# GENERALIZED DIFFERENTIATION WITH POSITIVELY HOMOGENEOUS MAPS: APPLICATIONS IN SET-VALUED ANALYSIS AND METRIC REGULARITY

C.H. JEFFREY PANG

Abstract. We propose a new concept of generalized differentiation of set-valued maps that captures first order information. This concept encompasses the standard notions of Fréchet differentiability, strict differentiability, calmness and Lipschitz continuity in single-valued maps, and the Aubin property and Lipschitz continuity in set-valued maps. We present calculus rules, sharpen the relationship between the Aubin property and coderivatives, and study how metric regularity and open covering can be refined to have a directional property similar to our concept of generalized differentiation. Finally, we discuss the relationship between the robust form of generalization differentiation and its one sided counterpart.

## Contents



## 1. Introduction

We say that $S$ is a *set-valued map* (or a *multi-valued function*, or *multifunction*) if for all $x \in X$, $S(x)$ is a subset of $Y$, and we denote set-valued maps by $S : X \rightrightarrows Y$. Many problems of feasibility, control, optimality and equilibrium are set-valued in nature, and are best treated with methods in set-valued analysis. The texts [4, 5, 21] contain much of the theory of set-valued analysis. Set-valued analysis serves as a foundation for the theory of differential inclusions [3, 2], control theory [13] and variational analysis [35, 27, 12], which in turn have many applications in applied

*Date*: February 16, 2010.

2000 *Mathematics Subject Classification.* 26E25, 46G05, 46T20, 47H04, 49J50, 49J52, 49J53, 54C05, 54C50, 54C60, 58C06, 58C07, 58C20, 58C25, 90C31

*OR/MS classification words*: Mathematics (Functions, Sets).

*Key words and phrases.* multi-function differentiability, metric regularity, coderivatives, calmness, Lipschitz continuity.

mathematics. We refer to these texts for the abundant bibliography on the history of set-valued analysis.

The main contribution of this paper is to introduce a new concept of generalized differentiation (Definitions 4.1 and 4.6) using positively homogeneous maps. Any reasonable definition of a derivative for set-valued maps has to describe changes in the set in terms of the input variables. Using the Pompieu-Hausdorff distance (a metric on the space of compact sets), one obtains the classical definition of Lipschitz continuity of set-valued maps. The concept introduced in this paper provides a more precise tool that incoporates first order information in a set-valued map, encompassing the standard notions of Fréchet differentiability, strict differentiability, calmness and Lipschitz continuity in single-valued maps, and the Aubin property and Lipschitz continuity in set-valued maps. We illustrate how this new concept relates to, and extends, existing methods in variational and set-valued analysis. To motivate our discussion, we revisit the relation between the Clarke subdifferential and Clarke Jacobian and the nonsmooth behavior of functions that can be traced back to [17].

Other than the first four sections which provide the necessary background for the rest of the paper, the last four sections can be read in any order. In Section 3, we recall generalized differentiation for single-valued functions, which was first proposed by Ioffe [16, 17]. We relate the generalized derivative to common notions in classical and variational analysis, paying particular attention to the Clarke subdifferential and the Clarke Jacobian. In Section 4, we define generalized differentiation for set-valued maps, and illustrate the lack of relation between our generalized derivatives and the notions of set-valued derivatives based on the tangent cones, namely semidifferentiability [29] and proto-differentiability [34]. We present calculus rules in Section 5.

The Aubin property (See Definition 4.7), which is commonly attributed to [1], is a method of analyzing local Lipschitz continuity of set-valued maps. In Section 6, we revisit the classical relationship between the Aubin property and the coderivatives of a set-valued map. This relationship is referred to as the Mordukhovich criterion in [35]. Since the coderivatives of a set-valued map can be calculated in many applications and enjoy an effective calculus, this relationship is an important tool in the study of the Lipschitz properties of set-valued maps. We will show that the coderivatives actually gives more information on the local Lipschitz continuity property in our language of generalized derivatives.

It is well known that the Aubin property is related to metric regularity and open covering [6, 26, 30]. Open covering is sometimes known as linear openness. Metric regularity is important in the analysis of solutions to $\bar{y} \in S(\bar{x})$, while open covering studies local covering properties of a set-valued map. Both metric regularity and open covering can be viewed as a study of set-valued maps whose inverse have the Aubin property. For more on metric regularity, we refer the reader to [35, 27, 19, 20]. In Section 7, we take a new look at metric regularity and open covering in view of our definitions of the generalized derivatives. We study metric regularity and open covering in a much broader framework, illustrating that a directional behavior similar to that in our definition of generalized derivatives is present in metric regularity and open covering.

In Section 8, we discuss how the (basic and strict) generalized derivatives defined in Sections 3 and 4 relate to each other. As particular cases, we obtain an equivalent criterion for strict differentiability of set-valued maps, and a relationship between calmness and Lipschitz continuity in both single-valued and set-valued maps. As far as we are aware, the relation between calmness and Lipschitz continuity in set-valued maps was first discussed in [24, 32].

## 2. Preliminaries and notation

Throughout this paper, we shall assume that $X$ and $Y$ are Banach spaces. In most cases, we follow the notation of [35]. Given two sets $A, B \subset X$, the notation $A+B$ stands for the *Minkowski sum* of two sets, defined by

$$A+B := \{a+b \mid a \in A, b \in B\}.$$

The notation $A-B$ is interpreted as $A+(-B)$. We use $\langle \cdot,\cdot \rangle : X^* \times X \to \mathbb{R}$, where $\langle \zeta, x\rangle := \zeta(x)$, to denote the usual dual relation. In Hilbert spaces (and hence in $\mathbb{R}^n$), $\langle \cdot,\cdot \rangle$ reduces to the usual inner product. The notation $x \underset{D}{\longrightarrow} \bar{x}$ means that $x \to \bar{x}$, with sequences lying in $D \subset X$. The closed ball with center $x$ and radius $r$ is denoted by $\mathbb{B}(x,r)$, while $\mathbb{B}$ denotes the ball with center $\mathbf{0}$ and radius 1.

We say that $S$ is a *set-valued map* from $X$ to $Y$ if $S(x) \subset Y$ for all $x \in X$, and a set-valued map is denoted by $S : X \rightrightarrows Y$. The set-valued map $S$ is *closed-valued* if $S(x)$ is closed for all $x \in X$, and it is *convex-valued* if $S(x)$ is convex for all $x \in X$. A *closed* set-valued map is a map whose graph is closed. We say that $C \subset X$ is a *cone* if $\mathbf{0} \in C$ and $\lambda x \in C$ for all $\lambda > 0$ and $x \in C$.

The *graph* of a set-valued map $\mathrm{gph}(S) \subset X \times Y$ is the set $\{(x,y) \mid y \in S(x)\}$. The set-valued map $S^{-1} : Y \rightrightarrows X$ is defined by $S^{-1}(y) := \{x \mid y \in S(x)\}$, and $\mathrm{gph}(S^{-1}) = \{(y,x) \mid (x,y) \in \mathrm{gph}(S)\}$.

**Definition 2.1.** A set-valued map $T : X \rightrightarrows Y$ is *positively homogeneous* if

$$T(\mathbf{0}) \text{ is a cone, and } T(kw) = kT(w) \text{ for all } k > 0 \text{ and } w \in X.$$

It is clear that $T$ is positively homogeneous if and only if $\mathrm{gph}(T)$ is a cone. If $T_1, T_2 : X \rightrightarrows Y$ are two set-valued maps such that $T_1(w) \subset T_2(w)$ for all $w \in X$, then we write this property as $T_1 \subset T_2$. We denote the set-valued map $T(-\cdot) : X \rightrightarrows Y$ to be $T(-\cdot)(w) := T(-w)$.

We recall the definition of inner limits of a set valued map.

**Definition 2.2.** When $S : X \rightrightarrows Y$ is a set-valued map, we say that

$$\liminf_{x \to \bar{x}} S(x) := \{y \in Y \mid \lim_{x \longrightarrow \bar{x}} d(y, S(x)) = 0\}$$

is the *inner limit* of $S(x)$ when $x \to \bar{x}$.

We recall the definitions of outer and inner semicontinuity.

**Definition 2.3.** For a closed-valued mapping $S : X \rightrightarrows Y$ and a point $\bar{x} \in X$:

(1) $S$ is *upper semicontinuous* at $\bar{x}$ if for any open set $U$ such that $S(\bar{x}) \subset U$, there exists some $\eta > 0$ such that if $|x - \bar{x}| < \eta$, then $S(x) \subset U$.
(2) $S$ is *outer semicontinuous at* $\bar{x}$ if for any open set $U$ such that $S(\bar{x}) \subset U$ and $U \cup \rho\mathbb{B} = X$ for some $\rho > 0$, there exists some $\eta > 0$ such that if $|x - \bar{x}| < \eta$, then $S(x) \subset U$.
(3) $S$ is *inner semicontinuous at* $\bar{x}$ if $S(\bar{x}) \subset \liminf_{x \to \bar{x}} S(x)$.
(4) $S$ is *continuous at* $\bar{x}$ if it is both outer and inner semicontinuous there.

We caution that the terminology used to denote upper semicontinuity and outer semicontinuity is not consistent in the literature.

Outer semicontinuity is better suited to handle set-valued maps with unbounded value sets $S(x)$. For example, the set-valued map $S:\mathbb{R}\rightrightarrows\mathbb{R}^2$ defined by

$$S(\theta):=\{(t\cos\theta,t\sin\theta)\mid t\geq 0\}$$

is not upper semicontinuous anywhere but is outer semicontinuous, and in fact continuous, everywhere. When $S(\bar{x})$ is bounded, upper and outer semicontinuity are equivalent. We will not use upper semicontinuity in this paper.

## 3. Generalized differentiability of single-valued maps

The emphasis of this section is the generalized differentiability of single-valued maps $f:X\to Y$. Much of the theory is already in [17], but we concentrate on the key results that we will extend for the set-valued case in later sections. We now begin with our first definition of generalized differentiability.

**Definition 3.1.** Let $T:X\rightrightarrows Y$ be a set-valued map. We say that $f:X\to Y$ is *$T$-differentiable* at $\bar{x}$ if for any $\delta>0$, there exists a neighborhood $V$ of $\bar{x}$ such that

$$f(x)\in f(\bar{x})+T(x-\bar{x})+\delta|x-\bar{x}|\mathbb{B}\text{ for all }x\in V.$$

We say that $f:X\to Y$ is *strictly $T$-differentiable* at $\bar{x}$ if $T:X\rightrightarrows Y$ is positively homogeneous, and for any $\delta>0$, there exists a neighborhood $V$ of $\bar{x}$ such that

$$f(x)\in f(x')+T(x-x')+\delta|x-x'|\mathbb{B}\text{ for all }x,x'\in V.$$

The map $T:X\rightrightarrows Y$ is referred to as a *prederivative* when $f$ is $T$-differentiable, and as the *strict prederivative* when $f$ is strictly $T$-differentiable in [17, Definition 9.1]. The definition of $T$-differentiability includes the familiar concepts of differentiability and Lipschitz continuity as special cases.

**Example 3.2.** (Examples of $T$-differentiability) Let $f:X\to Y$ be a single-valued map.

(1) When $T:X\to Y$ is a (single-valued) linear map, $T$-differentiability is precisely Fréchet differentiability with derivative $T$, and strict $T$-differentiability is precisely strict differentiability with derivative $T$.
(2) When $T:X\rightrightarrows Y$ is defined by $T(w)=\kappa|w|\mathbb{B}$, $T$-differentiability is precisely calmness with modulus $\kappa$, and strict $T$-differentiability is precisely Lipschitz continuity with modulus $\kappa$.

Strict $T$-differentiability is more robust than $T$-differentiability. The function $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x)=x^2\sin(\frac{1}{x^2})$ is Fréchet differentiable at 0 but not Lipschitz there.

With the right choice of $T$, $T$-differentiability can also handle inequalities.

**Example 3.3.** (Inequalities with $T$-differentiability) Let $f:X\to\mathbb{R}$ be a single-valued map.

(1) $f:X\to\mathbb{R}$ is *calm from below* at $\bar{x}$ with modulus $\kappa$ if and only if $f$ is $T$-differentiable there, where $T:X\rightrightarrows\mathbb{R}$ is defined by $T(w)=[-\kappa|w|,\infty)$.
(2) The vector $v\in X^*$ is a *Fréchet subgradient* of $f$ at $\bar{x}$ if and only if $f$ is $T$-differentiable there, where $T:X\rightrightarrows\mathbb{R}$ is defined by $T(w)=[\langle v,w\rangle,\infty)$.

(3) The convex set $C \subset X^*$ is a subset of the *Fréchet subdifferential* of $f$ at $\bar{x}$ if and only if $f$ is $T$-differentiable there, where $T : X \rightrightarrows \mathbb{R}$ is defined by $T(w) = [\sup_{v \in C} \langle v, w \rangle, \infty)$.

While the map $T : X \rightrightarrows Y$ need not be positively homogeneous, most results in this paper are obtained for this particular case. It is clear from the definitions that if $f$ is $T_1$-differentiable at $\bar{x}$, and $T_2$ satisfies $T_2 \supset T_1$, then $f$ is $T_2$-differentiable at $\bar{x}$ as well.

At this point, we mention connections to other notions of other generalized derivatives for single-valued functions close to the definition of (strict) $T$-differentiability. Semidifferentiability, as is recorded in [35, Definition 7.20], can be traced back to [28], and is equivalent to the case where $T : X \rightrightarrows Y$ for the case when $T$ is continuous and single-valued (see [35, Section 9D]). Semidifferentiability for single-valued maps is not to be confused with semidifferentiability for set-valued maps defined in Definition 4.10.

We now look at a slightly nontrivial example involving the Clarke subdifferential [7].

**Definition 3.4.** [9, Section 2.1] Let $X$ be a Banach space. Suppose $f : X \to \mathbb{R}$ is locally Lipschitz. The *Clarke generalized directional derivative* of $f$ at $\bar{x}$ in the direction $v \in X$ is defined by

$$f^\circ(\bar{x}; v) = \limsup_{t \searrow 0, x \to \bar{x}} \frac{f(x + tv) - f(x)}{t},$$

where $x \in X$ and $t$ is a positive scalar. The *Clarke subdifferential* of $f$ at $\bar{x}$, denoted by $\partial_C f(\bar{x})$, is the convex subset of the dual space $X^*$ given by

$$\{\zeta \in X^* \mid f^\circ(\bar{x}; v) \geq \langle \zeta, v \rangle \text{ for all } v \in X\}.$$

The Clarke subdifferential enjoys a mean value theorem. For $C \subset X^*$, define the set $\langle C, w \rangle$ by $\{\langle c, w \rangle \mid c \in C\}$. The following result is due to Lebourg [23], which has since been generalized in other ways.

**Theorem 3.5.** [23] *(Nonsmooth mean value theorem) Suppose $x_1, x_2 \in X$ and $f$ is Lipschitz on an open set containing the line segment $[x_1, x_2]$. Then there exists a point $u \in (x_1, x_2)$ such that*

$$f(x_2) - f(x_1) \in \langle \partial_C f(u), x_2 - x_1 \rangle.$$

We show how the Clarke subdifferential relates to $T$-differentiability. For extensions and a more general treatment of the following result, we refer the reader to [17, Sections 9, 10]

**Theorem 3.6.** *(Clarke subdifferential and $T$-differentiability) Let $f : X \to \mathbb{R}$ be a Lipschitz function, and $C$ be a convex subset of $X^*$. If $f$ is strictly $T$-differentiable at $\bar{x}$, where $T : X \rightrightarrows \mathbb{R}$ is defined by $T(w) = \langle C, w \rangle$, then $\partial_C f(\bar{x}) \subset C$. The converse holds if the map $x \mapsto \partial_C f(x)$ is outer semicontinuous at $\bar{x}$, which is the case when $X = \mathbb{R}^n$.*

*Proof.* If $f$ is strictly $T$-differentiable at $\bar{x}$, then for any direction $h \in X$, the Clarke directional derivative is

$$\begin{aligned}\max_{v\in\partial f(\bar{x})}\langle v,h\rangle &= f^{\circ}(\bar{x};h)\\ &= \limsup_{x\to\bar{x},t\searrow 0}\frac{f(x+th)-f(x)}{t}\\ &\leq \max_{v\in C}\langle v,h\rangle .\end{aligned}$$

Suppose on the contrary that $v' \in \partial_C f(\bar{x})\backslash C$. Then there exists a direction $h'$ and some $\alpha \in \mathbb{R}$ such that $\langle v', h'\rangle > \alpha$ but $\max_{v\in C}\langle v, h'\rangle \leq \alpha$. This is a contradiction, which shows that $\partial_C f(\bar{x}) \subset C$.

We now prove the converse. It is well-known that when $X = \mathbb{R}^n$, the Clarke subdifferential mapping of a Lipschitz function is outer semicontinuous. See [35] for example. Suppose that $\partial_C f(\bar{x}) \subset C$. By outer semicontinuity, given any $\delta > 0$, there is some $\epsilon > 0$ such that $\partial_C f(x) \subset C + \delta\mathbb{B}$ for all $x \in \mathbb{B}(\bar{x}, \epsilon)$. Theorem 3.5 states that for any points $x_1, x_2 \in \mathbb{B}(\bar{x}, \epsilon)$, there is an $x' \in (x_1, x_2)$ such that

$$f(x_1) - f(x_2) \subset \langle \partial_C f(x'), x_1 - x_2\rangle .$$

This immediately implies that $f(x_1) \in f(x_2) + \langle C, x_1 - x_2\rangle + \delta|x_1 - x_2|\mathbb{B}$, and hence strict $T$-differentiability. □

It is well known that in finite dimensions, the Clarke subdifferential is the limit of gradients taken over where the function is differentiable. We now recall Rademacher's theorem.

**Theorem 3.7.** *(Rademacher's Theorem) Let $O \subset \mathbb{R}^n$ be open, and let $f : O \to \mathbb{R}^m$ be Lipschitz. Let $D$ be the subset of $O$ consisting of the points where $F$ is differentiable. Then $O\backslash D$ is a set of measure zero in $\mathbb{R}^n$, In particular, $D$ is dense in $O$, i.e.,* $\operatorname{cl} D \supset O$.

Closely related to the Clarke subdifferential is the Clarke Jacobian that was first introduced in [8].

**Definition 3.8.** Let $f : O \to \mathbb{R}^m$ be Lipschitz, with $O \subset \mathbb{R}^n$ open, and let $D \subset O$ consist of the points where $f$ is differentiable. The *Clarke Jacobian* at $\bar{x}$ is defined by

$$\bar{\nabla} f(\bar{x}) := \operatorname{conv}\{A \in \mathbb{R}^{m\times n} : \exists x_i \to \bar{x} \text{ with } x_i \in D, \nabla f(x_i) \to A\}.$$

It is clear from Rademacher's Theorem that the Clarke Jacobian is a nonempty, compact set of matrices. If $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz, then it is well-known that the Clarke Jacobian reduces to the Clarke subdifferential. The following result is equivalent to [17, Proposition 10.9], and is a generalization of a result that is well known for $m = 1$.

**Theorem 3.9.** *(Clarke Jacobian and $T$-differentiability) Let $f : O \to \mathbb{R}^m$ be Lipschitz on an open set $O \subset \mathbb{R}^n$, and let $D$ be the subset of $O$ where $f$ is differentiable. Let $D' \subset D$ be such that $D\backslash D'$ is of measure zero. At each $\bar{x} \in O$, $f$ is $T$-differentiable at $\bar{x}$, where $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is defined by*

$$T(w) := \{Aw \mid A \in \bar{\nabla} f(\bar{x})\}.$$

*Proof.* In the case $O = \mathbb{R}^n$, a useful equivalent condition for a set $C \subset \mathbb{R}^n$ to be of measure zero is this: with respect to any vector $w \neq \mathbf{0}$, $C$ is of measure zero if and only if the set $\{\tau \mid x + \tau w \in C\} \subset \mathbb{R}$ is of measure zero (in the one-dimensional sense) for all $x$ outside a set of measure zero.

When $x \in O$ and $\tau$ is small enough that the line segment from $x$ to $x + \tau w$ lies in $O$, the function $\varphi(t) = f(x + tw)$ is Lipschitz continuous for $t \in [0, \tau]$. Lipschitz continuity guarantees that $\varphi(\tau) = \varphi(0) + \int_0^\tau \varphi'(t)dt$, and in this integral a negligible set of $t$ values can be disregarded. Thus, the integral is unaffected if we concentrate on $t$ values such that $x + tw \in D'$, in which case $\varphi'(t) = \nabla f(x + tw)(w)$.

For any $\epsilon > 0$, we can reduce $O$ so that whenever $A = \nabla f(x)$ at some $x \in O \backslash D'$, then $A \in \bar{\nabla} f(\bar{x}) + \epsilon\mathbb{B}$. We have

$$\begin{aligned} f(x + \tau w) &= \varphi(\tau) \\ &= \varphi(0) + \int_0^\tau \varphi'(t)dt \\ &= f(x) + \int_0^\tau \nabla f(x + tw)(w)dt \\ &\subset f(x) + \int_0^\tau T(w) + \epsilon|w|\mathbb{B}dt \\ &\subset f(x) + T(\tau w) + \epsilon|\tau w|\mathbb{B}. \end{aligned}$$

The case where $\nabla f(x+tw)$ does not exist for all $t$ can be treated easily by perturbing $x$. This establishes the $T$-differentiability of $f$. □

In Theorem 3.9, it is clear that there can be no closed convex valued positively homogeneous map $T' \subsetneq T$ such that $f$ is $T'$-differentiable at $\bar{x}$.

For convenience, we define strict $T$-differentiability on a set $U$ as follows.

**Definition 3.10.** We say that $f : X \to Y$ is *strictly $T$-differentiable on $U \subset X$* if for all $x$, $x' \in U$, we have

$$f(x) \in f(x') + T(x - x').$$

Given any map $T : X \rightrightarrows Y$ and constant $\delta > 0$, we shall denote $(T + \delta) : X \rightrightarrows Y$ to be the map

$$(T + \delta)(w) := T(w) + \delta|w|\mathbb{B}.$$

It is clear from the definitions that $f$ is strictly $T$-differentiable at $\bar{x}$ if and only if for any $\delta > 0$, we can find an open neighborhood $U_\delta$ of $\bar{x}$ such that $f$ is strictly $(T + \delta)$-differentiable on $U_\delta$.

For single-valued maps, $T$-differentiability behaves well under intersections.

**Proposition 3.11.** *(Intersections) Let $f : X \to Y$.*

1. *If $f : X \to Y$ is both strictly $T_1$-differentiable and strictly $T_2$-differentiable on an open set $U \subset X$, then $f$ is strictly $(T_1 \cap T_2)$-differentiable on $U$.*
2. *If $f : X \to Y$ is $T_1$-differentiable and $T_2$-differentiable at $\bar{x} \in X$, then $f$ is $(T_1 \cap T_2)$-differentiable at $\bar{x}$. An analogous statement holds for strict $T$-differentiability.*

A function strictly $T$-differentiable at a point has added structure.

**Proposition 3.12.** *($T = -T(-\cdot)$ in strict $T$-differentiability) Let $f : X \to Y$ be a single-valued function that is strictly $T$-differentiable on an open set neighborhood*

*$U \subset X$. Then $f$ is strictly $(T\cap -T(-\cdot))$-differentiable on $U$. Hence we can assume that $T$ satisfies $T = -T(-\cdot)$.*

*Similarly, $f$ is strictly $T$-differentiable at $\bar{x}$ implies $f$ is $(T\cap -T(-\cdot))$-differentiable there. We can again assume $T = -T(-\cdot)$.*

*Proof.* For any $x, x' \in U$, we have $f(x) - f(x') \in T(x - x')$. Reversing the role of $x$ and $x'$ gives $f(x) - f(x') \in -T(x' - x)$. This gives $f(x) - f(x') \in T(x - x') \cap -T(-(x - x'))$, which gives the required conclusion. The conclusion $T = -T(-\cdot)$ is also elementary. The statement on $T$-differentiability at a point has a similar proof. □

We refer the reader to Section 5 for calculus of $T$-differentiable functions, and to Corollary 8.4 for the relation between $T$-differentiability and strict $T$-differentiability. For more information on $T$-differentiability, we refer the reader to [17], especially Sections 9 and 10 there.

## 4. Generalized differentiability of set-valued maps

In this section, we move on to define the generalized differentiability of set-valued maps and state some basic properties. Here is the first definition of the differentiability of a set-valued map.

**Definition 4.1.** (a) Let $T : X \rightrightarrows Y$ be a set-valued map. We say that $S : X \rightrightarrows Y$ is *outer $T$-differentiable* at $\bar{x}$ if for any $\delta > 0$, there exists a neighborhood $V$ of $\bar{x}$ such that

$$S(x) \subset S(\bar{x}) + T(x - \bar{x}) + \delta|x - \bar{x}|\mathbb{B} \text{ for all } x \in V.$$

It is *inner $T$-differentiable* at $\bar{x}$ if the formula above is replaced by

$$S(\bar{x}) \subset S(x) - T(x - \bar{x}) + \delta|x - \bar{x}|\mathbb{B} \text{ for all } x \in V.$$

It is *$T$-differentiable* at $\bar{x}$ if it is both outer $T$-differentiable and inner $T$-differentiable.

(b) Let $T : X \rightrightarrows Y$ be a positively homogeneous set-valued map. We say that $S : X \rightrightarrows Y$ is *strictly $T$-differentiable* at $\bar{x}$ if for any $\delta > 0$, there exists a neighborhood $V$ of $\bar{x}$ such that

$$S(x) \subset S(x') + T(x - x') + \delta|x - x'|\mathbb{B} \text{ for all } x, x' \in V.$$

It is elementary that if $S : X \rightrightarrows Y$ is a single-valued map $S : X \to Y$, then the definitions of outer $T$-differentiability, inner $T$-differentiability and $T$-differentiability in Definition 4.1(a) coincide.

The case $T(w) := \kappa|w|\mathbb{B}$, where $\kappa \geq 0$ is finite, is well studied in variational analysis. In the definitions of calmness and Lipschitz continuity below, we recall the notation for $\kappa$ commonly used in variational analysis. Calmness was first referred to as "upper Lipschitzian" by Robinson [31], who established this property for polyhedral mappings.

**Definition 4.2.** (a) We say that $S : X \rightrightarrows Y$ is *calm* at $\bar{x}$ if there exists a neighborhood $V$ of $\bar{x}$ and $\kappa \geq 0$ such that

$$S(x) \subset S(\bar{x}) + \kappa|x - \bar{x}|\mathbb{B} \text{ for all } x \in V.$$

The infimum of all constants $\kappa$ is the *calmness modulus*, denoted by $\operatorname{calm} S(\bar{x})$. Calmness at $\bar{x}$ is equivalent to outer $T$-differentiability there, where $T : X \rightrightarrows Y$ is defined by $T(w) := \operatorname{calm} S(\bar{x})|w|\mathbb{B}$.

(b) We say that $S : X \rightrightarrows Y$ is *Lipschitz* at $\bar{x}$ if there exists a neighborhood $V$ of $\bar{x}$ and $\kappa > 0$ such that

$$S(x) \subset S(x') + \kappa|x - x'|\mathbb{B} \text{ for all } x, x' \in V.$$

The infimum of all constants $\kappa$ is the *Lipschitz modulus*, denoted by $\operatorname{lip} S(\bar{x})$. Lipschitz continuity at $\bar{x}$ is equivalent to strict $T$-differentiability there, where $T : X \rightrightarrows Y$ is defined by $T(w) := \operatorname{lip} S(\bar{x})|w|\mathbb{B}$.

Consider the case where there is some $\kappa \geq 0$ such that $T : X \rightrightarrows Y$ satisfies $T(w) \subset \kappa|w|\mathbb{B}$ for all $w \in X$. It follows straight from the definitions that if $S : X \rightrightarrows Y$ is outer $T$-differentiable and maps to compact sets, then it is outer semicontinuous. The same relations hold for inner $T$-differentiability and inner semicontinuity, and for $T$-differentiability and continuity. We remark that strict $T$-differentiability implies strict continuity in the sense of [35, Definition 9.28], but not vice versa. Their difference is analogous to the difference between upper semicontinuity and outer semicontinuity.

We motivate this definition of set-valued differentiability with Proposition 4.4, whose proof is straightforward. We now recall the Pompieu-Hausdorff distance.

**Definition 4.3.** For $C, D \subset X$ closed and nonempty, the *Pompieu-Hausdorff distance* between $C$ and $D$ is the quantity

$$\mathbf{d}(C, D) := \inf\{\eta \mid C \subset D + \eta\mathbb{B}, D \subset C + \eta\mathbb{B}\}.$$

The Pompieu-Hausdorff distance is a metric on compact subsets of $X$. In fact, the motivation of calmness and Lipschitz continuity in Definition 4.2 comes from the Pompieu-Hausdorff distance. To motivate the definition of $T$-differentiability, we note the following result.

**Proposition 4.4.** *(Single-valued $T$-differentiability) Suppose $S : X \rightrightarrows Y$ is closed-valued.*

1. *Let $T : X \to Y$ be a single-valued map. Then $S$ is $T$-differentiable at $\bar{x}$ if and only if for any $\delta > 0$, there is a neighborhood $V$ of $\bar{x}$ such that*
$$\mathbf{d}(S(\bar{x}) + T(x - \bar{x}), S(x)) < \delta|x - \bar{x}| \text{ for all } x \in V.$$
2. *Let $T : X \to Y$ be a single-valued map such that $T = -T(-\cdot)$. Then $S$ is strictly $T$-differentiable at $\bar{x}$ if and only if for any $\delta > 0$, there is a neighborhood $V$ of $\bar{x}$ such that*
$$\mathbf{d}(S(x') + T(x - x'), S(x)) < \delta|x - x'| \text{ for all } x, x' \in V.$$

We now make a remark on the Pompieu-Hausdorff distance that is in the spirit of the main idea in this paper.

*Remark* 4.5. (More precise measurement of sets) We can rewrite the Pompieu-Hausdorff distance as

$$\mathbf{d}(C, D) = \inf\{\eta \mid C \subset D + E_1, D \subset C + E_2, E_1 \subset \eta\mathbb{B}, E_2 \subset \eta\mathbb{B}\}.$$

In certain situations, it might be useful to study sets $E_1$ and $E_2$ for which $C \subset D + E_1$ and $D \subset C + E_2$ instead of just taking them to be $\eta\mathbb{B}$.

As is well-known in set-valued analysis, setting restrictions on the range gives a sharper analysis at the points of interest. We make the following definitions with this in mind.

**Definition 4.6.** Let $S : X \rightrightarrows Y$ be a set-valued map such that $\bar{y} \in S(\bar{x})$.

(1) Let $T : X \rightrightarrows Y$ be a set-valued map. We say that $S$ is *pseudo outer $T$-differentiable* at $\bar{x}$ for $\bar{y}$ if for any $\delta > 0$, there exists neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ such that
$$S(x) \cap W \subset S(\bar{x}) + T(x - \bar{x}) + \delta|x - \bar{x}|\mathbb{B} \text{ for all } x \in V.$$
It is *pseudo inner $T$-differentiable* at $\bar{x}$ for $\bar{y}$ if
$$S(\bar{x}) \cap W \subset S(x) - T(x - \bar{x}) + \delta|x - \bar{x}|\mathbb{B} \text{ for all } x \in V.$$
It is *pseudo $T$-differentiable* at $\bar{x}$ for $\bar{u}$ if it is both pseudo outer $T$-differentiable and pseudo inner $T$-differentiable there.

(2) Let $T : X \rightrightarrows Y$ be a positively homogeneous set-valued map. We say that $S$ is *pseudo strictly $T$-differentiable* at $\bar{x}$ for $\bar{y}$ if for any $\delta > 0$, there exists neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ such that
$$S(x) \cap W \subset S(x') + T(x - x') + \delta|x - x'|\mathbb{B} \text{ for all } x, x' \in V.$$

Again, the case $T(w) := \kappa|w|\mathbb{B}$ is of particular interest. The Aubin property was first introduced as pseudo-Lipschitzness in [1].

**Definition 4.7.** Let $S : X \rightrightarrows Y$ be a set-valued map such that $\bar{y} \in S(\bar{x})$.

(1) We say that $S$ is *calm* at $\bar{x}$ for $\bar{y}$ if there exists neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$, and $\kappa \geq 0$ such that
$$S(x) \cap W \subset S(\bar{x}) + \kappa|x - \bar{x}|\mathbb{B} \text{ for all } x \in V.$$
The infimum of all such constants $\kappa$ is the *calmness modulus*, denoted by $\operatorname{calm} S(\bar{x} \mid \bar{y})$. Calmness is precisely pseudo outer $T$-differentiability, where $T : X \rightrightarrows Y$ is defined by $T(w) := \operatorname{calm} S(\bar{x} \mid \bar{y})|w|\mathbb{B}$.

(2) We say that $S : X \rightrightarrows Y$ has the *Aubin Property* at $\bar{x}$ for $\bar{y}$ if there exists neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$, and $\kappa \geq 0$ such that
$$S(x) \cap W \subset S(x') + \kappa|x - x'|\mathbb{B} \text{ for all } x, x' \in V.$$
The infimum of all such constants $\kappa$ is the *graphical modulus*, denoted by $\operatorname{lip} S(\bar{x} \mid \bar{y})$. The Aubin property is also known as the *pseudo-Lipschitz property* and as the *Lipschitz-like property.* The Aubin property is precisely strict pseudo $T$-differentiability, where $T : X \rightrightarrows Y$ is defined by $T(w) := \operatorname{lip} S(\bar{x} \mid \bar{y})|w|\mathbb{B}$.

We present an example to show that Propositions 3.11 and 3.12 cannot be extended for set-valued maps in a straightforward manner.

**Example 4.8.** (Failure of intersections in set-valued $T$-differentiability) Let $S : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ be defined by $S(x, y) = \{x\} \times \mathbb{R}$. Let $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T_1(x, y) = (x, y)$ and $T_2(x, y) = (x, 0)$. It is clear that $S$ is (pseudo) strictly $T_1$-differentiable and $T_2$-differentiable at all points, but it is not (pseudo) $(T_1 \cap T_2)$-differentiable anywhere.

**Example 4.9.** (Failure of set-valued $-T(-\cdot)$-differentiability) Let $S : \mathbb{R} \rightrightarrows \mathbb{R}$ be defined by $S(x) := (-\infty, x]$, and let $T : \mathbb{R} \rightrightarrows \mathbb{R}$ be defined by
$$T(w) = \begin{cases} \{0\} & \text{if } w \leq 0 \\ [-w, w] & \text{if } w \geq 0. \end{cases}$$

Here, $S$ is strictly $T$-differentiable at $x$ for all $x \in \mathbb{R}$, and it is pseudo strictly $T$-differentiable at $x$ for $x$ for all $x \in \mathbb{R}$. However, $S$ is neither outer (or inner) $-T(-\cdot)$-differentiable anywhere, nor pseudo outer (or inner) $-T(-\cdot)$-differentiable at $x$ for $x$ for any $x \in \mathbb{R}$.

We remark that while inner $T$-differentiability is defined so that Proposition 4.4 holds, this definition of inner $T$-differentiability does not satisfy the property that strict $T$-differentiability implies inner $T$-differentiability in general. It actually implies inner $-T(-\cdot)$-differentiability. The same holds for pseudo strict $T$-differentiability and pseudo inner $T$-differentiability, or more correctly, pseudo inner $-T(-\cdot)$-differentiability. An example where this occurs is the function $S : \mathbb{R} \rightrightarrows \mathbb{R}$ as defined in Example 4.9. Fortunately, inner $T$-differentiability does not play a huge role in this paper.

Much of the current methods for set-valued differentiation are motivated by looking at the tangent cones of the graph of the set-valued map. See the discussion in [4, Chapter 5] on the different forms of set-valued differentiation obtained by taking different kinds of tangent cones of the graph. The notions of semidifferentiability [29] and proto-differentiability [34] are based on this idea. We shall only recall the definition of semidifferentiability. See also the techniques in [4, Chapter 5]. We now point out the lack of relation between pseudo $T$-differentiability and these methods by observing the finite dimensional case.

We remark on the definition of a limit. Let $S : D \rightrightarrows \mathbb{R}^m$ be a set-valued map, where $D \subset \mathbb{R}^n$. If $\bar{x} \in \mathrm{cl}(D)\backslash D$, then we say that $\lim_{x\to\bar{x}} S(x)$ is the value at $\bar{x}$ of a continuous extension of $S$ onto $D \cup \{\bar{x}\}$. Such an extension is unique. For a more complete treatment on limits in finite dimensional set-valued maps, see [35].

**Definition 4.10.** [29] For a set-valued map $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, let $\bar{x} \in \mathrm{dom}(F)$, $\bar{y} \in S(\bar{x})$ and $\bar{w} \in \mathbb{R}^n\backslash\{\mathbf{0}\}$. If the limit

$$\lim_{\tau\searrow 0, w\to\bar{w}} \frac{S(\bar{x}+\tau w)-\bar{y}}{\tau}$$

exists, then we say that it is the *semiderivative* at $\bar{x}$ for $\bar{y}$ and $\bar{w}$. If the semiderivative exists for every vector $\bar{w} \in \mathbb{R}^n\backslash\{\mathbf{0}\}$, then $F$ is *semidifferentiable* at $\bar{x}$ for $\bar{y}$ with derivative $DS(\bar{x} \mid \bar{y}) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ defined by $DS(\bar{x} \mid \bar{y})(\bar{w})$ being equal to the limit defined above.

The definition of semidifferentiability for set-valued maps is not to be confused with the semidifferentiability defined for single-valued maps earlier. Clearly, the limit in the definition above is a cone, and the derivative $DS(\bar{x} \mid \bar{y})$ is a positively homogeneous map. We present an example where $S : \mathbb{R} \rightrightarrows \mathbb{R}^2$ is not pseudo $DS(\bar{x} \mid \bar{y})$-differentiable at $\bar{x}$ for $\bar{y}$.

**Example 4.11.** (Semidifferentiability and $T$-differentiability 1) Consider the set-valued map $S : \mathbb{R} \rightrightarrows \mathbb{R}^2$ defined by

$$S(x) = \begin{cases} \{x\} \times [0,\infty) & \text{if } x \leq 0 \\ \{(\alpha,\sqrt{\alpha}) \mid 0 \leq \alpha \leq x\} \cup \big(\{x\} \times [\sqrt{x},\infty)\big) & \text{if } x \geq 0. \end{cases}$$

This map is semidifferentiable at 0 for $\mathbf{0}$, with semiderivative $DS(0 \mid \mathbf{0}) : \mathbb{R} \rightrightarrows \mathbb{R}^2$ defined by $DS(0 \mid \mathbf{0})(w) = \{0\} \times [0,\infty)$ for all $w \in \mathbb{R}$. However, $S$ is not $DS(0 \mid \mathbf{0})$-differentiable at 0 for $\mathbf{0}$, because at any point in $(0,\alpha) \in \{0\} \times (0,\infty)$, we can find a neighborhood $U_\alpha$ of $(0,\alpha)$ such that $S(x) \cap U_\alpha$ is $[\{x\} \times (-\infty,\infty)] \cap U_\alpha$.

One may expect that if $S$ is $T$-differentiable at $\bar{x}$ for $\bar{y}$, then $DS(\bar{x} \mid \bar{y}) \subset T$. The following example shows that this is not the case.

**Example 4.12.** (Semidifferentiability and $T$-differentiability 2) Consider $S : \mathbb{R} \rightrightarrows \mathbb{R}$ defined by $S(x) = \mathbb{R}$. Clearly $S$ is pseudo $T$-differentiable at $x$ for any $y \in \mathbb{R}$, where $T : \mathbb{R} \to \mathbb{R}$ is defined by $T(w) = 0$ for all $w \in \mathbb{R}$. But the semiderivative $DS(\bar{x} \mid \bar{y}) : \mathbb{R} \rightrightarrows \mathbb{R}$ is equal to $S$, which gives $DS(\bar{x} \mid \bar{y}) \not\subset T$ for all $\bar{x}, \bar{y} \in \mathbb{R}$.

In the finite dimensional case, semidifferentiability implies proto-differentiability [35, Exercise 8.43]. The examples showing the lack of relation between semidifferentiability and pseudo $T$-differentiability apply for proto-differentiability as well.

## 5. Calculus of $T$-differentiability

In this section, we prove a chain rule and a sum rule similar to the coderivative calculus in [35, Section 10H]. Let us first introduce the outer norm of a positively homogeneous set-valued map.

**Definition 5.1.** If $T : X \rightrightarrows Y$ is a positively homogeneous map, then the *outer norm of* $T$ is

$$|T|^+ := \sup_{|x|\leq 1} \sup_{y\in T(x)} |y| = \sup\{|y| \mid y \in T(x), |x| \leq 1\}.$$

Clearly $|T|^+ < \infty$ implies $T(\mathbf{0}) = \{\mathbf{0}\}$. When $T(\mathbf{0}) = \{\mathbf{0}\}$, $|T|^+ = \text{calm}\, T(\mathbf{0} \mid \mathbf{0}) = \text{calm}\, T(\mathbf{0})$.

Here is a chain rule for $T$-differentiable functions.

**Theorem 5.2.** *(Chain rule) Let $F : X \rightrightarrows Y$, $G : Y \rightrightarrows Z$, and $\bar{z} \in G \circ F(\bar{x})$. Suppose the following conditions hold*

1. *$F$ is pseudo outer $T_{\bar{x}\to y}$-differentiable at $\bar{x}$ for $y$ for all $y$ in $G^{-1}(\bar{z}) \cap F(\bar{x})$.*
2. *$G$ is pseudo strictly $T_{y\to\bar{z}}$-differentiable at $y$ for $\bar{z}$ for all $y$ in $G^{-1}(\bar{z}) \cap F(\bar{x})$.*
3. *The set $G^{-1}(\bar{z}) \cap F(\bar{x}) \subset Y$ is compact.*
4. *The map $(x, z) \mapsto G^{-1}(z) \cap F(x)$ is outer semicontinuous at $(\bar{x}, \bar{z})$.*
5. *$\alpha := \sup_{y\in G^{-1}(\bar{z})\cap F(\bar{x})} |T_{\bar{x}\to y}|^+$ is finite.*
6. *For all $y \in G^{-1}(\bar{z}) \cap F(\bar{x})$, $T_{y\to\bar{z}}(\mathbf{0}) = \{\mathbf{0}\}$, and $\beta$ is finite, where*

$$\beta := \sup_{y\in G^{-1}(\bar{z})\cap F(\bar{x})} \text{lip}\, T_{y\to\bar{z}}(\mathbf{0}).$$

*Then $G \circ F$ is pseudo outer $T$-differentiable at $\bar{x}$ for $\bar{z}$, where $T : X \rightrightarrows Z$ is defined by*

$$T := \bigcup_{y\in G^{-1}(\bar{z})\cap F(\bar{x})} T_{y\to\bar{z}} \circ T_{\bar{x}\to y}. \tag{5.1}$$

*The function $G \circ F$ is pseudo strictly $T$-differentiable for $T : X \rightrightarrows Y$ defined in* (5.1) *if in statement* (1)*, $F$ were pseudo strictly $T_{\bar{x}\to y}$-differentiable at $\bar{x}$ for $y$ instead.*

*Proof.* We shall prove only the result for $F$ being pseudo outer $T$-differentiable. The proof for pseudo strict $T$-differentiability is almost exactly the same. Choose any $\delta > 0$. Since (2) holds, for each $y \in G^{-1}(\bar{z}) \cap F(\bar{x})$, we can find some open convex neighborhoods $V'_y$ of $y$ and $W$ of $\bar{z}$ such that

$$G(y') \cap W \subset G(y'') + T_{y\to\bar{z}}(y' - y'') + \delta|y' - y''|\mathbb{B} \text{ for all } y', y'' \in V'_y.$$

Next, since (1) holds, for each $y \in G^{-1}(\bar{z}) \cap F(\bar{x})$, we can find open convex neighborhoods $U$ of $\bar{x}$ and $V_y \subset V'_y$ of $y$ such that

$$F(x) \cap V_y \subset [(F(\bar{x}) \cap V'_y) + T_{\bar{x}\to y}(x-\bar{x}) + \delta|x-\bar{x}|\mathbb{B}] \cap V_y \text{ for all } x \in U.$$

Since $G^{-1}(\bar{z}) \cap F(\bar{x})$ is compact by (3), there are finitely many $y_i \in G^{-1}(\bar{z}) \cap F(\bar{x})$ such that $G^{-1}(\bar{z}) \cap F(\bar{x}) \subset \bigcup_i V_{y_i}$. By taking finitely many intersections if necessary, the neighborhoods $U$ and $W$ can be assumed to be independent of $y_i$.

Let $V := \bigcup_i V_{y_i}$. Since the map $(x,z) \mapsto G^{-1}(z) \cap F(x)$ is outer semicontinuous at $(\bar{x}, \bar{z})$ by (4), we can reduce $U$ and $W$ if necessary so that $G^{-1}(W) \cap F(U) \subset V$. This implies that for any $x \in U$,

$$\begin{aligned} G(F(x)) \cap W &= G(F(x) \cap V) \cap W \\ &= \left( \bigcup_i G\big(F(x) \cap V_{y_i}\big) \right) \cap W. \end{aligned}$$

We have

$$\begin{aligned} & G(F(x) \cap V_{y_i}) \cap W \\ \subset\ & G\big([(F(\bar{x}) \cap V'_{y_i}) + T_{\bar{x}\to y_i}(x-\bar{x}) + \delta|x-\bar{x}|\mathbb{B}] \cap V_{y_i}\big) \cap W \\ \subset\ & G\big(F(\bar{x}) \cap V'_{y_i}\big) + T_{y_i\to\bar{z}}\big(T_{\bar{x}\to y_i}(x-\bar{x}) + \delta|x-\bar{x}|\mathbb{B}\big) \\ & \qquad +\delta\big|T_{\bar{x}\to y_i}(x-\bar{x}) + \delta|x-\bar{x}|\mathbb{B}\big|\mathbb{B} \\ \subset\ & G\circ F(\bar{x}) + T_{y_i\to\bar{z}} \circ T_{\bar{x}\to y_i}(x-\bar{x}) + \delta\beta|x-\bar{x}|\mathbb{B} + \delta(\alpha+\delta)|x-\bar{x}|\mathbb{B} \\ =\ & G\circ F(\bar{x}) + T_{y_i\to\bar{z}} \circ T_{\bar{x}\to y_i}(x-\bar{x}) + \delta(\alpha+\beta+\delta)|x-\bar{x}|\mathbb{B}. \end{aligned} \tag{5.2}$$

We may assume that $\delta$ is small enough so that $\alpha+\beta+\delta$ is bounded from above by some finite constant $\theta$. Choosing $y_i$ over all $i$ gives

$$\begin{aligned} & \bigcup_i G\big(F(x) \cap V_{y_i}\big) \cap W \\ \subset\ & G\circ F(\bar{x}) + \left( \bigcup_i T_{y_i\to\bar{z}} \circ T_{\bar{x}\to y_i}(x-\bar{x}) \right) + \delta\theta|x-\bar{x}|\mathbb{B} \\ \subset\ & G\circ F(\bar{x}) + \left( \bigcup_{y\in G^{-1}(\bar{z})\cap F(\bar{x})} T_{y\to\bar{z}} \circ T_{\bar{x}\to y}(x-\bar{x}) \right) + \delta\theta|x-\bar{x}|\mathbb{B}. \end{aligned}$$

This completes the proof of the theorem. □

If $G : Y \rightrightarrows Z$ were pseudo outer $T$-differentiable, then we can still obtain a result for the case where $F : X \rightrightarrows Y$ is single-valued.

**Proposition 5.3.** *(Chain rule) Let $f : X \to Y$ and $G : Y \rightrightarrows Z$, $\bar{y} = f(\bar{x})$, and $\bar{z} \in G(\bar{y})$. Suppose the following conditions hold*

1. *$f$ is $T_{\bar{x}\to\bar{y}}$-differentiable at $\bar{x}$ for $\bar{y}$.*
2. *$G$ is pseudo outer $T_{\bar{y}\to\bar{z}}$-differentiable at $\bar{y}$ for $\bar{z}$.*
3. *$|T_{\bar{x}\to\bar{y}}|^+$ is finite.*
4. *$T_{\bar{y}\to\bar{z}}(\mathbf{0}) = \{\mathbf{0}\}$, and $\operatorname{lip} T_{\bar{y}\to\bar{z}}(\mathbf{0})$ is finite.*

*Then $G \circ f$ is pseudo outer $T$-differentiable at $\bar{x}$ for $\bar{z}$, where $T : X \rightrightarrows Z$ is defined by $T = T_{\bar{y}\to\bar{z}} \circ T_{\bar{x}\to\bar{y}}$.*

*Proof.* Choose some $\delta > 0$. By condition (1), we can find a neighborhood $U$ of $\bar{x}$ such that

$$f(x) \in f(\bar{x}) + T(x-\bar{x}) + \delta|x-\bar{x}|\mathbb{B} \text{ for all } x \in U.$$

Let $V$ be such that $f(\bar{x}) + T(x-\bar{x}) + \delta|x-\bar{x}|\mathbb{B} \subset V$ for all $x \in U$. By (2), we can shrink $U$ and $V$ if necessary so that there is a neighborhood $W$ of $\bar{z}$ such that

$$G(y) \cap W \subset G(\bar{y}) + T_{\bar{y}\to\bar{z}}(y-\bar{y}) + \delta|y-\bar{y}|\mathbb{B} \text{ for all } y \in V.$$

A calculation similar to (5.2) concludes the proof. □

From the chain rule, we can infer a sum rule.

**Corollary 5.4.** *(Sum rule) Let $S_i : X \rightrightarrows Y$ for $i = 1, \ldots, p$, and $\bar{y} \in \sum_{i=1}^p S_i(\bar{x})$. Define $F : X \rightrightarrows Y^p$ by $F(x) = (S_1(x), \ldots, S_p(x))$, and define $g : Y^p \to Y$ to be the linear map mapping to the sum of the $p$ elements in $Y^p$. Suppose the following conditions hold.*

1. *$S_i$ is pseudo outer $T^i_{\bar{x}\to y_i}$-differentiable at $\bar{x}$ for $y_i$ whenever $y_i \in S_i(\bar{x})$ and $y_1 + \cdots + y_p = \bar{y}$.*
2. *The set $\{(y_1, \ldots, y_p) \mid y_i \in S_i(\bar{x}),\, y_1 + \cdots + y_p = \bar{y}\} \subset Y^p$ is compact.*
3. *The map $(x, y) \mapsto g^{-1}(y) \cap F(x)$ is outer semicontinuous at $(\bar{x}, \bar{y})$.*
4. *$\alpha_i := \sup_{y_i \in \Pi_i(g^{-1}(\bar{z}) \cap F(\bar{x}))} \left|T^i_{\bar{x}\to y_i}\right|^+$ is finite for each $i = 1, \ldots, p$. Here, $\Pi_i : Y^p \to Y$ is the projection onto the ith coordinate.*

*Then $g \circ F : X \rightrightarrows Y$, which is the sum of the maps $S_i$, is $T$-differentiable at $\bar{x}$ for $\bar{y}$, where $T : X \rightrightarrows Y$ is defined by*

$$T := \bigcup_{\substack{y_1 + \cdots + y_p = \bar{y} \\ y_i \in S_i(\bar{x})}} \left( \sum_i T^i_{\bar{x}\to y_i} \right). \tag{5.3}$$

*The function $g \circ F$ is pseudo strictly $T$-differentiable for $T : X \rightrightarrows Y$ defined in (5.3) if in statement (1), $S_i$ were pseudo strictly $T^i_{\bar{x}\to y_i}$-differentiable at $\bar{x}$ for $y_i$ instead.*

*Proof.* For $y_i \in S_i(\bar{x})$ for all $i$, define $T_{(y_1,\ldots,y_p)} : X \rightrightarrows Y^p$ by

$$T_{(y_1,\ldots,y_p)}(w) := \left(T^1_{\bar{x}\to y_1}(w), \ldots, T^p_{\bar{x}\to y_p}(w)\right).$$

Condition (1) implies that the map $F$ is pseudo outer $T_{(y_1,\ldots,y_p)}$-differentiable at $\bar{x}$ for $(y_1, \ldots, y_p)$. We now proceed to apply the chain rule in Theorem 5.2. Since $g$ is a linear function, the conditions for $g$ needed for the chain rule are satisfied. The rest of the conditions in this result are just the appropriate conditions in the chain rule rephrased. The case of pseudo strict $T$-differentiability is similar. □

Note that we have focused on pseudo (outer/ strict) $T$-differentiability so far in this section. The relation between pseudo (strict/ outer/ inner) $T$-differentiability and (strict/ outer/ inner) $T$-differentiability is illustrated by the following theorem. We say that $S : X \rightrightarrows Y$ is *locally compact* around $\bar{x} \in \mathrm{dom}(F)$ if there is a neighborhood $O$ of $\bar{x}$ and a compact set $C \subset Y$ such that $S(O) \subset C$.

**Theorem 5.5.** *($T$-differentiability from pseudo $T$-differentiability) Let $S : D \rightrightarrows Y$ be a closed-valued outer semicontinuous map on a closed domain $D \subset X$. Suppose $S$ is locally compact around $\bar{x} \in D$. Then $S$ is outer $T$-differentiable at $\bar{x}$ if and only if $S$ is pseudo outer $T$-differentiable at $\bar{x}$ for all $\bar{y} \in S(\bar{x})$.*

*An analogous statement holds for (strict/ inner) $T$-differentiability and pseudo (strict/ inner) $T$-differentiability.*

*Remark* 5.6. In fact, the hypothesis of outer semicontinuity in Theorem 5.5 can be weakened: We can assume that $S$ is *closed* at $\bar{x}$: For every $y \notin S(\bar{x})$, there are neighborhoods $U$ of $\bar{x}$ and $V$ of $y$ such that $S(x) \cap V = \emptyset$ for all $x \in U$. If $S$ is outer semicontinuous, then $\mathrm{gph}(S)$ is closed by [4, Proposition 1.4.8], which implies that $S$ is closed at $\bar{x}$. The proof of Theorem 5.5 can be easily adapted from the proof of [27, Theorem 1.42], which traces its roots in the finite dimensional case to [33].

Other calculus rules that are important are the Cartesian product of set-valued maps, and rules for unions. These two operations are simple to formulate and prove. For intersections of set-valued maps, we feel it is more effective to look at the normal cones of the intersections of the graph of the appropriate functions and apply the Mordukhovich criterion. See Section 6. We close this section by referring the reader to [33] for more applications of set-valued chain rules.

## 6. The Mordukhovich criterion

As we have seen in Section 4, the Aubin property gives a sharper analysis for the Lipschitz continuity of set-valued maps. An effective tool for calculating the graphical modulus (for the Aubin property) is the coderivative (Defintion 6.2), which is a generalization of the adjoint linear operator of linear functions. Coderivatives enjoy an effective calculus, and can be easily calculated for set-valued maps whose graphs are defined by smooth maps. The relationship between the Aubin property and the coderivatives is referred to as the Mordukhovich criterion in [35]. For a history of the Mordukhovich criterion, see the bibliography in [35], which in turn cited [10, 18, 22, 25, 37], and also Commentaries 1.4.6–1.4.9, 4.5.2 and 4.5.6 of [27]. The aim of this section is to show that coderivatives in fact give directional behavior that is captured in the language of pseudo strict $T$-differentiability.

We now recall the classical definition of the normal cones and coderivatives in finite dimensions.

**Definition 6.1.** Let $C \subset \mathbb{R}^n$ and $\bar{x} \in C$. A vector $v$ is *normal to $C$ at $\bar{x}$ in the regular sense*, or a *regular normal*, written $v \in \hat{N}_C(\bar{x})$, if

$$\langle v, x - \bar{x} \rangle \leq o(|x - \bar{x}|) \text{ for } x \in C.$$

It is *normal to $C$ at $\bar{x}$ in the general sense*, or simply a *normal* vector, written $v \in N_C(\bar{x})$, if there are sequences $x_i \underset{C}{\to} \bar{x}$ and $v_i \to v$ with $v_i \in \hat{N}_C(x_i)$.

**Definition 6.2.** Consider a mapping $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, a point $\bar{x} \in \mathrm{dom}(S)$ and $\bar{y} \in S(\bar{x})$. The *coderivative at $\bar{x}$ for $\bar{y}$* is the mapping $D^*S(\bar{x} \mid \bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ defined by

$$v \in D^*S(\bar{x} \mid \bar{y})(z) \iff (v, -z) \in N_{\mathrm{gph}(S)}(\bar{x}, \bar{y}).$$

In the case where $S$ is a smooth map, the coderivative $D^*S(\bar{x} \mid S(\bar{x}))$ is the adjoint of the derivative mapping there.

Next, we recall the definition of the regular subdifferential and general subdifferential, which are important in the proof in the main result of this section.

**Definition 6.3.** Consider a function $f : \mathbb{R}^n \to \mathbb{R}$ and a point $\bar{x} \in \mathbb{R}^n$. For a vector $v \in \mathbb{R}^n$,

(a) $v$ is a *regular subgradient* of $f$ at $\bar{x}$, written $v \in \hat{\partial} f(\bar{x})$, if

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle + o(|x - \bar{x}|);$$

(b) $v$ is a *(general) subgradient* of $f$ at $\bar{x}$, written $v \in \partial f(\bar{x})$, if there are sequences $x_i \to \bar{x}$ and $v_i \in \hat{\partial} f(x_i)$ with $v_i \to v$ and $f(x_i) \to f(\bar{x})$.

(c) The sets $\hat{\partial} f(\bar{x})$ and $\partial f(\bar{x})$ are the *regular subdifferential* (or *Fréchet subgradient*) and the *(general) subdifferential* (or *limiting subdifferential*, or *Mordukhovich subdifferential*) at $\bar{x}$ respectively.

We now present our main result of this section.

**Theorem 6.4.** *(Mordukhovich criterion revisited) Consider* $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, $\bar{x} \in \mathrm{dom}(S)$ *and* $\bar{y} \in S(\bar{x})$. *Suppose* $\mathrm{gph}(S)$ *is locally closed at* $(\bar{x}, \bar{y})$. *If* $D^*S(\bar{x} \mid \bar{y})(\mathbf{0}) = \{\mathbf{0}\}$, *or equivalently,* $|D^*S(\bar{x} \mid \bar{y})|^+ < \infty$, *then* $S$ *is pseudo strictly* $T$-*differentiable at* $\bar{x}$ *for* $\bar{y}$, *where* $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *is defined by*

$$T(w) = \kappa(w)\mathbb{B}^m \cup \{\mathbf{0}\},$$

*where* $\kappa : \mathbb{R}^n \to \mathbb{R}_+$ *is defined by*

$$\begin{aligned} \kappa(w) &:= \max\{\langle -v, w \rangle \mid v \in D^*S(\bar{x} \mid \bar{y})(\mathbb{B}^m)\} \\ &= \max\left\{\langle -v, w \rangle \mid (v, -z) \in N_{\mathrm{gph}(S)}(\bar{x}, \bar{y}) \text{ for some } |z| \leq 1\right\}. \end{aligned}$$

The assumptions of Theorem 6.4 are the same as that of the Mordukhovich Criterion as stated in [35, Theorem 9.40], but the conclusion is improved. The Mordukhovich criterion establishes the equivalence between the Aubin property (with graphical modulus $\kappa^* := \max_{|w| \leq 1} \kappa(w)$) and the outer norm of the coderivative. Theorem 6.4 shows that the worst case Lipschitzian behavior need not occur in all directions.

*Proof.* (of Theorem 6.4) We highlight only the parts we need to add to the proof of sufficiency as given in [35, Theorem 9.40], though with some changes to the variable names. Let $\bar{\kappa} = |D^*S(\bar{x} \mid \bar{y})|^+$. By the definition of $\bar{\kappa}$ and the coderivatives, for any $\theta > 0$, there exists $\delta_0 > 0$ and $\epsilon_0 > 0$ for which

$$N_{\mathrm{gph}(S)}(\hat{x}, \hat{y}) \cap (\mathbb{R}^n \times \mathbb{B}^m) \subset (\bar{\kappa} + \theta)\mathbb{B}^n \times \mathbb{B}^m \text{ for all } \hat{x} \in \mathbb{B}(\bar{x}, \delta_0), \hat{y} \in \mathbb{B}(\bar{y}, \epsilon_0).$$

(The above statement is just [35, 9(22)] rephrased.) The mapping to normal cones is outer semicontinuous, so

$$N_{\mathrm{gph}(S)}(\hat{x}, \hat{y}) \cap [(\bar{\kappa} + \theta)\mathbb{B}^n \times \mathbb{B}^m] \subset N_{\mathrm{gph}(S)}(\bar{x}, \bar{y}) + \theta\mathbb{B}^{n+m} \quad \text{for all } \hat{x} \in \mathbb{B}(\bar{x}, \delta_0), \hat{y} \in \mathbb{B}(\bar{y}, \epsilon_0). \tag{6.1}$$

Define the map $d_{\tilde{y}}(x) : \mathbb{R}^n \to \mathbb{R}$ by $d_{\tilde{y}}(x) := d(\tilde{y}, S(x))$. In their proof, the condition in [35, 9(23)] can be written more precisely as:

$$\begin{cases} \text{if } v \in \hat{\partial} d_{\tilde{y}}(\hat{x}) \text{ with } |\hat{x} - \bar{x}| \leq \delta_0 \\ \text{and } d_{\tilde{y}}(\hat{x}) + |\tilde{y} - \bar{y}| \leq \epsilon_0, \text{ then} \\ \text{there exists } z \in \mathbb{B} \text{ with } (v, -z) \in N_{\mathrm{gph}(S)}(\hat{x}, \hat{y}), \end{cases} \tag{6.2}$$

where $\hat{y}$ is any point of $S(\hat{x})$ nearest to $\tilde{y}$, that is $|\hat{y} - \tilde{y}| = d_{\tilde{y}}(\hat{x})$. Moving on, they proved that the map $(x, y) \mapsto d(y, S(x))$ is Lipschitz continuous near $(\bar{x}, \bar{y})$, or more precisely, there exists some $\lambda > 0$ and $\mu > 0$ such that

$$d(y, S(x)) \leq \lambda(|x - \bar{x}| + |y - \bar{y}|) \text{ when } |x - \bar{x}| \leq \mu \text{ and } |y - \bar{y}| \leq \mu. \tag{6.3}$$

Furnished with this, we can choose $\delta > 0$ and $\epsilon > 0$ small enough that $2\delta \leq \min\{\mu, \delta_0\}$, $\epsilon \leq \min\{\mu, \epsilon_0/2\}$ and $\lambda(2\delta + \epsilon) \leq \epsilon_0/2$. If $v \in \hat{\partial} d_{\tilde{y}}(\hat{x})$, $|\hat{x} - \bar{x}| \leq \delta$ and $|\tilde{y} - \bar{y}| \leq \epsilon$, then by (6.3), we have

$$d_{\tilde{y}}(\hat{x}) \leq \lambda(|\hat{x} - \bar{x}| + |\tilde{y} - \bar{y}|) \leq \lambda(2\delta + \epsilon) \leq \frac{\epsilon_0}{2}.$$

Formula (6.2) then tells us that for any $v \in \hat{\partial} d_{\tilde{y}}(\hat{x})$, there exists some $z \in \mathbb{B}$ such that $(v, -z) \in N_{\mathrm{gph}(S)}(\hat{x}, \hat{y})$.

By (6.1), there is some $(\tilde{v}, -\tilde{z}) \in N_{\mathrm{gph}(S)}(\bar{x}, \bar{y})$ such that $|(\tilde{v}, -\tilde{z}) - (v, -z)| < \theta$. This gives $\tilde{v} \in D^*S(\bar{x} \mid \bar{y})((1+\theta)\mathbb{B}^m)$, or $\frac{1}{1+\theta}\tilde{v} \in D^*S(\bar{x} \mid \bar{y})(\mathbb{B}^m)$. Further arithmetic gives $|v - \frac{1}{1+\theta}\tilde{v}| \leq |v - \tilde{v}| + |\frac{\theta}{1+\theta}\tilde{v}| \leq \theta(1 + \frac{\bar{\kappa}}{1+\theta})$, so $v \in D^*S(\bar{x} \mid \bar{y})(\mathbb{B}^m) + \theta(1 + \frac{\bar{\kappa}}{1+\theta})\mathbb{B}^n$. Let $\bar{\theta}$ be $\theta(1 + \frac{\bar{\kappa}}{1+\theta})$. We thus have $\partial d_{\tilde{y}}(\hat{x}) \subset D^*S(\bar{x} \mid \bar{y})(\mathbb{B}) + \bar{\theta}\mathbb{B}$. The function $d_{\tilde{y}}(\cdot)$ is Lipschitz, so the Clarke subdifferential equals $\mathrm{conv}(\partial d_{\tilde{y}}(\cdot))$ in $\mathbb{B}(\bar{y}, \epsilon)$.

Choose any two points $x, x' \in \mathbb{B}(\bar{x}, \delta)$. By the nonsmooth mean value theorem (Theorem 3.5), we have $d_{\tilde{y}}(x) - d_{\tilde{y}}(x') = \langle v, x - x' \rangle$ for some $v \in \partial_C d_{\tilde{y}}(x_\tau) = \mathrm{conv}(\partial d_{\tilde{y}}(x_\tau))$, where $x_\tau = \tau x + (1-\tau)x'$ for some $\tau \in (0,1)$. Now,

$$\begin{aligned} \langle v, x - x' \rangle &\leq \max\{\langle \tilde{v}, x - x' \rangle \mid \tilde{v} \in \partial_C d_{\tilde{y}}(x_\tau)\} \\ &\leq \max\{\langle \tilde{v}, x - x' \rangle \mid \tilde{v} \in D^*S(\bar{x} \mid \bar{u})(\mathbb{B}) + \bar{\theta}\mathbb{B}\} \\ &= \kappa(x' - x) + \bar{\theta}|x' - x|. \end{aligned}$$

This can be rephrased as $d_{\tilde{y}}(x') \geq d_{\tilde{y}}(x) - \kappa(x' - x) - \bar{\theta}|x' - x|$. As $\tilde{y}$ varies over all points in $\mathbb{B}(\bar{y}, \epsilon)$, this readily gives $S(x') \cap \mathbb{B}(\bar{y}, \epsilon) \subset S(x) + T(x' - x) + \bar{\theta}|x' - x|\mathbb{B}^m$, which is what we seek to prove. □

Rockafellar [33] established the relationship between the Aubin property of $S$ at $\bar{x}$ for $\bar{y}$ and the Lipschitz continuity properties of $d_{\tilde{y}}(\cdot)$. (See [35, Exercise 9.37].) The key to the proof of Theorem 6.4 is that the nonsmooth mean value theorem on $d_{\tilde{y}}(\cdot)$ gives us more information on the continuity properties of $S$.

We close this section with a remark on Theorem 6.4.

*Remark* 6.5. (More precise $T$-differentiability) Suppose $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, and $\bar{y} \in S(\bar{x})$. To obtain a better positively homogeneous map $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ than the one stated in Theorem 6.4, one applies Theorem 6.4 on the map $g \circ S + f : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, where $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^m$ are linear maps and $g$ is invertible. This approach is equivalent to looking at the normal cones of $\mathrm{gph}(g \circ S + f)$, which can also be obtained by performing the linear map $(x, y) \mapsto (x, g(y) + f(x))$ on $\mathrm{gph}(S) \subset \mathbb{R}^n \times \mathbb{R}^m$. By appealing to the calculus rules in Section 5, this gives another $\tilde{T} : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ for which $S$ is pseudo strictly $\tilde{T}$-differentiable. If we choose finitely many $\{(f_i, g_i)\}_i$, then we can define $T' : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ by

$$T'(w) = \tilde{T}_i(w), \text{ where } i \text{ is determined uniquely by } w.$$

It is easy to see that $S$ is $T'$-differentiable at $\bar{x}$ for $\bar{y}$.

## 7. Metric regularity and open covering

For $S : X \rightrightarrows Y$, it is well known that the Aubin property of $S^{-1}$ is related to the metric regularity and open covering properties of $S$. In this section, we study metric regularity and open covering in a more axiomatic manner with the help of

$T$-differentiability, proving new relations in these subjects. We caution that for much of this section, we need $T : Y \rightrightarrows X$ instead of $T : X \rightrightarrows Y$.

We begin with our definitions of generalized metric regularity and open covering.

**Definition 7.1.** Let $X$ and $Y$ be Banach spaces and let $S : X \rightrightarrows Y$ be a set-valued map where $\mathrm{gph}(S)$ is locally closed at $(\bar{x}, \bar{y})$, and $T : Y \rightrightarrows X$ be positively homogeneous.

(a) $S$ is *$T$-metrically regular* at $(\bar{x}, \bar{y}) \in \mathrm{gph}(S)$ if for any $\delta > 0$, there exist neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ and $r > 0$ such that, for any $x \in V$ and $A \subset r\mathbb{B}$, (or equivalently, for any set $A \subset r\mathbb{B}$ containing exactly one element),

$$y \in (S(x) + A) \cap W \text{ implies } x \in S^{-1}(y) + (T + \delta)(A). \tag{7.1}$$

(b) $S$ is *$(C, T)$-metrically regular* at $(\bar{x}, \bar{y}) \in \mathrm{gph}(S)$ if for any $\delta > 0$, there exist neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ and $r > 0$ such that, for any $x \in V$ and $t > 0$, (7.1) holds for any set $A \subset r\mathbb{B}$ of the form $A = tC$, where $C \subset Y$ is closed.

**Definition 7.2.** Let $X$ and $Y$ be Banach spaces and let $S : X \rightrightarrows Y$ be a set-valued map where $\mathrm{gph}(S)$ is locally closed at $(\bar{x}, \bar{y})$, and $T : Y \rightrightarrows X$ be positively homogeneous.

(a) $S$ is a *$T$-open covering* at $(\bar{x}, \bar{y}) \in \mathrm{gph}(S)$ if for any $\delta > 0$, there exist neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ and $r > 0$ such that, for any $x \in V$ and any set $A \subset r\mathbb{B}$ (or equivalently, for any set $A \subset r\mathbb{B}$ containing exactly one element),

$$(S(x) + A) \cap W \subset S\big(x + (T + \delta)(A)\big). \tag{7.2}$$

(b) $S$ is a *$(C, T)$-open covering* at $(\bar{x}, \bar{y}) \in \mathrm{gph}(S)$ if for any $\delta > 0$, there exist neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ such that, for any $x \in V$ and $t > 0$, (7.2) holds for any set $A \subset r\mathbb{B}$ of the form $A = tC$, where $C \subset Y$ is closed.

Open covering is sometimes known as *linear openness*. Setting $C = \mathbb{B}$ and $T(w) := \kappa|w|\mathbb{B}$ in both cases reduce the above definitions to the classical definitions of metric regularity and open covering with modulus $\kappa$, which will be clear by Proposition 7.8. Metric regularity with modulus $\kappa$ is often written compactly as: For all $\kappa' > \kappa$, there exists neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ such that

$$x \in V, y \in W \text{ implies } d\big(x, S^{-1}(y)\big) \leq \kappa' d\big(y, S(x)\big).$$

In practical problems, one might choose $C$ to be a set different from the unit ball $\mathbb{B}$, as this choice allows one to identify sensitive and less sensitive directions, and to use different norms.

We have our following theorem generalizing the classical equivalence of pseudo strict $T$-differentiability, $T$-metric regularity and $T$-open covering.

**Theorem 7.3.** *($T$-metric regularity and $T$-openness) Let $X$ and $Y$ be Banach spaces and $T : Y \rightrightarrows X$ be a positively homogeneous map. For $S : X \rightrightarrows Y$, the following are equivalent:*

(MR) *$S$ is $T(-\cdot)$-metric regular at $(\bar{x}, \bar{y}) \in \mathrm{gph}(S)$.*
(OC) *$S$ is a $(-T(-\cdot))$-open covering at $(\bar{x}, \bar{y}) \in \mathrm{gph}(S)$.*
(IT) *$S^{-1}$ is pseudo strictly $T$-differentiable at $\bar{y}$ for $\bar{x}$.*

*Proof.* We first note that $x \in S^{-1}(y') \cap V$ can be rewritten as $x \in V$, $y' \in S(x)$, so (IT) is equivalent to: For any $\delta > 0$, we can find neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ such that:

$$\underbrace{x \in V}_{(1)}, \underbrace{y' \in S(x)}_{(2)}, \underbrace{y \in W}_{(3)}, \underbrace{y' \in W}_{(4)} \text{ implies } \underbrace{x \in S^{-1}(y) + (T+\delta)(y'-y)}_{(5)}. \tag{7.3}$$

Take the set $A$ to be $A = \{y - y'\}$. Rewriting (2) and (3) as $y \in [S(x) + A] \cap W$, we get

$$\begin{aligned} & x \in V,\ A \subset y - W \text{ and } y \in [S(x)+A] \cap W \\ \text{implies} \quad & x \in S^{-1}(y) + (T+\delta)(A). \end{aligned}$$

We can assume $W = \mathbb{B}(\bar{y}, r')$. If $y \in \mathbb{B}(\bar{y}, \frac{r'}{2})$, then $\mathbb{B}(\mathbf{0}, \frac{r'}{2}) \subset y - \mathbb{B}(\bar{y}, r')$. So choosing $r = \frac{r'}{2}$ and writing (2) and (3) as $y \in [S(x) + A] \cap \mathbb{B}(\bar{y}, r)$, we have $T(-\cdot)$-metric regularity in (MR), i.e.,

$$\begin{aligned} & x \in V,\ A \subset \mathbb{B}(\mathbf{0}, r) \text{ and } y \in [S(x)+A] \cap \mathbb{B}(\bar{y}, r) \\ \text{implies} \quad & x \in S^{-1}(y) + (T+\delta)(A). \end{aligned} \tag{7.4}$$

For the converse, setting $A = \{y - y'\}$ in $A \subset \mathbb{B}(\mathbf{0}, r)$ gives $y' \in \mathbb{B}(y, r)$. The formula $y \in [S(x) + A] \cap \mathbb{B}(\bar{y}, r)$ can be written as $y' \in S(x)$, $y \in \mathbb{B}(\bar{y}, r)$. If $y \in \mathbb{B}(\bar{y}, \frac{r}{2})$, then $\mathbb{B}(\bar{y}, \frac{r}{2}) \subset \mathbb{B}(y, r)$. This implies statement (7.3) with $W = \mathbb{B}(\bar{y}, \frac{r}{2})$, which gives (IT) as needed.

For the equivalence of (OC) and (IT), note that condition (5) is equivalent to $y \in S(x - (T+\delta)(y'-y))$. The steps in the proof follow exactly with this change. □

At this point, we prove the equivalence of $(C, T)$-metric regularity and $T'$-metric regularity for some well chosen $T' : Y \rightrightarrows X$. The properties of $(C, T)$-metric regularity (and hence $(C, T)$-open coverings) can therefore be obtained from the corresponding properties of $T$-metric regularity.

**Proposition 7.4.** *(Reduction of $(C, T)$-metric regularity to $T$-metric regularity) Let $X$ and $Y$ be Banach spaces, $S : X \rightrightarrows Y$ be a set-valued map, and $T : Y \rightrightarrows X$ be positively homogeneous. Suppose $\epsilon\mathbb{B} \subset C \subset R\mathbb{B}$ for some $\epsilon, R > 0$. Then $S$ is $(C, T)$-metrically regular at $(\bar{x}, \bar{y})$ if and only if $S$ is $T'$-metrically regular there, where $T' : Y \rightrightarrows X$ is defined by $T'(w) = T(tC)$ for $t = \min\{\lambda \mid w \in \lambda C\}$.*

*Proof.* We recall from the proof of Theorem 7.3, $T'$-metric regularity is equivalent to: For all $\delta > 0$, there is a neighborhood $V$ of $\bar{x}$ and $r > 0$ such that

$$x \in V, y' \in S(x), y \in \mathbb{B}(\bar{y}, r), y - y' \in r\mathbb{B} \text{ implies } x \in S^{-1}(y) + (T'+\delta)(y-y'). \tag{7.5}$$

Next, we note that $(C, T)$-metric regularity is equivalent to: For any $\delta > 0$, there is a neighborhood $V$ of $\bar{x}$ and $r > 0$ such that

$$x \in V, y \in (S(x) + tC) \cap \mathbb{B}(\bar{y}, r), tC \subset r\mathbb{B} \text{ implies } x \in S^{-1}(y) + (T+\delta)(tC),$$

which is in turn equivalent to: For any $\delta > 0$, there is a neighborhood $V$ of $\bar{x}$ and $r > 0$ such that

$$\begin{aligned} & x \in V, y' \in S(x), y \in \mathbb{B}(\bar{y}, r), y - y' \in tC, tC \subset r\mathbb{B} \\ \text{implies} \quad & x \in S^{-1}(y) + (T+\delta)(tC). \end{aligned} \tag{7.6}$$

Suppose $S$ is $T'$-metrically regular at $(\bar{x}, \bar{y})$. Let $\delta > 0$, $V$ and $r$ be such that (7.5) holds, and suppose $x \in V$, $y' \in S(x)$ and $y \in \mathbb{B}(\bar{y}, r)$. Let $t > 0$ be such that $y - y' \in tC$. Then we have $x \in S^{-1}(y) + (T' + \delta)(y - y')$, which implies $x \in S^{-1}(y) + (T + \delta)(tC)$. This in turn means that (7.6) holds, so $S$ is $(C,T)$-metrically regular at $(\bar{x}, \bar{y})$.

For the converse, note that the assumptions implicitly assume that $C$ is bounded. We assume $\epsilon\mathbb{B} \subset C \subset R\mathbb{B}$ and examine the condition (7.6). If $y - y' \in \frac{r\epsilon}{R}\mathbb{B}$, then the minimum $t > 0$ such that $y - y' \in tC$ gives $tC \subset r\mathbb{B}$, so the condition $tC \subset r\mathbb{B}$ is superfluous in (7.6). This gives us the converse. □

We now look at the connection between pseudo $T$-outer differentiability and metric subregularity. For more on metric subregularity, we refer the reader to [11]. We define $T$-metric subregularity as follows.

**Definition 7.5.** Let $X$ and $Y$ be Banach spaces, $S : X \rightrightarrows Y$ be a set-valued map, and $T : Y \rightrightarrows X$ be positively homogeneous.

(1) $S$ is *$T$-metrically subregular* at $(\bar{x}, \bar{y}) \in \mathrm{gph}(S)$ if for any $\delta > 0$, there exists a neighborhood $V$ of $\bar{x}$ and $r > 0$ such that, for any $A \subset r\mathbb{B}$, (or equivalently, for any set $A \subset r\mathbb{B}$ containing exactly one element.)

$$x \in V, \bar{y} \in S(x) + A \text{ implies } x \in S^{-1}(\bar{y}) + (T + \delta)(A). \tag{7.7}$$

(2) $S$ is *$(C,T)$-metrically subregular* at $(\bar{x}, \bar{y}) \in \mathrm{gph}(S)$ provided that there exist a neighborhood $V$ of $\bar{x}$ such that, for $t > 0$, (7.7) holds for $A = tC \subset r\mathbb{B}$. (or equivalently, for any set $A \subset r\mathbb{B}$ containing exactly one element.)

Again, the special case of $C = \mathbb{B}$ and $T(w) := \kappa|w|\mathbb{B}$, $T$-metric subregularity and $(C,T)$-metric subregularity are equivalent to metric subregularity with modulus $\kappa$ (which will be clear by Proposition 7.8), and can be written compactly as: For any $\kappa' > \kappa$, there exists a neighborhood $V$ of $\bar{x}$ such that

$$x \in V \text{ implies } d\big(x, S^{-1}(\bar{y})\big) \leq \kappa' d\big(\bar{y}, S(x)\big).$$

Here is a result on the equivalences between $T$-metric subregularity and pseudo outer $T$-differentiability.

**Theorem 7.6.** *(Generalized metric subregularity) Let $X$ and $Y$ be Banach spaces, $S : X \rightrightarrows Y$ be a closed set-valued map, and $T : Y \rightrightarrows X$ be a positively homogeneous set-valued map. The following are equivalent:*

(MR′) *$S$ is $T(-\cdot)$-metric subregular at $(\bar{x}, \bar{y}) \in \mathrm{gph}(S)$.*

(IT′) *$S^{-1}$ is pseudo outer $T$-differentiable at $\bar{y}$ for $\bar{x}$.*

*Proof.* Condition (IT′) can be written as: For any $\delta > 0$, there exists a neighborhood $V$ of $\bar{x}$ and $r > 0$ such that

$$x \in V, x \in S^{-1}(y'), y' - \bar{y} \in r\mathbb{B} \implies x \in S^{-1}(\bar{y}) + (T + \delta)(y' - \bar{y}). \tag{7.8}$$

It is also clear that condition (MR′) can be written in this form. □

Here is a lemma on pseudo strict $T$-differentiability amended from [35, Lemma 9.39], which is in turn attributed to [14].

**Lemma 7.7.** *(Extended formulation of pseudo $T$-differentiability) Consider a mapping $S : X \rightrightarrows Y$, a pair $(\bar{x}, \bar{y}) \in \mathrm{gph}(S)$ where $\mathrm{gph}(S)$ is locally closed, and a set $D \subset X$ containing $\bar{x}$, and a positively homogeneous map $T : X \rightrightarrows Y$. Suppose that there is a $\delta > 0$ such that*

$$T(w) \supset \delta|w|\mathbb{B} \text{ for all } w \in X. \tag{7.9}$$

*Then the following two conditions are equivalent:*

(a) *there exist neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ such that*
$$S(x') \cap W \subset S(x) + T(x' - x) \text{ for all } x, x' \in D \cap V.$$

(b) *there exist neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ such that*
$$S(x') \cap W \subset S(x) + T(x' - x) \text{ for all } x \in D \cap V, x' \in D.$$

*The following two conditions are equivalent as well.*

(a$'$) *there exist neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ such that*
$$S(x') \cap W \subset S(\bar{x}) + T(x' - \bar{x}) \text{ for all } x' \in D \cap V.$$

(b$'$) *there exists a neighborhood $W$ of $\bar{y}$ such that*
$$S(x') \cap W \subset S(\bar{x}) + T(x' - \bar{x}) \text{ for all } x' \in D.$$

*Proof.* Trivially (b) implies (a), so assume that (a) holds for neighborhoods $V = \mathbb{B}(\bar{x}, \gamma)$ and $W = \mathbb{B}(\bar{y}, \epsilon)$. We will verify that (b) holds for $V' = \mathbb{B}(\bar{x}, \gamma')$ and $W' = \mathbb{B}(\bar{y}, \epsilon')$ when $0 < \gamma' < \gamma$, $0 < \epsilon' < \epsilon$, and $2\delta\gamma' + \epsilon' \leq \delta\gamma$.

Fix any $x \in D \cap \mathbb{B}(\bar{x}, \gamma')$. Our assumption gives us
$$S(x') \cap \mathbb{B}(\bar{y}, \epsilon') \subset S(x) + T(x' - x) \text{ when } x' \in D \cap \mathbb{B}(\bar{x}, \gamma'),$$
and our goal is to demonstrate that this holds also when $x' \in D \backslash \mathbb{B}(\bar{x}, \gamma)$. From applying (a) to $x' = \bar{x}$, we see that $\bar{y} \in S(x) + T(x - \bar{x})$ and consequently $\mathbb{B}(\bar{y}, \epsilon') \subset S(x) + (\delta\gamma' + \epsilon')\mathbb{B}$, where $\delta\gamma' + \epsilon' \leq \delta\gamma - \delta\gamma'$. But $|x' - x| > \gamma - \gamma'$ when $x' \in D \backslash \mathbb{B}(\bar{x}, \gamma)$, so for such points $x'$ we have $\delta\gamma - \delta\gamma' \leq \delta|x' - x|$. Hence
$$S(x') \cap \mathbb{B}(\bar{y}, \epsilon') \subset S(x) + \delta|x' - x|\mathbb{B} \subset S(x) + T(x' - x)$$
as required.

We now look at the equivalences of (a$'$) and (b$'$). Trivially (b$'$) implies (a$'$), so assume that (a$'$) holds for neighborhoods $V = \mathbb{B}(\bar{x}, \gamma)$ and $W = \mathbb{B}(\bar{y}, \epsilon)$. We can reduce $\epsilon$ so that $\epsilon \leq \delta\gamma$. Then $x' \in D \backslash \mathbb{B}(\bar{x}, \gamma)$ implies $W \subset S(\bar{x}) + T(x' - \bar{x})$, which gives us what we seek. □

If $T^{-1}(\mathbf{0}) = Y$, we have the following equivalent definitions for $T$-metric regularity. $T$-open covering and $T$-metric subregularity.

**Proposition 7.8.** *(Alternate definition of metric (sub)regularity and openness) Let $S : X \rightrightarrows Y$ and $T : Y \rightrightarrows X$ be set-valued maps, with $T$ positively homogeneous and $T^{-1}(\mathbf{0}) = Y$. The following conditions are equivalent to the $T$-metric regularity of $S$ at $(\bar{x}, \bar{y}) \in \operatorname{gph}(S)$:*

(MR$_1$) *For any $\delta > 0$, there exist a neighborhood $V$ of $\bar{x}$ and $r > 0$ such that, for any $x \in V$ and $A \subset Y$, (or equivalently, for any set $A \subset Y$ containing exactly one element),*
$$y \in [S(x) + A] \cap \mathbb{B}(\bar{y}, r) \text{ implies } x \in S^{-1}(y) + (T + \delta)(A).$$

*The following is equivalent to the $T$-open covering of $S$ at $(\bar{x}, \bar{y}) \in \operatorname{gph}(S)$:*

(OC$_1$) *For any $\delta > 0$, there exist a neighborhoods $V$ of $\bar{x}$ and $r > 0$ such that, for any $x \in V$ and any set $A \subset Y$ (or equivalently, for any set $A \subset Y$ containing exactly one element),*
$$[S(x) + A] \cap \mathbb{B}(\bar{y}, r) \subset S\big(x + (T + \delta)(A)\big).$$

*The following is equivalent to the $T$-metric subregularity of $S$ at $(\bar{x}, \bar{y}) \in \operatorname{gph}(S)$:*

($\mathrm{MR}_1'$) *For any $\delta > 0$, there exists a neighborhood $V$ of $\bar{x}$ such that, for any $x \in V$ and $A \subset Y$, (or equivalently, for any set $A \subset Y$ containing exactly one element)*

$$\bar{y} \in S(x) + A \text{ implies } x \in S^{-1}(\bar{y}) + (T+\delta)(A).$$

*Proof.* The condition $T^{-1}(\mathbf{0}) = Y$ ensures that Lemma 7.7 applies. Pseudo strict $T$-differentiability of $S^{-1}$ at $\bar{y}$ for $\bar{x}$ is equivalent to: For any $\delta > 0$, there exists neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ such that

$$\underbrace{x \in V}_{(1)}, \underbrace{y' \in S(x)}_{(2)}, \underbrace{y \in W}_{(3)} \implies \underbrace{x \in S^{-1}(y) + (T+\delta)(y'-y)}_{(5)}. \tag{7.10}$$

The difference here from (7.3) is that the condition $y' \in W$ is superfluous due to Lemma 7.7. We also note that $T(-\cdot)$-metric regularity is rewriting (2) and (3) as $y \in [S(x) + (y - y')] \cap W$ in the above, while $-T(-\cdot)$-open covering is rewriting (2) and (3) as $y \in [S(x) + (y - y')] \cap W$ and (5) as $S(x - (T+\delta)(y'-y))$. This proves the alternative definitions of $T$-metric regularity and $T$-open covering in ($\mathrm{MR}_1$) and ($\mathrm{OC}_1$). For the equivalence of $T$-metric subregularity and ($\mathrm{MR}_1'$), the proof is similar. □

## 8. Strict $T$-differentiability from outer $T$-differentiability

Suppose $S : X \rightrightarrows Y$ is such that $\bar{y} \in S(\bar{x})$, $T : X \rightrightarrows Y$ is positively homogeneous, and $S : X \rightrightarrows Y$ is pseudo strictly $T$-differentiable at $\bar{x}$ for $\bar{y}$. Then for any $\delta > 0$, there are neighborhoods $U$ of $\bar{x}$ and $V$ of $\bar{y}$ such that $S$ is pseudo outer $(T+\delta)$-differentiable at $x$ for $y$ for all $x \in U$ and $y \in V \cap S(x)$. The main theorem in this section is to show that the converse holds with additional assumptions.

We now study the relationship between pseudo outer $T$-differentiability and pseudo strict $T$-differentiability. The relation between pseudo (outer/ strict) $T$-differentiability and (outer/ strict) $T$-differentiability can be obtained from Theorem 5.5. First, here is a lemma on pseudo outer $T$-differentiable functions on a convex set that is comparable to the second part of [35, Theorem 9.2]. This has been extended from results in [24, 32].

**Lemma 8.1.** *(pseudo outer $T$-differentiability) Let $D \subset X$ be a convex set, and $S : D \rightrightarrows Y$ be an outer semicontinuous set-valued map satisfying $\bar{y} \in S(\bar{x})$. Assume that $T : X \rightrightarrows Y$ is a positively homogeneous set-valued map. Suppose that*

1. *there is some $r > 0$ such that $S(x) \cap \mathbb{B}(\bar{y}, r)$ is compact for all $x \in D$,*
2. *$\liminf_{x' \xrightarrow[D]{} x} S(x') \supset S(x) \cap \mathbb{B}(\bar{y}, r)$ with respect to $D$ for all $x \in D$,*
3. *$S$ is pseudo outer $T$-differentiable for all $x \in D$ and $y \in \mathbb{B}(\bar{y}, r) \cap S(x)$, and*
4. *$|T|^+ \leq \kappa$ for some $\kappa > 0$.*

*Then for all $x_0, x_1 \in D$ and $r'$ satisfying $\kappa|x_0 - x_1| + r' < r$, we have $S(x_1) \cap \mathbb{B}(\bar{y}, r') \subset S(x_0) + T(x_1 - x_0)$.*

*Proof.* By the definition of pseudo $T$-differentiability, for any $x \in D$ and $y \in \mathbb{B}(\bar{y}, r) \cap S(x)$, there are neighborhoods $V_{(x,y)}$ of $x$ and $U_{(x,y)}$ of $y$ such that

$$S(x') \cap U_{(x,y)} \subset S(x) + T(x' - x) \text{ for all } x' \in D \cap V_{(x,y)}.$$

Since $\mathbb{B}(\bar{y}, r) \cap S(x)$ is compact, we may choose $y_1, \ldots, y_k \in \mathbb{B}(\bar{y}, r)$ such that

$$\mathbb{B}(\bar{y}, r) \cap S(x) \subset \cup_{i=1}^k U_{(x,y_i)}.$$

Let $U_x$ be the right hand side of the above formula. Clearly, $U_x \cup (\mathbb{B}(\bar{y}, r))^c$ is an open set containing $S(x)$, where $(\mathbb{B}(\bar{y}, r))^c$ is the complement of $\mathbb{B}(\bar{y}, r)$. Since $S$ is outer semicontinuous, this implies that there is some neighborhood $V_x \subset \cap_{i=1}^k V_{(x,y_i)}$ such that $S(x') \subset U_x \cup (\mathbb{B}(\bar{y}, r))^c$ for all $x' \in V_x$, which implies that

$$S(x') \cap \mathbb{B}(\bar{y}, r) \subset U_x \text{ for all } x' \in V_x.$$

This gives us

$$S(x') \cap \mathbb{B}(\bar{y}, r) \subset S(x) + T(x' - x) \text{ for all } x' \in D \cap V_x. \tag{8.1}$$

Pick $x_0, x_1 \in D$. For each $t \in (0, 1)$, let $x_t = (1-t)x_0 + tx_1$. Formula (8.1) ensures that for each $t \in [0, 1]$, there is a ball $\mathbb{B}(x_t, \rho_t)$ such that for each $x'$ in $D \cap \mathbb{B}(x_t, \rho_t)$, $S(x') \cap \mathring{\mathbb{B}}(\bar{y}, r) \subset S(x_t) + T(x' - x_t)$. Here, $\mathring{\mathbb{B}}$ denotes an open ball. Define

$$\begin{aligned} \tau \quad := \quad & \sup\{t \in [0, 1] \mid \text{for each } s \in [0, t], \\ & \qquad S(x_s) \cap \mathring{\mathbb{B}}(\bar{y}, r - \kappa|x_s - x_0|) \subset S(x_0) + T(x_s - x_0)\}. \end{aligned}$$

We have $\tau > 0$ because $\rho_0$ is positive. We show first that

$$S(x_\tau) \cap \mathring{\mathbb{B}}(\bar{y}, r - \kappa|x_\tau - x_0|) \subset S(x_0) + T(x_\tau - x_0). \tag{8.2}$$

By assumption the set $S(x_0)$ is closed, and therefore so is $S(x_0) + T(x_\tau - x_0)$; let $Q$ be the complement of the latter set. If (8.2) were not true, then $S(x_\tau)$ would meet the open set $Q \cap \mathring{\mathbb{B}}(\bar{y}, r - \kappa|x_\tau - x_0|)$, at $\tilde{u}$ say. If $x_0, x_1$ were close enough to $\bar{x}$, then $\tilde{y} \in S(x_\tau) \cap \mathbb{B}(\bar{y}, r) \subset \liminf_{x' \to x_\tau} S(x')$. Therefore, for any $\tilde{r} > 0$, there exists some $\delta > 0$ such that

$$|x' - x_\tau| < \delta \text{ implies } S(x') \cap \mathbb{B}(\tilde{y}, \tilde{r}) \neq \emptyset.$$

If $\sigma \in (0, \tau)$ is such that $|x_\sigma - x_\tau| < \delta$, then $S(x_\sigma) \cap \mathbb{B}(\tilde{y}, \tilde{r}) \neq \emptyset$. This means that $(S(x_0) + T(x_\sigma - x_0)) \cap \mathbb{B}(\tilde{y}, \tilde{r}) \neq \emptyset$. We can find $\tilde{y}_\sigma$ such that $|\tilde{y} - \tilde{y}_\sigma| < \tilde{r}$, and $\tilde{y}_\sigma \in S(x_0) + T(x_\sigma - x_0)$. Next, recall that $S(x_0) + T(x_\tau - x_0) = S(x_0) + T(x_\tau - x_\sigma) + T(x_\sigma - x_0)$. This means that there exists $y_\sigma \in S(x_0) + T(x_\tau - x)$ for which $|\tilde{y}_\sigma - y_\sigma| < \kappa|x_\tau - x_\sigma|$. Combining the two gives

$$d\big(\tilde{y}, S(x_0) + T(x_\tau - x_0)\big) < \tilde{r} + \kappa|x_\tau - x_\sigma|.$$

Since the sum on the right hand side can be made arbitrarily small, we have $\tilde{y} \in S(x_0) + T(x_\tau - x_0)$, which contradicts $\tilde{y} \in Q$. This establishes (8.2).

If $\tau$ were less than 1 there would be $\lambda \in (\tau, 1)$ with $|x_\lambda - x_\tau| < \rho_\tau$, such that

$$S(x_\lambda) \cap \mathring{\mathbb{B}}(\bar{y}, r - \kappa|x_\lambda - x_0|) \not\subset S(x_0) + T(x_\lambda - x_0). \tag{8.3}$$

However, we would then have from (8.2) and the definition of $\rho_\tau$,

$$S(x_\lambda) \cap \mathring{\mathbb{B}}(\bar{y}, r - \kappa|x_\lambda - x_0|) \subset S(x_\tau) + T(x_\lambda - x_\tau),$$

and so

$$\begin{aligned} & S(x_\lambda) \cap \mathring{\mathbb{B}}(\bar{y}, r - \kappa|x_\lambda - x_0|) \\ \subset \quad & \big(S(x_\tau) + T(x_\lambda - x_\tau)\big) \cap \mathring{\mathbb{B}}(\bar{y}, r - \kappa|x_\lambda - x_0|) \\ = \quad & \Big(\big(S(x_\tau) \cap \mathring{\mathbb{B}}(\bar{y}, r - \kappa|x_\tau - x_0|)\big) + T(x_\lambda - x_\tau)\Big) \cap \mathring{\mathbb{B}}(\bar{y}, r - \kappa|x_\lambda - x_0|) \\ \subset \quad & \Big(\big(S(x_0) + T(x_\tau - x_0)\big) + T(x_\lambda - x_\tau)\Big) \cap \mathring{\mathbb{B}}(\bar{y}, r - \kappa|x_\lambda - x_0|) \\ \subset \quad & S(x_0) + T(x_\lambda - x_0), \end{aligned}$$

where the inclusion in the last line holds because $x_0, x_\tau$ and $x_\lambda$ are collinear and $T$ is a positively homogeneous convex valued map. This contradicts (8.3), so $\tau$ must be 1. Putting $\tau = 1$ in (8.2) shows that $S(x_1) \cap \mathring{\mathbb{B}}(\bar{y}, r - \kappa|x_1 - x_0|) \subset S(x_0) + T(x_1 - x_0)$. The required conclusion follows immediately. □

By decreasing $r$ and/or the size of $D$, we deduce that $S$ is pseudo strictly $T$-differentiable at $\bar{x}$ for $\bar{u}$. This is summarized in the theorem below.

**Theorem 8.2.** *(pseudo strict $T$-differentiability from pseudo outer $T$-differentiability) Suppose that $S : D \rightrightarrows Y$ and $\bar{y} \in S(\bar{x})$, where $D = \mathrm{dom}(S) \subset X$ is convex. Let $T : X \rightrightarrows Y$ be such that $|T|^+ < \kappa$ for some $\kappa > 0$. If for any $\delta > 0$, there exists open convex sets $U$ of $\bar{x}$ and $V$ of $\bar{y}$ such that*

(1) *$S$ is pseudo outer $(T + \delta)$-differentiable at $x$ for $y$ whenever $x \in U$, $y \in V$ and $y \in S(x)$,*
(2) *$\liminf_{x' \xrightarrow[D]{} x} S(x') \supset S(x) \cap V$ for all $x \in U$. (which is true when $S$ is inner semicontinuous.)*

*then $S$ is pseudo strictly $T$-differentiable at $\bar{x}$ for $\bar{y}$. Furthermore, the conclusion still holds if we amend condition (1) to:*

(1a) *$S$ is pseudo outer $(T + \delta)$-differentiable at $x$ for $y$ whenever $x \in U$, $y \in V$ and $y \in S(x)$ but $(x, y) \neq (\bar{x}, \bar{y})$, and*
(1b) *$S$ is outer semicontinuous at $\bar{x}$.*

*Proof.* Choose any $\delta > 0$. There are open sets $U$ containing $\bar{x}$ and $\mathbb{B}(\bar{y}, r)$ containing $\bar{y}$, such that $S$ is pseudo outer $(T + \delta)$-differentiable at $x$ for $y$ for all $x \in U$, $y \in \mathbb{B}(\bar{y}, r)$. We can reduce the size of $U$ so that the diameter of $U$, say $d$, satisfies $r - \kappa d > 0$. Let $r' = \frac{1}{2}(r - \kappa d)$. By Lemma 8.1,

$$S(x_1) \cap \mathbb{B}(\bar{y}, r') \subset S(x_0) + (T + \delta)(x_1 - x_0) \text{ for all } x_0, x_1 \in U \cap D.$$

Since $\delta$ is arbitrary, $S$ is pseudo strictly $T$-differentiable at $\bar{x}$ for $\bar{y}$ as needed.

To prove the second part, we only need to prove that if both conditions (1a) and (1b) hold, then $S$ is pseudo outer $(T + \delta)$-differentiable at $\bar{x}$ for $\bar{y}$. Again, suppose we have neighborhoods $U$ of $\bar{x}$ and $V = \mathbb{B}(\bar{y}, r)$ of $\bar{y}$ respectively such that condition 1(a) holds. First, we prove that for all $x \in \mathbb{B}(\bar{x}, \frac{r}{2\kappa}) \cap U \cap D$,

$$S(x) \cap \mathbb{B}\left(\bar{y}, \frac{r}{2}\right) \subset S(\bar{x}) + (T + \delta)(x - \bar{x}). \tag{8.4}$$

Since $S$ is outer semicontinuous at $\bar{x}$, for any $\epsilon > 0$, there exists a convex combination of $\{x, \bar{x}\}$ arbitrarily close enough to $\bar{x}$, say $\hat{x}$, such that

$$S(\hat{x}) \cap \mathbb{B}\left(\bar{y}, \frac{r}{2} + \kappa|x - \bar{x}|\right) \subset S(\bar{x}) + \epsilon\mathbb{B}.$$

Choose the domain $D'$ to be $\mathbb{B}(\frac{1}{2}(\hat{x} + x), \frac{1}{2}|\hat{x} - x|)$. Both $x, \hat{x}$ are in $D'$, which is convex, and $\frac{r}{2} + \kappa|x - \hat{x}| < r$, so the conditions for Lemma 8.1 are satisfied, and we have

$$S(x) \cap \mathbb{B}\left(\bar{y}, \frac{r}{2}\right) \subset S(\hat{x}) + (T + \delta)(x - \hat{x}).$$

Then

$$\begin{aligned}
S(x)\cap\mathbb{B}\left(\bar{y},\frac{r}{2}\right) &\subset (S(\hat{x})+(T+\delta)(x-\hat{x}))\cap\mathbb{B}\left(\bar{y},\frac{r}{2}\right)\\
&= \left(\left(S(\hat{x})\cap\mathbb{B}\left(\bar{y},\frac{r}{2}+\kappa|x-\bar{x}|\right)\right)+(T+\delta)(x-\hat{x})\right)\cap\mathbb{B}\left(\bar{u},\frac{r}{2}\right)\\
&\subset \left(S(\hat{x})\cap\mathbb{B}\left(\bar{y},\frac{r}{2}+\kappa|x-\bar{x}|\right)\right)+(T+\delta)(x-\hat{x})\\
&\subset (S(\bar{x})+\epsilon\mathbb{B})+(T+\delta)(x-\hat{x})\\
&= S(\bar{x})+(T+\delta)(x-\hat{x})+\epsilon\mathbb{B}.
\end{aligned}$$

For any fixed choice of $\epsilon>0$, we can always choose $\hat{x}$ close enough to $\bar{x}$ so that

$$(T+\delta)(x-\bar{x})\subset(T+\delta)(x-\hat{x})+\epsilon\mathbb{B}\subset(T+\delta)(x-\bar{x})+(\epsilon+\kappa|\hat{x}-\bar{x}|)\mathbb{B}.$$

Since $\epsilon$ is arbitrary, we have $S(x)\cap\mathbb{B}(\bar{y},\frac{r}{2})\subset S(\bar{x})+(T+\delta)(x-\bar{x})$. As $x$ is arbitrary in $\mathbb{B}(\bar{x},\frac{r}{2\kappa})\cap U\cap D$, this means that $S$ is pseudo outer $(T+\delta)$-differentiable at $\bar{x}$ for $\bar{y}$, and we are done. □

The corollary below addresses calmness and Lipschitz continuity. We did not explicitly treat the case where either the Lipschitz or calmness moduli could be infinity, but this is still easy.

**Corollary 8.3.** *(Calmness and Lipschitz moduli) Suppose that $S:D\rightrightarrows Y$ with $D=\mathrm{dom}(S)\subset X$ and $\bar{y}\in S(\bar{x})$. We have*

$$\mathrm{lip}\,S(\bar{x}\mid\bar{y})\geq\limsup_{(x,y)\xrightarrow[\mathrm{gph}(S)]{}(\bar{x},\bar{y})}\mathrm{calm}\ S(x\mid y).$$

*If $D$ is locally convex at $\bar{x}$, there is some $r>0$ such that $S(x)\cap\mathbb{B}(\bar{y},r)$ is compact for all $x\in D\cap V$, and there exist neighborhoods $V$ of $\bar{x}$ and $W$ of $\bar{y}$ such that $\liminf_{x'\xrightarrow[X]{}x}S(x')\supset S(x)\cap W$ for all $x\in V$, then equality holds. If in addition, $S$ is outer semicontinuous at $\bar{x}$, then we also have*

$$\mathrm{lip}\,S(\bar{x}\mid\bar{y})=\limsup_{\substack{(x,y)\xrightarrow[\mathrm{gph}(S)]{}(\bar{x},\bar{y})\\(x,y)\neq(\bar{x},\bar{y})}}\mathrm{calm}\ S(x\mid y).$$

In the single-valued case, we have the following corollary.

**Corollary 8.4.** *(Single-valued functions) Let $f:D\to Y$ be continuous, where $D\subset X$ is convex.*

1. *Let $T:X\rightrightarrows Y$ be a closed-convex-valued positively homogeneous map such that $|T|^+$ is finite. The function $f$ is strictly $T$-differentiable at $\bar{x}$ if and only if for all $\delta>0$, there is a convex neighborhood $U$ of $\bar{x}$ such that $f$ is $(T+\delta)$-differentiable at all points in $U\cap D$.*
2. *For any $\bar{x}\in D$,*

$$\mathrm{lip}\,f(\bar{x})=\limsup_{\substack{x\xrightarrow[D]{}\bar{x}\\x\neq\bar{x}}}\mathrm{calm}\,f(x).$$

*Acknowledgement.* I thank Asen Dontchev, Diethard Klatte, Bernd Kummer, Adrian Lewis, Boris Mordukhovich, Lionel Thibault for comments on a previous version of this paper which have led to improvements, and on the bibliography on the subject. In particular, I thank Lionel Thibault for bringing to my attention the papers [16, 17, 36]. Much of this paper was written while I was in the Fields Institute in Toronto, which has provided a wonderful environment for working on this paper.

## References


[1] J.-P. Aubin. *Lipschitz behavior of solutions to convex minimization problems*, Math. Oper. Res., 1, 225–245, 1984.
[2] J.-P. Aubin. *Viability Theory*, Bikhäuser, 1990. Republished as a Modern Bikhäuser Classic, 2009.
[3] J.-P. Aubin and A. Cellina. *Differential Inclusions: Set-Valued Maps and Viability Theory*, Grundlehren der mathematischen Wissenschaften, Vol 264, Springer, 1984.
[4] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*, Bikhäuser, 1990. Republished as a Modern Bikhäuser Classic, 2009.
[5] G. Beer. *Topologies on Closed and Convex Sets*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.
[6] J.M. Borwein and D.M. Zhuang. *Verifiable necessary and sufficient conditions for openness and regularity of set-valued and single-valued maps.* J. Math Anal. Appl., 134: 441–459, 1988.
[7] F.H. Clarke. *Generalized gradients and applications.* Trans. Am. Math. Soc., 205, 247–262, 1975.
[8] F. H. Clarke. *On the inverse function theorem.* Pacific Journal of Mathematics, 64, 97–102, 1976.
[9] F.H. Clarke. *Optimization and Nonsmooth Analysis.* Wiley, New York, 1983. Republished as Vol. 5, Classics in Applied Mathematics, SIAM, 1990.
[10] A. Dmitruk, A.A. Miliutin and N. Osmolovskii. *Liusternik's theorem and the theory of extrema, Russian Mathematical Surveys*, 35, 11–51, 1981.
[11] A.L. Dontchev and R.T. Rockafellar. *Regularity and conditioning of solution mappings in variational analysis*, Set-Valued Analysis, 12: 79–109, 2004.
[12] A.L. Dontchev and R.T. Rockafellar. *Implicit Functions and Solution Mappings: A View from Variational Analysis*, Springer Monographs in Mathematics, Springer, 2009.
[13] H. Frankowska. *Set-valued analysis and some control problems*, Proceedings of the International Conference: 30 Years of Modern Control Theory, Kingston, June 3-6, 1988. E. Roxin Editor, Marcel Dekker.
[14] R. Henrion. *The Approximate Subdifferential and Parametric Optimization*, Habilitation Thesis, Humboldt University, Berlin, 1997.
[15] J.-B. Hiriart-Urruty, *Tangent cones, generalized gradients and mathematical programming in Banach spaces*, Mathematics of Operations Research, 4, 79–97, 1979.
[16] A.D. Ioffe, *Différentielles généralisées d'applications localement Lipschitziennes d'un espace de Banach dans un autre*, C.r. hebd. Séanc. Acad. Sci. Paris. 289, 637-640 (1979).
[17] A.D. Ioffe. *Nonsmooth analysis: differential calculus of non-differentiable mappings*, Transactions of the AMS, 266, 1–56, 1981.
[18] A.D. Ioffe. *On the local surjection property*, Nonlinear Analysis: Theory, Methods and Applications, 11, 565–592, 1987.
[19] A.D. Ioffe. *Metric regularity and subdifferential calculus*, Russian Math. Surveys 55: 3 501–558, 2000.
[20] D. Klatte and B. Kummer. *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods and Applications*, Kluwer, 2002.
[21] E. Klein and A.C. Thompson. *Theory of Correspondences, Including Applications to Mathematical Economics*, Wiley, New York, 1984.
[22] A.Y. Kruger. *A covering theorem for set-valued mapping*s, Optimization, 19, 763–780, 1988.
[23] G. Lebourg. *Valeur moyenne pour gradient généralisé*, Comptes Rendus de l' Académie des Sciences de Paris, 281, 795–797, 1975.
[24] W. Li. *Sharp Lipschitz constants for basic optimal solutions and basic feasible solutions of linear programs*, SIAM J. Control Optim., 32, 1994, pp. 140–153.
[25] B.S. Mordukhovich. *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow, 1988. Russian.
[26] B.S. Mordukhovich. *Complete characterization of openness, metric regularity and Lipschitzian properties of multifunctions*, Trans. Amer. Math. Soc., 34: 1–35, 1993.
[27] B.S. Mordukhovich. *Variational Analysis and Generalized Differentiation I and II.*, Grundlehren der mathematischen Wissenschaften, Vols 330 & 331, Springer, Berlin, 2006.
[28] J.-P. Penot. *Calcul sous-différentiel et optimization*, Journal of Functional Analysis, 27, 248–276, 1978.

[29] J.-P. Penot. *Differentiability of relations and differential stability of perturbed optimization problems*, SIAM Journal on Control and Optimization, 22, 529–551, 1984.
[30] J.-P. Penot. *Metric regularity, openness and Lipschitzean behavior of multifunctions.* Nonlinear Analysis, 13: 629–643, 1989.
[31] S.M. Robinson. *Some continuity properties of polyhedral multifunctions*, Mathematical Programming Studies, 19, 200–221, 1981.
[32] S.M. Robinson. *Solution continuity in monotone affine variational inequalities*, SIAM J. Optim. Volume 18, Issue 3, pp. 1046–1060 (2007)
[33] R.T. Rockafellar. *Lipschitzian properties of multifunctions*, Nonlinear Analysis, 9, 867–885, 1985.
[34] R.T. Rockafellar. *Proto-differentiability of set-valued mappings and its applications in optimization*, in Analyse Non Linéaire, edited by H. Attouch, J.-P. Aubin, F. Clarke and I. Ekeland, pp. 449–482, Gauthier-Villars, Paris, 1989.
[35] R.T. Rockafellar and R.J.-B. Wets. *Variational Analysis*, Grundlehren der mathematischen Wissenschaften, Vol 317, Springer, Berlin, 1998.
[36] L. Thibault. *On Generalized differentials and subdifferentials of Lipschitz vector-valued functions*, Nonlinear Anal. 6 (1982), 1037-1053.
[37] J. Warga. *Fat homeomorphisms and unbounded derivate containers*, Journal of Mathematical Analysis and Applications, 81, 545–560, 1981.

*Current address*: Address for Fall 2009: Fields Institute, 222 College Street, Toronto, ON, M5T 3J1, Canada. For Spring 2010: University of Waterloo.
*E-mail address*: chj2pang@math.uwaterloo.ca