) + \varepsilon B \text{ for all } \omega \in \rho B,\ t \in (0, \tau), \tag{2.25}$$

$$D(\omega) \cap \rho B \subset D_t(\omega + \varepsilon B) + \varepsilon B \text{ for all } \omega \in \rho B,\ t \in (0, \tau). \tag{2.26}$$

Proof. *For $S_t = \operatorname{gph} D_t$ and $S = \operatorname{gph} D$ we invoke the characterization that $S_t \to S$ if and only if S is closed and for every $\varepsilon > 0$ and $\rho > 0$ there exists $\tau > 0$ such that*

$$S_t \cap \rho(B \times B) \subset S + \varepsilon(B \times B) \text{ and } S \cap \rho(B \times B) \subset S_t + \varepsilon(B \times B). \tag{2.27}$$

This is just a restatement of the property described in (2.6) in a form suitable for the product space $\mathbb{R}^d \times \mathbb{R}^n$. The two inclusions in (2.27) are equivalent to the ones in (2.25) and (2.26). □

For the sake of maintaining ties with other areas of nonsmooth analysis, where the Clarke tangent cone is fundamental, the following concept needs to be mentioned in comparison with proto-differentiability.

Definition 2.6. Let $G : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ be any multifunction, and let $u \in \operatorname{dom} G$ and $x \in G(u)$. We shall say that G is *strictly proto-differentiable* at u relative to x if it is proto-differentiable in the sense already defined and actually has the following, stronger property in place of formula (2.20) of Proposition 2.3. Consider any $\omega \in \operatorname{dom} G'_{u,x}$ and $\xi \in G'_{u,x}(\omega)$. Then there exist $\varepsilon > 0$ and $\tau > 0$ such that for each $u' \in \operatorname{dom} G$ with $|u' - u| \leq \varepsilon$ and $x' \in G(u')$ with $|x' - x| \leq \varepsilon$ (if any), and for each $t \in [0, \tau)$, it is possible to select $v(t, u', x') \in \operatorname{dom} G$ and $y(t, u', x') \in G(v(t, u', x'))$ in such a way that

$$v(0, u', x') = u' \text{ and } y(0, u', x') = x',$$

$$\lim_{\substack{(u',x') \to (u,x) \\ t \downarrow 0}} \frac{v(t, u', x') - v(0, u', x')}{t} = \omega,$$

$$\lim_{\substack{(u',x') \to (u,x) \\ t \downarrow 0}} \frac{y(t, u', x') - y(0, u', x')}{t} = \xi.$$

This concept certainly is much more complicated than plain proto-differentiability and indeed is not as easy to motivate in the present context. For such reasons along with considerations of space, we shall not devote attention to it here, except in stating the next proposition.

Proposition 2.7. *The multifunction $G : \mathbf{R}^d \rightrightarrows \mathbf{R}^n$ is strictly proto-differentiable at u relative to x, where $x \in G(u)$, if and only if the graph of G is tangentially regular at (u, x) in the sense of nonsmooth analysis, i.e. the contingent cone at (u, x) coincides with the Clarke tangent cone at (u, x). In this event $G'_{u,x}$ has convex graph, hence is a convex process.*

Proof. The Clarke tangent cone to $\operatorname{gph} G$ at (u, x) is the set

$$\liminf_{\substack{(u',x')\to(u,x)\\ t\downarrow 0}} t^{-1}[(\operatorname{gph} G) - (u', x')], \tag{2.28}$$

where the "lim inf" is restricted to elements (u', x') of $\operatorname{gph} G$. A pair (ω, ξ) belongs to this set if and only if it can be expressed in the manner described in Definition 2.6, as the reader can easily verify. Strict proto-differentiability thus requires the Clarke tangent cone, rather than just the derivative cone

$$\liminf t^{-1}[(\operatorname{gph} G) - (u, x)],$$

to agree with the contingent cone

$$\limsup_{t\downarrow 0} t^{-1}[(\operatorname{gph} G) - (u, x)].$$

The Clarke tangent cone is known always to be convex. □

Proposition 2.7 does make it possible to verify the strict proto-differentiability of various multifunctions (and thus their proto-differentiability) by applying known criteria for tangential regularity to their graphs.

3. Semi-differentiability and Differentiability.

To elucidate further the meaning of proto-differentiability, we explore connections with another concept that bridges the way toward the classical pattern of differentiability.

Definition 3.1. Let $G : \mathbf{R}^d \rightrightarrows \mathbf{R}^n$ be any multifunction, and let $u \in \operatorname{dom} G$, $x \in G(u)$. We shall say that G is *semi-differentiable at u relative to x* if there is a multifunction $D : \mathbf{R}^d \rightrightarrows \mathbf{R}^n$ such that the difference quotients $D_t(\omega)$ in (2.12) satisfy

$$\lim_{\substack{\omega'\to\omega\\ t\downarrow 0}} D_t(\omega') = D(\omega) \text{ for all } \omega \in \mathbf{R}^d. \tag{3.1}$$

We shall say that G is *differentiable at u* if, in addition, $G(u) = \{x\}$, i.e. G is single-valued at u itself, and D is a linear transformation.

When G is single-valued everywhere, i.e. a function, differentiability in the sense of Definition 3.1 reduces to the classical notion.

Theorem 3.2. *If G is semidifferentiable at u relative to x, then in particular G is proto-differentiable at u relative to x, and the multifunction D in the definition of semi-differentiability coincides with the proto-derivative $G'_{u,x}$.*

Proof. We stand on Proposition 2.3 as the characterization of proto-differentiability and also on the notation introduced in the proof of Proposition 2.3, namely the multifunctions D^+ and D^-. Since $D^+(\omega)$ denotes the right side of (2.19), it is clear from (3.1) that $D^+ = D$. We need only show that also $D^- = D$, where $D^-(\omega)$ is given by the right side of (2.20). Inasmuch as $D^+(\omega) \supset D^-(\omega)$ always, it is only the inclusion $D^-(\omega) \supset D(\omega)$ that needs justification. Consider any $\omega \in \operatorname{dom} D$ and $\xi \in D(\omega)$ in the case where $D(\omega)$ is given by (3.1). In particular we have

$$\liminf_{t \downarrow 0} D_t(\omega) = D(\omega)$$

by (3.1), so there must exist $\xi_t \in D_t(\omega)$ for all t in some interval $(0, \tau)$ such that $\xi_t \to \xi$ as $t \downarrow 0$. Let $v(t) = u + t\omega$, $y(t) = x + t\xi_t$. Then $v(0) = 0$ and $v'_+(0) = \omega$, while $y(0) = x$ and $y'_+(0) = \xi$. The condition $\xi_t \in D_t(\omega)$ can be written as

$$[y(t) - y(0)]/t \in [G(u + t\omega) - x]/t,$$

or equivalently as $y(t) \in G(v(t))$. Therefore $\xi \in D^-(\omega)$, and our goal has been achieved.□

A condition under which proto-differentiability conversely implies the stronger property of semi-differentiability will be presented in Theorem 4.3. Our attention at the moment turns instead toward the characteristics of semi-differentiability itself. We aim at identifying the domain in which this type of generalization might be viable. A notion of continuity will be used.

Definition 3.3. A multifunction $G : \mathbf{R}^d \rightrightarrows \mathbf{R}^n$ is *continuous* at u if

$$\lim_{u' \to u} G(u') = G(u). \tag{3.2}$$

It is *locally bounded* at u if there exist $\rho > 0$ and $\delta > 0$ such that

$$G(u') \subset \rho B \text{ for all } u' \text{ satisfying } |u' - u| \le \delta. \tag{3.3}$$

This concept of continuity reduces to the ordinary one in the case where G happens to be single-valued, i.e. a function rather than merely a multifunction. One must be cautious about its interpretation, however, in the multivalued case. The following example, where u is just a real variable, points out the pitfall:

$$G(u) = \begin{cases} \{0, 1/u\} & \text{when } u \ne 0, \\ \{0\} & \text{when } u = 0. \end{cases} \tag{3.4}$$

In this instance G is continuous at $u = 0$ and has $G(0) = \{0\}$. As a matter of fact, G is proto-differentiable at $u = 0$ relative to $x = 0$ and has $G'_{0,0} \equiv 0$. But G is not even locally bounded at $u = 0$. The discrepancy comes, of course, from the fact that set convergence only makes demands relative to an arbitrarily large bounded region at any one time. In the example one does have the property that for every $\rho > 0$ and $\varepsilon > 0$ there exists $\delta > 0$ such that

$$G(u) \cap \rho B \subset G(0) + \varepsilon B \text{ and } G(0) \cap \rho B \subset G(u) + \varepsilon B$$

when $|u| \leq \delta$. (Here $B = [-1, 1]$.)

Theorem 3.4. *If the multifunction $G : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ is semi-differentiable at u relative to x, then $G'_{u,x}$ is continuous everywhere on $\mathbb{R}^d$ with $\operatorname{dom} G'_{u,x} = \mathbb{R}^d$. Furthermore $u \in \operatorname{int} \operatorname{dom} G$ and*

$$x \in \liminf_{u' \to u} G(u'). \tag{3.5}$$

Proof. Fix any $\omega \in \mathbb{R}^d$. Let $D = G'_{u,x}$. Property (3.1) in the definition of semi-differentiability requires that for every $\rho > 0$ and $\varepsilon > 0$ there exist $\delta > 0$ and $\tau > 0$ such that

$$D_t(\omega') \cap \rho B \subset D(\omega) + \varepsilon B \text{ and } D(\omega) \cap \rho B \subset D_t(\omega') + \varepsilon B \tag{3.6}$$

when $|\omega' - \omega| < \delta$ and $t \in (0, \tau)$. It follows then, as seen through consideration of what happens as $t \downarrow 0$ with ω' fixed, that for every $\rho > 0$ and $\varepsilon > 0$ there exists $\delta > 0$ such that

$$D(\omega') \cap \rho B \subset D(\omega) + \varepsilon B \text{ and } D(\omega) \cap \rho B \subset D(\omega') + \varepsilon B \tag{3.7}$$

when $|\omega' - \omega| < \delta$. This means that D is continuous at ω. Applying the second inequality in (3.7) to the case where $\omega = 0$, we deduce that $D(\omega') \neq \emptyset$ for all ω' satisfying $|\omega'| < \delta$. (Recall here that $0 \in D(0)$, so $D(0) \cap \rho B \neq \emptyset$.) The positive homogeneity of $D = G'_{u,x}$ in (2.23) then gives us $D(\omega') \neq \emptyset$ for all ω'. Thus $\operatorname{dom} D = \mathbb{R}^d$. The second inclusion in (3.6) when applied to $\omega = 0$ tells us in like manner that for any $\varepsilon > 0$ there exist $\delta > 0$ and $\tau > 0$ such that

$$0 \in D_t(\omega') + \varepsilon B \text{ when } |\omega'| < \delta \text{ and } t \in (0, \tau).$$

We can write this equivalently as

$$G(u + t\omega') \cap (x + \varepsilon B) \neq \emptyset \text{ when } |\omega'| < \delta \text{ and } t \in (0, \tau).$$

Therefore u is an interior point of $\operatorname{dom} G$ and (3.5) is correct. □

The conclusions of Theorem 3.4 may be interpreted negatively as well as positively. They say that the concept of semi-differentiability, despite its natural appeal, is not suitable for the treatment for a multifunction G at a boundary point of the effective domain of G. Inasmuch as boundary points do play a crucial role in the case of some of the important multifunctions associated with problems of optimization, this observation provides motivation for why the more general concept of proto-differentiability is definitely needed.

Theorem 3.4 informs us likewise that semi-differentiability is inadequate for handling pairs $(u, x) \in \operatorname{gph} G$ for which (3.5) fails. However, in cases such as $G = \partial f$, where f is a closed proper convex function, (3.5) fails for every $x \in G(u)$ unless $G(u)$ happens to be a singleton! (See [15, Thm. 24.6].)

True differentiability is a special type of semi-differentiability by Definition 3.1. Therefore, according to Theorem 3.2, it falls within the larger realm of proto-differentiability. When G is actually a function, one can sensibly inquire further about the circumstances in which proto-differentiability will be the same as differentiability. An immediate conjecture is that this holds whenever the proto-derivative is a linear transformation. The conjecture is false, however.

The kind of situation to be wary of is demonstrated by an example closely related to the one in (3.4). Let u be a real variable and define

$$G(u) = \begin{cases} 1/u & \text{if } u \text{ is irrational,} \\ 0 & \text{if } u \text{ is rational.} \end{cases} \tag{3.8}$$

There is no multivaluedness here, and G is not even continuous at $u = 0$, much less differentiable there. Nonetheless G is proto-differentiable at $u = 0$ relative to $x = 0 = G(0)$, and the proto-derivative $G'_{0,0}$ is linear (the constant 0).

When discontinuities such as seen in this example are excluded, everything does fall into place, however.

Theorem 3.5. *Suppose $G : \mathbf{R}^d \rightarrow \mathbf{R}^n$ is actually a function (single-valued). Then G is differentiable at u if and only if G is continuous at u and at the same time proto-differentiable at u relative to $x = G(u)$, with $G'_{u,x}$ linear.*

The proof of Theorem 3.5 is postponed until §4, just after the proof of Theorem 4.1, because it will then be much easier to carry out.

A comment at this stage may head off some possible confusion in the treatment of the special case of a function $g : \mathbf{R}^d \rightarrow \mathbf{R}$. One could consider such a function as a multifunction that happens to be single-valued: in present notation with

$$G(u) = \{g(u)\}. \tag{3.9}$$

But one could also handle it in terms of

$$G(u) = \{x \in \mathbb{R} \mid x \geq g(u)\} \quad \text{(epigraphical framework)} \tag{3.10}$$

or instead

$$G(u) = \{x \in \mathbb{R} \mid x \leq g(u)\} \quad \text{(hypographical framework)} \tag{3.11}$$

All three choices lead to useful concepts of generalized differentiation. In the case of (3.10), for instance, one can speak of *epigraphical* proto-derivatives in order to keep matters straight. An important advantage of (3.10) and (3.11) is that they are not limited to *real*-valued functions. They furnish viable approaches even when g can take on $\pm\infty$ as values, as often turns out to be convenient in optimization theory. There is a natural tie-in with first-order (epigraphical) epi-differentiation of extended-real-valued functions as developed in Rockafellar [10], but we shall not look at this further here.

4. Derivative Bounds and Lipschitz Properties.

The proto-derivative multifunction $G'_{u,x}$, when it exists, is always positively homogeneous in the sense of (2.23). One can therefore define its *(outer) norm* by

$$|G'_{u,x}| := \min\{\mu \in [0,\infty) \mid \xi \in G'_{u,x}(\omega) \Rightarrow |\xi| \leq \mu|\omega|\}, \tag{4.1}$$

with the convention that $|G'_{u,x}| = \infty$ if no such $\mu \in [0,\infty)$ exists. The first result in this section is a characterization of the case where $|G'_{u,x}| < \infty$ in terms of a kind of pointwise Lipschitz growth property holding for G at u relative to x. The property in question will be used subsequently in the verification of Theorem 3.5.

Theorem 4.1. *Suppose $G : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ is proto-differentiable at u relative to x. Then the following conditions are equivalent:*

(a) $|G'_{u,x}| < \infty$;

(b) the cone $G'_{u,x}(0)$ consists only of 0;

(c) there exist $\mu > 0$, $\rho > 0$ and $\tau > 0$ such that

$$G(u + t\omega) \cap (x + \rho B) \subset x + t\mu|\omega|B \text{ for all } \omega \in B,\ t \in (0,\tau). \tag{4.2}$$

If these properties are present, x must be an isolated point of $G(u)$. Moreover $|G'_{u,x}|$ is then the infimum of the values μ for which (c) holds.

Proof. We begin with condition (c) and observe by taking $\omega = 0$ in (4.2) that it implies $G(u) \cap (x + \rho B) = \{x\}$. Then x is an isolated point of $G(u)$. If we write (4.2) next in the equivalent form

$$D_t(\omega) \cap (\rho/t)B \subset \mu|\omega|B \text{ for all } \omega \in B,\ t \in (0,\tau). \tag{4.3}$$

where D_t is as before the difference quotient multifunction in (2.12), we see that

$$G'_{u,x}(\omega) \subset \mu|\omega|B \text{ for all } \omega \in B. \tag{4.4}$$

This follows from the formula

$$G'_{u,x}(\omega) = \lim\sup_{\substack{\omega' \to \omega \\ t \downarrow 0}} D_t(\omega') \tag{4.5}$$

in (2.19). We deduce from (4.4) and the positive homogeneity of $G'_{u,x}$ in (2.23) that $|G'_{u,x}| \leq \mu$. In particular, (a) holds.

We assume next that (a) holds and that $\overline{\mu}$ is a number satisfying

$$|G'_{u,x}| < \overline{\mu} < \infty. \tag{4.6}$$

We need to show that (c) holds, i.e. that (4.3) is valid for $\mu = \overline{\mu}$ and some choice of $\rho > 0$ and $\tau > 0$. This will also establish that $|G'_{u,x}|$ is the infimum of the values for which (c) holds.

In the contrary case, where (4.6) is satisfied and yet (4.3) fails to hold for any choice of $\rho > 0$ or $\tau > 0$, we can take arbitrary sequences $\rho^\nu \downarrow 0$ and $\tau^\nu \downarrow 0$ and somehow select $\omega^\nu \in B$, $t^\nu \in (0, \tau^\nu)$, and

$$\xi^\nu \in D_{t^\nu}(\omega^\nu) \cap (\rho^\nu / t^\nu) \text{ with } \xi^\nu \notin \overline{\mu}|\omega^\nu|B.$$

Then $\xi^\nu \neq 0$. If we set

$$\overline{\xi}^\nu = \xi^\nu / |\xi^\nu|, \; \overline{\omega}^\nu = \omega^\nu / |\xi^\nu|, \; \overline{t}^\nu = t^\nu |\xi^\nu|,$$

so that $t^\nu \xi^\nu = \overline{t}^\nu \overline{\xi}^\nu$, $t^\nu \omega^\nu = \overline{t}^\nu \overline{\omega}^\nu$, we get

$$\overline{\xi}^\nu \in D_{\overline{t}^\nu}(\overline{\omega}^\nu) \text{ with } \overline{\xi}^\nu \notin \overline{\mu}|\overline{\omega}^\nu|B,$$

where $|\overline{\xi}^\nu| = 1$ and $\overline{t}^\nu \leq \rho^\nu$. Then $\overline{t}^\nu \to 0$ and $|\overline{\omega}^\nu| < |\overline{\xi}^\nu|/\overline{\mu} = 1/\overline{\mu}$. Passing to subsequences if necessary we can suppose that $\overline{\xi}^\nu$ converges to some $\overline{\xi}$ and $\overline{\omega}^\nu$ to some $\overline{\omega}$. Then $\overline{\xi} \in G'_{u,x}(\overline{\omega})$ by (4.5), and yet $|\overline{\xi}| = 1$, $|\overline{\omega}| \leq 1/\overline{\mu}$, so that $\overline{\mu}|\overline{\omega}| \leq |\overline{\xi}| \neq 0$. This contradicts the strict inequality in (4.6): there could not be a number $\mu \in [0, \overline{\mu})$ for which

$$|\xi| \leq \mu|\omega| \text{ whenever } \xi \in G'_{u,x}(\omega). \tag{4.7}$$

So far we have verified the equivalence between (a) and (c) as well as the corresponding assertions about x and $|G'_{u,x}|$. To finish the proof of Theorem 4.1 it will suffice to demonstrate the equivalence between (a) and (b). The implication from (a) to (b) is quite trivial: under (4.7), the set $G'_{u,x}(0)$ cannot contain any $\xi \neq 0$. The implication from (b) to (a) is not much harder. If (a) is untrue, we can take any sequence $\mu^\nu \uparrow \infty$ and select vectors $\omega^\nu \in \operatorname{dom} G'_{u,x}$ and $\xi^\nu \in G'_{u,x}(\omega^\nu)$ such that $|\xi^\nu| > \mu^\nu |\omega^\nu|$. By setting $\overline{\xi}^\nu = \xi^\nu / |\xi^\nu|$ and $\overline{\omega}^\nu = \omega^\nu / |\xi^\nu|$, we can transform this (in view of (2.23)) into

$$\overline{\xi}^\nu \in G'_{u,x}(\overline{\omega}^\nu) \text{ with } |\overline{\xi}^\nu| = 1 \text{ and } |\overline{\omega}^\nu| < 1/\mu^\nu \to 0.$$

Passing to subsequences if necessary, we can obtain $\overline{\xi}^\nu \to \overline{\xi}$ and $\overline{\omega}^\nu \to 0$. Then $\overline{\xi} \in G'_{u,x}(0)$ by the closed graph property in Proposition 2.4, but $|\overline{\xi}| = 1$, so $\overline{\xi} \neq 0$. This contradicts (b). □

Proof of Theorem 3.5. If G, which is now just a function, is differentiable at u, then from Definition 3.1 it is in particular semi-differentiable and we may conclude using Theorem 3.2 that G is proto-differentiable with $G'_{u,x}$ linear. The continuity of G at u follows then from property (3.5) in Theorem 3.4: of course this is also known classically.

Assume next, on the other hand, that G is continuous at u and proto-differentiable relative to $x = G(u)$ with $G'_{u,x}$ linear. Then for $D = G'_{u,x}$ we have $|D| < \infty$, so the properties in Theorem 4.1 hold. In particular there exist $\rho_0 > 0$ and $\tau_0 > 0$ such that

$$\begin{aligned} &G(u + t\omega) \in x + t(1 + |D|)|\omega| B \text{ for all} \\ &\qquad \omega \in B \text{ and } t \in (0, \tau_0) \text{ such that } G(u + t\omega) \in \rho_0 B. \end{aligned}$$

Since $G(u') \to G(u) = x$ as $u' \to u$, we can replace τ_0 by a smaller value if necessary and write this as

$$D_t(\omega) \in (1 + |D|)|\omega| B \text{ for all } \omega \in B \text{ and } t \in (0, \tau_0).$$

Let $\rho = (1 + |D|)$. Then in particular

$$D_t(\omega) \in \rho B \text{ for all } \omega \in B \text{ and } t \in (0, \tau_0).$$

Because of proto-differentiability, we can obtain from Proposition 2.5 for any $\varepsilon > 0$ a $\tau > 0$ such that

$$D_t(\omega) \in D(\omega + \varepsilon B) + \varepsilon B \text{ for all } \omega \in B \text{ and } t \in (0, \tau).$$

But

$$D(\omega + \varepsilon B) + \varepsilon B \subset D(\omega) + (\varepsilon |D| + \varepsilon) B = D(\omega) + \varepsilon \rho B.$$

Thus we are able to find for any $\varepsilon > 0$ a $\tau > 0$ such that

$$|D_t(\omega) - D(\omega)| \leq \varepsilon\rho \text{ for all } \omega \in B,\ t \in (0,\tau).$$

This means that G is differentiable at u with derivative D. □

A sufficient condition for proto-differentiability to imply semi-differentiability will be developed next in terms of a generalized Lipschitz property that was first introduced by Aubin [**2**].

Definition 4.2. A multifunction $G : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ *with closed graph* is said to be *pseudo-Lipschitzian* at u relative to $x \in G(u)$ if there exist $\varepsilon > 0$, $\delta > 0$ and $\mu > 0$ such that

$$G(u') \cap [x + \varepsilon B] \subset G(u'') + \mu|u' - u''|B$$
$$\text{for all } u', u'' \in [u + \delta B].$$

Sufficient conditions for G to be pseudo-Lipschitzian in this sense have been provided in many forms in Rockafellar [**16**] through the apparatus of subdifferential calculus.

Theorem 4.3. *Suppose $G : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ has closed graph and is pseudo-Lipschitzian at u relative to x, where $x \in G(u)$. Then G is proto-differentiable at u relative to x if and only if the limit*

$$D(\omega) = \lim_{t \downarrow 0} D_t(\omega) = \lim_{t \downarrow 0}[G(u + t\omega) - x]/t \tag{4.9}$$

exists for every ω, in which event the multifunction D that is defined in this manner is $G'_{u,x}$. Then, moreover, G turns out to be semi-differentiable at u relative to x, and $G'_{u,x}$ is itself globally Lipschitzian in the sense that

$$G'_{u,x}(\omega') \subset G'_{u,x}(\omega'') + \mu|\omega' - \omega''|B \text{ for all } \omega', \omega'' \in \mathbb{R}^d, \tag{4.10}$$

where μ is the modulus of pseudo-Lipschitz continuity for G in Definition 4.2.

Proof. Let ε, δ and μ be as in Definition 4.2. For $t > 0$ we have

$$G(u + t\omega') \cap [x + \varepsilon B] \subset G(u + t\omega'') + \mu t|\omega'' - \omega'|B$$

as long as $u + t\omega'$ and $u + t\omega''$ both lie in the ball $u + \delta B$. In other words,

$$D_t(\omega') \cap (\varepsilon/t)B \subset D_t(\omega'') + \mu|\omega'' - \omega'|B \text{ when } \omega', \omega'' \in (\delta/t)B. \tag{4.11}$$

Suppose G is proto-differentiable at u relative to x. For any $\rho > 0$ and $\varepsilon' > 0$ there exists by Proposition 2.5 a $\tau > 0$ such that for $\bar{\rho} = 2\rho(1+\mu)$ one has

$$D_t(\omega) \cap \bar{\rho}B \subset G'_{u,x}(\omega + \varepsilon' B) + \varepsilon' B \text{ for } \omega \in \bar{\rho}B,\ t \in (0,\tau), \tag{4.12}$$

$$G'_{u,x}(\omega) \cap \bar{\rho}B \subset D_t(\omega + \varepsilon' B) + \varepsilon' B \text{ for } \omega \in \bar{\rho}B,\ t \in (0,\tau). \tag{4.13}$$

Require $\varepsilon' < \rho$, so that $\rho + \varepsilon' < \bar{\rho}$ in particular. Take $\tau' \in (0,\tau)$ such that $\varepsilon/t \geq \bar{\rho}$ and $\delta/t \geq \bar{\rho}$ when $t \in (0,\tau')$. Then for $\omega', \omega'' \in \rho B$ and $t \in (0,\tau')$ one has by (4.13) that

$$\begin{aligned} G'_{u,x}(\omega') \cap \rho B &\subset [D_t(\omega' + \varepsilon' B) + \varepsilon' B] \cap \rho B \\ &\subset [D_t(\omega' + \varepsilon' B) \cap (\rho + \varepsilon')B] + \varepsilon' B. \end{aligned}$$

But also for $\omega', \omega'' \in \rho B$ and $t \in (0,\tau')$ one has by (4.11) that

$$D_t(\omega' + \varepsilon'\varsigma) \cap (\rho + \varepsilon')B \subset D_t(\omega'' + \varepsilon'\varsigma) + \mu|\omega'' - \omega'|B \text{ for all } \varsigma \in B,$$

because $\omega' + \varepsilon\varsigma$ and $\omega'' + \varepsilon'\varsigma$ belong to $(\rho + \varepsilon')B$, and

$$(\rho + \varepsilon')B \subset \bar{\rho}B \subset (\varepsilon/t)B \cap (\delta/t)B.$$

Thus for $\omega', \omega'' \in \rho B$ and $t \in (0,\tau')$ one has

$$D_t(\omega' + \varepsilon B) \cap (\rho + \varepsilon')B \subset D_t(\omega'' + \varepsilon' B) + \mu|\omega'' - \omega'|B$$

and therefore

$$\begin{aligned} G'_{u,x}(\omega') \cap \rho B &\subset [D_t(\omega'' + \varepsilon' B) + (\mu|\omega'' - \omega'| + \varepsilon')B] \cap \rho B \\ &\subset [D_t(\omega'' + \varepsilon' B) \cap (\rho + \mu|\omega'' - \omega'| + \varepsilon')B] + (\mu|\omega'' - \omega'| + \varepsilon')B. \end{aligned}$$

Here $\rho - \mu| + \varepsilon'\omega'' + \omega'| \leq \rho + \mu(2\rho) + \rho = \bar{\rho}$ by the definition of $\bar{\rho}$, so that

$$\begin{aligned} D_t(\omega'' + \varepsilon' B) &\cap (\rho + \mu|\omega'' - \omega'| + \varepsilon')B \\ &\subset \bigcup_{\omega \in [\omega'' + \varepsilon' B]} D_t(\omega) \cap \bar{\rho}B \subset \bigcup_{\omega \in [\omega'' + \varepsilon' B]} [G'_{u,x}(\omega + \varepsilon' B) + \varepsilon' B] \end{aligned}$$

by (4.12). It follows that for $\omega', \omega'' \in \rho B$ one has

$$G'_{u,x}(\omega') \cap \rho B \subset G'_{u,x}(\omega'' + 2\varepsilon' B) + 2\varepsilon' B + \mu|\omega'' - \omega'|B.$$

We have demonstrated this for arbitrary $\rho > 0$ and $\varepsilon' \in (0, \rho)$, and we so may conclude (because $G'_{u,x}$ has closed graph according to Proposition 2.4) that $G'_{u,x}$ is globally Lipschitzian with modulus μ in the sense of (4.10).

Using this fact we argue by (4.12) that when $\omega' \in \rho B$, $\tau \in (0, \tau')$, one has by (4.10)

$$D_t(\omega') \cap \rho B \subset G'_{u,x}(\omega') + \mu\varepsilon' B + \varepsilon' B$$

and consequently also by (4.10) for arbitrary $\omega \in \mathbf{R}^d$ that

$$(4.14) \qquad D_t(\omega') \cap \rho B \subset G'_{u,x}(\omega) + (\mu\varepsilon' + \varepsilon' + \mu|\omega' - \omega|)B \text{ for all } \omega' \in \rho B,\ t \in (0, \tau').$$

By (4.13), on the other hand, one has (since $\rho < \overline{\rho} - \varepsilon'$) that

$$(4.15) \qquad \begin{aligned} G'_{u,x}(\omega) \cap \rho B &\subset [D_t(\omega + \varepsilon' B) + \varepsilon' B] \cap \rho B \\ &\subset [D_t(\omega + \varepsilon' B) \cap (\rho + \varepsilon')B] + \varepsilon' B \\ &\subset [D_t(\omega' + [\varepsilon' + |\omega' - \omega|]B) \cap \overline{\rho} B] + \varepsilon' B \\ &\qquad \text{when } |\omega| \le \rho,\ t \in (0, \tau). \end{aligned}$$

Here $\overline{\rho} B \subset (\varepsilon / t)B$ when actually $t \in (0, \tau')$, and in that case by (4.11) one has

$$(4.16) \qquad \begin{aligned} D_t(\omega' + [\varepsilon' + |\omega' - \omega|]B) \cap \overline{\rho} B &\subset D_t(\omega') + \mu[\varepsilon' + |\omega' - \omega|]B \\ &\qquad \text{when } |\omega'| + \varepsilon' + |\omega' - \omega| \le \delta / t. \end{aligned}$$

Choose $\tau'' \in (0, \tau')$ small enough that

$$|\omega'| + \varepsilon + |\omega' - \omega| \le \delta / t \text{ when } |\omega| \le \rho,\ |\omega'| \le \rho,\ t \in (0, \tau''.$$

We then obtain from the combination of (4.15) and (4.16) the result that as long as $|\omega| \le \rho$, one has

$$(4.17) \qquad \begin{aligned} G'_{u,x}(\omega) \cap \rho B &\subset D_t(\omega') + (\mu\varepsilon' + \varepsilon' + |\omega' - \omega|)B \\ &\qquad \text{for all } \omega' \in \rho B,\ t \in (0, \tau''). \end{aligned}$$

In summary, for any fixed ω we can take arbitrary $\rho \ge |\omega|$ and $\varepsilon' \in (0, \rho)$ and then have both (4.14) and (4.17) holding over sufficiently small intervals $(0, \tau')$ and $(0, \tau'')$. This means that

$$\lim_{\substack{\omega' \to \omega \\ t \downarrow 0}} D_t(\omega') = G'_{u,x}(\omega) \text{ for all } \omega.$$

In other words G is semi-differentiable at u relative to x, as claimed. In particular the limit multifunction D in (4.9) does exist and equals $G'_{u,x}$.

For the final part of the proof of Theorem 4.3, we start merely from the assumption that the limits (4.9) exist and demonstrate that this implies proto-differentiability. We rely this time on the characterization of proto-differentiability in Proposition 2.3, as well as on property (4.11), which represents in this context our hypothesis of pseudo-Lipschitz continuity. Let $D^+(\omega)$ denote, as earlier in this paper, the set defined by the right side of (2.19). Let $D^-(\omega)$ be the corresponding set on the right side of (2.20). Our task is to verify that $D^+(\omega) \subset D(\omega) \subset D^-(\omega)$ when $D(\omega)$ is defined by (4.9).

The inclusion $D^+(\omega) \subset D(\omega)$ can be seen from the expression

$$D^+(\omega) = \limsup_{\substack{\omega' \to \omega \\ t \downarrow 0}} D_t(\omega')$$

together with (4.11) as an estimate for $D_t(\omega')$ in terms of $D_t(\omega)$. For the inclusion $D(\omega) \subset D^-(\omega)$, we consider an arbitrary pair (ω, ξ) with $\xi \in D(\omega)$. Because

$$D(\omega) = \liminf_{t \downarrow 0} D_t(\omega)$$

in particular, there is an interval $[0, \tau)$ such that for each $t \in (0, \tau)$ we can choose $\xi_t \in D_t(\omega)$ and do so in such a way that $\xi_t \to \xi$ as $t \downarrow 0$. The arcs $v(t) = u + t\omega$ and $y(t) = x + t\xi_t$ over $[0, \tau)$ then meet the requirement on the right side of (3.20) and establish for us that $\xi \in D^-(\omega)$. This was the last thing to prove. □

An auxiliary result which is complementary to Theorem 4.3 will be obtained next. It merely assumes a Lipschitz property for $G'_{u,x}$.

Theorem 4.4. *Let $G : \mathbf{R}^d \rightrightarrows \mathbf{R}^n$ be proto-differentiable at u relative to x, where $x \in G(u)$. If the multifunction $G'_{u,x}$ is pseudo-Lipschitzian at 0 relative to the pont $0 \in G'_{u,x}(0)$, then it actually has the global Lipschitzian property in (4.10) for some $\mu > 0$. In this case there exists for every $\rho > 0$ and $\varepsilon > 0$ a $\tau > 0$ such that*

$$(4.18) \qquad G(u + t\omega) \cap [x + t\rho B] \subset x + tG'_{u,x}(\omega) + t\varepsilon B \text{ for all } \omega \in \rho B,\ t \in (0, \tau).$$

If in addition $G'_{u,x}(0) = \{0\}$, this conclusion holds in a stronger form: for some $\overline{\rho} > 0$, one can find for every $\rho > 0$ and $\varepsilon > 0$ a $\tau > 0$ such that

$$(4.19) \qquad G(u + t\omega) \cap [x + \overline{\rho} B] \subset x + tG'_{u,x}(\omega) + t\varepsilon B \text{ for all } \omega \in \rho B,\ t \in (0, \tau).$$

Proof. The pseudo-Lipschitzian property for $G'_{u,x}$ at 0 relative to 0 means the existence of $\mu > 0$ such that for some $\varepsilon > 0$ and $\delta > 0$ one has

$$G'_{u,x}(\omega') \cap \varepsilon B \subset G'_{u,x}(\omega'') + \mu|\omega' - \omega''|B \text{ when } \omega', \omega'' \in \delta B.$$

When this holds, it can be applied for arbitrary $\omega', \omega'' \in \mathbb{R}^n$ and $\rho \geq \max\{|\omega'|, |\omega''|\}$ to the vectors

$$\overline{\omega}' = (\delta/\rho)\omega' \text{ and } \overline{\omega}'' = (\delta/\rho)\omega''.$$

These are in δB and therefore give

$$G'_{u,x}(\overline{\omega}') \cap \varepsilon B \subset G'_{u,x}(\overline{\omega}'') + \mu|\overline{\omega}' - \overline{\omega}''|B,$$

which because of the positive homogeneity of $G'_{u,x}$ in (2.23) is equivalent to

$$G'_{u,x}(\omega') \cap (\rho\varepsilon/\delta)B \subset G'_{u,x}(\omega'') + \mu|\omega' - \omega''|B.$$

Inasmuch as this holds for arbitrary $\rho \geq \max\{|\omega'|, |\omega''|\}$, the global Lipschitz property in (4.10) is valid for $G'_{u,x}$.

Continuing now from property (4.10) for $D = G'_{u,x}$, we invoke the proto-differentiability of G at u relative to x and draw from Proposition 2.5 the conclusion that for arbitrary $\rho > 0$ and $\varepsilon' > 0$ we can find $\tau > 0$ with

$$D_t(\omega) \cap \rho' B \subset D(\omega + \varepsilon' B) + \varepsilon' B \text{ for all } \omega \in \rho B,\ t \in (0, \tau),$$

where

$$D(\omega + \varepsilon' B) + \varepsilon' B \subset D(\omega) + (\mu\varepsilon' + \varepsilon')B.$$

Starting with an arbitrary $\varepsilon > 0$ and choosing $\varepsilon' \leq \varepsilon/(1 + \mu)$, we obtain in this way the inclusion

(4.20) $$D_t(\omega) \cap \rho B \subset D(\omega) + \varepsilon B \text{ for all } \omega \in \rho B,\ t \in (0, \tau),$$

which is equivalent to (4.18).

When $G'_{u,x}(0) = \{0\}$ we can make use of condition (c) in Theorem 4.1: there exist $\overline{\rho} > 0$, $\overline{\mu} > 0$ and $\overline{\tau} > 0$, such that

(4.21) $$G(u + t\omega) \cap [x + \overline{\rho} B] \subset x + t\overline{\mu}|\omega|B \text{ for all } \omega \in B,\ t \in (0, \overline{\tau}).$$

Again we consider arbitrary $\rho > 0$ and $\varepsilon > 0$. With the change of variables $t' = \rho t$, $\omega' = \omega/\rho$ we can write (4.21) instead as

$$(4.22) \qquad G(u + t'\omega') \cap [x + \overline{\rho} B] \subset x + t'\overline{\mu}|\omega'|B \text{ for all } \omega' \in \rho B,\ t' \in (0, \rho\overline{\tau}).$$

This being true, we can just as well convert notation from t' and ω' back to t and ω and express (4.22) in terms of D_t as

$$(4.23) \qquad D_t(\omega) \cap (\overline{\rho}/t)B \subset \overline{\mu}|\omega|B \text{ for all } \omega \in \rho B,\ t \in (0, \rho\overline{\tau}).$$

Taking $\rho' = (1 + \overline{\mu})\rho$ we call forth the property of G that has already been established in our proof, specifically that for this value ρ' and the given ε the corresponding version of (4.18) holds, or more conveniently for the moment, the equivalent statement in (4.20): there exists $\tau' > 0$ such that

$$(4.24) \qquad D_t(\omega) \cap \rho' B \subset D(\omega) + \varepsilon B \text{ for all } \omega \in \rho' B,\ t \in (0, \tau').$$

Since $\overline{\mu}|\omega| \leq \overline{\mu}\rho < \rho'$ when $\omega \in \rho B$ by the choice of ρ, we have from (4.23) the estimate

$$D_t(\omega) \cap (\overline{\rho}/t)B \subset D_t(\omega) \cap \overline{\mu}|\omega|B \subset D_t(\omega) \cap \rho' B$$
$$\text{for all } \omega \in \rho B,\ t \in (0, \rho\overline{\tau}),$$

Applying (4.24) and remembering that $\rho < \rho'$, we get for $\tau_1 = \min(\tau', \rho\overline{\tau})$ that

$$D_t(\omega) \cap (\overline{\rho}/t)B \subset D(\omega) + \varepsilon B \text{ for all } \omega \in \rho B,\ t \in (0, \tau_1).$$

But this is exactly the assertion of (4.19), except for the notation τ_1 in place of τ. □

5. Applications in Optimization.

The proto-differentiability of a number of multifunctions that are of central importance in the theory of optimization will be proved in this section. We do not try to cover the territory with thoroughness, but content ourselves for the purposes of this paper with presenting cases that demonstrate the depth and variety of the applications without getting us into further technical developments.

Three elementary results will help in verifying proto-differentiability.

Proposition 5.1. *A multifunction $G : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ is proto-differentiable at u relative to the element $x \in G(u)$ if and only if its inverse G^{-1} is proto-differentiable at x relative to the element $u \in G^{-1}(x)$, in which case $(G'_{u,x})^{-1} = (G^{-1}_{x,u})'$. The same holds for strict proto-differentiability.*

Proof. This is obvious because proto-differentiability is a property of the graph of a multifunction, cf. Proposition 2.2. □

Proposition 5.2. *Suppose $G = \overline{G} + g$, where $\overline{G} : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ is proto-differentiable at u relative to the element $\overline{x} \in \overline{G}(u)$ and $g : \mathbb{R}^d \to \mathbb{R}^n$ is a function (single-valued) that is differentiable at u. Then G is proto-differentiable at u relative to the element $x = \overline{x} + g(u) \in G(u)$ with $G'_{u,x} = \overline{G}'_{u,\overline{x}} + g'_u$.*

Proof. The characterization of proto-differentiability in Proposition 3.3 serves quickly to verify this. □

Proposition 5.3. *Suppose that $G : \mathbb{R}^d \rightrightarrows \mathbb{R}^n$ is both proto-differentiable and pseudo-Lipschitzian at u relative to the element $x \in G(u)$. Write $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ and accordingly $u = (u_1, u_2)$, and let $G_1 : \mathbb{R}^{d_1} \rightrightarrows \mathbb{R}^n$ be the multifunction defined by $G_1(\cdot) = G(\cdot, u_2)$. Then G_1 is both proto-differentiable and pseudo-Lipschitzian at u_1 relative to x.*

Proof. The combination of pseudo-Lipschitzian plus proto-differentiable is equivalent by Theorems 3.2 and 4.3 to pseudo-Lipschitzian plus the special property that

$$\lim_{t \downarrow 0} [G(u + t\omega) - x]/t \text{ exists for all } \omega.$$

This second combination is obviously preserved when forming G_1 as a restriction of G. □

In the remainder of this section we use the notion that $N_C(z)$ and $T_C(z)$ denote for a convex set C the normal cone and tangent cone to C at z in the sense of convex analysis. We also denote by $\nabla F(z)$ for a differentiable mapping $F : \mathbb{R}^r \to \mathbb{R}^m$ the $m \times r$ Jacobian matrix at z.

Theorem 5.4. *Let $G : \mathbf{R}^d \rightrightarrows \mathbf{R}^n$ have the form*

$$G(u) = \{x \in D \mid F(u,x) \in C\}, \tag{5.1}$$

where $F : \mathbf{R}^d \times \mathbf{R}^n \to \mathbf{R}^m$ is a mapping (single-valued) of class C' and the sets $C \subset \mathbf{R}^m$ and $D \subset \mathbf{R}^n$ are closed and convex. Suppose for a particular u and element $x \in G(u)$ that the following constraint qualification holds:

(5.2) *The only vector $y \in N_C(F(u,x))$ satisfying $-y\nabla_x F(u,x) \in N_D(x)$ is $y = 0$.*

Then G is both proto-differentiable and pseudo-Lipschitzian at u relative to x, in fact semi-differentiable there. The proto-derivative is given by

$$G'_{u,x}(\omega) = \{\xi \in T_D(x) \mid \nabla_u F(u,x)\omega + \nabla_x F(u,x)\xi \in T_C(F(x,u))\}. \tag{5.3}$$

Proof. Our strategy will be to use an alternative representation of G. Temporarily consider u and x again as variables rather than fixed, and let

$$H(u,z) = \{(u,x) \in \mathbf{R}^d \times D \mid F(u,x) - z \in C\} \tag{5.4}$$

and

$$A(u) = H(u,0) + h(u) \text{ where } h(u) = (-u,0). \tag{5.5}$$

Then A is virtually another copy of G: one has

$$A(u) = \{(0,x) \in \mathbf{R}^d \times \mathbf{R}^n \mid x \in G(u)\} \tag{5.6}$$

and also

$$A'_{u,(0,x)}(\omega) = \{(0,\xi) \mid \xi \in G'_{u,x}(\omega)\}$$

if such proto-derivatives exist. It will suffice therefore to prove that, under our hypotheses, A is proto-differentiable and pseudo-Lipschitzian at u relative to $(0,x)$ with

$$A'_{u,(0,x)}(\omega) = \{(0,\xi) \in \mathbf{R}^d \times T_D(x) \mid \nabla_u F(u,x)\omega + \nabla_x F(u,x)\xi \in T_C(F(x,u))\}. \tag{5.7}$$

(Semi-differentiability then follows from Theorem 4.3.) In fact by Propositions 5.2 and 5.3 we can reduce this to showing that H itself is proto-differentiable and pseudo-Lipschitzian at $(u,0)$ relative to (u,x), with

$$\begin{aligned} &H'_{(u,0),(u,x)}(\omega,\varsigma) \\ &= \{(\omega,\xi) \in \mathbf{R}^d \times T_D(x) \mid \nabla_u F(u,x)\omega + \nabla_x F(u,x)\xi - \varsigma \in T_C(F(u,x))\}. \end{aligned} \tag{5.8}$$

Consider now the mapping

$$\tilde{F}(v,x) = (v, F(v,x)) \in \mathbf{R}^d \times \mathbf{R}^m \text{ for } (v,x) \in \mathbf{R}^d \times \mathbf{R}^n \tag{5.9}$$

and translate (5.1) into

$$H(u,z) = \{(v,x) \in \tilde{D} \mid \tilde{F}(v,x) - (u,z) \in \tilde{C}\}, \tag{5.10}$$

where

$$\tilde{D} = \mathbf{R}^d \times D \text{ and } \tilde{C} = \{0\} \times C. \tag{5.11}$$

This gives us

$$H^{-1} = \tilde{F} + S, \tag{5.12}$$

where S is the multifunction defined by

$$S(v,x) = \begin{cases} -\tilde{C} & \text{if } (v,x) \in \tilde{D}, \\ \emptyset & \text{if } (v,x) \notin \tilde{D}. \end{cases} \tag{5.13}$$

The sets $\tilde{C}$ and $\tilde{D}$ are obviously closed and convex, and so also is the set gph $S = \tilde{D} \times (-\tilde{C})$. A convex set is tangentially regular everywhere (the Clarke tangent cone coinciding with the contingent cone, cf. [8]) and in particular therefore is approximable everywhere. Thus by Proposition 2.2, S is proto-differentiable everywhere. Furthermore $\tilde{F}$ is a mapping which is differentiable everywhere. Proposition 5.2 tells us that H^{-1} is in this case proto-differentiable everywhere. Specifically, the proto-derivative of H^{-1} at (u,x) relative to the element $(u,0) \in H^{-1}(u,x)$ is

$$(H^{-1})'_{(u,x),(u,0)}(\theta,\xi) = \nabla\tilde{F}(u,x)(\theta,\xi) + S'_{(u,x),(0,-w)}(\theta,\xi), \tag{5.14}$$

where $w = F(u,x)$ and

$$\begin{aligned} \text{gph } S'_{(u,x),(0,-w)} &= T_{\text{gph } S}((u,x),(0,-w)) \\ &= T_{\tilde{D}}(u,x) \times T_{-\tilde{C}}(0,-w) = T_{\tilde{D}}(u,x) \times (-T_{\tilde{C}}(0,w)). \end{aligned} \tag{5.15}$$

It follows then from Proposition 5.1 that H is proto-differentiable at $(u,0)$ relative to the element $(u,x) \in H(u,0)$ with

$$H'_{(u,0),(u,x)}(\omega,\zeta) = \{(\theta,\xi) \in T_{\tilde{D}}(u,x) \mid \nabla\tilde{F}(u,x)(\theta,\xi) - (\omega,\zeta) \in T_{\tilde{C}}(0,w)\}. \tag{5.16}$$

We calculate now that

$$T_{\tilde D}(u,x) = \mathbf{R}^d \times T_D(x) \text{ and } T_{\tilde C}(0,w) = \{0\} \times T_C(F(u,x)), \tag{5.17}$$

$$\nabla \tilde F(u,x) = (I, \nabla F(u,x)). \tag{5.18}$$

Formula (5.16) reduces therefore to (5.8).

At this stage we have taken care of the proto-differentiability properties of H but still have to establish the pseudo-Lipschitzian property. Let us rewrite the formula for H one more time, starting from (5.10), as

$$H(u,z) = \{(v,x) \mid \Phi(u,z,v,x) \in \tilde C,\ (v,x) \in \tilde D\}, \tag{5.19}$$

where

$$\Phi(u,z,v,x) = \tilde F(v,x) - (u,z) = (v-u, F(v,x) - z). \tag{5.20}$$

This representation fits the general pattern in the studies of pseudo-Lipschitz continuity in Rockafellar [16]. The sufficient condition given by [16, Theorem 3.2] for H to be pseudo-Lipschitzian at $(u,0)$ relative to the element (u,x) is the following constraint qualification: There should be no *nonzero* multiplier element $\tilde y$ satisfying

$$\tilde y \in N_{\tilde C}(\Phi(u,0,u,x)), \quad -\tilde y \nabla_{v,x}\Phi(u,0,u,x) \in N_{\tilde D}(u,x). \tag{5.21}$$

All we have to do is translate this back into our original notation using (5.20), (5.21), and (5.11). Obviously

$$\begin{aligned}
\Phi(u,0,u,x) &= (0, F(u,x)),\\
\nabla_{(v,x)}\Phi(u,0,u,x) &= \begin{bmatrix} I & 0 \\ \nabla_u F(u,x) & \nabla_x F(u,x) \end{bmatrix}\\
N_{\tilde D}(u,x) &= \{0\} \times N_D(x).\\
N_{\tilde C}(\Phi(u,0,u,x)) &= \mathbf{R}^d \times N_C(F(u,x)).
\end{aligned}$$

A vector $\tilde y$ satisfying (5.21) is a pair (y', y) such that

$$y' \in \mathbf{R}^d,\ y \in N_C(F(u,x)),\ -y' - y\nabla_u F(u,x) = 0,\ -y\nabla_x F(u,x) \in N_D(x).$$

Assumption (5.2) clearly ensures that the only such pair is $(y', y) = (0, 0)$. This completes the proof. $\square$

Example 5.5. Here is the version of Theorem 5.4 that corresponds to the standard formulation of a system of smooth constraints dependent on a parameter vector u. Let $G(u)$ denote the set of all $x \in \mathbf{R}^n$ satisfying

$$f_i(u,x) \begin{cases} \leq 0 & \text{for } i = 1, \dots, s \\ = 0 & \text{for } i = s+1, \dots, m, \end{cases} \tag{5.22}$$

where $f_i : \mathbf{R}^d \times \mathbf{R}^n \to \mathbf{R}$ is continuously differentiable for $i = 1, \dots, m$. This corresponds to the case of (5.1) where $D = \mathbf{R}^n$, $C = \mathbf{R}^s_- \times \mathbf{R}^{m-s}$, and

$$F(u,x) = (f_1(u,x), \dots, f_m(u,x)).$$

For a given u, the vectors x for which condition (5.2) is fulfilled are precisely the ones at which the constraint system satisfies the Mangasarian-Fromovitz constraint qualification. For such u and x, Theorem 5.4 tells us in particular that

$$\lim_{\substack{\omega' \to \omega \\ t \downarrow 0}} [G(u + t\omega') - x]/t = D(\omega) \text{ for all } \omega \in \mathbb{R}^d, \tag{5.23}$$

where $D(\omega)$ is the set of all $\xi \in \mathbb{R}^n$ satisfying the linearized system

$$\begin{aligned} \nabla_u f_i(u,x)\omega + \nabla_x f_i(u,x)\xi &\leq 0 \text{ for all } i \in I(u,x), \\ &= 0 \text{ for } i = s+1, \dots, m, \end{aligned}$$

in which $I(u,x)$ denotes the indices of the inequality constraints in (5.22) that are active at x, i.e. the indices $i \in \{1, \dots, s\}$ such that $f_i(u,x) = 0$.

We turn now to the type of multifunction that corresponds to optimality conditions and the like. For motivation, let us recall that the relation

$$x \in D, \ -F(x) \in N_D(x), \tag{5.24}$$

where D is a closed convex set in $\mathbf{R}^n$ and F a mapping from $\mathbf{R}^n$ into itself, is a so-called *variational inequality*, or in the terminology of Robinson, a *generalized equation*. First-order optimality conditions of all sorts in convex and nonconvex programming can be put into this form, usually with F smooth and D polyhedral. Here x could stand for a vector of primal variables or it could be comprised of both primal and dual variables. In the latter case, F would be obtained from the gradient mapping associated with a certain Lagrangian function, and (5.24) would represent "Kuhn-Tucker conditions". There is too much to say here for the confines of this paper. We refer the reader to the representations described in Robinson [17], [18].

Theorem 5.6. *Let $G : \mathbf{R}^d \times \mathbf{R}^n \rightrightarrows \mathbf{R}^n$ have the form*

$$G(u,z) = \{x \in D \mid -F(u,x) + z \in N_D(x)\}. \tag{5.25}$$

where D is a polyhedral convex set in $\mathbf{R}^n$ and $F : \mathbf{R}^d \times \mathbf{R}^n \to \mathbf{R}^n$ is a differentiable mapping. Consider any $(u,z) \in \operatorname{dom} G$ and $x \in G(u,z)$. Then G is proto-differentiable at (u,z) relative to x with

$$G'_{(u,z),x}(\omega,\varsigma) = \{\xi \in D'(u,z,x) \mid -\nabla_u F(x,u)\omega - \nabla_x F(u,x)\xi + \varsigma \in N_{D'(u,z,x)}(\xi)\}, \tag{5.26}$$

where

$$D'(u,z,x) = \{\xi \in T_D(x) \mid \xi \cdot [z - F(u,x)] = 0\}. \tag{5.27}$$

Proof. We make a notational maneuver similar to the one in the proof of Theorem 5.4 and introduce

$$H(u,z) = \{(v,x) \in \tilde{D} \mid (u,z) - \tilde{F}(v,x) \in N_{\tilde{D}}(v,x)\}, \tag{5.28}$$

where $\tilde{F}$ and $\tilde{D}$ again are given by (5.9) and (5.11). Then in terms of

$$A(u,z) = H(u,z) + g(u,z), \quad \text{with } g(u,z) = (-u,0) \tag{5.29}$$

we have

$$A(u,z) = \{(0,x) \mid x \in G(u,z)\}. \tag{5.30}$$

In order to demonstrate that G is proto-differentiable with the formula (5.26), it is enough to demonstrate that A is proto-differentiable with the obviously corresponding formula. By applying Proposition 5.2 to (5.29), we see this amounts to showing that H is proto-differentiable with

$$\begin{aligned} H'_{(u,z),(u,x)}(\omega,\varsigma) = \{(\omega,\varsigma) \in \mathbf{R}^d \times D'(u,z,x) \\ - \nabla_u F(u,x)\omega - \nabla_x F(u,x)\xi + \varsigma \in N_{D'(u,z,x)}(\xi)\}. \end{aligned} \tag{5.31}$$

Define the multifunction $S : \mathbf{R}^d \times \mathbf{R}^n \rightrightarrows \mathbf{R}^d \times \mathbf{R}^n$ by

$$S(v,x) = \begin{cases} N_{\tilde{D}}(v,x) & \text{if } (v,x) \in \tilde{D}, \\ \emptyset & \text{if } (v,x) \notin \tilde{D}. \end{cases} \tag{5.33}$$

Formula (5.28) is equivalent to

$$H^{-1} = \tilde{F} + S. \tag{5.34}$$

We wish to apply Proposition 5.2 to $\tilde{F} + S$ and then return to H by Proposition 5.1. The crucial task is the verification of the proto-differentiability of S. We shall accomplish this through the graphical approach in Proposition 2.2.

From the formula $\tilde{D} = \mathbf{R}^d \times D$, it is clear that

$$\operatorname{gph} S = \{(v, x, w, p) \mid v \in \mathbf{R}^d,\ x \in D,\ w = 0,\ p \in N_D(x)\}. \tag{5.35}$$

We must show that this set is approximable everywhere, i.e. that

$$\lim_{t \downarrow 0} t^{-1}[(\operatorname{gph} S) - (v, x, w, p)] \text{ exists for all } (v, x, w, p) \in \operatorname{gph} S.$$

The description in (5.35) reveals that the set

$$M = \{(x, p) \in \mathbf{R}^n \times \mathbf{R}^n | x \in D,\ p \in N_D(x)\} \tag{5.36}$$

is the key. We need to establish that at each $(x, p) \in M$ the contingent cone and derivative cone coincide, and we further need eventually, for the sake of the calculation of the proto-derivative of H, an expression for this common cone.

The set M is the graph of $\partial \delta_D$, the subdifferential of the indicator function δ_D for the polyhedral convex set $D \subset \mathbf{R}^n$. Robinson has shown in [19, Proposition 3] that such a subdifferential is *polyhedral*, which means that M is the union of a finite collection of polyhedral convex sets in $\mathbf{R}^n \times \mathbf{R}^n$. Any polyhedral convex set is, of course, everywhere approximable, the tangent cone in the sense of convex analysis serving both as the contingent cone and the derivative cone. Let M be expressed as the union of polyhedral convex sets M_j, $j = 1, \ldots, q$. For any $(x, p) \in M$, let $J(x, p)$ denote the set of indices j such that $(x, p) \in M_j$. Then

$$\begin{aligned} \lim_{t \downarrow 0} t^{-1}[M - (x, p)] &= \bigcup_{j \in J(x,p)} \lim_{t \downarrow 0} t^{-1}[M_j - (x, p)] \\ &= \bigcup_{j \in J(x,p)} T_{M_j}(x, p). \end{aligned} \tag{5.37}$$

The pairs (ξ, π) belonging to this set are the ones such that for some $\tau > 0$, one has $(x, p) + t(\xi, \pi) \in M$ for all $t \in [0, \tau)$. Thus M is approximable at (x, p) for any $x \in D$ and $p \in N_D(x)$, and the corresponding cone is

$$\{(\xi, \pi) \mid \exists \tau > 0 \text{ with } x + t\xi \in D \text{ and } p + t\pi \in N_D(x + t\xi) \text{ for all } t \in [0, \tau)\}. \tag{5.38}$$

The polyhedral nature of D implies that no matter what the choice of $x \in D$ and $\xi \in \mathbf{R}^n$, the cone $N_D(x + t\xi)$ will be constant relative to t in some sufficiently small interval $(0, \tau)$. In fact $N_D(x + t\xi) = K(x, \xi)$ for small $t > 0$, where

$$K(x, \xi) := \{q \in N_D(x) \,|\, q \cdot \xi = 0\}, \text{ where } \xi \in T_D(x).$$

This set is a polyhedral convex cone. If we have $p + t\pi \in K(x, \xi)$ for all sufficiently small $t > 0$ as in (5.38), this means that

$$p \in K(x, \xi) \text{ and } \pi \in T_{K(x,\xi)}(p) = T_{N_D(x)}(p) \cap \xi^{\perp}, \tag{5.39}$$

where $\xi^{\perp}$ is the set of all vectors orthogonal to ξ. Using the fact that $N_D(x)$ is a polyhedral convex cone containing p, we obtain

$$T_{N_D(x)}(p) = \{q + \lambda p \,|\, q \in N_D(x),\ \lambda \in \mathbf{R}\}. \tag{5.40}$$

Since $N_D(x)$ and $T_D(x)$ are polar to each other, the polyhedral convex cone in (5.40) is polar to

$$C(x, p) := \{\xi' \in T_D(x) \,|\, p \cdot \xi' = 0\}. \tag{5.41}$$

It follows that the conditions (5.39) are equivalent to

$$\xi \in C(x, p), \ \pi \cdot \xi = 0, \ \text{ and } \pi \cdot \xi' \leq 0 \text{ for all } \xi' \in C(x, p),$$

or in other words

$$\xi \in C(x, p) \text{ and } \pi \in N_{C(x,p)}(\xi).$$

This shows that the cone (5.38) is identical to

$$\{(\xi, \pi) \,|\, \xi \in C(x, p),\ \pi \in N_{C(x,p)}(\xi)\}. \tag{5.42}$$

Recalling that this was the cone

$$\lim_{t \downarrow 0} t^{-1}[M - (x, p)],$$

where M is given by (5.36), we are able to conclude in the notation of the multifunction S in (5.33) and (5.35), given by

$$S(v, x) = \begin{cases} \{0\} \times N_D(x) & \text{if } x \in D, \\ \emptyset & \text{if } x \notin D, \end{cases}$$

that S is proto-differentiable at any $(v,x) \in \operatorname{dom} S$ relative to any element of $S(v,x)$, and

$$S'_{(v,x),(0,p)}(\theta,\xi) = \begin{cases} \{(0,\pi) \mid \pi \in N_{C(x,p)}(\xi)\} & \text{if } \xi \in C(x,p), \\ \emptyset & \text{if } \xi \notin C(x,p). \end{cases} \tag{5.43}$$

We are prepared now to return to the calculation of proto-derivatives of H^{-1} in (5.34). We want to do this for (u,z) and (v,x) satisfying $(u,z) \in H^{-1}(v,x)$, which in terms of (5.34) requires

$$(v,x) \in \operatorname{dom} S \text{ and } (u,z) - \tilde{F}(v,x) \in S(v,x). \tag{5.44}$$

These conditions reduce by (5.43) and (5.9) to

$$x \in D, \quad v = u, \quad (u,z) - \tilde{F}(v,x) = (0,p),$$

where

$$p = z - F(u,x) \in N_D(x). \tag{5.45}$$

For such elements Proposition 5.2 conveys the information that H^{-1} is proto-differentiable at (u,x) relative to (u,z) with

$$(H^{-1})'_{(u,x),(u,z)}(\theta,\xi) = \nabla\tilde{F}(u,x)(\theta,\xi) + S'_{(u,x),(0,p)}(\theta,\xi). \tag{5.46}$$

Here

$$\nabla\tilde{F}(u,x)(\theta,\xi) = (\theta, \nabla_u F(u,x)\theta + \nabla_x F(u,x)\xi),$$

while the proto-derivative of S is given by (5.43). When p has the form given by (5.45), the set $C(x,p)$ in (5.43), which was defined in (5.41), becomes the set $D'(u,z,x)$ in (5.27). Thus from (5.46) we have

$$\begin{aligned} &(\omega,\varsigma) \in (H^{-1})'_{(u,x),(u,z)}(\theta,\xi) \Leftrightarrow \\ &\qquad \omega = \theta,\ \xi \in D'(u,z,x), \text{ and } \varsigma - \nabla_u F(u,x)\theta - \nabla_x F(u,x)\xi \in N_{D'(u,z,x)}(\xi). \end{aligned}$$

Invoking Proposition 5.1 we conclude that $H'_{(u,z),(u,x)}$ exists and is given by (5.31). This was all we needed to show to wind up the proof. □

In comparing Theorems 5.4 and 5.5, the reader may be struck by the fact that 5.4 gets away with a general parameterization in terms of u, while 5.5 has z as well as u. The introduction of z does not, of course, add new possibilities for parameterization in (5.25) than could already be handled by u. Rather this is a sort of *restriction* in the formulation:

we are requiring at the minimum that *all* the perturbations of the form z are present, in addition to which we allow arbitrary perturbations of the form u. The result we then obtain in terms of $G(u, z)$ is in truth *more special* than a result simply for

$$G_0(u) = G(u, 0), \tag{5.47}$$

which we do not know how to establish at present in such a framework without severe restrictions of other kinds.

The difficulty, of course, is that $G(u, z)$ is not necessarily *pseudo-Lipschitzian*. If it were, we could apply Proposition 5.3, pass to the context of (5.47) and obtain a better result. To make G be pseudo-Lipschitzian at $(u, 0)$, we would in particular (because of the nature of variational inequalities/generalized equations) have to make G be single-valued at $(u, 0)$, which we prefer to avoid. See Rockafellar [16, p. 876–877], however, for more on this possible approach.

Theorem 4.5 can be compared with the various results of Robinson [16], [17] on multifunctions of the form (5.25) or (5.47). These results are complementary. Robinson assumes only the continuity of F in u (he is able to avoid the introduction of z) and works with the linearization of F in x alone. He aims at deducing bounds for the behavior of G relative to this linearization, especially bounds of Lipschitz type that are based on verifiable assumptions about the properties of the linearization. There is no attempt in his work to perform any kind of differentiation of G with respect to u.

Theorem 4.5 does concern differentiation, for which it provides exact formulas. In this way one also obtains estimates and approximations for G, by way of Propositions 2.5, 4.1 and 4.4. But these are generally different in nature from Robinson's. The closest is the estimate in condition (c) of Theorem 4.1, which is an upper Lipschitz property of the sort Robinson treats, but Theorem 4.1 can not come into play unless x is an isolated point of $G(u)$, which is something Robinson does not need to suppose.

This subject is, of course, still in the making. One can hope in the future for a better understanding of how bounds on proto-derivatives might provide other kinds of estimates.

References.

1. J.P. Aubin, "Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions", Mathematical Analysis and Applications, Part A: Advances in Math. Supplementary Studies, Vol. **7A** (Academic Press, 1981), 159–229.
2. J.P. Aubin, "Lipschitz behavior of solutions to convex minimization problems", Math. of Op. Research **9**(1984), 87–111.
3. J.P. Aubin and I, Ekeland, *Applied Nonlinear Analysis*, Wiley-Interscience, 1984.
4. J.P. Aubin and A. Cellina, *Differential Inclusions*, Springer-Verlag, 1984.
5. H. Frankowska, "Inclusions adjointes associées aux trajectoires minimales d'inclusions différentielles", C.R. Acad. Sc. Paris **297**(1983), 461–464.
6. H. Frankowska, "Adjoint differential inclusions in necessary conditions for the minimal trajectories of differential inclusions", Ann. Inst. H. Poincaré: Analyse Non Linéaire **2**(1985), 75–99.
7. H. Frankowska, "Local controllability and infinitesimal generators of semigroups of set-valued maps", SIAM J. Control Opt. **25**(1987). 412–432.
8. F.H. Clarke, "Generalized gradients and applications", Trans. Amer. Math. Soc. **205**(1975), 247-262.
9. F.H. Clarke, *Nonsmooth Analysis and Optimization*, Wiley-Interscience, 1983.
10. R.T. Rockafellar, "First and second-order epi-differentiability in nonlinear programming", Trans. Amer. Math. Soc., to appear.
11. R.T. Rockafellar, "Maximal monotone relations and the second derivatives of nonsmooth functions", Ann. Inst. H. Poincaré: Analyse Non Linéaire **2**(1985), 167–184.
12. R.T. Rockafellar, "Generalized second derivatives of convex functions and saddle functions", forthcoming.
13. G. Bouligand, *Introduction à la Géométrie Infinitesimale Directe*, Gauthier-Villars, Paris. 1932.
14. A. Shapiro, "Second-order sensitivity anlysis and asymptotic theory of parameterized nonlinear programs", Math. Prog. **33**(1985), 280–299.
15. R.T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, 1970.
16. R.T. Rockafellar, "Lipschitzian properties of multifunctions", Nonlin. Anal. Th. Math. Appl. **9**(1985), 867–885.
17. S.M. Robinson, "Generalized equations and their solution, part I: basic theory", Math. Programming Study **10**(1979). 128–141.

18. S.M. Robinson, "Generalized equations and their solutions, part II: applications to nonlinear programming", Math. Programming Study **19**(1982), 200–221.

19. S.M. Robinson, "Some continuity properties of polyhedral multifunctions", Math. Programming Study **14**(1981), 206–214.

20. C. Ursescu, "Tangent sets' calculus and necessary conditions for extremality", SIAM J. Control Opt. **20**(1982), 563–574.

Imprimerie GAUTHIER-VILLARS, France
Dépôt légal, Imprimeur, n° 3177
Dépôt légal : août 1989

Imprimé en France