Warga derivate containers and other generalized differentials *

Dedicated to Jack Warga on his 80th birthday

Héctor J. Sussmann
Department of Mathematics
Rutgers, the State University of New Jersey
Hill Center—Busch Campus
Piscataway, NJ 08854-8019, USA
sussmann@math.rutgers.edu

1. Introduction

This is the first of two papers devoted to recent ideas on the theory of generalized differentials with good open mapping properties. Here we will discuss "generalized differentiation theories" (abbr. GDTs), with special emphasis on the series of developments initiated by Jack Warga's pioneering work on derivate containers. In the second paper, we will focus on the most recent theory, of "path-integral generalized differentials," and prove that it has the crucial properties required for a version of the Pontryagin Maximum Principle (abbr. PMP) to exist, namely, the chain rule and the directional open mapping property.

Our work continues the study of general smooth, nonsmooth, high-order, and hybrid versions of the PMP for finite-dimensional deterministic optimal control problems without state space constraints by means of a method developed by us in recent years. As explained in [11, 12, 13, 14], such versions can be derived in a unified way, by using a modified version of the approach of the classical book [6] by Pontryagin *et al.*. In the classical approach, one constructs "packets of needle variations," linearly approximates these packets at the base value of the variation parameter, and propagates the resulting linear approximations to the terminal point of the trajectory by means of the differentials of the reference flow maps. This technique must be modified by (a) replacing the the theory of the classical differential by other GDTs, (b) replacing the time-varying vector fields that occur in the classical MP by *flows*, and (c) replacing the needle variations by *abstract variations*. Every GDT yields a version of the MP provided it satisfies some natural properties such as the chain rule and an appropriate "directional open mapping property" (abbr. DOMP). (Details are provided in [11, 12, 13, 14].)

In these two papers we will not explicitly discuss flows, variations, and the PMP. How these topics are dealt with in the theory, and how one gets versions of the PMP for general GDTs, has been extensively disucssed in other papers, and for lack of space we choose not to address these questions here. We will, instead, give a detailed discussion of GDTs, and show how several important GDTs are defined and how they are related to one another. It will turn out that the complicated nature of these interrelationships, and in particular the fact that none of the "natural" theories contains all the others, suggests the question whether a theory that truly contains all other GDTs can be defined. The answer to this question is given in the second paper of the series, where a theory that does the desired job is defined and studied.

Here we propose an axiomatic definition of the concept of a GDT and a precise statement of the DOMP. We then outline the definitions of some GDTs, such as the Warga derivate containers (abbr. WDCs), and our more recent theories of "multidifferentials" and "generalized differential quotients" (abbbr. GDQs). Multidifferentials are generalizations of Warga's derivate containers, which in turn include the Clarke generalized Jacobians (abbr. CGJs) as a special case. The GDQs, on the other hand, generalize the CGJs in a different direction, and constitute a theory that neither contains nor is contained in that of the WDCs or in that of multidifferentials. We will make this precise by comparing these two types of theories by means of examples. This will set the stage for the second paper, in which we will presents a theory that achieves the desired unification.

Remark 1.1 Generalized differentials that extend the classical differential and the Warga derivate containers were studied by H. Halkin in the 1970s in three remarkably insightful papers ([3, 4, 5]) that, unfortunately, do not seem to have attracted the attention they deserved.

The work presented here is a continuation and extension of that of Halkin. In particular, our machinery makes it possible to deal systematically with set-valued maps, which appear when one studies flows of continuous but not necessarily Lipschitz vector fields (as in the "Lojasiewicz maximum principle," cf. [7]) and also in the analysis of differential inclusions (cf. [8]). $\diamondsuit$

2. GDTs

A "generalized differentiation theory" (abbr. GDT) is, roughly, a way of assigning a "differential" $\mathcal{D}(F;x,y;S)$ to each 4-tuple (F,x,y,S) consisiting of (a) a set valued map $F : M \to\!\!\!\to N$ whose source M and target N are manifolds of class C^1, (b) a point x of M, (c) a point y of N, and (d) a subset S of M. The object $\mathcal{D}(F;x,y;S)$ is required to be a set of nonempty compact subsets of $\mathrm{Lin}(T_xM, T_yN)$, the space of linear maps from T_xM to T_yN, where, for a manifold Q and a $q \in Q$, T_qQ denotes the tangent space to Q at q. The members of $\mathcal{D}(F;x,y;S)$ are the "$\mathcal{D}$-differentials of F at (x,y) in the direction of S." (The set $\mathcal{D}(F;x,y;S)$ could be empty, in which case we say that F is not D-differentiable at (x,y) in the direction of S.) The correspondence $\mathcal{D}$ is required to satisfy the chain rule, the Cartesian product rule, locality, and invariance under C^1 diffeomorphisms. Furthermore, it is required to be an extension of the classical theory of differentials of maps of class C^1.

To make all this precise, we introduce some notations and definitions.

*The author was supported in part by NSF Grant DMS98-03411-00798.

Set-valued maps. By a *set-valued map* (abbr. SVM) we mean a triple $F = (A,B,G)$ such that A and B are sets and G is a subset of $A \times B$. The sets A, B, G are, respectively, the *source*, *target*, and *graph* of F, and we write $A = \mathrm{So}(F)$, $B = \mathrm{Ta}(F)$, $G = \mathrm{Gr}(F)$. If x is any object, we write $F(x) = \{y : (x,y) \in \mathrm{Gr}(F)\}$. (Hence $F(x) = \emptyset$ unless $x \in \mathrm{So}(F)$.) The sets $\mathrm{Do}(F) = \{x \in \mathrm{So}(F) : F(x) \neq \emptyset\}$, $\mathrm{Im}(F) = \bigcup_{x\in\mathrm{So}(F)} F(x)$, are, respectively, the *domain* and *image* of F. If $F = (A,B,G)$ is an SVM, we say that F is an *SVM from A to B*, and write $F : A \to\!\!\!\to B$. We use $SVM(A,B)$ to denote the set of all SVMs from A to B. The expression "ppd map" stands for "possibly partially defined (that is, not necessarily everywhere defined) ordinary (that is, single-valued) map," and we write $f : A \dashrightarrow B$ to indicate that f is a ppd map from A to B. A *time-varying ppd map* from a set A to a set B is a ppd map from $A \times \mathbb{R}$ to B.

Classes of manifolds. We use $\mathbb{Z}_+$ to denote the set of all nonnegative integers. If $k \in \mathbb{Z}_+$, then $\mathcal{M}^k$, $\mathcal{SVM}(\mathcal{M}^k)$ will denote, respectively, the class of all finite-dimensional Hausdorff manifolds of class C^k without boundary, and the class of all SVMs F such that $\mathrm{So}(F)\in\mathcal{M}^k$ and $\mathrm{Ta}(F)\in\mathcal{M}^k$. If $k \geq 1$ and $x\in M\in\mathcal{M}^k$, then T_xM, T_x^*M, TM, T^*M will denote, respectively, the tangent and cotangent spaces to M at x, and the tangent and cotangent bundles of M. Clearly, TM and T^*M belong to $\mathcal{SVM}(\mathcal{M}^{k-1})$.

The Bouligand tangent cone. If $M \in \mathcal{M}^1$, $x \in M$, and $S \subseteq M$, then T_x^BS will denote the *Bouligand tangent cone* to S at x. By definition, T_x^BS is the set of all tangent vectors $v \in T_xM$ such that there exist a sequence $\{(x_j,h_j)\}_{j\in\mathbb{N}}$ of points of $S \times \mathbb{R}$ having the property that $h_j > 0$ for all j, $h_j \downarrow 0$ as $j \to \infty$, and $\lim_{j\to\infty} h_j^{-1}(\varphi(x_j) - \varphi(x)) = v\varphi$ for all functions $\varphi \in C^1(M,\mathbb{R})$.

Classes of linear spaces, linear multimaps. We write $\mathcal{RLS}$, $\mathcal{FDRLS}$, to denote, respectively, the class of all linear spaces over $\mathbb{R}$, and the class of all $X \in \mathcal{RLS}$ that are finite-dimensional. If $X,Y \in \mathcal{RLS}$, then $\mathrm{Lin}(X,Y)$ denotes the set of all $\mathbb{R}$-linear maps from X to Y. A subset of $\mathrm{Lin}(X,Y)$ will be called a *linear multimap* from X to Y. We use $\mathrm{MLin}(X,Y)$, $\mathrm{MLin}_c(X,Y)$, to denote, orespectively, the set of all linear multimaps from X to Y and (if $X,Y \in \mathcal{FDRLS}$) the set of all nonempty compact linear multimaps from X to Y.

Definition 2.1 A *generalized differentiation theory* (abbr. GDT) is a correspondence $\mathcal{D}$ that assigns to every $F \in \mathcal{SVM}(\mathcal{M}^1)$, every $(x,y) \in \mathrm{So}(F) \times \mathrm{Ta}(F)$, and every subset S of the source $\mathrm{So}(F)$, a set $\mathcal{D}(F;x,y;S) \subseteq \mathrm{MLin}_c(T_xM,T_yY)$ of nonempty compact linear multimaps from T_xM to T_yY, in such a way that the following axioms are satisfied:

1. If $m,n \in \mathbb{Z}_+$, $f : \mathbb{R}^m \mapsto \mathbb{R}^n$ is a map of class C^1, and S is a polyhedral convex cone in $\mathbb{R}^m$, then $\{Df(0)\}\in\mathcal{D}(f;0,0;S)$.
2. *(The chain rule)* If $F_i \in \mathcal{SVM}(\mathcal{M}^1)$ and $S_i\subseteq M_i$ for $i = 1,2$, $M_1 = \mathrm{So}(F_1)$, $M_2=\mathrm{Ta}(F_1)=\mathrm{So}(F_2)$, $M_3=\mathrm{Ta}(F_2)$, $x_i \in M_i$ for $i = 1,2,3$, $F_1(S_1) \subseteq S_2$, $\Lambda_1(T_{x_1}^BS_1)\subseteq T_{x_2}^BS_2$, and $\Lambda_2\in\mathcal{D}(F_2;x_2,x_3;S_2)$, then $\Lambda_2 \circ \Lambda_1\in\mathcal{D}(F_2 \circ F_1;x_1,x_3;S_1)$.
3. *(The product rule)* If, for $i = 1,2$, $F_i \in \mathcal{SVM}(\mathcal{M}^1)$, $x_i \in M_i = \mathrm{So}(F_i)$, $y_i\in N_i=\mathrm{Ta}(F_i)$, $S_i \subseteq M_i$, and $\Lambda_i \in \mathcal{D}(F_i;x_i,y_i;S_i)$, then $\Lambda_1 \times \Lambda_2$ belongs to $\mathcal{D}(F_1 \times F_2;(x_1,x_2),(y_1,y_2);S_1 \times S_2)$.
4. *(C^1 invariance)* Assume that (a) for $i = 1,2$, $M_i,N_i \in \mathcal{M}^1$, $F_i \in \mathcal{SVM}(\mathcal{M}^1)$, $M_i = \mathrm{So}(F_i)$, $N_i=\mathrm{Ta}(F_i)$, $S_i \subseteq M_i$, $x_i \in M_1$, $y_i \in N_i$, U_i,V_i are open subsets of M_i,N_i such that $x_i \in U_i$ and $y_i \in V_i$, (b) $\Phi : U_1 \mapsto U_2$ and $\Psi : V_1 \mapsto V_2$ are diffeomorphisms of class C^1 such that $\Phi(x_1)=x_2$ and $\Psi(y_1) = y_2$, (c) $\Phi(S_1 \cap U_1) = S_2 \cap V_2$, and (d) the set $(\Phi\times\Psi)\Big(\mathrm{Gr}(F_1)\cap((S_1 \cap U_1)\times V_1)\Big)$ is equal to $\mathrm{Gr}(F_2) \cap ((S_2 \cap U_2) \times V_2)$. Then the set $\mathcal{D}(F_2;x_2,y_2;S_2)$ is equal to $D\Psi(y_1) \circ \mathcal{D}(F_1;x_1,y_1;S_1) \circ D\Phi^{-1}(x_2)$. ♢

Remark 2.2 The C^1 invariance property implies in particular that a GDT $\mathcal{D}$ is *local*, in the sense that $\mathcal{D}$ satisfies

4'. *(Locality)* Assume that (a) $M,N \in \mathcal{M}^1$, $x \in M$, $y \in N$, U,V is are open subsets of M,N such that $x \in U$, $y \in V$, (b) for $i = 1,2$, $F_i \in \mathcal{SVM}(\mathcal{M}^1)$, $M = \mathrm{So}(F_i)$, $N=\mathrm{Ta}(F_i)$, $S_i \subseteq M_i$, (c) $S_1 \cap U = S_2 \cap U$ and $\mathrm{Gr}(F_1) \cap ((S_1 \cap U) \times V)$ is equal to $\mathrm{Gr}(F_2) \cap ((S_2 \cap U) \times V)$. Then $\mathcal{D}(F_1;x,y;S_1) = \mathcal{D}(F_2;x,y;S_2)$. ♢

Remark 2.3 To construct a GDT $\mathcal{D}$, it suffices to specify $\mathcal{D}(F;0,0;S)$, for all k,ℓ, whenever $F : \mathbb{R}^k \to\!\!\!\to \mathbb{R}^\ell$ and $S \subseteq \mathbb{R}^k$. Let $\mathcal{D}_0$ be a correspondence of the kind described above, but restricted to $M = \mathbb{R}^k$, $N = \mathbb{R}^k$, $x = 0$, $y = 0$. It is then a routine matter to write down a list of axioms that will guarantee that $\mathcal{D}_0$ "gives rise in a canonical way" to a GDT in the sense of Definition 2.1.

The crucial point is C^1 invariance and locality, i.e., the version of the C^1 invariance condition of Definition 2.1 restricted to $\mathbb{R}^k$, $\mathbb{R}^\ell$. We omit the details of the general theory, but refer the reader to our discussion of weak multidifferentials below (in §6, especially the remarks following Corollary 6.3), where we explain how this is done in one particular case. ♢

3. The directional open mapping property

We say that a GDT $\mathcal{D}$ has the *directional open mapping property* if the following statement is true:.

(DOMP) *Assume that $n,m \in \mathbb{Z}_+$, $F\in\mathcal{SVM}(\mathbb{R}^n,\mathbb{R}^m)$, $v \in \mathbb{R}^m$, C is a closed convex cone in $\mathbb{R}^n$, Λ belongs to $\mathcal{D}(F;0,0;C)$, and $v\in\bigcap_{L\in\Lambda}\mathrm{Int}(LC)$. Then there exists a closed convex cone D in $\mathbb{R}^n$ such that $v \in \mathrm{Int}(D)$, having the property that for every $\delta \in]0,\infty[$ there exists an $\varepsilon(\delta) \in]0,\infty[$ such that* $\boxed{D\cap\{y\in\mathbb{R}^m:\|y\|\leq\varepsilon(\delta)\}\subseteq F(C\cap\{x\in\mathbb{R}^n:\|x\|\leq\delta\})}$. ♢

4. Warga derivate containers (WDCs)

Definition 4.1 Let f be a map from an open subset Ω of $\mathbb{R}^m$ to $\mathbb{R}^n$, and let $x_* \in \Omega$. A *Warga derivate container* of f at x is a nonempty compact subset Λ of $\mathrm{Lin}(\mathbb{R}^m,\mathbb{R}^n)$ such that

(*) for every open neighborhood Λ' of Λ in the space $\mathrm{Lin}(\mathbb{R}^m,\mathbb{R}^n)$ there exist an open neighborhood

$U = U_{\Lambda'}$ of x_* in Ω and a sequence $\{f_j\}_{j=1}^{\infty}$ of maps of class C^1 from U to $\mathbb{R}^n$, such that (a) $f_j \to f$ uniformly on U as $j \to \infty$, and (b) the differential $Df_j(x)$ belongs to Λ' for every $x \in U$ and every j.

It follows from the definition of a WDC that any map f that admits a WDC at a point x_* must be Lipschitz continuous on a neighborhood of x_*, since we can always choose Λ' to be bounded, in which case f will be a uniform limit, on a neighborhood U of x_*, of maps of class C^1 with uniformly bounded derivatives.

Associated to the Warga derivate containers we can define GDTs by extending the definition of the WDCs to "WDCs along sets." This can be done in more than one way. For example, suppose C is a closed convex cone in $\mathbb{R}^n$, U is a neighborhood of 0 in $\mathbb{R}^n$, and $f : U \cap C \mapsto \mathbb{R}^m$ is a Lipschitz continuous map. We could then define a WDC of f at 0 in the direction of C to be a WDC of any extension $\tilde{f}$ of f to a Lipschitz map defined on a full neighborhood V of 0 in $\mathbb{R}^n$. Alternatively, we could mimic Definition 4.1, using uniform approximations of f by maps of class C^1 on relative neighborhoods $V \cap C$ of 0 in C. These two approaches lead to different theories.

We will not pursue this matter here any further and will, instead, move on directly to a more general theory called WMD ("weak multidifferentials") that will be defined *ab initio* as theory of differentials in the direction of sets.

5. Regular set-valued maps

If X, Y are metric spaces, then $SVM_{comp}(X, Y)$ will denote the subset of $SVM(X, Y)$ whose members are the set-valued maps from X to Y that have a compact graph. We say that a sequence $\{F_j\}_{j\in\mathbb{N}}$ of members of $SVM_{comp}(X, Y)$ *inward graph-converges* to an $F \in SVM_{comp}(X, Y)$—and write $F_j \xrightarrow{\text{igr}} F$—if for every open subset Ω of $X \times Y$ such that $\mathrm{Gr}(F) \subseteq \Omega$ there exists a $j_\Omega \in \mathbb{N}$ such that $\mathrm{Gr}(F_j) \subseteq \Omega$ whenever $j \geq j_\Omega$.

Definition 5.1 Assume that X, Y are metric spaces. A *regular set-valued map* from X to Y is a set-valued map $F \in SVM(X, Y)$ such that

- for every compact subset K of X, the restriction $F \lceil K$ of F to K belongs to $SVM_{comp}(K, Y)$ and is a limit—in the sense of inward graph-convergence—of a sequence of continuous single-valued maps from K to Y.

We use $\mathrm{REG}(X; Y)$ to denote the set of all regular set-valued maps from X to Y. $\diamondsuit$

It is easy to see that if $F : X \mapsto Y$ is an ordinary (that is, single-valued and everywhere defined) map, then F belongs to $\mathrm{REG}(X; Y)$ if and only if F is continuous.

An important class of examples of regular maps is provided by the following two results, whose proof we omit.

Theorem 5.2 *Assume that K is a compact metric space, Y is a normed space, and C is a convex subset of Y. Let $\Phi : K \to\!\!\!\to C$ be a set-valued map such that the graph of Φ is compact and the value $\Phi(x)$ is a nonempty convex set for every $x \in K$. Then Φ is regular as a map from K to C.*

Theorem 5.3 *Assume that X is a metric space, Y is a normed space, and C is a convex subset of Y. Let $\Phi : K \to\!\!\!\to C$ be an upper semicontinuous set-valued map with nonempty compact convex values. Then $\Phi \in$* $\mathrm{REG}(X; C)$. $\diamondsuit$

In addition, it is not hard to prove the following.

Theorem 5.4 *Assume that X, Y, Z are metric spaces. Let $F \in \mathrm{REG}(X; Y)$, $G \in \mathrm{REG}(Y; Z)$. Then the composite map $G \circ F$ belongs to* $\mathrm{REG}(X; Z)$. $\diamondsuit$

6. Weak multidifferentials

"Multidifferentials" were studied in [10]. Here we introduce a slightly more general theory, of objects that we call "weak multidifferentials."

As a preliminary, we need some notations. If X, Y are finite-dimensional real spaces, $\Omega \subseteq X$, and $\Lambda \subseteq \mathrm{Lin}(X, Y)$, we use $C^1_\Lambda(\Omega, Y)$ to denote the set of all maps $h : \Omega \mapsto Y$ of class C^1 such that $Dh(x) \in \Lambda$ for all $x \in \Omega$. (The precise meaning of "h is of class C^1" is clear if Ω is open. We also need to assign a meaning to this expression when Ω is closed. This is done in the usual Whitney way, which turns out to be equivalent to the following: a map of class C^1 from Ω to Y is a pair (h_1, H) consisting of continuous maps $h_1 : \Omega \mapsto Y$, $H : \Omega \mapsto \mathrm{Lin}(X, Y)$, having the property that, for every $x \in \Omega$, $h_1(x') - h_1(x) - H(x) \cdot (x' - x) = o(\|x' - x\|)$ as $x' \to x$ via values in Ω. If $h = (h_1, H)$ is such an object, then Dh is defined to be H, and if $x \in \Omega$ we define $h(x)$ to be $h_1(x)$.)

Definition 6.1 Let X, Y be finite-dimensional normed real linear spaces. Define $\bar{B}_X = \{x \in X : \|x\| \leq 1\}$, $Z = \mathrm{Lin}(X, Y)$. Let $F : X \to\!\!\!\to Y$ be a set-valued map. Let $(x_*, y_*) \in X \times Y$, and let Λ be a compact subset of Z. Let C be a closed convex cone in X. We say that Λ is a *weak multidifferential of F at (x_*, y_*) in the direction of C*, and write

$$\Lambda \in WMD(F; x_*, y_*; C)\,,$$

if the following is true:

(WMD) for every neighborhood Λ' of Λ in Z there exists a pair (R, Θ) such that

- (WMD.1) $R > 0$, and Θ is a function on the interval $]0, R]$ with values in $[0, \infty[$ and such that $\lim_{s\downarrow 0} \Theta(s) = 0$,
- (WMD.2) for every $r \in]0, R]$ there exist f, h such that
  - (WMD.2.1) $f \in \mathrm{REG}(x_* + (C \cap r\bar{B}_X), Y)$;
  - (WMD.2.2) $\mathrm{Gr}(f) \subseteq \mathrm{Gr}(F)$;
  - (WMD.2.3) $h \in C^1_{\Lambda'}(x_* + (C \cap r\bar{B}_X), Y)$;
  - (WMD.2.4) $h(x_*) = y_*$;
  - (WMD.2.5) the inequality

$$\sup\Big\{\|y - h(x)\| : y \in f(x)\Big\} \leq r.\Theta(r) \tag{1}$$

holds whenever $x - x_* \in C \cap r\bar{B}_X$. $\diamondsuit$

Weak multidifferntials satisfy the following version of the chain rule:

Theorem 6.2 *Let X_1, X_2, X_3 be finite-dimensional real linear spaces, and let $\bar{x}_i \in X_i$, $i = 1, 2, 3$. For $i = 1, 2$, let $F_i : X_i \to\to X_{i+1}$ be a set-valued map from X_i to X_{i+1}, let C_i be a closed convex cone in X_i, and let Λ_i be such that $\Lambda_i \in WMD(F_i; \bar{x}_i, \bar{x}_{i+1}, C_i)$. Let S be the linear subspace of X_2 spanned by C_2, and let Π be a linear projection from X_2 onto S. Assume that $F_1(\bar{x}_1 + C_1)$ is a subset of $\bar{x}_2 + C_2$. Then*

$$\Lambda_2 \circ \Pi \circ \Lambda_1 \in WMD(F_2 \circ F_1; \bar{x}_1, \bar{x}_3; C_1)\,. \qquad (2)$$

Theorem 6.2 was proved in [10] for multidifferentials, but the same proof applies to weak multidifferentials as well and we will not repeat it here.

The following is a trivial consequence of Theorem 6.2.

Corollary 6.3 *Let X_1, X_2, X_3 be finite-dimensional real linear spaces, and let $\bar{x}_i \in X_i$, $i = 1, 2, 3$. For $i = 1, 2$, let $F_i : X_i \to\to X_{i+1}$ be a set-valued map from X_i to X_{i+1}, let C_i be a closed convex cone in X_i, and let Λ_i be such that $\Lambda_i \in WMD(F_i; \bar{x}_i, \bar{x}_{i+1}, C_i)$. Let S be the linear subspace of X_2 spanned by C_2, and let Π be a linear projection from X_2 onto S. Assume that $F_1(\bar{x}_1 + C_1)$ is a subset of $\bar{x}_2 + C_2$ and $LC_1 \subseteq S$ for all $L \in \Lambda_1$. Then*

$$\Lambda_2 \circ \Lambda_1 \in WMD(F_2 \circ F_1; \bar{x}_1, \bar{x}_3; C_1)\,. \qquad (3)$$

Using Corollary 6.3, we can extend WMD to a GDT by defining $WMD(F; x_*, y_*; S)$, if M, N are manifolds of class C^1, $F : M \to\to N$, $x_* \in M$, $y_* \in N$, and $S \subseteq M$, as follows. We call the set S *conic at* x_* if there exists a coordinate chart κ near x_* such that $\kappa(x_*) = 0$ and $\kappa(\mathrm{Do}(\kappa) \cap S) = \mathrm{Im}(\kappa) \cap C$ for some closed convex cone C in $\mathbb{R}^{\dim M}$. If S is conic at x_*, then we define $WMD(F; x_*, y_*; S)$ as a set of nonempty compact subsets of $\mathrm{Lin}(T_{x_*}M, T_{y_*}N)$ using a chart κ as above, together with a chart for N near y_*. Corollary 6.3 then tells us that $WMD(F; x_*, y_*; S)$ does not depend on the choice of charts, so it is intrinsically defined. We complete the definition by taking $WMD(F; x_*, y_*; S)$ to be empty if S is not conic at x_*.

It is then easy to show that

Theorem 6.4 *WMD is a GDT that has the DOMP.* $\diamondsuit$

7. Generalized differential quotients (GDQs)

Definition 7.1 Let $m, n \in \mathbb{Z}_+$, let $F : \mathbb{R}^m \to\to \mathbb{R}^n$ be a set-valued map, and let Λ be a nonempty compact subset of $\mathbb{R}^{n\times m}$. Let S be a subset of $\mathbb{R}^m$. We say that Λ is a *generalized differential quotient* (abbreviated "GDQ") *of* F *at* $(0, 0)$ *in the direction of* S, and write $\Lambda \in GDQ(F; 0, 0; S)$, if for every positive real number δ there exist U, G such that

1. U is a compact neighborhood of 0 in $\mathbb{R}^m$ and $U \cap S$ is compact;
2. G is a regular set-valued map from $U \cap S$ to the δ-neighborhood Λ^δ of Λ in $\mathbb{R}^{n\times m}$;
3. $G(x) \cdot x \subseteq F(x)$ for every $x \in U \cap S$. $\diamondsuit$

Remark 7.2 The definition of the GDQ is a generalization of a characterization of classical differentials due to Botsko and Gosser, [2]. $\diamondsuit$

If M, N are C^1 manifolds, $\bar{x} \in M$, $\bar{y} \in N$, $S \subseteq M$, and $F : M \to\to N$, then we can define a set $GDQ(F; \bar{x}, \bar{y}; S)$ of nonempty compact linear multimaps from $T_{\bar{x}}M$ to $T_{\bar{y}}N$ by picking coordinate charts $M \ni x \to \xi(x) \in \mathbb{R}^m$, $N \ni y \to \eta(y) \in \mathbb{R}^n$—where $m = \dim M$, $n = \dim N$—defined near $\bar{x}$, $\bar{y}$ such n that $\xi(\bar{x}) = 0$, $\eta(\bar{y}) = 0$, and declaring a subset Λ of $\mathrm{Lin}(T_{\bar{x}}M, T_{\bar{y}}N)$ to belong to $GDQ(F; \bar{x}, \bar{y}; S)$ if the multimap $D\eta(\bar{y}) \circ \Lambda \circ D\xi(\bar{x})^{-1}$ is in $GDQ(\eta \circ F \circ \xi^{-1}; 0, 0; \xi(S))$. It turns out that, with this definition, *the set $GDQ(F; \bar{x}, \bar{y}; S)$ does not depend on the choice of the charts ξ, η.* Moreover, the following three results can be proved.

Theorem 7.3 *GDQ is a generalized differentiation theory that has the DOMP.* $\diamondsuit$

Theorem 7.4 *If $F : \mathbb{R}^n \mapsto \mathbb{R}^m$ is a continuous map, $x \in \mathbb{R}^n$, and F is classically differentiable at x, then $\{DF(x)\} \in GDQ(F; x, F(x); \mathbb{R}^n)$.* $\diamondsuit$

Theorem 7.5 *If $F : \mathbb{R}^n \mapsto \mathbb{R}^m$ is Lipschitz-continuous, and $x \in \mathbb{R}^n$, then the Clarke generalized Jacobian $\partial F(x)$ belongs to $GDQ(F; x, F(x); \mathbb{R}^n)$.* $\diamondsuit$

8. Comparison of WDC, WMD, and GDQ

Theorem 8.1 *If $F : \mathbb{R}^n \mapsto \mathbb{R}^m$ is Lipschitz-continuous, $x_* \in \mathbb{R}^n$, and Λ is a Warga derivate container of F at x_*, then $\Lambda \in WMD(F\ x_*, F(x_*) : \mathbb{R}^n)$.*

Proof. Let Λ' be a neighoborhood of Λ in $\mathbb{R}^{m\times n}$. Let $\{F_j\}_{j=1}^\infty$ be a sequence of $\mathbb{R}^m$-valued maps of class C^1 defined on a closed ball B in $\mathbb{R}^n$ with center x_*, such that $F_j \to F$ uniformly on B and $DF_j(x) \in \Lambda'$ for all j and all $x \in B$. Pick R and Θ in an arbitrary fashion so that (WMD.1) holds and $\Theta(s) > 0$ when $s > 0$. Given $r \in]0, R]$, let f be the restriction of F to B. Then f is a single-valued continuous map, so f is *a fortiori* a regular set-valued map.

The maps $B \ni x \mapsto F_j(x) - F_j(x_*) + F(x_*) = h_j(x)$ converge uniformly to f on B and satisfy $h_j(x_*) = F(x_*)$. Pick j such that $\|h_j\|_{sup} \le r.\Theta(r)$, and let $h = h_j$, $C = \mathbb{R}^n$. It is then clear that all the conditions of (WMD.2) are satisfied. $\diamondsuit$

Theorems 7.4 and 7.5 show that GDQ is a common generalization of both the classical theory of differentials and the CGJ. Furthermore, it is easy to exhibit maps that have GDQs at a point $\bar{x}$ but are not classically differentiable at $\bar{x}$ and do not have a CGJ or a WDC. (A simple example is provided by the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x \sin 1/x$ if $x \ne 0$, $f(0) = 0$. Then f is not differentiable at 0 and does not have a WDC of f at 0, since f is not Lipschitz continuous near 0. On the other hand, the interval $[-1, 1]$ clearly belongs to $GDQ(f; 0, 0, \mathbb{R})$.)

The previous remarks show that GDQ is "much more general than the classical differential and the CGJ combined." It would then be natural to conjecture that GDQ is also more general than WMD, or at least more general than WDC. This, however, it not true, because *there are Warga derivate containers that are not GDQs.*

For a simple example, consider the function $f : \mathbb{R} \mapsto \mathbb{R}^2 \sim \mathbb{C}$ given by $f(x) = \int_0^x e^{\frac{i}{t}}\, dt$. Then f is Lipschitz continuous. Let $h_j(t) = t^{-1}$ when $|t| \ge j^{-1}$, and $h_j(t) = j^2 t$ when $|t| \le j^{-1}$. Let $f_j(x) = \int_o^x e^{i\, h_j(t)}\, dt$. Then $f_j \in C^1(\mathbb{R}, \mathbb{C})$, and $f_j \to f$ uniformly on compact sets as $j \to \infty$. Furthermore, for every x and every j, the derivative $f_j'(x)$ belongs to the unit circle $\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$. It follows that $\mathbb{S}^1$ *is a Warga derivate container of* f *at* 0. On the other hand, $\mathbb{S}^1$ *is not a GDQ of* f *at* 0. We will actually show that *if* Λ *is any GDQ of* f *at* 0, *then* $0 \in \Lambda$. To see this, assume that $0 \notin \Lambda$, and pick a compact neighborhood Λ' of Λ such that $0 \notin \Lambda'$. Observe that, if $x > 0$, then we can make the change of variables $t = 1/s$ and conclude that $f(x) = \int_{1/x}^{\infty} \frac{e^{is}}{s^2}\, ds$. Integration by parts, with $\xi = 1/x$, then yields $f(x) = -i\xi^{-2}e^{i\xi} + 2i\int_{\xi}^{\infty} \frac{e^{is}}{s^3}\, ds$. A second integration by parts then gives

$$f(x) = -i\xi^{-2}e^{i\xi} + 2\xi^{-3}e^{i\xi} - 6\int_{\xi}^{\infty} \frac{e^{is}}{s^4}\, ds\,.$$

But $6|\int_{\xi}^{\infty} \frac{e^{is}}{s^4}\, ds| \le 6\int_{\xi}^{\infty} s^{-4}\, ds = 2\xi^{-3}$. Therefore $|f(x)| = O(\xi^{-2}) = O(x^2)$ as $x \downarrow 0$. Now, since Λ is a GDQ of f at 0, f must have a factorization $f(x) = g(x)x$ such that g is continuous and Λ'-valued on some punctured neighborhood $]-\alpha, \alpha[\backslash\{0\}$ of 0. Hence there exists a $\delta \in \mathbb{R}$ such that $\delta > 0$ and $|g(x)| \ge \delta$ whenever $0 < x \le \alpha$. On the other hand, the bound $|f(x)| = O(x^2)$ implies that $|g(x)| = O(x)$ as $x \downarrow 0$, and we have reached a contradiction.

9. Conclusion

The remarks at the end of the previous section show that the two "most general" GDTs described so far—that is, GDQ and WMD—do not contain each other. In the second paper of this series we introduce a new GDT—the "path-integral generalized differential"—that contains both.

References

[1] Berkovitz, L. D., *Optimal Control Theory.* Springer-Verlag, New York, 1974.

[2] Botsko, M. W., and R. A. Gosser, *On the differentiability of functions of several variables.* Amer. Math. Monthly **92**, No. 9, 1985, pp. 663–665.

[3] Halkin, H., *Necessary conditions in mathematical programming and optimal control theory.* In *Optimal control theory and its applications, Part I,* Lect. Notes in Econom. and Math. Systems **105**, Springer-Verlag, Berlin, 1974, pp. 113–165.

[4] Halkin, H., *Interior mapping theorem with set-valued derivatives.* Journal d'Analyse Mathématique **30**, 1976, pp. 200–207.

[5] Halkin, H., *Necessary conditions for optimal control problems with differentiable or nondifferentiable data.* In *Mathematical Control Theory,* Lect. Notes in Math. **680**, Springer-Verlag, Berlin, 1978, pp. 77–118.

[6] Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko, *The Mathematical Theory of Optimal Processes.* Wiley, New York, 1962.

[7] Sussmann, H. J., *A strong version of the Lojasiewicz maximum principle.* In *Optimal Control of Differential Equations,* N. H. Pavel Ed., Marcel Dekker Inc., New York, 1994, pp. 293–309.

[8] Sussmann, H. J., *A strong maximum principle for systems of differential inclusions.* In *Proc. 35th IEEE Conference on Decision and Control, Kobe, Japan, Dec. 1996,* IEEE publications, 1996, pp. 1809–1814.

[9] Sussmann, H. J., *Some recent results on the maximum principle of optimal control theory.* In *Systems and Control in the Twenty-First Century,* C. I. Byrnes, B. N. Datta, D. S. Gilliam and C. F. Martin Eds., Birkhäuser, Boston, 1997, pp. 351–372.

[10] Sussmann, H. J., *Multidifferential calculus: chain rule, open mapping and transversal intersection theorems.* In *Optimal Control: Theory, Algorithms, and Applications,* W. W. Hager and P. M. Pardalos, Editors, Kluwer Academic Publishers, 1998, pp. 436–487.

[11] Sussmann, H.J., "A maximum principle for hybrid optimal control problems." In *Proc. 38th IEEE Conf. Decision and Control, Phoenix, AZ, Dec. 1999.* IEEE publications, New York, 1999, pp. 425–430.

[12] Sussmann, H.J., "Résultats récents sur les courbes optimales." In *15^e Journée Annuelle de la Société Mathémathique de France (SMF)*, Publications de la SMF, Paris, 2000, pp. 1-52.

[13] Sussmann, H.J., "New theories of set-valued differentials and new versions of the maximum principle of optimal control theory." In *Nonlinear Control in the year 2000*, A. Isidori, F. Lamnabhi-Lagarrigue and W. Respondek Eds., Springer-Verlag, London, 2000, pp. 487-526.

[14] Sussmann, H.J., "Set-valued differentials and the hybrid maximum principle." In *Proc. 39th IEEE Conf. Decision and Control, Sydney, Australia, December 12-15, 2000*, IEEE publications, New York, 2000.

[15] Tougeron, J.-C., *Idéaux de fonctions différentiables.* Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 71. Springer-Verlag, 1972.

[16] Warga, J., *Fat homeomorphisms and unbounded derivate containers.* J. Math. Anal. Appl. **81**, 1981, pp. 545–560.

[17] Warga, J., *Controllability, extremality and abnormality in nonsmooth optimal control.* J. Optim. Theory Applic. **41**, 1983, pp. 239–260.

[18] Warga, J., *Optimization and controllability without differentiability assumptions.* SIAM J. Control and Optimization **21**, 1983, pp. 837–855.

[19] Warga, J., *Homeomorphisms and local C^1 approximations.* Nonlinear Anal. TMA **12**, 1988, pp. 593–597.

[20] Warga, J., *An extension of the Kaskosz maximum principle.* Applied Math. Optim. **22**, 1990, pp. 61–74.