"Optimal control: Theory, Algorithms, and Applications, pp. 1-2
W. W. Hager and P. M. Pardalos, Editors


# Multidifferential Calculus: Chain Rule, Open Mapping and Transversal Intersection Theorems*

Hector J. Sussmann†
*Department of Mathematics*
*Rutgers University, New Brunswick, NJ 08903*
sussmann@hamilton.rutgers.edu

Received June 27, 1997; Final version September 17, 1997

## 1 Introduction

The purpose of this note is to develop a calculus of "multidifferentials" of set-valued maps between finite-dimensional real linear spaces. This will be done by generalizing the concept of the classical differential (CD), following the ideas of F. Clarke's theory of generalized Jacobians (CGJ's) and J. Warga's theory of derivate containers (WDC's), cf. Clarke [4] and Warga [17, 18, 19, 20].

The calculus proposed here has the properties needed for a proof of a general version of the Pontryagin Maximum Principle (PMP) of optimal control theory, which generalizes the classical version (cf. [1, 8]), the non-smooth one (cf. [3]), and the one —due to S. Łojasiewicz Jr., cf. [7]— for control systems with a continuous right-hand side and a Lipschitz reference vector field, and in addition applies to differential inclusions as well as to vector field systems, and even to many systems with discontinuous vector fields.

The two most important properties needed for the proof are the *chain rule* and the *open mapping and transversal intersection theorems*.

*This research was supported in part by NSF Grant DMS95-00798 and AF0SR Grant 0923.
†The author is grateful to two anonymous referees for their very helpful remarks.

The statement and proof of the PMP itself will not be discussed here, but the chain rule, open mapping and transversal intersection theorems will be stated and proved in full detail. A paper giving a complete self-contained proof of the PMP is now in preparation. Several preliminary announcements have appeared, outlining various steps of the proof (cf. [13, 14, 15, 16]).

**Remark**. *The definition of "multidifferential" proposed here is slightly different from that of our earlier paper [15].* The concept introduced in [15] suffices to establish the weaker open mapping and separation theorems stated there (Theorems 2.3, 2.4 and 3.1 of [15]), but is not enough to guarantee the truth of the transversal intersection result of this paper (Theorem 4.3.3). For this reason, we now believe that the more restrictive definition proposed here ought to supersede that of [15]. Since we still feel that the word used in [15] is the most adequate name for the concept, we have taken the liberty of retaining the word while slightly changing its meaning, rather than creating a new word. □

The proof of the open mapping theorem is an elaboration of ideas originally due to S. Lojasiewicz Jr., and is based on a method introduced by Wazewski (cf. [21]) for the study of implicitly defined mappings.

The multidifferentials at a point $(\bar{x}, \bar{y}) \in X \times Y$ of a set-valued map $F : X \to 2^Y$ —where $X$, $Y$ are finite-dimensional real linear spaces— will be, by definition, compact subsets of the set $L(X, Y)$ of linear maps from $X$ to $Y$.

We will use $MD(F; \bar{x}, \bar{y})$ to denote the set of all multidifferentials of $F$ at $(\bar{x}, \bar{y})$. Since we are dealing with set-valued maps from $X$ to $Y$, we will be talking about multidifferentials *at a point* $(\bar{x}, \bar{y})$ *of* $X \times Y$ rather than at a point $\bar{x} \in X$. Our definitions will be such that the multidifferential, when it exists, is *never* unique, so we will never be allowed to talk about "the" multidifferential of $F$ at $(\bar{x}, \bar{y})$. It will also turn out that $MD(F; \bar{x}, \bar{y}) = \emptyset$ unless $(\bar{x}, \bar{y})$ belongs to the closure of the graph $G(F)$ of $F$. In particular, when $F$ is single-valued and continuous, $MD(F; \bar{x}, \bar{y})$ can only be nonempty if $\bar{y} = F(\bar{x})$, so in this case we can talk about "multidifferentials of $F$ at a point $\bar{x} \in X$," and it is not necessary to specify $\bar{y}$.

In the special case when $X = Y = \mathbb{R}$, $L(X, Y)$ is of course naturally identified with $\mathbb{R}$, so for a multifunction $F : \mathbb{R} \to 2^{\mathbb{R}}$ the set $MD(F; \bar{x}, \bar{y})$ will just be a set of compact subsets of $\mathbb{R}$.

In the proof of the PMP one needs to study composites of maps, some of which have a CD at a point, while others have a CGJ or a WDC. So there is a need for a calculus that combines these three theories into one unified setting, within which it should be possible, for example, to apply the Chain

Rule to a composite $F_2 \circ F_1$ where $F_1$ has a "nice" WDC at a point $p$, and $F_2$ is differentiable in the classical sense at $F_1(p)$ but fails to be Lipschitz near $F_1(p)$. In addition, by combining Lipschitz maps —the natural domain for the CGJ and WDC theories— with classically differentiable ones, one can easily produce maps that fit in none of these categories, so we will of necessity have to consider maps in a class much larger than the union of the domains of the other three theories.

Consider, for example, the functions $f_i : \mathbb{R} \to \mathbb{R}$, $i = 1, 2, 3, 4$, given by

$$\begin{aligned} f_1(x) &= x + 5x^2 \sin\frac{1}{x}\,, \\ f_2(x) &= x + |x|^{3/2} \sin\frac{1}{x}\,, \\ f_3(x) &= 3x + |x| + 5x^2 \sin\frac{1}{x}\,, \\ f_4(x) &= 3x + |x| + |x|^{3/2} \sin\frac{1}{x}\,. \end{aligned}$$

The $f_i$ are continuous, satisfy $f_i(0) = 0$, and are such that

(1) $f_1$ is both differentiable at 0 in the classical sense and Lipschitz near 0; the CD of $f_1$ at 0 is equal to 1, and the CGJ is the interval $[-4, 6]$;

(2) $f_2$ is differentiable at 0 in the classical sense but is not Lipschitz near 0; the CD of $f_2$ at 0 is equal to 1, but the CGJ of $f_2$ at 0 does not exist;

(3) $f_3$ is not differentiable at 0 in the classical sense, but is Lipschitz near 0; the CD of $f_3$ at 0 does not exist, but the CGJ is the interval $[-3, 9]$;

(4) $f_4$ is neither differentiable at 0 in the classical sense nor Lipschitz near 0; the CD and the CGJ of $f_4$ at 0 both fail to exist.

In all four cases, it will turn out that *$f_i$ is multidifferentiable at 0 in the sense of our theory.* Moreover, $\{1\} \in MD(f_1; 0, 0)$, $\{1\} \in MD(f_2; 0, 0)$, $[2, 4] \in MD(f_3; 0, 0)$ and $[2, 4] \in MD(f_4; 0, 0)$.

In general, our theory will assign a multidifferential $\{D\}$ at $\bar{x}$ to any continuous map near $\bar{x}$ which has a CD equal to $D$ at $\bar{x}$, and will also assign a multidifferential $\mathbf{D}$ at $\bar{x}$ to a map $f$ which is Lipschitz near $\bar{x}$ and has CGJ equal to $\mathbf{D}$ at $\bar{x}$. When a function has both a CD and a CGJ at a point, then it may happen, as in the case of our function $f_1$, that the CGJ does not coincide with the CD. In that case, our theory will give *both* the CD and the CGJ as possible multidifferentials.

Moreover, our theory will also yield multidifferentials for maps such as $f_4$, which has neither a CD nor a CGJ at 0. Finally, even when the CD does not exist but the CGJ does, it may happen that our theory yields a multidifferential which is strictly smaller than the CGJ, as shown by our function $f_3$.

The Open Mapping Theorem will say, roughly, that if $\mathbf{D} \in MD(F; \bar{x}, \bar{y})$ is such that every member $D$ of $\mathbf{D}$ is surjective as a map from $X$ to $Y$, then $F$ maps neighborhoods of $\bar{x}$ to neighborhoods of $\bar{y}$.

Notice that, to apply this result to a particular $F$, it is always convenient to select a multidifferential $\mathbf{D}$ which is as small as possible. In the particular case of our four functions $f_i$, if we use the Open Mapping Theorem with the CGJ, then we are *not* able to see that the maps $f_i$ are all open at 0. (For $f_1$ and $f_3$ the CGJ is too large, and in particular contains 0. For $f_2$ and $f_4$ there is no CGJ.) On the other hand, for $f_1$ and $f_2$ we can see the openness at 0 by using the Open Mapping Theorem with the classical differential. For $f_3$ and $f_4$ the CD does not exist and, as we have seen, the CGJ is either insufficient (for $f_3$) or nonexistent (for $f_4$). Our theory, however, yields in all four cases a multidifferential which is small enough to enable us to see on the infinitesimal level that the map $f_i$ is open at 0.

More generally, we will want to consider "directional multidifferentials" of a set-valued map $F : X \to 2^Y$, at a point $(\bar{x}, \bar{y}) \in X \times Y$, in the direction of a convex cone $C$ in $X$. We will use $MD_C(F; \bar{x}, \bar{y})$ to denote the set of all such multidifferentials. The Open Mapping Theorem will say, in this setting, that if $\mathbf{D} \in MD_C(F; \bar{x}, \bar{y})$ is such that $DC = Y$ for all $D \in \mathbf{D}$, then $F$ maps relative neighborhoods of $\bar{x}$ in $\bar{x} + C$ to neighborhoods of $\bar{y}$ in $Y$.

Even more generally, we will be interested in *transversal intersections* of sets of the form $S_1 = F_1(\bar{x}_1 + C_1)$, $S_2 = F_2(\bar{x}_2 + C_2)$, where $F_i : X_i \to 2^Y$, $i = 1, 2$, are set-valued maps having multidifferentials $\mathbf{D}_i \in MD_{C_i}(F_i; \bar{x}_i, \bar{y})$. In this case, our goal will be to prove that a "transversality property" of the infinitesimal objects $(C_i, \mathbf{D}_i)$ implies that the intersection $S_1 \cap S_2$ contains at least a sequence of points $y_j$ such that $y_j \neq \bar{y}$ and $y_j \to \bar{y}$. Even more strongly, we will show that the conclusion holds if $S_2$ is replaced by a set $S_2^* = f_2(\bar{x}_2 + C_2^*)$, where $C_2^*$ is a relatively closed subset of $C_2$ that "separates 0 from the rest of $C_2$," in a sense to be made precise below.

We illustrate the transversal intersection result with a simple example. Let $C_1$ be the cone in $\mathbb{R}^3$ given by $C_1 = \{(x, y, z) : z = 0, |y| \leq x\}$. Let $C_2$ be given by $C_2 = \{(x, y, z) : y = 0, |z| \leq x\}$. Then it is intuitively clear that $C_1$ and $C_2$ "intersect transversally along the half-line $L = \{(x, 0, 0) : x \geq 0\}$" (cf. Figure 1).

The precise mathematical definition of *transversality* of two cones $C_1$,

$C_2$ in a finite-dimensional real linear space $X$ turns out to be as follows: we say that $C_1$ and $C_2$ are *transversal*, and write $C_1\overline{\pitchfork}C_2$, if $C_1 - C_2 = X$. Equivalently, $C_1\overline{\pitchfork}C_2$ iff there exists no nontrivial linear functional $\zeta$ on $X$ such that $\zeta \leq 0$ on $C_1$ and $\zeta \geq 0$ on $C_2$, i.e. iff $C_1^\dagger \cap (-C_2)^\dagger = \{0\}$, where $C^\dagger$ stands for "polar cone of $C$." (When $C_1$ and $C_2$ are linear subspaces this reduces to the usual condition that $C_1 + C_2 = X$.)

We then call $C_1$ and $C_2$ *strongly transversal* —and write $C_1\overline{\overline{\pitchfork}}C_2$— if $C_1\overline{\pitchfork}C_2$ and in addition $C_1 \cap C_2 \neq \{0\}$.

With these definitions, the "intuitively clear" statement made earlier about our two cones in $\mathbb{R}^3$ can be made precise. To begin with, the two cones are strongly transversal. In addition, one can show that the strong transversality of $C_1$ and $C_2$ implies that, even if $C_1$ and $C_2$ are "slightly perturbed," the resulting sets have an infinite intersection.

The following is one example of a rigorous version of the above "slight perturbation" statement: suppose $f_i : C_i \to \mathbb{R}^3$ are, for $i = 1, 2$, continuous maps such that $f_i(v) = v + o(\|v\|)$ as $v \to_{C_i} 0$. Then $f_1(C_1) \cap f_2(C_2)$ contains an infinite sequence converging to 0.

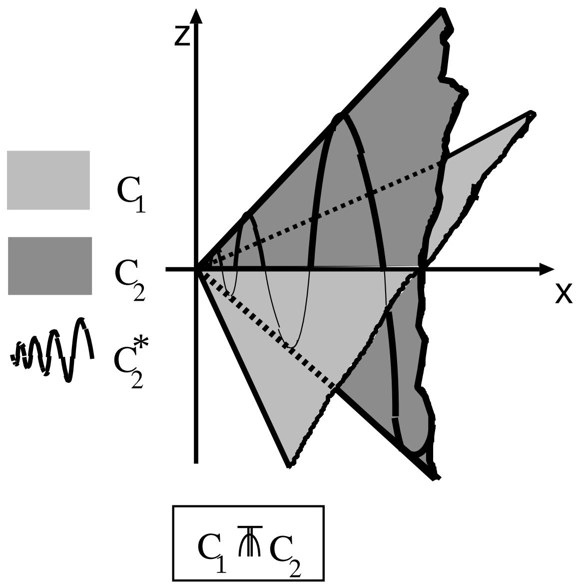


Figure 1: A Separating Set and a Transversal Intersection

It turns out that a much stronger "transversal intersection" result is true. Not only can it be proved that $f_1(C_1) \cap f_2(C_2)$ contains an infinite sequence, but the same will be true for $f_1(C_1) \cap f_2(C_2^*)$, even if $C_2^*$ is a "small" subset of $C_2$, such as, for example,

$$C_2^* = \left\{ (x, y, z) : x \geq 0, y = 0, z = x \sin \frac{1}{x} \right\}.$$

It is intuitively clear that, even though $C_2^*$ is only 1-dimensional, it "crosses back and forth" from one side of $C_1$ to the other infinitely many times, and the existence of infinitely many crossings is a "robust" phenomenon, which will not be affected by a small perturbation. The *transversal intersection theorem for classically differentiable maps* (Theorem 2.2.4) makes this intuition precise, and then our *transversal intersection theorem for multidifferentiable maps* (Theorem 4.3.3) extends the result to a much larger class of situations. In both the classical differential and the multidifferential case, the transversal intersection theorem trivially implies the open mapping theorem, as shown below in Remark 2.2.5 (for the CD case) and in §4.4 (for the multidifferential case). So the logical relationship among the four theorems is as follows:

$$\begin{tikzcd}
\boxed{\begin{array}{c}\text{Transversal Intersection}\\ \text{Theorem for}\\ \text{Multidifferentials}\\ \text{(Theorem 4.3.3)}\end{array}} \arrow[r, Rightarrow] \arrow[d, Rightarrow] & \boxed{\begin{array}{c}\text{Transversal Intersection}\\ \text{Theorem for}\\ \text{Classical Differentials}\\ \text{(Theorem 2.2.4)}\end{array}} \arrow[d, Rightarrow] \\
\boxed{\begin{array}{c}\text{Open Mapping}\\ \text{Theorem for}\\ \text{Multidifferentials}\\ \text{(Theorem 4.4.1)}\end{array}} \arrow[r, Rightarrow] & \boxed{\begin{array}{c}\text{Open Mapping}\\ \text{Theorem for}\\ \text{Classical Differentials}\\ \text{(Theorems 2.1.1, 2.1.3, and 2.1.4)}\end{array}}
\end{tikzcd}$$

This means that, in order to establish the four theorems, it would suffice to prove Theorem 4.3.3, since the other three results would then follow easily.

We choose, however, to make the exposition clearer by first presenting, with proofs, the theorems for classical differentials. We do this both because the proofs in this case are very simple, and because these proofs display rather clearly the *topological* and *finite-dimensional* nature of the results, since the proof of Theorem 2.1.1 makes essential use of the Brouwer Fixed Point Theorem, and that of Theorem 2.2.4 depends crucially on a result about the existence of certain connected sets of zeros of homotopies proved by F. Browder in [2] (cf. also Rabinowitz [9]).

This will help the reader appreciate the extent to which the proof of Theorem 4.3.3 is built on that of Theorem 2.2.4, and how it diverges from it. The key new ideas are (a) the use of Wazewski's method to construct nonlinear substitutes for the linear inverses used in the proofs of the CD theorems, and (b) the use of a set-valued generalization (Theorem 3.5.4, of which we give a self-contained proof for completeness) of the result of Browder [2] on connected zero sets of homotopies

The paper is organized as follows. In §2 we state and prove the results for CD's that we will want to generalize later, namely, the open mapping and transversal intersection theorems. We then define and study the class of "regular set-valued maps" in §3 proving, in particular, the set-valued generalization of the theorem on connected zero sets of homotopies. We then define multidifferentials in §4.1, state and prove the chain rule in §4.2, and then prove the transversal intersection theorem in §4.3, and use it in §4.4 to derive the open mapping and directional open mapping theorems for multidifferentials.

# 2 Maps with Classical Differentials

*Throughout this section, we assume that $X$ and $Y$ are finite-dimensional real linear spaces, and use $Z$ to denote the space $L(X, Y)$ of all linear maps from $X$ to $Y$.*

## 2.1 The Classical Open Mapping Theorems

**Theorem 2.1.1** *Let $X, Y$ be finite-dimensional real linear spaces, and write $Z = L(X, Y)$. Let $U$ be a neighborhood of $\bar{x} \in X$, and let $f : U \to Y$ be continuous and such that $f$ is differentiable at $\bar{x}$ and the differential $Df(\bar{x}) \in Z$ is surjective. Then $f$ is open at $\bar{x}$, in the sense that, if $U'$ is any neighborhood of $\bar{x}$ in $X$ then $f(U')$ is a neighborhood of $f(\bar{x})$ in $Y$.*

This elementary result is well known. Moreover, it is also a special case of the much more general theorems to be proved later, as explained in §1. So there really is no need to give a proof at this point. We prefer to give the proof anyhow, since it is extremely simple and provides a good introduction to the ideas involved in the generalizations to be discussed later.

*Proof of Theorem 2.1.1.* Assume, as we clearly may without loss of generality, that $\bar{x} = 0$, $f(\bar{x}) = 0$. Let $L = Df(0)$, so $L$ is a surjective linear map from $X$ to $Y$. Let $M \in L(Y, X)$ be such that $L \circ M = \mathrm{id}_Y$. Let $B_Y$ denote the closed unit ball in $Y$, relative to some norm. Let $U'$ be a neighborhood of 0 in $X$.

For $\varepsilon > 0$ sufficiently small, define a map $\psi_\varepsilon : B_Y \to Y$ by letting

$$\psi_\varepsilon(y) = \frac{1}{\varepsilon} f(\varepsilon M y) \,. \tag{2.1.1}$$

Then $\psi_\varepsilon$ is continuous, and $\psi_\varepsilon(y) = y + o(1)$ as $\varepsilon \downarrow 0$, uniformly on $B_Y$. Fix $\alpha$ such that $0 < \alpha < 1$, and choose $\varepsilon$ such that $\varepsilon M B_Y \subseteq U'$ and $\|\psi_\varepsilon(y) - y\| \leq 1 - \alpha$ for all $y \in B_Y$. For $y \in \alpha B_Y$, let $\rho_y : B_Y \to Y$ be the map given by $\rho_y(y') = y + y' - \psi_\varepsilon(y')$. Then $\rho_y$ is a continuous map from $B_Y$ to $B_Y$, so there exists $y' \in B_Y$ such that $\rho_y(y') = y'$, by the Brouwer Fixed Point Theorem (cf., e.g., [11]). This $y'$ then satisfies $\psi_\varepsilon(y') = y$, so $y \in \frac{1}{\varepsilon} f(\varepsilon M B_Y)$, and then $\varepsilon y \in f(U')$. Since $y \in \alpha B_Y$ was arbitrary, we see that $f(U')$ contains the ball $\varepsilon \alpha B_Y$, and the proof is complete. □

**Remark 2.1.2** Theorem 2.1.1 is *not* a special case of the results usually known as the Implicit Function Theorem or Inverse Function Theorem. Indeed, these theorems require $f$ to be of class $C^1$, and would imply, if applicable, the existence of a right-inverse of $f$ near $\bar{y}$, i.e. a continuous $U$-valued map $g$, defined on a neighborhood $V$ of $\bar{y}$, such that $g(\bar{y}) = \bar{x}$ and $f \circ g$ is the identity map of $V$. On the other hand, *under the hypothesis of Theorem 2.1.1 a right inverse need not exist.* To see this, consider the example of the function

$f : \mathbb{R} \to \mathbb{R}$ given by $f(0) = 0$ and $f(x) = x + 2x^2 \sin \frac{1}{x}$ when $x \neq 0$. Then $f$ is continuous

on $\mathbb{R}$, differentiable at 0 with derivative equal to 1, but not of class $C^1$ near 0. Clearly, the derivative of $f$ takes negative and positive values on every neighborhood of 0, from which it follows that $f$ is not monotonic on any neighborhood of 0, so $g$ cannot exist.

In addition, *Theorem 2.1.1 is a truly finite-dimensional result*, whereas its analogue for maps of class $C^1$ is valid under more general conditions, when $X$ and $Y$ are Banach spaces. Actually, the analogue of Theorem 2.1.1 is valid on Banach spaces under the condition that $f$ is Fréchet differentiable at $\bar{x}$, with a surjective differential $L$, provided that

$$\lim_{x \to \bar{x}, x' \to \bar{x}, x \neq x'} \frac{f(x) - f(x') - L.(x - x')}{\|x - x'\|} = 0 \,. \tag{2.1.2}$$

(This result is known as "Graves' theorem," cf. [5].) Clearly, (2.1.2) holds when $f$ is of class $C^1$ near $\bar{x}$, but need not hold if $f$ is just continuous near $\bar{x}$ and Fréchet-differentiable at $\bar{x}$ with differential $L$.

The finite-dimensionality of $Y$ was used essentially in the proof, since the argument depends crucially on the Brouwer Fixed Point Theorem, which is not valid in infinite dimensions. It is not hard to show that the natural infinite-dimensional analogue of Theorem 2.1.1 on Banach spaces is in fact false. □

There is a simple generalization of Theorem 2.1.1 to maps defined on conic neighborhoods. Recall that a *cone* in a finite-dimensional real linear space $X$ is, by definition, a nonempty subset $C$ of $X$ with the property that $(\forall x \in C)(\forall r \in \mathbb{R})(r \geq 0 \Rightarrow r.x \in C)$. (In particular, if $C$ is a cone in $X$ then $0 \in C$.) Recall also that if a $Y$-valued map $f$ is defined on a set of the form $U \cap (\bar{x} + C)$, where $U$ is a neighborhood of $\bar{x}$ in $X$, and $L : X \to Y$ is a linear map, then $f$ is *differentiable at $\bar{x}$ in the direction of $C$ with differential* $L$ if

$$f(\bar{x} + h) = f(\bar{x}) + L.h + o(\|h\|) \quad \text{as} \quad h \to 0\,, \ h \in C\,.$$

(In that case, $L$ is uniquely determined by $f$ if the cone $C$ linearly spans the whole space $X$, and we can then use a notation such as $L = D_C f(\bar{x})$. If $C$ does not span $X$, then only the restriction of $L$ to the linear span of $C$ is determined by $f$.)

We then have

**Theorem 2.1.3** *Let $X, Y$ be finite-dimensional real linear spaces, and write $Z = L(X, Y)$. Let $U$ be a neighborhood of $\bar{x} \in X$, let $C$ be a convex cone in $X$, and let $f : U \cap (\bar{x} + C) \to Y$ be continuous and such that $f$ is differentiable at $\bar{x}$ in the direction of $C$, with a differential $L \in Z$ such that $LC = Y$. Then $f$ is open at $\bar{x}$ along $C$, in the sense that, if $U'$ is any neighborhood of $\bar{x}$ in $X$ then $f(U' \cap (\bar{x} + C))$ is a neighborhood of $f(\bar{x})$ in $Y$.*

This result can be proved by a fairly simple modification of the argument that we used to prove Theorem 2.1.1. We omit the proof, since in any case Theorem 2.1.3 is clearly contained in the more general Theorem 4.4.1 for multidifferentiable maps. □

Finally, we point out that there is also a "directional" conic open mapping theorem:

**Theorem 2.1.4** *Let $X, Y$ be finite-dimensional real linear spaces, and write $Z = L(X, Y)$. Let $v \in Y$, let $U$ be a neighborhood of $\bar{x} \in X$, let $C$ be a convex cone in $X$, and suppose that $f : U \cap (\bar{x} + C) \to Y$ is continuous and such that $f$ is differentiable at $\bar{x}$ in the direction of $C$, with a differential $L \in Z$ such that $v$ is an interior point of $LC$. Then $f$ is $v$-directionally* *open at $\bar{x}$ along $C$, in the sense that, if $U'$ is any neighborhood of $\bar{x}$ in $X$ then $f(U' \cap (\bar{x}+C))$ contains a directional $v$-neighborhood of $f(\bar{x})$ in $Y$, i.e. a set of the form $V \cap (\bar{x}+C')$, where $C'$ is a convex cone in $Y$ such that $v \in \mathrm{Int}_Y(C')$, and $V$ is a neighborhood of $f(\bar{x})$.*

Theorem 2.1.3 is just a special case of Theorem 2.1.4, corresponding to $v = 0$. In addition, Theorem 2.1.4, like Theorem 2.1.3, can be easily proved by adapting the method used to prove Theorem 2.1.1. Moreover, Theorem 2.1.4 is a special case of the directional conic open mapping theorem 4.4.1 for multidifferentiable maps, so once again we omit the proof. □

## 2.2 The Transversal Intersection Theorem

*Throughout this subsection, we assume that $X$, $X_1$, $X_2$ and $Y$ are finite-dimensional real linear spaces, and use $Z$, $Z_1$, $Z_2$ to denote the spaces $L(X,Y)$, $L(X_1,Y)$, $L(X_2,Y)$ of $Y$-valued linear maps on $X$, $X_1$, $X_2$.*

**Definition 2.2.1** If $C_1$, $C_2$ are convex cones in a finite-dimensional real linear space $X$, we say that $C_1$ and $C_2$ are *transversal*, and write $C_1 \pitchfork C_2$, if $C_1 - C_2 = X$. (Here $C_1 - C_2$ is, of course, the set of all differences $c_1 - c_2$, $c_i \in C_i$.) We say that $C_1$ and $C_2$ are *strongly transversal*, and write $C_1 \pitchfork\!\!\!\pitchfork C_2$, if $C_1 \pitchfork < C_2$ and $C_1 \cap C_2 \neq \{0\}$. □

**Lemma 2.2.2** *Let $C_1$, $C_2$ be convex cones in $X$ such that $C_1 \pitchfork C_2$. Then either (i) $C_1 \pitchfork\!\!\!\pitchfork C_2$ or (ii) $C_1$ and $C_2$ are linear subspaces of $X$ such that $X = C_1 \oplus C_2$.*

*Proof.* Suppose that the cones $C_1$ and $C_2$ are transversal but not strongly transversal. Then $C_1 \cap C_2 = \{0\}$. We prove that $C_1$ is a linear subspace. Let $c \in C_1$. Since $C_1 - C_2 = X$, we can write $-c = c_1 - c_2$, $c_i \in C_i$. Then $c + c_1 = c_2$, so $c + c_1 \in C_1 \cap C_2$. Therefore $c + c_1 = 0$. So $-c = c_1$, showing that $-c \in C_1$. Therefore $-C_1 = C_1$, so $C_1$ is a subspace as stated. The same argument obviously works for $C_2$. □

**Definition 2.2.3** If $C$ is a convex cone in a finite-dimensional real linear space $X$, and $C^*$ is a relatively closed subset of $C$, we say that $C^*$ *separates* $0$ *from the rest of* $C$ if the following condition holds:

(SEP) *for every neighborhood $U$ of $0$ in $X$ there is a neighborhood $V$ of $0$ such that, if $\gamma : [0,1] \to C$ is continuous, $\gamma(0) \in V$ and $\gamma(1) \notin U$, it follows that $\gamma(t) \in C^*$ for some $t \in [0,1]$.*

(In simpler terms, (SEP) says, roughly, that one cannot go from points close to 0 to points far from 0 without crossing $C^*$.) $\square$

**Theorem 2.2.4** *For $i = 1, 2$, let $X_i$ be a finite-dimensional real linear space, let $\bar{x}_i \in X_i$, let $U_i$ be a neighborhood of $\bar{x}_i$ in $X_i$, and let $C_i$ be a convex cone in $X_i$. Let $S_i = U_i \cap (\bar{x}_i + C_i)$. Let $f_i : S_i \to Y$ be continuous maps, differentiable at $\bar{x}_i$ in the direction of $C_i$ with differential $L_i$, and such that $f_1(\bar{x}_1) = f_2(\bar{x}_2)$. Let $\bar{y} = f_1(\bar{x}_1)$. Let $C_2^*$ be a relatively closed subset of $C_2$ that separates 0 from the rest of $C_2$. Let $S_2^* = U_2 \cap (\bar{x}_2 + C_2^*)$. Assume that $L_1C_1 \overline{\pitchfork} L_2C_2$. Then the set $f_1(S_1) \cap f_2(S_2^*)$ contains a sequence of points $y_j$ such that $y_j \neq \bar{y}$ and $y_j \to \bar{y}$.*

*Proof.* Assume, without loss of generality, that $\bar{x}_1 = 0$, $\bar{x}_2 = 0$, $\bar{y} = 0$. Fix a $w \in L_1C_1 \cap L_2C_2$ such that $w \neq 0$, and then a linear functional $\mu : Y \to \mathbb{R}$ such that $\mu(w) = 1$. Write $w = L_1v_1 = L_2v_2$, $v_i \in C_i$.

Let $n = \dim Y$. Fix a norm $\|\ldots\|$ on $Y$. Pick $w^0, \ldots, w^n \in Y$ such that the closed unit ball $B_Y$ of $Y$ is contained in the interior of the convex hull of the $w^k$. Then each $w^k$ can be expressed as

$$w^k = L_1v_1^k - L_2v_2^k\,, \quad v_i^k \in C_i\,. \tag{2.2.1}$$

Given any $r \geq 0$, we can always substitute the pair $(v_1^k + rv_1, v_2^k + rv_2)$ for the pair $(v_1^k, v_2^k)$. In particular, we can always choose the representations (2.2.1) so that $\mu(L_1v_1^k) \geq 1$ and $\mu(L_2v_2^k) \geq 1$.

Given any $y \in B_Y$, we can represent $y$ uniquely as a linear combination $y = \sum_{k=0}^n \theta_k(y)w^k$ such that $\sum_k \theta_k(y) = 1$ and $\theta_k(y) \geq 0$ for $k = 0, \ldots, n$. For $\varepsilon > 0$, $y \in B_Y$, define

$$\psi_\varepsilon(y) = \frac{1}{\varepsilon}\left(f_1\Big(\varepsilon \sum_k \theta_k(y)v_1^k\Big) - f_2\Big(\varepsilon \sum_k \theta_k(y)v_2^k\Big)\right). \tag{2.2.2}$$

Then $\psi_\varepsilon(y)$ is defined for $y \in B_Y$, $\varepsilon \in ]0, \bar{\varepsilon}]$, for some $\bar{\varepsilon} > 0$, and is jointly continuous in $(y, \varepsilon)$. Moreover, $\psi_\varepsilon(y) = y + o(1)$ as $\varepsilon \downarrow 0$, uniformly with respect to $y$. So there is a continuous map $\Xi : B_Y \times [0, \bar{\varepsilon}] \to Y$ such that $\Xi(y, \varepsilon) = \psi_\varepsilon(y)$ for $\varepsilon > 0$, and $\Xi(y, 0) = y$. By making $\bar{\varepsilon}$ smaller, if necessary, we may assume that $\Xi(y, \varepsilon) \neq 0$ whenever $\|y\| = 1$ and $\varepsilon \in [0, \bar{\varepsilon}]$.

It then follows from the theorem of F. Browder [2] (which is also a special case of a slightly more general result that will be given a self-contained proof in Theorem 3.5.4 below) that for every $\alpha$, $\beta$ such that $0 \leq \alpha \leq \beta \leq \bar{\varepsilon}$ there exists a compact connected subset $\Sigma_{\alpha,\beta}$ of $B_Y \times [\alpha, \beta]$ such that $\Xi(y, \varepsilon) = 0$ whenever $(y, \varepsilon) \in \Sigma_{\alpha,\beta}$, $\Sigma_{\alpha,\beta} \cap (B_Y \times \{\alpha\}) \neq \emptyset$, and $\Sigma_{\alpha,\beta} \cap (B_Y \times \{\beta\}) \neq \emptyset$.

Now, let $\nu_i : B_Y \times [0,\bar{\varepsilon}] \to C_i$ be, for $i = 1,2$, the maps defined by

$$\nu_i(y,\varepsilon) = \varepsilon \sum_k \theta_k(y) v_i^k \,.$$

Then, if $c \in \nu_2(B_Y \times [0,\bar{\varepsilon}])$, and $c = \nu_2(y,\varepsilon) = \varepsilon \sum_k \theta_k(y) v_2^k$, we have the bounds

$$\varepsilon \le \mu(L_2 c) \le \varepsilon K \,, \tag{2.2.3}$$

where $K = \max\{\mu(L_2 v_2^k) : k = 0,\ldots,n\}$. Since $f_2(x) = L_2 x + o(\|x\|)$ as $x \to 0$ via values in $C_2$, we have an estimate $\|f_2(c)\| \ge \|L_2.c\| + o(\varepsilon)$. On the other hand, (2.2.3) implies that $\|L_2.c\| \ge K'\varepsilon$ for some constant $K' > 0$. So, if we choose $\bar{\varepsilon}$ small enough, we can guarantee that $\|f_2(c)\| > 0$, i.e. that $f_2(c) \ne 0$, whenever $c = \nu_2(y,\varepsilon)$ and $0 < \varepsilon \le \bar{\varepsilon}$. Also, we can choose $\bar{\varepsilon}$ so small that $\nu_i(B_Y \times [0,\bar{\varepsilon}]) \subseteq U_i$ for $i = 1,2$.

It is clear that $f_1(\nu_1(y,\varepsilon)) = f_2(\nu_2(y,\varepsilon))$ whenever $(y,\varepsilon) \in \Sigma_{\alpha,\beta}$ for some pair $(\alpha,\beta)$. We will show that

(*) *given any $\beta$ such that $0 < \beta \le \bar{\varepsilon}$, there exists $\alpha$ such that $0 < \alpha < \beta$ and $\nu_2(\Sigma_{\alpha,\beta}) \cap C_2^* \ne \emptyset$.*

Let us first prove that (*) implies our desired conclusion. Choose a sequence $\{\beta_j\}$ such that $\beta_j \downarrow 0$, and then pick for each $j$ an $\alpha_j$ with the property of (*). We can then choose points $(z_j,\varepsilon_j) \in \Sigma_{\alpha_j,\beta_j}$ such that, if we let $c_j = \nu_2(z_j,\varepsilon_j)$, then $c_j \in C_2^*$. If we then write $y_j = f_2(c_j)$, we have $\|y_j\| \le \beta_j$.constant, so $y_j \to 0$. Also, $y_j = f_2(c_j) = f_1(\nu_1(z_j,\varepsilon_j))$, since $(z_j,\varepsilon_j) \in \Sigma_{\alpha_j,\beta_j}$, so $y_j \in f_1(C_1 \cap U_1) \cap f_2(C_2^* \cap U_2)$, since $\nu_i(z_j,\varepsilon_j) \in U_i$. Finally, it is clear that $y_j = f_2(c_j) \ne 0$, since $\varepsilon_j > 0$. So the sequence $\{y_j\}$ satisfies all the requirements of our conclusion.

To conclude our proof, we show that (*) holds. Given $\beta$, all the sets $\nu_2(\Sigma_{\alpha,\beta})$ are compact and connected, and contain at least one point in $C_2 \backslash W(\beta)$, where

$$W(\beta) = \{x \in X_2 : \mu(L_2.x) < \beta\} \,. \tag{2.2.4}$$

(Indeed, pick $y \in B_Y$ such that $(y,\beta) \in \Sigma_{\alpha,\beta}$. Then $c = \nu_2(y,\beta)$ satisfies $c \in C_2$ and $\mu(L_2.c) \ge \beta$, so $c \notin W(\beta)$.)

Since $C_2^*$ separates 0 from the rest of $C_2$, there is a neighborhood $W'(\beta)$ of 0 in $X_2$ with the property that every continuous path $\gamma : [0,1] \to C_2$ such that $\gamma(0) \in W'(\beta)$ and $\gamma(1) \notin W(\beta)$ necessarily intersects $C_2^*$. It then follows that no path-connected component of $C_2 \backslash C_2^*$ can intersect both sets $W'(\beta)$ and $C_2 \backslash C_2^*$. Since $C_2 \backslash C_2^*$ is a relatively open subset of a convex subset of $X_2$, the path-connected components of $C_2 \backslash C_2^*$ are exactly the connected components of $C_2 \backslash C_2^*$. It then follows that no connected component of $C_2 \backslash C_2^*$ can intersect both sets $W'(\beta)$ and $C_2 \backslash W(\beta)$. Now choose $\alpha > 0$ small enough so that $\nu_2(y, \alpha) \in W'(\beta)$ for all $y \in B_Y$. Then $\nu_2(\Sigma_{\alpha,\beta})$ is a connected subset of $C_2$ which intersects $C_2 \backslash W(\beta)$. Moreover, since $\Sigma_{\alpha,\beta} \cap (B_Y \times \{\alpha\}) \neq \emptyset$, we can pick $y \in B_Y$ such that $(y, \alpha) \in \Sigma_{\alpha,\beta}$, and then $\nu_2(y, \alpha) \in W'(\beta)$. So $\nu_2(\Sigma_{\alpha,\beta})$ is a connected subset of $C_2$ which intersects both $C_2 \backslash W(\beta)$ and $W'(\beta)$. It then follows that $\nu_2(\Sigma_{\alpha,\beta})$ cannot be entirely contained in $C_2 \backslash C_2^*$. So $\nu_2(\Sigma_{\alpha,\beta}) \cap C_2^* \neq \emptyset$, and our proof is complete. □

**Remark 2.2.5** Theorem 2.2.4 contains the open mapping theorems 2.1.1, 2.1.3 and 2.1.4 as special cases. Indeed, suppose $X$, $\bar{x}$, $U$, $C$, $Y$, $Z$, $L$, $v$ satisfy the hypotheses of Theorem 2.1.4. Assume that $\bar{x} = 0$ and $f(\bar{x}) = 0$. Let $U'$ be a neighborhood of 0 in $X$. Assume $||v|| = 1$. (This is possible if $v \neq 0$. The case $v = 0$ follows easily if Theorem 2.1.4 is proved for $v \neq 0$.) Suppose $f(U' \cap C)$ does not contain a set of the form $\tilde{C} \cap V$, with $V$ a neighborhood of 0 and $\tilde{C}$ a convex cone containing $v$ in its interior. Then there exists a sequence $\{y_j\}$ of points of $Y$ such that $y_j \to 0$, $y_j \neq 0$, $y_j \notin f(U' \cap C)$, $v = \lim_{j \to \infty} \frac{y_j}{\|y_j\|}$, and the the sequence $\{\|y_j\|\}_{j=1}^{\infty}$ is strictly decreasing. Let $X_2 = \mathbb{R}$, and $C_2$ be the cone $[0, +\infty[$. Let $U_2$ be the interval $[-1, \|y_1\|]$. For $r \in U_2 \cap C_2$, define $f_2(r)$ by letting $f_2(0) = 0$ and

$$f_2(r) = \alpha y_{j+1} + (1-\alpha) y_j \ \text{ if } \ r = \alpha \|y_{j+1}\| + (1-\alpha)\|y_j\| \, , \ 0 \leq \alpha \leq 1 \, .$$

Then $f_2$ is clearly continuous. Moreover, if $r = \alpha\|y_{j+1}\| + (1-\alpha)\|y_j\|$, $0 \leq \alpha \leq 1$, then

$$\begin{aligned} f_2(r) - rv &= (\alpha y_{j+1} + (1-\alpha) y_j) - (\alpha \|y_{j+1}\| + (1-\alpha)\|y_j\|) v \\ &= \alpha z_{j+1} + (1-\alpha) z_j \end{aligned}$$

where $z_j = y_j - \|y_j\| v$. Since $z_j = \|y_j\| w_j$, with $w_j \to 0$, it is clear that $\|f_2(r) - rv\| \leq r\theta(r)$, where $\theta(r) = \max(\|w_j\|, \|w_{j+1}\|)$. Since $\theta(r) \to 0$ as $r \downarrow 0$, $f_2$ is differentiable at 0 along $C_2$ with differential $L_2$, where $L_2$ is the linear map $\mathbb{R} \ni r \to rv \in Y$. The set $C_2^* = \{\|y_j\| : j = 1, 2, \ldots\}$ obviously separates 0 from the rest of $C_2$. Moreover, $LC$ and $L_2 C_2$ are clearly strongly transversal. So Theorem 2.2.4 implies that $f(U' \cap C) \cap f_2(C_2^*)$ contains a sequence of nonzero points converging to zero. Since $f_2(C_2^*)$ is exactly the set of the $y_j$, it follows that some $y_j$ must belong to $f(U' \cap C)$, contradicting our choice of the sequence $\{y_j\}$. □

# 3 Regular Set-Valued Maps

*Throughout this section we use* $SVM(A,B)$ *to denote the set of all set-valued maps from a set* $A$ *to a set* $B$*, so* $F \in SVM(A,B)$ *iff* $F$ *is an ordinary map from* $A$ *to* $2^B$*. The* **graph** *of a set-valued map* $F \in SVM(A,B)$ *is the set*

$$G(F) = \{(a,b) \in A \times B : b \in F(a)\} .$$

*The domain* $\mathrm{Dom}(F)$ *of* $F$ *is the set* $\{a \in A : F(a) \neq \emptyset\}$*. We use* $d_X$ *to denote the distance function of a metric space* $X$*. We assume, unless otherwise stated, that* $K$ *is a compact metric space and* $Y$ *is a metric space. The product* $K \times Y$ *is equipped with the product metric, in which*

$$d_{K\times Y}((k,y),(k',y')) = d_K(k,k') + d_Y(y,y') \quad \text{if} \quad k,k' \in K\ ,\ \ y,y' \in Y\ .$$

We now develop in detail the theory of "regular maps," i.e. the class of set-valued maps that results by taking all possible limits of ordinary single-valued continuous maps from $K$ to $Y$ in the sense of graph convergence. We first define graph-convergence of set-valued maps, and then proceed to define regular maps and to prove some of their most important properties, such as fixed-point and homotopy theorems. Since these theorems will play a crucial role in the proof of the the transversal intersection theorem for multidifferentiable maps, we give a completely self-contained proof.

## 3.1 Set Convergence

If $T$ is an arbitrary topological space, then there is a natural topology $\mathcal{T}_T$ on the set $2^T$ of all subsets of $T$, defined by letting $\mathcal{T}_T$ be the topology generated by the power sets of open sets, i.e. by the sets $2^\Omega$, for all open subsets $\Omega$ of $T$. Since $2^{\Omega_1} \cap 2^{\Omega_2} = 2^{\Omega_1 \cap \Omega_2}$, the sets $2^\Omega$ form a basis for a topology. This means that a subset $S$ of $2^T$ is open iff it is a union $\cup_\alpha 2^{\Omega_\alpha}$ of a family of power sets of open sets, and $S$ is a neighborhood of a subset $Y$ of $T$ iff there exists $\Omega$, open in $T$, such that $Y \subseteq \Omega$ and $2^\Omega \subseteq S$.

It follows that a net $\{S_\alpha\}_{\alpha\in A,\preceq}$ in $2^T$ converges in $\mathcal{T}_T$ to a set $S \in 2^T$ if and only if for every neighborhood $\Omega$ of $S$ there exists $\alpha(\Omega) \in A$ such that $S_\alpha \subseteq \Omega$ whenever $\alpha(\Omega) \preceq \alpha$.

The topology $\mathcal{T}_T$ is never Hausdorff. (For example: (a) the empty set belongs to every nonempty open set of $\mathcal{T}_T$, and (b) if $S_1$ is a proper subset of $S_2$ then every neighborhood of $S_2$ contains $S_1$.)

Now let $X$ be a metric space. We will be particularly interested in the topological subspace $\mathrm{Comp}(X)$ of all nonempty compact subsets of $X$, endowed with the subspace topology of $(2^X, \mathcal{T}_X)$. We can define the "distance"

$\Delta(L_1, L_2)$ between two nonempty subsets of $X$ by letting

$$\Delta(L_1, L_2) = \sup_{p \in L_1} \operatorname{dist}(p, L_2)\,. \tag{3.1.1}$$

This function is not in general symmetric (for example, if $L_1 \subseteq L_2$, $L_1 \neq L_2$, and $L_1$ is closed, then $\Delta(L_1, L_2) = 0$ but $\Delta(L_2, L_1) > 0$), and it can take infinite values. On the other hand, $\Delta$ satisfies the "triangle inequality"

$$\Delta(L_1, L_3) \leq \Delta(L_1, L_2) + \Delta(L_2, L_3)\,, \tag{3.1.2}$$

if $L_1$, $L_2$, $L_3$ are arbitrary nonempty subsets of $X$.

If $L \in \operatorname{Comp}(X)$, then $L$ has a fundamental family of neighborhoods given by the sets $V(L, \varepsilon)$, for $\varepsilon > 0$, where

$$V(L, \varepsilon) = \{L' \in \operatorname{Comp}(X) : \Delta(L', L) \leq \varepsilon\}\,. \tag{3.1.3}$$

In particular, $L$ has a countable fundamental system of neighborhoods, given by the sets $V(L, \frac{1}{n})$. So $\operatorname{Comp}(X)$ is a first-countable topological space. It follows in particular that the topology of $\operatorname{Comp}(X)$ is completely determined by sequential convergence. (For example, if $L \in \operatorname{Comp}(X)$, and $S \subseteq \operatorname{Comp}(X)$, then $L \in \operatorname{Clos}(S)$ iff there exists a sequence of members of $S$ that converges to $L$ in $\operatorname{Comp}(X)$.)

Convergence of sequences in $\operatorname{Comp}(L)$ can be easily characterized:

**Fact 3.1.1** *Let $X$ be a metric space, let $\{L_j\}$ be a sequence in* $\operatorname{Comp}(X)$*, and let $L \in \operatorname{Comp}(X)$. Let*

$$\mathcal{L} = L \cup \Big(\bigcup_{j=1}^{\infty} L_j\Big)\,, \quad \tilde{\mathcal{L}} = \bigcap_{k=1}^{\infty} \operatorname{Clos}\Big(\bigcup_{j=k}^{\infty} L_j\Big)\,. \tag{3.1.4}$$

*Then the following six conditions are equivalent:*

*(3.1.1.i)* $\{L_j\}$ *converges to $L$ in* $\operatorname{Comp}(X)$*;*

*(3.1.1.ii)* $\lim_{j\to\infty} \Delta(L_j, L) = 0$*;*

*(3.1.1.iii)* $\mathcal{L}$ *is compact and* $\tilde{\mathcal{L}} \subseteq L$*;*

*(3.1.1.iv)* $\mathcal{L}$ *is compact and*

*(3.1.1.*) if $J$ is an infinite subset of $\mathbb{N}$, and $\{x_j\}_{j\in J}$ is a family of points of $X$ such that $x_j \in L_j$ for all $j \in J$ and the limit*

$$x = \lim_{j\to\infty, j\in J} x_j \tag{3.1.5}$$

*exists, then $x \in L$;*

*(3.1.1.v)* *there exists a compact subset $M$ of $X$ such that $\mathcal{L} \subseteq M$, and (3.1.1.*) holds;*

*(3.1.1.vi)* *if $J$ is an infinite subset of $\mathbb{N}$, and $\{x_j\}_{j\in J}$ is a family of points of $X$ such that $x_j \in L_j$ for all $j \in J$, then there exists an infinite subset $J'$ of $J$ such that the limit*

$$x = \lim_{j\to\infty, j\in J'} x_j \tag{3.1.6}$$

*exists and belongs to $L$.*

*Proof.* For $\varepsilon > 0$, write $U_\varepsilon = \{x \in X : \mathrm{dist}(x, L) \leq \varepsilon\}$. Then the $U_\varepsilon$ are neighborhoods of $L$ in $X$, and every neighborhood of $L$ in $X$ contains $U_\varepsilon$ for some $\varepsilon$. The sets $V(L, \varepsilon)$ defined by by (3.1.3) are given by

$$V(L, \varepsilon) = \{L' \in \mathrm{Comp}(X) : L' \subseteq U_\varepsilon\}\,.$$

So if (3.1.1.i) holds then for every $\varepsilon > 0$ there exists a $j(\varepsilon)$ such that $L_j \subseteq U_\varepsilon$ for $j \geq j(\varepsilon)$, and this implies that $\Delta(L_j, L) \leq \varepsilon$ for $j \geq j(\varepsilon)$. So (3.1.1.ii) holds.

If (3.1.1.ii) holds, then for every $\varepsilon > 0$ there is $j(\varepsilon)$ such that $L_j \in V(L, \varepsilon)$ for $j \geq j(\varepsilon)$. This implies that $L_j \subseteq U(\varepsilon)$ for $j \geq j(\varepsilon)$, so (3.1.1.i) holds, and also that $\bigcup_{j\geq j(\varepsilon)} L_j \subseteq U_\varepsilon$, and then $\mathrm{Clos}\Big(\bigcup_{j\geq j(\varepsilon)} L_j\Big) \subseteq U_\varepsilon$, since $U_\varepsilon$ is closed. Therefore $\tilde{\mathcal{L}} \subseteq \cap_{n=1}^{\infty} U_{2^{-n}} = L$. In addition, if we are given a sequence $\{x_k\}_{k=1}^{\infty}$ in $\mathcal{L}$, then either infinitely many $x_k$'s belong to one of the sets $L$, $L_1$, $L_2$, ... —in which case $\{x_k\}$ has a subsequence converging to a point in $\mathcal{L}$, because $L$ and the $L_j$ are compact— or we may assume, by passing to a subsequence, that none of the $x_k$'s belong to $L$, and that, if $j_k = \min\{j : x_k \in L_j\}$, then $j_k \to \infty$ as $k \to \infty$. Since $\Delta(L_j, L) \to 0$, we have $\mathrm{dist}(x_k, L) \leq \Delta(L_{j_k}, L) \to 0$. So there are points $x'_k \in L$ such that $\mathrm{dist}(x_k, x'_k) \to 0$. Since $L$ is compact, we may assume, by passing to a subsequence, that $\{x'_k\}$ converges to a point $x \in L$. Then $x_k \to x$. This shows that $\mathcal{L}$ is compact, so (3.1.1.iii) holds. So we have proved the implications (i)$\Leftrightarrow$(ii)$\Rightarrow$(iii).

If (3.1.1.iii) holds and $J$, $\{x_j\}_{j\in J}$, $x$ are as in (3.1.1.iv), then it is clear that $x \in \tilde{\mathcal{L}}$, so $x \in L$. Therefore (3.1.1.iv) holds.

The implication (3.1.1.iv)$\Rightarrow$(3.1.1.v) is trivial.

If (3.1.1.v) holds, and $J$, $\{x_j\}_{j\in J}$ are as in (3.1.1.vi), then the compactness of $M$ implies that there is an infinite subset $J'$ of $J$ such that $x = \lim_{j\to\infty, j\in J'} x_j$ exists, and then (3.1.1.v) tells us that $x \in L$, so (3.1.1.vi) holds.

Finally, we show that if (3.1.1.vi) holds then (3.1.1.ii) must hold as well. Assume (3.1.1.vi) holds but $\Delta(L_j, L)$ does not go to zero. Then there exist an $\varepsilon > 0$ and an infinite subset $J$ of $\mathbb{N}$ such that $\Delta(L_j, L) > \varepsilon$ for $j \in J$. So we can pick $x_j \in L_j$, for $j \in J$, such that $\operatorname{dist}(x_j, L) \geq \varepsilon$. By (3.1.1.vi), there is an infinite subset $J'$ of $J$ such that $x = \lim_{j\to\infty, j\in J'} x_j$ exists and belongs to $L$. On the other hand, $\operatorname{dist}(x, L) = \lim \operatorname{dist}(x_j, L) \geq \varepsilon$, so $x \notin L$. This contradiction shows that (3.1.1.i) holds. $\square$

## 3.2 Graph Convergence

Let $K, Y$ be metric spaces, with $K$ compact. We will write $SVM_{comp}(K, Y)$ to denote the set of all nonempty set-valued maps $F \in SVM(K, Y)$ whose graph is compact. (Recall that if $F \in SVM(K, Y)$ then $F$ has a compact graph if and only if it has compact values and is upper semicontinuous.) Clearly, $SVM_{comp}(K, Y)$ can be canonically identified with the set $\operatorname{Comp}(K \times Y)$ of all compact subsets of $K \times Y$. This enables us to define a topology on $SVM_{comp}(K, Y)$, called the *graph topology*, or *topology of graph convergence*, by endowing $SVM_{comp}(K, Y)$ with the topology resulting from the identification with $\operatorname{Comp}(K \times Y)$, which is a topological subspace of $(2^{K\times Y}, \mathcal{T}_{K\times Y})$, as explained before. Convergence in this topology will be called *graph convergence*, or *convergence in the graph sense*. The graph topology is clearly not Hausdorff, since, for example, if $G(F)$ is a proper subset of $G(F')$, then $F$ belongs to every neighborhood of $F'$.

Since $SVM_{comp}(K, Y)$ is none other than $\operatorname{Comp}(K \times Y)$, all the properties described above of the topology of the space of compact subsets of a metric space can be restated as properties of the topology of $SVM_{comp}(K, Y)$. Specifically, there is a nonsymmetric "distance function" $\Delta_{gr}$ —called the "graph distance"— between members of $SVM_{comp}(K, Y)$, given by

$$\begin{aligned} \Delta_{gr}(F, F') &= \Delta\Big(G(F), G(F')\Big) \\ &= \sup\Big\{\operatorname{dist}\Big((x, y), G(F')\Big) : (x, y) \in G(F)\Big\}, \end{aligned}$$

that satisfies the "triangle inequality"

$$\Delta_{gr}(F_1, F_3) \leq \Delta_{gr}(F_1, F_2) + \Delta_{gr}(F_2, F_3).$$

Every $F$ has a countable fundamental system of neighborhoods, given by the sets $V(F, \frac{1}{n})$, where $V(F, \varepsilon) = \{F' \in SVM_{comp}(K, Y) : \Delta_{gr}(F', F) \leq \varepsilon\}$. So $SVM_{comp}(K, K')$ is a first-countable topological space. In particular, if

$F \in SVM_{comp}(K, Y)$ and $S \subseteq SVM_{comp}(K, Y)$, then $F \in \text{Clos}(S)$ iff there exists a sequence of members of $S$ that converges to $F$ in the graph sense.

We call a subset $S$ of $SVM(K, Y)$ *c-bounded* if there exists a compact subset $K'$ of $Y$ such that $F(K) \subseteq K'$ for all $F \in S$.

We will write $F_j \xrightarrow{\text{gr}} F$ to indicate that $\{F_j\}$ converges to $F$ in the graph sense. Graph convergence of sequences in $SVM_{comp}(K, Y)$ can be easily characterized, as is done in the following statement, which is just a rewording of Fact 3.1.1 for the special case of a product space $K \times Y$:

**Fact 3.2.1** *Let $\{F_j\}$ be a sequence in $SVM_{comp}(K, Y)$, and let $F$ belong to $SVM_{comp}(K, Y)$. Let*

$$\mathcal{F} = G(F) \cup \Big( \bigcup_{j=1}^{\infty} G(F_j) \Big) , \tag{3.2.1}$$

$$\tilde{\mathcal{F}} = \bigcap_{k=1}^{\infty} \text{Clos}\Big( \bigcup_{j=k}^{\infty} G(F_j) \Big) . \tag{3.2.2}$$

*Then the following seven conditions are equivalent:*

*(3.2.1.i)* $F_j \xrightarrow{\text{gr}} F$;

*(3.2.1.ii)* $\lim_{j\to\infty} \Delta_{gr}(F_j, F) = 0$;

*(3.2.1.iii)* $\mathcal{F}$ *is compact and* $\tilde{\mathcal{F}} \subseteq G(F)$;

*(3.2.1.iv)* $\mathcal{F}$ *is compact, and*

*(3.2.1.*) if $J$ is an infinite subset of $\mathbb{N}$, and $\{(x_j, y_j)\}_{j\in J}$ is a family of points of $K \times Y$ such that $y_j \in F_j(x_j)$ for all $j \in J$ and the limit*

$$(x, y) = \lim_{j\to\infty, j\in J} (x_j, y_j) \tag{3.2.3}$$

*exists, then $y \in F(x)$;*

*(3.2.1.v) there exists a compact subset $M$ of $K \times Y$ such that $\mathcal{F} \subseteq M$, and (3.2.1.*) holds;*

*(3.2.1.vi) if $J$ is an infinite subset of $\mathbb{N}$, and $\{(x_j, y_j)\}_{j\in J}$ is a family of points of $Y$ such that $y_j \in F_j(x_j)$ for all $j \in J$, then there exists an infinite subset $J'$ of $J$ such that the limit*

$$(x, y) = \lim_{j\to\infty, j\in J'} (x_j, y_j) \tag{3.2.4}$$

*exists and satisfies $y \in F(x)$;*

*(3.2.1.vii) the set $\{F_j : j \in \mathbb{N}\}$ is c-bounded and (3.2.1.*) holds.*

*Proof.* The equivalence of the first six conditions follows from Fact (3.1.1). If (3.2.1.v) holds, then we can take $K' = \pi(M)$, where $\pi$ is the projection $K \times Y \ni (k, y) \to y \in Y$, and it is clear that $F_j(K) \subseteq K'$ for all $j$, so (3.2.1.vii) holds. Finally, if (3.2.1.vii) holds, and $K' \subseteq Y$ is compact and such that $F_j(K) \subseteq K'$ for all $j$, then we can take $M = K \times (K' \cup F(K))$, and (3.2.1.v) holds. □

## 3.3 Regular Maps

An important subset of $SVM_{comp}(K, Y)$ is the set $SVM_{comp,sv,ed}(K, Y)$ of all set-valued maps from $K$ to $Y$ with compact graph that are in fact single-valued and everywhere defined. It follows from the following lemma that $SVM_{comp,sv,ed}(K, Y)$ is exactly $C^0(K, Y)$, the set of all continuous maps from $K$ to $Y$.

**Lemma 3.3.1** *Let $F \in SVM_{comp}(K, Y)$ be such that $F$ is single-valued at a point $x \in K$. Then $F$ is continuous at $x$, in the sense that*

$$\lim_{x' \to x} \sup\{\mathrm{dist}_Y(y, F(x)) : y \in F(x')\} \to 0\,. \tag{3.3.1}$$

*In particular, $F$ is single-valued at every point of $K$ iff $F \in C^0(K, Y)$.*

*Proof.* Let $\{x_j\}$ be a sequence in $K$ converging to $x$, and let $y_j \in F(x_j)$, $y = F(x)$. We want to show that $y = \lim_j y_j$. If this is not true then we may assume, after passing to a subsequence, that $\lim_j y_j = y'$ exists and $y' \neq y$. (This is possible because $G(F)$ is compact.) But then $(x, y') \in G(F)$, since $(x_j, y_j) \in G(F)$ and $G(F)$ is closed. Then $y' = y$, and we have reached a contradiction. □

**Definition 3.3.2** A *regular* set-valued map from $K$ to $Y$ is a set-valued map $F \in SVM_{comp}(K, Y)$ that belongs to the closure in $SVM_{comp}(K, Y)$ of the space $C^0(K, Y)$ of continuous single-valued maps from $K$ to $Y$.

We use $\mathrm{REG}(K, Y)$ to denote the set of all regular set-valued maps from $K$ to $Y$. □

**Definition 3.3.3** If $F \in SVM_{comp}(K, Y)$, a *single-valued approximating sequence* (abbr. SVAS) for $F$ is a sequence $\{F_j\}_{j=1}^{\infty}$ of continuous maps from $K$ to $Y$ such that $F_j \stackrel{\mathrm{gr}}{\to} F$. □

So, if $F \in SVM_{comp}(K,Y)$, then $F$ is regular iff there exists a SVAS $\{F_j\}_{j=1}^{\infty}$ for $F$. Moreover, Fact 3.2.1 implies the following characterizations of SVAS's:

**Fact 3.3.4** *Let $\{F_j\}$ be a sequence of continuous maps from $K$ to $Y$, and let $F$ belong to $SVM_{comp}(K,Y)$. Let $\mathcal{F}$, $\tilde{\mathcal{F}}$ be given by (3.2.1), (3.2.2). Then $F$ is a SVAS for $F$ if and only if any one of the seven equivalent conditions of Fact 3.2.1 holds.* □

The following observations are then obvious. (To prove (OB.ii.b), pick a SVAS $\{F_j\}$, with values in the compact subset $K'$ of $Y$. If $x \in K$, then the sequence $\{F_j(x)\}$ must have a subsequence $\{F_{j(k)}(x)\}$ that converges to a limit $y$. The definition of SVAS then implies that $y \in F(x)$, so $F(x) \neq \emptyset$.)

- (OB.i) $C^0(K,Y) \subseteq \mathrm{REG}(K,Y)$ *(i.e. every continuous map from $K$ to $Y$ is regular)*;
- (OB.ii) *if $F \in \mathrm{REG}(K,Y)$, then*
  - (a) *$G(F)$ is compact,*
  - (b) *$\mathrm{Dom}(F) = K$ (i.e. $F(x) \neq \emptyset$ for all $x \in K$),*
  - (c) *$F(K)$ is compact,*
  - (d) *$F \in C^0(K,Y)$ iff $F$ is single-valued;*
- (OB.iii) *if $F \in SVM(K,Y)$, then $F \in \mathrm{REG}(K,Y)$ iff $G(F)$ is compact (i.e., equivalently, iff $F$ is upper semicontinuous with compact values) and $F$ has a SVAS;*
- (OB.iv) *if $F \in SVM(K,Y)$, then $F \in C^0(K,Y)$ iff $G(F)$ is compact and $F$ is everywhere defined and single-valued.* □

The space $\mathrm{REG}(K,Y)$ has an obvious "compactness" property, namely, *every c-bounded sequence in* $\mathrm{REG}(K,Y)$ *converges.* (In particular, if $Y$ is compact, then every sequence in $\mathrm{REG}(K,Y)$ converges.) This is true for trivial reasons: if $\{F_j\}$ is a sequence and $K'$ is such that $F_j(K) \subseteq K'$ for all $j$, we can let $F \in S_{comp}(K,Y)$ be the map whose graph is $K \times K'$. Then $F_j \overset{\mathrm{gr}}{\to} F$.

Not surprisingly, this trivial property is not of much use, because the limit $F$ is too large. On the other hand, there do exist much smaller limits, and some of these are truly useful.

**Definition 3.3.5** Given a sequence $\{F_j\}$ in $\mathrm{REG}(K,Y)$, the *minimal limit* $\min\lim_{j\to\infty} F_j$ is the set-valued function $F \in SVM(K,Y)$ whose graph is the set $\tilde{\mathcal{F}}$ defined by (3.2.2). □

The following elementary observation justifies the name "minimal limit":

**Fact 3.3.6** *If $\{F_j\}$ is a c-bounded sequence of members of* $\mathrm{REG}(K,Y)$*, and* $F = \min\lim_{j\to\infty} F_j$*, then*

*(3.3.6.i)* $F \in \mathrm{REG}(K,Y)$ *and* $F_j \overset{\mathrm{gr}}{\to} F$*;*

*(3.3.6.ii) if* $F' \in SVM_{comp}(K,Y)$*, then* $F_j \overset{\mathrm{gr}}{\to} F'$ *iff* $G(F) \subseteq G(F')$*.* $\square$

**Remark 3.3.7** An important example of regular maps is provided by the following

**Fact 3.3.8** *Assume that $Y$ is a finite-dimensional real linear space, and $F \in SVM(K,Y)$ has a compact graph and nonempty convex values. Then $F \in \mathrm{REG}(K,Y)$.*

*Proof.* Let $L$ be the convex hull of $F(K)$, so $L$ is a compact subset of $Y$. Let $n = \dim(Y)$.

For each $\delta > 0$, pick a covering of $K$ by open subsets $U_1(\delta),\ldots,U_{m(\delta)}(\delta)$ of diameter $< \delta$, and pick a point $x_{i,\delta} \in U_i(\delta)$, and then a $y_{i,\delta} \in F(x_{i,\delta})$ for each $i \in \{1,\ldots,m(\delta)\}$. Let $\varphi_1,\ldots,\varphi_{m(\delta)}$ be continuous functions from $K$ to $[0,1]$ such that $\sum_{i=1}^{m(\delta)} \varphi_{i,\delta} \equiv 1$ and $\{x : \varphi_{i,\delta}(x) > 0\} \subseteq U_{i,\delta}$. (For example, let $\psi_{i,\delta}(x) = \mathrm{dist}(x, K\backslash U_{i,\delta})$. Then use

$$\varphi_{i,\delta}(x) \quad = \quad \frac{\psi_{i,\delta}(x)}{\sum_{j=1}^{m(\delta)} \psi_{j,\delta}(x)}$$

to define the $\varphi_{i,\delta}$.) Then define

$$H_\delta(x) \quad = \quad \sum_{i=1}^{m(\delta)} \varphi_{i,\delta}(x) y_{i,\delta}\,.$$

Clearly, the $H_\delta$ are continuous maps from $K$ to $Y$ with values in $L$. Define $F_j = H_{1/j}$. We claim that $F_j \overset{\mathrm{gr}}{\to} F$. To see this, we pick an infinite subset $J$ of $\mathbb{N}$ and points $x_j \in K$, for $j \in J$, such that the limit $(x,y) = \lim_{j\to\infty, j\in J}(x_j, F_j(x_j))$ exists, and prove that $y \in F(x)$. Let $y_j = F_j(x_j)$. Then, if we let $\delta = 1/j$, we have $y_j = \sum \varphi_{i,\delta}(x_j) y_{i,\delta}$, and the sum can obviously be taken over the set $I_j$ of those $i \in 1,\ldots,m(\delta)$ such that $\varphi_{j,\delta}(x_j) > 0$. If $i \in I_j$, then $x_j \in U_{i,\delta}$, so $\mathrm{dist}(x_j, x_{i,\delta}) < 2\delta$. Since $y_j \in F(x_{i,\delta})$, it follows that $y_j$ belongs to the convex hull of $Q_j = \bigcup_{x\in V_j} F(x)$, where $V_j = \{x' \in K : \mathrm{dist}(x', x_j) < 2\delta\}$. Then Carathéodory's Theorem implies that $y_j$ is a convex combination of $n+1$ points of $Q_j$. So there exist

$z_0^j, \ldots, z_n^j \in V_j$, $w_i^j \in F(z_i^j)$ for $i = 0, \ldots, n$, $c_0^j, \ldots, c_n^j \in \mathbb{R}$ such that $c_i^j \geq 0$ and $\sum_{i=0}^n c_i^j = 1$, for which $y_j = \sum_{i=0}^n c_i^j w_i^j$. Then all the $w_i^j$ belong to the compact set $L$. So there is an infinite subset $J'$ of $J$ such that the limits $c_i = \lim_{j\to\infty, j\in J'} c_i^j$ and $w_i = \lim_{j\to\infty, j\in J'} w_i^j$ exist for each $i$, and belong to $\mathbb{R}$, $L$, respectively. Clearly, $x = \lim_{j\to\infty, j\in J'} z_i^j$, since $z_i^j \in V_j$. Since $w_i^j \in F(z_i^j)$, the compactness of the graph of $F$ implies that $w_i \in F(x)$ for $i = 0, \ldots, n$. Clearly, $y = \sum_{i=0}^n c_i w_i$, so $y$ is a convex combination of the $w_i$. Since $F(x)$ is convex, we conclude that $y \in F(x)$, as desired. □

## 3.4 Connectedness Properties

Recall that a topological space $T$ is *connected* if the only nonempty subset of $T$ which is both open and closed is $T$ itself. A subset $S$ of $T$ is *connected* if it is connected as a topological subspace of $T$. The closure of a connected set is connected. If $p \in T$, then the union of any family of connected sets containing $p$ is also connected. It follows that the binary relation $\sim$ on $T$ defined by "$p \sim q$ iff there is a connected subset of $T$ containing both $p$ and $q$" is an equivalence relation. The equivalence classes are the *connected components* of $T$. The connected component $CC(p, T)$ through a point $p \in T$ is the union of all the connected subsets of $T$ containing $p$. Clearly, $CC(p, T)$ is connected and closed. A space $T$ is *locally connected* if every point of $T$ has a connected neighborhood. Clearly, $T$ is locally connected iff all the connected components of $T$ are open.

**Definition 3.4.1** We call an $F \in \mathrm{REG}(K, Y)$ *connected* if its graph is connected. We call $F$ *r-connected* if all the connected components of $G(F)$ are graphs of regular maps from $K$ to $Y$. □

Since a regular set-valued map is everywhere defined, and a continuous image of a connected space is connected, a necessary condition for the existence of r-connected members of $\mathrm{REG}(K, Y)$ is that $K$ itself be connected. In that case, a continuous map $F : K \to Y$ is obviously connected.

The following two observations say that a minimal limit of connected regular maps is connected, and a minimal limit of r-connected regular maps is r-connected.

**Fact 3.4.2** *Assume that $K$ is connected, and let $\{F_j\}$ be a c-bounded sequence of connected members of* $\mathrm{REG}(K, Y)$. *Assume that there exists a convergent sequence $\{(x_j, y_j)\}_{j=1}^{\infty}$ such that $y_j \in F_j(x_j)$ for each $j$. Let $F = \min\lim_{j\to\infty} F_j$. Then $F$ is connected.*

*Proof.* We know that $F \in \mathrm{REG}(K,Y)$ and $F_j \overset{\mathrm{gr}}{\to} F$. Let $S = G(F)$. We have to show that $S$ is connected. Suppose it is not, and let $S_1$, $S_2$ be nonempty relatively open disjoint subsets of $S$ such that $S = S_1 \cup S_2$. Let $(\bar{x},\bar{y}) = \lim_{j\to\infty}(x_j,y_j)$. Then $(\bar{x},\bar{y}) \in S$. So, after relabeling $S_1$ and $S_2$, if necessary, we may assume that $(\bar{x},\bar{y}) \in S_1$. Then $S_1$, $S_2$ are compact, so there exist disjoint open subsets $U_1$, $U_2$ of $X \times Y$ such that $S_i \subseteq U_i$ for $i = 1,2$. Then $G(F_j) \subseteq U_1 \cup U_2$ for $j$ sufficiently large. Since $G(F_j)$ is connected, it must be entirely contained in one of the sets $U_1$ or $U_2$. But $(x_j,y_j) \in U_1$ for large $j$, since $(\bar{x},\bar{y}) \in S_1$. So for large $j$ the graph of $F_j$ is contained in $U_1$. Therefore $G(F_j) \subseteq Y \backslash U_2$ for large $j$. It then follows that $S \subseteq U_1$, so $S_2 = \emptyset$, which is a contradiction. □

**Fact 3.4.3** *Assume that $K$ is connected, and let $\{F_j\}$ be a c-bounded sequence of r-connected members of* $\mathrm{REG}(K,Y)$*. Let $F = \min\lim_{j\to\infty} F_j$. Then $F$ is r-connected.*

*Proof.* Let $S = G(F)$, $S_j = G(F_j)$. We have to prove that every connected component of $S$ is the graph of a regular set-valued map from $K$ to $Y$. Let $S'$ be a connected component of $S$. Pick a point $(\bar{x},\bar{y}) \in S'$. Since $(\bar{x},\bar{y}) \in S = \bigcap_{k=1}^{\infty} \mathrm{Clos}(\bigcup_{j=k}^{\infty} S_j)$, we can pick, for every $k$, a $j(k)$ such that $j(k) \geq k$, and a point $(x_k,y_k) \in S_{j(k)}$ such that $\mathrm{dist}((x_k,y_k),(\bar{x},\bar{y})) < 2^{-k}$. Let $\tilde{S}_k$ be the connected component of $S_k$ that contains $(x_k,y_k)$. Then $\tilde{S}_k = G(H_k)$ for an $H_k \in \mathrm{REG}(K,Y)$. Let $H = \min\lim_{k\to\infty} H_k$, and let $S'' = G(H)$. Then $(\bar{x},\bar{y}) \in S''$, and $(\bar{x},\bar{y}) = \lim_{k\to\infty}(x_k,y_k)$. So by Fact 3.4.2, $H$ is connected. So $S''$ is a connected set. Since $(\bar{x},\bar{y}) \in S'' \subseteq S$, we conclude that $S'' \subseteq S'$. Since $H \in \mathrm{REG}(K,Y)$ by Fact 3.3.6, and $S'$ is compact, it follows that $S'$ is the graph of a regular map from $K$ to $Y$. □

**Corollary 3.4.4** *Let $K$ be connected, and let $F \in$* $\mathrm{REG}(K,Y)$*. Then there exists a connected $F' \in$* $\mathrm{REG}(K,Y)$ *such that $G(F') \subseteq G(F)$.*

*Proof.* Let $\{F_j\}$ be a SVAS for $F$. Then each $F_j$ is connected, hence r-connected. So $\hat{F} = \min\lim_{j\to\infty} F_j$ is r-connected. If $S$ is any connected component of $G(\hat{F})$, and $S = G(F')$, then $F'$ satisfies all our conditions. □

## 3.5 Fixed Point Theorems

Many important properties of continuous maps are also valid for regular maps. For example, the following statement —in which $\mathbb{B}^n$ denotes the closed unit ball in $\mathbb{R}^n$— is the well known *Brouwer Fixed Point Theorem:*

**Fact 3.5.1** *Let $F \in \mathrm{REG}(\mathbb{B}^n, \mathbb{R}^n)$ be such that $F(\mathbb{B}^n) \subseteq \mathbb{B}^n$. Then $F$ has a fixed point, that is, there exists an $x \in \mathbb{B}^n$ such that $x \in F(x)$.* □

This is an easy corollary of the usual Brouwer theorem. (Let $\{F_j\}$ be a SVAS for $F$, and let $\rho : \mathbb{R}^n \to \mathbb{B}^n$ be a retraction. Let $G_j = \rho \circ F_j$. Then $\{G_j\}$ is also a SVAS for $F$, and $G_j$ is a continuous map from $\mathbb{B}^n$ to $\mathbb{B}^n$. By the Brouwer Fixed Point Theorem, there is an $x_j$ such that $G_j(x_j) = x_j$. By passing to a subsequence, if necessary, we may assume that $x_j \to x \in \mathbb{B}^n$. Then $G_j(x_j) \to x$ as well, so $x \in F(x)$.)

In addition, Fact 3.5.1 also follows from Theorem 3.5.4 below, whose proof will be self-contained (cf. Remark 3.5.11).

**Remark 3.5.2** Fact 3.5.1 contains as a particular case the well known *Kakutani Fixed Point Theorem*: if $K$ is a nonempty compact convex subset of $\mathbb{R}^n$ and $F : K \to 2^K$ is upper semicontinuous and has nonempty compact convex values, then there is an $x \in K$ such that $x \in F(x)$. To see this, observe that Fact 3.3.8 implies that $F \in \mathrm{REG}(K, \mathbb{R}^n)$. Then $F \in \mathrm{REG}(K, K)$, because $K$ is a retract of $\mathbb{R}^n$. Since $K$ is homeomorphic to a ball $\mathbb{B}^k$ for some $k$, and regularity of a map from $K$ to $K$ is invariant under homeomorphisms, we can assume that $K = \mathbb{B}^k$, and then the conclusion follows from Fact 3.5.1. □

**Fact 3.5.3** *Let $F \in \mathrm{REG}(\mathbb{B}^n, \mathbb{R}^n)$ and $\alpha \in \mathbb{R}$ be such that $0 \le \alpha \le 1$ and*

$$\|y - x\| \le \alpha \quad \text{for all} \quad x \in \partial\mathbb{B}^n\,,\ y \in F(x)\,. \tag{3.5.1}$$

*Then $F^{-1}(y) \neq \emptyset$ for all $y \in \mathbb{B}^n$ such that $\|y\| \le 1 - \alpha$.*

If we make the stronger assumption that

$$\|y - x\| \le \alpha \quad \text{for all} \quad x \in \mathbb{B}^n\,,\ y \in F(x)\,, \tag{3.5.2}$$

then this follows easily from Fact 3.5.1. (Pick $y$ such that $\|y\| \le 1-\alpha$. Define $G$ by letting $G(x) = x + y - F(x)$, i.e. $G(x) = \{x + y - z : z \in F(x)\}$. Then $G \in \mathrm{REG}(\mathbb{B}^n, \mathbb{R}^n)$ and $G$ takes values in $\mathbb{B}^n$. So there is an $x$ such that $x \in G(x)$, and then there is a $z \in F(x)$ such that $x = x + y - z$. Therefore $z = y$, so $y \in F(x)$.)

Fact 3.5.3, under the weaker assumption (3.5.1), will follow from Theorem 3.5.4 below, as explained in Remark 3.5.11.

**Theorem 3.5.4** *Let $-\infty < a < b < \infty$, and let $H \in \mathrm{REG}([a, b] \times \mathbb{B}^n, \mathbb{R}^n)$ and $\alpha \in \mathbb{R}$ be such that $0 \le \alpha \le 1$ and*

$$\|y - x\| \le \alpha \quad \text{for all} \quad x \in \partial\mathbb{B}^n \quad \text{and all} \quad y \in H(a, x)\,. \tag{3.5.3}$$

*Let $y_0 \in \mathbb{B}^n$ be such that $\|y_0\| \leq 1 - \alpha$, and assume that*

$$y_0 \notin H(t,x) \ \text{ for all } \ (t,x) \in [a,b] \times \partial\mathbb{B}^n \,. \tag{3.5.4}$$

*Then there exists a $\gamma \in \mathrm{REG}([0,1],[a,b]\times\mathbb{B}^n)$ such that*

*(3.5.4.i)* $\gamma(0) \cap \Big(H^{-1}(y_0) \cap (\{a\} \times \mathbb{B}^n)\Big) \neq \emptyset$,

*(3.5.4.ii)* $\gamma(1) \cap \Big(H^{-1}(y_0) \cap (\{b\} \times \mathbb{B}^n)\Big) \neq \emptyset$,

*(3.5.4.iii)* $\gamma([a,b]) \subseteq H^{-1}(y_0)$,

*and $\gamma$ is connected.*

*Proof.* For any $h \in SVM([a,b]\times\mathbb{B}^n, \mathbb{R}^n)$, define $h_t \in SVM(\mathbb{B}^n,\mathbb{R}^n)$, for $t \in [a,b]$, by letting $h_t(x) = h(t,x)$. For $h \in \mathrm{REG}([a,b]\times\mathbb{B}^n,\mathbb{R}^n)$ and $y \in \mathbb{R}^n$, let

$$\begin{aligned} \delta(h,y) &= \inf\Big\{\|y'-y\| : y' \in h(t,x)\,,\ (t,x) \in [a,b]\times \in \partial\mathbb{B}^n\Big\}\,, \\ \alpha^*(h) &= \sup\{\|y'-x\| : y' \in h(a,x)\,,\ x \in \partial\mathbb{B}^n\}\,. \end{aligned}$$

Then our hypotheses say that $\alpha^*(H) \leq \alpha$ and $\delta(H,y_0) > 0$. (The inequality $\delta(H,y_0) > 0$ follows because $\delta(H,y_0)$ is the infimum of the continuous function $(t,x,y) \to \|y - y_0\|$ on the compact set $G(H) \cap ([a,b]\times\partial\mathbb{B}^n\times\mathbb{R}^n)$, and all the values of this function are $> 0$.)

It is easy to see that

(*) *If $h_j \overset{\mathrm{gr}}{\to} h$ and $y_j \to y$, then*

$$\liminf \delta(h_j,y_j) \geq \delta(h,y) \quad \text{and} \quad \limsup \alpha^*(h_j) \leq \alpha^*(h)\,.$$

Indeed, if $\liminf \delta(h_j,y_j) = \rho < \delta(h,y)$, then we can find an infinite subset $J$ of $\mathbb{N}$ such that $\lim_{j\to\infty, j\in J} \delta(h_j,y_j) = \rho$. By shrinking $J$ if necessary, we may assume that there are $(t_j,x_j) \in [a,b]\times\partial\mathbb{B}^n$, $z_j \in h_j(t_j,x_j)$, such that $\|z_j - y_j\| \leq \rho'$, where $\rho'$ is such that $\rho < \rho' < \delta(h,y)$. By shrinking $J$ further, we may assume that $t_j \to t$, $x_j \to x$, $z_j \to z$. But then $z$ belongs to $h(t,x)$, and $\|z-y\| \leq \rho' < \delta(h,y)$, which is a contradiction. The proof that $\limsup \alpha^*(h_j) \leq \alpha^*(h)$ is similar.

Next, we show that

(#) if $h_j \in \mathrm{REG}([a,b]\times\mathbb{B}^n,\mathbb{R}^n)$ are such that $h_j \overset{\mathrm{gr}}{\to} h$ and the conclusion of our theorem holds for the $h_j$, then it holds for $h$.

To prove this, we let $\gamma^j \in \mathrm{REG}([0,1],[a,b]\times \mathbb{B}^n)$ be connected and such that $\gamma_j([a,b]) \subseteq h_j^{-1}(y_0)$, and

$$\begin{aligned} \gamma_j(0)\cap h_j^{-1}(y_0)\cap \Big(\{a\}\times \mathbb{B}^n\Big) &\neq \emptyset\,, \\ \gamma_j(1)\cap h_j^{-1}(y_0)\cap \Big(\{b\}\times \mathbb{B}^n\Big) &\neq \emptyset\,. \end{aligned}$$

Pick $\bar{x}_j \in \mathbb{B}^n$ such that $y_0 \in h_j(a,\bar{x}_j)$ and $(a,\bar{x}_j) \in \gamma_j(0)$. Let $\{\bar{x}_{j(k)}\}$ be a convergent subsequence of $\{\bar{x}_j\}$, and write $p_k = (0,a,\bar{x}_{j(k)})$. Then let $\gamma = \min\lim_{k\to\infty}\gamma_{j(k)}$. Then $\gamma \in \mathrm{REG}([0,1],[a,b]\times\mathbb{B}^n)$ and, since $p_k \in G(\gamma_{j(k)})$ and $\gamma_{j(k)}$ is connected, Fact 3.4.2 implies that $\gamma$ is connected as well.

If $(t,x) \in \gamma(s)$, then we can find for each $\ell$ a $k(\ell) \geq \ell$ and $s_\ell$, $t_\ell$, $x_\ell$ such that $(t_\ell,x_\ell) \in \gamma_{j(k(\ell))}(s_\ell)$ and $s_\ell \to s$, $t_\ell \to t$, $x_\ell \to x$. It then follows that $y_0 \in h_{j(k(\ell))}(t_\ell,x_\ell)$, so $y_0 \in h(t,x)$. Therefore $\gamma([a,b]) \subseteq h^{-1}(y_0)$. If $\bar{x} = \lim \bar{x}_{j(k)}$, then $(a,\bar{x}) \in \gamma(0)$. Since $y_0 \in h_{j(k)}(a,\bar{x}_{j(k)})$, we see that $y_0 \in h(a,\bar{x})$, so $\gamma(0)\cap h^{-1}(y_0)\cap\Big(\{a\}\times\mathbb{B}^n\Big) \neq \emptyset$. Finally, we may pick $\tilde{x}_k$ such that $(b,\tilde{x}_k) \in \gamma_{j(k)}(1)$ and $y_0 \in h_{j(k)}(b,\tilde{x}_k)$. Then we can find a subsequence $\{\tilde{x}_{k(\ell)}\}$ that converges to a limit $\tilde{x}$. Since $(b,\tilde{x}_{k(\ell)}) \in \gamma_{j(k(\ell))}(1)$, and $\gamma_{j(k)}(1)\xrightarrow{\mathrm{gr}}\gamma$, we have $(b,\tilde{x}) \in \gamma(1)$. Since $y_0 \in h_{j(k(\ell))}(b,\tilde{x}_{k(\ell)})$ and $h_{j(k(\ell))}\xrightarrow{\mathrm{gr}}h$, we have $(b,\tilde{x}) \in h^{-1}(y_0)$. So $\gamma_j(1)\cap h_j^{-1}(y_0)\cap\Big(\{b\}\times\mathbb{B}^n\Big) \neq \emptyset$. This completes the proof of (#). □

In view of (#), we can make several extra assumptions on $H$ without loss of generality. First, we can assume that

(A1) $\alpha > 0$.

To see that this is possible, assume that our hypothesis holds with $\alpha = 0$. Then $H_a\lceil\partial\mathbb{B}^n$ is the identity map of $\partial\mathbb{B}^n$, so $H_a$ takes on all the values in $\partial\mathbb{B}^n$. Since $H_a$ does not take the value $y_0$, we conclude that $\|y_0\| < 1$, so our hypotheses are also satisfied if we take instead $\alpha = 1 - \|y_0\|$, in which case $\alpha > 0$.

Next, we show that we can assume

(A2) $\alpha^*(H) < \alpha$.

To see this, define

$$H_j(t,x) = (1-2^{-j})H(t,x) + 2^{-j}x = \{(1-2^{-j})y + 2^{-j}x : y \in H(t,x)\}\,,$$

for $a \leq t \leq b$, $x \in \mathbb{B}^n$. Then $H_j \in \mathrm{REG}([a,b]\times\mathbb{B}^n,\mathbb{R}^n)$, and $H_j\xrightarrow{\mathrm{gr}}H$. Also,

$$\begin{aligned} \alpha^*(H_j) &= \sup\{\|y-x\| : y \in H_j(a,x)\,,\, x \in \partial\mathbb{B}^n\} \\ &\leq (1-2^{-j})\alpha^*(H) \leq (1-2^{-j})\alpha < \alpha\,, \end{aligned} \tag{3.5.5}$$

where the last inequality follows because $\alpha > 0$.

Since $H_j \stackrel{\mathrm{gr}}{\to} H$, it is clear from (*) that $\delta(H_j, y_0) > 0$ for large $j$. So $H_j$ satisfies all our hypotheses for large $j$, as well as (A2). If our conclusion holds for the $H_j$, then it holds for $H$ by (#). So it suffices to prove our conclusion for the $H_j$, i.e. to assume that (A2) holds.

Next, we add the assumption that

(A3) $H$ is single-valued.

(To see that this is possible, let $\{H_j\}_{j=1}^{\infty}$ be a SVAS for $G$. Then $H_j \stackrel{\mathrm{gr}}{\to} H$, and $H_j$ satisfies all our assumptions (including (A2)), for large $j$, in view of (*).)

Next, we can make the even stronger assumption that

(A4) $H$ is a polynomial map.

(This follows from the fact that continuous maps from $[a, b] \times \mathbb{B}^n$ to $\mathbb{R}^n$ can be uniformly approximated by polynomial ones.)

Finally, we can also assume that

(A5) $y_0$ is a regular value of the maps $H$, $H_a$ and $H_b$, regarded as maps on $\mathbb{R} \times \mathbb{R}^n$, $\mathbb{R}^n$ and $\mathbb{R}^n$, respectively. (Recall that at this stage we are assuming that $H$ is a polynomial map.)

We recall that a *regular value* of a smooth map $\mu$ from a smooth manifold $M$ to a smooth manifold $N$ is a point $q \in N$ such that the differential $D\mu(p)$ is surjective —as a map from the tangent space $T_pM$ of $M$ at $p$ to the tangent space $T_qN$ of $N$ at $q$— for every $p \in M$ such that $\mu(p) = q$. Sard's Theorem —cf. Sard [10] or Sternberg [12]—says that the complement of the set of regular values is of measure zero. Applying this to the maps $H$, $H_a$ and $H_b$, we find that there exist $v_j$ such that $v_j \to y_0$ as $j \to \infty$ that are regular values of $H$, $H_a$ and $H_b$. If we let $H^j(t, x) = H(t, x) + y_0 - v_j$, we see that $H^j \stackrel{\mathrm{gr}}{\to} H$ and $y_0$ is a regular value of $H^j$, $H^j_a$ and $H^j_b$.

Since $y_0$ is a regular value of $H_a$, the set $Z_a = \{x \in \mathbb{B}^n : H_a(x) = y_0\}$ is finite. We need the stronger fact that $Z_a$ has an odd number of elements. This follows from

**Lemma 3.5.5** *Let $0 < \alpha \leq 1$ and let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a smooth map such that $\|F(x) - x\| < \alpha$ for all $x \in \partial\mathbb{B}^n$. Let $y_0 \in \mathbb{R}^n$ be such that $\|y_0\| \leq 1 - \alpha$ and $y_0$ is a regular value of $F$. Let $W = F^{-1}(y_0) \cap \mathbb{B}^n$. Then $W$ is a finite set with an odd number of elements.*

*Proof.* This lemma is a well known fact from topology, since the hypothesis implies that $F\lceil\partial\mathbb{B}^n$ is homotopic to the identity map $\mathrm{id}_{\partial\mathbf{B}_n}$ of $\partial\mathbb{B}^n$ via a homotopy that never takes the value $y_0$. For completeness, we include the proof.

Write $K=[0,1]\times\mathbb{B}^n$. Then $K$ is a compact subset of $\mathbb{R}\times\mathbb{R}^n$, whose interior $\mathrm{Int}(K)$ is the set $(0,1)\times(\mathbb{B}^n\backslash\partial\mathbb{B}^n)$. Clearly, $K=\mathrm{Clos}(\mathrm{Int}(K))$, and the boundary $\partial K=K\backslash\mathrm{Int}(K)$ is the union of the sets $\{0\}\times\mathbb{B}^n$, $\{1\}\times\mathbb{B}^n$ and $[0,1]\times\partial\mathbb{B}^n$.

The fact that $W$ is finite is trivial, since $F^{-1}(y_0)$ is discrete (because $F$ is a local diffeomorphism near every point of $F^{-1}(y_0)$) and $\mathbb{B}^n$ is compact. Moreover, $W\cap\partial\mathbb{B}^n=\emptyset$, because for $x\in\partial\mathbb{B}^n$ we have $\|F(x)-x\|<\alpha$ but $\|y_0-x\|\geq\alpha$, so $F(x)\neq y_0$.

Let us now show that

($) the cardinality of the set $W(y)=\{x\in\mathbb{B}^n: F(x)=y\}$ is independent of $y$ for $y$ in a neighborhood of $y_0$.

To see this, let $W=\{x_1,\ldots,x_\nu\}$, where the $x_i$ are all different. Let $U_1,\ldots,U_\nu$ be pairwise disjoint neighborhoods of $x_1,\ldots,x_\nu$ in $\mathbb{R}^n$ such that $U_i\subseteq\mathbb{B}^n$ (recall that $W\cap\partial\mathbb{B}^n=\emptyset$) chosen so that $F$ is a diffeomorphism on each $U_i$. Then for $y$ close enough to $y_0$ each $U_i$ contains a point $x_i(y)$ such that $F(x_i(y))=y$. So $\mathrm{card}(W(y))\geq\nu$ for $y$ near $y_0$. If there are points $y$ arbitrarily close to $y_0$ for which $\mathrm{card}(W(y))>\nu$, then there is a sequence $\{y_k\}$ converging to $y_0$ and, for each $k$, $\nu+1$ different points $x_k^0,\ldots,x_k^\nu$ such that $F(x_k^i)=y_k$. By passing to a subsequence, we may assume that $x'^i=\lim_{k\to\infty}x_k^i$ exists for each $i\in\{0,\ldots,\nu\}$. Clearly, each $x'^i$ is one of the $x_\ell$, and since there are exactly $\nu$ $x_\ell$'s, there must exist $i_1,i_2,\ell$ such that $x'_{i_1}=x'_{i_2}=x_\ell$ but $i_1\neq i_2$. Then $x_{i_1}^k\in U_\ell$ and $x_{i_2}^k\in U_\ell$ for large $k$. Since $x_{i_1}^k\neq x_{i_2}^k$ but $F(x_{i_1}^k)=y_k=F(x_{i_2}^k)$, we have reached a contradiction, since $F$ is one-to-one on $U_\ell$. This proves ($).

Now define $\mathcal{F}(s,x)=x+s(F(x)-x)$, for $s\in\mathbb{R}$, $x\in\mathbb{R}^n$. Then $\mathcal{F}$ is a smooth map, and $\mathcal{F}_1=F$, $\mathcal{F}_0=\mathrm{id}_{\mathbf{R}^n}$. To prove that $\nu$ is odd, it suffices to take $y$ close enough to $y_0$, to guarantee that $\mathrm{card}(W(y))=\nu$, and prove that $\mathrm{card}(W(y))$ is odd. By Sard's Theorem, we can take $y$ to be a regular value of $\mathcal{F}$ and $\mathcal{F}_1=F$. (Clearly, $y$ is a regular value of $\mathcal{F}_0$ as well.) We can require in addition that $\|y\|\leq 1-\alpha$. (Since $\|F(x)-x\|<\alpha$ for all $x\in\partial\mathbb{B}^n$, we have $\alpha'=\sup\{\|F(x)-x\|:x\in\partial\mathbb{B}^n\}<\alpha$, and then we can pick $\beta$ such that $\alpha'<\beta<\alpha$, choose $y$ so that $\|y\|\leq 1-\beta$, and rename $\beta$ to make it the new $\alpha$.)

Then $Q=\mathcal{F}^{-1}(y)$ is a closed subset of $\mathbb{R}\times\mathbb{R}^n$, and is a smooth embedded 1-dimensional submanifold. Moreover, the fact that $y$ is a regular value of

$\mathcal{F}_0$ and $\mathcal{F}_1$ implies that $Q$ is transversal to the submanifolds $S_0 = \{0\} \times \mathbb{R}^n$ and $S_1 = \{1\} \times \mathbb{R}^n$. (If $q = (0, x) \in Q$, then the differential $D\mathcal{F}(q)$ vanishes on the tangent space $T_qQ$ to $Q$ at $q$, because $\mathcal{F}$ is constant on $Q$. If $T_qQ$ was contained in $T_qS_0$, then $D\mathcal{F}_0(x)$ would vanish on a 1-dimensional subspace of $\mathbb{R}^n$, if we identify $\mathbb{R}^n$ with $T_qS_0$ via the diffeomorphism $x \to (0, x)$. Since $\mathcal{F}_0(x) = y$, this would contradict the fact that $y$ is a regular value of $\mathcal{F}_0$. The proof that $Q$ is transversal to $S_1$ is similar.) In addition, $Q$ does not intersect the set $[0,1] \times \partial\mathbb{B}^n$, because if $0 \le s \le 1$ and $x \in \partial\mathbb{B}^n$, then

$$\|\mathcal{F}(s, x) - x\| = s\|F(x) - x\| < \alpha\,,$$

whereas $1 = \|x\| \le \|x - y\| + \|y\| \le \|x - y\| + 1 - \alpha$, so $\|x - y\| \ge \alpha$, and then $\mathcal{F}(s, x) \ne y$. So $Q \cap \partial K = Q \cap (\{0,1\} \times \mathbb{B}^n) = \{(0, y)\} \cup (\{1\} \times W(y))$, which is a finite set consisting of $\nu + 1$ elements.

It follows that each connected component of $Q$ is diffeomorphic to $\mathbb{R}$ or to the unit circle $S^1$.

**Fact 3.5.6** *If $Q'$ is any connected component of $Q$, then $Q' \cap \partial K$ is a finite set, with an even number of elements.*

To see this, define $g : Q' \backslash (Q' \cap \partial K) \to \{-1, 1\}$ by letting $g(q) = 1$ if $q \notin K$ and $g(q) = -1$ for $q \in \mathrm{Int}(K)$. Then $g$ changes sign at each $q \in Q' \cap \partial K$, because the crossing occurs transversally in $S_0 \cup S_1$. If $Q'$ is diffeomorphic to $S^1$, this implies that $\mathrm{card}(Q' \cap \partial K)$ is even. If $Q'$ is diffeomorphic to $\mathbb{R}$, and $\Phi : \mathbb{R} \to Q'$ is a diffeomorphism, then the fact that $\mathrm{card}(Q' \cap \partial K)$ is even follows because $g(\Phi(t)) = 1$ for large $|t|$, which is a consequence of

$$\lim_{|t| \to \infty} \|\Phi(t)\| = \infty\,. \tag{3.5.6}$$

To see that (3.5.6) holds, assume it does not. Then there is a sequence $\{t_k\}$ such that $|t_k| \to \infty$ and $\Phi(t_k)$ has a limit $p$. Then $p \in Q$ because $Q$ is closed, and $p \in \mathrm{Clos}(Q')$, so $p \in Q'$ because $Q'$ is a connected component of $Q$. Then $p = \Phi(\tau)$ for some $\tau$, and $t_k \to \tau$, because $\Phi(t_k) \to \Phi(\tau)$ and $\Phi$ is a diffeomorphism. But this contradicts the fact that $|t_k| \to \infty$. This completes the proof of fact 3.5.6.

Adding the result of fact 3.5.6 over all connected components of $Q$, we find that $\nu + 1 = \mathrm{card}(Q)$ is even. So $\nu$ is odd, and the proof of Lemma 3.5.5 is complete. □

Returning now to the proof of Theorem 3.5.4, we recall the fact that $H : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is a smooth map which does not take the value $y_0$ anywhere on $[a, b] \times \partial\mathbb{B}^n$, and is such that $y_0$ is a regular value of $H$, $H_a$

and $H_b$. In addition, the finite set $Z_a = \{x \in \mathbb{B}^n : H(a,x) = y_0\}$ has odd cardinality. From these facts, we want to conclude that there is a smooth curve $\gamma : [0,1] \to [a,b] \times \mathbb{B}^n$ such that (i) $H(\gamma(t)) = y_0$ for all $t$, (ii) $\gamma(0) \in \{a\} \times \mathbb{B}^n$, and (iii) $\gamma(1) \in \{b\} \times \mathbb{B}^n$. The proof is similar to the argument used to prove Lemma 3.5.5. We let $Q = H^{-1}(y_0)$, so $Q$ is a closed embedded 1-dimensional submanifold of $\mathbb{R} \times \mathbb{R}^n$, and observe that each connected component $Q'$ of $Q$ is diffeomorphic to $\mathbb{R}$ or to $S^1$, and that the cardinality of $Q' \cap (\{a,b\} \times \mathbb{B}^n)$ must be even. Since, on the other hand, $\mathrm{card}(Q \cap (\{a\} \times \mathbb{B}^n))$ is odd, there must exist at least one component $Q'$ such that $\mathrm{card}(Q' \cap (\{a\} \times \mathbb{B}^n))$ is odd. But then $\mathrm{card}(Q' \cap (\{b\} \times \mathbb{B}^n))$ is odd as well, so $Q'$ intersects both $\{a\} \times \mathbb{B}^n$ and $\{b\} \times \mathbb{B}^n$. If $Q'$ is diffeomorphic to $\mathbb{R}$, we let $\Phi : \mathbb{R} \to Q'$ be a diffeomorphism. If $Q'$ is diffeomorphic to $S^1$, then we pick a point $q \in Q'$ such that $q \notin [a,b] \times \mathbb{B}^n$, and let $\Phi : \mathbb{R} \to Q' \backslash \{q\}$ be a diffeomorphism. Then there must exist real numbers $t_1$, $t_2$ such that $t_1 < t_2$, $\Phi(t) \in ]a,b[ \times (\mathbb{B}^n \backslash \partial \mathbb{B}^n)$ for $t_1 < t < t_2$, and one of $\Phi(t_1)$, $\Phi(t_2)$ belongs to $\{a\} \times \mathbb{B}^n$ while the other one is in $\{b\} \times \mathbb{B}^n$. Then the restriction of $\Phi$ to $[t_1, t_2]$, suitably reparametrized, is our desired curve $\gamma$, and our proof is complete. □

**Remark 3.5.7** Theorem 3.5.4 essentially says that if $H = \{H_t\}_{a \le t \le b}$ is a "regular homotopy" (i.e. a homotopy in the usual sense, except that instead of being required to be continuous and single-valued it is allowed to be multivalued and regular) of maps $H_t : \mathbb{B}^n \to \mathbb{R}^n$ such that $H_t$ maps $\partial \mathbb{B}^n$ to $\mathbb{R}^n \backslash \{y_0\}$ with degree one, then (i) each map $H_t - y_0$ has a zero and, in addition, (ii) there is a *connected* subset of $[a,b] \times \mathbb{B}^n$ —precisely, the set $\gamma([0,1])$— that joins a zero of $H_a - y_0$ to a zero of $H_b - y_0$. So this is a weak form of the statement that "there is a path joining a zero of $H_a - y_0$ to a zero of $H_b - y_0$." □

**Remark 3.5.8** In Remark (3.5.7), it is not possible to make the stronger assertion that "there is a path joining a zero of $H_a - y_0$ to a zero of $H_b - y_0$," if by "path" we truly mean "continuous curve," rather than just "connected set." □

Indeed, the following example shows that this stronger statement can fail to be true, even if $H$ is a smooth single-valued map.

**Example 3.5.9** Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a $C^\infty$ function that vanishes of infinite order at 0, and is strictly positive elsewhere. Let $\psi : \mathbb{R} \to \mathbb{R}$ be of class $C^\infty$ and such that $\psi(x) = 0$ for $-1 \le x \le 1$, $\psi(x) < 0$ for $x < -1$, $\psi(x) > 0$ for $x > 1$, $\psi(2) = 2$ and $\psi(-2) = -2$. Finally, let $\rho \in C^\infty(\mathbb{R}, \mathbb{R})$ be such that $0 \leq \rho(x) \leq 1$ for all $x$, $\rho(x) = 1$ for $-1 \leq x \leq 1$, $0 < \rho(x) < 1$ for $1 < |x| < 2$, and $\rho(x) = 0$ for $|x| \geq 2$. Define $H : \mathbb{R}^2 \to \mathbb{R}$ by

$$H(t,x) = \rho(x)\varphi(t)\left(x - \sin\left(\frac{\pi}{t}\right)\right) + (1-\rho(x))\psi(x)\,.$$

Then $H \in C^\infty$, $H(t,x) = x$ for $x \in \partial([-2,2])$, and the zero set $Z$ of $H$ is the set

$$Z = \left\{(t,x) : t \neq 0\,,\, x = \sin\left(\frac{\pi}{t}\right)\right\} \cup \left(\{0\} \times [-1,1]\right).$$

Clearly, $(-1,0)$ is the unique zero of $H(t,x)$ for $t = -1$, and $(1,0)$ is the unique zero for $t = -1$. It is easy to see that there is no continuous path in $Z$ joining these two points. On the other hand, $Z$ is connected, and the set-valued map $\gamma \in SVM([-1,1], \mathbb{R}^2)$ given by $\gamma(t) = \left\{(t, \sin\left(\frac{\pi}{t}\right))\right\}$ for $t \neq 0$, $\gamma(0) = \{0\} \times [-1,1]$ is regular and such that $\gamma([-1,1]) = Z$, consistently with Theorem 3.5.4. (To see that $\gamma$ is regular, define $\gamma_k(t) = \left(1 - \chi_{I_k}(t)\right)\sin\left(\frac{\pi}{t}\right)$, where $\chi_I$ is the indicator function of the set $I$, and $I_k = [-\frac{1}{k}, \frac{1}{k}]$. Then the $\gamma_k$ are continuous, and $\gamma_k \xrightarrow{\mathrm{gr}} \gamma$.) □

**Remark 3.5.10** In Example 3.5.9, the map $\gamma$ cannot be taken to be continuous, but it does satisfy $\gamma(t) \subseteq \{t\} \times \mathbb{R}^n$, i.e. the "time" parameter along $\gamma$ is the parameter of the homotopy. It is easy to construct examples where a $\gamma$ with this property does not exist. For example, let

$$H(t,x) = \rho(t,x)(x^3 - 3x - t)\,,$$

where the factor $\rho$ is strictly positive and chosen so that $H(t,4) = 4$ and $H(t,-4) = -4$ for $t \in [-50,50]$. Then $H(t,\cdot)$ has three zeros for $-2 < t < 2$, and one zero for $|t| > 2$, and all the zeros for $t \in [-50,50]$ lie in $[-4,4]$. It is clear that the points $(-16,-3)$ and $(16,3)$ lie in the zero set $Z$ of $H$, but there is no $\gamma \in \mathrm{REG}([-16,16], \mathbb{R}^2)$ taking values in $Z$ and such that

$$\gamma(t) \subseteq \{t\} \times \mathbb{R}^2 \quad \text{for} \quad t \in [-16,16]\,. \tag{3.5.7}$$

On the other hand, if we do not require (3.5.7), then we can take $\gamma$ to be a smooth curve by letting, for example $\gamma(s) = (s^3 - 3s, s)$, for $s \in [-3,3]$. □

**Remark 3.5.11** For completeness, we observe that Fact 3.5.3 is a trivial special case of Theorem 3.5.4 (just taking $H_t$ to be independent of $t$), and Fact 3.5.3 easily implies Fact 3.5.1, i.e. the Brouwer Fixed Point Theorem for regular maps. To see this, we observe that Fact 3.5.3 implies the no-retraction theorem, according to which a continuous map $F : \mathbb{B}^n \to \partial\mathbb{B}^n$ cannot be such that $F(x)=x$ for $x\in\partial\mathbb{B}^n$. (Indeed, such an $F$, regarded as a map from $\mathbb{B}^n$ to $\mathbb{B}^n$, satisfies the hypothesis of Fact 3.5.3, with $\alpha=1$, so $F^{-1}(0)\neq\emptyset$, contradicting the fact that $F(\mathbb{B}^n) \subseteq \partial\mathbb{B}^n$.) A well known elementary argument shows that the no-retraction theorem implies the Brouwer theorem for continuous maps. Finally, the Brouwer theorem for continuous maps implies the theorem for regular maps by an elementary limiting argument, as shown above following the statement of Fact 3.5.1. □

# 4 Multidifferential Calculus

*Throughout this section, we assume that $X$ and $Y$ are finite-dimensional real linear spaces, and use $Z$ to denote the space $L(X,Y)$ of all linear maps from $X$ to $Y$. We assume that norms —denoted in all cases by $\|\ldots\|$— are fixed on $X$, $Y$ and $Z$, and use $B_X$, $B_Y$, $B_Z$ to denote the corresponding closed unit balls.*

## 4.1 Multidifferentials

If $S$ is a convex subset of $X$, $\mathbf{D} \subseteq Z$, and $k = 1, 2, \ldots$, we define $C^k_{\mathbf{D}}(S,Y)$ to be the set of all maps $H : S \to Y$ that admit an extension to a $C^k$ map $\tilde{H} : \Omega \to Y$, defined on an open subset $\Omega$ of $X$ such that $\mathrm{Clos}(S) \subseteq \Omega$, having the property that

$$D\tilde{H}(x) \in \mathbf{D} \;\text{ for all }\; x \in S\,. \tag{4.1.1}$$

If $\mathbf{D}$ is closed in $Z$, then of course (4.1.1) is equivalent to the requirement that $D\tilde{H}(x) \in \mathbf{D}$ for all $x \in \mathrm{Clos}(S)$.

**Definition 4.1.1** Let $F : X \to 2^Y$ be a set-valued map. Let $(\bar{x},\bar{y}) \in X\times Y$, and let $\mathbf{D}$ be a compact subset of $Z$. Let $C$ be a convex cone in $X$. We say that $\mathbf{D}$ is a *multidifferential of $F$ at $(\bar{x},\bar{y})$ in the direction of $C$*, and write

$$\mathbf{D} \in MD_C(F;\bar{x},\bar{y})\,,$$

if the following is true:

(MD) for every neighborhood $\mathbf{D}'$ of $\mathbf{D}$ in $Z$ there exists a pair $(\rho,\Theta)$ such that

(MMD.1) $\rho > 0$, and $\Theta$ is a function on the interval $]0, \rho]$ with values in $[0, \infty[$ and such that $\lim_{s \downarrow 0} \Theta(s) = 0$,

(MMD.2) for every $\varepsilon > 0$ there exist $f, H$ such that

(MMD.2.1) $f \in \mathrm{REG}(\bar{x} + (C \cap \rho B_X), Y)$;

(MMD.2.2) $\mathrm{G}(f) \subseteq \mathrm{G}(F)$;

(MMD.2.3) $H \in C^1_{\mathbf{D}'}(\bar{x} + (C \cap \rho B_X), Y)$;

(MMD.2.4) $H(\bar{x}) = \bar{y}$;

(MMD.2.5) the inequality

$$\sup\left\{\|y - H(x)\| : y \in f(x)\right\} \leq \varepsilon + \Theta(\|x - \bar{x}\|)\|x - \bar{x}\| \qquad (4.1.2)$$

holds whenever $x - \bar{x} \in C \cap \rho B_X$. □

**Definition 4.1.2** A pair $(\rho, \Theta)$ for which (MMD.1,2) hold will be called a *modulus of multidifferentiability* for the 7-tuple $(X, Y, \bar{x}, \bar{y}, C, F, \mathbf{D}')$. □

So $\mathbf{D} \in MD_C(F; \bar{x}, \bar{y})$ iff for every neighborhood of $\mathbf{D}$ in $L(X, Y)$ there exists a modulus of multidifferentiability for $(X, Y, \bar{x}, \bar{y}, C, F, \mathbf{D}')$.

**Definition 4.1.3** A modulus of multidifferentiability $(\rho, \Theta)$ will be called *nice* if $\Theta$ is continuous and strictly increasing on $[0, \rho]$ and the function $s \to s\Theta(s)$ is of class $C^1$ on $[0, \rho]$ and has a derivative that vanishes at $s = 0$. □

The following observation will be important in the proof of the chain rule:

**Fact 4.1.4** *If a 7-tuple $\mathcal{X} = (X, Y, \bar{x}, \bar{y}, C, F, \mathbf{D}')$ has a modulus of multidifferentiability, then $\mathcal{X}$ has a nice modulus of multidifferentiability.*

*Proof.* Let $(\rho, \Theta)$ be a modulus of multidifferentiability for $\mathcal{X}$. Then it is clear that, if $0 < \rho' \leq \rho$, and $\Theta' : ]0, \rho'] \to \mathbb{R}$ satisfies $\Theta(s) \leq \Theta'(s)$ for all $s \in ]0, \rho']$, then $(\rho', \Theta')$ is also a modulus of multidifferentiabilty for $\mathcal{X}$.

By making $\rho$ smaller, if necessary, we may assume that $\Theta$ is bounded. Let $r_n = 2^{-n}\rho$ for $n = 0, 1, 2, \ldots$. For each strictly positive integer $n$, let $y_n = \sup\{\Theta(s) : 0 < s \leq 2r_n\}$. Then $y_n \downarrow 0$ as $n \to \infty$. Define $\Theta_1 : ]0, \rho] \to \mathbb{R}$ by letting $\Theta_1(r_n) = y_n$ for $n = 1, 2, \ldots$, $\Theta_1(\rho) = y_1$, and interpolating linearly on each interval $[r_{n+1}, r_n]$. Then $\Theta_1$ is continuous on $]0, \rho]$ and $\Theta_1(s)$ goes to 0 as $s \downarrow 0$, so we may extend $\Theta_1$ to a continuous function on $[0, \rho]$ that vanishes at 0. It is then clear that $\Theta_1(s) \geq \Theta(s)$ for all $s \in ]0, \rho]$, and $\Theta_1$ is monotonically nondecreasing.

Next define

$$\Theta_2(s) = \frac{1}{s}\int_s^{2s} \Theta_1(r)dr \quad \text{for} \quad s > 0\,, \tag{4.1.3}$$

and let $\Theta_2(0) = 0$. Once again, $\Theta_2$ is continuous. Moreover,

$$\Theta_2(s) \geq \frac{1}{s}\int_s^{2s} \Theta_1(s)dr = \Theta_1(s)\,,$$

due to the monotonicity of $\Theta_1$. So $\Theta_2(s) \geq \Theta_1(s) \geq \Theta(s)$ for all $s$.

Clearly, $s\Theta_2(s) = \int_s^{2s}\Theta_1(r)dr$, so $s \to s\Theta_2(s)$ is a function of class $C^1$, whose derivative equals $2\Theta_2(2s) - \Theta_2(s)$. In particular, this derivative vanishes at $s = 0$.

Finally, the derivative $\dot{\Theta}_2$ of $\Theta_2$ is given by

$$\dot{\Theta}_2(s) = s^{-1}\Big(2\Theta_1(2s) - \Theta_1(s)\Big) - s^{-2}\int_s^{2s}\Theta_1(r)dr\,.$$

Since $\Theta_1(r) \leq \Theta_1(2s)$ for $s \leq r \leq 2s$, the derivative is bounded below by $s^{-1}\Big(2\Theta_1(2s) - \Theta_1(s)\Big) - s^{-1}\Theta_1(2s)$, i.e. by $s^{-1}\Big(\Theta_1(2s) - \Theta_1(s)\Big)$, which is nonnegative. Therefore $\Theta_2$ is also monotonically nondecreasing. We then take $\Theta_3(s) = s^2 + \Theta_2(s)$ and get a strictly increasing continuous function such that $\Theta_3(0) = 0$, $\Theta_3(s) \geq \Theta(s)$ for all $s$, and $s \to s\Theta_3(s)$ is of class $C^1$, with a vanishing derivative at $s = 0$.

Clearly, then, $(\rho, \Theta_3)$ is a nice modulus of multidifferentiability for $\mathcal{X}$. □

## 4.2 The Chain Rule

**Theorem 4.2.1** *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 $F_i : X_i \to 2^{X_{i+1}}$ 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 $\mathbf{D}_i$ be a compact subset of $L(X_i, X_{i+1})$ such that $\mathbf{D}_i \in MD_{C_i}(F_i; \bar{x}_i, \bar{x}_{i+1})$. Let $S$ be the linear subspace of $X_2$ spanned by $C_2$, and let $\Pi$ be a linear projection from $X_2$ onto $S$. Assume that*

$$F_1(\bar{x}_1 + C_1) \subseteq \bar{x}_2 + C_2\,. \tag{4.2.1}$$

*Then*

$$\mathbf{D}_2 \circ \Pi \circ \mathbf{D}_1 \in MD_{C_1}(F_2 \circ F_1; \bar{x}_1, \bar{x}_3)\,, \tag{4.2.2}$$

*where*

$$\mathbf{D}_2 \circ \Pi \circ \mathbf{D}_1 = \{D_2 \circ \Pi \circ D_1 : D_2 \in \mathbf{D}_2\,,\, D_1 \in \mathbf{D}_1\}\,. \tag{4.2.3}$$

*Proof.* Without loss of generality, we assume that $\bar{x}_i = 0$ for $i = 1, 2, 3$. We fix Euclidean inner products and their corresponding norms —and use for them the notations $\langle \cdot, \cdot \rangle$, $\| \cdot \|$— on the spaces $X_i$, such that $\Pi$ is the orthogonal projection from $X_2$ onto $S$. We then endow the spaces $L(X_i, X_j)$ with their corresponding operator norms. We use $B_i$ to denote the closed unit ball in $X_i$.

We fix a vector $v$ belonging to the interior of $C_2$ relative to $S$ and such that $\|v\| = 1$, and a $\Delta > 0$ such that $v + \Delta(B_2 \cap S) \subseteq C_2$. For $\alpha \in \mathbb{R}$, we let $T_\alpha : S \to S$ be the translation map given by $T_\alpha(x) = x + \alpha v$. The following trivial fact will be important in our proof:

**Fact 4.2.2** *If $x \in S$, $\alpha \geq 0$, and* $\operatorname{dist}(x, C_2) \leq \alpha\Delta$, *then $T_\alpha(x) \in C_2$.*

Indeed, let $x' \in C_2$ be such that $\|x - x'\| \leq \alpha\Delta$. Write $x - x' = \alpha y$, where $\|y\| \leq \Delta$. Then $T_\alpha(x) = x - x' + \alpha v + x' = \alpha(y + v) + x'$. The facts that $\|y| \leq \Delta$ and $y \in S$ imply that $y + v \in C_2$. Since $x' \in C_2$ and $\alpha \geq 0$, our conclusion follows. □

Let $F = F_2 \circ F_1$, $\mathbf{D} = \mathbf{D}_2 \circ \Pi \circ \mathbf{D}_1$. To prove that $\mathbf{D} \in MD_{C_1}(F; 0, 0)$, we pick a neighborhood $\mathbf{D}'$ of $\mathbf{D}$ in $L(X_1, X_3)$, and construct a modulus of multidifferentiability $(\rho, \Theta)$ for $(X_1, X_3, 0, 0, C_1, F, \mathbf{D}')$. For this purpose, we choose, for $i = 1, 2$, neighborhoods $\mathbf{D}'_i$ of $\mathbf{D}_i$ in $L(X_i, X_{i+1})$, and a neighborhood $\mathbf{N}$ in $L(X_1, X_2)$ of the zero map, such that

$$\mathbf{D}'_2 \circ \Pi \circ \Big(\mathbf{D}'_1 + \mathbf{N}\Big) \subseteq \mathbf{D}' , \tag{4.2.4}$$

and then fix nice moduli of multidifferentiability $(\rho_i, \Theta_i)$, for $i = 1, 2$, for the 7-tuples $(X_i, X_{i+1}, 0, 0, C_i, F_i, \mathbf{D}'_i)$. (The $(\rho_i, \Theta_i)$ can be taken to be nice thanks to Fact 4.1.4.)

We let $\psi(s) = s\Theta_1(s)$, so $\psi$ is a strictly increasing function of class $C^1$ on $[0, \rho_1]$ such that $\psi(0) = 0$ and $\dot{\psi}(0) = 0$. We define $\Psi(x) = \psi(\|x\|)$, and observe that $\Psi$ is a real-valued function of class $C^1$ on the ball $\rho B_1$, whose gradient vanishes at $x = 0$.

Also, we let

$$M_i = \sup\Big\{\|D\| : D \in \mathbf{D}'_i\Big\} . \tag{4.2.5}$$

After making $\rho_1$ smaller, if necessary, we assume that

$$2\Big(\rho_1 + \Theta_1(\rho_1) + M_1\Big)\rho_1 \leq \rho_2 , \tag{4.2.6}$$

and

$$\Big(\rho_1 + \Theta_1(\rho_1)\Big)\rho_1 \leq \frac{\Delta\rho_2}{4} . \tag{4.2.7}$$

If $\eta$ is a real-valued function of class $C^1$ on $\rho_1 B_1$, we can consider the map $P_\eta : \rho_1 B_1 \to X_2$ given by

$$P_\eta(x) = \eta(x)v\,. \tag{4.2.8}$$

Then $P_\eta \in C^1$. We are particularly interested in the function $\eta$ given by

$$\eta(x) = \frac{\Psi(x)}{\Delta} = \frac{\Theta_1(\|x\|).\|x\|}{\Delta}\,, \tag{4.2.9}$$

so from now we fix $\eta$ to be this function. Notice that $\eta$ satisfies $\nabla\eta(0) = 0$. So $DP_\eta(0) = 0$. Since $DP_\eta$ is continuous, we can make $\rho_1$ even smaller, if necessary, and assume that

$$DP_\eta(x) \in \mathbf{N} \quad \text{whenever} \quad x \in \rho_1 B_1\,. \tag{4.2.10}$$

We now specify $\rho$ and $\Theta$. We take $\rho = \rho_1$, and then define $\Theta$ by

$$\Theta(s) = M_2 \frac{\Delta{+}1}{\Delta}\Theta_1(s) + (M_1{+}\Theta_1(\rho))\Theta_2\Big(s(M_1{+}\Theta_1(s))\Big)\,. \tag{4.2.11}$$

(Notice that if $s \in [0,\rho]$ then $s(M_1 + \Theta_1(s)) \in [0,\rho_2]$, so $\Theta_2\Big(s(M_1{+}\Theta_1(s))\Big)$ is defined.) Then $\Theta$ is continuous and nonnegative, and $\Theta(0) = 0$.

We now prove that $(\rho,\Theta)$ is a modulus of multidifferentiability for the 7-tuple $(X_1, X_3, 0, 0, C_1, F, \mathbf{D}')$. It is clear that (MMD.1) holds, so we have to prove (MMD.2). Let $\varepsilon > 0$ be arbitrary. Choose $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, such that

$$\begin{aligned} \varepsilon_1 &\leq \rho^2\,, & (4.2.12)\\ M_2\,\frac{\Delta+1}{\Delta}\,\varepsilon_1 + \varepsilon_2 + \Theta_2(\rho_2)\varepsilon_1 &\leq \frac{\varepsilon}{3}\,, & (4.2.13)\\ \sup\left\{\Theta_2(\varepsilon_1{+}r) - \Theta_2(r) : 0 \leq r \leq \frac{\rho_2}{2}\right\} &\leq \frac{\varepsilon}{3\rho(M_1{+}\Theta_1(\rho))}\,. & (4.2.14) \end{aligned}$$

Then use the fact that the $(\rho_i,\Theta_i)$ are moduli of multidifferentiability for $(X_i, X_{i+1}, 0, 0, C_i, F_i, \mathbf{D}'_i)$ to obtain maps $f_i \in \mathrm{REG}(C_i \cap \rho_i B_i, X_{i+1})$ and $H_i \in C^1_{\mathbf{D}'_i}(C_i \cap \rho_i B_i, X_{i+1})$, such that $\mathrm{G}(f_i) \subseteq \mathrm{G}(F_i)$, $H_i(0) = 0$, and

$$\sup\Big\{\|y - H_i(x)\| : y \in f_i(x)\Big\} \leq \varepsilon_i + \Theta_i(\|x\|)\|x\| \ \text{ if } \ x \in C_i \cap \rho_i B_i\,. \tag{4.2.15}$$

We then observe that

$$f_1(C_1 \cap \rho B_1) \subseteq C_2 \cap \rho_2 B_2\,. \tag{4.2.16}$$

Indeed, if $x_1 \in C_1 \cap \rho B_1$ and $x_2 \in f_1(x_1)$, we have

$$\|x_2 - H_1(x_1)\| \leq \varepsilon_1 + \Theta_1(\|x_1\|)\|x_1\| \leq \rho^2 + \Theta_1(\rho)\rho\,. \tag{4.2.17}$$

Moreover, since $DH_1(x) \in \mathbf{D}_1'$ for all $x \in C_1 \cap \rho B_1$, and $H_1(0) = 0$, we have $\|H_1(x_1)\| \leq M_1\|\rho\|$. So

$$\|x_2\| \leq \Big(\rho + \Theta_1(\rho) + M_1\Big)\rho \leq \frac{\rho_2}{2}\,. \tag{4.2.18}$$

In addition, we are assuming that $F_1(C_1) \subseteq C_2$. So $x_2 \in C_2$, and then (4.2.16) is proved.

We now define $f = f_2 \circ f_1$. Since $G(f_1) \subseteq G(F_1)$ and $G(f_2) \subseteq G(F_2)$, it is clear that $G(f) \subseteq G(F)$.

Next we show that $f \in \mathrm{REG}(C_1 \cap \rho B_1, X_3)$. To begin with, we have to prove that $G(f)$ is compact. Let $\{(x_1^j, x_3^j)\}$ be a sequence of points in $G(f)$. Pick $x_2^j \in f_1(x_1^j)$ such that $x_3^j \in f_2(x_2^j)$. Since $G(f_1)$ is compact, there is an infinite subset $J$ of $\mathbb{N}$ such that $(x_1, x_2) = \lim_{j\to\infty, j\in J}(x_1^j, x_2^j)$ exists and belongs to $G(f_1)$. Then, using the compactness of $G(f_2)$, we can find an infinite subset $J'$ of $\mathbb{N}$ such that $(x_2', x_3)) = \lim_{j\to\infty, j\in J'}(x_2^j, x_3^j)$ exists and belongs to $G(f_2)$. Since $J' \subseteq J$, we have $x_2' = x_2$. Then $(x_1, x_3) \in G(f)$, and $\lim_{j\to\infty, j\in J'}(x_1^j, x_3^j) = (x_1, x_3)$. So $G(f)$ is compact, as stated.

We now have to find a sequence $\{f^j\}$ of continuous $X_3$-valued functions on $C_1 \cap \rho B_1$ such that $f^j \xrightarrow{\mathrm{gr}} f$. Using the fact that $f_1$ and $f_2$ are regular, we find sequences $\{f_i^j\}_{j=1}^{\infty}$, for $i = 1, 2$, such that $f_i^j \in C^0(C_i \cap \rho_i B_i, X_{i+1})$ and $f_i^j \xrightarrow{\mathrm{gr}} f_i$ as $j \to \infty$. Let

$$\alpha_i^j = \sup\left\{\mathrm{dist}\Big((x_i, f_i^j(x_i)), G(f_i)\Big) : x_i \in C_i \cap \rho_i B_i\right\}, \tag{4.2.19}$$

so $\alpha_i^j \to 0$ as $j \to \infty$.

We would like to define $f^j = f_2^j \circ f_1^j$, but we cannot do this because it need not be true that $f_1^j(C_1 \cap \rho B_1) \subseteq C_2 \cap \rho_2 B_2$. On the other hand, if we define $\sigma : X_2 \to C_2 \cap \rho_2 B_2$ by letting $\sigma(x)$ be the point of $C_2 \cap \rho_2 B_2$ closest to $x$, then $\sigma$ is well defined (because the norm on $X_2$ is Euclidean, so that the balls in $X_2$ are strictly convex) and is a continuous (actually, Lipschitz with constant 1) retraction from $X_2$ to $C_2 \cap \rho_2 B_2$. We then define $f^j = f_2^j \circ \sigma \circ f_1^j$, so now the $f^j$ are continuous maps from $C_1 \cap \rho B_1$ to $X_3$.

We now show that $f^j \xrightarrow{\mathrm{gr}} f$. If this were not so, we would be able to find an infinite subset $J$ of $\mathbb{N}$, a constant $c > 0$, and a family $\{x_1^j\}_{j\in J}$ of points of $C_1 \cap \rho B_1$ such that, if we let $x_3^j = f^j(x_1^j)$, $q^j = (x_1^j, x_3^j)$, then

$$\mathrm{dist}(q^j, G(f)) \geq c \quad \text{for} \quad j \in J\,. \tag{4.2.20}$$

Let $x_2^j = f_1^j(x_1^j)$, $\tilde{x}_2^j = \sigma(x_2^j)$. Then $x_3^j = f_2^j(\tilde{x}_2^j)$. Since $(x_1^j, x_2^j) \in G(f_1^j)$, we can find $(\hat{x}_1^j, \hat{x}_2^j) \in G(f_1)$ such that $\|x_1^j - \hat{x}_1^j\| + \|x_2^j - \hat{x}_2^j\| \le 2\alpha_1^j$. By shrinking $J$ if necessary, we can assume, using the compactness of $G(f_1)$, that the limit $(\hat{x}_1, \hat{x}_2)$ of $(\hat{x}_1^j, \hat{x}_2^j)$ as $j \to \infty$ via values in $J$ exists and belongs to $G(f_1)$. Then $\lim_{j\to\infty, j\in J} x_1^j = \hat{x}_1$, and $\lim_{j\to\infty, j\in J} x_2^j = \hat{x}_2$. Since $\hat{x}_2 \in f_1(\hat{x}_1)$, we have $\hat{x}_2 \in C_2 \cap \rho_2 B_2$. So $\sigma(\hat{x}_2) = \hat{x}_2$. Since $\sigma$ is continuous, we have $\lim_{j\to\infty, j\in J} \sigma(x_2^j) = \sigma(\hat{x}_2)$, i.e. $\lim_{j\to\infty, j\in J} \tilde{x}_2^j = \hat{x}_2$.

We now find $(\check{x}_2^j, \hat{x}_3^j) \in G(f_2)$ such that $\|\tilde{x}_2^j - \check{x}_2^j\| + \|x_3^j - \hat{x}_3^j\| \le 2\alpha_2^j$. By shrinking $J$ even more, if necessary, we can assume, using the compactness of $G(f_2)$, that the limit $(\check{x}_2, \hat{x}_3)$ of $(\check{x}_2^j, \hat{x}_3^j)$ as $j \to \infty$ via values in $J$ exists and belongs to $G(f_2)$. Then $\lim_{j\to\infty, j\in J} \tilde{x}_2^j = \check{x}_2$, and $\lim_{j\to\infty, j\in J} x_3^j = \hat{x}_3$. Since $\lim_{j\to\infty, j\in J} \tilde{x}_2^j = \hat{x}_2$, we have $\check{x}_2 = \hat{x}_2$. Since $\hat{x}_3 \in f_2(\check{x}_2) = f_2(\hat{x}_2)$, and $\hat{x}_2 \in f_1(\hat{x}_1)$, we have $\hat{x}_3 \in f(\hat{x}_1)$. So $(\hat{x}_1, \hat{x}_3) \in G(f)$. On the other hand, $x_1^j \to \hat{x}_1$ and $x_3^j \to \hat{x}_3$ as $j \to \infty$ via values in $J$. So $\mathrm{dist}(q^j, G(f)) \to 0$ as $j \to \infty$ via $J$, contradicting (4.2.20). This completes the proof that $f \in \mathrm{REG}(C_1 \cap \rho B_1)$.

We now have to construct the $C^1$ map $H$. Once again, we would like to define $H = H_2 \circ H_1$, but this is not possible because $H_1$ need not take values in $C_2 \cap \rho_2 B_2$. It will turn out, however, that $H_1(x)$ is close to $C_2$, so that if we add to $H_1(x)$ a small multiple $\alpha(x)v$ of $v$, the resulting point $H_1(x) + \alpha(x)v$ is in $C_2$. Moreover, we will show that the function $\alpha$ can be chosen to be of class $C^1$ and with a vanishing gradient at 0. This will imply that, if $\rho_1$ is small enough, then the differentials of the modified $H_1$ are close to those of the original one.

To make all this precise, we first observe that, if $x \in C_1 \cap \rho B_1$, then

$$\|H_1(x)\| \le M_1\|x\| \le \frac{\rho_2}{2} \quad \text{and} \quad \mathrm{dist}(H_1(x), C_2) \le \varepsilon_1 + \Psi(x)\,. \tag{4.2.21}$$

(The first bound follows because $H_1$ is Lipschitz with constant $M_1$, and $H_1(0) = 0$. To prove the second bound, pick $x \in C_1 \cap \rho B_1$, $y \in f_1(x)$, and observe that $y \in C_2$ and $\|H_1(x) - y\| \le \varepsilon_1 + \Theta_1(\|x\|)\|x\|$.)

Define

$$\alpha(x) = \frac{\varepsilon_1 + \Psi(x)}{\Delta} = \frac{\varepsilon_1}{\Delta} + \eta(x)\,. \tag{4.2.22}$$

Then if $x \in C_1 \cap \rho B_1$ we have $\mathrm{dist}(H_1(x), C_2) \le \Delta\alpha(x)$, so

$$\mathrm{dist}(\Pi(H_1(x)), C_2) \le \Delta\alpha(x) \tag{4.2.23}$$

as well. Therefore Fact 4.2.2 tells us that the point

$$\hat{H}_1(x) = T_{\alpha(x)}(\Pi(H_1(x))) = \Pi(H_1(x)) + \alpha(x)v \tag{4.2.24}$$

is in $C_2$. Moreover, the first bound of (4.2.21) implies that

$$\|\hat{H}_1(x)\| \le M_1\|x\| + |\alpha(x)| \le \Big(M_1 + \frac{\rho+\Theta_1(\rho)}{\Delta}\Big)\rho \le \frac{3\rho_2}{4}\,, \qquad (4.2.25)$$

using (4.2.6), (4.2.7), and( 4.2.12).

Then $\hat{H}_1$ maps $C_1 \cap \rho B_1$ into $C_2 \cap \rho_2 B_2$, so the composite $H = H_2 \circ \hat{H}_1$ is well defined as a map from $C_1 \cap \rho B_1$ to $X_3$. Clearly, $H$ is the composite $H_2 \circ \Pi \circ (H_1 + P_\alpha)$ of three maps of class $C^1$, so $H$ is of class $C^1$ as well. (Cf. (4.2.8) for the definition of $P_\eta$ for any function $\eta$.)

If $x \in C_1 \cap \rho B_1$, then the derivative $DH(x)$ is given by

$$DH(x) = DH_2\Big(\Pi(H_1(x) + P_\alpha(x))\Big) \circ \Pi \circ (DH_1 + DP_\alpha)(x)\,. \qquad (4.2.26)$$

Since $DH_1(x) \in \mathbf{D}'_1$, $DP_\alpha(x) \in \mathbf{N}$ (because of (4.2.10), together with the fact that $P_\alpha$ and $P_\eta$ differ by a constant), and $DH_2(\Pi(H_1(x)+P_\alpha(x))) \in \mathbf{D}'_2$, we have $DH(x) \in \mathbf{D}'_2 \circ \Pi \circ \Big(\mathbf{D}'_1 + \mathbf{N}\Big)$, so $DH(x) \in \mathbf{D}'$. We have therefore proved that

$$H \in C^1_{\mathbf{D}'}(C_1 \cap \rho B_1; X_3)\,. \qquad (4.2.27)$$

We now have to estimate $f - H$. Precisely, we pick $x \in C_1 \cap \rho B_1$ and $y \in f(x)$, and estimate $y - H(x)$.

By the definition of $f$, there exists $z \in f_1(x)$ such that $y \in f_2(z)$. We then have

$$\|z - H_1(x)\| \le \varepsilon_1 + \Theta_1(\|x\|)\|x\|\,, \qquad (4.2.28)$$

$$\|y - H_2(z)\| \le \varepsilon_2 + \Theta_2(\|z\|)\|z\|\,. \qquad (4.2.29)$$

Since $z \in S$, and $\Pi$ is the orthogonal projection on $S$, (4.2.28) implies in particular that

$$\|z - \Pi(H_1(x))\| \le \varepsilon_1 + \Theta_1(\|x\|)\|x\|\,. \qquad (4.2.30)$$

In addition, $\hat{H}_1(x) - \Pi(H_1(x)) = \alpha(x)v$, so

$$\|z - \hat{H}_1(x))\| \le \frac{\Delta+1}{\Delta}\Big(\varepsilon_1 + \Theta_1(\|x\|)\|x\|\Big)\,. \qquad (4.2.31)$$

Since $H_2$ is Lipschitz with constant $M_2$, (4.2.31) implies

$$\|H_2(z) - H(x)\| \le M_2\frac{\Delta+1}{\Delta}\Big(\varepsilon_1 + \Theta_1(\|x\|)\|x\|\Big)\,. \qquad (4.2.32)$$

Combining this with (4.2.29), we get the bound

$$\|y - H(x)\| \le M_2 \frac{\Delta+1}{\Delta}\Big(\varepsilon_1 + \Theta_1(\|x\|)\|x\|\Big) + \varepsilon_2 + \Theta_2(\|z\|)\|z\| . \qquad (4.2.33)$$

On the other hand, (4.2.21), together with (4.2.28), imply

$$\|z\| \le \varepsilon_1 + \Big(M_1 + \Theta_1(\|x\|)\Big)\|x\| \le \rho_2 . \qquad (4.2.34)$$

Therefore

$$\Theta_2(\|z\|)\|z\| \le \Theta_2(\rho_2)\varepsilon_1 + \Theta_2(\|z\|)(M_1 + \Theta_1(\rho))\|x\| . \qquad (4.2.35)$$

Plugging this into (4.2.33), we get the bound

$$\|y - H(x)\| \le E + \hat{\Theta}(\|x\|)\|x\| , \qquad (4.2.36)$$

where

$$E = M_2 \frac{\Delta+1}{\Delta}\varepsilon_1 + \varepsilon_2 + \Theta_2(\rho_2)\varepsilon_1 , \qquad (4.2.37)$$

and

$$\hat{\Theta}(s) = M_2 \frac{\Delta+1}{\Delta}\Theta_1(s) + (M_1 + \Theta_1(\rho))\Theta_2\Big(\varepsilon_1 + s(M_1 + \Theta_1(s))\Big) . \qquad (4.2.38)$$

It follows from (4.2.13) that

$$E \le \frac{\varepsilon}{3} . \qquad (4.2.39)$$

Moreover, (4.2.11) implies that

$$\hat{\Theta}(s) = \Theta(s) + (M_1 + \Theta_1(\rho))\bigg(\Theta_2\Big(\varepsilon_1 + s(M_1 + \Theta_1(s))\Big) - \Theta_2\Big(s(M_1 + \Theta_1(s))\Big)\bigg) ,$$

so (4.2.14) tells us that

$$\hat{\Theta}(s) \le \Theta(s) + \frac{\varepsilon}{3\rho} \ \text{ for } \ s \in [0, \rho] , \qquad (4.2.40)$$

since $s \in [0, \rho]$ implies $s(M_1 + \Theta_1(s)) \in [0, \frac{\rho_2}{2}]$. Then,

$$\hat{\Theta}(\|x\|)\|x\| \le \Theta(\|x\|)\|x\| + \frac{\varepsilon\|x\|}{3\rho} \le \Theta(\|x\|)\|x\| + \frac{\varepsilon}{3} . \qquad (4.2.41)$$

Combining (4.2.36), (4.2.39) and (4.2.41), we get the bound

$$\|y - H(x)\| \le \frac{2\varepsilon}{3} + \Theta(\|x\|)\|x\| , \qquad (4.2.42)$$

valid for $x \in C_1 \cap \rho B_1$, $y \in f(x)$.

So $H$ satisfies all the desired conditions, except only for the fact that $H(0)$ need not be equal to $0$. In fact, our definition of $H$ implies that $H(0) = H_2(\varepsilon_1 v)$. But then $\|H(0)\| \leq M_2\varepsilon_1$. Clearly, (4.2.13) implies in particular that $M_2\varepsilon_1 \leq \frac{\varepsilon}{3}$. So, if we let $\tilde{H}(x) = H(x) - H(0)$, we now have the bound

$$\|y - \tilde{H}(x)\| \leq \varepsilon + \Theta(\|x\|)\|x\|. \tag{4.2.43}$$

The addition of the constant $-H(0)$ to $H$ does not change its other properties, so the new function $\tilde{H}$, together with $f$, satisfy all our conditions. The proof it thus complete. □

## 4.3 The Transversal Intersection Theorem

**Definition 4.3.1** If $\mathbf{C}_1$, $\mathbf{C}_2$ are sets of convex cones in a finite-dimensional real linear space $X$, we say that $\mathbf{C}_1$ and $\mathbf{C}_2$ are *transversal*, and write $\mathbf{C}_1 \overline{\pitchfork} \mathbf{C}_2$, if $C_1 \overline{\pitchfork} C_2$ whenever $C_1 \in \mathbf{C}_1$ and $C_2 \in \mathbf{C}_2$. We call $\mathbf{C}_1$ and $\mathbf{C}_2$ *strongly transversal*, and write $\mathbf{C}_1 \overline{\overline{\pitchfork}} \mathbf{C}_2$, if $\mathbf{C}_1 \overline{\pitchfork} \mathbf{C}_2$, and there exists a nontrivial linear functional $\mu : X \to \mathbb{R}$ such that $C_1 \cap C_2 \cap \{x : \mu(x) > 0\} \neq \emptyset$ for every $C_1 \in \mathbf{C}_1$, $C_2 \in \mathbf{C}_2$. □

Notice that we do *not* define $\mathbf{C}_1 \overline{\overline{\pitchfork}} \mathbf{C}_2$, to mean "$C_1 \overline{\overline{\pitchfork}} C_2$ whenever $C_1 \in \mathbf{C}_1$ and $C_2 \in \mathbf{C}_2$." Our definition imposes the stronger requirement that the family of vectors $v_{C_1,C_2} \in (C_1 \cap C_2)\backslash\{0\}$, whose existence is guaranteed by the condition that $C_1 \overline{\overline{\pitchfork}} C_2$ for all $C_1$, $C_2$, can be chosen so that all the $v_{C_1,C_2}$ belong to a fixed open half-space $H = \{x : \mu(x) > 0\}$. Naturally, when both $\mathbf{C}_i$'s consist of a single cone $C_i$, we have $\mathbf{C}_1 \overline{\overline{\pitchfork}} \mathbf{C}_2$ iff $C_1 \overline{\overline{\pitchfork}} C_2$.

The following trivial observation will be important later:

**Lemma 4.3.2** *If $\mathbf{C}_1$, $\mathbf{C}_2$ are sets of cones in $X$ such that $\mathbf{C}_2$ consists of a single cone $C_2$ which is not a linear subspace, then $\mathbf{C}_1 \overline{\overline{\pitchfork}} \mathbf{C}_2$ iff $\mathbf{C}_1 \overline{\pitchfork} \mathbf{C}_2$.*

*Proof.* Suppose $\mathbf{C}_1 \overline{\pitchfork} \mathbf{C}_2$. Let $\mu : X \to \mathbb{R}$ be a linear functional such that $\mu(c) \geq 0$ for all $c \in C_2$ and $\mu(c) > 0$ for some $c \in C_2$. (Such a $\mu$ exists because $C_2$ is not a linear subspace. Indeed, if $S$ is the linear span of $C_2$, and $C_2'$ is the closure of $C_2$, then it is clear that $C_2' \neq S$, so the Hahn-Banach theorem yields a $\mu : S \to \mathbb{R}$ with the desired properties, and this $\mu$ can then be extended to $X$.)

Fix a $c \in C_2$ such that $\mu(c) > 0$. Let $C_1 \in \mathbf{C}_1$. Then $C_1 - C_2 = X$, because $C_1 \pitchfork C_2$. Then we can write $c = c_1 - c_2$, $c_i \in C_i$. So the vector $v = c_1 = c_2 + c$ satisfies $v \in C_1 \cap C_2$ and $\mu(v) > 0$. $\square$

If $C$ is a cone in $X$, and $\mathbf{D}$ is a set of linear maps from $X$ to another linear space $Y$, then we use $\mathbf{D}.C$ to denote the set of cones $\{D.C : D \in \mathbf{D}\}$.

**Theorem 4.3.3** *Let $X_1$, $X_2$, $Y$ be finite-dimensional real linear spaces, and let $\bar{x}_1 \in X_1$, $\bar{x}_2 \in X_2$, $\bar{y} \in Y$. For $i = 1, 2$, let $U_i$ be a neighborhood of $\bar{x}_i$ in $X_i$, and let $C_i$ be a closed convex cone in $X_i$. Define $S_i = U_i \cap (\bar{x}_i + C_i)$. Let $F_i : X_i \to 2^Y$ be set-valued maps, having multidifferentials*

$$\mathbf{D}_i \in MD_{C_i}(F_i; \bar{x}_i, \bar{y}).$$

*Let $C_2^*$ be a relatively closed subset of $C_2$ that separates $0$ from the rest of $C_2$. Let $S_2^* = U_2 \cap (\bar{x}_2 + C_2^*)$. Assume that*

$$\mathbf{D}_1 C_1 \pitchfork \mathbf{D}_2 C_2.$$

*Then the set $F_1(S_1) \cap F_2(S_2^*)$ contains a sequence of points $y_j$ such that $y_j \neq \bar{y}$ and $y_j \to \bar{y}$.*

*Proof.* Let $n = \dim Y$. Fix norms on $X_1$, $X_2$, $Y$, and use $\|\ldots\|$ to denote them.

Assume, without loss of generality, that $\bar{x}_1 = 0$, $\bar{x}_2 = 0$, $\bar{y} = 0$. Fix a linear functional $\mu : Y \to \mathbb{R}$ such that, whenever $D_1 \in \mathbf{D}_1$ and $D_2 \in \mathbf{D}_2$, it follows that there exists $w_{D_1,D_2} \in D_1 C_1 \cap D_2 C_2$ such that $\mu(w_{D_1,D_2}) > 0$. After multiplying $\mu$ by a positive constant, if necessary, we may assume that $\|\mu\| \leq 1$.

Pick $w^0, \ldots, w^n \in Y$ such that the closed unit ball $B_Y$ of $Y$ is contained in the interior of the convex hull $K$ of the $w^k$. Then pick $\tilde{w}^0, \ldots, \tilde{w}^n \in Y$ such that $K$ is contained in the interior of the convex hull $\tilde{K}$ of the $\tilde{w}^k$.

Fix $\bar{D}_i \in \mathbf{D}_i$, for $i = 1, 2$. Then each $\tilde{w}^k$ can be expressed as

$$\tilde{w}^k = \bar{D}_1 v_1^k - \bar{D}_2 v_2^k, \quad v_i^k \in C_i. \tag{4.3.1}$$

Given any $r \geq 0$, we can always substitute the pair $(v_1^k + r v_1, v_2^k + r v_2)$ for the pair $(v_1^k, v_2^k)$, where $v_1$, $v_2$ are such that $\bar{D}_1 v_1 = \bar{D}_2 v_2 = w_{\bar{D}_1,\bar{D}_2}$. So, by picking $r$ large enough, we can choose the representations (4.3.1) so that $\mu(\bar{D}_1 v_1^k) \geq 2$ and $\mu(\bar{D}_2 v_2^k) \geq 2$.

We now let $D_1$, $D_2$ vary in open, bounded neighborhoods $Q_1$, $Q_2$ of $\bar{D}_1$, $\bar{D}_2$ in $Z_1$, $Z_2$. Define

$$\tilde{w}^k(D_1, D_2) = D_1 v_1^k - D_2 v_2^k. \tag{4.3.2}$$

Let $\tilde{K}(D_1, D_2)$ be the convex hull of the vectors $\tilde{w}^k(D_1, D_2)$, $k = 0, \ldots, n$. Choose $Q_1$, $Q_2$ small enough, so that the three conditions $K \subseteq \tilde{K}(D_1, D_2)$, $\mu(D_1 v_1^k) \geq 1$, and $\mu(D_2 v_2^k) \geq 1$ hold for all $(D_1, D_2) \in Q_1 \times Q_2$.

For $(D_1, D_2) \in Q_1 \times Q_2$, we can write each $w^k$ as a convex combination

$$w^k = \sum_{\ell=0}^{n} \sigma_\ell^k(D_1, D_2) \tilde{w}^\ell(D_1, D_2) \,, \tag{4.3.3}$$

where the coefficients $\sigma_\ell^k(D_1, D_2)$ are smooth nonnegative functions of the pair $D_1, D_2$, such that $\sum_{\ell=0}^{n} \sigma_\ell^k(D_1, D_2) = 1$.

Define

$$\hat{v}_i^k(D_1, D_2) = \sum_{\ell=0}^{n} \sigma_\ell^k(D_1, D_2) v^\ell \,. \tag{4.3.4}$$

Then the $\hat{v}_i^k$ are smooth $C_i$-valued maps on $Q_1 \times Q_2$, such that, whenever $(D_1, D_2)$ belongs to $Q_1 \times Q_2$, the conditions

$$w^k = D_1 \hat{v}_1^k(D_1, D_2) - D_2 \hat{v}_2^k(D_1, D_2) \,, \tag{4.3.5}$$

$$\mu(D_i \hat{v}_i^k(D_1, D_2)) \geq 1 \ \text{ for } \ i = 1, 2 \tag{4.3.6}$$

are satisfied.

So far, the pair $(\bar{D}_1, \bar{D}_2)$ was fixed. We now choose a finite set $\mathcal{D}$ of such pairs, corresponding open bounded neighborhoods $Q_1(\bar{D}_1, \bar{D}_2)$, $Q_2(\bar{D}_1, \bar{D}_2)$ of $\bar{D}_1$, $\bar{D}_2$ in $Z_1$, $Z_2$ such that

$$\mathbf{D}_1 \times \mathbf{D}_2 \subseteq \bigcup_{(\bar{D}_1, \bar{D}_2) \in \mathcal{D}} Q_1(\bar{D}_1, \bar{D}_2) \times Q_2(\bar{D}_1, \bar{D}_2) \,, \tag{4.3.7}$$

and smooth maps

$$\hat{v}_{i, \bar{D}_1, \bar{D}_2}^k : Q_1(\bar{D}_1, \bar{D}_2) \times Q_2(\bar{D}_1, \bar{D}_2) \to C_i$$

such that, whenever $(D_1, D_2) \in Q_1(\bar{D}_1, \bar{D}_2) \times Q_2(\bar{D}_1, \bar{D}_2)$, it follows, for $k = 0, \ldots, n$, that

$$w^k = D_1 \hat{v}_{1, \bar{D}_1, \bar{D}_2}^k(D_1, D_2) - D_2 \hat{v}_{2, \bar{D}_1, \bar{D}_2}^k(D_1, D_2)$$

and

$$\mu\Big(D_i \hat{v}_{i, \bar{D}_1, \bar{D}_2}^k(D_1, D_2)\Big) \geq 1 \ \text{ for } \ i = 1, 2 \,.$$

We then let $\mathbf{D}_1'$, $\mathbf{D}_2'$ be compact neighborhoods of $\mathbf{D}_1$, $\mathbf{D}_2$ in $Z_1$, $Z_2$, such that

$$\mathbf{D}_1' \times \mathbf{D}_2' \subseteq \bigcup_{(\bar{D}_1, \bar{D}_2) \in \mathcal{D}} Q_1(\bar{D}_1, \bar{D}_2) \times Q_2(\bar{D}_1, \bar{D}_2) \,. \tag{4.3.8}$$

We let $\mathcal{Q}$ be the open covering of the product space $Z_1 \times Z_2$ whose members are the sets $Q_1(\bar{D}_1, \bar{D}_2) \times Q_2(\bar{D}_1, \bar{D}_2)$, for $(\bar{D}_1, \bar{D}_2) \in \mathcal{D}$, together with the set $(Z_1 \times Z_2) \backslash (\mathbf{D}'_1 \times \mathbf{D}'_2)$. We pick a smooth partition of unity for $\mathcal{Q}$, i.e. a collection of smooth nonnegative functions $\varphi_{\bar{D}_1, \bar{D}_2} : Z_1 \times Z_2 \to \mathbb{R}$, for $(\bar{D}_1, \bar{D}_2) \in \mathcal{D}$, together with a smooth nonnegative real-valued function $\varphi_\#$ on $Z_1 \times Z_2$, such that

$$\varphi_\# + \sum_{(\bar{D}_1, \bar{D}_2) \in \mathcal{D}} \varphi_{\bar{D}_1, \bar{D}_2} \equiv 1 \,,$$

$$\mathrm{support}(\varphi_{\bar{D}_1, \bar{D}_2}) \subseteq Q_1(\bar{D}_1, \bar{D}_2) \times Q_2(\bar{D}_1, \bar{D}_2) \ \text{ for } \ (\bar{D}_1, \bar{D}_2) \in \mathcal{D} \,,$$

and

$$\mathrm{support}(\varphi_\#) \subseteq (Z_1 \times Z_2) \backslash (\mathbf{D}'_1 \times \mathbf{D}'_2) \,.$$

We then define

$$v_i^k(D_1, D_2) = \sum_{(\bar{D}_1, \bar{D}_2) \in \mathcal{D}} \varphi_{\bar{D}_1, \bar{D}_2}(D_1, D_2) \hat{v}^k_{i, \bar{D}_1, \bar{D}_2}(D_1, D_2) \tag{4.3.9}$$

for $i = 1, 2$, $k = 0, \ldots, n$, $D_1 \in Z_1$, $D_2 \in Z_2$. It then follows that the $v_i^k$ are smooth compactly supported $C_i$-valued maps on $Z_1 \times Z_2$, such that

$$w^k = D_1 v_1^k(D_1, D_2) - D_2 v_2^k(D_1, D_2) \,, \tag{4.3.10}$$

and

$$\mu(D_i v_i^k(D_1, D_2)) \geq 1 \ \text{ for } \ i = 1, 2 \tag{4.3.11}$$

whenever $D_1 \in \mathbf{D}'_1$, $D_2 \in \mathbf{D}'_2$.

Since the $v_i^k$ are compactly supported, we can fix a constant $\kappa > 0$ such that $\|v_i^k(D_1, D_2)\| \leq \kappa$ and $\|D_i v_i^k(D_1, D_2)\| \leq \kappa$ for all $D_1, D_2, i, k$.

Given any $y \in Y$, we can represent $y$ uniquely as a linear combination

$$y = \sum_{k=0}^{n} \theta_k(y) w^k \,, \quad \sum_k \theta_k(y) = 1 \,, \tag{4.3.12}$$

and the coefficient functions $\theta_k : Y \to \mathbb{R}$ are clearly smooth. If $y \in B_Y$ then $\theta_k(y) \geq 0$ for $k = 0, \ldots, n$, so (4.3.12) is in fact a convex combination.

We now use the fact that $\mathbf{D}_i \in MD_{C_i}(F_i; 0, 0)$. Since each $\mathbf{D}'_i$ is a neighborhood of $\mathbf{D}_i$ in $Z_i$, we can find a common nice modulus of multidifferentiability $(\rho, \Theta)$ for the 7-tuples $(X_i, Y, 0, 0, C_i, F_i, \mathbf{D}'_i)$, $i = 1, 2$. Then the function $\Theta : [0, \rho] \to [0, \infty[$ is strictly increasing, continuous, and such that $\Theta(0) = 0$. Moreover, for every $\varepsilon > 0$ there exist $f_1, H_1, f_2, H_2$ such that, for $i = 1, 2$, the following properties are satisfied:

(P1) $f_i \in \mathrm{REG}(C_i \cap \rho B_{X_i}, Y)$;

(P2) $\mathrm{G}(f_i) \subseteq \mathrm{G}(F_i)$;

(P3) $H_i \in C^1_{\mathbf{D}'_i}(C_i \cap \rho B_{X_i}, Y)$;

(P4) $H_i(0) = 0$;

(P5) the inequality

$$\sup\Big\{\|y - H_i(x)\| : y \in f_i'(x)\Big\} \leq \varepsilon + \Theta(\|x\|)\|x\| \tag{4.3.13}$$

holds for each $i \in \{1, 2\}$, $x \in C_i \cap \rho B_{X_i}$.

It then follows easily that the functions $H_i$ can be assumed to be of class $C^2$. (Indeed, we can apply the conditions of the definition of multidifferentiability using, instead of the $\mathbf{D}'_i$, smaller neighborhoods $\mathbf{D}''_i$ of the $\mathbf{D}_i$, chosen so that $\mathbf{D}'_i$ is a neighborhood of $\mathbf{D}''_i$. For a given $\varepsilon$, we construct $f_i$, $H_i$ so that (P5) holds with $\frac{\varepsilon}{2}$ instead of $\varepsilon$. We then approximate the $H_i$ in the $C^1$ topology by functions $H_i^j$ that are of class $C^2$ and such that $H_i^j(0) = 0$. By choosing $j$ sufficiently large, we can assume that $DH_i^j(x) \in \mathbf{D}'_i$ and $\|H_i^j(x) - H_i(x)\| \leq \frac{\varepsilon}{2}$ for all $x \in C_i \cap \rho B_{X_i}$. Then we can take $H_i^j$ to be our new $H_i$.) Moreover, we can also assume that the $H_i$ are defined on the whole space $X_i$ (although, of course, the condition that $DH_i(x) \in \mathbf{D}'_i$ is only required to hold for $x \in C_i \cap \rho B_{X_i}$). In addition, we can assume that the $H_i$ have compact supports. Also, we can make $\rho$ smaller, if necessary, and assume that the balls $\rho B_{X_i}$ are contained in the neighborhoods $U_i$.

We now consider the system of ordinary differential equations

$$\dot{x}_1 = \sum_k \theta_k(y) v_1^k(DH_1(x_1), DH_2(x_2)), \tag{4.3.14}$$

$$\dot{x}_2 = \sum_k \theta_k(y) v_2^k(DH_1(x_1), DH_2(x_2)). \tag{4.3.15}$$

Since the right-hand side is a compactly supported $C^1$ function of $(x_1, x_2)$ on $X_1 \times X_2$, depending smoothly on the parameter $y \in Y$, this system has global existence and uniqueness of solutions, and the solutions are of class $C^1$ with respect to $y$.

We let $\mathbb{R} \ni t \to (\xi_1^y(t), \xi_2^y(t))$ denote the solution of (4.3.14), (4.3.15) with initial condition $\xi_1^y(0) = 0$, $\xi_2^y(0) = 0$.

If $y \in B_Y$, then the coefficients $\theta_k(y)$ are nonnegative and add up to 1. So $\dot{\xi}_i^y(t) \in C_i$ for all $t$, and then $\xi_i^y(t) \in C_i$ for $t \geq 0$. Moreover, the fact that $\|v_i^k(D_1, D_2)\| \leq \kappa$ for all $D_1, D_2$ implies the bounds $\|\dot{\xi}_i^y(t)\| \leq \kappa$, so

$$\|\xi_i^y(t)\| \leq \kappa t \quad \text{for} \quad y \in B_Y\,,\ t \geq 0\,. \tag{4.3.16}$$

This implies in particular that

$$\xi_i^y(t) \in C_i \cap \rho B_{X_i} \quad \text{whenever} \quad 0 \leq t \leq \frac{\rho}{\kappa}\,. \tag{4.3.17}$$

We now fix $\rho$ so that

$$8\Theta(\rho) \leq 1\,. \tag{4.3.18}$$

Having chosen $\rho$, we pick $\alpha > 0$ as follows. We let $M$ be such that $\|D\| \leq M$ whenever $D \in \mathbf{D}'_2$. We then let $W_+$ be the open ball in $X_2$ with center 0 and radius

$$r_+ = \frac{\rho}{\kappa M}\,. \tag{4.3.19}$$

Using the fact that $C_2^*$ separates 0 from the rest of $C_2$, we pick a neighborhood $W_-$ of 0 in $X_2$ with the property that, if $\gamma : [0,1] \to C_2$ is continuous, $\gamma(0) \in W_-$, and $\gamma(1) \notin W_+$, then $\gamma([0,1]) \cap C_2^* \neq \emptyset$. Clearly, we can pick $W_-$ to be an open ball. We let $r_-$ be its radius, and then choose $\alpha$ such that

$$0 < \alpha < \frac{r_-}{\rho}\,. \tag{4.3.20}$$

With these choices, the following is true:

(#) *if $\Gamma$ is any connected subset of $C_2$ that contains points $p_-$, $p_+$ such that $\|p_-\| \leq \alpha\rho$ and $\|p_+\| \geq \frac{\rho}{\kappa M}$, then $\Gamma \cap C_2^* \neq \emptyset$.*

(Indeed, suppose $\Gamma$, $p_-$, $p_+$ are as in (#) but $\Gamma \cap C_2^* = \emptyset$. Let $S = C_2 \backslash C_2^*$, so $\Gamma \subseteq S$. Then $p_-$ and $p_+$ both belong to the same connected component $S_c$ of $S$. Since $S$ is a relatively open subset of the convex set $C_2$, the set $S_c$ is path-connected. So there is a curve $\gamma : [0,1] \to S_c$ satisfying the endpoint conditions $\gamma(0) = p_-$ and $\gamma(1) = p_+$. Clearly, $\gamma(0) \in W_-$ and $\gamma(1) \notin W_+$. So $\gamma([0,1]) \cap C_2^* \neq \emptyset$, contradicting the fact that $\gamma$ is a curve in $S$.)

We then choose $\varepsilon$ such that

$$8\varepsilon \leq \alpha\rho \quad \text{and} \quad 8\varepsilon \leq \frac{\alpha\rho}{\kappa}\,. \tag{4.3.21}$$

We let $\tau = \frac{\rho}{\kappa}$ (so $8\varepsilon \leq \alpha\tau$), and define

$$\Psi_s(y) = \frac{1}{s\tau}\Big(H_1(\xi_1^y(s\tau)) - H_2(\xi_2^y(s\tau))\Big) \tag{4.3.22}$$

for $0 < s \le 1$, $y \in Y$.

It is then clear that the $t$-derivative of $H_i(\xi_i^y(t))$, for $t \in [0,\tau]$, is equal to $DH_i(\xi_i^y(t)).\dot{\xi}_i^y(t)$, i.e. to

$$DH_i(\xi_i^y(t)).\sum_k \theta_k(y) v_i^k\Big(DH_1(\xi_1^y(t)), DH_2(\xi_2^y(t))\Big)\,.$$

Since $\xi_i^y(t) \in C_i \cap \rho B_{X_i}$, we have

$$\|DH_i(\xi_1^y(t)).v_i^k(DH_1(\xi_1^y(t)), DH_2(\xi_2^y(t)))\| \le \kappa \tag{4.3.23}$$

and

$$\begin{array}{rcl} DH_1(\xi_1^y(t)).v_1^k(DH_1(\xi_1^y(t)), DH_2(\xi_2^y(t))) & & \\ -DH_2(\xi_2^y(t)).v_2^k(DH_1(\xi_1^y(t)), DH_2(\xi_2^y(t))) & = & w^k\,, \end{array}$$

so

$$\left\|\frac{d}{dt}\Big(H_1(\xi_1^y(t))\Big)\right\| \le \kappa. \tag{4.3.24}$$

and

$$\frac{d}{dt}\Big(H_1(\xi_1^y(t)) - H_2(\xi_2^y(t))\Big) = y\,. \tag{4.3.25}$$

Since $H_i(\xi_i^y(0)) = H_i(0) = 0$, we have

$$H_1(\xi_1^y(t)) - H_2(\xi_2^y(t)) = ty \quad \text{for} \quad 0 \le t \le \tau\,, \tag{4.3.26}$$

and

$$\|H_i(\xi_i^y(t))\| \le t\kappa \quad \text{for} \quad 0 \le t \le \tau\,,\ i = 1,2\,. \tag{4.3.27}$$

Then $\Psi_s(y) = y$. So $\Psi_s$, restricted to $B_Y$, is the identity map of $B_Y$. Also, (4.3.11) implies that

$$\begin{array}{rl} & \dfrac{d}{dt}\left(\mu\Big(H_2(\xi_2^y(t))\Big)\right) \\ = & \mu\left(DH_2(\xi_2^y(t)).\displaystyle\sum_k \theta_k(y) v_2^k(DH_1(\xi_1^y(t)), DH_2(\xi_2^y(t)))\right) \\ = & \displaystyle\sum_k \theta_k(y).\mu\Big(DH_2(\xi_2^y(t)).v_2^k(DH_1(\xi_1^y(t)), DH_2(\xi_2^y(t)))\Big) \\ \ge & 1\,, \end{array}$$

so that

$$\mu(H_2(\xi_2^y(t)) \ge t \ \text{ if } \ y \in B_Y\,,\ \ 0 \le t \le \tau\,. \tag{4.3.28}$$

We now define

$$\Phi_s(y) = \frac{1}{s\tau}\Big(f_1(\xi_1^y(s\tau)) - f_2(\xi_2^y(s\tau))\Big) \tag{4.3.29}$$

for $0 < s \leq 1$. (In agreement with our general conventions, the set difference occurring in the right-hand side is to be interpreted as the set of all differences $z_1 - z_2$, $z_i \in f_i(\xi_i^y(s\tau))$.) Using (P4), together with the fact that $\Psi_s(y) = y$, we get the estimate

$$\sup\Big\{\|y' - y\| : y' \in \Phi_s(y)\Big\} \leq \frac{2\varepsilon + 2s\rho\Theta(s\rho)}{s\rho} \tag{4.3.30}$$

for $0 < s \leq 1$, $y \in B_Y$. In view of (4.3.18) and (4.3.21), we can conclude that

$$\sup\Big\{\|y' - y\| : y' \in \Phi_s(y)\Big\} \leq \frac{1}{2} \ \text{ for } \ \alpha \leq s \leq 1\,. \tag{4.3.31}$$

The set-valued map $B_Y \times [\alpha, 1] \ni (y, s) \to \Phi_s(y) \in Y$ is regular. So (4.3.31), together with Theorem 3.5.4, imply that there exists a compact connected subset $\Sigma$ of $B_Y \times [\alpha, 1]$ such that $0 \in \Phi_s(y)$ whenever $(y, s) \in \Sigma$, and in addition $\Sigma \cap (B_Y \times \{\alpha\}) \neq \emptyset$ and $\Sigma \cap (B_Y \times \{1\}) \neq \emptyset$.

Now, let $\nu_i : B_Y \times [0, 1] \to C_i$ be, for $i = 1, 2$, the maps defined by $\nu_i(y, s) = \xi_i^y(s\tau)$. We then have the bounds

$$\|\nu_2(y, s)\| \leq \kappa s\tau \ \text{ and } \ \mu(H_2(\nu_2(y, s))) \geq s\tau \ \text{ if } \ s \in [0, 1], y \in B_Y\,. \tag{4.3.32}$$

If $z \in f_2(\nu_2(y, s))$, then

$$\|z - H_2(\nu_2(y, s))\| \leq \varepsilon + \Theta\Big(\|\nu_2(y, s)\|\Big).\|\nu_2(y, s)\|\,. \tag{4.3.33}$$

If $s$ belongs to $[\alpha, 1]$, then the inequalities $8\varepsilon \leq \alpha\tau \leq s\tau$, $\|\nu_2(y, s)\| \leq \kappa s\tau$, and $\Theta\Big(\|\nu_2(y, s)\|\Big) \leq \Theta(\rho) \leq \frac{1}{8}$ must hold. So

$$\|z - H_2(\nu_2(y, s))\| \leq \frac{s\tau}{4}\,. \tag{4.3.34}$$

Since $\mu(H_2(\nu_2(y, s))) = \mu(H_2(\xi_2^y(s\tau))) \geq s\tau$ (because of (4.3.28)), and in addition $\|\mu\| \leq 1$, we have

$$\mu(z) \geq \frac{3s\tau}{4}\,. \tag{4.3.35}$$

So $\mu(z) \neq 0$, and then $z \neq 0$. We have therefore shown that

$$0 \notin f_2(\nu_2(y, s)) \ \text{ if } \ \alpha \leq s \leq 1\,, \ y \in B_Y\,. \tag{4.3.36}$$

Now suppose that $(y,s) \in \Sigma$. Then $0 \in \Phi_s(y)$, so there exists a point $z$ in $f_1(\nu_1(y,s)) \cap f_2(\nu_2(y,s))$. Then this $z$ satisfies

$$z \neq 0\,,\quad z \in F_1(S_1) \cap F_2(S_2)\,,\quad \text{and}\quad \|z\| \leq \hat{\kappa}\rho\,, \tag{4.3.37}$$

where $\hat{\kappa} = 1 + \frac{\kappa}{4}$. (The last bound follows from (4.3.27), which gives $\|H_2(\nu_2(y,s))\| \leq s\tau\kappa = s\varepsilon$, together with (4.3.34).)

Since $\Sigma$ is nonempty, a $z$ with the properties of (4.3.37) exists. We now show that $z$ can actually be chosen to belong to $F_2(S_2^*)$. For this purpose, we let $\tilde{\Sigma} = \nu_2(\Sigma)$, so $\tilde{\Sigma}$ is a compact connected subset of $C_2$. Using the facts that $\Sigma \cap (B_Y \times \{\alpha\}) \neq \emptyset$, and $\Sigma \cap (B_Y \times \{1\}) \neq \emptyset$, pick $y_\alpha, y_1 \in B_Y$ such that $(y_\alpha, \alpha) \in \Sigma$ and $(y_1, 1) \in \Sigma$. Then (4.3.28) implies that $\mu\Big(H_2(\nu_2(y_1,1))\Big) \geq \tau$. Since $\|\mu\| \leq 1$, we have $\|H_2(\nu_2(y_1,1))\| \geq \tau$. Since $H_2$ is Lipschitz with constant $M$, and $H_2(0) = 0$, we have

$$\|\nu_2(y_1,1)\| \geq \frac{\tau}{M} = \frac{\rho}{\kappa M}\,. \tag{4.3.38}$$

On the other hand, (4.3.32) yields

$$\|\nu_2(y_\alpha,\alpha)\| \leq \kappa\alpha\tau = \alpha\rho\,. \tag{4.3.39}$$

So, if we let $p_- = \nu_2(y_\alpha,\alpha)$, $p_+ = \nu_2(y_1,1)$, $\Gamma = \tilde{\Sigma}$, it is clear that (#) applies, yielding the conclusion that $\tilde{\Sigma} \cap C_2^* \neq \emptyset$. If we now pick $(y,s) \in \Sigma$ such that $\nu_2(y,s) \in C_2^*$, and then choose $z \in f_1(\nu_1(y,s)) \cap f_2(\nu_2(y,s))$, we get a $z$ such that $z \neq 0$, $z \in F_1(S_1) \cap F_2(S_2^*)$, and $\|z\| \leq \hat{\kappa}\rho$.

So far, $\rho$ was arbitrary, subject only to the requirement that it be small enough for (4.3.18) to hold. In particular, we can apply our conclusion to a sequence $\{\rho_j\}$ of values of $\rho$, converging towards zero. The corresponding points $z_j$ then satisfy $z_j \in F_1(S_1) \cap F_2(S_2^*)$, $z_j \neq 0$, and $z_j \to 0$. This completes our proof. □

## 4.4 The Open Mapping Theorems

Theorem 4.3.3 trivially implies the *open mapping* and *directional open mapping theorems* for multidifferentiable maps. Precisely, let $X$, $Y$ be finite-dimensional real linear spaces, let $F : X \to 2^Y$ be a set-valued map, let $\bar{x} \in X$, $\bar{y} \in Y$. Let $v \in Y$. We say that $F$ is *v-directionally open at* $(\bar{x},\bar{y})$ if for every neighborhood $U$ of $\bar{x}$ in $X$ there exists a closed convex cone $C$ in $Y$ such that $v \in \mathrm{Int}_Y(C)$, with the property that the image $F(U)$ contains a relative neighborhood of $\bar{y}$ in $\bar{y} + C$.

In the special case when $v = 0$, a set-valued map which is $v$-directionally open at $(\bar{x}, \bar{y})$ is just said to be *open at* $(\bar{x}, \bar{y})$. In that case, $C$ must necessarily equal $Y$. So $F$ is open at $(\bar{x}, \bar{y})$ iff for every neighborhood $U$ of $\bar{x}$ in $X$ the image $F(U)$ is a neighborhood of $\hat{y}$ in $Y$.

The directional open mapping theorem then says:

**Theorem 4.4.1** *Let $X$ and $Y$ be finite-dimensional real linear spaces, and let $F$ be a set-valued map from $X$ to $Y$. Let $C_0$ be a closed convex cone in $X$. Let $\bar{x} \in X$, $\bar{y} \in Y$, $v \in Y$. Let $\mathbf{D} \in MD_{C_0}(F; \bar{x}, \bar{y})$ be such that $v \in \mathrm{Int}_Y(DC_0)$ for every $D \in \mathbf{D}$. Then the restriction of $F$ to the set $\bar{x} + C_0$ is $v$-directionally open at $(\bar{x}, \bar{y})$.*

*Proof.* To prove this using Theorem 4.3.3 we assume, without loss of generality, that $\bar{x} = 0$ and $\bar{y} = 0$. We also assume that $\|v\| = 1$. (This is possible with no loss of generality if $v \neq 0$. Moreover, a trivial compactness argument shows that our theorem is true for $v = 0$ if it is true for all nonzero $v$.)

Fix a neighborhood $U$ of 0 in $X$, and suppose that $F(U \cap C_0)$ does not contain a set of the form $V \cap C$, with $V$ a neighborhood of 0 and $C$ a cone containing $v$ in its interior. It is then easy to conclude that there exists a sequence of points $y_j \in Y$, not belonging to $F(U \cap C_0)$, such that $y_j \to 0$ and $\frac{y_j}{\|y_j\|} \to v$. We can assume that the sequence of the norms $\|y_j\|$ is strictly decreasing.

Let $C_1$ be the cone $[0, \infty[$, in the space $\mathbb{R}$. Then let $C_1^*$ be the subset $\{\|y_j\| : j = 1, 2, \ldots\}$ of $C_1$. Eactly as in Remark 2.2.5, construct a continuous map $F_1 : C_1 \to Y$ such that $F_1(\|y_j\|) = y_j$ for every $j$, which is differentiable at 0 and has the map $\Delta : s \to sv$ as its differential.

The hypothesis that $v$ is an interior point of $DC_0$ for every $D \in \mathbf{D}$ implies that the cones $DC_0$ and $\Delta C_1$ are transversal. Hence $\mathbf{D}C_0 \pitchfork \Delta C_1$. Theorem 4.3.3 then implies that $F(U \cap C_0) \cap F_1(C_1^*)$ contains a sequence of points converging to zero. But the members of $F_1(C_1^*)$ are the $y_j$, which do not belong to $F(U \cap C_0)$. So we have reached a contradiction and the proof is complete. □

## References


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

[2] Browder, F. E., "On the fixed point index for continuous mappings of locally connected spaces," *Summa Bazil. Math. 4*, 253-293.

[3] Clarke, F. H. (1976), "The Maximum Principle under minimal hypotheses," *S.I.A.M. J. Control Optim. 14*, 1078-1091.

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

[5] Graves, L. M. (1950), "Some mapping theorems," *Duke Math. J. 17*, 111-114.

[6] Kaskosz, B., and S. Lojasiewicz Jr. (1985), "A Maximum Principle for generalized control systems," *Nonlinear Anal. TMA 9*, 109-130.

[7] Lojasiewicz Jr., S., "Local controllability of parametrized differential equations," in preparation.

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

[9] Rabinowitz, P. (1971), "Some global results for nonlinear eigenvalue problems," *J. Funct. Analysis 7*, 487-513.

[10] Sard, A. (1942), "The measure of critical values of differentiable maps," *Bull. Amer. Math. Soc. 48*, 883-890.

[11] Spanier, E. H. (1966), *Algebraic topology*, Springer-Verlag, New York.

[12] Sternberg, S. (1964), *Lectures on differential geometry*, Prentice-Hall, Englewood Cliffs, NJ.

[13] Sussmann, H.J. (1994), "A strong version of the Lojasiewicz Maximum Principle," in *Optimal Control of Differential Equations,* Nicolai H. Pavel Ed., M. Dekker Inc., New York.

[14] Sussmann, H.J. (1994), "A strong version of the Maximum Principle under weak hypotheses," *Proc. 33rd IEEE Conference on Decision and Control*, Orlando, FL, 1950-1956.

[15] Sussmann, H.J. (1996), "Some recent results on the maximum principle of optimal control theory." In the *Proceedings of the 1996 Mathematical Theory of Networks and Systems (MTNS) Symposium,* held in St. Louis, Mo., in June 1996. Birkhäuser, Boston.

[16] Sussmann, H.J. (1996), "A strong maximum principle for systems of differential inclusions." in *Proceedings of the 35th IEEE Conference on Decision and Control*, Kobe, Japan, Dec. 1996. IEEE publications.

[17] Warga, J. (1981), "Fat homeomorphisms and unbounded derivate containers," *J. Math. Anal. Appl. 81*, 545-560.

[18] Warga, J. (1983), "Controllability, extremality and abnormality in nonsmooth optimal control," *J. Optim. Theory Applic. 41*, 239-260.

[19] Warga, J. (1983), "Optimization and controllability without differentiability assumptions," *SIAM J. Control and Optimization 21*, 837-855.

[20] Warga, J. (1988), "Homeomorphisms and local $C^1$ approximations," *Nonlinear Anal. TMA 12*, 593-597.

[21] Wazewski, T. (1947), "Sur l'évaluation du domaine d'existence des fonctions implicites réelles ou complexes," *Ann. Soc. Polon. Math. 20*.