TANGENT CONES, GENERALIZED SUBDIFFERENTIAL CALCULUS, AND OPTIMIZATION

by

© Douglas Eric Ward

Department of Mathematics,
Statistics and Computing Science
Dalhousie University
Halifax, Nova Scotia
B3H 4H8

Submitted in partial fulfillment
of the requirements for
the Degree of Doctor of Philosophy
at Dalhousie University
August 1984

Table of Contents

ABSTRACT

Tangent cones play a fundamental role in nonsmooth analysis and optimization theory. In recent years, a plethora of tangent cones have been defined, each having its strengths and limitations. In this dissertation, a comparative study of these many tangent cones is undertaken. The results of this study are applied to the consideration of generalized directional derivatives defined via tangent cones of epigraphs of functions - in particular the upper subderivative of Rockafellar. The calculus of the subgradients associated with these directional derivatives is developed in detail for real-valued and vector-valued functions and applied to obtain new necessary optimality conditions for non-linear programs.

PREFACE

This dissertation got its start from a seminar conducted at Carnegie-Mellon University in the spring of 1982. The participants in the seminar took turns lecturing on parts of Rockafellar's monograph [Ro5]. That seminar has proven to be the springboard for a great deal of research in nonsmooth analysis by J.M. Borwein, H.M. Strojwas and myself, including the research presented here. This dissertation actually owes a great deal to [Ro5] in its organization and approach. Tangent cones are viewed here as the building blocks of nonsmooth analysis. Directional derivatives and subgradients are defined in terms of tangent cones to epigraphs, and a knowledge of the properties of these cones is applied to establish subdifferential calculus formulae. The subdifferential calculus is in turn applied in the development of necessary conditions for optimality in mathematical programs.

Some comments on the subject matter of each chapter are in order. As with (perhaps too) much mathematical research, the material in chapter one is presented in an order opposite to that in which it was originally studied. Chapter one grew out of an attempt to rigorously establish the geometric intuition that no tangent cone could have certain combinations of properties. Not sure how to approach this problem, I decided to write down every possible permutation of the quantifications in tangent cone definitions and examine the properties of each such "tangent cone". The eventual result was Table 1.9.3. Coincidentally, Treiman's valuable new characterization of the Clarke tangent cone was announced while this research was going on and added further motivation to the project.

The examples considered and insights gained in the construction of Table 1.9.3 made it easy to put together the impossibility theorems in section 1.2. The rest of the first chapter is an attempt to make precise the correspondences between the properties of a tangent cone and the quantifications in its definition that became clear in the process of filling in the boxes of Table 1.9.3. The particular properties discussed in chapter one are the ones which seem to be the most important in the study of subdifferential calculus and optimization.

One main conclusion of chapter one, as might be expected, is that (with one notable exception) the tangent cones worth studying are the ones that are already used most often.

In the second chapter, known results are refined in three directions. First, the generalized subdifferential calculus of Rockafellar [Ro3] is developed in finite dimensions under weaker assumptions. This is in analogy with the convex case, where interiority hypotheses can be replaced in finite dimensions by assumptions about relative interiors. In fact, the strongest finite-dimensional convex subdifferential calculus results are corollaries of the theorems of this chapter.

Interestingly enough, Theorem 2.3.1 and its consequences have been derived independently by Ioffe and by Rockafellar with methods that are apparently completely different from each other and completely different from the approach taken here. Each of the three approaches seems to have its own advantages and limitations.

Second, a chain rule formulation not previously studied outside the convex or Lipschitzian settings is examined. Corollaries include new product and quotient rules. Third, tangent cone information from the first chapter is exploited to refine the usual subdifferential regularity conditions for equality in subdifferential calculus inclusions.

It should be noted that most of the "hard work" in chapter two is hidden in Theorem 2.2.2, the starting point of the proofs. The slick method of proof used in Theorem 2.3.1 is due to Jon Borwein; I was able to successfully apply it several times in chapters two and three (in Theorems 2.3.15, 2.4.11, 2.4.12, and Proposition 3.4.6). The flexibility of this approach is demonstrated in section 2.4, where other tangent cones are substituted for the Clarke tangent cone in the proofs of Theorems 2.3.1 and 2.3.15 to deduce new directional derivative inequalities (Theorems 2.4.11 and 2.4.12).

The general chain rule formulation presented in chapter three has some attractive features. It encompasses the two ostensibly distinct chain rule formulations studied in chapter two; in addition, it yields new results for real-valued functions without too much effort. There are, however, some pitfalls in trying to find vector-valued analogues of subdifferential calculus formulae for real-valued functions. It is very easy to fallaciously assume that relationships that are true in the real-valued case carry over without a hitch to an ordered vector space setting. The hypothesis of epigraph regularity was pressed into service to repair one such mistake. It is also easy to pile up assumptions on the ordering until results are rendered trivial. I have endeavored to avoid this pitfall.

The results for vector-valued functions obtained in chapter three are less satisfactory than I had originally hoped. In particular, a stronger subgradient sum formula can be derived by other, more direct means. The epigraph regularity assumption in Theorem 3.2.6 pinpoints the major difficulty and suggests a direction for further research.

The fourth chapter is a survey of the important new "lim inf" inclusions that establish the relationship between the Clarke tangent cone and contingent cone and have far-reaching applications. Recent results of this kind due to Jon Borwein are employed to extend a finite-dimensional result of Ioffe to a reflexive Banach space setting. A product rule for Ioffe's approximate subdifferentials is proved in section 4.3.

The purpose of chapter five is to demonstrate the broad applicability in optimization theory of the tangent cone and subdifferential calculus theory amassed in chapters one through four. Necessary conditions for optimality in mathematical programs involving subgradients, approximate subdifferentials, and upper convex approximates are systematically developed. Section 5.4 elaborates on an idea suggested in several papers by Vlach (references [V1] through [V4])but never fully carried out in those papers. A brief discussion of Pareto optimization appears in section 5.5.

I have attempted to make the presentation self-contained. Many of the proofs are elementary epsilon-delta type arguments (although often neighbourhoods are used rather than actual epsilons and deltas). One of my goals was to demonstrate the elementary nature of this material; even the upper subderivative with its intimidating "lim sup inf" definition can be handled by this tangent cone based approach.

The reader should quickly locate and become familiar with the tables at the end of chapter 1. These tables give definitions of the many tangent cones considered in chapter 1 and compare the properties of these cones. It is hoped that these tables will help the reader navigate the sea of technical results presented in this dissertation.

Acknowledgments

I would first like to express my gratitude to Dr. Jonathan Borwein, my supervisor in this research. He introduced me to convex analysis, and many of the ideas in this dissertation originated in stimulating lectures and personal conversation with him over the past four years. His continual advice and encouragement got this research going, and his invaluable suggestions on how to correct and improve the rough drafts of the manuscript greatly enhanced the quality of the work presented here. He has also generously assisted in several "extramathematical" ways by, for example, giving help and moral support when I was faced with harassment from Canada Immigration.

This research was greatly assisted by the financial support of a Killam Scholarship. My father generously provided funds for the typing of the manuscript.

I very much appreciate the help of Paula Flemming, who contributed hours of time and effort in doing a wonderful job of typing the bulk of this weighty tome; and Laura Smith, who pitched in to help finish the typing on time.

Finally, I thank my dear wife Sherry, whose loving support and sacrifice made this dissertation possible.

CHAPTER I

Tangent Cones

1.1. Tangency operators

Let E be a real, locally convex, Hausdorff topological vector space (hereafter abbreviated l.c.s. or l.c.t.v.s.).

Definition 1.1.1: A tangency operator (on E) is a mapping $R: 2^E \times E \longrightarrow 2^E$.

The class of tangency operators, of course, includes many mappings which bear no resemblance to tangential approximants. Actually, all of the specific tangency operators that we will consider in detail here satisfy in addition at least the following three conditions:

(1.1.1) $R(C,x_0) = \emptyset$ if $x_0 \notin \text{cl } C$ (where "cl C" denotes the closure of $C \subset E$).

(1.1.2) $R(C,x_0)$ is a cone for every C and $x_0 \in \text{cl } C$.

(1.1.3) $R(\alpha C, \alpha x_0) = \alpha R(C,x_0)$ for all $\alpha > 0$ and $(C,x_0) \in 2^E \times E$.

(Here $\alpha S = \{\alpha s | s \in S\}$, and by convention $\alpha\emptyset = \emptyset$.) However, because (1.1.1) is never used in the proofs in section 1.2, we do not include it in Definition 1.1.1. Properties (1.1.2) and (1.1.3) will sometimes be invoked in section 1.2, and we will keep track of when they are needed.

We call a tangency operator which satisfies (1.1.3) a <u>homogeneous tangency operator</u> and a cone-valued tangency operator a <u>tangent cone</u>.

There are other similar definitions in the literature. Vlach ([V3], [V4]) defines a mapping $R: 2^E \to 2^E$ to be an <u>approximation operator</u> if $R(E) = E$, $R(\emptyset) = \emptyset$, and R is isotone with respect to set inclusion - i.e., $R(C_1) \subset R(C_2)$ whenever $C_1 \subset C_2$. These papers also compare approximation operators with related, previously-defined notions.

Notice that Definition 1.1.1 excludes concepts of tangential approximation which can assign more than one set to a given set at a given point - for example, the "tents" of Boltyanskii and indicating cones of Martin, Gardner, and Watkins. A thorough discussion and comparison of these and other such tangential approximants can be found in [Ma1] and [Ma2].

Having made Definition 1.1.1, it is our purpose here to single out desirable properties for tangency operators and to determine which combinations of these properties a tangency operator may simultaneously possess. We will focus in on the following six properties:

(1) $\operatorname{cl} R(C,x_0) = \operatorname{cl} \operatorname{cone}(C-x_0) := \operatorname{cl} \bigcup_{\lambda \geq 0} \lambda(C-x_0)$ whenever C is closed and convex and $x_0 \in C$.

Notice that if R has property (1), $\operatorname{cl} R$ is convex whenever C is.

(2) If $h: \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable near

$x_0 \in h^{-1}(0)$ and $\nabla h(x_0)$ has rank m, then

$$R(h^{-1}(0), x_0) = (\nabla h(x_0))^{-1}(0) .$$

Properties (1) and (2) are expected of a "sensible" tangential approximant. We will call a tangency operator which satisfies (1) and (2) a standard tangency operator.

(3) If $C_1 \subset C_2 \subset E$, $R(C_1, x_0) \subset R(C_2, x_0)$.

This is the isotonicity with respect to set inclusion mentioned earlier. As Vlach ([V3], [V4]) points out, the following are each equivalent to property (3):

(3a) $R(C_1 \cap C_2, x_0) \subset R(C_1, x_0) \cap R(C_2, x_0)$ for all $C_1, C_2 \subset E$.

(3b) $R(C_1, x_0) \cup R(C_2, x_0) \subset R(C_1 \cup C_2, x_0)$ for all $C_1, C_2 \subset E$.

(3c) $R(C_1, x_0) \cap E \backslash R(E \backslash C_2, x_0) = \emptyset$ whenever $C_1 \cap C_2 = \emptyset$.

Property (3c) has some application in optimization theory. For example, let $f: E \to \mathbb{R}$ and $C \subset E$ and suppose

$$f(x_0) = \min\{f(x) \mid x \in C\} .$$

If we define $C_1 := C$ and $C_2 := \{x \in E \mid f(x) < f(x_0)\}$, then $C_1 \cap C_2 = \emptyset$ and (3c) gives a necessary condition for optimality. Vlach develops some examples of this approach to optimality conditions in [V1].

(4) $R(C, x_0)$ is convex for all $(C, x_0) \in 2^E \times E$.

(Remember here that $\emptyset$ is vacuously convex.)

Tangent cones which satisfy (4) are useful in that the machinery of convex analysis can be applied in connection with them, as we will see in Chapter 2.

(5) $R(C_1,x_0) \cap R(C_2,x_0) \subset R(C_1 \cap C_2,x_0)$ for all $C_1,C_2 \subset E$ and $x_0 \in E$.

Guaranteeing that the inclusion in (5) holds for some large class of subsets of E is important in optimization theory, as is demonstrated in Chapters 2 and 5 and in [Ro3], [W1].

(6) $R(C_1 \cup C_2,x_0) \subset R(C_1,x_0) \cup R(C_2,x_0)$ for all $C_1,C_2 \subset E$ and $x_0 \in E$.

This property seems to be of less importance than properties (1) through (5) but is satisfied by a few important tangent cones.

Notice that for each of properties (1.1.1), (1.1.2), (1.1.3) and (1) through (6), cl R possesses the property whenever R does. Later in this chapter we will discuss other properties of tangency operators which come into play in subsequent chapters.

Now let us consider several specific tangency operators and compare which of properties (1) through (6) they satisfy. All of these examples are in fact homogeneous tangent cones. The assertions we make now about what properties they possess will be verified later in this chapter. More information on these and other tangent cones can be found in [Bo5], [U1], [Do1], [V2], [Pe1], [Bo6], [Bo8] and their references.

In the definition below, we denote by $N(y)$ the family of neighbourhoods of $y \in E$.

<u>Definition 1.1.2</u>: The <u>contingent</u> <u>cone</u> of C at x_0 is the set

$$K_C(x_0) := \{y \in E \mid \forall\, Y \in N(y),\ \forall\, \lambda > 0, \tag{1.1.4}$$
$$\exists\, t \in (0,\lambda),\ \exists\, y' \in Y \text{ with } x_0 + ty' \in C\} .$$

The contingent cone is standard, isotone, and preserves unions (properties (3b) and (6)). It is always a closed cone but is not always convex. For example, if C is a closed cone with vertex x_0, $K_C(x_0) = C - x_0$ and is thus convex exactly when C is. The contingent cone is the most widely used tangent cone and is in fact often referred to simply as "the tangent cone".

<u>Definition 1.1.3</u>: The <u>pseudotangent</u> <u>cone</u> of C at x_0 is the set

$$P_C(x_0) = \text{cl conv } K_C(x_0) , \tag{1.1.5}$$

where "conv", as usual, is short for "convex hull".

By definition, $P_C(x_0)$ is always closed and convex, and it inherits properties (1) through (3) from $K_C(x_0)$. However, it does not satisfy (6). For example, in $\mathbb{R}^2$

define $C_1 := \{(x,0) \mid x \in \mathbb{R}\}$ and $C_2 := \{(0,y) \mid y \in \mathbb{R}\}$. Then $P_{C_1}((0,0)) = C_1$ and $P_{C_2}((0,0)) = C_2$, but $P_{C_1 \cup C_2}((0,0)) = \mathbb{R}^2$. Like $K_C(x_0)$, $P_C(x_0)$ lacks property 5.

In section 1.6, we will discuss which properties of a tangency operator are inherited by the closure of its convex hull.

<u>Definition 1.1.4</u>: The <u>DM</u> (Dubovitskii-Milyutin) <u>tangent cone</u> is the set

$$(1.1.6) \qquad k_C(x_0) := \{y \in E \mid \forall\, Y \in N(y),\ \exists\, \lambda > 0,\ \forall\, t \in (0,\lambda),\ \exists\, y' \in Y \text{ with } x_0 + ty' \in C\}.$$

Here we borrow the name "DM tangent cone" from [Ma2] and the notation "k" for this cone from [U1]. Like the contingent cone, the DM cone possesses properties (1) through (3), but not (4) or (5). Unlike the contingent cone, it also does not satisfy (6). For example, in $\mathbb{R}$ define $C_1 := \bigcup_{n \in \mathbb{Z}} [2^{2n}, 2^{2n+1}] \cup \{0\}$ and $C_2 := \bigcup_{n \in \mathbb{Z}} [2^{2n-1}, 2^{2n}] \cup \{0\}$. Then $K_{C_1}(0) = K_{C_2}(0) = \mathbb{R}$ and $k_{C_1 \cup C_2}(0) = k_{\mathbb{R}}(0) = \mathbb{R}$, while $k_{C_1}(0) = k_{C_2}(0) = \{0\}$.

In applications, the fact that the DM cone lacks property (6) is more than offset by the fact that it, unlike the contingent cone, is "product-preserving" (see section 1.4). The DM cone may win the award for "most underrated tangent cone".

Definition 1.1.5: The Clarke tangent cone is the set

$$(1.1.7) \qquad T_C(x_0) := \{y \in E \mid \forall\, Y \in N(y)\,,\ \exists\, X \in N(x_0),\ \exists\, \lambda > 0,\ \forall\, x \in X \cap C\,,\ \forall\, t \in (0,\lambda)\,,\ \exists\, y' \in Y \text{ with } x + ty' \in C\}.$$

The Clarke tangent cone is standard and always convex (see [Ro2], [Ro5] or [Do1]). Its main disadvantage is that it is not isotone - e.g., let $C_1 := \{(x,y) \in \mathbb{R}^2 \mid x = 0\}$, $C_2 := \{(x,y) \mid x = 0 \text{ or } y = 0\}$, and $x_0 = (0,0)$. Then $C_1 \subset C_2$, but $T_{C_1}(x_0) = C_1$ while $T_{C_2}(x_0) = \{(0,0)\}$. The Clarke tangent cone does satisfy the inclusion in (5) for a large class of sets, as we shall see in chapter 2.

The tangent cones defined in (1.1.4) through (1.1.7) are the ones we will use most often in subsequent chapters. For a summary of their properties, along with the properties of other tangent cones we will discuss in Chapter 1, see Table 1.9.3.

It is clear from (1.1.4) through (1.1.7) that $T_C(x_0) \subset k_C(x_0) \subset K_C(x_0) \subset P_C(x_0)$ for all $(C,x_0) \in 2^E \times E$. A good example for comparing $K_C(x_0)$, $k_C(x_0)$ and $T_C(x_0)$ can be found in [Ma2]. Here is another instructive example:

<u>Example 1.1.6:</u> For $\alpha \geq 0$, define a class C_α of subsets of $\mathbb{R}^2$ by

$$C_\alpha := \{(x,y) \mid y \geq 0\} \cup \{(x,y) \mid y \geq \alpha x\} .$$

Then C_0 is simply the upper halfplane and $P_{C_0}((0,0)) = K_{C_0}((0,0)) = k_{C_0}((0,0)) = T_{C_0}((0,0)) = C_0$, an illustration of the fact that all four cones are standard. However, if we let α increase,

$$K_{C_\alpha}((0,0)) = k_{C_\alpha}((0,0)) = C_\alpha ,$$

while $P_{C_\alpha}((0,0))$ and $T_{C_\alpha}((0,0))$ go their separate ways. Indeed, $P_{C_\alpha}((0,0)) = \mathbb{R}^2$ for any $\alpha > 0$, but

$$T_{C_\alpha}((0,0)) = \{(x,y) \mid y \geq 0\} \cap \{(x,y) \mid y \geq \alpha x\} ,$$

becoming smaller as α (and C_α) become larger. This example constitutes the starting point for the discussion in section 1.2.

1.2. Impossibility theorems

As the graphic Example 1.1.6 illustrates, $K_C(x_0)$ and $k_C(x_0)$ are isotone but not always convex, while $T_C(x_0)$ is always convex but not isotone. The pseudotangent cone is both convex and isotone, but it has two disadvantages in applications:

(i) It is less tractable analytically than the other cones; in particular, it does not seem possible to base a complete subdifferential calculus upon it.

(ii) While useful in the formulation of optimality conditions in differentiable programming [Gul], it is too large to play a similar role in nonsmooth programming, where a subcone of the contingent cone is often needed (e.g. [Bo5, Proposition 6.1]).

Specifically, suppose $f: \mathbb{R}^n \to \mathbb{R}$ is convex and continuous, $C \subset \mathbb{R}^n$, and $f(x_0) = \min\{f(x) \mid x \in C\}$. Define the <u>subgradient of</u> f <u>at</u> x_0 by

$$(1.2.1) \qquad \partial f(x_0) := \{x^* \in \mathbb{R}^n \mid (x^*, x - x_0) \le f(x) - f(x_0)\} .$$

For $S \subset \mathbb{R}^n$, define the <u>dual cone of</u> S by

$$(1.2.2) \qquad S^+ := \{x^* \in \mathbb{R}^n \mid (x^*, x) \ge 0 \quad \forall\, x \in S\} ,$$

and for $h \in \mathbb{R}^n$, define the <u>upper directional derivative</u> of f at x_0 with respect to h by

$$(1.2.3) \qquad f'(x_0; h) = \lim_{t \downarrow 0} \frac{f(x_0 + th) - f(x_0)}{t} .$$

Then $f'(x_0, h) \ge 0$ for all $h \in K_C(x_0)$. If f is in fact differentiable, then $f'(x_0; \cdot)$ is linear and $f'(x_0; h) \ge 0$ will hold for all $h \in P_C(x_0)$, implying that (e.g. [Gul, Theorem 1])

(1.2.4) $\qquad \partial f(x_0) \cap P_C^{+}(x_0) \neq \emptyset$.

However, if f is not differentiable, (1.2.4) does not necessarily hold. Consider $f: \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x,y) = \max(|x-1|,|y-1|)$ and $C := \{(x,y) \mid x = 0 \text{ or } y = 0\}$. Then $1 = f(0,0) = \min\{f(x,y) \mid (x,y) \in C\}$ and $P_C((0,0)) = \mathbb{R}^2$, but $(0,0) \notin \partial f(0,0)$, so (1.2.4) does not hold.

Given these facts, it is natural to ask if there exists a tangent cone which is isotone and always convex but is also contained in $K_C(x_0)$ for all $(C,x_0) \in 2^E \times E$. Looking at Example 1.1.6, we see that if such a tangent cone $R_C(x_0)$ does exist, $R_{C_\alpha}((0,0)) = C_0$ for all $\alpha \geq 0$. Notice that C_0 is a maximal convex subcone of $K_{C_\alpha}((0,0))$ for each $\alpha \geq 0$. A possible candidate for the tangent cone we are seeking would be some specified maximal convex subcone of the contingent cone, or perhaps the intersection of all maximal convex subcones of the contingent cone.

Example 1.1.6 shows that this latter idea will not work. Defining $M_C(x_0)$ to be the intersection of all maximal convex subcones of $K_C(x_0)$, we see that although $C_\alpha \subset C_\beta$ for $\alpha \leq \beta$,

$$T_{C_\alpha}(x_0) = M_{C_\alpha}(x_0) \supsetneq M_{C_\beta}(x_0) = T_{C_\beta}(x_0)$$

for $\alpha \leq \beta$. Thus $M_C(x_0)$ is not isotone. In fact, Martin and Watkins have shown that for a large class of sets C , $M_C(x_0) = T_C(x_0)$ [Ma2], as we see below.

<u>Definition 1.2.1</u>: The set $C \subset E$ is said to be <u>tangentially</u> <u>regular</u> at $x_0 \in C$ if $K_C(x_0) = T_C(x_0)$.

<u>Theorem 1.2.2</u> ([Ma2, Theorem 4]): Suppose $C \subset E$ can be written as a finite union $C = \cup C_i$, where each C_i is, tangentially regular at $x_0 \in \cap C_i$ and each $K_{C_i}(x_0)$ is a maximal convex subcone of $K_C(x_0)$. Then

$$T_C(x_0) = \cap K_{C_i}(x_0) = M_C(x_0) .$$

<u>Proof</u>: It is always true that

$$\cap T_{C_i}(x_0) \subset T_{\cup C_i}(x_0) ,$$

so
$$\begin{aligned} M_C(x_0) &\subset \cap K_{C_i}(x_0) \\ &= \cap T_{C_i}(x_0) \end{aligned}$$

by tangential regularity of each C_i at x_0

$$\begin{aligned} &\subset T_{\cup C_i}(x_0) \\ &= T_C(x_0) . \end{aligned}$$

Now notice that $K_C(x_0) + T_C(x_0) \subset K_C(x_0)$ ([Ma2, Theorem 1]). It follows that for any maximal convex subcone $\hat{K}$ of $K_C(x_0)$, $\hat{K} + T_C(x_0)$ is a convex subcone of $K_C(x_0)$ containing $\hat{K}$, so $T_C(x_0) \subset \hat{K}$. Hence $T_C(x_0) \subset M_C(x_0)$, completing the proof. □

Remark 1.2.3: (a) The setting in [Ma2] is finite-dimensional, but the argument above, which is the same as that given in [Ma2], is valid in any l.c.s.

(b) Theorem 1.2.2 can help one calculate $T_C(x_0)$ in certain special cases. For example, if C is the union of two convex cones with vertex 0 which intersect only at 0, then $T_C(0) = \{0\}$.

We now demonstrate that there is no "sensible" tangent cone which is isotone and always a convex subcone of the contingent cone. Specifically, a by now familiar example will establish the following result:

Theorem 1.2.4: Suppose E has dimension greater than or equal to 2. Then there is no tangency operator R on E having all of the following properties:

(i) Property (3) (isotonicity);

(ii) Property (4) (convexity);

(iii) $R(C,x_0) \subset K_C(x_0)$ for all $(C,x_0) \in 2^E \times E$;

(iv) $R(C,x_0) \supset C$ for every one-dimensional subspace $C \subset E$ and $x_0 \in C$.

Proof: Define $C_1 := \{(x,y) \in \mathbb{R}^2 | x=0\}$ and $C_2 := \{(x,y) \in \mathbb{R}^2 | y=0\}$, and let $x_0 = (0,0)$. By (iv), $C_1 \subset R(C_1,x_0)$ and $C_2 \subset R(C_2,x_0)$, so by property (3b), which is equivalent to (i),

$$C_1 \cup C_2 \subset R(C_1,x_0) \cup R(C_2,x_0)$$
$$\subset R(C_1 \cup C_2,x_0) \; .$$

Now by (iii),

$$R(C_1 \cup C_2,x_0) \subset K_{C_1 \cup C_2}(x_0)$$
$$= C_1 \cup C_2 \; .$$

Hence $R(C_1 \cup C_2,x_0) = C_1 \cup C_2$, contradicting (ii). □

<u>Corollary 1.2.5</u>: If E has dimension greater than or equal to 2, there is no standard tangency operator on E satisfying (i), (ii) and (iii) of Theorem 1.2.4.

<u>Proof</u>: Standard tangency operators have property (1) and thus satisfy condition (iv) of Theorem 1.2.4. □

If one of the conditions (i) - (iv) is removed, there exist tangent cones satisfying the other three. Here are some examples:

$P_C(x_0)$ satisfies (i), (ii) and (iv).

$T_C(x_0)$ satisfies (ii), (iii) and (iv).

$K_C(x_0)$ satisfies (i), (iii) and (iv).

$$R(C,x_0) = \begin{cases} \{0\} & \text{if } x_0 \in \operatorname{cl} C \\ \emptyset & \text{otherwise} \end{cases}$$

satisfies (i), (ii) and (iii).

Are any other combinations of properties (1) through (6) incompatible for a large class of tangency operators? The answer is yes. Here are some more "impossibility theorems", beginning with the one which seems to be the most significant.

Theorem 1.2.6: If E has dimension ≥ 2, there is no tangency operator R on E satisfying both of the following conditions:

(i) Property (2) (inversion)

(ii) $R(C_1,x_0) \cap R(C_2,x_0) \subset R(C_1 \cap C_2, x_0)$ whenever $x_0 \in C_1 \cap C_2$, C_1 and C_2 are closed, and $\operatorname{ri} R(C_1,x_0) \cap \operatorname{ri} R(C_2,x_0) \neq \emptyset$. (For $S \subset E$, $\operatorname{ri} S$ denotes the relative interior of S, i.e., the interior of S relative to $\operatorname{aff} S$, the affine span of S [Rol].)

Proof: Define $h: \mathbb{R}^2 \to \mathbb{R}$ by $h(x,y) := y-x^2$ and $g: \mathbb{R}^2 \to \mathbb{R}$ by $g(x,y) := y + x^2$, and let $C_1 := h^{-1}(0)$, $C_2 := g^{-1}(0)$, and $x_0 = (0,0)$. The functions h and g satisfy the hypotheses of property (2), so if R has property (2),

$$R(C_1,x_0) = \nabla h(x_0)^{-1}(0) = \{(x,y) \mid y=0\}$$

and

$$R(C_2,x_0) = \nabla g(x_0)^{-1}(0) = \{(x,y) \mid y=0\} .$$

If R also satisfies (ii),

$$R(\{(0,0)\},(0,0)) = R(C_1 \cap C_2, x_0)$$
$$\supset \{(x,y) \mid y = 0\} .$$

Now define $f: \mathbb{R}^2 \to \mathbb{R}^2$ by $f(x,y) = (x,y)$. Then $\nabla f(x_0)^{-1}(0) = \{(0,0)\}$, and by property (2), $R(\{(0,0)\},(0,0)) = \{(0,0)\}$, a contradiction. □

<u>Corollary 1.2.7</u>: If E has dimension ≥ 2, no standard tangency operator on E possesses property (5). In particular, $K_C(x_0)$, $k_C(x_0)$, and $T_C(x_0)$ do not have property (5). Conversely, no tangency operator with property (5) has property (2). In particular, the <u>hypertangent</u> cone ([Ro4], [Ro5])

$$H_C(x_0) := \{y \in E \mid \exists\ X \in N(x_0),\ \exists\ \lambda > 0,\ \forall\ x' \in X \cap C,$$
$$\forall\ t \in (0,\lambda),\ x' + ty \in C\}$$

and the <u>internal tangent cone</u> ([Pe2], [Ma2])

$$L_C(x_0) := \{y \in E \mid \exists\ Y \in N(y),\ \exists\ \lambda > 0,\ \forall\ t \in (0,\lambda),$$
$$\forall\ y' \in Y,\ x_0 + ty' \in C\}$$

do not have property (2).

Theorem 1.2.6 shows that it is necessary to look for some "constraint qualification" under which the important inclusion of property (5) holds - and in fact, that a condition stronger than

$$\text{ri } R(C_1,x_0) \cap \text{ri } R(C_2,x_0) \neq \emptyset$$

is required. We will return to these considerations in Chapter 2.

Theorem 1.2.8: If E has dimension ≥ 2, there is no tangency operator R on E which satisfies all of the following conditions:

(i) $R(C_1 \cup C_2, x_0) = R(C_1,x_0) \cup R(C_2,x_0)$ for all one-dimensional subspaces $C_1, C_2 \subset E$.

(ii) Property (4) (convexity).

(iii) $C = R(C,x_0)$, whenever $C \subset E$ is a one-dimensional subspace and $x_0 \in C$.

Proof: Define C_1, C_2 and x_0 as in the proof of Theorem 1.2.4. By (i) and (iii), $R(C_1 \cup C_2, x_0) = C_1 \cup C_2$, contradicting (ii). $\square$

Again, if any of the conditions in Theorem 1.2.8 is removed, there are tangent cones which satisfy the remaining ones:

$K_C(x_0)$ satisfies (i) and (iii).

$T_C(x_0)$ satisfies (ii) and (iii).

$R(C,x_0) = 0$ for all $(C,x_0) \in 2^E \times E$ satisfies (i) and (ii).

Corollary 1.2.9: If E has dimension ≥ 2, there is no standard tangency operator R on E satisfying (i) and (ii) of Theorem 1.2.8.

Proof: Any standard tangency operator satisfies (iii) of Theorem 1.2.8. □

Theorem 1.2.10: If E has dimension ≥ 1, there is no homogeneous tangency operator R on E satisfying all of the following conditions:

(i) Property (5) (intersection);

(ii) Property (6) (union);

(iii) $R(C,x_0) = C$ if $C \subset E$ is a subspace of dimension ≤ 1 and $x_0 \in C$;

(iv) Property (1.1.2) (cone valuedness).

Proof: Let $C := \{0\} \cup \bigcup_{n \in \mathbb{Z}} (2^{2n}, 2^{2n+1}]$, and define $C_1 := C \cup (-C)$ and $C_2 := 2C_1$. Let $x_0 := 0$. Then $R(C_2,0) = R(2C_1,0) = 2R(C_1,0) = R(C_1,0)$ since R is homogeneous and cone-valued. If $R(C_1,0) = R(C_2,0) = \mathbb{R}$, then by (i), $R(C_1 \cap C_2,0) = \mathbb{R}$. However $C_1 \cap C_2 = \{0\}$, so $R(C_1 \cap C_2,0) = \{0\}$, a contradiction. We reach a similar contradiction if $R(C_1,0) = R(C_2,0) = \pm\mathbb{R}_+$. If $R(C_1,0) = R(C_2,0) = \{0\}$, $R(C_1 \cup C_2,0) = \{0\}$ by (ii), again contradicting (iii) since $C_1 \cup C_2 = \mathbb{R}$. □

Corollary 1.2.11: If E has dimension ≥ 1, there is no standard homogeneous tangent cone with both properties (5) and (6).

Here are examples of homogeneous tangency operators satisfying three of the conditions of Theorem 1.2.10:

$R(C,x_0) = \{0\}$ for all $(C,x_0) \in 2^E \times E$ satisfies (i), (ii) and (iv).

$K_C(x_0)$ satisfies (ii), (iii) and (iv).

$H_C(x_0)$ satisfies (i), (iii) and (iv).

$R(C,x_0) = C$ satisfies (i), (ii) and (iii).

In applications, one is rarely interested in sets like those in the proof of Theorem 1.2.10, and it is more important to identify classes of sets for which the inclusions in properties (5) and (6) hold. For example, if C_1 and C_2 are <u>hypertangentially regular</u> at x_0 - i.e., $K_{C_i}(x_0) = H_{C_i}(x_0)$ for $i = 1,2$ - then

$$P_{C_i}(x_0) = K_{C_i}(x_0) = k_{C_i}(x_0) = T_{C_i}(x_0) = H_{C_i}(x_0)$$

and the inclusions in (5) and (6) hold for all five of these tangent cones. The cones P, K, k and T also satisfy (5) and (6) for sets C_1, C_2 which are tangentially regular at x_0 and for which $T_{C_1}(x_0) \cap \operatorname{int} T_{C_2}(x_0) \neq \emptyset$. (If E is finite-dimensional, the intersection condition may be weakened to $T_{C_1}(x_0) - T_{C_2}(x_0) = E$, as we will see in Chapter 2.)

<u>Theorem 1.2.12:</u> If E has dimension ≥ 1, there is no tangency operator R on E satisfying all of the following conditions:

(i) Property (5) (intersection)

(ii) $R(\{0\},0) = \{0\}$

(iii) $R(C,x_0) \supset T_C(x_0)$ for all $(C,x_0) \in 2^E \times E$.

Proof: Let $C_1 := Q$, the set of rational numbers, and let $C_2 := (\mathbb{R}\setminus Q) \cup \{0\}$ and $x_0 := 0$. Since both C_1 and C_2 are dense in $\mathbb{R}$, it is not hard to see that $T(C_1,0) = T(C_2,0) = \mathbb{R}$, and so (iii) and (i) imply that $R(C_1 \cap C_2,0) = \mathbb{R}$. But $C_1 \cap C_2 = \{0\}$, so this contradicts (ii). □

Theorem 1.2.12 shows that a tangency operator has to be fairly "small" in order to have property (5). Here are examples of tangent cones which satisfy two of the conditions of Theorem 1.2.12:

$H_C(x_0)$ satisfies (i) (Corollary 1.3.25) and (ii).

$T_C(x_0)$ satisfies (ii) and (iii).

$R(C,x_0) = E$ for all $(C,x_0) \in 2^E \times E$ satisfies (i) and (iii).

Corollary 1.2.13: If E has dimension ≥ 1, there is no standard tangency operator R on E satisfying (i) and (iii) of Theorem 1.2.12.

Theorem 1.2.14: If E has dimension ≥ 1, there is no homogeneous tangency operator R on E satisfying all of the following conditions:

(i) Property (6) (union)

(ii) $R(C,x_0) = C$ for all one-dimensional subspaces $C \subset E$ and all $x_0 \in C$.

(iii) $R(C,x_0) \subset k_C(x_0)$, for all $(C,x_0) \in 2^E \times E$.

Proof: Take C_1, C_2 and x_0 as in the proof of Theorem 1.2.10. Since $k_{C_1}(0) = k_{C_2}(0) = \{0\}$, (iii) and (i) imply that $R(C_1 \cup C_2, 0) = \{0\}$. However, $C_1 \cup C_2 = \mathbb{R}$, so this contradicts (ii). □

Theorem 1.2.14 shows that a homogeneous tangency operator has to be fairly "large" to possess property (6).

Below are examples of homogeneous tangent cones satisfying two of the conditions of Theorem 1.2.14.

$K_C(x_0)$ satisfies (i) and (ii).

$k_C(x_0)$ satisfies (ii) and (iii).

$$R(C,x_0) = \begin{cases} \{0\} & \text{if } x_0 \in \text{cl } C \\ \emptyset & \text{otherwise} \end{cases}$$

satisfies (i) and (iii).

Corollary 1.2.15: If E has dimension ≥ 1, there is no standard homogeneous tangency operator on E satisfying conditions (i) and (iii) of Theorem 1.2.14.

Remark 1.2.16: In the statements of the impossibility theorems of this section, we have implicitly assumed that $R(C,x_0)$ is unaffected by any increase in the dimension of the space E in which C is considered to lie. Almost all important tangent cones have this property.

1.3. The quantificational tangent cones

Most of the specific tangent cones we have so far considered have definitions containing a number of "$\forall$'s" and "$\exists$'s". With the goal of fitting these various tangent cones into one general framework, we make the following definition:

<u>Definition 1.3.1</u>: A <u>quantificational tangent cone</u> (or "q-cone") is a tangency operator R (on a l.c.s. E) of the form

$$
\begin{aligned}
(1.3.1) \qquad R(C,x_0) = \{y \in E \mid\ & * \ Y \in \mathcal{N}(y),\ \#\ X \in \mathcal{N}(x_0),\ \$\ \lambda > 0, \\
& \exists\ W \in \mathcal{K}(X),\ \exists\ Z \in \mathcal{M}_C(Y),\ \#'\ x \in C \cap W, \\
& \$'\ t \in (0,\lambda),\ *'\ y' \in Z,\ x + ty' \in C\}
\end{aligned}
$$

where

(i) $\mathcal{K}(X)$ and $\mathcal{M}_C(Y)$ are given classes of nonempty subsets of X and Y, respectively.

(ii) $*, *', \#, \#', \$, \$' \in \{\forall, \exists\}$.

(iii) $* \neq *'$, $\# \neq \#'$, $\$ \neq \$'$.

If $\mathcal{M}_C(\cdot)$ is independent of C or if there is no risk of confusion, we will suppress the subscript and simply write $\mathcal{M}(\cdot)$.

We will especially concerned with these specific cases:

(1.3.2) (a) $K(X)$ (or resp. $M(Y)$) consists only of the neighbourhood $X(Y)$ itself. We will denote this case by $K(X) = X$ ($M(Y) = Y$) .

(b) $K(X)$ ($M(Y)$) consists only of the set $\{x_0\}$ ($\{y\}$) . We will denote this case by $K(X) = x_0$ ($M(Y) = y$) .

(c) $M(Y)$ is the class of all nonempty compact subsets of Y .

(d) $M_C(Y)$ consists only of the set $Y \cap (\text{aff } C - x_0)$. We will denote this case by $M_C(Y) = Y \cap (\text{aff } C - x_0)$.

Most of the tangent cones we have defined so far are special cases of (1.3.1). For example, to obtain $K_C(x_0)$, let $K(X) = x_0$, $M(Y) = Y$, $* = \forall$, $\$ = \forall$, and let $\#$ be either $\forall$ or $\exists$. To obtain $k_C(x_0)$, let $K(X) = x_0$, $M(Y) = Y$, $* = \forall$, $\$ = \exists$, and let $\#$ be either $\forall$ or $\exists$. To obtain $T_C(x_0)$, let $K(X) = X$, $M(Y) = Y$, $* = \forall$, $\$ = \exists$, and $\# = \exists$.

Some further comments on the motivation behind the various parts of Definition 1.3.1 are in order. The reason the quantifications are "paired" (assumption (iii) is that our intention is to study <u>local</u> approximations (see Theorem 1.4.6). If $*$ and $*'$ or $\#$ and $\#'$ or $\$$ and $\$'$ were allowed to coincide, the resulting q-cones would be determined by all of C , not just the part of C near x_0 . The family of sets $K(\cdot)$ is not assumed to depend on C

since dependence on C is already built into the definition with "$\#' \ x \in C \cap W$".

It would be possible to extend the definition by allowing the quantifications on W and Z to be "$\forall$" as well as "$\exists$". We will not pursue here the possibilities created by this extension; the definition as it stands is sufficient to encompass all previously-defined major special cases. Notice that in cases (a), (b) and (d) of (1.3.2), the quantification on W and Z does not matter anyway.

It is important to observe that there exist a largest and a smallest q-cone. Indeed, the <u>interiorly tangent cone</u> ([Th1], [Th2])

$$(1.3.3) \qquad \begin{aligned} T_C(x_0) := \{y \in E \mid\ & \exists\, Y \in N(y),\ \exists\, X \in N(x_0),\\ & \exists\, \lambda > 0,\ \forall\, x \in X \cap C,\ \forall\, t \in (0,\lambda),\ \forall\, y' \in Y,\\ & x + ty' \in C\} \end{aligned}$$

is contained in every q-cone, and the cone

$$(1.3.4) \qquad \begin{aligned} D_C(x_0) := \{y \in E \mid\ & \forall\, Y \in N(y),\ \forall\, X \in N(x_0),\\ & \forall\, \lambda > 0,\ \exists\, x \in X \cap C,\ \exists\, t \in (0,\lambda),\ \exists\, y' \in Y,\\ & x + ty' \in C\} \end{aligned}$$

contains every q-cone. To see these facts, notice that the most restrictive quantification on the variable x (respectively, y) in (1.3.1) is to set $\# := \exists$ ($* := \exists$) and $K(X) := X$ ($M(Y) := Y$), while the least restrictive quantification on the variable $x(y)$ is to set $\# := \forall$ ($* := \forall$)

and $K(X) := X$ ($M(Y) := Y$) . We will use this sort of reasoning several times in this chapter to assert that some q-cone is the largest or smallest in a given class of q-cones.

With regard to cases (a) through (d) of (1.3.2), notice that $M(Y) = Y$ and case (c) coincide in finite dimensions. In case (d), $M_C^{\circ}(\cdot)$ does not depend on x_0 , since aff $C - x_0$ = aff $C - x$ whenever x_0 and x are elements of C . Also observe that if $* = \forall$, $M(Y) = Y$ and case (d) coincide. We have already noted that if $K(X) = x_0$ (respectively, $M(Y) = y$) , then the same tangent cone is obtained for $\# = \forall$ or $\exists$ ($* = \forall$ or $\exists$) .

One main reason for concentrating on this particular class of tangent cones is that the quantificational tangent cones are intrinsically defined, so that they are often easier to calculate than other tangent cones - e.g., $P_C(x_0)$ or $M_C(x_0)$. They are also especially well-suited for the purpose of defining directional derivatives and subgradients, as we will see in section 1.5.

In this and subsequent sections we undertake a detailed study of q-cones, with several purposes in mind:

(i) In doing so, we may unearth a new tangent cone of some value.

(ii) We can hope to learn which properties of a q-cone result from which parts of its definition.

(iii) We may find new ways of characterizing familiar tangent cones.

Some earlier experimentation with quantification in tangent cone definitions can be found in Vlach's papers ([V2], [V3] and [V4]).

In many proofs in this chapter, we will use the symbols * , # and \$ as in Definition 1.3.1 in order to avoid having to do a case-by-case analysis of specific tangent cones. This will also be a way of pinpointing which quantifications are important and which do not matter in determining whether a given tangent cone has a given property.

We observe first that for any q-cone R , $R(C,x_0) = \emptyset$ whenever $x_0 \notin \mathrm{cl}\, C$. In addition, all "reasonable" q-cones, including those in cases (a) through (d) of (1.3.2), actually are cone-valued.

<u>Proposition 1.3.2</u>: Suppose R is a q-cone on E such that $\alpha M(Y) = M(\alpha Y)$ for all $\alpha > 0$. Then R satisfies (1.1.2) (cone-valuedness).

<u>Proof</u>: Let $y \in R(C,x_0)$ and $\alpha > 0$ be given. Then $* \frac{1}{\alpha} Y \in \mathcal{N}(y)$, $\# X \in \mathcal{N}(x_0)$, $\$ \alpha\lambda > 0$, there exist $Z \in M(\frac{1}{\alpha} Y)$ and $W \in K(X)$, $\#' x \in C \cap W$, $\$' t \in (0,\alpha\lambda)$, $*' y' \in Z$, $x + ty' \in C$. Now $\alpha y' \in \alpha Z$, and since $\alpha M(Y) = M(\alpha Y)$, $\alpha Z \in M(Y)$. Therefore $* Y \in \mathcal{N}(\alpha y)$, $\# X \in \mathcal{N}(x_0)$, $\$ \lambda > 0$, there exist $W \in K(X)$ and $\alpha Z \in M(Y)$, $\#' x \in C \cap W$, $\$' \frac{t}{\alpha} \in (0,\lambda)$, $*' \alpha y' \in \alpha Z$, $x + (\frac{t}{\alpha})\alpha y' = x + ty' \in C$. Thus $\alpha y \in R(C,x_0)$ and $R(C,x_0)$ is a cone. □

It is also easy to show that many q-cones are homogeneous.

Proposition 1.3.3: Suppose R is a q-cone on E such that $\mathcal{M}_C(\alpha Y) = \alpha\mathcal{M}_C(Y) = \mathcal{M}_{\alpha C}(\alpha Y)$ and $\alpha\mathcal{K}(X) = \mathcal{K}(\alpha X)$ for all $\alpha > 0$. Then R is a homogeneous tangent cone.

Proof: The fact that R is a cone follows from Proposition 1.3.2. Let $y \in R(C,x_0)$ and $\alpha > 0$ be given. Then $* \; Y \in \mathcal{N}(y)$, $\# \; X \in \mathcal{N}(x_0)$, $\$ \; \lambda > 0$, there exist $W \in \mathcal{K}(X)$ and $Z \in \mathcal{M}_C(Y)$, $\#' \; x \in C \cap W$, $\$' \; t \in (0,\lambda)$, $*' \; y' \in Z$, $x + ty' \in C$. By hypothesis, $\alpha W \in \mathcal{K}(\alpha X)$ and $\alpha Z \in \mathcal{M}_{\alpha C}(\alpha Y)$. Thus $* \; \alpha Y \in \mathcal{N}(\alpha y)$, $\# \; \alpha X \in \mathcal{N}(\alpha x_0)$, $\$ \; \lambda > 0$, $\#' \; \alpha x \in \alpha C \cap \alpha W$, $\$' \; t \in (0,\lambda)$, $*' \; \alpha y' \in \alpha Z$,

$$\alpha x + t\alpha y' = \alpha(x+ty') \in \alpha C .$$

Hence $\alpha y \in R(\alpha C,\alpha x_0)$ and $\alpha R(C,x_0) \subset R(\alpha C,\alpha x_0)$. It follows that

$$\frac{1}{\alpha} R(\alpha C,\alpha x_0) \subseteq R(\frac{1}{\alpha}\alpha C, \frac{1}{\alpha}\alpha x_0) = R(C,x_0) ,$$

and so $R(\alpha C,\alpha x_0) \subseteq \alpha R(C,x_0)$ also. Therefore R is homogeneous. □

When $* = \forall$ and $\mathcal{M}(Y)$ is as in case (a) or (c), it is not hard to show that $R(C,x_0)$ is a closed set. (A specific example is [Bo8, Proposition 2.1].) We generalize this fact below.

<u>Theorem 1.3.4:</u> Suppose that $* := \forall$ in (I.3.1) and that $M(\cdot)$ is such that if $Y_1 \in N(y_1)$, $Y_2 \in N(y_2)$, $Y_1 \subset Y_2$ and $Z_1 \in M(Y_1)$, there exists $Z_2 \supset Z_1$ with $Z_2 \in M(Y_2)$. Then $R(C,x_0)$ is a closed set for all $(C,x_0) \in 2^E \times E$.

<u>Proof:</u> Let $y \in \mathrm{cl}\ R(C,x_0)$ be given. Let $Y = y + V \in N(y)$, where $V \in N(0)$, and let $U \in N(0)$ be such that $U + U \subset V$. Then for some $u \in U$, $\bar{y} := y + u \in R(C,x_0)$. Hence $\# X \in N(x_0)$, $\$ \lambda > 0$; there exist $W \in K(X)$ and $Z_1 \in M(\bar{y} + U)$, $\#' x \in C \cap W$, $\$' t \in (0,\lambda)$, there exists $y' \in Z_1$ with $x + ty' \in C$. Now $y' \in y + V$ and there exists $Z_2 \in M(y+V)$ with $Z_2 \supset Z_1$, so y must actually be in $R(C,x_0)$. Therefore $R(C,x_0)$ is a closed set. □

<u>Corollary 1.3.5:</u> The tangent cones $K_C(x_0)$, $k_C(x_0)$, and $T_C(x_0)$ are always closed sets.

<u>Example 1.3.6:</u> Let $M(Y)$ be as in case (c), and define

$F_C(x_0) := \{y \in E \mid \forall Y \in N(y), \exists X \in N(x_0), \exists \lambda > 0,$
$\exists Z \in M(Y), \forall x \in X \cap C, \forall t \in (0,\lambda),$
$\exists y' \in Z$ with $x + ty' \in C\}$.

This is a q-cone with $* = \forall$, $\# = \exists$, $\$ = \exists$, and $K(X) = X$. Since case (c) satisfies the hypothesis of Theorem 1.3.4 with $Z_2 = Z_1$, $F_C(x_0)$ is always a closed set (see also [Bo8, Proposition 2.1]). This tangent cone is thoroughly discussed in [Bo8].

When $* = \exists$ and $M(Y)$ is as in cases (a) or (c), $R(C,x_0)$ is an open set. Our next result generalizes this fact.

Theorem 1.3.7: Suppose that $* := \exists$ in (1.3.1) and that $M(\cdot)$ is such that if $Y_1 \in N(y_1)$, $Y_2 \in N(y_2)$, $Y_2 \supset Y_1$ and $Z_2 \in M(Y_2)$, there exists $Z_1 \in M(Y_1)$ with $Z_1 \subset Z_2$. Then $R(C,x_0)$ is an open set for all $(C,x_0) \in 2^E \times E$.

Proof: Let $y \in R(C,x_0)$. There exists $Y = y + V \in N(y)$, where $V \in N(0)$, $\# X \in N(x_0)$, $\$ \lambda > 0$, there exist $W \in K(X)$ and $Z \in M(Y)$, $\#' x \in C \cap W$, $\$' t \in (0,\lambda)$ with

$$x + tZ \subset C .$$

Now let $U \in N(0)$ be such that $U + U \subset V$, and let $\bar{y} := y + \bar{u} \in y + U$ for a given $\bar{u} \in U$. By hypotheses, there exists $Z' \subset Z$ with $Z' \in M(\bar{y} + U)$. Thus $\# X \in N(x_0)$, $\$ \lambda > 0$, there exists $W \in K(X)$, $\#' x \in C \cap W$, $\$' t \in (0,\lambda)$, for all $y' \in Z'$,

$$\begin{aligned} x + ty' &\in x + tZ' \\ &\subset x + tZ \subset C . \end{aligned}$$

Therefore $y + U \subset R(C,x_0)$ and so $R(C,x_0)$ is an open set. □

In the remainder of this section, we analyze which classes of q-cones satisfy various of the properties (1)-(6)

given in section 1.1. We begin with property (1), commenting that a tangent cone must have property (1) in order to be widely applicable. On this basis we can rule out much further study of q-cones with $\# := \forall$ and $K(X) := X$, as the following example illustrates:

Example 1.3.8: Define the q-cone

$G_C(x_0) := \{y \in E \mid \exists\, Y \in N(y),\ \forall\, X \in N(x_0),\ \exists\, \lambda > 0,$
$\exists\, x \in X \cap C,\ \forall\, t \in (0,\lambda),\ \forall\, y' \in Y,\ x + ty' \in C\}$,

and consider the case $E := \mathbb{R}$, $C := \mathbb{R}_+$, $x_0 := 0$. Let $\varepsilon > 0$, $\delta > 0$ be given, let $y \in \mathbb{R}$, and define $X := [-\delta,\delta]$ and $Y := [y-\varepsilon, y+\varepsilon]$. Then if $y \geq \varepsilon$, $\delta + (0,r)Y \subset C$ for any $r > 0$, and if $y \leq \varepsilon$, $\delta + (0, \frac{-\delta}{y-\varepsilon})Y \subset C$. Letting $x := \delta$ in the definition of $G_C(x_0)$, we conclude that $G_C(x_0) = \mathbb{R}$, and hence that $G_C(x_0)$ does not have property (1). (In fact, it can be shown that in general $G_C(x_0)$ can take on only two possible values – $\emptyset$ and the whole space E.) From this example, we deduce the following result:

Proposition 1.3.9: Suppose R is a q-cone with $\# := \forall$ and $K(X) := X$. Then R does not have property (1).

Proof: Let E, C and x_0 be as in Example 1.3.8. The smallest q-cone with $\# = \forall$ and $K(X) = X$ is $G_C(x_0)$; so for any such q-cone R, $R(C,x_0) \supset G_C(x_0)$ for all $(C,x_0) \in 2^E \times E$. For this example, then, we have $R(C,x_0) = \mathbb{R}$. Thus $R(C,x_0)$ does not have property (1). □

Example 1.3.10: Let $E := \mathbb{R}^2$, $x_0 := (0,0)$, and $C := \{(x,y) \mid x = 0\}$. Then $G_C(x_0) = \emptyset$. This is a special case of the following fact.

Proposition 1.3.11: Suppose $C \subset E$ with $\operatorname{int} C = \emptyset$, and suppose R is a q-cone on E with $* := \exists$ and $M(Y) = Y$. Then $R(C,x_0) = \emptyset$ for all $x_0 \in E$.

Proof: If $y \in R(C,x_0)$, there exists $x \in C$, $t > 0$, and $Y \in N(y)$ with $x + tY \subset C$. The set $x + tY$ has nonempty interior, so if $\operatorname{int} C = \emptyset$, $R(C,x_0) = \emptyset$. □

By Proposition 1.3.11, the internal tangent cone $L_C(x_0)$ and the interiorly tangent cone $I_C(x_0)$ do not have property (1), and similarly, neither does the cone

$$\ell_C(x_0) := \{y \in E \mid \exists\, Y \subset N(y),\ \forall\, \lambda > 0,\ \exists\, t \in (0,\lambda),\ \forall\, y' \in Y,\ x_0 + ty' \in C\}.$$

What will happen, though, if we use $M_C(Y) = Y \cap (\operatorname{aff} C - x_0)$ instead of $M(Y) = Y$ in these tangent cones?

Definition 1.3.12: Suppose R is a q-cone with $* := \exists$ and $M(Y) = Y$. Define the relative q-cone by

$$R^{rel}(C,x_0) := \{y \in E \mid \exists\, Y \in N(y),\ \#\, X \in N(x_0),\ \$\, \lambda > 0,\ \exists\, W \in K(X),\ \#'\, x \in C \cap W,\ \$'\, t \in (0,\lambda),\ \forall\, y' \in Y \cap (\operatorname{aff} C - x_0),\ x + ty' \subset C\}. \tag{1.3.5}$$

<u>Proposition 1.3.13</u>: If $E = \mathbb{R}^n$, the q-cones $H_C(x_0)$, $I_C^{rel}(x_0)$, $L_C^{rel}(x_0)$ and $\ell_C^{rel}(x_0)$ have property (1).

<u>Proof</u>: It is well known (see for example [Ro5, Chapter 2]) that the contingent cone and Clarke tangent cone have property (1) in any l.c.s. E. In addition, it is shown in [Ro2] that $\operatorname{int} T_C(x_0) = I_C(x_0)$ for a closed set $C \subset \mathbb{R}^n$, if $\operatorname{int} T_C(x_0) \neq \emptyset$. Keeping these facts in mind, suppose $C \subset \mathbb{R}^n$ is closed and convex and $x_0 \in C$. The convex set $T_C(x_0)$ has nonempty relative interior [Ro1, Section 6], and it follows from the result mentioned above that

$$\operatorname{ri} T_C(x_0) = I_C^{rel}(x_0) .$$

Since $I_C^{rel}(x_0) \subset H_C(x_0) \subset T_C(x_0)$, we conclude that

$$\begin{aligned} T_C(x_0) &= \operatorname{cl}(\operatorname{ri} T_C(x_0)) \subset \operatorname{cl} I_C^{rel}(x_0) \\ &\subset \operatorname{cl} H_C(x_0) \subset T_C(x_0) , \end{aligned}$$

and so $I_C^{rel}(x_0)$ and $H_C(x_0)$ have property (1). Similarly,

$$\begin{aligned} \operatorname{ri} T_C(x_0) &= I_C^{rel}(x_0) \\ &\subset L_C^{rel}(x_0) \\ &= \ell_C^{rel}(x_0) \subset K_C(x_0) , \end{aligned}$$

so $L_C^{rel}(x_0)$ and $\ell_C^{rel}(x_0)$ also have property (1). □

In infinite dimensions, examples in [Bo7] and [Bo8] show that $F_C(x_0)$ and $H_C(x_0)$, and thus $I_C^{rel}(x_0)$, do not have property (1).

Since $T_C(x_0)$ and $K_C(x_0)$ have property (1), so do all tangent cones lying between them - e.g. $k_C(x_0)$. The pseudotangent cone inherits property (1) from $K_C(x_0)$. The <u>cone of feasible directions</u>

$$(1.3.6) \qquad E_C(x_0) := \{y \in E \mid \exists\, \lambda > 0,\ \forall\, t \in (0,\lambda),\ x_0+ty \in C\}$$

and the q-cone

$$(1.3.7) \qquad A_C(x_0) := \{y \in E \mid \forall\, \lambda > 0,\ \exists\, t \in (0,\lambda),\ x_0+ty \in C\}$$

also have property (1). (For more information and references on these cones, see [V1], [V2] and [V3].) This follows from the observation that when C is convex,

$$(1.3.8) \qquad A_C(x_0) = E_C(x_0) = \{y \in E \mid \exists\, t > 0,\ x_0 + ty \in C\} = \mathrm{cone}(C-x_0)\ .$$

We turn now to property (2). First, we make a definition:

<u>Definition 1.3.14</u> [Ro3]: Let E, F be l.c.s. The function $h: E \to F$ is <u>strictly differentiable</u> at $x_0 \in E$ if there is a continuous linear mapping $\nabla h(x_0): E \to F$ such that

$$\lim_{\substack{x \to x_0 \\ y' \to y \\ t \downarrow 0}} \frac{h(x+ty')-h(x)}{t} = \nabla h(x_0)y \tag{1.3.9}$$

for all $y \in E$.

Notice in particular that a function which is C^1 near x_0 is strictly differentiable at x_0.

One inclusion of property (2) is easy to establish for any q-cone.

Proposition 1.3.15: Suppose R is a q-cone on E and $h: E \to F$ is strictly differentiable at $x_0 \in h^{-1}(0)$. Then

$$R(h^{-1}(0),x_0) \subset \nabla h(x_0)^{-1}(0) . \tag{1.3.10}$$

Proof: Let $y \in D_{h^{-1}(0)}(x_0)$ be given (see 1.3.4) for the definition of $D_C(x_0)$). Then for all $Y \in N(y)$, $X \in N(x_0)$ and $\lambda > 0$, there exist $x \in X \cap h^{-1}(0)$, $t \in (0,\lambda)$, and $y' \in Y$ with $h(x+ty') = 0$ and thus also $\frac{h(x+ty')-h(x)}{t} = 0$. Since h is strictly differentiable at x_0, $\nabla h(x_0)y = 0$, and so $y \in \nabla h(x_0)^{-1}(0)$. Hence $D_{h^{-1}(0)}(x_0) \subset \nabla h(x_0)^{-1}(0)$, and it follows that $R(h^{-1}(0),x_0) \subset D_{h^{-1}(0)}(x_0) \subset \nabla h(x_0)^{-1}(0)$ for any q-cone R. □

The reverse inclusion holds much less often. We know by Theorem 1.2.6 that it will not hold for any q-cone with property (5). More generally, it will not hold if $* := \exists$ in any of cases (a) - (d).

Example 1.3.16: Consider $h: \mathbb{R}^2 \to \mathbb{R}$ defined by $h(x,y) := y - x^2$ and $x_0 := (0,0)$. Then $\nabla h(x_0)^{-1}(0) = \{(x,y) \mid y = 0\}$, while $A_{h^{-1}(0)}(x_0) = H_{h^{-1}(0)}(x_0) = E_{h^{-1}(0)}(x_0) = \{(0,0)\}$ and $L^{rel}_{h^{-1}(0)}(x_0) = I^{rel}_{h^{-1}(0)}(x_0) = \emptyset$.

It is well known, however (see for example [Bo5, Theorem 4.1]), that the reverse inclusion holds for the Clarke tangent cone if $\nabla h(x_0)$ is of full rank. We can thus deduce the following result:

Proposition 1.3.17: Suppose R is a q-cone with $* := \forall$ and $M(Y) = Y$, and suppose $h: \mathbb{R}^n \to \mathbb{R}^m$ is strictly differentiable at $x_0 \in h^{-1}(0)$ with $\nabla h(x_0)$ of rank m. Then

$$(1.3.11) \qquad R(h^{-1}(0), x_0) = \nabla h(x_0)^{-1}(0) .$$

Proof: By Proposition 1.3.15, $R(h^{-1}(0),x_0) \subset \nabla h(x_0)^{-1}(0)$; and since $\nabla h(x_0)$ is of full rank, $\nabla h(x_0)^{-1}(0) \subset T_{h^{-1}(0)}(x_0)$. Now for any q-cone with $* := \forall$ and $M(Y) = Y$, it is not hard to see that

$$T_{h^{-1}(0)}(x_0) \subset R(h^{-1}(0), x_0) .$$

Therefore (1.3.11) holds. $\square$

<u>Corollary 1.3.18</u>: The q-cones $T_C(x_0)$, $K_C(x_0)$, $k_C(x_0)$, $F_C(x_0)$ and $D_C(x_0)$ all possess property (2).

<u>Proof</u>: In finite dimensions $F_C(x_0) = T_C(x_0)$, so $F_C(x_0)$ automatically has property (2). The cones $K_C(x_0)$, $k_C(x_0)$ and $D_C(x_0)$ all have $* := \forall$ and $M(Y) = Y$. □

We next consider property (3) - isotonicity with respect to set inclusion.

<u>Theorem 1.3.19</u>: Suppose R is q-cone with $\# := \forall$. Then $R(C_1,x_0) \subset R(C_2,x_0)$ whenever $C_1 \subset C_2 \subset E$ and $x_0 \in E$.

<u>Proof</u>: Suppose $x_0 \in E$, $C_1 \subset C_2 \subset E$ and $y \in R(C_1,x_0)$. Then $* \, Y \in N(y)$, for all $X \in N(x_0)$, $\$ \, \lambda > 0$, there exist $W \in K(X)$, $Z \in M(Y)$ and $x \in C_1 \cap W \subset C_2 \cap W$, $\$' \, t \in (0,\lambda)$, $*' \, y' \in Z$, $x + ty' \in C_1 \subset C_2$. Hence $y \in R(C_2,x_0)$ and $R(C_1,x_0) \subset R(C_2,x_0)$. □

<u>Corollary 1.3.20</u>: The q-cones $K_C(x_0)$, $k_C(x_0)$, $\ell_C(x_0)$, $L_C(x_0)$, $\ell_C^{rel}(x_0)$, $L_C^{rel}(x_0)$, $A_C(x_0)$, $E_C(x_0)$, $G_C(x_0)$, $G_C^{rel}(x_0)$ and $D_C(x_0)$ have property (3).

Unless $K(X) = x_0$, R will not in general be isotone if $\# := \exists$. Such an R would be isotone if in (1.3.1), "$\#' \, x \in C \cap W$" were replaced by "$\#' \, x \in W$". However, such a definition would not give useful tangential approximants. For example, define

$R(C,x_0) := \{y \in E \mid \forall\, Y \in N(y),\ \exists\, X \in N(x),\ \exists\, \lambda > 0,\ \forall\, x \in X,\ \forall\, t \in (0,\lambda),\ \exists\, y' \in Y \text{ with } x + ty' \in C\}$. It is easy to prove that $R(C,x_0)$ is isotone and convex, so it might seem at first glance to be a potentially useful tangent cone. But $R(C,x_0) \subset T(C,x_0) \subset K(C,x_0)$, so by Theorem 1.2.4, R is not standard. For example, if $E := \mathbb{R}^2$, $C := \mathbb{R}^2_+$, and $x_0 := (0,0)$; then $R(C,x_0) = \emptyset$. This example demonstrates that we can exclude such definitions of tangent cone from further consideration.

Of the q-cones we have defined so far, three are always convex - $T_C(x_0)$, $H_C(x_0)$ and $F_C(x_0)$ (for proofs, see [Ro5, Chapter 2], [Dol] and [Bo8]). We now give a generalization of these facts.

<u>Theorem 1.3.21</u>: Suppose R is a q-cone on E with $* := \forall$, $\# := \exists$ and $\$:= \exists$ and assume the following conditions hold:

(1.3.12) $\alpha M(Y) = M(\alpha Y)$ for all $\alpha > 0$.

(1.3.13) If $Y_1 \in N(y_1)$, $Y_2 \in N(y_2)$, $Y \in N(y_1+y_2)$ with $Y_1 + Y_2 \subset Y$, $Z_1 \in M(Y_1)$ and $Z_2 \in M(Y_2)$, there exists $Z \subset M(Y)$ with $Z_1 + Z_2 \subset Z$.

(1.3.14) $K(X_1 \cap X_2) = K(X_1) \cap K(X_2)$.

(1.3.15) If $X' \in N(x_0)$, $W' \in K(X')$, $y' \in E$, and $Y \in N(y')$, there exist $X \in N(x_0)$, $W \in K(X)$, and $\lambda > 0$ such that

$$W + (0,\lambda)Y \subset W' .$$

Then $R(C,x_0)$ is a convex cone for all $(C,x_0) \in 2^E \times E$.

Proof: Let $y_1, y_2 \in R(C,x_0)$. By (1.3.12) and Proposition 1.3.2, R is cone-valued, so it suffices to show that $y := y_1 + y_2 \in R(C,x_0)$. Let $Y \in N(y)$ be given, and let $Y_1 \in N(y_1)$, $Y_2 \in N(y_2)$ be such that $Y_1 + Y_2 \subset Y$. Since $y_i \in R(C,x_0)$, there exist $X_i \in N(x_0)$, $\lambda_i > 0$, $W_i \in K(X_i)$ and $Z_i \in M(Y_i)$ such that for all $x \in W_i \cap C$ and $t \in (0,\lambda_i)$, $(x + tZ_i) \cap C \neq \emptyset$ for $i = 1,2$. By (1.3.14) and (1.3.15), choose $\lambda \leq \min(\lambda_1,\lambda_2)$, $X \subset X_1 \cap X_2$, $X \in N(x_0)$ and $W \in K(X)$, $W \subset W_1 \cap W_2$ such that $W + (0,\lambda)Y_1 \subset W_2$. Now let $x' \in C \cap W$, $t \in (0,\lambda)$. Then $x' \in C \cap W_1$, so $(x' + tZ_1) \cap C \neq \emptyset$. Pick $x'' \in (x' + tZ_1) \cap C \subset C \cap W_2$. Then $(x'' + tY_2) \cap C \neq \emptyset$. By (1.3.13), choose $Z \in M(Y)$ with $Z_1 + Z_2 \subset Z$. Then $x'' + tZ_2 \subset x' + tZ_1 + tZ_2 \subset x' + tZ$. Hence $x' + tZ \cap C \neq \emptyset$, and so $y \in R(C,x_0)$ and $R(C,x_0)$ is convex. □

Corollary 1.3.22: The q-cones $T_C(x_0)$, $H_C(x_0)$ and $F_C(x_0)$ are always convex.

Proof: $K(X) = X$, $M(Y) = Y$, $M(Y) = y$ and $M(Y) =$ "nonempty compact subsets of Y" all satisfy hypotheses (1.3.12) through (1.3.15) of Theorem 1.3.21. □

We next give a companion result to Theorem 1.3.21 on the convexity of certain q-cones.

Theorem 1.3.23: Suppose R is a q-cone on E with $* := \exists$, $\# := \exists$ and $\$:= \exists$ and assume that conditions (1.3.12), (1.3.14) and (1.3.15) hold, along with the following condition:

(1.3.13)' If $Y_1 \in N(y_1)$; $Y_2 \in N(y_2)$ and $Y \in N(y_1+y_2)$ with $Y \subset Y_1 + Y_2$ and if $Z_1 \in M(Y_1)$, $Z_2 \in M(Y_2)$, there exists $Z \subset Z_1 + Z_2$ with $Z \in M(Y)$.

Then $R(C,x_0)$ is a convex cone for all $(C,x_0) \in 2^E \times E$.

Proof: Suppose $y_1, y_2 \in R(C,x_0)$. It again suffices to show that $y := y_1 + y_2 \in R(C,x_0)$. Since $y_i \in R(C,x_0)$, there exist $X_i \in N(x_0)$, $\lambda_i > 0$ $W_i \in K(X_i)$, convex neighbourhoods $Y_i \in N(y_i)$, $Z_i \in M(Y_i)$ such that $(W_i \cap C) + (0,\lambda_i)Z_i \subset C$ for $i = 1,2$. By (1.3.14) and (1.3.15), choose $\lambda \leq \min(\lambda_1,\lambda_2)$, $X \in N(x_0)$ with $X \subset X_1 \cap X_2$ and $W \in K(X)$ with $W \subset W_1 \cap W_2$ such that

$$W + (0,\lambda)Y_1 \subset W_2 .$$

Now let $x' \in C \cap W$ and $t \in (0,\lambda)$. Then $x' \in W_1 \cap C$, so $x' + tZ_1 \subset C$. But $x' + tZ_1 \subset W_2$, so $x' + tZ_1 + tZ_2 \subset C$. Now choose $Y \in N(y)$ with $Y \subset Y_1 + Y_2$ and by (1.3.13)', choose $Z \in M(Y)$ with $Z \subset Z_1 + Z_2$. Then

$x' + tZ \subset x' + t(Z_1+Z_2) \subset C$, and so $y \in R(C,x_0)$ and $R(C,x_0)$ is convex. □

The fact that $H_C(x_0)$ is always convex follows from Theorem 1.3.23 as well as from Theorem 1.3.21, since $M(Y) = y$ for $Y \in N(y)$ satisfies both (1.3.13) and (1.3.13)'. It also follows from Theorem 1.3.23 that $I_C(x_0)$ is always convex.

We conclude this section by giving conditions sufficient for q-cones to possess property (5) or property (6).

<u>Theorem 1.3.24</u>: Suppose R is a q-cone on E with $* := \exists$, $\# := \exists$ and $\$:= \exists$ and suppose the following conditions hold:

(1.3.16) If $X_1, X_2 \in N(x_0)$, then

$$K(X_1 \cap X_2) \supset \{W_1 \cap W_2 \mid W_1 \in K(X_1),\ W_2 \in K(X_2)\} .$$

(1.3.17) If $Y_1, Y_2 \in N(y)$, then

$$M_{C_1 \cap C_2}(Y_1 \cap Y_2) \supset \{Z_1 \cap Z_2 \mid Z_1 \in M_{C_1}(Y_1), Z_2 \in M_{C_2}(Y_2)\} .$$

Then R has property (5).

<u>Proof</u>: Suppose $y \in R(C_1,x_0) \cap R(C_2,x_0)$. There exist $Y_i \in N(y)$, $X_i \in N(x_0)$, $W_i \in K(X_i)$, $Z_i \in M_{C_i}(Y_i)$ and $\lambda_i > 0$ such that

$$(C_i \cap W_i) + (0,\lambda_i)Z_i \subset C_i \quad \text{for } i = 1,2 .$$

Let $X := X_1 \cap X_2 \in N(x_0)$ and $Y := Y_1 \cap Y_2 \in N(y)$. By

(1.3.16), $W_1 \cap W_2 \in K(X)$, and by (1.3.17), $Z_1 \cap Z_2 \in M_{C_1 \cap C_2}(Y)$. Now $C_1 \cap C_2 \cap W_1 \cap W_2 + (0,\min(\lambda_1,\lambda_2))(Z_1 \cap Z_2) \subset C_1 \cap C_2$ and so $y \in R(C_1 \cap C_2, x_0)$ and R possesses property (5). □

Corollary 1.3.25: The q-cones $E_C(x_0)$, $L_C(x_0)$, $L_C^{rel}(x_0)$, $H_C(x_0)$, $I_C(x_0)$ and $I_C^{rel}(x_0)$ have property (5).

Theorem 1.3.26: Suppose R is a q-cone on E with $K(X) = x_0$, $* = \forall$, $\$ = \forall$ and suppose the following condition holds:

(1.3.18) If $Y_1, Y_2 \in N(y)$, then

$$M_{C_1 \cup C_2}(Y_1 \cap Y_2) \subset \{Z_1 \cap Z_2 \mid Z_1 \in M_{C_1}(Y_1), Z_2 \in M_{C_2}(Y_2)\}.$$

Then R has property (6).

Proof: Suppose $y \in (E \setminus R(C_1,x_0)) \cap (E \setminus R(C_2,x_0))$. Then there exist $Y_i \in N(y)$ and $\lambda_i > 0$ such that for all $Z_i \in M_{C_i}(Y_i)$, $y' \in Z_i$ and $t \in (0,\lambda_i)$, $x_0 + ty' \notin C_i$, $i = 1,2$. Now let $Y := Y_1 \cap Y_2 \in N(y)$, and let $Z \in M_{C_1 \cup C_2}(Y)$. By (1.3.18), there exist $Z_1 \in M_{C_1}(Y_1)$, $Z_2 \in M_{C_2}(Y_2)$ with $Z = Z_1 \cap Z_2$. Then for all

$$y' \in Z = Z_1 \cap Z_2$$

and $t \in (0,\min(\lambda_1,\lambda_2))$, $x_0 + ty' \notin C_1 \cup C_2$. Hence $y \notin R(C_1 \cup C_2, x_0)$, and so $R(C_1 \cup C_2, x_0) \subset R(C_1,x_0) \cup R(C_2,x_0)$. □

<u>Corollary 1.3.27:</u> The q-cones $K_C(x_0)$ and $A_C(x_0)$ possess property (6).

<u>Proof:</u> Both $M(Y) = Y$ and $M(Y) = y$ satisfy (1.3.18). □

It is a general rule of thumb that if a q-cone cannot be readily proven to have a certain property, there is a counterexample which shows that it does not have that property. This is true at least for cases (a) through (d) and properties (1) through (6) - see Table 1.9.3. In particular, in cases (a) through (d), there are no nontrivial q-cones having both properties (3) and (4) (compare Theorems 1.3.19, 1.3.21 and 1.3.23), none having both properties (4) and (6) (compare Theorems 1.3.21, 1.3.23 and 1.3.26), and none having both properties (5) and (6) (compare Theorem 1.3.24 and 1.3.26).

The only q-cones in cases (a) or (b) which have both properties (1) and (2) are the contingent cone, DM tangent cone, and Clarke tangent cone. This adds extra justification to the fact that $K_C(x_0)$ and $T_C^*(x)$ are the tangent cones which have received the most attention in optimization theory. The potential of $k_C(x_0)$ has not up to now been fully exploited; it plays an important subsidiary role and can be used to sharpen certain results in which the contingent cone has previously been used, as we will see in later chapters.

1.4. Additional Tangent Cone Properties

In this section we discuss additional important tangent cone properties and establish which quantificational tangent cones satisfy each of them.

<u>Definition 1.4.1</u>: Let E, F be l.c.s. A tangency operator is <u>product-preserving</u> if

$$(1.4.1) \qquad R(C_1 \times C_2, (x_1, x_2)) = R(C_1, x_1) \times R(C_2, x_2)$$

whenever $x_1 \in C_1 \subset E$ and $x_2 \in C_2 \subset F$.

This property will prove to be an essential one in establishing the subdifferential calculus results of chapter 2. The following theorem gives conditions sufficient for a q-cone to be product-preserving.

<u>Theorem 1.4.2</u>: Suppose E, F are l.c.s. and $x_1 \in C_1 \subset E$, $x_2 \in C_2 \subset F$. Suppose the topology on $E \times F$ is the product topology and make the following assumptions:

$$(1.4.2) \qquad K(X_1 \times X_2) \subset K(X_1) \times K(X_2) \quad \text{for all} \quad X_1 \in N(x_1),\ X_2 \in N(x_2).$$

$$(1.4.3) \qquad M_{C_1 \times C_2}(Y_1 \times Y_2) \subset M_{C_1}(Y_1) \times M_{C_2}(Y_2) \quad \text{for all} \quad Y_1 \in N(y_1),\ Y_2 \in N(y_2), \text{ where } (y_1, y_2) \in R(C_1 \times C_2, (x_1, x_2)).$$

Then

(1.4.4) $R(C_1 \times C_2, (x_1,x_2)) \subset R(C_1,x_1) \times R(C_2,x_2)$.

Moreover, if $\$:= \exists$ and equality holds in (1.4.2) and (1.4.3), then (1.4.1) is satisfied.

<u>Proof</u>: Suppose $(y_1,y_2) \in R(C_1 \times C_2,(x_1,x_2))$. Since $E \times F$ is topologized with the product topology, $* \, Y_1 \in N(y_1)$, $* \, Y_2 \in N(y_2)$, $\# \, X_1 \in N(x_1)$, $\# \, X_2 \in N(x_2)$, $\$ \, \lambda > 0$, there exist $W \in K(X_1 \times X_2)$ and $Z \in M_{C_1 \times C_2}(Y_1 \times Y_2)$, $\#'$ $(x_1',x_2') \in (C_1 \times C_2) \cap W$, $\$' \, t \in (0,\lambda)$, $*' \, (y_1',y_2') \in Z$, $(x_1',x_2') + t(y_1',y_2') \in C_1 \times C_2$. By (1.4.2) and (1.4.3), $W = W_1 \times W_2$ where $W_1 \in K(X_1)$, $W_2 \in K(X_2)$ and $Z = Z_1 \times Z_2$ where $Z_1 \in M_{C_1}(Y_1)$, $Z_2 \in M_{C_2}(Y_2)$. It follows that $y_1 \in R(C_1,x_0)$ and $y_2 \in R(C_2,x_0)$, so (1.4.4) holds.

Conversely, suppose $\$:= \exists$ and $y_1 \in R(C_1,x_1)$, $y_2 \in R(C_2,x_2)$. Then $* \, Y_i \in N(y_i)$, $\# \, X_i \in N(x_i)$, there exist $\lambda_i > 0$, $W_i \in K(X_i)$ and $Z_i \in M_{C_i}(Y_i)$, $\#' \, x_i' \in C_i \cap W_i$, for all $t \in (0,\lambda_i)$, $*' \, y_i' \in Z_i$, $x_i' + ty_i' \in C_i$, $i = 1,2$. Let $\lambda := \min(\lambda_1,\lambda_2)$, $W := W_1 \times W_2$, $Z := Z_1 \times Z_2$, $Y := Y_1 \times Y_2$, $X := X_1 \times X_2$. By (1.4.2) and (1.4.3), $W \in K(X)$ and $Z \in M_{C_1 \times C_2}(Y)$. Then $* \, Y \in N((y_1,y_2))$, $\# \, X \in N((x_1,x_2))$, there exist $W \in K(X)$ and $Z \in M_{C_1 \times C_2}(Y)$, $\#' \, (x_1',x_2') \in (C_1 \times C_2) \cap W$, for all $t \in (0,\lambda)$, $*' \, (y_1',y_2') \in Y \cap Z$, $(x_1',x_2') + t(y_1',y_2') \in C_1 \times C_2$. Hence

$(y_1, y_2) \in R(C_1 \times C_2, (x_1, x_2))$, and so (1.4.1) holds. □

Corollary 1.4.3: The q-cones $k_C(x_0)$, $T_C(x_0)$, $E_C(x_0)$, $I_C(x_0)$, $L_C(x_0)$, $H_C(x_0)$, $F_C(x_0)$ and $G_C(x_0)$ are product-preserving.

Proof: Assumptions (1.4.2) and (1.4.3) hold with equality in cases (a) through (c). In the definition of each of these cones, $\$:= \exists$. □

If $\$ = \forall$ in the definition of a q-cone, that cone will not in general be product preserving. (There is a counter-example for the contingent cone in [Bo5, Section 2]. Counterexamples for other q-cones are given in Table 1.9.3.) For q-cones with $\$:= \forall$, the best obtainable "product inclusion" is the following result, the proof of which we leave to the reader.

Proposition 1.4.4: Suppose E , F are l.c.s. and $E \times F$ has the product topology. Suppose R is a q-cone, and suppose R' is a q-cone on F in whose definition $\$:= \exists$. Assume that $R'(C, x_0) \subset R(C, x_0)$ for all $(C, x_0) \in 2^F \times F$, that $M(\cdot)$ and $K(\cdot)$ are identical in the definition of R and R' , and that equality holds in (1.4.2) and (1.4.3). Then for all $x_1 \in C_1 \subset E$ and $x_2 \in C_2 \subset F$,

$$(1.4.5) \qquad R(C_1 \times C_2, (x_1, x_2)) \supset R(C_1, x_1) \times R'(C_2, x_2) .$$

Corollary 1.4.5. (See [U1] for (1.4.6) and (1.4.8)): For all $x_1 \in \mathrm{cl}\, C_1 \subset E$ and $x_2 \in \mathrm{cl}\, C_2 \subset F$,

(1.4.6) $$K_{C_1}(x_1) \times K_{C_2}(x_2) \subset K_{C_1 \times C_2}((x_1,x_2))$$

(1.4.7) $$A_{C_1}(x_1) \times E_{C_2}(x_2) \subset A_{C_1 \times C_2}((x_1,x_2))$$

(1.4.8) $$\ell_{C_1}(x_1) \times L_{C_2}(x_2) \subset \ell_{C_1 \times C_2}((x_1,x_2))$$

We next mention a localization property possessed by many q-cones.

Theorem 1.4.6: Suppose R is a q-cone on E satisfying the condition below:

(1.4.9) If $X_1, X_2 \in \mathcal{N}(x_0)$ and $W \in \mathcal{K}(X_1)$, then $W \cap X_2 \in \mathcal{K}(X_1 \cap X_2)$.

Then

(1.4.10) $R(C \cap X, x_0) = R(C, x_0)$ for all $(C, x_0) \in 2^E \times E$ and $X \in \mathcal{N}(x_0)$.

Proof: Let $y \in R(C, x_0)$. We observe that for any $Y \in \mathcal{N}(y)$, there exist $\lambda_1 > 0$ and $X_1 \in \mathcal{N}(x_0)$ such that $X_1 \subset X$ and $X_1 + (0,\lambda_1)Y \subset X$. Then $* \; Y \in \mathcal{N}(y)$, $\# \; X' \in \mathcal{N}(x_0)$, $\$ \; \lambda > 0$, there exist $W \in \mathcal{K}(X')$ and $Z \in \mathcal{M}(Y)$, $\#' \; x \in X_1 \cap C \cap W = (C \cap X_1) \cap (W \cap X_1)$, $\$' \; t \in (0, \min(\lambda,\lambda_1))$, $*' \; y' \in Z \subset Y$, $x + ty' \in C$. But since $X_1 + (0,\lambda_1)Y \subset X$, $x + ty' \in C \cap X$. In addition,

$W \cap X_1 \in K(X' \cap X_1)$ by (1.4.9). Hence $y \in R(C \cap X, x_0)$ and

$$R(C,x_0) \subset R(C \cap X, x_0) \ .$$

Conversely, suppose $y \in R(C \cap X, x_0)$. Then $* \ Y \in N(y)$, $\# \ X' \in N(x_0)$, $\$ \ \lambda > 0$, there exist $W \in K(X')$ and $Z \in K(Y)$, $\#' \ x \in (C \cap X) \cap W = C \cap (W \cap X)$, $\$' \ t \in (0,\lambda)$, $*' \ y' \in Z$, $x + ty' \in C \cap X$. Since $X \cap X' \in N(x_0)$ and $W \cap X \subset K(X' \cap X)$, $y \in R(C,x_0)$. Hence $R(C \cap X, x_0) \subset R(C,x_0)$ and (1.4.10) holds. $\square$

In particular, hypothesis (1.4.9) holds whenever $K(X) := X$ or $K(X) := x_0$, so (1.4.10) is true for any of the q-cones we have defined so far. This property is also discussed in [Bo8]. Observe that (1.4.10) does not hold for $R(C,x_0) := \mathrm{cone}(C-x_0)$, which is not a q-cone.

Another property important in our future considerations is the following inclusion for $g: E \to F$ strictly differentiable at $x_0 \in C \subset E$:

$$(1.4.11) \qquad \nabla g(x_0) R(C,x_0) \subset R(g(C), g(x_0)) \ .$$

In Propositions 1.4.7 and 1.4.10 below, we give conditions under which (1.4.11) holds for various classes of q-cones.

<u>Proposition 1.4.7</u>: Suppose R is a q-cone with $\# := \forall$, $M(Y) = Y$ and $* := \forall$ and suppose $g: E \to F$ is strictly differentiable at $x_0 \in C \subset E$. Make the following assumption:

(1.4.12) If $V \in N(g(x_0))$, $X \in N(x_0)$ with $g(X) \subset V$, then $g(K(X)) \subset K(V)$.

Then (1.4.11) holds.

<u>Proof</u>: Let $y \in R(C,x_0)$, $U \in N(\nabla g(x_0)y)$ and $V \in N(g(x_0))$. Since g is strictly differentiable at x_0, there exist $Y \in N(y)$, $X \in N(x_0)$ and $\lambda_1 > 0$ such that for all $t \in (0,\lambda_1)$, $y' \in Y$ and $x \in X$, $\frac{g(x+ty')-g(x)}{t} \in U$ and $g(x) \in V$.

Since $y \in R(C,x_0)$, $\# := \forall$, $M(Y) = Y$ and $* := \forall$, $\$ \lambda > 0$, there exist $W \in K(X)$ and $x \in C \cap W$, $\$' t \in (0,\min(\lambda,\lambda_1))$, there exists $y' \in Y$, $x + ty' \in C$. In addition, there exists $\bar{y} \in U$ with $g(x) + t\bar{y} = g(x+ty') \in g(C)$. Now if $x \in C \cap W$, $g(x) \in g(C) \cap g(W)$ and $g(W) \in K(V)$ by (1.4.12). Hence $y \in R(g(C),g(x_0))$, and (1.4.11) holds. □

<u>Corollary 1.4.8</u>: Suppose $g: E \to F$ is strictly differentiable at $x_0 \in C \subset E$. Then (1.4.11) holds for the q-cones $K_C(x_0)$, $k_C(x_0)$, $D_C(x_0)$, and

$$a_C(x_0) := \{y \in E \mid \forall Y \in N(y), \forall X \in N(x_0), \exists \lambda > 0, \exists x \in X \cap C, \forall t \in (0,\lambda), \exists y' \in Y ; x + ty' \in C\}.$$

<u>Proof</u>: The cases $K(X) = X$ and $K(X) = x_0$ both satisfy condition (1.4.12). □

In [A1], Aubin establishes (1.4.11) for $K_C(x_0)$ in the case where E and F are Hilbert spaces. See also [Bo2], [Gul] for applications of this well-known inclusion.

<u>Corollary 1.4.9</u>: Suppose $A: E \to F$ is linear and continuous and $x_0 \in C \subset E$. Then if R satisfies the hypotheses of Proposition 1.4.7,

$$A(R(C,x_0)) \subset R(A(C),Ax_0) \tag{1.4.13}$$

In particular,

$$A(K_C(x_0)) \subset K_{A(C)}(Ax_0) \tag{1.4.14}$$

and

$$A(k_C(x_0)) \subset k_{A(C)}(Ax_0) . \tag{1.4.15}$$

Inclusions (1.4.14) and (1.4.15) will be invoked in Chapter 2 in the proofs of subdifferential calculus formulae.

<u>Proposition 1.4.10</u>: Suppose R is a q-cone with $\# := \exists$, $M(Y) = Y$ and $* := \forall$ and suppose $g: E \to F$ is strictly differentiable at $x_0 \in C \subset E$. Make the following relative openness assumption:

(1.4.16) For each $X \in N(x_0)$ and $W_1 \in K(X)$, there exist $V \in N(g(x_0))$ and $W \in K(V)$ such that

$$g(C) \cap W \subset g(C \cap W_1) \, .$$

Then (1.4.11) holds.

Proof: Let $y \in R(C,x_0)$ and $U \in N(\nabla g(x_0)y)$. Since g is strictly differentiable at x_0, there exist $Y_1 \in N(y)$, $X_1 \in N(x_0)$ and $\lambda_1 > 0$ such that for all $t \in (0,\lambda_1)$, $y' \in Y_1$, and $x \in X_1$,

$$\frac{g(x+ty')-g(x)}{t} \in U$$

Since $y \in R(C,x_0)$, there exists $X \in N(x_0)$ with $X \subset X_1$, $\$ \lambda > 0$, there exists $W_1 \in K(X)$, for all $x \in C \cap W_1$, $\$' t \in (0,\min(\lambda_1,\lambda))$, there exists $y' \in Y_1$, $x + ty' \in C$. By (1.4.16), there exists $V \in N(g(x_0))$ and $W \in K(V)$ such that $g(C) \cap W \subset g(C \cap W_1)$. Now for all $y' \in Y$, $x \in C \cap W_1$, and $t \in (0,\min(\lambda,\lambda_1))$, there exists $\bar{y} \in U$ with

$$g(x+ty') = g(x) + t\bar{y} \, .$$

Thus there exists $V \in N(g(x_0))$, $\$ \lambda > 0$, there exists $W \in K(V)$, for all $g(x) \in g(C) \cap W$, $\$' t \in (0,\min(\lambda,\lambda_1))$, there exists $\bar{y} \in U$, $g(x) + t\bar{y} \in g(C)$. Hence $y \in R(g(C),g(x_0))$ and (1.4.11) holds. □

Corollary 1.4.11: Inclusion (1.4.11) holds for the q-cones $T_C(x_0)$ and

$$T_C^*(x_0) := \{y \in E \mid \forall Y \in N(y), \exists X \in N(x_0), \forall \lambda > 0, \forall x \in X \cap C, \exists t \in (0,\lambda), \exists y' \in Y, x + ty' \in C\}$$

under the following assumption:

(1.4.17) For each $X \in N(x_0)$, there exists $V \in N(g(x_0))$ with $V \cap g(C) \subset g(X \cap C)$.

<u>Proof</u>: Set $K(X) = X$ in Proposition 1.4.10. Then (1.4.16) reduces to (1.4.17). □

<u>Corollary 1.4.12</u>: Suppose $A: E \to F$ is linear and continuous and $x_0 \in C \subset E$. Then if R satisfies the hypotheses of Proposition 1.4.10, inclusion (1.4.13) holds. In particular, if (1.4.17) holds with $g := A$,

(1.4.18) $$A(T_C(x_0)) \subset T_{A(C)}(Ax_0) .$$

and

(1.4.19) $$A(T_C^*(x_0)) \subset T_{A(C)}^*(Ax_0) .$$

Inclusion (1.4.18) will be used in the proofs of sub-differential calculus formulae in Chapter 2.

In our discussion of property (2) in section 1.3, we saw that (1.4.11) does not hold for q-cones with $* := \Xi$ and sets of the form $C := g^{-1}(0)$, where $g: \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable near x_0 and $\nabla g(x_0)$ is of rank m (see Example 1.3.16). However, inclusion (1.4.13) holds more generally than does (1.4.11), as we now show in Propositions 1.4.13, 1.4.15, 1.4.17 and 1.4.19.

<u>Proposition 1.4.13</u>: Suppose that R is a q-cone with $\# := \forall$ and $* := \forall$ and let $x_0 \in C \subset E$. Suppose $A: E \to F$ is linear and continuous and that (1.4.12) holds with $g = A$. Make the following additional assumption:

(1.4.20) If $V \in \mathcal{N}(Ay)$, $Y \in \mathcal{N}(y)$ with $A(Y) \subset V$, then

$$A(\mathcal{M}(Y)) \subset \mathcal{M}(V) .$$

Then (1.4.13) holds.

<u>Proof</u>: Let $y \in R(C,x_0)$, $V \in \mathcal{N}(Ay)$ and $U \in \mathcal{N}(Ax_0)$. Then $Y := A^{-1}(V) \in \mathcal{N}(y)$ and $X := A^{-1}(U) \in \mathcal{N}(x_0)$, and $\$ \lambda > 0$, there exist $W \in \mathcal{K}(X)$, $Z \in \mathcal{M}(Y)$, and $x \in C \cap W$, $\$' \ t \in (0,\lambda)$, there exists $y' \in Z$, $x + ty' \in C$.

By (1.4.12) and (1.4.20), $A(W) \in \mathcal{K}(U)$ and $A(Z) \in \mathcal{M}(V)$. Then $\$ \lambda > 0$, there exists $Ax \in A(C \cap W) \subset A(C) \cap A(W)$, $\$' \ t \in (0,\lambda)$, there exists $Ay' \in A(Z)$ with $A(x) + tA(y') = A(x+ty') \in A(C)$. Hence $Ay \in R(A(C),Ax_0)$ and (1.4.13) holds. □

<u>Corollary 1.4.14</u>: The q-cones $E_C(x_0)$ and $A_C(x_0)$ satisfy (1.4.13) for all $x_0 \in C \subset E$.

<u>Proof</u>: The cases $\mathcal{M}(Y) = y$ and $\mathcal{M}(X) = x_0$ satisfy (1.4.20) and (1.4.12). □

<u>Proposition 1.4.15</u>: Suppose $A: E \to F$ is linear and continuous and $x_0 \in C \subset E$. Suppose R is a q-cone with

$\# := \exists$, and $* := \forall$ and assume that (1.4.20) holds and that (1.4.16) holds with $g = A$. Then (1.4.13) holds.

<u>Proof:</u> Let $y \in R(C,x_0)$ and $U \in N(Ay)$. Again $Y := A^{-1}(U) \in N(y)$ and there exists $X \in N(x_0)$, $\$ \lambda > 0$, there exist $W_1 \in K(X)$ and $Z \in M(Y)$, for all $x \in W_1 \cap C$, $\$' t \in (0,\lambda)$, there exists $y' \in Z$ with $x + ty' \in C$. By (1.4.16), there exist $V \in N(Ax_0)$ and $W \in K(V)$ with

$$A(C) \cap W \subset A(C \cap W_1) .$$

By (1.4.20), $A(Z) \in M(U)$. Combining these facts, we see that there exists $V \in N(Ax_0)$, $\$ \lambda > 0$, there exist $W \in K(V)$ and $A(Z) \in M(U)$, for all $Ax \in A(C) \cap W$, $\$' t \in (0,\lambda)$, there exists $A(y') \in A(Z)$ with $Ax + tAy' = A(x+ty') \in A(C)$. Thus $y \in R(A(C),Ax_0)$ and (1.4.13) holds. □

<u>Corollary 1.4.16:</u> Suppose that $A: E \to F$ is linear and continuous, that $x_0 \in C \subset E$, and that (1.4.17) holds with $g = A$. Then (1.4.13) holds for the q-cones $H_C(x_0)$, $F_C(x_0)$, and

$$H_C^*(x_0) := \{y \mid \exists\, x \in N(X_0), \forall \lambda > 0, \forall x \in X \cap C , \\ \exists\, t \in (0,\lambda), x + ty \in C\} .$$

<u>Proof:</u> Assumption (1.4.17) holds with $M(Y)$ as in cases (b) or (c). □

<u>Proposition 1.4.17</u>: Suppose that $A: E \to F$ is linear, open and continuous and that $x_0 \in C \subset E$. Suppose R is a q-cone with $* := \exists$ and $\# := \forall$. Assume (1.4.12) holds with $g = A$ and make the following additional assumption:

(1.4.21) If $Y \in \mathcal{N}(y)$ and $V \in \mathcal{N}(Ay)$ with $V \subset A(Y)$,
then $\mathcal{M}(V) \supset A(\mathcal{M}(Y))$.

Then (1.4.13) holds.

<u>Proof</u>: Let $y \in R(C,x_0)$ and $U \in \mathcal{N}(Ax_0)$. Then $X := A^{-1}(U) \in \mathcal{N}(x_0)$ and there exists $Y \in \mathcal{N}(y)$, $\$\, \lambda > 0$, there exist $W \in \mathcal{K}(X)$, $Z \in \mathcal{M}(Y)$, and $x \in C \cap W$, $\$'\, t \in (0,\lambda)$, for all $y' \in Z$, $x + ty' \in C$. Since A is open, there exists $V \in \mathcal{N}(Ay)$ with $V \subset A(Y)$, so by (1.4.21), $A(Z) \in \mathcal{M}(V)$. By (1.4.12) with $g := A$, $A(W) \in \mathcal{K}(U)$. Combining these facts, we see that there exists $V \in \mathcal{N}(Ay)$ such that for all $U \in \mathcal{N}(Ax_0)$, $\$\, \lambda > 0$, there exist $A(W) \in \mathcal{K}(U)$, $A(Z) \in \mathcal{M}(V)$ and $A(x) \in A(C) \cap A(W)$, $\$'\, t \in (0,\lambda)$ for all $A(y') \in A(Z)$, $A(x) + tA(y') = A(x+ty') \in A(C)$. Hence $A(y) \in R(A(C),A(x_0))$ and (1.4.13) holds. $\square$

<u>Corollary 1.4.18</u>: Suppose that $A: E \to F$ is linear, continuous, and open, and that $x_0 \in C \subset E$. Then the q-cones $G_C(x_0)$, $L_C(x_0)$, $\ell_C(x_0)$, $L_C^{rel}(x_0)$ and $\ell_C^{rel}(x_0)$ satisfy (1.4.13).

Proof: Assumption (1.4.12) holds for $K(X) = X$ or $K(X) = x_0$, and assumption (1.4.21) holds for $M(Y) = Y$ or $M(Y) = Y \cap (\text{aff } C - x_0)$. □

Proposition 1.4.19: Suppose that $A: E \to F$ is linear, continuous and open, and that $x_0 \in C \subset E$. Suppose R is a q-cone with $* := \exists$; $\# := \exists$, and assume that (1.4.21) holds and that (1.4.16) holds with $g = A$. Then (1.4.13) holds.

Proof: The proof is similar to that of Proposition 1.4.17, using (1.4.16) instead of (1.4.12). □

Corollary 1.4.20: Suppose $A: E \to F$ is continuous, linear and open and let $x_0 \in C \subset E$. Assume (1.4.16) holds with $g = A$. Then the q-cones $I_C(x_0)$, $I_C^{rel}(x_0)$ and

$$I_C^*(x_0) := \{y \in E \mid \exists\, Y \in N(y),\ \exists\, X \in N(x_0),\ \forall\, \lambda > 0,$$
$$\forall\, x \in X \cap C,\ \exists\, t \in (0,\lambda),\ \forall\, y' \in Y \cap C,\ x + ty' \in C\}$$

satisfy (1.4.13).

Proof: If $M(Y)$ is as in cases (a) or (d), (1.4.21) is satisfied. □

A summary of the conditions under which (1.4.11) and (1.4.13) hold for various q-cones is given in Table 1.9.5.

1.5. Directional derivatives, subgradients, and necessary conditions for optimality

<u>Definition 1.5.1</u>: Let $f\colon E \to \overline{\mathbb{R}}$ be an extended real valued function. The (effective) <u>domain</u> of f is the set

(1.5.1) $$\operatorname{dom} f := \{x \in E \mid f(x) < +\infty\}$$

The <u>epigraph</u> of f is the set

(1.5.2) $$\operatorname{epi} f := \{(x,r) \in E \times \mathbb{R} \mid f(x) \leq r\} .$$

One of the most important uses of tangent cones is in defining directional derivatives and subgradients of functions via tangent cones of epigraphs. This is an especially fruitful idea for q-cones. Directional derivatives associated with q-cones are much more analytically tractable than those associated with other tangent cones, and their subgradients often admit a versatile calculus. We will therefore concentrate on directional derivatives and subgradients associated with q-cones.

<u>Definition 1.5.2</u>: Let R be a q-cone, and denote by $\mathcal{F}(E,\overline{\mathbb{R}})$ the set of extended real valued functions on E. Define $\hat{R}\colon \mathcal{F}(E,\overline{\mathbb{R}}) \times E \to 2^{E \times \mathbb{R}}$ by

(1.5.3) $$\hat{R}(f,x_0) := \{(y,d) \in E \times \mathbb{R} \mid \forall\, \delta > 0,\ *\, Y \in \mathcal{N}(y),\ \#\, X \in \mathcal{N}(x_0),\ \#\, u > 0,\ \$\, \lambda > 0,\ \exists\, W \in \mathcal{K}(X) \times \mathcal{K}((f(x_0)-u, f(x_0)+u)),\ \exists\, Z \in \mathcal{M}(Y),\ \#'\, (x,r) \in \operatorname{epi} f \cap W,\ \$'\, t \in (0,\lambda)\ ,$$

$*' y' \in Z, \exists d' \leq d + \delta, (x,r) + t(y',d') \in \text{epi } f\}$.

Remark 1.5.3: (a) Notice that the expression "$(x,r) + t(y',d') \in \text{epi } f$" in (1.5.3) can also be written "$\frac{f(x+ty')-r}{t} \leq d'$". It follows that if $(y,d) \in \hat{R}(f,x_0)$, then $(y,d+p) \in \hat{R}(f,x_0)$ for any $p \geq 0$.

(b) If $* := \forall$ and $M(Y) = Y$, then $\hat{R}(f,x_0) = R(\text{epi } f,(x_0,f(x_0)))$. This is true in particular for $T_C(x_0)$, $K_C(x_0)$ and $k_C(x_0)$, the three examples we will examine in the greatest detail.

Definition 1.5.4: Let R be a q-cone, $f: E \to \overline{\mathbb{R}}$, and $x_0 \in \text{dom } f$. The R directional derivative of f at x_0 with respect to Y is defined by

$$(1.5.4) \quad f^R(x_0;y) := \inf\{r \mid (y,r) \in \hat{R}(f,x_0)\} .$$

The R subgradient of f at x_0 is the set

$$(1.5.5) \quad \partial^R f(x_0) := \{x' \in E' \mid \langle y,x' \rangle \leq f^R(x_0;y) \text{ for all } y \in E\}.$$

Remark 1.5.5: (a) This idea of defining directional derivatives and subgradients via tangent cones of epigraphs of functions has been widely discussed in recent years - see for example [A1], [A2], [Bo7], [C14], [Do1], [Hi1], [Hi2], [Hi3], [Pe2], [Ro4], [Ro5] and [Th1].

(b) It is also possible to define directional derivatives via graphs, rather than epigraphs, of functions - see [U1],

[Dol]. These derivatives, in order to be finite, require the function f to be more "smooth" than do the directional derivatives defined in (1.5.4). Since our purpose here is to develop versions of differential calculus and optimization theory for nonsmooth functions, we will not investigate these concepts here.

(c) Definitions analogous to (1.5.3), (1.5.4) and (1.5.5) have been made for vector-valued functions (e.g. [Th1]) and set-valued mappings (e.g. [Th2], [Pe2], [Al]). We will discuss the vector-valued case in Chapter 3.

(d) Suppose R_1 and R_2 are q-cones satisfying $\hat{R}_1(f,x_0) \subset \hat{R}_2(f,x_0)$ for all $f: E \to \overline{\mathbb{R}}$ and $x_0 \in E$. It follows from (1.5.4) that

$$f^{R_2}(x_0;\cdot) \leq f^{R_1}(x_0;\cdot) ,$$

and so by (1.5.5)

$$\partial^{R_2} f(x_0) \subset \partial^{R_1} f(x_0) .$$

In particular,

$$f^K(x_0;\cdot) \leq f^k(x_0;\cdot) \leq f^T(x_0;\cdot)$$

and

$$\partial^K f(x_0) \subset \partial^k f(x_0) \subset \partial^T f(x_0) .$$

Relationships among the various R directional derivatives can thus be deduced from relationships among the corresponding tangent cones.

(e) It is not hard to show, taking into account Remark 1.5.3(a), that $\hat{R}(f,x_0) = \text{epi } f^R(x_0;\cdot)$. We will apply this fact below.

We have already noted that if R is a q-cone with $* := \forall$ and $M(Y) = Y$, (1.5.4) becomes

$$(1.5.6) \qquad f^R(x_0;y) = \inf\{r \mid (y,r) \in R(\text{epi } f,(x_0,f(x_0)))\} .$$

We now demonstrate that (1.5.6) also holds if $M(Y) = y$.

<u>Proposition 1.5.6</u>: Let R be a q-cone with $M(Y) = y$, $f: E \to \overline{\mathbb{R}}$, and $x_0 \in \text{dom } f$. Then (1.5.6) holds.

<u>Proof</u>: It is clear that for such a q-cone,

$$R(\text{epi } f,(x_0,f(x_0))) \subset \hat{R}(f,x_0) ,$$

so $f^R(x_0;y) \le \inf\{r \mid (y,r) \in R(\text{epi } f,(x_0,f(x_0)))\}$. On the other hand, suppose $f^R(x_0;y) \le d$. Since $\hat{R}(f,x_0) = \text{epi } f^R(x_0;\cdot)$, $(y,d) \in \hat{R}(f,x_0)$, and it follows that $(y,d+\delta) \in R(\text{epi } f,(x_0,f(x_0)))$ for all $\delta > 0$. Therefore

$$\inf\{r \mid (y,r) \in R(\text{epi } f,(x_0,f(x_0)))\} \le d ,$$

and so $f^R(x_0;d) \ge \inf\{r \mid (y,r) \in R(\text{epi } f,(x_0,f(x_0)))\}$. Hence (1.5.6) holds. □

<u>Corollary 1.5.7</u>: Let $f: E \to \overline{\mathbb{R}}$ and $x_0 \in \operatorname{dom} f$. Then

$$(1.5.7) \qquad f^H(x_0;y) = \inf\{r \mid (y,r) \in H_{\operatorname{epi} f}((x_0,f(x_0)))\},$$

$$(1.5.8) \qquad f^A(x_0;y) = \inf\{r \mid (y,r) \in A_{\operatorname{epi} f}((x_0,f(x_0)))\}$$

and

$$(1.5.9) \qquad f^E(x_0;y) = \inf\{r \mid (y,r) \in E_{\operatorname{epi} f}((x_0,f(x_0)))\}.$$

<u>Proof</u>: In the definition of the tangent cones $H_C(\cdot)$, $E_C(\cdot)$ and $A_C(\cdot)$, $M(Y) = y$, so these are special cases of Proposition 1.5.6. □

There is an alternate, more direct way to define the directional derivatives of (1.5.4). With this alternate characterization, we will see that many of the R directional derivatives are familiar objects.

<u>Theorem 1.5.8</u>: Let R be a q-cone, $f: E \to \overline{\mathbb{R}}$. and $x_0 \in \operatorname{dom} f$. Then

$$(1.5.10) \qquad f^R(x_0;y) = \mathop{\square}_{Y\in N(y)} \ \mathop{\Delta}_{\substack{X\in N(x_0)\\ u>0}} \ \mathop{O}_{\lambda>0} \ \inf_{\substack{W\in K(X)\\ \times K(f(x_0)-u,\,f(x_0)+u)}}$$

$$\inf_{Z\in M(Y)} \ \mathop{\Delta'}_{(x,r)\in \operatorname{epi} f\cap W} \ \mathop{O'}_{t\in(0,\lambda)} \ \mathop{\square'}_{y'\in Z} \ \frac{f(x+ty')-r}{t}$$

where $\square = \begin{cases} \sup & \text{if } * = \forall \\ \inf & \text{if } * = \exists \end{cases},$

$$\Delta = \begin{cases} \sup & \text{if } \# = \forall \\ \inf & \text{if } \# = \exists \end{cases} ,$$

$$O = \begin{cases} \sup & \text{if } \$ = \forall \\ \inf & \text{if } \$ = \exists \end{cases} , \quad \text{and}$$

$$\square' , \Delta' , O' \in \{\sup,\inf\} ,$$

$$\square' \neq \square , \Delta \neq \Delta' , O' \neq O .$$

<u>Proof</u>: Denote by S the right hand side of (1.5.10). Then $S \leq d$ if and only if for all $\delta > 0$; $* \; Y \in N(y)$, $\# \; X \in N(x_0)$, $\# \; u > 0$, $\$ \; \lambda > 0$, there exist $W \in K(X) \times K(f(x_0)-u, f(x_0)+u)$ and $Z \in M(Y)$, $\#' \; (x,r) \in \operatorname{epi} f \cap W$, $\$' \; \lambda > 0$, $*' \; y' \in Y \cap Z$, $\frac{f(x+ty')-r}{t} \leq d + \delta$. In other words, $S \leq d$ if and only if $(y,d) \in \hat{R}(f,x_0)$. As was mentioned in Remark 1.5.5(e), $\hat{R}(f,x_0) = \operatorname{epi} f^R(x_0;\cdot)$, so $(y,d) \in \hat{R}(f,x_0)$ if and only if $f^R(x_0;y) \leq d$. Hence $S \leq d$ if and only if $f^R(x_0;y) \leq d$ and (1.5.10) holds. □

We next list some familiar special cases of (1.5.10).

<u>Example 1.5.9</u>: The upper and lower one-sided directional derivatives of f at x_0 with respect to y are given, respectively, by

$$(1.5.11) \qquad f^E(x_0;y) = \inf_{\lambda>0} \sup_{t\in(0,\lambda)} \frac{f(x_0+ty)-f(x_0)}{t}$$

and

$$(1.5.12) \qquad f^A(x_0;y) = \sup_{\lambda>0} \inf_{t\in(0,\lambda)} \frac{f(x_0+ty)-f(x_0)}{t} .$$

Example 1.5.10: The upper and lower one-sided Hadamard derivatives of f at x_0 with respect to y are given, respectively, by

$$(1.5.13) \qquad f^+(x_0;y) := f^L(x_0;y) = \inf_{\substack{Y\in N(y)\\ \lambda>0}} \sup_{\substack{y'\in Y\\ t\in(0,\lambda)}} \frac{f(x_0+ty')-f(x_0)}{t}$$

and

$$(1.5.14) \qquad f_+(x_0;y) := f^K(x_0;y) = \sup_{\substack{Y\in N(y)\\ \lambda>0}} \inf_{\substack{y'\in Y\\ t\in(0,\lambda)}} \frac{f(x_0+ty')-f(x_0)}{t}$$

Example 1.5.11: The Clarke derivative of f at x_0 with respect to y (e.g. [Hi3], [Cl4]) is defined by

$$(1.5.15) \qquad f^0(x_0;y) := f^H(x_0;y)$$

$$= \inf_{\substack{X\in N(x_0)\\ u>0\\ \lambda>0}} \sup_{\substack{(x,r)\in X\times(f(x_0)-u,f(x_0)+u)\\ \cap \text{ epi } f\\ t\in(0,\lambda)}} \frac{f(x+ty)-r}{t} .$$

When f is continuous at x_0, the Clarke derivative simplifies to the more familiar-looking

$$(1.5.16) \qquad f^0(x_0;y) = \inf_{\substack{X\in N(x_0)\\ \lambda>0}} \sup_{\substack{x\in X\\ t\in(0,\lambda)}} \frac{f(x+ty)-f(x)}{t} .$$

We will discuss this in more detail later.

Example 1.5.12: The upper subderivative of f at x_0 with respect to y ([Ro3], [Ro4], [Ro5]) is defined by

$$(1.5.17)\quad f^{\uparrow}(x_0;y) = f^{T}(x_0;y) = \sup_{Y\in N(y)} \inf_{\substack{X\in N(x_0)\\ \lambda>0\\ u>0}} \sup_{\substack{(x,r)\in X\times(f(x_0)-u,f(x_0)+u)\\ \cap \operatorname{epi} f\\ t\in(0,\lambda)}} \inf_{y'\in Y} \frac{f(x+ty')-r}{t} .$$

The expression in (1.5.17) becomes simpler in special cases, as we will see shortly.

By Remark 1.5.5(d) and Table 1.9.4, the inequalities

$$f^{A}(x_0;\cdot) \le f^{E}(x_0;\cdot) ,$$

$$f_{+}(x_0;\cdot) \le f^{+}(x_0;\cdot) ,$$

$$f_{+}(x_0;\cdot) \le f^{A}(x_0;\cdot) ,$$

$$f^{E}(x_0;\cdot) \le f^{+}(x_0;\cdot) , \text{ and}$$

$$f_{+}(x_0;\cdot) \le f^{\uparrow}(x_0;\cdot) \le f^{0}(x_0;y) \text{ hold.}$$

In important special cases, various of these directional derivatives coincide. We begin by examining the case in which $f: E \to \mathbb{R}$ is convex. In this case,

$$A_{\operatorname{epi} f}((x_0,f(x_0))) = E_{\operatorname{epi} f}((x_0,f(x_0))) ,$$

and so $f^{A}(x_0;\cdot) = f^{E}(x_0;\cdot) = f'(x_0;\cdot)$, where $f'(x_0;\cdot)$ is the one-sided directional derivative defined in (1.2.3).

In addition, as we saw in section 1.3, $K_{\mathrm{epi}\ f}((x_0,f(x_0))) = k_{\mathrm{epi}\ f}((x_0,f(x_0))) = T_{\mathrm{epi}\ f}((x_0,f(x_0))) = \mathrm{cl}\ A_{\mathrm{epi}\ f}((x_0,f(x_0)))$ for convex f, so $f_+(x_0;\cdot) = f^k(x_0;\cdot) = f^\dagger(x_0;\cdot) = \mathrm{cl}\ f'(x_0;\cdot)$. (For the definition of the "closure" of a function, see [Ro1, Section 7] or Definition 1.6.4(ii).) It follows that $\partial^A f(x_0) = \partial^E f(x_0) = \partial^K f(x_0) = \partial^k f(x_0) = \partial^T f(x_0)$ for f convex. In addition, if f is bounded above on a neighbourhood of x_0, then $T_{\mathrm{epi}\ f}((x_0,f(x_0))) = \mathrm{cl}\ H_{\mathrm{epi}\ f}(x_0,f(x_0))$ and so $f^\dagger(x_0;\cdot) = \mathrm{cl}\ f^0(x_0;\cdot)$ and $\partial^T f(x_0) = \partial^H f(x_0)$. All of these R subgradient sets, in fact, coincide with the "classical" subgradient set defined in (1.2.1); since

$$\partial f(x_0) = \{x'\in E' \mid \langle y,x'\rangle \le f'(x_0;y),\ \forall\ y \in E\} .$$

This fact is the motivation behind the notation in (1.5.5).

We can now see the importance of tangent cone Property (1). Tangent cones with this property are those whose associated subgradient sets coincide with the ordinary subgradient for convex functions.

Another important case is that in which $f: E \to \overline{\mathbb{R}}$ is strictly differentiable at $x_0 \in E$. In this case (see [Ro3], [Ro5]), $\nabla f(x_0)y = f^A(x_0;y) = f^K(x_0;y) = f^k(x_0;y) = f^\dagger(x_0;y) = f^0(x_0;y) = f^E(x_0;y) = f^L(x_0;y)$, and $\nabla f(x_0) = \partial^R f(x_0)$ for each of the R subgradients corresponding with these directional derivatives. This fact is essential, since the goal of subdifferential calculus, is, after all, the extension of the differential calculus.

We next consider simplifications of (1.5.10) which can be obtained when f is not necessarily differentiable but satisfies some continuity property.

<u>Proposition 1.5.13</u>: Let $f: E \to \overline{\mathbb{R}}$ be l.s.c. at $x_0 \in \operatorname{dom} f$, and let R be a q-cone with $\# := \Xi$ and $K(X) = X$. Then (1.5.10) becomes

$$(1.5.18) \qquad f^R(x_0;y) = \mathop{\square}_{Y\in N(y)} \ \inf_{\substack{X\in N(x_0)\\ u>0}} \ \mathop{O}_{\lambda>0} \ \inf_{Z\in M(Y)}$$

$$\sup_{\substack{x\in X\\ f(x)\le f(x_0)+u}} \ \mathop{O'}_{t\in(0,\lambda)} \ \mathop{\square'}_{y'\in Y\cap Z} \ \frac{f(x+ty') - f(x)}{t} .$$

<u>Proof</u>: Denote the right hand side of (1.5.18) by S, and suppose $f^R(x_0;y) \le d$. Then for all $\delta > 0$, $*\ Y \in N(y)$, there exists $X \in N(x_0)$ and $u > 0$, $\$\ \lambda > 0$, there exists $Z \in M(Y)$, for all $(x,r) \in X \times (f(x_0)-u, f(x_0)+u) \cap \operatorname{epi} f$, $\$'\ t \in (0,\lambda)$, $*'\ y' \in Z$, $\frac{f(x+ty')-r}{t} \le d + \delta$. Choose $X_1 \subset X$, if necessary, such that $f(x) \ge f(x_0) - u$ for all $x \in X_1$. Then for any $y' \in Z$ and $t \in (0,\lambda)$, (x,r) satisfies "$x \in X_1$, $r \in (f(x_0)-u, f(x_0)+u)$, $f(x) \le r$, implies $\frac{f(x+ty')-r}{t} \le d + \delta$" if and only if x satisfies "$x \in X_1$, $f(x) \le f(x_0) + u$ implies $\frac{f(x+ty')-f(x)}{t} \le d + \delta$". Thus $f^R(x_0;y) \le d$ if and only if $S \le d$, and (1.5.18) holds. $\square$

<u>Corollary 1.5.14</u> ([Ro3], [Ro5]): Let $f: E \to \overline{\mathbb{R}}$ be l.s.c. at $x_0 \in \operatorname{dom} f$. Then

$$(1.5.19) \qquad f^{\uparrow}(x_0;y) = \sup_{Y\in N(y)} \inf_{\substack{X\in N(x_0)\\ \mu>0\\ \lambda>0}} \sup_{\substack{x\in X\\ f(x)\le f(x_0)+\mu\\ t\in(0,\lambda)}} \inf_{y'\in Y} \frac{f(x+ty')-f(x)}{t} .$$

<u>Proposition 1.5.15</u>: Let $f: E \to \overline{\mathbb{R}}$ be continuous at x_0, and suppose R is a q-cone with $\# := \exists$ and $K(X) = X$. Then (1.5.10) reduces to

$$(1.5.20) \qquad f^{R}(x_0;y) = \Box_{Y\in N(y)} \inf_{X\in N(x_0)} O_{\lambda>0} \inf_{Z\in M(Y)} \sup_{x\in X} O'_{t\in(0,\lambda)} \Box'_{y'\in Y\cap Z} \frac{f(x+ty')-f(x)}{t} .$$

<u>Proof</u>: As usual, denote the right-hand side of (1.5.20) by S, and suppose $f^R(x_0;y) \le d$. Then for all $\delta > 0$, $*\ Y \in N(y)$, there exist $X \in N(x_0)$ and $\mu > 0$, $\$\ \lambda > 0$, $\exists\ Z \in M(Y)$, for all $(x,r) \in X \times (f(x_0)-\mu, f(x_0)+\mu) \cap \operatorname{epi} f$, $\$'\ t \in (0,\lambda)$, $*'\ y' \in Z$, $\frac{f(x+ty')-r}{t} \le d + \delta$. Since f is continuous, we may assume, by choosing a smaller neighborhood X if necessary, that $f(x) \in (f(x_0)-\mu, f(x_0)+\mu)$ for all $x \in X$. Then for any $y' \in Z$ and $t \in (0,\lambda)$, (x,r) satisfies "$x \in X$, $r \in (f(x_0)-\mu, f(x_0)+\mu)$, $f(x) \le r$ implies $\frac{f(x+ty')-r}{t} \le d + \delta$" if and only if x satisfies

"$x \in X$ implies $\frac{f(x+ty')-f(x)}{t} < d + \delta$". Thus $f^R(x_0;y) < d$ if and only if $S \le d$, and (1.5.20) holds. □

Corollary 1.5.16: Let $f: E \to \bar{\mathbb{R}}$ be continuous at x_0. Then $f^0(x_0;\cdot)$ can be expressed as in (1.5.16) and

$$(1.5.21) \quad f^{\uparrow}(x_0;y) = \sup_{Y \in N(y)} \inf_{\substack{X \in N(x_0) \\ \lambda > 0}} \sup_{\substack{x \in X \\ t \in (0,\lambda)}} \inf_{y' \in Y} \frac{f(x+ty')-f(x)}{t}$$

Definition 1.5.17: Let E be a normed space. The function $f: E \to \bar{\mathbb{R}}$ is said to be locally Lipschitzian near $x_0 \in E$ if there exist $X \in N(x_0)$ and a constant $K > 0$ such that $|f(x)-f(x')| \le K||x-x'||$ whenever $x,x' \in X$.

Proposition 1.5.18: Let E be a normed space, and suppose $f: E \to \bar{\mathbb{R}}$ is locally Lipschitzian near $x_0 \in E$. Let R be a q-cone with $* := \forall$, $M(Y) = Y$, $\# := \exists$ and either $K(X) = X$ or $K(X) = x_0$. Then

$$(1.5.22) \quad f^R(x_0;y) = \inf_{X \in N(x_0)} \ \underset{\lambda>0}{0} \ \inf_{W \in K(X)} \ \sup_{x \in X \cap W} \ \underset{t \in (0,\lambda)}{0'} \ \frac{f(x+ty)-f(x)}{t} .$$

Proof: Denote the right hand side of (1.5.22) by S. It is clear that $f^R(x_0;y) \le S$. Suppose $f^R(x_0;y) \le d$. Then for all $\delta > 0$ and $Y \in N(y)$, there exists $X \in N(x_0)$, \$ $\lambda > 0$, there exists $W \in K(X)$ such that for all $x \in W$, \$' $t \in (0,\lambda)$, there exists $y' \in Y$ with $\frac{f(x+ty')-f(x)}{t} \leq d + \frac{\delta}{2}$. (For the case $K(X) = X$, this characterization of "$f^R(x_0;y) \leq d$" relies on Proposition 1.5.15 and the fact that f is continuous at x_0.) Since f is locally Lipschitzian near x_0, there exist $X_1 \in N(x_0)$ and $K > 0$ such that for all x, $x' \in X_1$, $|f(x)-f(x')| \leq K||x-x'||$. We may choose $X \in N(x_0)$ to be smaller, if necessary, so that there exist $\lambda_1 > 0$ and $Y_1 \in N(y)$ with $X + [0,\lambda_1)Y_1 \subset X_1$ and $||y'-y''|| \leq \frac{\delta}{2K}$ for all $y',y'' \in Y_1$.

Now let $\delta > 0$ and $Y \in N(y)$ be given. There exist $X \in N(x_0)$, \$ $\lambda > 0$, there exists $W \in K(X)$, for all $x \in X \cap W$, \$' $t \in (0,\min(\lambda,\lambda_1))$, there exists $y' \in Y \cap Y_1$ with

$$\begin{aligned}\frac{f(x+ty)-f(x)}{t} &\leq \frac{f(x+ty') + t\,\frac{\delta}{2} - f(x)}{t} \\ &\leq \frac{f(x'+ty)-f(x)}{t} + \frac{\delta}{2} \\ &\leq r + \delta .\end{aligned}$$

Thus $f^R(x_0;y) \leq d$ implies that $S \leq d$, and (1.5.22) holds. □

Corollary 1.5.19: Let E be a normed space, and suppose $f: E \to \bar{\mathbb{R}}$ is locally Lipschitzian near $x_0 \in E$. Then $f^{\uparrow}(x_0;\cdot) = f^0(x_0;\cdot)$, $f^{T*}(x_0;\cdot) = f^{H*}(x_0;\cdot)$,

$f_+(x_0;\cdot) = f^A(x_0;\cdot)$, and $f^k(x_0;\cdot) = f^E(x_0;\cdot)$.

Remark 1.5.20: (a) We will see in section 1.7 that if E is a Banach space and $f\colon E \to \overline{\mathbb{R}}$ is locally Lipschitzian near $x_0 \in E$, then $f^{T*}(x_0;\cdot) = f^{\dagger}(x_0;\cdot)$ in addition to the relationships listed in Corollary 1.5.19.

(b) An examination of the proof of Proposition 1.5.18 shows that (1.5.22) will also hold if $* := \exists$ and $M(Y) = Y$, and so $f^+(x_0;\cdot) = f^E(x_0;\cdot)$, $f^I(x_0;\cdot) = f^0(x_0;\cdot)$ and $f^{I*}(x_0;\cdot) = f^{H*}(x_0;\cdot)$ under the hypotheses of Corollary 1.5.19. Moreover, we will see in section 1.7 that $f^I(x_0;\cdot) = f^{I*}(x_0;\cdot) = f^0(x_0;\cdot) = f^{H*}(x_0;\cdot)$.

(c) The assumption that $\# := \exists$ in Propositions 1.5.15 and 1.5.18 can also be dropped, but little of interest can be deduced from the case $\# := \forall$ and $K(X) = X$. We will see (as was mentioned in section 1.3) that such q-cones are too "large" to be useful in applications.

For our applications to optimization in Chapter 5, we are interested in q-cones R for which $0 \in \partial^R f(x_0)$ whenever $x_0 \in E$ is a local minimum of $f\colon E \to \overline{\mathbb{R}}$. We now determine which q-cone subgradients satisfy such a condition.

Theorem 1.5.21: Suppose $f\colon E \to \overline{\mathbb{R}}$ has a local minimum at $x_0 \in E$. Then

$$0 \in \partial^K f(x_0) . \tag{1.5.23}$$

Proof: Let $y \in E$ be given. Since x_0 is a local minimum for f, there exists $X \in N(x_0)$ such that $f(x) \geq f(x_0)$ for all $x \in X$. Choose $Y_1 \in N(y)$ and $\lambda_1 > 0$ such that $x_0 + [0,\lambda_1)Y \subset X$. Then for all $t \in (0,\lambda)$ and $y' \in Y$, $f(x_0+ty') - f(x_0) \geq 0$, and so $\frac{f(x_0+ty')-f(x_0)}{t} \geq 0$. Since y was arbitrarily chosen, it follows that

$$f_+^\uparrow(x_0;y) = \sup_{\substack{Y \in N(y) \\ \lambda > 0}} \inf_{\substack{y' \in Y \\ t \in (0,\lambda)}} \frac{f(x_0+ty')-f(x_0)}{t} \geq 0$$

for all $y \in E$. Thus (1.5.23) holds. □

Corollary 1.5.22: Suppose $f\colon E \to \bar{\mathbb{R}}$ has a local minimum at $x_0 \in E$. Then

$$(1.5.24) \qquad 0 \in \partial^R f(x_0)$$

for any q-cone R which is a subset of the contingent cone for all $(C,x_0) \in 2^E \times E$. In particular, (1.5.24) holds for the q-cones $A_C(x_0)$, $E_C(x_0)$, $k_C(x_0)$, $T_C(x_0)$, $T_C^*(x_0)$, $H_C(x_0)$, $H_C^*(x_0)$, $I_C(x_0)$, $I_C^*(x_0)$, $\ell_C(x_0)$, and $L_C(x_0)$.

Proof: Suppose R is a q-cone such that $R(C,x_0) \subset K_C(x_0)$ for all $(C,x_0) \in 2^E \times E$. It is not hard to see that $\hat{R}(f,x_0)$ is then contained in $K_{\text{epi } f}(x_0,f(x_0))$, and so

$0 \le f_+(x_0;y) \le f^R(x_0;y)$ for all $y \in E$ by Remark 1.5.5(d).
Hence (1.5.24) holds. □

Remark 1.5.23: If f is convex, the condition $0 \in \partial f(x_0)$ is sufficient as well as necessary for optimality. This will not be true in general of condition (1.5.24). The best we can hope for in general is that (1.5.24) be a necessary condition for optimality. (See [Ro5, Chapter 5] for more discussion and examples.)

Example 1.5.24: Condition (1.5.24) is not even a necessary condition for minimality for q-cones with $\# := \forall$ and $K(X) = X$. For instance, consider the q-cone $G_C(x_0)$ and the function $f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = |x|$. It is not hard to verify that $\hat{G}(f,0) = \mathbb{R}^2$, so that $f^G(0,y) = -\infty$ for all $y \in \mathbb{R}$, and $\partial^R f(0) = \emptyset$. As was mentioned in section 1.3, $G_C(x_0)$ is the "smallest" of the q-cones with $\# = \forall$ and $K(X) = X$, so $\partial^R f(0) = \emptyset$ for all other such q-cones.

1.6. Convex hulls of tangent cones

If a tangency operator or tangent cone has a certain property, will the closure of its convex hull possess the same property? There are important reasons for asking this question. In particular, suppose we have a result involving the contingent cone. If replacing the contingent cone by the pseudotangent cone would strengthen the result, the answer to the above question is of interest.

For most of the properties we have examined in Chapter 1, the answer to this question is quite easy. In this short section, we compile the answers for the various properties we have studied so far.

Properties (1) and (2) are clearly preserved when closures of convex hulls are taken, since cl cone$(C-x_0)$ is convex for C convex (and is of course always closed) and since $\nabla h(x_0)^{-1}(0)$ is closed and convex. If $C_1 \subseteq C_2 \subset E$, then cl conv $C_1 \subset$ cl conv C_2, so the closure of the convex hull of an isotone tangency operator is still isotone. The answer for property (4) is immediate - the closure of the convex hull of any tangency operator is convex since closures of convex sets are convex.

Properties (5) and (6) are not preserved by closures of convex hulls, however. A counterexample for property (6) was given in section 1.1. Here is a counterexample for property (5):

Example 1.6.1: In $\mathbb{R}^2$, define $C_1 := \{(x,y) \mid \frac{1}{2}x \le y \le 2x\}$ and $C_2 := \{(x,y) \mid y \le \frac{1}{2}x \text{ or } y \ge 2x\}$, and let $x_0 := (0,0)$. Then $L_{C_1}(x_0) = \text{int } C_1$, and $L_{C_2}(x_0) = \text{int } C_2$, while cl conv $L_{C_1}(x_0) = C_1$ and cl conv $L_{C_2}(x_0) = \mathbb{R}^2$. Thus cl conv $L_{C_1}(x_0) \cap$ cl conv $L_{C_2} = C_1$, but $L_{C_1 \cap C_2}(x_0) = \emptyset$. The tangent cone $L_C(x_0)$ has property (5), but its convex hull and closure of its convex hull do not.

This example illustrates a major drawback in achieving convexity of a tangent cone by simply taking convex hulls. The inclusion in property (5) is very important in applications, and it is not easy to establish whether the convex hull of a tangent cone satisfies it.

<u>Proposition 1.6.2</u>: Suppose R is a tangent cone which is product-preserving. Then $\text{cl conv } R$ is also product-preserving.

<u>Proof</u>: Suppose $x_1 \in C_1 \subset E$ and $x_2 \in C_2 \subset F$, and let $y \in \text{conv } R(C_1 \times C_2, (x_1,x_2))$. Then there exist $p > 0$, $d_i \in R(C_1 \times C_2, (x_1,x_2))$ and $\lambda_i \geq 0$, $i = 1,\ldots,p$, with $\sum_{i=1}^{p} \lambda_i = 1$ and $y = \sum_{i=1}^{p} \lambda_i d_i$. By (1.4.1), each $d_i = (a_i, b_i)$ with $a_i \in R(C_1,x_1)$ and $b_i \in R(C_2,x_2)$, so $y = (\sum_{i=1}^{p} \lambda_i a_i , \sum_{i=1}^{p} \lambda_i b_i) \in \text{conv } R(C_1,x_1) \times \text{conv } R(C_2,x_2)$.
Hence $\text{conv } R(C_1 \times C_2, (x_1,x_2)) \subset \text{conv } R(C_1,x_1) \times \text{conv } R(C_2,x_2)$.
Conversely, suppose $y_1 \in \text{conv } R(C_1,x_1)$ and $y_2 \in \text{conv } R(C_2,x_2)$. Then $y_1 = \sum_{i=1}^{n} \lambda_i a_i$ for some $n > 0$, $a_i \in R(C_1,x_1)$, $\lambda_i > 0$, $i = 1,\ldots,n$, $\sum_{i=1}^{n} \lambda_i = 1$, and $y_2 = \sum_{i=1}^{m} \lambda_i b_i$ for some $m > 0$, $b_i \in R(C_2,x_2)$, $\mu_i > 0$,

$\sum_{i=1}^{m} \mu_i = 1$. Assume without loss of generality that $m \le n$. Let $\mu_i = 0$, $m+1 \le i \le n$, and let b_i, $m+1 \le i \le n$, be arbitrary elements of $R(C_2,x_2)$. Then $y_2 = \sum_{i=1}^{n} \mu_i b_i$ and $(y_1,y_2) = \sum_{i=1}^{n} \lambda_i (a_i, \frac{\mu_i}{\lambda_i} b_i) \in \operatorname{conv}(R(C_1,x_1) \times R(C_2,x_2))$, since $R(C_2,x_2)$ is a cone. Thus $\operatorname{conv} R(C_1,x_1) \times \operatorname{conv} R(C_2,x_2) \subset \operatorname{conv}(R(C_1,x_1) \times R(C_2,x_0))$ $= \operatorname{conv} R(C_1 \times C_2, (x_1,x_2))$ by (1.4.1). Hence (1.4.1) holds for $\operatorname{conv} R$, and so it also must hold for $\operatorname{cl} \operatorname{conv} R$. □

<u>Proposition 1.6.3</u>: Suppose $g: E \to F$ is strictly differentiable at $x_0 \in C \subset E$, and suppose R is a tangent cone. Assume (1.4.11) holds. Then

(1.6.1) $\quad \nabla g(x_0) \operatorname{cl} \operatorname{conv} R(C,x_0) \subset \operatorname{cl} \operatorname{conv} R(g(C),g(x_0))$.

In particular,

(1.6.2) $\quad \nabla g(x_0) P_C(x_0) \subset P_{g(C)}(g(x_0))$:

<u>Proof</u>: By (1.4.11), $\nabla g(x_0) R(C,x_0) \subset R(g(C),g(x_0))$, so $\nabla g(x_0) \operatorname{conv} R(C,x_0) = \operatorname{conv} \nabla g(x_0) R(C,x_0) \subset \operatorname{cl} \operatorname{conv} R(g(C),g(x_0))$, and hence $\nabla g(x_0) \operatorname{cl} \operatorname{conv} R(C,x_0) \subset \operatorname{cl} \operatorname{conv} R(g(C),g(x_0))$. Assertion (1.6.2) follows from (1.6.1) and Corollary 1.4.8. □

The inclusion (1.6.2) is well known - see for example [Bo3].

We finish this section by showing that the necessary condition for minimality in (1.5.24) will hold for cl conv R whenever it holds for R.

Definition 1.6.4 ([Rol, Sections 5 and 7]): Let $f: E \to \overline{\mathbb{R}}$ be an extended real-valued function on E.

(i) The convex hull of f, denoted conv f, is the function whose epigraph is the convex hull of the epigraph of f.

(ii) The closure of f, denoted cl f, is the function whose epigraph is the closure of the epigraph of f.

(iii) Suppose R is a q-cone and $x_0 \in \operatorname{dom} f$. Define

$$(1.6.3) \qquad f^{\overline{co}\,R}(x_0;\cdot) := \operatorname{cl}\operatorname{conv} f^R(x_0;\cdot) .$$

Proposition 1.6.5: Suppose R is a q-cone, $f: E \to \overline{\mathbb{R}}$, and $x_0 \in \operatorname{dom} f$. Then

$$(1.6.4) \quad \operatorname{epi}(f^{\overline{co}\,R}(x_0;\cdot)) = \operatorname{cl}\operatorname{conv} \hat{R}(f,x_0) .$$

Proof: By Definition 1.6.4(iii),

$$\operatorname{epi}(f^{\overline{co}\,R}(x_0;\cdot)) = \operatorname{cl}\operatorname{conv} \hat{R}(f,x_0) ,$$

and by Remark 1.5.5(e),

$$\operatorname{epi} f^R(x_0;\cdot) = \hat{R}(f,x_0) .$$

Combining these facts gives (1.6.4). □

<u>Proposition 1.6.6</u>: Suppose R is a q-cone, and suppose x_0 is a local minimizer of $f: E \to \overline{\mathbb{R}}$. If condition (1.5.24) is satisfied for $\partial^R f$, then

$$(1.6.5) \quad 0 \in \partial^{\overline{co}\, R} f(x_0) := \{x' \in E' \mid \langle y, x' \rangle \le f^{\text{cl conv } R}(x_0; y) \text{ for all } y \in E\} .$$

<u>Proof</u>: Since $0 \in \partial^R f(x_0)$, it follows that $r \ge 0$ for every $(y,r) \in \hat{R}(f,x_0)$. Thus $r \ge 0$ for every $(y,r) \in \text{cl conv } \hat{R}(f,x_0)$, and $0 \in \partial^{\overline{co}\, R} f(x_0)$ by Proposition 1.6.5. □

<u>Corollary 1.6.7</u>: Suppose R is a q-cone such that $R(C,x_0) \subset K(C,x_0)$ for all $(C,x_0) \in 2^E \times E$, and suppose x_0 is a local minimizer for $f: E \to \overline{\mathbb{R}}$. Then (1.6.5) holds. In particular,

$$(1.6.6) \qquad 0 \in \partial^P f(x_0) := \partial^{\overline{co}\, K} f(x_0) .$$

<u>Proof</u>: This follows immediately from Proposition 1.6.6 and Corollary 1.5.22. □

1.7. Alternate characterizations of tangent cones.

Earlier in this chapter, we investigated the effect of changing $\$$ from $\forall$ to $\exists$ in the definition of $K_C(x_0)$. The tangent cone resulting from this change in quantification, $k_C(x_0)$, is product-preserving, unlike $K_C(x_0)$. On the other hand, $k_C(x_0)$ lacks property (6), a property possessed by the contingent cone.

In this section, we examine the consequences of some other slight changes in quantification in the definitions of q-cones. We begin by discussing the effect of changing $\$$ from $\exists$ to $\forall$.

If we change $\$$ from $\exists$ to $\forall$ in the definition of $T_C(x_0)$, we obtain the cone $T_C^*(x_0)$, defined in Corollary 1.4.11. Treiman [Tr1, Lemma 2.1] and Penot [Pe3, proof of Theorem 1] have shown that $T_C(x_0) = T_C^*(x_0)$ when E is a Banach space, $C \subset E$ is closed, and $x_0 \in C$. (This result has important consequences that we will discuss in Chapter 4.) An examination of Treiman's proof shows that a similar result holds for other q-cones with $\# := \exists$ and $K(X) = X$. We prove this in Propositions 1.7.1 and 1.7.4.

<u>Proposition 1.7.1:</u> Suppose R is a q-cone with $\# := \exists$, $K(X) = X$, $* := \forall$, $\$:= \exists$ and with $M(\cdot)$ such that for any closed, bounded, and convex $Y \in N(y)$ and any $Z \in M(Y)$, Z is closed and convex. Then if $x_0 \in C$, a closed subset of a Banach space E,

$$(1.7.1)\quad R(C,x_0) = R^*(C,x_0) := \{y \in E \mid \forall\, Y \in N(y)\ ,\ \exists\, X \in N(x_0)\ ,\ \forall\, \lambda > 0\ ,\ \exists\, Z \in M(Y)\ ,\ \forall\, x \in X \cap C\ ,\ \exists\, t \in (0,\lambda)\ ,\ \exists\, y' \in Y \cap Z\ ,\ x + ty' \in C\}\ .$$

Proof: The inclusion $R(C,x_0) \subset R^*(C,x_0)$ is immediate from the definitions of R and R^*. Suppose $y \notin R(C,x_0)$. Then there exists $Y \in N(y)$ such that for all $X \in N(x_0)$, $\lambda > 0$ and $Z \in M(Y)$, there exist $x \in X$ and $t \in (0,\lambda)$ with $(x+tZ) \cap C = \emptyset$. Replacing Y by a smaller neighbourhood if necessary, we may assume that Y is closed, bounded, and convex. We may also assume X is convex. Let such an X be given, along with some $Z \in M(Y)$, and choose $x \in X \cap C$ and $\lambda > 0$ such that $x + \lambda Y \subset X$ and

$$(x + \lambda Z) \cap C = \emptyset\ .$$

As in the proof of [Tr1, Lemma 2.1], order the points of $A := [C \cap (x + [0,\lambda]Z)^{\circ}]$ as follows:
For any $x_1, x_2 \in A$, $x_2 \geq x_1$ if and only if $x_2 \in x_1 + tZ$ for some $t \geq 0$. Since Z is convex by hypothesis, this order relation is transitive. Now define a sequence $\{x_m\}_{m=1}^{\infty}$ inductively. Let $x_1 := x$. Given x_m, let $\alpha_m := \sup\{\alpha \in \mathbb{R} \mid (x_m + \alpha Z) \cap A \neq \emptyset\}$. Choose x_{m+1} such that $x_{m+1} \in (x_m + \beta_m Z) \cap A$ where $\alpha_m - \beta_m \leq 2^{-(m+2)}$. Thus if $z \geq x_m$, $z \in x_m + \beta Z$ with $\beta \leq 2^{-(m+1)}$. Let $d = \sup\{\|y'\| \mid y' \in Y\}$. Then for all $z \geq x_m$, $\|z - x_m\| \leq d2^{-(m+1)}$. Since $x_k \geq x_m$ for any $k \geq m$,

$\{x_m\}_{m=1}^{\infty}$ is Cauchy, and since E is complete, the sequence has a limit $\bar{x}$. Since C and $x + [0,\lambda]Z$ are closed, $\bar{x} \in C$ and $\bar{x} \in x + tZ$ for some $t \in [0,\lambda)$. Now if $z \geq \bar{x}$, $z \geq x_m$ for all m, and

$$\begin{aligned} ||z-\bar{x}|| &\leq ||z-x_m|| + ||x_m-\bar{x}|| \\ &\leq d2^{-(m+1)} + d2^{-(m+1)} = d2^{-m} \end{aligned}$$

for all m. Hence $z = \bar{x}$, and $\bar{x}$ is thus maximal under $\leq$. We conclude that for some $r > 0$, $(\bar{x} + (0,r)Z) \cap C = \emptyset$. Therefore $y \notin R^*(C,x_0)$, and so $R^*(C,x_0) \subset R(C,x_0)$ and (1.7.1) holds. □

<u>Corollary 1.7.2</u>: Suppose $x_0 \in C$, a closed subset of a Banach space E. Then $T_C(x_0) = T_C^*(x_0)$ and $H_C(x_0) = H_C^*(x_0)$. If R in Proposition 1.7.1 has $M(Y)$ equal to the class of all non-empty convex compact subsets of Y, $R(C,x_0) = R^*(C,x_0)$.

<u>Corollary 1.7.3</u>: Suppose E is a Banach space, $f: E \to \bar{\mathbb{R}}$ is l.s.c., and $x_0 \in \text{dom}\, f$. Then for all $y \in E$,

$$(1.7.2) \qquad f^{\uparrow}(x_0;y) = \sup_{Y \in \bar{N}(y)} \inf_{\substack{X \in N(x_0) \\ \delta > 0}} \sup_{\substack{\lambda > 0 \\ f(x) \leq f(x_0)+\delta}} \inf_{t \in (0,\lambda)} \inf_{y' \in Y} \frac{f(x+ty')-f(x)}{t}.$$

If f is continuous at x_0, then for all $y \in E$,

$$(1.7.3) \quad f^0(x_0;y) = \inf_{X \in N(x_0)} \sup_{\lambda>0} \sup_{x \in X} \inf_{t \in (0,\lambda)} \frac{f(x+ty)-f(x)}{t} .$$

Proof: In either case, epi f is a closed subset of the Banach space $E \times \mathbb{R}$. Then $T_{\text{epi } f}((x_0,f(x_0))) = T^*_{\text{epi } f}((x_0,f(x_0)))$, and so $f^\uparrow(x_0;\cdot) = f^{T^*}(x_0;\cdot)$. The expression in (1.7.2) is that for $f^{T^*}(x_0;y)$ for an l.s.c. function f by (1.5.18). Similarly, $H_{\text{epi } f}((x_0,f(x_0))) = H^*_{\text{epi } f}((x_0,f(x_0)))$, so $f^0(x_0;\cdot) = f^{H^*}(x_0;\cdot)$ and (1.7.3) follows from (1.5.20). □

Proposition 1.7.4: Suppose R is a q-cone with $\# := \exists$, $K(X) = X$, $* := \exists$, and $\$:= \exists$. Then if C is a closed subset of a l.c.s. and $x_0 \in C$,

$$(1.7.4) \quad R(C,x_0) = R^*(C,x_0) = \{y \in E \mid \exists\, Y \in N(y),\ \exists\, X \in N(x_0),$$
$$\forall\, \lambda > 0,\ \exists\, Z \in M(Y),\ \forall\, x \in X \cap C,\ \exists\, t \in (0,\lambda),$$
$$\forall\, y' \in Y \cap Z,\ x + ty' \in C\} .$$

Proof: The inclusion $R(C,x_0) \subset R^*(C,x_0)$ is immediate from the definition of R and R^*. Suppose $y \notin R(C,x_0)$. Then for all $Y \in N(y)$, $X \in N(x_0)$, $\lambda > 0$ and $Z \in M(Y)$, there exist $y' \in Z$, $x \in X \cap C$ and $t \in (0,\lambda)$ such that $x + ty' \notin C$. So let $Y \in N(y)$, $X \in N(x_0)$ and $Z \in M(Y)$ be given. We may again assume that X is convex.

Choose $x \in X \cap C$, $y' \in Z$ and $\lambda > 0$ with $x + \lambda y' \in X$ and $x + \lambda y' \notin C$. Order the points of $A := C \cap [x + [0,\lambda]y']$ as follows: For any $x_1, x_2 \in A$, $x_2 \geq x_1$ if and only if $x_2 = x_1 + ty'$ for some $t \geq 0$. This order relation is again transitive. Arguing as in Proposition 1.7.1, we conclude that there exists $\bar{x} \in X$ and $r > 0$ with

$$(\bar{x} + (0,r)y') \cap C = \emptyset .$$

We can reach this conclusion whether or not E is normed or complete, since our Cauchy sequence in this case is actually a sequence of real numbers. Therefore $y \notin R^*(C,x_0)$ and so $R(C,x_0) \subset R^*(C,x_0)$ and (1.7.4) holds. □

Corollary 1.7.5: Suppose C is a closed subset of E and $x_0 \in C$. Then $I_C(x_0) = I^*_C(x_0)$ and $H_C(x_0) = H^*_C(x_0)$. Moreover, (1.7.3) holds in any normed space.

Remark 1.7.6: (a) An examination of the proof of Proposition 1.7.4 shows that in fact

$$(1.7.5) \quad R(C,x_0) = R^*(C,x_0) = \{y \in E \mid \exists\ Y \in N(y),\ \exists\ X \in N(x_0),\ \forall\ \lambda > 0,\ \forall\ x \in X \cap C,\ \forall\ y' \in Y \cap Z,\ \exists\ t \in (0,\lambda),\ x + ty' \in C\},$$

a slightly stronger conclusion. In (1.7.4), the same $t \in (0,\lambda)$ must "work" uniformly for $y' \in Y \cap Z$, while in (1.7.5), a different t may be used for each $y' \in Y \cap Z$.

(b) If the set C is not closed, the tangent cones $R(C,x_0)$ and $R^*(C,x_0)$ will not in general be equal in either Proposition 1.7.1 or Proposition 1.7.4. For example, if $C = \mathbb{Q}$, $H_C(x_0) = 0$ and $H_C^*(x_0) = \mathbb{R}$ for any $x_0 \in \mathbb{R}$. If $C := \mathbb{R}^+ \backslash \{2^{-n} | n \geq 0\}$, $I_C(0) = \{0\}$ while $I_C^*(0) = \mathbb{R}^+$. If

$$C := \{\mathbb{Q} \cap [\bigcup_{n=1}^{\infty} (\frac{1}{2^{2n}\sqrt{2}}, \frac{1}{2^{2n-1}\sqrt{2}})]\} \cup \{0\},$$

$$T_C(0) = \{0\} \text{ while } T_C^*(0) = \mathbb{R}^+ .$$

(c) In $\mathbb{R}^2$, there are nonclosed sets which show that H^*, T^* and I^* are not convex or product-preserving in general. For $C := \mathbb{Q} \times \mathbb{Q}$, $H_C^*(0,0) = (\mathbb{Q} \times \mathbb{Q}) \cup ((\mathbb{R}\backslash\mathbb{Q}) \times (\mathbb{R}\backslash\mathbb{Q}))$, which is not convex and not equal to $H_{\mathbb{Q}}^*(0) \times H_{\mathbb{Q}}^*(0) = \mathbb{R}^2$. For $C := C_1 \times C_2$, where $C_1 := \bigcup_{n=0}^{\infty} (\frac{1}{2^{4n+1}}, \frac{1}{2^{4n}}) \cup \{0\}$ and $C_2 := \bigcup_{n=0}^{\infty} (\frac{1}{2^{4n+3}}, \frac{1}{2^{4n+2}}) \cup \{0\}$, $T_C^*((0,0))$ is not convex. Points of the form $(y,sy) \in T_C^*((0,0))$ for $s \in \{0,[2,8],[32,128],\ldots,[2^{4n+1},2^{4n+3}]\}$, but not for any other positive values of s. Thus $(1,4)$ and $(1,64)$ are in $T_C^*((0,0))$, while $\frac{3}{4}(1,4) + \frac{1}{4}(1,64) = (1,19) \notin T_C^*((0,0))$. This example also shows that T^* is not product-preserving, since $T_{C_1}^*(0) = T_{C_2}^*(0) = \mathbb{R}$. This same example demonstrates that I^* is neither convex nor product-preserving in general.

(d) Examples are given in [Bo8] of closed sets in incomplete spaces for which $T_C^*(x_0) \neq T_C(x_0)$ and $T_C^*(x_0)$ is not convex.

Another consequence of Proposition 1.7.4 is the following variation of [Ro2, Theorem 2]:

<u>Corollary 1.7.7</u>: Suppose C is a closed subset of $\mathbb{R}^n$ and $x_0 \in C$, and suppose $\operatorname{int} T_C(x_0) \neq \emptyset$. Then $\operatorname{int} T_C(x_0) = I_C^*(x_0)$.

<u>Proof</u>: By [Ro2, Theorem 2], $\operatorname{int} T_C(x_0) = I_C(x_0)$, so by Proposition 1.7.4, $\operatorname{int} T_C(x_0) = I_C^*(x_0)$. $\square$

Our definition of g-cone could have been expanded to include tangency operators of the form

$$\begin{aligned}(1.7.6)\qquad R'(C,x_0) := \{y \in E \mid {*}\ Y \in N(y),\ \#\ X \in N(x_0),\ \$\ \lambda > 0,\\ \exists\ W \in K(X),\ \exists\ Z \in M(Y),\ \$'\ t \in (0,\lambda),\\ \#'\ x \in X \cap C \cap W,\ {*}'\ y' \in Y \cap Z,\ x + ty' \in C\}.\end{aligned}$$

An examination of the proofs in sections 1.3 through 1.6 shows that the results in these sections still hold for R' defined as in (1.7.6). Including cones of the form (1.7.6) in our discussion adds a few new tangent cones to our collection; however, none of these new cones is of any great import. Still, it is of some interest to compare cones of the form (1.7.6) with closely related ones of the form (1.3.1). We consider here three examples, beginning with

$$(1.7.7) \qquad H_C^{!}(x_0) := \{y \in E \mid \exists\, X \in N(x_0),\ \forall\, \lambda > 0,\ \exists\, t \in (0,\lambda),\ \forall\, x \in X \cap C,\ x + ty \in C\}.$$

In (1.7.7), the same t must work uniformly for all $x \in X \cap C$, in contrast to the definition of $H_C^*(x_0)$, in which a different t may be used for each $x \in X \cap C$. It is clear that the inclusions

$$H_C(x_0) \subset H_C^{!}(x_0) \subset H_C^*(x_0)$$

always hold, and that these three tangent cones are equal whenever C is a closed subset of a l.c.s. (by Proposition 1.7.4 with $M(Y) = y$). There are nonclosed sets, however, for which these three cones are distinct. For example, in $\mathbb{R}^2$ let $C := Q^+ \times \mathbb{R}\backslash\{2^{-n} \mid n \geq 0\}$. Then the points (0,1) and (1,0) are in $H_C^*(x_0)$, while $(0,1) \notin H_C^{!}(x_0)$ and neither (0,1) nor (1,0) is in $H_C(x_0)$.

The example $C = Q \times Q$ from Remark 1.7.6(c) shows that $H_C^{!}$ is in general neither convex nor product-preserving, since for this example $H_C^{!}((0,0)) = H_C^*((0,0))$.

We next consider the tangent cone

$$(1.7.8)\quad I_C^*(x_0) := \{y \in E \mid \exists\, Y \in N(y),\ \exists\, X \in N(x_0),\ \forall \lambda > 0,\ \exists\, t \in (0,\lambda),\ \forall\, x \in X \cap C,\ \forall\, y' \in Y,\ x + ty' \in C\}\ .$$

Here again it is clear that

$$I_C(x_0) \subset I_C'(x_0) \subset I_C^*(x_0)\ ,$$

and that these three tangent cones are equal whenever C is a closed subset of a Banach space E. In fact, it turns out that $I_C(x_0)$ and $I_C'(x_0)$ are always the same.

<u>Proposition 1.7.8</u>: Suppose $x_0 \in C \subset E$. Then $I_C(x_0) = I_C'(x_0)$.

<u>Proof</u>: Suppose $y \notin I_C(x_0)$ and let $Y \in N(y)$, $X \in N(x_0)$ and $\lambda > 0$. We may assume without loss of generality that Y is closed and convex. Let $X_1 \in N(x_0)$ and $\lambda_1 \in (0,\lambda)$ be such that $X_1 + (0,\lambda_1)Y \subset X$. Choose $Y_1 \in N(y)$ such that there exists $\varepsilon > 0$ with $(1-\varepsilon,1+\varepsilon)y' \in Y$ whenever $y' \in Y_1$. There exist $x' \in X_1 \cap C$, $t_0 \in (0,\lambda_1)$ and $y' \in Y_1$ with $x' + t_0y' \notin C$. Define $\bar{t} := \sup\{t \mid t \le t_0,\ x' + ty' \in C\}$, and call $a := x' + \bar{t}y'$.

Case 1: Suppose $\bar{t} < t_0$. Then $a + (0,t_0-\bar{t})y' \cap C = \emptyset$. Let $t \in (0,t_0-\bar{t})$. There exists $\bar{x} \in X \cap C$ such that $\bar{x} = x' + sy'$ for some $s \in (\bar{t}-t,\bar{t}]$. Then $s+t \in (\bar{t},t_0)$, and $\bar{x}+ty' = x' + (s+t)y' \in a + (0,t_0-\bar{t})y'$, so $\bar{x}+ty' \notin C$.

Case 2: Suppose $\bar{t} = t_0$. There exists $\varepsilon > 0$ with $(1-\varepsilon,1+\varepsilon)y' \in Y$. Let $t \in (0, \frac{\bar{t}}{1+\varepsilon})$. Then $x' + (\bar{t}-t(1+\varepsilon), \bar{t}-t(1-\varepsilon))y' \subset X$. If there exists $\bar{x} \in [x'+\bar{t}-t(1+\varepsilon), \bar{t}-t(1-\varepsilon)y] \cap C$, then there exists $\bar{y} \in (1-\varepsilon,1+\varepsilon)y'$ with $\bar{x}+t\bar{y} \notin C$. If not, call $m := \sup\{s \mid x'+sy' \in C,\ s \le \bar{t}-t(1+\varepsilon)\}$. Call $b := x'+my'$. Then there exists $\lambda_0 \in (0,\lambda_1)$ such that $b + (0,\lambda_0)y' \cap C = \emptyset$. Let $t \in (0,\lambda_0)$. There exists $\bar{x} \in X \cap C$ such that $\bar{x} = x' + sy'$ for some $s \in (m-t,m]$. Then $s+t \in (m,m+\lambda_0)$, and $\bar{x}+ty' = x'+sy'+ty' \in b + (0,\lambda_0)y'$, so $\bar{x}+ty' \notin C$.

In either case, $y \notin I_C'(x_0)$ and so $I_C'(x_0) \subset I_C(x_0)$. Therefore $I_C(x_0) = I_C'(x_0)$. □

Finally we consider

$$(1.7.9) \qquad T_C'(x_0) := \{y \in E \mid \forall\, Y \in N(y),\ \exists\, X \in N(x_0),\ \forall\, \lambda > 0,\ \exists\, t \in (0,\lambda),\ \forall\, x \in X \cap C,\ \exists\, y' \in Y,\ x+ty' \in C\}.$$

Once more it is true in general that $T_C(x_0) \subset T_C'(x_0) \subset T_C^*(x_0)$, and Proposition 1.7.1 shows that these three tangent cones are equal whenever C is a closed subset of a Banach space.

If $C \subset \mathbb{R}$, it is not hard to establish (by an argument similar to that used in Proposition 1.7.8) that $T_C(x_0) = T'_C(x_0)$. Whether those two tangent cones are equal more generally is an open question.

There is only one q-cone with $K(X) = X$ and $M(Y) = Y$, that we have not yet mentioned. It is

$$G^*_C(x_0) := \{y \in E \mid \exists\, Y \in N(y),\ \forall\, X \in N(x_0),\ \forall\, \lambda > 0,$$

$$\exists\, x \in X \cap C,\ \exists\, t \in (0,\lambda),\ \forall\, y' \in Y,\ x + ty' \in C\}.$$

The reason we have not mentioned $G^*_C(x_0)$ is that it, like $G_C(x_0)$, is a very "trivial" tangent cone which can only take on the values $\emptyset$ and E. In fact, one can show that $G^*_C(x_0) = G_C(x_0)$ for all $(C,x_0) \in 2^E \times E$.

1.8. Weak tangent cones

One class of tangent cones we have not yet considered are the weak tangent cones defined in [Ba1] and [Bo3]. Although we will not deal with weak tangent cones in subsequent chapters, we will discuss them briefly in this section and define a similar concept which fits into our c-cone framework.

In this section, let E be a normed space. As in [Bo3], denote by s the norm topology on E and denote by τ another locally convex (Hausdorff) topology on E such that

(1.8.1) (i) τ is coarser than s and

(ii) τ-convergent sequences are s-bounded.

Let "$\to$" stand for s-convergence and "$\rightharpoonup(\tau)$" stand for τ-convergence.

Definition 1.8.1. [Bo3]: Let $x_0 \in E$ and $C \subset E$. The sequential τ-contingent cone of C at x_0 is the set

$$(1.8.2) \quad K_\tau(C,x_0) := \{y \in E \mid \exists \text{ a sequence } x_n \to x_0,\ x_n \in C,\ \exists\ \lambda_n > 0 \text{ such that } \lambda_n(x_n - x_0) \rightharpoonup (\tau) y\} .$$

The sequential τ-pseudotangent cone is the set

$$(1.8.3) \quad P_\tau(C,x_0) := \text{cl conv } K_\tau(c,x_0) .$$

We begin by deriving an equivalent form of (1.8.2) which is easier to compare with our q-cone definitions.

Proposition 1.8.2: An equivalent form of (1.8.2) is

$$(1.8.4) \quad K_\tau(C,x_0) = \{y \in E \mid \exists \text{ sequences } t_n \downarrow 0 \text{ and } y_n \rightharpoonup (\tau) y \text{ such that } x_0 + t_n y_n \in C\} .$$

Proof: Denote the right-hand side of (1.8.4) by S, and let $y \in K_\tau(C,x_0)$ with $y \neq 0$. Then there exists $x_n \in C$ with $x_n \to x_0$ and $\lambda_n > 0$ such that $\lambda_n(x_n - x_0) \rightharpoonup (\tau) y$. Since $y \neq 0$, we may assume λ_n is a nondecreasing sequence.

Set $y_n := \lambda_n(x_n - x_0)$ and $t_n := \lambda_n^{-1}$. Then $y_n \rightharpoonup (\tau) y$ and $x_0 + t_n y_n = x_n \in C$. Since $x_n \to x_0$, it follows that $t_n y_n \to 0$, and so $t_n \downarrow 0$. Hence $y \in S$.

Conversely, let $y \in S$. Then there exist $t_n \downarrow 0$ and $y_n \rightharpoonup (\tau) y$ such that $x_0 + t_n y_n \in C$. Set $x_n := x_0 + t_n y_n$ and $\lambda_n := t_n^{-1}$. Now

$$\lambda_n(x_n - x_0) = y_n \rightharpoonup (\tau) y ,$$

and since $\{y_n\}$ is s-bounded and $t_n \downarrow 0$, it follows that $x_n \to x_0$. Hence $y \in K_\tau(C, x_0)$. Noting that both sets contain 0, we have proven (1.8.4). □

<u>Remark 1.8.3</u>: It is well known that if E is a normed space,

$$(1.8.5) \qquad K_C(x_0) := \{y \in E \mid \exists\ t_n \downarrow 0,\ \exists\ y_n \to y \text{ with } x_0 + t_n y_n \in C\} .$$

Comparing (1.8.5) with (1.8.4), one sees immediately that $K_C(x_0) \subset K_\tau(C, x_0)$. It is shown in [Bo3] that the sequential τ-contingent and τ-pseudotangent cones satisfy inclusions analogous to (1.4.11) and (1.6.2) and can thus be used to sharpen the Guignard Kuhn-Tucker conditions [Gu1]. These tangent cones were defined with such applications in mind.

We now define and briefly discuss a similar tangent cone notion that fits into our q-cone framework. It has previously been examined in [Pel].

<u>Definition 1.8.4</u>: Let $x_0 \in E$ and $C \subset E$ and let $N_\tau(\cdot)$ denote τ-neighbourhood. The τ-contingent cone of C at x_0 is the set

(1.8.6) $K_C^\tau(x_0) := \{y \in E \mid \exists\ \underset{\sim}{Z} \subset E$ bounded, $\forall\ Y \in N_\tau(y)$, $\forall\ \lambda > 0$, $\exists\ t \in (0,\lambda)$, $\exists\ y' \in Y \cap Z$, $x_0 + ty' \quad C\}$.

The τ-contingent cone is not quite a q-cone since the order of its quantifications is different from that in (1.3.1). Clearly $K_C(x_0) \subset K_\tau(C,x_0) \subset K_C^\tau(x_0)$, and the three sets are equal if $\tau = s$.

Referring to results of sections 1.3 and 1.4, we can readily deduce a number of properties of $K_C^\tau(x_0)$. As in Proposition 1.3.3, it is a homogeneous tangent cone. It is not always a closed set - the proof of Theorem 1.3.4 does not work in this case. A direct argument shows that if C is convex and $x_0 \in C$, then

$$\text{cl cone}(C-x_0) = \tau - \text{cl cone}(C-x_0) = K_C^\tau(x_0) .$$

If $h = \mathbb{R}^n \to \mathbb{R}^m$ is strictly differentiable with $x_0 \subset h^{-1}(0)$ and $\nabla h(x_0)$ is of rank m , then $\nabla h(x_0)^{-1}(0) \subset K_{h^{-1}(0)} \subset K^\tau_{h^{-1}(0)}(x_0)$. However, $K_C^\tau(x_0)$ does not have property (2) since the inclusion $K_{h^{-1}(0)}(x_0) \subset \nabla h(x_0)^{-1}(0)$ does not necessarily hold. It is immediate as in Theorem 1.3.19 that $K_C^{\tau}(x_0)$ is isotone with respect to set inclusion. Like the contingent cone, $K_C^{\tau}(x_0)$ is not always convex, does not have property (5) and is not product-preserving. It is easy to see, arguing as in Theorem 1.3.26, that $K_C^{\tau}(x_0)$ has property (6).

We conclude this section by demonstrating that $K_C^{\tau}(x_0)$ satisfies generalizations of (1.4.11) and [Bo3, Theorem 1].

<u>Definition 1.8.5</u>: Let E and F be normed spaces and $x_0 \in E$. The function $g: E \to F$ is Frechet differentiable at x_0 if there exists a linear continuous $\nabla g(x_0): E \to F$ satisfying

$$(1.8.7) \qquad \lim_{t \downarrow 0} \frac{g(x_0+th)-g(x_0)}{t} = \nabla g(x_0)h ,$$

where the limit in (1.8.7) is approached uniformly through bounded subsets of E.

<u>Proposition 1.8.6</u>: Suppose $g: E \to F$ is Frechet differentiable at $x_0 \in C \subset E$ and $\nabla g(x_0)$ is $\tau_1 - \tau_2$ continuous. Then

$$(1.8.8) \qquad \nabla g(x_0) K_C^{\tau_1}(x_0) \subset K_{g(C)}^{\tau_2}(g(x_0)) .$$

<u>Proof</u>: Let $y \in K_C^{\tau_1}(x_0)$, $\lambda > 0$ and $U \in N_{\tau_2}(\nabla g(x_0)y)$ be given. There exists $V \in N_{\tau_2}(0)$ such that $\nabla g(x_0)y + V + V \subset U$. Since g is Frechet differentiable

at x_0, there exists $Y_1 \in N_\tau(y)$ and $\lambda_1 \in (0,\lambda)$ such that for all bounded $Z \subset E$ containing y,

$$\frac{g(x_0+ty')-g(x_0)}{t} - \nabla g(x_0)y' \in V$$

whenever $y' \in Y_1 \cap Z$ and $t \in (0,\lambda_1)$. Since $\nabla g(x_0)$ is $\tau_1 - \tau_2$ continuous, there exists $Y_2 \in N_{\tau_1}(y)$ such that

$$g(x_0)y' \in \nabla g(x_0)y + V \quad \text{for all } y' \in Y_2 .$$

Thus for all bounded Z containing y,

$$(1.8.9) \qquad \frac{g(x_0+ty')-g(x_0)}{t} \in \nabla g(x_0)y' + V$$
$$\subset \nabla g(x_0)y + V + V \subset U$$

whenever $y' \in Y_1 \cap Y_2 \cap Z$ and $t \in (0,\lambda_1)$. Now since $y \in K_C^{\tau_1}(x_0)$, there exist a bounded $Z \subset Y_1 \cap Y_2$, $t \in (0,\lambda_1)$, and $y' \in Z$ with $x_0 + ty' \in C$. Hence by (1.8.9), there exists $\bar{y} \in U \cap \nabla g(x_0)Z$ such that

$$g(x_0) + t\bar{y} = g(x_0+ty') \in g(C) .$$

Therefore $\nabla g(x_0)y \in K_{g(C)}^{\tau_2}(g(x_0))$ and (1.8.8) holds. □

<u>Corollary 1.8.7</u> (cf. [Bo3, Proposition 6]): Under the hypotheses of Proposition 1.8.6, $P_C^\tau(x_0) := \text{cl conv } K_C^\tau(x_0)$

satisfies

$$\nabla g(x_0) P_C^{\tau_1}(x_0) \subset P_{g(C)}^{\tau_2}(g(x_0)) \tag{1.8.10}$$

Proposition 1.8.8 (cf. [Bo3, Theorem 1]): Let $f: E \to \overline{\mathbb{R}}$ be Frechet differentiable at $x_0 \in C \subset E$, and suppose that $\nabla f(x_0)$ is τ-continuous. A necessary condition for x_0 to minimize f over C is that $\nabla f(x_0) \in P_C^\tau(x_0)^+$.

Proof: Suppose x_0 minimizes f over C, and let $y \in K_C^\tau(x_0)$. Then there exists a bounded set $Z \subset E$ such that for all $Y \in N_\tau(y)$ and $\lambda > 0$, there exist $t \in (0,\lambda)$ and $y' \in Y \cap Z$ with $x_0 + ty' \in C$ and

$$\frac{f(x_0+ty')-f(x_0)}{t} \geq 0 .$$

Since F is Frechet differentiable and $\nabla f(x_0)$ is τ-continuous, we conclude that $\nabla f(x_0)y \geq 0$. Hence $\nabla f(x_0) \in K_C^\tau(x_0)^+$. Finally, since $\nabla f(x_0)$ is linear and τ-continuous,

$$\nabla f(x_0) \in P_C(x_0)^+ . \qquad \square$$

One could go further and make a general definition of $R^\tau(C,x_0)$. The properties of such tangent cones could be easily established by the results of sections 1.3 and 1.4. We leave the details of such a development to the reader.

It was observed earlier that $K_\tau(C,x_0) \subset K_C^\tau(x_0)$ in general. We now note that in the case where E is reflexive and $\tau = w$, the weak topology, the two are in fact equal.

Proposition 1.8.9: If E is a reflexive normed space and $x_0 \in C \subset E$, then $K_w(C,x_0) = K_C^w(x_0)$.

Proof: Let $y \in K_C^w(x_0)$. Then there exist $M > 0$ and nets $t_\alpha \downarrow 0$ and $y_\alpha \to (w)y$ such that $c_\alpha := x_0 + t_\alpha y_\alpha \in C$ and $||y_\alpha|| \le M$. Now let α_0 be such that $t_\alpha \le 1$ whenever $\alpha \ge \alpha_0$, and consider

$$D := \{(t_\alpha, t_\alpha^{-1}(c_\alpha - x_0)) \ , \ \alpha \ge \alpha_0\} .$$

Then $(0,y) \in w \text{ cl } C$. It is a consequence of "Whitley's construction" [Hol, p. 148] that there exist sequences t_{α_n} and c_{α_n} such that $t_{\alpha_n} \downarrow 0$ and $t_{\alpha_n}^{-1}(c_{\alpha_n} - x_0) \to (w)y$. Therefore $y \in K_w(C,x_0)$. □

One can apply Corollary 1.8.7 and Proposition 1.8.8 to prove an analogue of the Kuhn-Tucker conditions given in [Bo3, Theorem 2]. By Proposition 1.8.9, such a result would coincide with [Bo3, Theorem 2] if E is reflexive and would be an extension of it otherwise.

1.9. Conclusions

In this section, the results of Chapter 1 are summarized in a series of tables. Table 1.9.1 lists the major properties studied in Chapter 1, and in Table 1.9.2, the combinations of properties that do not occur in general (by results of section 1.2) are compiled, along with combinations which never occur for q-cones. It would be interesting to have further impossibility theorems to rule out more generally some of the combinations which never occur for q-cones. One other pair of properties, (3) and (4), only occurs among q-cones for $G_C(x_0)$, which is trivially convex since it always equals either E or $\emptyset$.

Table 1.9.3 collects together the definitions and properties of the specific tangent cones discussed in Chapter 1. Counterexamples are exhibited for the properties a given cone does not possess. Table 1.9.4 lists a number of inclusions relating these various tangent cones, all of which can be deduced quickly from the definitions of the cones. In Table 1.9.5, the results of sections 1.3 and 1.4 are summarized, making them easier to compare and apply.

With the information in this chapter, one can presumably, as in section 1.8, establish readily the properties of a given q-cone. It is also possible to examine known results involving q-cones and determine whether they might be sharpened by substituting other tangent cones for those given.

Our results so far indicate that the most important tangent cones are those lying between the pseudotangent cone and the Clarke tangent cone, in particular $P_C(x_0)$, $K_C(x_0)$; $k_C(x_0)$, $T_C^*(x_0)$ and $T_C(x_0)$. In the sequel we will focus in on these tangent cones, adding comments about other tangent cones when appropriate.

Table 1.9.1: Tangent cone properties

(1) If $C \subset E$ is closed and convex and $x_0 \in C$,
$\text{cl } R(C,x_0) = \text{cl} \bigcup_{\lambda \geq 0} \lambda(C-x_0)$.

(2) If $h: \mathbb{R}^n \to \mathbb{R}^m$ is continuously differentiable near
$x_0 \in h^{-1}(0)$ and $\nabla h(x_0)$ has rank m , then
$R(h^{-1}(0),x_0) = \nabla h(x_0)^{-1}(0)$.

(3) If $C_1 \subset C_2 \subset E$, $R(C_1,x_0) \subset R(C_2,x_0)$ for all
$x_0 \in E$.

(4) $R(C,x_0)$ is convex for all $(C,x_0) \in 2^E \times E$.

(5) $R(C_1,x_0) \cap R(C_2,x_0) \subset R(C_1 \cap C_2,x_0)$ for all $C_1,C_2 \subset E$
and $x_0 \in E$.

(6) $R(C_1 \cup C_2,x_0) \subset R(C_1,x_0) \cup R(C_2,x_0)$ for all $C_1,C_2 \subset E$
and $x_0 \in E$.

(X) $R(C_1 \times C_2,(x_1,x_2)) = R(C_1,x_1) \times R(C_2,x_2)$ whenever
$x_1 \in C_1 \subset E$ and $x_2 \in C_2 \subset F$.

(NC) If x_0 is a local minimizer for $f: E \to \bar{\mathbb{R}}$,
$0 \in \partial^R f(x_0)$.

Table 1.9.2: Impossible combinations of tangent cone properties

Combination	Reason
$R(C,x_0) \subset K_C(x_0)$, (1), (3), (4)	Theorem 1.2.4
$R(C,x_0) \subset K_C(x_0)$, (2), (3), (4)	
(2), (5)	Theorem 1.2.6
(1), (3), (4), (6)	Theorem 1.2.8
(2), (3), (4), (6)	
(1), (5), (6)	for a homogeneous cone valued tangency operator by Theorem 1.2.10
(2), (5), (6)	
(1), (5), $R(C,x_0) \supset T_C(x_0)$	Theorem 1.2.12
(2), (5), $R(C,x_0) \supset T_C(x_0)$	
(1), (6), $R(C,x_0) \subset k_C(x_0)$	for a homogeneous tangency operator by Theorem 1.2.14
(2), (6), $R(C,x_0) \subset k_C(x_0)$	

Additional combinations which never occur for quantificational tangent cones, cases (a), (b), (c), (d) of (1.3.2):

(i)* (6), (X)

(ii) (5), not (X)

(iii)* (5), (6)

(iv)* (4), (6)

(v) (6), not (3)

(vi) (4), (2), not (1)

(vii)* (6), not (1)

(viii) (5), not (4), not (3)

*satisfied by the tangency operator $R(C,x_0) \equiv \{0\}$.

Table 1.9.3. Tangent cones

$R(C,x_0) := \{y \in E \mid *\ Y \in N(y)\ ,\ \#\ X \in N(x_0)\ ,\ \$\ \lambda > 0\ ,\ \exists\ W \in K(X)\ ,$

$\exists\ \dot{Z} \in M_C(Y)\ ,\ \#'\ x \in W \cap C\ ,\ \$'\ t \in (0,\lambda)\ ,\ *'\ y' \in Z\ ,\ x + ty' \in C\}.$

Name	*	#	\$	K	M	f^R	cl,op	1	2	3	4	5	6	X	NC
A	–	–	∀	x_0	y	f^A	n	+	0 e10	+	0 e2	0 e3	+	0 e8	+
E	–	–	∃	x_0	y	f^E	n	+	0 e10	+	0 e2	+	0 e3	+	+
K	∀	–	∀	x_0	Y	f_+	cl	+	+	+	0 e2	0 e3	+	0 e8	+
k	∀	–	∃	x_0	Y	$f_\square$	cl	+	+	+	0 e2	0 e3	0 e11	+	+
ℓ	∃	–	∀	x_0	Y	f^ℓ	op	0 rel	0 e10	+	0 e2	0 e12	0 e3	0 e12	+
L	∃	–	∃	x_0	Y	f^+	op	0 rel	0 e10	+	0 e2	+	0 e3	+	+
H*	–	∃	∀	X	y	f^{H*}	n	0 fd	0 e10	0 e1	0 e7	0 e3	0 e13	0 e6	+
H	–	∃	∃	X	y	f^0	n	0 fd	0 e10	0 e1	+	+	0 e3	+	+
D	∀	∀	∀	X	Y	f^D	cl	0 1.3.8	+	+	0 e8	0 e3	0 e4	0 e8	1.5.24
a	∀	∀	∃	X	Y	f^a	cl	0 1.3.8	+	+	0 e4	0 e3	0 e11	+	+ 1.5.24
G	∃	∀	∃	X	Y	f^G	op	0 1.3.8	0 e10	+	+	0 e5	0 e3	+	0 1.5.24
T*	∀	∃	∀	X	Y	f^{T*}	cl	+	+	0 e1	0 e5	0 e3	0 e13	0 e5	+

Table 1.9.3 continued on following page.

Table 1.9.3 (continued)

Name	*	#	$	K	M	f^R	cl(osed) or op(en)	1	2	3	4	5	6	X	NC
T	∀	∃	∃	X	Y	$f^{\uparrow}$	cl	+	+	0 e1	+	0 e3	0 e13	+	+
I*	∃	∃	∀	X	Y	f^{I*}	op	0 re1	0 e10	0 e1	0 e5	0 e5	0 e3	0 e5	+
I	∃	∃	∃	X	Y	f^*	op	0 re1	0 e10	0 e1	+	+	0 e3	+	+
F	∀	∃	∃	X	case c 1.3.2	f^F	cl	0 fd	+	0 e1	+	0 e3	0 e13	+	+
P	not a q-cone					f^P	cl	+	+	+	+	0 e3	0 e4	0 e9	+
cone$(C-x_0)$		not a q-cone				f'	n	+	0 e10	+	0 e2	0 e3	+	0 e8	+

Table 1.9.3 (continued)

Abbreviations in Table 1.9.3:

Symbol	Meaning
-	Either $\forall$ or $\exists$.
+	R has that property.
0	R does not have that property.
cl	R is always closed.
op	R is always open.
n	R is neither open nor closed in general.
fd	R has Property 1 in finite dimensions, but not in general (see section 1.3).
rel	In finite dimensions, R^{rel} has Property 1. A counterexample for R is example 1.3.10.
1.3.8	See Example 1.3.8.
1.5.24	See Example 1.5.24.
case c	$M(Y)$ = nonempty compact subsets of Y.

Examples referred to in Table 1.9.3:

e1: $C_1 = \{(x,y)\,|\,x=0\}$, $C_2 = \{(x,y)\,|\,x=0 \text{ or } y=0\}$, $x_0 = (0,0)$

e2: $C = \{(x,y)\,|\,y \geq -|x|\}$, $x_0 = (0,0)$.

e3: $C_1 = \mathbb{Q}$, $C_2 = (\mathbb{R}\setminus\mathbb{Q}) \cup \{0\}$, $x_0 = 0$.

e4: $C_1 = \{(x,y)\,|\,x=0\}$, $C_2 = \{(x,y)\,|\,y=0\}$, $x_0 = (0,0)$.

e5: $C_1 = \bigcup_{n=0}^{\infty} \left(\frac{1}{2^{4n+1}}, \frac{1}{2^{4n}}\right)$, $C_2 = \bigcup_{n=0}^{\infty} \left(\frac{1}{2^{4n+3}}, \frac{1}{2^{4n+2}}\right)$,

$C = C_1 \times C_2$, $x_0 = (0,0)$, $x_1 = 0$, $x_2 = 0$.

e6: $C_1 = \mathbb{Q}$, $C_2 = \mathbb{Q}$, $x_1 = 0$, $x_2 = 0$.

e7: $C = \mathbb{Q} \times \mathbb{Q}_2$, $x_0 = (0,0)$.

e8: $C_1 = \{2^{-2n}\,|\,n \geq 0\} \cup \{0\}$, $x_1 = 0$

$C_2 = \{2^{-2n+1}\,|\,n \geq 0\} \cup \{0\}$, $x_2 = 0$

$C = C_1 \cup C_2$, $x_0 = (0,0)$.

e9: $C_1 = \{2^{-\frac{n(n+1)}{2}}\ |\ n \text{ even}\} \cup \{0\}$, $x_1 = 0$

$C_2 = \{2^{-\frac{n(n+1)}{2}}\ |\ n \text{ odd}\} \cup \{0\}$, $x_2 = 0$.

e10: $C = \{(x,y)\,|\,y = x^2\}$, $x_0 = (0,0)$.

e11: $C_1 = \{\frac{1}{m}\ |\ \exists\ n \geq 0,\ 2^{2n} \leq m < 2^{2n+1}\} \cup \{0\}$, $x_1 = 0$

$C_2 = \{\frac{1}{m}\ |\ \exists\ n \geq 0,\ 2^{2n-1} \leq m < 2^{2n}\} \cup \{0\}$, $x_2 = 0$

e12: $C_1 = \mathbb{R}^+\setminus\{\frac{1}{m}\ |\ \exists\ n \geq 0,\ 2^{2n} \leq m < 2^{2n+1}\} \cup \{0\}$, $x_1 = 0$

$C_2 = \mathbb{R}^+\setminus\{\frac{1}{m}\ |\ \exists\ n \geq 0,\ 2^{2n-1} \leq m < 2^{2n}\} \cup \{0\}$, $x_2 = 0$.

e13: $C_1 = \bigcup_{n=0}^{\infty} \left(\frac{1}{2^{2n+1}}, \frac{1}{2^{2n}}\right] \cup \{0\}$

$C_2 = \bigcup_{n=1}^{\infty} \left(\frac{1}{2^{2n}}, \frac{1}{2^{2n-1}}\right] \cup \{0\}$, $x_0 = 0$.

Table 1.9.4. Tangent cone inclusions

(i)	$I_C(x_0) \subset R(C,x_0)$ for any q-cone R
(ii)	$R(C,x_0) \subset D_C(x_0)$ for any q-cone R
(iii)	$H_C(x_0) \subset E_C(x_0) \subset A_C(x_0) \subset \mathrm{cone}(C-x_0)$
(iv)	$L_C(x_0) \subset E_C(x_0) \subset k_C(x_0) \subset a_C(x_0)$
(v)	$L_C(x_0) \subset \ell_C(x_0) \subset A_C(x_0) \subset K_C(x_0)$
(vi)	$T_C(x_0) \subset k_C(x_0) \subset K_C(x_0) \subset P_C(x_0)$
(vii)	$T_C(x_0) \subset T^*_C(x_0) \subset K_C(x_0)$
(viii)	$I^*_C(x_0) \subset H^*_C(x_0) \subset T^*_C(x_0)$
(ix)	$H_C(x_0) \subset H'_C(x_0) \subset H^*_C(x_0) \subset A_C(x_0)$
(x)	$H_C(x_0) \subset F_C(x_0) \subset T_C(x_0)$
(xi)	$L_C(x_0) \subset G_C(x_0) \subset a_C(x_0)$
(xii)	$\left.\begin{array}{l} H_C(x_0) = H^*_C(x_0) \\ I_C(x_0) = I^*_C(x_0) \end{array}\right\}$ if C is closed in a l.c.s.
(xiii)	$T_C(x_0) = T^*_C(x_0)$ if C is closed in a Banach space

Table 1.9.5. Summary of results of Sections 1.3 and 1.4

*	#	$	K	M	imply	by Prop. or Thm.
–	–	–	–	all	cone-valuedness	1.3.2
–	–	–	all	all	homogeneity	1.3.3
$\forall$	–	–	–	all	closure	1.3.4
$\exists$	–	–	–	all	openness	1.3.7
–	$\forall$	–	–	–	isotonicity	1.3.19
$\forall$	$\exists$	$\exists$	X	all	convexity	1.3.21
$\exists$	$\exists$	$\exists$	X	a,b,c	convexity	1.3.23
$\exists$	$\exists$	$\exists$	all	all	property (5)	1.3.24
$\forall$	–	$\forall$	x_0	all	property (6)	1.3.26
–	–	$\exists$	all	a,b,c	products preserved	1.4.2

*	#	$	K	M	additional assumption	imply	by Prop., Thm.
$\forall$	$\forall$	–	all	Y		(1.4.11)	1.4.7
$\forall$	$\exists$	–	X	Y	relative openness	(1.4.11)	1.4.10
$\forall$	$\forall$	–	all	all		(1.4.13)	1.4.13
$\forall$	$\exists$	–	X	all	relative openness	(1.4.13)	1.4.15
$\exists$	$\forall$	–	all	all	A open	(1.4.13)	1.4.17
$\exists$	$\exists$	–	X	all	relative openness, A open	(1.4.13)	1.4.19

(1.4.11) $\quad \nabla g(x_0) R(C, x_0) \subset R(g(C), g(x_0))$

(1.4.13) $\quad A(R(C, x_0)) \subset R(A(C), Ax_0)$.

In the table above,

"–" denotes "no restriction".

"all" denotes "assumption holds for all of cases (a),(b),(c),(d)" of (1.3.2).

"a,b,c" denotes "assumption holds for cases (a),(b),(c)".

"relative openness" is condition (1.4.16).

CHAPTER II

Generalized Subdifferential Calculus: the Finite-dimensional Case

2.1. Introduction

The concept of subgradient was introduced originally for convex functions $f: E \to \overline{\mathbb{R}}$ by the definition

$$(2.1.1) \qquad \partial f(x_0) := \{x' \in E' \mid \langle x-x_0, x'\rangle \le f(x)-f(x_0) \text{ for all } x \in E\} .$$

These subgradients admit a versatile calculus (see [Ro1], [E2] and their references), the cornerstone of which is the "subgradient sum formula":

Theorem 2.1.1: Let $g_1: E \to \overline{\mathbb{R}}$, $g_2: E \to \overline{\mathbb{R}}$ be proper convex functions, and let $x_0 \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$. Assume

$$(2.1.2) \qquad \operatorname{dom} g_1 \cap \operatorname{int} \operatorname{dom} g_2 \ne \emptyset .$$

Then

$$(2.1.3) \qquad \partial(g_1+g_2)(x_0) = \partial g_1(x_0) + \partial g_2(x_0) .$$

Note that the inclusion $\partial g_1(x_0) + \partial g_2(x_0) \subset \partial(g_1+g_2)(x_0)$ follows immediately from (2.1.1). The proof of the opposite inclusion in (2.1.3) uses (2.1.2) to invoke some equivalent of the Hahn-Banach theorem ([E2], [Hol]). If $E := \mathbb{R}^n$, assumption (2.1.2) can be replaced by the weaker requirement [Ro1, Theorem 23.8].

$$\text{ri dom } g_1 \cap \text{ri dom } g_2 \neq \emptyset . \tag{2.1.4}$$

In section 1.5, we saw that for a convex function $f: E \to \overline{\mathbb{R}}$ and $x_0 \in \text{dom } f$,

$$\partial f(x_0) = \partial^A f(x_0) = \partial^E f(x_0) = \partial^K f(x_0)$$
$$= \partial^k f(x_0) = \partial^T f(x_0) ,$$

and if in addition $x_0 \in \text{int dom } f$,

$$\partial^H f(x_0) = \partial f(x_0) .$$

Do there exist analogues of Theorem 2.1.1 involving R-subgradients which are valid for wider classes of functions? The answer is yes. For locally Lipschitzian functions defined on a normed space, Clarke ([Cl2], [Cl3], and [Cl4, chapter 2]) and Hiriart-Urruty [Hi1] have developed an extensive calculus for $\partial^H f$. More recently, Rockafellar [Ro3] has derived a calculus for $\partial^T f$ which generalizes that for ∂f and $\partial^H f$ and is valid for functions which are not necessarily convex or even continuous (see also [Ro5], [Cl4, section 2.9]). In this latter development, the analogue of Theorem 2.1.1 is the following result:

Theorem 2.1.2 [Ro3, Theorem 2]: Let E be a l.c.s., $f_1: E \to \bar{\mathbb{R}}$ and $f_2: E \to \bar{\mathbb{R}}$, and suppose f_1 and f_2 are finite at x_0. Assume

$$\text{(2.1.5)} \qquad \operatorname{dom} f_1^{\uparrow}(x_0;\cdot) \cap \operatorname{int} \operatorname{dom} f_2^{\uparrow}(x_0;\cdot) \neq \emptyset .$$

Then

$$\text{(2.1.6)} \qquad \partial^T(f_1 + f_2)(x_0) \subset \partial^T f_1(x_0) + \partial^T f_2(x_0) .$$

Remark 2.1.3: (a) Equality does not in general hold in (2.1.6), as demonstrated by the following example, given in [Ro5]: Define $f_1 : \mathbb{R} \to \mathbb{R}$ by $f_1(x) := |x|$ and $f_2: \mathbb{R} \to \mathbb{R}$ by $f_2(x) := -|x|$, and let $x_0 := 0$. Then $f_1 + f_2 \equiv 0$ and $\partial^T(f_1+f_2)(0) = 0$, while $\partial^T f_1(0) = \partial^T f_2(0) = [-1,1]$, so that $\partial^T f_1(0) + \partial^T f_2(0) = [-2,2]$. Conditions sufficient for equality in (2.1.6) are given in [Ro3]. We will discuss them in section 2.4.

(b) Statements (2.1.5) and (2.1.6) may be written equivalently [Bo7] as

$$\text{(2.1.7)} \qquad \operatorname{dom} f_1^{\uparrow}(x_0;\cdot) \cap \operatorname{dom} f_2^{I}(x_0;\cdot) \neq \emptyset$$

and

$$\text{(2.1.8)} \qquad \partial^T(f_1+f_2)(x_0) \subset \partial^T f_1(x_0) + \partial^I f_2(x_0) .$$

The proof of Theorem 2.1.2 relies in a fundamental way upon the convexity of the Clarke tangent cone. Recall (Remark 1.5.3(b)) that

$$(2.1.9) \qquad \text{epi } f^{\dagger}(x_0;\cdot) = T_{\text{epi } f}((x_0,f(x_0))) .$$

Through (2.1.9), $f^{\dagger}(x_0;\cdot)$ inherits important properties from the Clarke tangent cone. Since $T_C(x_0)$ is closed, $f^{\dagger}(x_0;\cdot)$ is l.s.c., and since $T_C(x_0)$ is a convex cone, $f^{\dagger}(x_0;\cdot)$ is convex and positively homogeneous. As a result, either

(2.1.10) (a) $f^{\dagger}(x_0;0) = 0$, in which case $f^{\dagger}(x_0,\cdot)$ is proper

or (b) $f^{\dagger}(x_0;0) = -\infty$, in which case $f^{\dagger}(x_0;\cdot)$ is equal to $-\infty$ throughout its domain and $\partial^T f(x_0) = \emptyset$.

Using these facts, Rockafellar's proof proceeds by the following strategy: First, the direct characterization of $f^{\dagger}(x_0;y)$ given in (1.5.17) and the fact that $f^{\dagger}(x_0;y) = f^{I}(x_0;y)$ for $y \in \text{int dom } f\,(x_0;\cdot)$ are applied to prove that if (2.1.5) holds, then

$$(2.1.11) \qquad (f_1+f_2)^{\dagger}(x_0;y) \le f_1^{\dagger}(x_0;y) + f_2^{\dagger}(x_0;y)$$
$$\text{for all } y \in E .$$

Secondly, it is observed that if either $f_1^{\dagger}(x_0;0)$ or $f_2^{\dagger}(x_0;0) = -\infty$, then $(f_1+f_2)^{\dagger}(x_0;0) = -\infty$ by (2.1.11) and $\partial^T(f_1+f_2)(x_0) = \emptyset$ by (2.1.10)(b). Otherwise, for $i = 1,2$, $f_i^{\dagger}(x_0;\cdot)$ are proper convex functions, $\partial^T f_i(x_0) =$

$\partial f_i^{\uparrow}(x_0;\cdot)^{\vee}(0)$, and (2.1.5) is just (2.1.2) with $g_i = f_i^{\uparrow}(x_0;\cdot)$. Finally, (2.1.11) and Theorem 2.1.1 are combined to prove (2.1.6).

It is interesting to note that Theorem 2.1.1, which is used to prove Theorem 2.1.2, is in turn a corollary of Theorem 2.1.2. This becomes easier to see after it is observed that (2.1.2) can be written equivalently as

(2.1.12) $$\operatorname{dom} g_1^{\vee}(x_0;\cdot) \cap \operatorname{int} \operatorname{dom} g_2^{\prime}(x_0;\cdot) \neq \emptyset .$$

Unfortunately, Theorem 2.1.2 is not strong enough to recapture the finite-dimensional version of Theorem 2.1.1, in which assumption (2.1.2) is replaced by (2.1.4). One of our purposes in this chapter is to produce a finite-dimensional calculus for $\partial^T f$ which is strong enough to encompass as a special case the strongest version of the finite-dimensional subdifferential calculus for convex functions ([Ro1 , Section 23]). We do so by a method in some ways similar to, and in some ways different from, the method of proof for Theorem 2.1.2 outlined above. In our version of the subgradient sum formula, we first establish (2.1.11), then use [Ro1, Theorem 23.8] to prove (2.1.6), the overall pattern followed in [Ro3]. However, rather than work directly with (1.5.17), as in [Ro3], we rely on an "inversion theorem" of Borwein [Bo5], inclusion (1.4.18) of Corollary 1.4.12, and the relationship in (2.1.9) to prove (2.1.11) more easily than in [Ro3] and with an assumption less demanding than (2.1.5).

Here is an outline of Chapter 2: In section 2.2, we collect the preliminary results that we will employ in proving subdifferential calculus formulae. We prove our two main subdifferential calculus formulae in section 2.3 and give a number of corollaries. This section contains finite-dimensional strengthenings of many of the results in [Ro3], as well as a few formulae - product and quotient rules - that have no analogue in [Ro3]. In section 2.4, we prove directional derivative inequalities involving $f_+(x_0;\cdot)$ and $f^k(x_0;\cdot)$ and use them to derive conditions guaranteeing equality in the inequalities and inclusions of section 2.3. These conditions are somewhat weaker than the usual "subdifferential regularity" conditions [Ro3]. Some of our inequalities involving $f_+(x_0;\cdot)$ and $f^k(x_0;\cdot)$ will have further application in Chapter 5 in generalizing previous results about "upper convex approximates". Finally, in section 2.5, we indicate briefly how the methods of this chapter can be used to derive subdifferential calculus formulae for extended real valued functions defined on Banach spaces.

In subsequent sections, we will simplify notation by writing simply "∂f" in place of "$\partial^T f$", since this causes no ambiguity. The notation "∂f" has previously been used in the locally Lipschitzian case (in place of our $\partial^H f$) as well as in the convex case; however, as we saw in section 1.5, $\partial^T f$ coincides with ∂f in the convex case (since the Clarke tangent cone has property (1)) and in the locally Lipschitzian case (by Corollary 1.5.19). One further notational note: in section 2.4, we will use $f_{\square}(x_0;\cdot)$ as an alternate notation for $f^k(x_0;\cdot)$.

Very recently, Ioffe, in the significant paper [I2], has used an entirely different approach to prove some of the results of section 2.3. Specifically, he has derived Corollaries 2.3.5 and 2.3.7 and Proposition 2.3.14 as special cases of results in the calculus of "approximate subdifferentials". In contrast to our "convex analysis based" approach, Ioffe uses a penalty type argument to establish an analogue of Theorem 2.1.2 for approximate subdifferentials, then passes to a corresponding result for ∂f by means of the important inclusion ([Pe2], [Pe3], [Tr1], 4.1.1))

$$\liminf_{\substack{x \to x_0 \\ C}} K_C(x_0) \subset T_C(x_0)$$

We will discuss this inclusion - and approximate subdifferentials - in chapter 4. While our methods do not yield results about approximate subdifferentials, they do seem to provide a simpler and more versatile method of establishing generalized subdifferential calculus rules of the type presented in section 2.3.

Even more recently, Rockafellar [Ro7] has derived inclusions (2.3.3), (2.3.6), (2.3.16), and (2.3.33) by an approach that is "dual" to that taken here. The methods used in [Ro7] center around the normal cone $N_C(x_0)$ (see Definition 2.3.2) and earlier results of Rockafellar on proximal normals. The methods of [Ro7] can be very fruitfully applied to the study of perturbed optimization problems, however, they do not seem to yield any information on conditions for equality in the subgradient inclusions.

2.2 Preliminaries

<u>Definition 2.2.1</u>: Let E be a l.c.s. and $x_0 \in C \subseteq E$.

(a) The set C is said to be <u>closed near x_0</u> if there exists $X \in N(x_0)$ such that $X \cap C$ is closed.

(b) The function $f: E \to \bar{\mathbb{R}}$ is <u>strictly l.s.c. at</u> x_0 if for some $\alpha > f(x_0)$, the function $\min\{f,\alpha\}$ is l.s.c.

(c) The function $f: E \to \bar{\mathbb{R}}$ is <u>l.s.c. at</u> x_0 if for any $\varepsilon > 0$, there exists $X \in N(x_0)$ such that $f(x) \geq f(x_0) - \varepsilon$ for all $x \in X$.

It is observed in [Ro6] that if f is strictly l.s.c. at x_0, then the set epi f is closed near $(x_0, f(x_0))$ and f is l.s.c. at x_0. We make the further observation that Propositions 1.7.1 and 1.7.4 still hold for C closed near x_0 and (1.7.2) still holds for f merely strictly l.s.c. at x_0, by the localization property of q-cones (Theorem 1.4.16).

The result that will allow us to weaken assumption (2.1.5) is the special case of [Bo5, Theorem 4.1] given below. Notice that a corollary of Theorem 2.2.2 is the fact that $T_C(x_0)$ has property (2).

<u>Theorem 2.2.2</u>: Let C be a closed subset of $\mathbb{R}^p$, and let $G: \mathbb{R}^p \to \mathbb{R}^q$ be strictly differentiable at $x_0 \in C \cap G^{-1}(0)$. Assume

(2.2.1) $$\nabla G(x_0) T_C(x_0) = \mathbb{R}^q .$$

Then

$$(2.2.2) \qquad T_C(x_0) \cap \nabla G(x_0)^{-1}(0) \subset T_{C \cap G^{-1}(0)}(x_0) .$$

The proof of Theorem 2.2.2 depends on Ekeland's variational principle [El], an important tool of nonsmooth analysis. In a sense, all the "hard work" in the proofs in section 2.3 is contained in Theorem 2.2.2.

Because of the localization property of q-cones, we actually need only assume in Theorem 2.2.2 that C is closed near x_0, as we now demonstrate:

<u>Corollary 2.2.3</u>: In Theorem 2.2.2, replace the hypothesis that C is closed by the hypothesis that C is closed near x_0. Then (2.2.2) still holds.

<u>Proof</u>: Since C is closed near x_0, there exists $X \in N(x_0)$ such that $C \cap X$ is closed. By Theorem 1.4.6 and (2.2.1),

$$\nabla G(x_0) T_{C \cap X}(x_0) = \nabla G(x_0) T_C(x_0) = \mathbb{R}^q .$$

Hence $T_{C \cap X}(x_0) \cap \nabla G(x_0)^{-1}(0) \subset T_{C \cap X \cap G^{-1}(0)}(x_0)$. Applying Theorem 1.4.6 again, we conclude that

$$T_C(x_0) \cap \nabla G(x_0)^{-1}(0) \subset T_{C \cap G^{-1}(0)}(x_0) . \qquad \square$$

Another key ingredient in the proofs in section 2.3 is the following special case of Corollary 1.4.12 (see also [Bo5, Corollary 4.2]):

<u>Proposition 2.2.4</u>: Let E, E^1 be l.c.s. and $A: E \to E^1$ be linear and continuous. Let $z_0 \in C \subset E$. Suppose that A is relatively open on C at z_0; i.e.,

$$\text{(2.2.3)} \qquad \text{For each } X \in \mathcal{N}(z_0)\text{, there exists } Z \in \mathcal{N}(Az_0) \text{ such that } Z \cap A(C) \subset A(X \cap C).$$

Then

$$\text{(2.2.4)} \qquad A(T_C(z_0)) \subset T_{A(C)}(Az_0).$$

Condition (2.2.3) is simply condition (1.4.17) with $g := A$. Notice that (2.2.3) holds in particular whenever A is open and one-to-one on C.

<u>Example 2.2.5</u>: (a) Let $A: \mathbb{R}^2 \to \mathbb{R}$ be defined by $A(x,y) = y$, $C = \{(x,y) \mid x \geq 0, y \leq 0\} \cup \{(0, \frac{1}{n}) \mid n = 1,2,\ldots\}$, and $z_0 = (1,0)$. Then neither (2.2.3) nor (2.2.4) holds, since $A(T_C(z_0)) = \{y \mid y \leq 0\}$, while

$$T_{A(C)}(Az_0) = \{0\}.$$

(b) It is possible, though, for (2.2.4) to hold without (2.2.3) being satisfied. For example, consider A and z_0 as in (a), and let $C := \{(x,y) \mid x \geq 0, y \leq 0\} \cup \{(x,y) \mid x = 0\}$. Here (2.2.3) does not hold, but (2.2.4) does, since

$$A(T_C(z_0)) = \{y \mid y \leq 0\} \quad \text{and} \quad T_{A(C)}(Az_0) = \mathbb{R}.$$

In section 2.3, we will establish calculus rules involving functions of two forms:

(2.2.5) (a) $h := f_1 + f_2 \circ F$, where $f_1 \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is strictly l.s.c. at x_0 , $f_2 \colon \mathbb{R}^m \to \overline{\mathbb{R}}$ is strictly l.s.c. at $F(x_0)$, and $F \colon \mathbb{R}^n \to \mathbb{R}^m$ is strictly differentiable at $x_0 \in \operatorname{dom} f_1 \cap F^{-1}(\operatorname{dom} f_2)$.

(b) $h := F \circ f$, where $f = (f_1, \ldots, f_n)$; each $f_i \colon \mathbb{R}^m \to \overline{\mathbb{R}}$ is strictly l.s.c. at x_0 , and $F \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is isotone (nondecreasing) with respect to the coordinate ordering in $\mathbb{R}^n$.

The proofs of these calculus rules will consist of two stages, as was mentioned in section 2.1:

Stage 1: Establish a directional derivative inequality with the help of (2.1.9), inclusions (2.2.2) and (2.2.4), and the fact that the Clarke tangent cone is product-preserving (Corollary 1.4.3).

Stage 2: Use that inequality and a convex subdifferential calculus formula to establish a corresponding subgradient inclusion.

In order to carry out stage 1, we need to establish that condition (2.2.3) holds for the appropriate A , C and z_0 . We do so below in two technical lemmas. In the proofs of these lemmas, for $x = (x_1,\ldots,x_p) \in \mathbb{R}^p$ and $\varepsilon > 0$, we define $B_\varepsilon(x) := \{y = (y_1,\ldots,y_p) \in \mathbb{R}^p \,|\, |y_i - x_i| \le \varepsilon,\ i = 1,\ldots,p\}$.

<u>Lemma 2.2.6:</u> Let E, E^1 be l.c.s., and let $F: E \to E^1$ be continuous at x_0, $f_1: E \to \overline{\mathbb{R}}$ be finite and l.s.c. at x_0, and $f_2: E^1 \to \overline{\mathbb{R}}$ be finite and l.s.c. at $F(x_0)$. Define $A: E \times \mathbb{R} \times E^1 \times \mathbb{R} \to E \times \mathbb{R}$ by $A(x,y,z,r) := (x,y+r)$ and $G: E \times \mathbb{R} \times E^1 \times \mathbb{R} \to E^1$ by $G(x,y,z,r) := F(x) - z$. Then (2.2.3) is satisfied with A as above, $C := (\text{epi } f_1 \times \text{epi } f_2) \cap G^{-1}(0)$, and $z_0 := (x_0, f_1(x_0), F(x_0), f_2(F(x_0)))$.

<u>Proof:</u> Let $\varepsilon > 0$, $U \in N(x_0)$, and $Y \in N(F(x_0))$ be given, and let $X := U \times B_\varepsilon(f_1(x_0)) \times Y \times B_\varepsilon(f_2(F(x_0))) \in N(z_0)$. By hypothesis, there exists $U_1 \subset U$, $U_1 \in N(x_0)$ such that for all $x \in U_1$, we have

$$f_1(x) \ge f_1(x_0) - \varepsilon/3 ,$$

$$F(x) \in Y \text{ , and}$$

$$(f_2 \circ F)(x) \ge (f_2 \circ F)(x_0) - \varepsilon/3 .$$

Define $N := U_1 \times B_{\varepsilon/3}(f_1(x_0)) \times Y \times B_{\varepsilon/3}(f_2(F(x_0)))$. Since A is surjective, $Z := A(N) \in N(Az_0)$. We will now verify that $Z \cap A(C) \subset A(X \cap C)$. To do so, let $(\bar{x},\bar{r}) \in Z \cap A(C)$. Since $(\bar{x},\bar{r}) \in Z$, there exists $(x,y,z,r) \in N$ with $x = \bar{x}$, $y + r = \bar{r}$. Since $(\bar{x},\bar{r}) \in A(C)$, there exists

$(x',y',z',r') \in C$ with $x' = \bar{x}$, $y' + r' = \bar{r}$, and $f_1(x') \le y'$, $z' = F(x')$, $f_2(z') \le r'$. Now $x' \in U_1$, so $z' \in Y$. Also $\bar{r} \in B_{\varepsilon/3}(f_1(x_0)) + B_{\varepsilon/3}(f_2(F(x_0)))$, so $\bar{r} \in B_{2\varepsilon/3}(f_1(x_0) + f_2(F(x_0)))$. Finally, since $y' + r' = \bar{r}$ and $y' \ge f_1(x_0) - \varepsilon/3$, we conclude that $y' \in B_\varepsilon(f_1(x_0))$ and $r' \in B_\varepsilon(f_2(F(x_0)))$. Thus $(x',y',z',r') \in X$, and so $(\bar{x},\bar{r}) \in A(X \cap C)$. We conclude that $Z \cap A(C) \subset A(X \cap C)$. □

<u>Definition 2.2.7</u>: $F: \mathbb{R}^n \to \bar{\mathbb{R}}$ is <u>isotone</u> on $D \subset \mathbb{R}^n$ if $F(x) \le F(y)$ whenever $x,y \in D$ and $x \le y$ (with respect to the coordinate ordering). F is strictly isotone in the i^{th} coordinate at $x := (x_1,\ldots,x_n) \in \mathbb{R}^n$ if $F(x) < F(y)$ whenever $x \le y$ and $x_i < y_i$.

<u>Lemma 2.2.8:</u> Let $f_i: E \to \bar{\mathbb{R}}$, $i = 1,\ldots,n$ be finite and l.s.c. at x_0, and call $f := (f_1,\ldots,f_n)$. Let $F: \mathbb{R}^n \to \bar{\mathbb{R}}$ be finite at $f(x_0)$, isotone on $B_{t_0}(f(x_0)) + \mathbb{R}^n_+$ for some $t_0 > 0$, and l.s.c. Define $A: (E \times \mathbb{R})^n \times \mathbb{R}^{n+1} \to E \times \mathbb{R}$ by $A(x_1,y_1,\ldots,x_n,y_n,z_1,\ldots,z_n,r) := (x_1,r)$. (Here $x_i \in \mathbb{R}^m$, $y_i, z_i \in \mathbb{R}$.) Define $G: (E \times \mathbb{R})^n \times \mathbb{R}^{n+1} \to E^{n-1} \times \mathbb{R}^n$ by $G(x_1,y_1,\ldots,x_n,y_n,z_1,\ldots,z_n,r) := (x_1-x_2,\ldots,x_1-x_n,y_1-z_1,\ldots,y_n-z_n)$. Assume for each $j \in \{1,\ldots,n\}$ that either (1) f_j is continuous at x_0, or (2) F is strictly isotone in the jth coordinate at $f(x_0)$. Then (2.2.3) is satisfied with A as above,

$C := (\text{epi } f_1 \times \ldots \times \text{epi } f_n \times \text{epi } F) \cap G^{-1}(0)$, and

$z_0 := (x_0, f_1(x_0),\ldots,x_0,f_n(x_0),f_1(x_0),\ldots,f_n(x_0),F(f(x_0)))$.

Proof: Let $\varepsilon > 0$ and $U \in N(x_0)$ be given, and define
$X := \prod_{i=1}^{n} (U \times B_\varepsilon(f_i(x_0))) \times B_\varepsilon(f(x_0)) \times B_\varepsilon(F(f(x_0))) \in N(z_0)$.
For $t > 0$, define $U_t := \{(y_1,\ldots,y_n) \in \mathbb{R}^n | y_i \geq f_i(x_0)-t$, $i = 1,\ldots,n\}$. By hypothesis, F is isotone on U_{t_0} for some $t_0 > 0$. Let $I \subset \{1,\ldots,n\}$ be the set of coordinates in which F is strictly isotone at $f(x_0)$. Suppose $i \in I$. We claim that there exists $\mu_i \in (0,t_0)$ such that $F(f_1(x_0) - \mu_i,\ldots,f_i(x_0) + \varepsilon,\ldots,f_n(x_0)-\mu_i) > F(f(x_0))$. If not, then for all $t \in (0,t_0)$,

$$F(f_1(x_0)-t,\ldots,f_i(x_0) + \varepsilon,\ldots,f_n(x_0)-t) \leq F(f(x_0)) .$$

But since F is l.s.c., it follows that

$$F(f_1(x_0),\ldots,f_i(x_0) + \varepsilon,\ldots,f_n(x_0)) \leq F(f(x_0)) .$$

This contradicts our assumption that $i \in I$. So for each $i \in I$, we can choose $\mu_i \in (0,t_0)$ with

$$\delta_i := F(f_1(x_0)-\mu_i,\ldots,f_i(x_0) + \varepsilon,\ldots,f_n(x_0)-\mu_i)-F(f(x_0))>0.$$

Let $\mu = \min_I \mu_i$, and let $\delta = \frac{1}{2} \min_I \delta_i$. Then if $r \in B_\delta(F(f(x_0))) \cap F(U_\mu)$, it follows from the isotonicity of F that $r = F(y)$ for some $y = (y_1,\ldots,y_n)$ with $y_i \leq f_i(x_0) + \varepsilon$ for all $i \in I$. Now by assumptions (1) and (2), there exists $V \in N(x_0)$, $V \subset U$ such that $f_j(x) \in B_\mu(f_j(x_0))$ whenever $x \in V$, $j \notin I$, and $f_j(x) \geq f_j(x_0) - \mu$ whenever $x \in V$, $j \in I$. Let

$N := \prod_{i=1}^{n} (V \times B_\varepsilon(f_i(x_0))) \times B_\varepsilon(f(x_0)) \times B_\delta(F(f(x_0)))$. Again $Z := A(N) \in \mathcal{N}(Az_0)$. We will now show that $Z \cap A(C) \subset A(X \cap C)$. Suppose $(\bar{x},\bar{r}) \in Z \cap A(C)$. Since $(\bar{x},\bar{r}) \in A(N)$, $\bar{x}$ must be in V and $\bar{r}$ in $B_\delta(F(f(x_0)))$. Since $(\bar{x},\bar{r}) \in A(C)$, there exists $y \in \mathbb{R}^n$ with $f(\bar{x}) \le y$ and $F(y) \le \bar{r}$. Thus $\bar{y} := f(\bar{x})$ satisfies $F(\bar{y}) \le \bar{r}$, $\bar{y}_i \in B_\mu(f_j(x_0))$ for all $j \notin I$, and $f_j(x_0) - \mu \le \bar{y}_j \le f_j(x_0) + \varepsilon$ for all $j \in I$. We conclude that $(\bar{x},\bar{y}_1,\ldots,\bar{x},\bar{y}_n,\bar{y}_1,\ldots,\bar{y}_n,\bar{r}) \in X \cap C$, and so $(\bar{x},\bar{r}) \in A(X \cap C)$. □

Remark 2.2.9: The hypotheses of Lemma 2.2.8 hold in the following important cases:

(a) $F(y_1,\ldots,y_n) := \sum_{i=1}^{n} y_i$ and each f_j l.s.c. at x_0 .

(b) $F(y_1,\ldots,y_n) := \prod_{i=1}^{n} y_i$ and each f_j positive and l.s.c. at x_0 .

(c) $F(y_1,\ldots,y_n) := \max_{1\le j\le n} y_j$ and each f_j l.s.c. at x_0 , with f_j continuous at x_0 for all $j \notin I(x_0)$, where $I(x_0) := \{i \in \{1,\ldots,n\} \mid f_i(x_0) = \max_{1\le j\le n} f_j(x_0)\}$.

We will also make use of the following lemma concerning preservation of isotonicity by R-directional derivatives:

Lemma 2.2.10: Let E be a l.c.s., and let R be a q-cone with $K(X) = X$ or $K(X) = x_0$. Let $F: \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at

$x_0 \in \mathbb{R}^n$ and isotone on a neighbourhood of x_0. Then $F^R(x_0;\cdot)$ is isotone on $\mathbb{R}^n$.

<u>Proof</u>: Let $y_1, y_2 \in \mathbb{R}^n$ with $y_1 \leq y_2$, and suppose $(y_2,d) \in \hat{R}(F,x_0)$. It suffices to show that $(y_1,d) \in \hat{R}(F,x_0)$. To this end, let $\delta > 0$ be given. There exists $X_1 \in N(x_0)$ such that F is isotone on X_1. For a given $U \in N(0)$, there exists $X_2 \in N(x_0)$ and $\lambda_0 > 0$ such that $X_2 + (0,\lambda_0)(y_i+U) \subset X_1$, $i = 1,2$. Keeping in mind that $K(X) = X$ or $K(X) = x_0$, we conclude that $* \, U \in N(0)$, $\# \, X \in N(x_0)$, $\# \, \mu > 0$, $\$ \, \lambda > 0$, there exist $W \in K(X) \times K((F(x_0)-\mu, F(x_0)+\mu))$ and $Z \in M(U)$, $\#' \, (x,r) \in W \cap (X_2 \times (F(x_0)-\mu, F(x_0)+\mu)) \cap \operatorname{epi} F$, $\$' \, t \in (0,\min(\lambda,\lambda_0))$, $*' \, h \in Z$,

$$\frac{F(x+t(y_2+h))-r}{t} \leq d + \varepsilon .$$

Now $x+t(y_2+h) \geq x+t(y_1+h)$ and both $x+t(y_2+h)$ and $x+t(y_1+h)$ are in X_1. By isotonicity of F on X_1,

$$\frac{F(x+t(y_1+h))-r}{t} \leq \frac{F(x+t(y_2+h))-r}{t} ,$$

and hence $(y_1,d) \in \hat{R}(F,x_0)$, as required. □

For stage 2 of the proofs in section 2.3, we require two subdifferential calculus rules from convex analysis. The first is a hybrid version of the subgradient sum formula (Theorem 2.1.1) and a well-known chain rule ([Rol, Theorem

23.9], [E2, Proposition 5.7]). It is easily proven by successively applying these two rules. We provide a proof of the second result.

<u>Theorem 2.2.11</u>: Let E , E^1 be l.c.s., and let $f_1: E \to \bar{\mathbb{R}}$ and $f_2: E^1 \to \bar{\mathbb{R}}$ be proper convex functions. Let $A: E \to E^1$ be continuous and linear, and suppose $x_0 \in \mathrm{dom}\, f_1 \cap A^{-1}(\mathrm{dom}\, f_2)$. Assume

$$(2.2.6) \qquad A(\mathrm{dom}\, f_1) \cap \mathrm{int}\,\mathrm{dom}\, f_2 \neq \emptyset .$$

Then

$$(2.2.7) \qquad \partial(f_1 + f_2 \circ A)(x_0) = \partial f_1(x_0) + A^* \partial f_2(Ax_0)$$

If $E := \mathbb{R}^m$ and $E^1 := \mathbb{R}^n$, then (2.2.6) may be replaced by

$$(2.2.8) \qquad A(\mathrm{ri}\,\mathrm{dom}\, f_1) \cap \mathrm{ri}\,\mathrm{dom}\, f_2 \neq \emptyset .$$

<u>Theorem 2.2.12</u>: Let $f_i: E \to \bar{\mathbb{R}}$, $i = 1,\ldots,n$ be proper convex functions, and let $F: \mathbb{R}^n \to \bar{\mathbb{R}}$ be proper, convex and isotone on $\mathbb{R}^n$. Call $f := (f_1,\ldots,f_n)$. Suppose $x_0 \in \bigcap_{i=1}^{n} \mathrm{dom}\, f_i$ and $f(x_0) \in \mathrm{dom}\, F$. Assume

$$(2.2.9) \qquad \mathrm{int}(\mathrm{dom}\, F) \cap (\mathrm{range}\, f + \mathbb{R}^n) \neq \emptyset .$$

Then

$$(2.2.10) \qquad \partial(F \circ f)(x_0) = \{\partial(\lambda \cdot f)(x_0) \mid \lambda \geq 0,\ \lambda \in \partial F(f(x_0))\} .$$

If in addition

$$(2.2.11) \qquad \operatorname{dom} f_1 \cap \bigcap_{i=2}^{n} \operatorname{int} \operatorname{dom} f_i \neq \emptyset ,$$

then

$$(2.2.12) \qquad \partial(F\circ f)(x_0) = \{\lambda\cdot(\partial f_1(x_0),\ldots,\partial f_n(x_0)) \,|\, \lambda \geq 0 , \quad \lambda \in \partial F(f(x_0))\} .$$

In the special case in which $E = \mathbb{R}^m$, (2.2.11) may be replaced by

$$(2.2.13) \qquad \bigcap_{i=1}^{n} \operatorname{ri} \operatorname{dom} f_i \neq \emptyset .$$

Proof: Let $\lambda \in \partial F(f(x_0))$, $\lambda \geq 0$, and $x^* \in \partial(\lambda\cdot f)(x_0)$. Then

$$\begin{aligned} \langle x-x_0, x^*\rangle &\leq (\lambda\cdot f)(x) - (\lambda\cdot f)(x_0) \\ &= \lambda\cdot(f(x)-f(x_0)) \\ &\leq F(f(x))-F(f(x_0)) \\ &= (F\circ f)(x)-(F\circ f)(x_0) . \end{aligned}$$

Thus $x^* \in \partial(F\circ f)(x_0)$, so

$$\begin{aligned} &\{\lambda\cdot(\partial f_1(x_0),\ldots, f_n(x_0)) \,|\, \lambda \in \partial F(f(x_0)), \lambda \geq 0\} \\ &\subset \{\partial(\lambda\cdot f)(x_0) \,|\, \lambda\in\partial F(f(x_0)), \lambda\geq 0\} \subset \partial(F\circ f)(x_0) . \end{aligned}$$

Conversely, suppose $x_0^* \in \partial(F\circ f)(x_0)$.

Call $h := F\,f$. Since F is isotone,
$h(x) = \min\{F(y) \mid f(x) \leq y\}$ for all $x \in \bigcap_{i=1}^{n} \operatorname{dom} f_i$. Then
$h^*(x_0^*) := \sup_x \langle x, x_0^* \rangle - h(x) = \sup_{x,y} \{\langle x, x_0^* \rangle - F(y) \mid y - f(x) \geq 0\}$.
Assumption (2.2.9) guarantees the existence of a Slater point for the above concave program, so we may apply the Lagrange multiplier theorem (e.g. [Hol, 14G], [Rol, Theorem 28.2]): There exists $\lambda \geq 0$ such that

$$h^*(x_0^*) = \sup_{x,y} \langle x, x_0^* \rangle - F(y) + \lambda \cdot (y - f(x))$$
$$= F^*(\lambda) + (\lambda \cdot f)^*(x_0^*) .$$

Then

$$\langle x, x_0^* \rangle = h(x_0) + h^*(x_0^*)$$
$$= F(f(x_0)) + (\lambda \cdot f)^*(x_0^*) + F^*(\lambda) ,$$

and so

$$\langle x, x_0^* \rangle + \lambda \cdot f(x_0) = \lambda \cdot f(x_0) + (\lambda \cdot f)^*(x_0^*) + F(f(x_0)) + F^*(\lambda) .$$

We now have

$$\langle x, x_0^* \rangle = \lambda \cdot f(x_0) + (\lambda \cdot f)^*(x_0^*)$$

and

$$\lambda \cdot f(x_0) = F(f(x_0)) + F^*(\lambda) ,$$

implying that

$$\lambda \in \partial F(f(x_0)) \quad \text{and} \quad x_0^* \in \partial(\lambda \cdot f)(x_0) .$$

Thus $\partial(F\circ f)(x_0) \subset \{\partial(\lambda\cdot f)(x_0) \mid \lambda \in \partial F(f(x_0)), \lambda \geq 0\}$. If in addition (2.2.11) holds (or (2.2.13) if $E := \mathbb{R}^m$), then the subgradient sum formula for n functions (see [Rol, Theorem 23.8] for the finite-dimensional version) gives $x_0^* \in \partial(\lambda\cdot f)(x_0) = \lambda\cdot(\partial f_1(x_0),\ldots,\partial f_n(x_0))$, so that (2.2.12) holds. □

Remark 2.2.13: Since $\operatorname{int}(\operatorname{dom} F) \neq \emptyset$ (because F is proper and isotone), (2.2.9) is equivalent in the case $E := \mathbb{R}^m$ to the seemingly weaker condition

$$\operatorname{ri}(\operatorname{dom} F) \cap \operatorname{ri}(\operatorname{range} f) \neq \emptyset$$

2.3. The main theorems and their corollaries

Theorem 2.3.1: Let $F\colon \mathbb{R}^n \to \mathbb{R}^m$ be strictly differentiable at x_0, $f_1\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ finite and strictly l.s.c. at x_0, and $f_2\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ finite and strictly l.s.c. at $F(x_0)$. Assume

$$(2.3.1) \qquad \nabla F(x_0)\operatorname{dom} f_1^{\uparrow}(x_0;\cdot) - \operatorname{dom} f_2^{\uparrow}(x_0;\cdot) = \mathbb{R}^m .$$

Then for all $y \in \mathbb{R}^n$,

$$(2.3.2) \qquad (f_1+f_2\circ F)^{\uparrow}(x_0;y) \leq f_1^{\uparrow}(x_0;y) + f_2^{\uparrow}(F(x_0);\nabla F(x_0)y) .$$

Moreover,

$$(2.3.3) \qquad \partial(f_1+f_2\circ F)(x_0) \subset \partial f_1(x_0) + (\nabla F(x_0))^T \partial f_2(F(x_0)) .$$

Proof: Call $f := f_1 + f_2 \circ F$. Then

$$\text{epi } f = \{(x_1,r) \times \mathbb{R}^n \times \mathbb{R} \mid f_1(x_1) \le r_1, f_2(x_2) \le r_2 ,$$
$$r = r_1 + r_2 , F(x_1) - x_2 = 0, \text{ for some } x_2 \in \mathbb{R}^m, r_1, r_2 \in \mathbb{R}\} .$$

Define A , G and z_0 as in Lemma 2.2.6, and define $C := \text{epi } f_1 \times \text{epi } f_2$. Then $A(C \cap G^{-1}(0)) = \text{epi } f$. Now

$$\text{epi } f^{\uparrow}(x_0;\cdot) = T_{A(C\cap G^{-1}(0))}(x_0,f(x_0)) \quad \text{by } (2.1.9)$$
$$\supset A(T_{C\cap G^{-1}(0)}(z_0)) \quad \text{by Lemma 2.2.6 and Proposition 2.2.4.}$$

Next observe that (2.3.1) and Corollary 1.4.3 ensure that $\nabla G(z_0) T_C(z_0) = \mathbb{R}^m$. We can therefore apply Corollary 2.2.3 to obtain

$$T_{C\cap G^{-1}(0)}(z_0) \supset T_C(z_0) \cap \nabla G(z_0)^{-1}(0) .$$

Thus

$$A(T_{C\cap G^{-1}(0)}) \supset A(T_C(z_0) \cap \nabla G(z_0)^{-1}(0))$$
$$= A((\text{epi } f_1^{\uparrow}(x_0;\cdot) \times \text{epi } f_2^{\uparrow}(x_0;\cdot)) \cap \nabla G(z_0)^{-1}(0))$$
$$\text{(by Corollary 1.4.3)}$$
$$= A(\{(h_1,r_1,h_2,r_2) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \times \mathbb{R} \mid f_1^{\uparrow}(x_0;h_1) \le r_1 ,$$
$$f_2^{\uparrow}(F(x_0);h_2) \le r_2 , \nabla F(x_0)h_1 = h_2\})$$
$$= \{(h, r_1+r_2) \in \mathbb{R}^n \times \mathbb{R} \mid f_1^{\uparrow}(x_0;h) \le r_1 ,$$
$$f_2^{\uparrow}(F(x_0); \nabla F(x_0)h) \le r_2\} .$$
$$= \text{epi}[f_1^{\uparrow}(x_0;\cdot) + f_2^{\uparrow}(F(x_0); \nabla F(x_0)(\cdot))] .$$

Therefore $\operatorname{epi} f^{\dagger}(x_0;\cdot) \supset \operatorname{epi}[f_1^{\dagger}(x_0;\cdot) + f_2^{\dagger}(F(x_0);\nabla F(x_0)(\cdot))]$, and so (2.3.2) holds.

The rest of the proof is much as in [Ro3]. Set $p_1 := f_1^{\dagger}(x_0;\cdot)$ and $p_2^{\dagger} := f_2(F(x_0);\cdot)$. If either $p_1(0)$ or $p_2(0)$ is $-\infty$, equality holds in (2.3.3), since (2.3.2) shows that both sides of the inclusion are then empty. Assume, then, that $p_1(0) = p_2(0) = 0$. Then

$$\begin{aligned}\partial f(x_0) &= \{z \in \mathbb{R}^n \mid (f_1+f_2\circ F)^{\dagger}(x_0;y) \geq \langle y,z\rangle \quad \forall y \in \mathbb{R}^n\} \\ &\subset \{z \mid p_1(y) + \nabla F(x_0)^T p_2(y) \geq \langle y,z\rangle \quad \forall y \in \mathbb{R}^n\} \\ &= \partial(p_1 + \nabla F(x_0)^T p_2)(0) .\end{aligned}$$

Since p_1 and p_2 are convex and proper, and since (2.3.1) holds, we may apply Theorem 2.2.11 to obtain

$$\begin{aligned}\partial(p_1 + \nabla F(x_0)^T p_2)(0) &= \partial p_1(0) + \nabla F(x_0)^T \partial p_2(0) \\ &= \partial f_1(x_0) + \nabla F(x_0)^T \partial f_2(F(x_0))\end{aligned}$$

by definition, and so (2.3.3) holds. □

We can now obtain improved versions, in the finite-dimensional case, of results in [Ro2] and [Ro3].

<u>Definition 2.3.2</u>: Let $C \subset E$. (a) The <u>indicator function</u> of C , denoted i_C , is defined by $i_C(x) := \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{otherwise} \end{cases}$.
(b) The <u>normal cone</u> to C at x_0 is the set

$$N_C(x_0) := T_C(x_0)^0 := \{z \in E' \mid \langle y,z\rangle \leq 0 \quad \forall y \in T_C(x_0)\} .$$

Observe that for $x_0 \in C$, $i_C^{\uparrow}(x_0;\cdot) = i_{T_C(x_0)}(\cdot)$ and $\partial i_C(x_0) = N_C(x_0)$.

Our first corollary of Theorem 2.3.1 is a strengthening of [Ro2, Theorem 5] which has also been proved by Aubin [A3].

Corollary 2.3.3: Let $F: \mathbb{R}^n \to \mathbb{R}^m$ be strictly differentiable at x_0, and let $C_1 \subset \mathbb{R}^n$ be closed near x_0 and $C_2 \subset \mathbb{R}^m$ be closed near $F(x_0)$. Suppose

$$(2.3.4) \qquad \nabla F(x_0) T_{C_1}(x_0) - T_{C_2}(F(x_0)) = \mathbb{R}^m .$$

Then

$$(2.3.5) \qquad T_{C_1 \cap F^{-1}(C_2)}(x_0) \supset T_{C_1}(x_0) \cap \nabla F(x_0)^{-1} T_{C_2}(F(x_0))$$

and

$$(2.3.6) \qquad N_{C_1 \cap F^{-1}(C_2)}(x_0) \subset N_{C_1}(x_0) + \nabla F(x_0)^T N_{C_2}(F(x_0)) .$$

Proof: Let $f_1 := i_{C_1}$, $f_2 := i_{C_2}$ in Theorem 2.3.1. Then (2.3.1) becomes (2.3.4), and (2.3.5) and (2.3.6) follow from (2.3.2) and (2.3.3), respectively. □

Remark 2.3.4: If $F := A$ is linear, (2.3.4) can be weakened to

$$(2.3.7) \qquad A(T_{C_1}(x_0)) - T_{C_2}(Ax_0) = \mathrm{span}(AC_1 - C_2) .$$

To see this, suppose $C \subset \mathbb{R}^p$, $x_0 \in C$. For a given affine set S with $\text{aff}\, C \subset S \subset \mathbb{R}^p$, denote by $T_C^S(x_0)$ the Clarke tangent cone of C at x_0 where C is considered as a subset of S rather than as a subset of $\mathbb{R}^p$. It is easy to see that $T_C^S(x_0) = T_C(x_0)$. Without loss of generality, we may assume $x_0 = 0$, since $T_{C-x_0}(0) = T_C(x_0)$ and $x_0 \in C_1 \cap A^{-1}(C_2)$ if and only if $0 \in (C-x_0) \cap A^{-1}(C_2-Ax_0)$. Call $V := \text{span}\, C_1$ and $W := \text{span}(AC_1-C_2)$. If (2.3.7) holds, then $A(T_{C_1}(x_0))-T_{C_2}^W(A(x_0)) = W$ and A may be considered as $A: V \to W$. By Corollary 2.3.3, $T_{C_1 \cap A^{-1}(C_2)}^V(x_0) \subset T_{C_1}^V(x_0) \cap A^{-1}T_{C_2}^W(Ax_0)$, which is the same as (2.3.5).

<u>Corollary 2.3.5</u>: Let $f_1: \mathbb{R}^n \to \bar{\mathbb{R}}$, $f_2: \mathbb{R}^n \to \bar{\mathbb{R}}$ be finite and strictly l.s.c. at x_0. Assume

$$\text{(2.3.8)} \qquad \text{dom}\, f_1^{\uparrow}(x_0;\cdot) - \text{dom}\, f_2^{\uparrow}(x_0;\cdot) = \mathbb{R}^n .$$

Then for all $y \in \mathbb{R}^n$,

$$\text{(2.3.9)} \qquad (f_1+f_2)^{\uparrow}(x_0;y) \le f_1^{\uparrow}(x_0;y) + f_2^{\uparrow}(x_0;y) .$$

Moreover,

$$\text{(2.3.10)} \qquad \partial(f_1+f_2)(x_0) \subset \partial f_1(x_0) + \partial f_2(x_0) .$$

<u>Proof</u>: Set $m = n$ and $F := I$ in Theorem 2.3.1. □

Remark 2.3.6: (a) Corollary 2.3.5 is more general than the specialization of [Ro3, Theorem 2] to strictly l.s.c. functions with finite-dimensional domains, since Rockafellar's assumption

$$(2.3.11) \qquad \operatorname{dom} f_1^{\uparrow}(x_0;\cdot) \cap \operatorname{int} \operatorname{dom} f_2^{\uparrow}(x_0;\cdot) \neq \phi$$

implies, but is not implied by, (2.3.8). Here is an example which satisfies (2.3.8) but not (2.3.11): Define $f_1: \mathbb{R}^2 \to \mathbb{R}$, $f_2: \mathbb{R}^2 \to \mathbb{R}$ by $f_1(x,y) := |x|^{1/2}$, $f_2(x,y) := |y|^{1/2}$. Then $\operatorname{dom} f_1^{\uparrow}(0;\cdot) = 0 \times \mathbb{R}$ and $\operatorname{dom} f_2^{\uparrow}(0;\cdot) = \mathbb{R} \times 0$, both of which have empty interior.

(b) Condition (2.3.8) can actually be weakened to

$$(2.3.12) \qquad \operatorname{dom} f_1^{\uparrow}(x_0;\cdot) - \operatorname{dom} f_2^{\uparrow}(x_0;\cdot) = \operatorname{span}(\operatorname{dom} f_1 - \operatorname{dom} f_2).$$

To see this, simply observe that we can again assume that $x_0 = 0$, and we can then replace $\mathbb{R}^n$ by $S := \operatorname{span}(\operatorname{dom} f_1 - \operatorname{dom} f_2)$, considering f_1 and f_2 as $f_1: S \to \overline{\mathbb{R}}$ and $f_2: S \to \overline{\mathbb{R}}$.

It is not possible, however, to weaken the hypothesis still further to

$$(2.3.13) \qquad \operatorname{ri} \operatorname{dom} f_1^{\uparrow}(x_0;\cdot) \cap \operatorname{ri} \operatorname{dom} f_2^{\uparrow}(x_0;\cdot) \neq \phi \,.$$

This is a consequence of Theorem 1.2.6. If Corollary 2.3.5 held under hypothesis (2.3.13), then for any closed sets

$C_1, C_2 \subset \mathbb{R}^n$ with $\mathrm{ri}\ T_{C_1}(x_0) \cap \mathrm{ri}\ T_{C_2}(x_0) \neq \phi$, the inclusion $T_{C_1}(x_0) \cap T_{C_2}(x_0) \subset T_{C_1 \cap C_2}(x_0)$ would hold, contradicting Theorem 1.2.6.

(c) If f_1 and f_2 are not strictly l.s.c. at x_0, then (2.3.8) may hold without (2.3.9) or (2.3.10) being satisfied. For example, let $C_1 := \mathbb{Q}$, the set of rational numbers, and let $C_2 := (\mathbb{R}\backslash\mathbb{Q}) \cup \{0\}$. Define $f_1 := i_{C_1}$, $f_2 := i_{C_2}$, and let $x_0 := 0$. Then $T_{C_1}(0) = T_{C_2}(0) = \mathbb{R}$, so (2.3.8) holds, and $f_1^\uparrow(x_0;\cdot) = f_2^\uparrow(x_0;\cdot) = i_{\mathbb{R}}(\cdot)$. However, $C_1 \cap C_2 = \{0\}$, so $(f_1+f_2)^\uparrow(x_0;\cdot) = i_{\{0\}}(\cdot)$, and (2.3.9) does not hold for $y \neq 0$.

<u>Corollary 2.3.7</u>: Let $F: \mathbb{R}^n \to \mathbb{R}^m$ be strictly differentiable at $x_0 \in \mathbb{R}^n$, and let $f: \mathbb{R}^m \to \overline{\mathbb{R}}$ be finite and strictly l.s.c. at $F(x_0)$. Assume

(2.3.14) $$\nabla F(x_0)\mathbb{R}^n - \mathrm{dom}\ f^\uparrow(x_0;\cdot) = \mathbb{R}^m .$$

Then for all $y \in \mathbb{R}^n$,

(2.3.15) $$(f \circ F)^\uparrow(x_0;y) \le f_2^\uparrow(F(x_0), \nabla F(x_0)y) .$$

Moreover,

(2.3.16) $$\partial(f \circ F)(x_0) \subset (\nabla F(x_0))^T \partial f(F(x_0)) .$$

<u>Proof</u>: Set $f_1 := i_{\mathbb{R}^n}$, $f_2 := f$ in Theorem 2.3.1 . □

Remark 2.3.8: (a) It is also possible (but rather complicated) to derive Corollary 2.3.7 from Corollary 2.3.5 (see the proof of [Ro3, Theorem 3]).

(b) Corollary 2.2.3, which we used to derive Theorem 2.3.1, can itself be derived from Corollary 2.3.3 by setting $m := q$, $n := p$, $F := G$, $C_1 := C$, and $C_2 := \{0\}$.

We will also in this section derive subdifferential calculus rules for n functions. To do so, we must first find n-function analogues of conditions like (2.3.8).

Definition 2.3.9: Let $C \subset \mathbb{R}^p$. Define

$$\Delta^n C := \{(x_1,\ldots,x_n) \mid x_i \in C,\ x_1 = x_2 = \ldots = x_n\} .$$

Definition 2.3.10: Let $C_i \subset \mathbb{R}^m$, $i = 1,\ldots,n$ be convex sets. The sets C_i are said to be in strong general position [Za1] if

$$(2.3.17) \qquad 0 \in \operatorname{int}[\Delta^{n-1} C_1 - \prod_{j=2}^{n} C_j] .$$

An equivalent way to write (2.3.17) is

$$(2.3.18) \qquad 0 \in \operatorname{int}[\Delta^n \mathbb{R}^m - \prod_{j=1}^{n} C_j] .$$

Writing the condition as in (2.3.18) shows that the order in which the sets in Definition 2.3.10 are listed is irrelevant. The equivalence of (2.3.17) and (2.3.18) is proven in [Za1], which gives a thorough discussion of this concept.

If the sets C_i are cones, (2.3.17) is of course equivalent to

$$\Delta^{n-1} C_1 - \prod_{j=2}^{n} C_j = \mathbb{R}^{(n-1)m} . \tag{2.3.19}$$

The concept of convex cones in strong general position turns out to be just what we need for the results in the remainder of this section. Our first such result is an important special case of Corollary 2.3.3:

<u>Proposition 2.3.11</u>: Let $D_i \subset \mathbb{R}^m$, $i = 1,\ldots,n$ be locally closed near $y_0 \in \bigcap_{i=1}^{n} D_i$. Assume $T_{D_i}(y_0)$, $i = 1,\ldots,n$ are in strong general position. Then

$$T_{D_1\cap\ldots\cap D_n}(y_0) \supset \bigcap_{i=1}^{n} T_{D_i}(y_0) \tag{2.3.20}$$

and

$$N_{D_1\cap\ldots\cap D_n}(y_0) \subset \sum_{i=1}^{n} N_{D_i}(y_0) . \tag{2.3.21}$$

<u>Proof</u>: Call $C_1 := D_1 \times\ldots\times D_n$, and let C_2 be the origin in $\mathbb{R}^{(n-1)m}$. Define $F: \mathbb{R}^{nm} \to \mathbb{R}^{(n-1)m}$ by $F(x_1,\ldots,x_n) := (x_1-x_2,\ldots,x_1-x_n)$. Now apply Corollary 2.3.3 with $x_0 = (y_0,\ldots,y_0)$. In this case, condition (2.3.4) is equivalent to the sets $T_{D_i}(y_0)$, $i = 1,\ldots,n$ being in strong general position. By (2.3.5),

$$T_{\Delta^n(D_1 \cap \dots \cap D_n)}(y_0,\dots,y_0) \supset (T_{D_1}(y_0)\times\dots\times T_{D_n}(y_0)) \cap \Delta^n \mathbb{R}^m ,$$

or $\Delta^n T_{D_1 \cap \dots \cap D_n}(y_0) \supset \Delta^n \bigcap_{i=1}^{n} T_{D_i}(y_0)$. Thus (2.3.20) holds.

Finally, (2.3.21) follows from (2.3.6). □

Remark 2.3.12: (a) See Watkins' paper [W1] for an alternate proof of inclusion (2.3.20).

(b) By (2.3.7), the strong general position assumption can be weakened to

$$(2.3.22) \qquad \Delta^{n-1} T_{D_1}(y_0) - \prod_{j=2}^{n} T_{D_j}(y_0) = \prod_{j=2}^{n} \operatorname{span}(D_1 - D_j) .$$

We can derive from Proposition 2.3.11 a formula for the subgradient of $f(x) := \max_{1\le i\le n} f_i(x)$, where $f_i : \mathbb{R}^m \to \overline{\mathbb{R}}$ are strictly l.s.c. We start with a lemma which we will use to show that the condition "dom $f_i^{\uparrow}(x_0;\cdot)$, $i = 1,\dots,n$, are in general position" is sufficient to guarantee this subgradient formula.

Lemma 2.3.13: Suppose $f_i : \mathbb{R}^m \to \overline{\mathbb{R}}$, $i = 1,\dots,n$ are such that

$$\Delta^{n-1} \operatorname{dom} f_1 - \prod_{i=2}^{n} \operatorname{dom} f_i = \mathbb{R}^{m(n-1)} .$$

Then

$$\Delta^{n-1} \operatorname{epi} f_1 - \prod_{i=2}^{n} \operatorname{epi} f_i = (\mathbb{R}^m \times \mathbb{R})^{n-1} .$$

<u>Proof</u>: We must show $(\mathbb{R}^m \times \mathbb{R})^{n-1} \subset \Delta^{n-1}\text{epi } f_1 - \prod_{i=2}^{n} \text{epi } f_i$.
Let $(y_1, s_1, \ldots, y_{n-1}, s_{n-1}) \in (\mathbb{R}^m \times \mathbb{R})^{n-1}$. Then there exist $x_i \in \text{dom } f_i$, $i = 1,\ldots,n$ such that $x_1 - x_i = y_{i-1}$, $i = 2,\ldots,n$. Choose r_i, $i = 1,\ldots,n$, such that $r_i \geq f_i(x_i)$. If $r_1 - r_2 < s_1$, replace r_1 by $s_1 + r_2$. If $r_1 - r_2 > s_1$, replace r_2 by $r_1 - s_1$. Now if $r_1 - r_3 > s_2$, replace r_3 by $r_1 - s_2$. If $r_1 - r_3 < s_2$, replace r_1 by $s_2 + r_3$ and r_2 by $s_2 + r_3 - s_1$. Proceed in this manner. After k-1 steps, we have $r_1 - r_i = s_{i-1}$, $r = 2,\ldots,k$. If $r_1 - r_{k+1} > s_k$ replace r_{k+1} by $r_1 - s_k$. If $r_1 - r_{k+1} < s_k$, replace r_1 by $s_k + r_{k+1}$ and replace r_j, $j = 2,\ldots,k$ by $s_k + r_{k+1} - s_{j-1}$. After n-1 steps, we obtain $r_1 - r_i = s_{i-1}$, $i = 2,\ldots,n$, and $(x_i, r_i) \in \text{epi } f_i$. Thus $(y_j, s_j) \in \Delta^{n-1}\text{epi } f_1 - \prod_{i=2}^{n} \text{epi } f_i$ and so $(\mathbb{R}^m \times \mathbb{R})^{n-1} \subset \Delta^{n-1}\text{epi } f_1 - \prod_{i=2}^{n} \text{epi } f_i$. □

We will also use the fact that if $f: E \to \overline{\mathbb{R}}$ and $x \in E$ are such that $\partial f(x) \neq \phi$, then

$$(2.3.23) \qquad N_{\text{epi } f}(x, f(x)) = \bigcup_{\lambda > 0} \lambda(\partial f(x), -1) \cup 0^+(\partial f(x), -1),$$

where "0^+" denotes "recession cone" ([Ro1, Section 8], [Ro3, page 350]). Following Rockafellar, we denote the right hand side of (2.3.23) by $\bigcup_{\lambda \geq 0^+} \lambda(\partial f(x), -1)$.

<u>Proposition 2.3.14</u>: Let $f_i : \mathbb{R}^m \to \overline{\mathbb{R}}$, $i = 1,\dots,n$ be finite and strictly l.s.c. at x_0. Define $f(x) := \max_{1\le i\le n} f_i(x)$ and $I(x) := \{i\in\{1,\dots,n\} \mid f_i(x) = f(x)\}$. Suppose that f_i is continuous at x_0 for each $i \notin I(x_0)$, and suppose that $\operatorname{dom} f_i^{\uparrow}(x_0;\cdot)$, $i \in I(x_0)$, are in strong general position. Then for all $y \in \mathbb{R}^m$,

$$(2.3.24) \qquad f^{\uparrow}(x_0;y) \le \max_{I(x_0)} f_i^{\uparrow}(x_0;y) .$$

If also $\partial f_i(x_0)$, $i \in I(x_0)$, are nonempty, then

$$(2.3.25) \qquad \partial f(x_0) \subset \cup\{ \sum_{I(x_0)} \lambda_i \partial f_i(x_0) \mid \lambda_i \ge 0^+,\ \sum_{I(x_0)} \lambda_i = 1\}.$$

<u>Proof</u>: Call $D_i := \operatorname{epi} f_i$, $i = 1,\dots,n$. Since f_j, $j \in I(x_0)$, are strictly l.s.c. at x_0, and f_j, $j \notin I(x_0)$, are continuous at x_0, $T_{D_j}((x_0,f(x_0))) = \mathbb{R}^{m+1}$ for all $j \notin I(x_0)$. Thus we only need to consider D_j with $j \in I(x_0)$. By Lemma 2.3.13, our assumption that $\operatorname{dom} f_i^{\uparrow}(x_0;\cdot)$, $i = 1,\dots,n$, are in strong general position implies that $T_{D_i}((x_0,f(x_0)))$, $i \in I(x_0)$, are also in strong general position. We may then apply Proposition 2.3.11. By (2.3.20),

$$\operatorname{epi} f^{\uparrow}(x_0;\cdot) \supset \operatorname{epi} \max_{i\in I(x_0)} f_i^{\uparrow}(x_0;\cdot) ,$$

giving (2.3.24). If also $\partial f_i(x_0)$, $i \in I(x_0)$, are nonempty, we have by (2.3.21) and (2.3.23) that

$$\bigcup_{\lambda \geq 0^+} \lambda(\partial f(x_0), -1) \subset \sum_{i \in I(x_0)} \bigcup_{\lambda_i \geq 0^+} \lambda_i(\partial f_i(x_0), -1) .$$

To obtain (2.3.25), set $\lambda = 1$ on the left hand side of the above inclusion. □

Our second main result is the following chain rule:

<u>Theorem 2.3.15</u>: Let $f_i : \mathbb{R}^m \to \overline{\mathbb{R}}$, $i = 1, \ldots, n$ be finite and strictly l.s.c. at x_0, and define $f := (f_1, \ldots, f_n)$. Let $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at $f(x_0)$, isotone on the union of a neighbourhood of $f(x_0)$ and Range $f + \mathbb{R}^n_+$, and l.s.c. Assume that $f_j^\uparrow(x_0;0) = 0$, $j = 1, \ldots, n$, and that for all $j \in \{1, \ldots, n\}$, either

(2.3.26) f_j is continuous at x_0,

or

(2.3.27) F is strictly isotone in the jth coordinate at $f(x_0)$.

Assume also that the following condition holds:

(2.3.28) For each $(s_1, \ldots, s_{n-1}, t_1, \ldots, t_n)$ in $\mathbb{R}^{(n-1)m+n}$, there exist $d_i \in \mathbb{R}^m$, $r_i \in \mathbb{R}$, $i = 1, \ldots, n$, satisfying $f_i^\uparrow(x_0;d_i) \leq r_i$ and $(h_1, \ldots, h_n) \in \operatorname{dom} F^\uparrow(f(x_0);\cdot)$ such that

(i) $d_1 - d_i = s_{i-1}$, $i = 2,\dots,n$ and

(ii) $r_i - h_i = t_i$, $i = 1,\dots,n$.

Then

(2.3.29) $(F\circ f)^{\uparrow}(x_0;y) \le F^{\uparrow}(f(x_0); f_1^{\uparrow}(x_0;y),\dots,f_n^{\uparrow}(x_0;y))$ for all $y \in \mathbb{R}^m$

and

(2.3.30) $\partial(F\circ f)(x_0) \subset \{\lambda\cdot(\partial f_1(x_0),\dots,\partial f_n(x_0)) \mid \lambda \in \partial F(f(x_0)), \lambda \ge 0\}$.

Proof: Call $h := F\circ f$. Since F is isotone on Range $f + \mathbb{R}^n_+$,

$$\text{epi } h = \{(x,r) \in \mathbb{R}^m\times\mathbb{R} \mid (y_1,\dots,y_n) \in \mathbb{R}^n \text{ with } F(y_1,\dots,y_n) \le r,\ f_i(x) \le y_i,\ 1\le i\le n\}.$$

Define the functions G and A as in Lemma 2.2.8, and define

$$D := \text{epi } f_1 \times\dots\times \text{epi } f_n \times \text{epi } F$$

$$C := D \cap G^{-1}(0).$$

Then epi $h = A(C)$. Assumptions (2.3.26) and (2.3.27) allow us to apply Lemma 2.2.8 and Proposition 2.2.4. Thus

$$\text{epi } h^{\uparrow}(x_0;\cdot) = T_{A(C)}(x_0,h(x_0)) \quad \text{(by (2.1.9))},$$

$\supset A(T_C(z_0))$, where $z_0 := (x_0, f_1(x_0), \ldots, x_0, f_n(x_0),$ $f_1(x_0), \ldots, f_n(x_0), h(x_0))$ (by Lemma 2.2.8 and Proposition 2.2.4).

$\supset A(T_D(z_0) \cap \nabla G(z_0)^{-1}(0))$, by Corollary 2.2.3, since (2.3.28) says exactly that $\nabla G(z_0) T_D(z_0) = \mathbb{R}^{(n-1)m+n}$,

$= \{(x,r) \in \mathbb{R}^m \times \mathbb{R} \mid \exists y \in \mathbb{R}^n \text{ with } f_i^{\uparrow}(x_0;x) \le y_i,\ F^{\uparrow}(f(x_0);y) \le r\}$

$= \{(x,r) \mid F^{\uparrow}(f(x_0);\ f_1^{\uparrow}(x_0;x), \ldots, f_n^{\uparrow}(x_0;x)) \le r\}$,

since $F^{\uparrow}(f(x_0);\cdot)$ is itself isotone (Lemma 2.2.10). Thus $\text{epi } h^{\uparrow}(x_0;\cdot) \supset \text{epi } F^{\uparrow}(f(x_0);\ f_1^{\uparrow}(x_0;\cdot), \ldots, f_n^{\uparrow}(x_0;\cdot))$, and so (2.3.29) holds.

Now if $F^{\uparrow}(f(x_0);0) = -\infty$, (2.3.29) shows that both sides of (2.3.20) are empty. Assume, then, that $F^{\uparrow}(f(x_0);\cdot)$ is proper. Since each $f_i^{\uparrow}(x_0;\cdot)$ is convex and proper and (2.3.28) (i) and (ii) imply that (2.2.9) and (2.2.13) hold, we may apply Theorem 2.2.12. By (2.3.29) and Theorem 2.2.12, we have

$$\begin{aligned}
\partial(F\circ f)(x_0) &= \{z \in \mathbb{R}^m \mid (F\circ f)^{\uparrow}(x_0;y) \ge \langle y,z\rangle \quad \forall y \in \mathbb{R}^m\} \\
&\subset \{z \in \mathbb{R}^m \mid F^{\uparrow}(f(x_0);\ f_1^{\uparrow}(x_0;y), \ldots, f_n^{\uparrow}(x_0;y)) \ge \langle y,z\rangle \\
&\qquad\qquad \forall\, y \in \mathbb{R}^m\} \\
&= \partial((F^{\uparrow}(f(x_0);\cdot)\circ(f_1^{\uparrow}(x_0;\cdot), \ldots, f_n^{\uparrow}(x_0;\cdot)))(0) \\
&= \{\lambda\cdot(\partial f_1^{\uparrow}(x_0;\cdot)(0), \ldots, \partial f_n^{\uparrow}(x_0;\cdot)(0)) \mid \lambda \ge 0 \text{ and} \\
&\qquad\qquad \lambda \in \partial F^{\uparrow}(f(x_0);\cdot)(0)\} \\
&= \{\lambda\cdot(\partial f_1(x_0), \ldots, \partial f_n(x_0)) \mid \lambda \ge 0,\ \lambda \in \partial F(f(x_0))\} .
\end{aligned}$$

Therefore (2.3.30) holds, and the proof is complete. $\square$

Remark 2.3.16: (a) Assumption (2.3.28) reduces to a simple, familiar-looking form in important special cases. For example, if $n = 1$, (2.3.28) becomes.

$$\text{(2.3.31)} \qquad \text{Range } f_1^{\uparrow}(x_0;\cdot) - \text{dom } F^{\uparrow}(f_1(x_0);\cdot) = \mathbb{R},$$

and if F is as in Remark 2.2.9 (a) or (b), (2.3.28) reduces to the assumption that dom $f_i^{\uparrow}(x_0;\cdot)$, $i = 1,\ldots,n$ are in strong general position.

(b) The assumption that $f_i^{\uparrow}(x_0;0) = 0$, $i = 1,\ldots,n$ is not needed in important special cases. For example, if F is as in Remark 2.2.9 (a) or (b) and $f_i^{\uparrow}(x_0;0) = -\infty$ for some i, then both sides of (2.3.30) will equal ϕ.

(c) Corollaries of Theorem 2.3.15 include Corollary 2.3.5, Proposition 2.3.11, and Proposition 2.3.14 (without having to use Lemma 2.3.13).

(d) Any isotone function $F: \mathbb{R}^n \to \overline{\mathbb{R}}$ which is finite at $x \in \mathbb{R}^n$ is directionally Lipschitzian at x with respect to any $y \leq 0$; i.e., $F^I(x;y) < +\infty$ for all $y \leq 0$ (see [Ro4, Proposition 4]). This property plays a key role in the subdifferential calculus results of [Ro3]. However, there is no analogue of Theorem 2.3.15 if F is merely assumed to be directionally Lipschitzian, as we will see in Chapter 3.

From Theorem 2.3.15 we can derive an extension of Corollary 2.3.5 to n functions and a product rule for functions which are nonnegative on $\mathbb{R}^m$ and strictly l.s.c. and positive at x_0.

<u>Corollary 2.3.17</u>: Let $f_i : \mathbb{R}^m \to \bar{\mathbb{R}}$, $i = 1,\ldots,n$ be finite and strictly l.s.c. at x_0 , and suppose that $\operatorname{dom} f_i^\uparrow(x_0;\cdot)$, $i = 1,\ldots,n$ are in strong general position. Then for all $y \in \mathbb{R}^m$,

$$(2.3.22) \qquad (f_1+\ldots+f_n)^\uparrow(x_0;y) \le \sum_{i=1}^n f_i^\uparrow(x_0;y) .$$

Moreover,

$$(2.3.33) \qquad \partial(f_1+\ldots+f_n)(x_0) \subset \sum_{i=1}^n \partial f_i(x_0) .$$

<u>Proof</u>: Let $F(x_1,\ldots,x_n) = \sum_{i=1}^n x_i$ in Theorem 2.3.15. F is continuous and strictly isotone in each coordinate, so (2.3.27) is satisfied. Assumption (2.3.28) in this case reduces to $\operatorname{dom} f_i^\uparrow(x_0;\cdot)$, $i = 1,\ldots,n$ being in strong general position. As explained in Remark 2.3.16 (b), the assumption that each $f_i^\uparrow(x_0;0) = 0$ is not needed in this case. Then (2.3.22) follows from (2.3.29) and (2.3.33) follows from (2.3.30). □

<u>Corollary 2.3.18</u>: Let $f_i : \mathbb{R}^m \to \bar{\mathbb{R}}$, $i = 1,\ldots,n$ be non-negative on $\mathbb{R}^m$ and strictly l.s.c. and positive at $x_0 \in \bigcap_{i=1}^n \operatorname{dom} f_i$. Suppose $\operatorname{dom} f_i^\uparrow(x_0;\cdot)$, $i = 1,\ldots,n$ are in strong general position. Then for all $y \in \mathbb{R}^m$,

$$(2.3.34) \qquad \Big(\prod_{i=1}^n f_i\Big)^\uparrow(x_0;y) \le \sum_{i=1}^n \Big(\prod_{j\ne i} f_j(x_0)\Big) f_i^\uparrow(x_0;y) .$$

Moreover,

$$\partial\Big(\prod_{i=1}^{n} f_i\Big)(x_0) \subset \sum_{i=1}^{n}\Big(\prod_{j\neq i} f_j(x_0)\Big)\partial f_i(x_0) . \tag{2.3.25}$$

Proof: Let $F(x_1,\ldots,x_n) = \prod_{i=1}^{n} x_i$ in Theorem 2.3.15. Condition (2.3.28) again reduces to $\operatorname{dom} f_i^{\uparrow}(x_0;\cdot)$, $i = 1,\ldots,n$ being in strong general position. As explained in Remark 2.3.16 (b), the assumption that each $f_i^{\uparrow}(x_0;0) = 0$ is not needed. Then (2.3.24) follows from (2.3.29) and (2.3.25) from (2.3.30). □

We can also derive a quotient rule as a special case of Corollary 2.3.18. The first step is to obtain a formula for the subgradient of a function of the form $\frac{1}{g}$.

Lemma 2.3.19: Suppose $g: E \to \overline{\mathbb{R}}$ is continuous and positive at $x_0 \in \operatorname{dom} g$. Then for all $y \in E$,

$$\Big(\frac{1}{g}\Big)^{\uparrow}(x_0;y) = \frac{(-g)^{\uparrow}(x_0;y)}{(g(x_0))^2} . \tag{2.3.36}$$

Moreover,

$$\partial\Big(\frac{1}{g}\Big)(x_0) = \frac{1}{(g(x_0))^2}\,\partial(-g)(x_0) . \tag{2.3.37}$$

Proof: First observe that (2.3.37) follows immediately from (2.3.36) because $\frac{1}{(g(x_0))^2} > 0$. To prove (2.3.36), use (1.5.21) and the notation of [Ro3] and section 3.5 to write

$$\left(\frac{1}{g}\right)^{\uparrow}(x_0;y) = \limsup_{\substack{x\to x_0\\ t\downarrow 0}} \inf_{y'\to y} \frac{\frac{1}{g(x+ty')} - \frac{1}{g(x)}}{t}$$

$$= \limsup_{\substack{x\to x_0\\ t\downarrow 0}} \inf_{y'\to y} \frac{(-g)(x+ty')-(-g)(x)}{t} \cdot \frac{1}{g(x+ty')g(x)}$$

$$= \limsup_{\substack{x\to x_0\\ t\downarrow 0}} \inf_{y'\to y} \frac{(-g)(x+ty')-(-g)(x)}{t} \lim_{\substack{x\to x_0\\ t\downarrow 0\\ y'\to y}} \frac{1}{g(x+ty')g(x)}$$

(by Lemma 3.5.5, which we prove in chapter 3)

$$= \frac{1}{(g(x_0))^2}(-g)^{\uparrow}(x_0;y) \quad . \qquad \square$$

<u>Proposition 2.3.20</u>: Let $f: \mathbb{R}^m \to \overline{\mathbb{R}}$ and $g: \mathbb{R}^m \to \overline{\mathbb{R}}$ be non-negative on $\mathbb{R}^m$, with f strictly l.s.c. and positive at x_0 and g continuous and positive at $x_0 \in \operatorname{dom} f \cap \operatorname{dom} g$. Assume

(2.3.38) $$\operatorname{dom} f^{\uparrow}(x_0;\cdot) - \operatorname{dom}(-g)^{\uparrow}(x_0;\cdot) = R^m .$$

Then

(2.3.39) $$\left(\frac{f}{g}\right)^{\uparrow}(x_0;y) \le \frac{f(x_0)(-g)^{\uparrow}(x_0;y)+g(x_0)f^{\uparrow}(x_0;y)}{(g(x_0))^2}$$

for all $y \in \mathbb{R}^m$, and

(2.3.40) $$\partial\left(\frac{f}{g}\right)(x_0) \subset \frac{f(x_0)\partial(-g)(x_0)+g(x_0)\partial f(x_0)}{(g(x_0))^2} .$$

Proof: By Lemma 2.3.19, $\operatorname{dom}(\frac{1}{g})^{\uparrow}(x_0;\cdot) = \operatorname{dom}(-g)^{\uparrow}(x_0;\cdot)$, so (2.3.38) ensures that we can apply Corollary 2.3.18 with $n = 2$, $f_1 := f$, and $f_2 := \frac{f}{g}$. By (2.3.34) and (2.3.36),

$$\begin{aligned}(\tfrac{f}{g})^{\uparrow}(x_0) &\le f(x_0)(\tfrac{1}{g})^{\uparrow}(x_0;y) + \frac{1}{g(x_0)} f^{\uparrow}(x_0;y)\\ &= \frac{f(x_0)}{(g(x_0))^2}(-g)^{\uparrow}(x_0;y) + \frac{1}{g(x_0)} f^{\uparrow}(x_0;y)\\ &= \frac{f(x_0)(-g)^{\uparrow}(x_0;y) + g(x_0)f^{\uparrow}(x_0;y)}{(g(x_0))^2}\end{aligned}$$

Similarly, by (2.3.35) and (2.3.37),

$$\begin{aligned}\partial(\tfrac{f}{g})(x_0) &\subset f(x_0)\partial(\tfrac{1}{g})(x_0) + \frac{1}{g(x_0)}\partial f(x_0)\\ &= \frac{f(x_0)}{(g(x_0))^2}\partial(-g)(x_0) + \frac{1}{g(x_0)}\partial f(x_0)\\ &= \frac{f(x_0)\partial(-g)(x_0) + g(x_0)\partial f(x_0)}{(g(x_0))^2}.\end{aligned}$$

□

Remark 2.3.21: If g is locally Lipschitzian near x_0, $(-g)^{\uparrow}(x_0;\cdot) = -g^{\uparrow}(x_0;\cdot)$, $\partial(-g)(x_0) = -\partial g(x_0)$, and (2.3.38) is automatically satisfied. Proposition 2.3.20 then reduces to a quotient rule similar to that of [Hil, Chapter 8], [Cl4, Chapter 2].

2.4. Directional derivative inequalities and subdifferential regularity

Under what circumstances will equality hold in the inequalities and inclusions of section 2.3? From our discussion in sections 1.5 and 2.1, we know that one way to ensure equality in these formulae is to assume that the functions involved are convex and proper. In this section, we derive more general conditions for equality, slightly weakening the conditions given in [Ro3]. Our strategy is to prove directional derivative inequalities involving $f_+(x_0;\cdot) := f^K(x_0;\cdot)$ and $f_\square(x_0;\cdot) := f^k(x_0;\cdot)$, then compare them with the inequalities of section 2.3 to find conditions for equality. The results of the convex subdifferential calculus come out as corollaries.

We begin with a counterpart to Theorem 2.3.15. Notice that it holds quite generally, with no need for a "constraint qualification".

<u>Proposition 2.4.1</u>: Let E be a l.c.s., let $f_i : E \to \overline{\mathbb{R}}$, $i = 1,\dots,n$ be finite at x_0, and call $f := (f_1,\dots,f_n)$. Let $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite at $f(x_0)$ and isotone on the union of Range $f + \mathbb{R}^n_+$ and $B_\varepsilon(f(x_0))$ for some $\varepsilon > 0$. Then if $y \in E$ is such that $(f_1)_+(x_0;y)$, $i = 1,\dots,n$ are finite,

$$(2.4.1) \qquad (F\circ f)_+(x_0;y) \geq F_+(f(x_0);(f_1)_+(x_0;y),\dots,(f_n)_+(x_0;y)).$$

Alternatively, if $(f_1)_\square(x_0;y)$, $(f_i)_+(x_0;y)$, $i = 2,\ldots,n$ are finite, then

$$(2.4.2)\qquad (F\circ f)_\square(x_0;y) \geq F_+(f(x_0);(f_1)_\square(x_0;y),(f_2)_+(x_0;y),\ldots,(f_n)_+(x_0;y)) .$$

Proof: Let $(y,r) \in K_{\mathrm{epi}(F\circ f)}((x_0,(F\circ f)(x_0)))$. To prove (2.4.1), it suffices to show that

$$((f_1)_+(x_0;y),\ldots,(f_n)_+(x_0;y),r) \in K_{\mathrm{epi}\ F}((f(x_0),F(f(x_0)))).$$

To this end, let $\lambda > 0$ and $\delta > 0$ be given. Call $z_i = (f_i)_+(x_0;y)-\delta$, and choose $\lambda_0 \in (0,\lambda)$ such that $f_i(x_0) + [0,\lambda_0]z_i \geq f_i(x_0)-\varepsilon$. Then there exist $Y_i \in N(y)$, $\lambda_1 \in (0,\lambda_0)$ such that for all $t \in (0,\lambda_1)$ and all $y' \in Y_i$,

$$\frac{f_i(x_0+ty')-f_i(x_0)}{t} \geq z_i \ , \ i = 1,\ldots,n \ ,$$

In addition, there exist $\bar{y} \in \bigcap_{i=1}^{n} Y_i$, $\bar{t} \in (0,\min_i \lambda_i)$ such that

$$\frac{(F\circ f)(x_0+\bar{t}\bar{y})-(F\circ f)(x_0)}{\bar{t}} \leq r + \delta \ .$$

Now since $f_i(x_0)+\bar{t}z_i \leq f_i(x_0+t\bar{y})$ and F is isotone,

$$\frac{F(f(x_0)+\bar{t}(z_1,\ldots,z_n))-F(f(x_0))}{\bar{t}} \leq r + \delta \ .$$

Thus $((f_1)_+(x_0;y),\ldots,(f_n)_+(x_0;y),r) \in K_{\text{epi } F}((f(x_0),$ $F(f(x_0))))$, and so (2.4.1) holds.

Now let $(y,r) \in K_{\text{epi}(F\circ f)}((x_0,(F\circ f)(x_0)))$. To prove (2.4.2), it suffices to show that $((f_1)_{\Box}(x_0;y),(f_2)_+(x_0;y),$ $\ldots,(f_n)_+(x_0;y),r) \in K_{\text{epi } F}((f(x_0),F(f(x_0)))))$. To this end, let $\delta > 0$ and $\lambda > 0$ be given. Call $z_1 = (f_1)_{\Box}(x_0;y)-\delta$, $z_i = (f_i)_+(x_0;y) - \delta, i = 2,\ldots,n$. Again choose $\lambda_0 \in (0,\lambda)$ such that $f_i(x_0) + [0,\lambda_0]z_i \geq f_i(x_0)-\varepsilon$. There exist $Y_i \in N(y)$, $\lambda_i \in (0,\lambda_0)$ such that for all $t \in (0,\lambda_i)$ and all $y' \in Y_i$, $\frac{f_i(x_0+ty')-f(x_0)}{t} \geq z_i$, $i = 2,\ldots,n$, and there exists $Y_1 \in N(y)$ such that for all $t > 0$, there exists $t' \in (0,t)$ with $\frac{f_1(x_0+t'y')-f(x_0)}{t'} \geq z_1$ for all $y' \in Y_1$. In addition, there exists $\bar{y} \in \bigcap_{i=1}^{n} Y_i$ and $\lambda_1 \in (0,\lambda_0)$ such that for all $t \in (0,\lambda_1)$, $\frac{(F\circ f)(x_0+t\bar{y})-(F\circ f)(x_0)}{t} \leq r + \delta$. Thus for some $\bar{t} \in (0,\min_i \lambda_i)$, $f_i(x_0) + \bar{t}z_i \leq f_i(x_0+\bar{t}\bar{y})$. By the isotonicity of F ,

$$\frac{F(f(x_0)+\bar{t}(z_1,\ldots,z_n))-F(f(x_0))}{\bar{t}} \leq r + \delta .$$

Hence $((f_1)_{\Box}(x_0;y),(f_2)_+(x_0;y),\ldots,(f_n)_+(x_0;y),r) \in$ $K_{\text{epi } F}((f(x_0),F(f(x_0)))))$, and so (2.4.2) holds. $\Box$

<u>Proposition 2.4.2:</u> Let E , E^1 be l.c.s., let $F: E \to E^1$ be strictly differentiable at $x_0 \in E$, and let $f: E^1 \to \bar{\mathbb{R}}$ be finite at $F(f(x_0))$. Then for all $y \in E^1$,

$$(2.4.3) \qquad (f\circ F)_+(x_0;y) \geq f_+(F(x_0);\nabla F(x_0)y)$$

and

$$(2.4.4) \qquad (f\circ F)_\square(x_0;y) \geq f_\square(F(x_0);\nabla F(x_0)y) .$$

<u>Proof</u>: Let $(y,r) \in K_{\text{epi}(f\circ F)}((x_0,(f\circ F)(x_0)))$. To prove (2.4.3), it suffices to show that $(\nabla F(x_0)y,r) \in K_{\text{epi } f}((F(x_0),f(F(x_0))))$. Let $Z \in N(\nabla F(x_0)y)$, $\lambda > 0$, and $\delta > 0$ be given. There exist $Y \in N(y)$ and $\lambda_1 \in (0,\lambda)$ such that $\dfrac{F(x_0+ty')-F(x_0)}{t} \in Z$ for all $y' \in Y$ and $t \in (0,\lambda_1)$. In addition, there exist $\bar{y} \in Y$ and $\bar{t} \in (0,\lambda_1)$ such that $\dfrac{(f\circ F)(x_0+\bar{t}\bar{y})-(f\circ F)(x_0)}{\bar{t}} \leq r + \delta$. Call $\bar{z} := \dfrac{F(x_0+\bar{t}\bar{y})-F(x_0)}{\bar{t}}$. Then $\bar{z} \in Z$ and

$\dfrac{f(F(x_0)+\bar{t}\bar{z})-f(F(x_0))}{\bar{t}} \leq r + \delta$. Hence $(\nabla F(x_0)y,r) \in K_{\text{epi } f}((F(x_0),f(F(x_0))))$, and (2.4.3) holds. The proof of (2.4.4) is very similar to the proof of (2.4.3), and we leave it to the reader. $\square$

We can now give conditions under which equality holds in Theorems 2.3.1 and 2.3.15.

<u>Definition 2.4.3</u>: $f: E \to \bar{\mathbb{R}}$ is said to be <u>subdifferentiably regular</u> at $x_0 \in \text{dom } f$ if $f_+(x_0;y) = f^\uparrow(x_0;y)$ for all $y \in E$. f is said to be <u>subdifferentiably weakly regular</u> at x_0 if $f_\square(x_0;y) = f^\uparrow(x_0;y)$ for all $y \in E$.

<u>Proposition 2.4.4</u>: Suppose the hypotheses of Theorem 2.3.15 are satisfied. Assume in addition that F is subdifferentiably regular at $f(x_0)$, $f_2,\dots,f_n$ are subdifferentiably regular at x_0, and f_1 is subdifferentiably weakly regular at x_0. Then equality holds in (2.3.29) and (2.3.30).

<u>Proof</u>: By (2.4.2) and our regularity hypothesis, for all $y \in \mathbb{R}^n$,

$$
\begin{aligned}
(F\circ f)^{\uparrow}(x_0;y) &\geq (F\circ f)_{\square}(x_0;y) \\
&\geq F_+(f(x_0);(f_1)_{\square}(x_0;y),(f_2)_+(x_0;y),\dots,(f_n)_+(x_0;y)) \\
&= F^{\uparrow}(f(x_0);f_1^{\uparrow}(x_0;y),\dots,f_n^{\uparrow}(x_0;y)) ,
\end{aligned}
$$

so equality holds in (2.3.29). Equality in (2.3.30) follows from equality in (2.3.29), as an inspection of the proof of Theorem 2.3.15 shows. □

<u>Proposition 2.4.5</u>: Suppose the hypotheses of Theorem 2.3.1 are satisfied, and that either

(2.4.5) f_1 is subdifferentiably regular at x_0 and f_2 is subdifferentiably weakly regular at $F(x_0)$,

or

(2.4.6) f_1 is subdifferentiably weakly regular at x_0 and f_2 is subdifferentiably regular at $F(x_0)$.

Then equality holds in (2.3.3). If in addition $f_1^{\uparrow}(x_0,0) = f_2^{\uparrow}(F(x_0);0) = 0$, equality holds in (2.3.2) also.

Proof: If either $f_1^{\uparrow}(x_0;0)$ or $f_2^{\uparrow}(F(x_0;0)$ is $-\infty$, (2.3.2) implies that $(f_1+f_2\circ F)^{\uparrow}(x_0;0) = -\infty$, and both sides of (2.3.3) are empty. Otherwise apply Proposition 2.4.1 with $n = 2$ and $F:\mathbb{R}^2 \to \mathbb{R}$ defined by $F(x_1,x_2) = x_1 + x_2$. If (2.4.5) holds, then for all $y \in \mathbb{R}^n$,

$$\begin{aligned}(f_1+f_2\circ F)^{\uparrow}(x_0;y) &\geq (f_1+f_2\circ F)_{\Box}(x_0;y) \\ &\geq (f_1)_{+}(x_0;y)+(f_2\circ F)_{\Box}(x_0;y) && \text{(by (2.4.2))} \\ &\geq (f_1)_{+}(x_0;y)+(f_2)_{\Box}(F(x_0);\nabla F(x_0)y) && \text{(by (2.4.4))} \\ &= (f_1)^{\uparrow}(x_0;y)+f_2^{\uparrow}(F(x_0);\nabla F(x_0)y) .\end{aligned}$$

If (2.4.6) holds, then similarly, for all $y \in \mathbb{R}^n$,

$$\begin{aligned}(f_1+f_2\circ F)^{\uparrow}(x_0;y) &\geq (f_1+f_2\circ F)_{\Box}(x_0;y) \\ &\geq (f_1)_{\Box}(x_0;y)+(f_2\circ F)_{+}(x_0;y) && \text{(by (2.4.2))} \\ &\geq (f_1)_{\Box}(x_0;y)+(f_2)_{+}(F(x_0);\nabla F(x_0)y) && \text{(by (2.4.3))} \\ &= (f_1)^{\uparrow}(x_0;y)+f_2^{\uparrow}(F(x_0);\nabla F(x_0)y) .\end{aligned}$$

In either case, equality holds in (2.3.2), and equality in (2.3.3) follows. □

Remark 2.4.6: The conditions in Propositions 2.4.4 and 2.4.5 guaranteeing equality in the subdifferential calculus formulae of Section 2.3 are less stringent than those given in [Ro3]. For example, let $C_1 := \bigcup_{n=0}^{\infty} [2^{2n}, 2^{2n+1}]$ and

$C := C_1 \cup (-C_1) \cup \{0\}$. Then $T_C(0) = k_C(0) = \{0\}$, but $K_C(0) = \mathbb{R}$. In Proposition 2.4.5, let $F := I$ and $x_0 := 0$, Define $f_1: \mathbb{R} \to \mathbb{R}$ by $f_1(x) := 0$, and let $f_2: \mathbb{R} \to \mathbb{R}$ be i_C. Since (2.3.8) and (2.4.5) are satisfied, equality holds in (2.3.9) and (2.3.10), even though the condition for equality given in [Ro3, Theorem 2] — that both f_1 and f_2 be subdifferentially regular at x_0 — is not satisfied. Observe also that the regularity conditions in Propositions 2.4.4 and 2.4.5 are still valid in the infinite dimensional setting of [Ro3].

As we noted in section 1.5, convex functions are subdifferentiably regular, so Propositions 2.4.4 and 2.4.5 enable us to rederive convex subdifferential calculus results, in particular the finite-dimensional versions of Theorems 2.2.11 and 2.2.12. Thus Theorems 2.3.1 and 2.3.15, which are proved by means of Theorems 2.2.11 and 2.2.12, combine with Propositions 2.4.4 and 2.4.5 to yield Theorems 2.2.11 and 2.2.12 as special cases. We list below one further convex subdifferential calculus result.

<u>Proposition 2.4.7</u>: Let $F: \mathbb{R}^n \to \mathbb{R}^m$ be strictly differentiable at x_0 , let $f_1: \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex, proper, and strictly l.s.c. at x_0, and let $f_2: \mathbb{R}^m \to \overline{\mathbb{R}}$ be convex, proper, and strictly l.s.c. at $F(x_0)$, where $x_0 \in \operatorname{dom} f_1 \cap F^{-1}(\operatorname{dom} f_2)$. Assume

$$(2.4.7) \qquad \nabla F(x_0)x_0 - F(x_0) \in \operatorname{int}(\nabla F(x_0)\operatorname{dom} f_1 - \operatorname{dom} f_2) .$$

Then for all $y \in \mathbb{R}^n$,

$$(2.4.8) \qquad (f_1+f_2\circ F)'(x_0;y) = f_1'(x_0;y)+f_2'(F(x_0);\nabla F(x_0)y) .$$

Moreover,

$$(2.4.9) \qquad \partial(f_1+f_2\circ F)(x_0) = \partial f_1(x_0)+\nabla F(x_0)^T\partial f_2(x_0) .$$

Proof: By Theorem 2.3.1 and Proposition 2.4.5, equality will hold in (2.4.8) and (2.4.9) as long as $\nabla F(x_0)\operatorname{dom} f_1^{\uparrow}(x_0;\cdot) - \operatorname{dom} f_2^{\uparrow}(F(x_0);\cdot) = \mathbb{R}^m$, which in this case is equivalent to

$$\nabla F(x_0)\operatorname{cone}(\operatorname{dom} f_1-x_0) - \operatorname{cone}(\operatorname{dom} f_2-F(x_0)) = \mathbb{R}^m$$

(see [Rol, Section 23]). This reduces to

$$\operatorname{cone}(F(x_0)\operatorname{dom} f_1 - \operatorname{dom} f_2 - \nabla F(x_0)x_0 + F(x_0)) = \mathbb{R}^m ,$$

which is equivalent to (2.4.7). □

Corollary 2.4.8 (cf. [Rol, Theorem 23.8]): Let $f_1: \mathbb{R}^m \to \overline{\mathbb{R}}$ and $f_2: \mathbb{R}^m \to \overline{\mathbb{R}}$ be convex, proper, and l.s.c., and let $x_0 \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$. Assume

$$(2.4.10) \qquad \operatorname{ri}\operatorname{dom} f_1 \cap \operatorname{ri}\operatorname{dom} f_2 \neq \phi$$

Then for all $y \in \mathbb{R}^m$,

$$(2.4.11) \qquad (f_1+f_2)'(x_0;y) = f_1'(x_0;y) + f_2'(x_0;y) .$$

Moreover,

(2.4.12) $$\partial(f_1+f_2)(x_0) = \partial f_1(x_0) + \partial f_2(x_0) .$$

Proof: By Proposition 2.4.7 with $m = n$ and $F := I$, (2.4.11) and (2.4.12) will hold if $\text{cone}(\text{dom } f_1 - \text{dom } f_2) = \mathbb{R}^m$. By (2.3.12) of Remark 2.3.6 (b), this assumption can be weakened to $\text{cone}(\text{dom } f_1 - \text{dom } f_2) = \text{span}(\text{dom } f_1 - \text{dom } f_2)$, which is equivalent to (2.4.10). □

Remark 2.4.9: We saw in Remark 2.3.6 (b) that hypothesis (2.3.8) of Corollary 2.3.5 cannot be weakened to (2.3.13). That is because $\text{aff dom } f^{\dagger}(x_0;\cdot)$ can be of smaller dimension than $\text{aff dom } f$, as in the example $C := \{(x,y) \in \mathbb{R}^2 | y = x^2\}$, $f: \mathbb{R}^2 \to \overline{\mathbb{R}}$ defined by $f := i_C$, and $x_0 := (0,0)$. If f is convex and proper, however, $\text{aff dom } f^{\dagger}(x_0;\cdot) = \text{aff dom } f$, and assumption (2.4.10) is sufficient to give (2.4.11) and (2.4.12).

It is possible, by the techniques of section 2.3, to prove analogues of (2.3.2) and (2.3.29) involving $f_+(x_0;\cdot)$ and $f_{\square}(x_0;\cdot)$. To do so, we first need two counterparts of Corollary 2.2.3.

Theorem 2.4.10: Let $G: \mathbb{R}^p \to \mathbb{R}^q$ be strictly differentiable at $x_0 \in G^{-1}(0)$, and let $C \subset \mathbb{R}^p$ be closed near x_0. Assume

(2.4.13) $$\nabla G(x_0) T_C(x_0) = \mathbb{R}^q .$$

Then

$$(2.4.14) \qquad K_C(x_0) \cap \nabla G(x_0)^{-1}(0) \subset K_{C\cap G^{-1}(0)}(x_0)$$

and

$$(2.4.15) \qquad k_C(x_0) \cap \nabla G(x_0)^{-1}(0) \subset k_{C\cap G^{-1}(0)}(x_0) .$$

Proof: The inclusion (2.4.14) is a special case of [Bo5, Theorem 4.1(a)]. The proof of (2.4.15) is entirely analogous to that of [Bo5, Theorem 4.1]. □

Theorem 2.4.11: Let $f_i : \mathbb{R}^m \to \overline{\mathbb{R}}$, $i = 1,\ldots,n$ be finite and strictly l.s.c. at x_0 , and call $f := (f_1,\ldots,f_n)$. Let $F: \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite and strictly l.s.c. at $f(x_0)$ and isotone on the union of Range $f + \mathbb{R}^n_+$ and $B_\varepsilon(f(x_0))$ for some $\varepsilon > 0$. Assume condition (2.3.28) of Theorem 2.3.15 holds. Then for all $y \in \mathbb{R}^m$,

$$(2.4.16) \qquad (F\circ f)_\square(x_0;y) \le F_\square(f(x_0);(f_1)_\square(x_0;y), \ldots,(f_n)_\square(x_0;y))$$

and

$$(2.4.17) \qquad (F\circ f)_+(x_0;y) \le F_\square(f(x_0);(f_1)_+(x_0;y),(f_2)_\square(x_0;y), \ldots,(f_n)_\square(x_0;y)) .$$

Equality holds in (2.4.16) and (2.4.17) if $(f_i)_\square(x_0;y) = (f_i)_+(x_0;y)$, $i = 2,\ldots,n$ for all $y \in \mathbb{R}^m$, $F_\square(f(x_0);z) = F_+(f(x_0);z)$ for all $z \in \mathbb{R}^n$, and $(f_i)_+(x_0;y)$, $i = 1,\ldots,n$ are never $-\infty$.

Proof of (2.4.16): Let $h := F\circ f$, and define A and G as in Lemma 2.2.8 and Theorem 2.3.15. Define $D := \text{epi } f_1 \times \ldots \times \text{epi } f_n \times \text{epi } F$ and $C := D \cap G^{-1}(0)$. As in Theorem 2.3.15, $\text{epi } h = A(C)$. Then

$$\begin{aligned}\text{epi } h_{\square}(x_0;\cdot) &= k_{A(C)}((x_0,h(x_0)))\\ &\supset A(k_C(z_0))\end{aligned}$$

where $z_0 := (x_0,f_1(x_0),\ldots,x_0,f_n(x_0),f_1(x_0),\ldots,f_n(x_0),h(x_0))$, by inclusion (1.4.15) of Corollary 1.4.9. Now assumption (2.3.28) guarantees that $\nabla G(x_0)T_D(z_0) = \mathbb{R}^{(n-1)m+n}$, so we may apply Theorem 2.4.10. By inclusion (2.4.15) and Corollary 1.4.3,

$$\begin{aligned}A(k_C(z_0)) &\supset A(k_D(z_0) \cap \nabla G(z_0)^{-1}(0))\\ &= \{(x,r) \in \mathbb{R}^m \times \mathbb{R} \mid \exists\, y \in \mathbb{R}^n \text{ with}\\ &\quad (f_i)_{\square}(x_0;x) \le y_i,\ F_{\square}(f(x_0);y) \leqq r\}\\ &= \text{epi } F_{\square}(f(x_0);(f_1)_{\square}(x_0;\cdot),\ldots,(f_n)_{\square}(x_0;\cdot))\end{aligned}$$

since $F_{\square}(f(x_0);\cdot)$ is itself isotone by Lemma 2.2.10. Hence $\text{epi } h_{\square}(x_0;\cdot) \supset \text{epi } F_{\square}(f(x_0);(f_1)_{\square}(x_0;\cdot),\ldots,(f_n)_{\square}(x_0;\cdot))$, and (2.4.16) holds. The proof of (2.4.17) is similar and uses (1.4.14), (2.4.14), and (1.4.6). That equality holds in (2.4.16) and (2.4.17) under the given assumptions is a consequence of Proposition 2.4.1. □

Theorem 2.4.12: Let $F: \mathbb{R}^n \to \mathbb{R}^m$ be strictly differentiable at x_0, let $f_1: \mathbb{R}^m \to \mathbb{R}$ be finite and strictly l.s.c. at

x_0 , and let $f_2 : \mathbb{R}^m \to \bar{\mathbb{R}}$ be finite and strictly l.s.c. at $F(x_0)$. Assume (2.3.1) holds. Then for all $y \in \mathbb{R}^n$,

$$\begin{aligned}(2.4.18)\qquad (f_1+f_2\circ F)_{\Box}(x_0;y) &\le (f_1)_{\Box}(x_0;y) \\ &\quad + (f_2)_{\Box}(F(x_0);\nabla F(x_0)y)\end{aligned}$$

and

$$\begin{aligned}(2.4.19)\qquad (f_1+f_2\circ F)_{+}(x_0;y) &\le (f_1)_{\Box}(x_0;y) \\ &\quad + (f_2)_{+}(F(x_0);\nabla F(x_0)y)\end{aligned}$$

$$\begin{aligned}(2.4.20)\qquad (f_1+f_2\circ F)_{+}(x_0;y) &\le (f_1)_{+}(x_0;y) \\ &\quad + (f_2)_{\Box}(F(x_0);\nabla F(x_0)y) .\end{aligned}$$

If $(f_1)_+(x_0;\cdot)$ and $(f_2)_+(F(x_0);\nabla F(x_0)(\cdot))$ are never $-\infty$, equality holds in (2.4.18) and (2.4.19) if $(f_1)_{\Box}(x_0;\cdot) = (f_1)_+(x_0;\cdot)$ and in (2.4.20) if $(f_2)_{\Box}(F(x_0);\nabla F(x_0)(\cdot)) = (f_2)_+(F(x_0);\nabla F(x_0)(\cdot))$.

<u>Proof</u>: The proofs of (2.4.18), (2.4.19), and (2.4.20) are analogous to that of Theorem 2.3.1, using Theorem 2.4.10, Corollary 1.4[illegible], and Corollaries 1.4.3 and 1.4.5. That equality holds under the given assumptions is a consequence of Propositions 2.4.1 and 2.4.2. □

We give here two consequences of Theorem 2.4.12.

<u>Corollary 2.4.13</u>: Let $C_1 \subset \mathbb{R}^n$ be closed near x_0 , let $F: \mathbb{R}^n \to \mathbb{R}^m$ be strictly differentiable at x_0 , and let $C_2 \subset \mathbb{R}^m$ be closed near $F(x_0)$. Assume that (2.3.4) holds.

Then

$$(2.4.21) \qquad k_{C_1 \cap F^{-1}(C_2)}(x_0) = k_{C_1}(x_0) \cap \nabla F(x_0)^{-1} k_{C_2}(F(x_0)) ,$$

$$(2.4.22) \qquad K_{C_1}(x_0) \cap \nabla F(x_0)^{-1} K_{C_2}(F(x_0)) \supset K_{C_1 \cap F^{-1}(C_2)}(x_0) \supseteq k_{C_1}(x_0) \cap \nabla F(x_0)^{-1} K_{C_2}(F(x_0)) ,$$

and

$$(2.4.23) \qquad K_{C_1 \cap F^{-1}(C_2)}(x_0) \supset K_{C_1}(x_0) \cap \nabla F(x_0)^{-1} k_{C_2}(F(x_0)).$$

<u>Proof</u>: In (2.4.21), $k_{C_1 \cap F^{-1}(C_2)}(x_0) \subset k_{C_1}(x_0) \cap \nabla F(x_0)^{-1} k_{C_2}(F(x_0))$ always holds. The opposite inclusion follows from (2.4.18) with $f_1 := i_{C_1}$ and $f_2 := i_{C_2}$. The first inclusion in (2.4.22) always holds, while the second is (2.4.19) with $f_1 := i_{C_1}$ and $f_2 := i_{C_2}$. Similarly, (2.4.23) follows from (2.4.20).

<u>Corollary 2.4.14</u>: Let $f_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$, $f_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite and strictly l.s.c. at $x_0 \in \mathbb{R}^n$. Assume (2.3.8) holds. Then for all $y \in \mathbb{R}^n$,

$$(2.4.24) \qquad (f_1+f_2)_{\square}(x_0;y) \le (f_1)_+(x_0;y) + (f_2)_{\square}(x_0;y)$$

and

$$(2.4.25) \qquad (f_1+f_2)_+(x_0;y) \le (f_1)_+(x_0;y) + (f_2)_{\square}(x_0;y) .$$

Equality holds in (2.4.24) if $(f_1)_{\square}(x_0;\cdot) = (f_1)_+(x_0;y)$ and if $(f_1)_+(x_0;\cdot)$ and $(f_2)_+(x_0;\cdot)$ are never equal to $-\infty$.

<u>Proof</u>: Take $m = n$ and $F := I$ in Theorem 2.4.12. □

In the special case in which $f_{\square}(x_0;\cdot)$ is convex, we can obtain results like the following:

<u>Proposition 2.4.15</u>: Let $f_1: \mathbb{R}^n \to \overline{\mathbb{R}}$ and $f_2: \mathbb{R}^n \to \overline{\mathbb{R}}$ be finite and strictly l.s.c. at x_0. Assume that $(f_1)_{\square}(x_0;\cdot)$ and $(f_2)_{\square}(x_0;\cdot)$ are convex, and that (2.3.8) holds. Then

$$(2.4.26) \qquad \partial^k(f_1+f_2)(x_0) \subset \partial^k f_1(x_0) + \partial^k f_2(x_0) .$$

Equality holds in (2.4.26) if $(f_1)_{\square}(x_0;\cdot) = (f_1)_+(x_0;\cdot)$ or $(f_2)_{\square}(x_0;\cdot) = (f_2)_+(x_0;\cdot)$ - in particular, if f_1 or f_2 is convex.

<u>Proof</u>: Since (2.3.8) holds, so does (2.4.24), and (2.4.26) follows from (2.4.24). Equality under the given assumptions follows from Corollary 2.4.14 and the fact that if either $(f_1)_{\square}(x_0;\cdot)$ or $(f_2)_{\square}(x_0;\cdot)$ is not proper, both sides of (2.4.26) are empty. □

It is interesting to note that assumptions involving $f^{\uparrow}(x_0;\cdot)$ are required in these results on $f_+(x_0;\cdot)$ and $f_{\square}(x_0;\cdot)$. The same is true of Ioffe's results on approximate subdifferentials [I2].

We now present an example which satisfies the hypotheses of Proposition 2.4.15 even though f_1 and f_2 are not convex. A similar example is given in [Ma2].

Example 2.4.16: Define $f_1: \mathbb{R} \to \mathbb{R}$ by

$$f_1(x) := \begin{cases} 0 & \text{if } x \le 0 \\ \frac{1}{2^{n+1}} & \text{if } \frac{1}{2^{n+1}} < x \le \frac{1}{2^n}, \quad n = 0, \pm 1, \pm 2, \ldots, \end{cases}$$

and define $f_2: \mathbb{R} \to \mathbb{R}$ by

$$f_2(x) := \begin{cases} 0 & \text{if } x \le 0 \\ \frac{1}{4^{(n+1)}} & \text{if } \frac{1}{4^{n+1}} < x \le \frac{1}{4^n}, \quad n = 0, \pm 1, \pm 2, \ldots, \end{cases}$$

It is not hard to see that

$$f_1^{\uparrow}(0;y) = f_2^{\uparrow}(0;y) = \begin{cases} 0 & \text{if } y \le 0 \\ +\infty & \text{if } y > 0 \end{cases},$$

$$(f_1)_{\square}(0;y) = (f_2)_{\square}(0;y) = \begin{cases} 0 & \text{if } y \le 0 \\ y & \text{if } y > 0 \end{cases},$$

$$(f_1+f_2)_{\square}(0;y) = \begin{cases} 0 & \text{if } y \le 0 \\ 2y & \text{if } y > 0 \end{cases}, \text{ and}$$

$$(f_1)_{+}(0;y) = \begin{cases} 0 & \text{if } y \le 0 \\ y/2 & \text{if } y > 0 \end{cases},$$

$$(f_2)_{+}(0;y) = \begin{cases} 0 & \text{if } y \le 0 \\ y/4 & \text{if } y > 0 \end{cases},$$

$$(f_1+f_2)_+(0;y) = \begin{cases} 0 & \text{if } y \le 0 \\ \frac{3y}{4} & \text{if } y \ge 0 \end{cases} .$$

Then $\operatorname{dom} f_1^{\uparrow}(0;\cdot) = \operatorname{dom} f_2^{\uparrow}(0;\cdot) = \{y \mid y \le 0\}$, so (2.3.8) is satisfied and (2.4.26) holds. In fact $\partial_{\square}(f_1+f_2)(0) = [0,2]$ while $\partial_{\square}f_1(0) = \partial_{\square}f_2(0) = [0,1]$, so equality holds in (2.4.26) even though $(f_i)_+(x_0;\cdot) \ne (f_i)_{\square}(x_0;\cdot)$, $i = 1,2$.

If $f_{\square}(x_0;\cdot)$ is not convex, it is still possible to define a notion of subgradient for f by means of upper convex approximates.

<u>Definition 2.4.17</u> (cf. [Ps1]): The function $h: E \to \overline{\mathbb{R}}$ is an <u>upper convex approximate</u> for $f: E \to \overline{\mathbb{R}}$ at $x_0 \in \operatorname{dom} f$ if

(a) $h(y) \ge f_{\square}(x_0;y)$ for all $y \in E$ and

(b) $h(\cdot)$ is convex, l.s.c., and positively homogeneous.

The set

$$\partial_h^k(x_0) := \{x' \in E' \mid h(y) \ge \langle y,x'\rangle \quad \forall y \in E\}$$

is called a <u>h-subgradient</u> of f at x_0.

Specific examples of upper convex approximates include $f_{\square}(x_0;\cdot)$, $f^0(x_0;\cdot)$, $f^I(x_0;\cdot)$, and the upper convex approximations of [Ps1]. As one might expect, not as much can be said about general upper convex approximates as can be determined about specific ones like $f^0(x_0;\cdot)$ or $f^{\uparrow}(x_0;\cdot)$; however, necessary conditions for minimality can be expressed in terms of upper convex approximates, as we will show in Chapter 5. We postpone further discussion of them until then.

Remark 2.4.18: By comparing Proposition 2.4.1 and Corollary 2.4.14, one can see the implications of the slight difference in quantification in the definitions of $K_C(x_0)$ and $k_C(x_0)$. From Proposition 2.4.1, we know that the inequality

$$(f_1+f_2)_+(x_0;\cdot) \geq (f_1)_+(x_0;\cdot) + (f_2)_+(x_0;\cdot)$$

always holds; however, it does not necessarily hold with $f_+(x_0;\cdot)$ replaced by $f_\Box(x_0;\cdot)$, as the following example illustrates: Consider $f_1: \mathbb{R} \to \mathbb{R}$ and $f_2: \mathbb{R} \to \mathbb{R}$ defined by

$$f_1(x) := \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } \frac{1}{2^{n+1}} < x \leq \frac{1}{2^n} \text{ and } n \text{ is odd} \\ 2x & \text{if } \frac{1}{2^{n+1}} \leq x \leq \frac{1}{2^n} \text{ and } n \text{ is even} \end{cases}$$

$$f_2(x) := \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } \frac{1}{2^{n+1}} < x \leq \frac{1}{2^n} \text{ and } n \text{ is even} \\ 2x & \text{if } \frac{1}{2^{n+1}} < x \leq \frac{1}{2^n} \text{ and } n \text{ is odd} \end{cases}$$

Then if $y \geq 0$,

$$(f_1)_\Box(0;y) = (f_2)_\Box(0;y) = 2y ,$$

while

$$(f_1+f_2)(x) = \begin{cases} 0 & \text{if } x \le 0 \\ 3x & \text{if } x > 0 \end{cases},$$

so that $(f_1+f_2)_\Box(0;y) = 3y$. Thus for $y > 0$,

$$(f_1+f_2)_\Box(0;y) = 3y < 4y = (f_1)_\Box(0;y) + (f_2)_\Box(0;y) .$$

On the other hand, Corollary 2.4.14 shows that the inequality

$$(f_1+f_2)_\Box(x_0;y) \le (f_1)_\Box(x_0;y) + (f_2)_\Box(x_0;y)$$

holds under assumption (2.3.8), but the fact that $K_C(x_0)$ is not product-preserving prevents us from proving the same inequality for $f_+(x_0;\cdot)$ under that assumption.

Does there exist a q-cone R such that

$$(f_1+f_2)^R(x_0;\cdot) = f_1{}^R(x_0;\cdot) + f_2{}^R(x_0;\cdot)$$

holds in general? The answer is no. The only R for which

$$(f_1+f_2)^R(x_0;\cdot) \ge f_1{}^R(x_0;\cdot) + f_2{}^R(x_0;\cdot)$$

holds in general satisfy Property (6) and can be expressed as "lim inf's" of difference quotients. On the other hand, the only q-cones for which

$$(f_1+f_2)^R(x_0;\cdot) \le f_1{}^R(x_0;\cdot) + f_2{}^R(x_0;\cdot)$$

holds in general satisfy Property (5) and can be expressed as "lim sups" of difference quotients. The regularity conditions given in this section seem to be the best possible involving q-cone directional derivatives.

We saw in Theorem 1.2.10 that Properties (5) and (6) are incompatible for a broad class of tangency operators. The fact that

$$(f_1+f_2)^R(x_0;\cdot) = f_1^R(x_0;\cdot) + f_2^R(x_0;\cdot)$$

cannot be made to hold in general seems to be a reflection of Theorem 1.2.10. It also reflects the fact that the approximations we are studying are <u>one-sided</u>. On the other hand, for <u>two-sided</u> approximations to be effective, much more is required of the functions than mere lower semicontinuity (see [Do1], [U1]).

2.5. Subdifferential calculus - the Banach space case

The key finite-dimensional results of this chapter - Theorems 2.3.1, 2.3.15, 2.4.11, and 2.4.12 - rely heavily upon Theorem 2.2.2. There is also a Banach space version of Theorem 2.2.2 given in [Bo5], which will enable us to apply the methods of this chapter to derive subdifferential calculus formulae for extended real valued functions defined on a Banach space. In this brief section we state this theorem and give a few of its applications to subdifferential calculus.

Definition 2.5.1: Let E be a l.c.s. and $C \subset E$ a convex set. The core (or algebraic interior) of C is the set

$$\text{core } C := \{x \in C \mid \forall\ y \in E,\ \exists\ \delta > 0 \text{ such that } x + \delta y \in C\} .$$

Definition 2.5.2 [Ro4]: Let E be a l.c.s. (a) The set $C \subseteq E$ is said to be epi-Lipschitzian at $x_0 \in C$ if $I_C(x_0) \neq \phi$. (b) The function $f: E \to \bar{R}$ is said to be directionally Lipschitzian at $x_0 \in E$ if epi f is epi-Lipschitzian at $(x_0, f(x_0))$.

If f is directionally Lipschitzian at x_0, then $f^I(x_0;y) < +\infty$ for some $y \in E$, and the set of all such y comprises the interior of dom $f^{\dagger}(x_0;\cdot)$ (see [Ro3], [Ro4], [Ro5]).

Theorem 2.5.3 [Bo5]: Let E and E^1 be Banach spaces, let $G: E \to E^1$ be strictly differentiable at $x_0 \in G^{-1}(0)$, and let C be closed near x_0 . Assume either

(2.5.1) $\quad C$ is convex and $0 \in \text{core}(\nabla G(x_0)(C-x_0))$

or

(2.5.2) $\quad C$ is epi-Lipschitzian at x_0 and $\nabla G(x_0) T_C(x_0) = E^1$.

Then

(2.5.3) $\quad T_C(x_0) \cap \nabla G(x_0)^{-1}(0) \subset T_{C \cap G^{-1}(0)}(x_0)$.

Using (2.5.1) and (2.5.3), we can readily derive formulae from the convex subdifferential calculus. We give as an example an extension of Proposition 2.4.7.

<u>Proposition 2.5.4</u>: Let E and E^1 be Banach spaces, and let $F: E \to E^1$ be strictly differentiable at $x_0 \in E$. Let $f_1: E \to \overline{\mathbb{R}}$ be convex, proper, and strictly l.s.c. at x_0, and let $f_2: E^1 \to \overline{\mathbb{R}}$ be convex, proper, and strictly l.s.c. at $F(x_0)$, where $x_0 \in \operatorname{dom} f_1 \cap F^{-1}(\operatorname{dom} f_2)$. Assume

$$(2.5.4) \quad \nabla F(x_0)x_0 - F(x_0) \in \operatorname{core}(\nabla F(x_0)\operatorname{dom} f_1 - \operatorname{dom} f_2) .$$

Then

$$(2.5.5) \quad \partial(f_1 + f_2 \circ F)(x_0) = \partial f_1(x_0) + \nabla F(x_0)^* \partial f_2(x_0) .$$

<u>Proof</u>: The proof here parallels that of Theorem 2.3.1, except that we apply Theorem 2.5.3 instead of Corollary 2.2.3 with G as in Lemma 2.2.6, $C := \operatorname{epi} f_1 \times \operatorname{epi} f_2$, and $x_0 := (x_0, f_1(x_0), F(x_0), f_2(F(x_0)))$. Condition (2.5.1) in this case reduces to (2.5.4). □

Unfortunately, Theorem 2.5.3 is not strong enough to enable us to prove the best generalized subdifferential calculus formulae for functions defined on a Banach space. For example, in using Theorem 2.5.2 to prove a subgradient sum formula, we need to assume that each of the two functions is directionally Lipschitzian, while in Theorem 2.1.2, it is only assumed that one of the two functions is directionally Lipschitzian. The best analogue of Theorem 2.3.1 obtainable via Theorem 2.5.2 is the following result. We omit the proof, which parallels that of Theorem 2.3.1.

Proposition 2.5.5: Let E and E^1 be Banach spaces, and let $F: E \to E^1$ be strictly differentiable at x_0. Let $f_1: E \to \overline{\mathbb{R}}$ be finite, epi-Lipschitzian, and strictly l.s.c. at x_0, and let $f_2: E^1 \to \overline{\mathbb{R}}$ be finite, epi-Lipschitzian, and strictly l.s.c. at $F(x_0)$. Assume

$$(2.5.6) \qquad F(x_0)\,\mathrm{dom}\, f_1^{\uparrow}(x_0;\cdot) - \mathrm{dom}\, f_2^{\uparrow}(x_0;\cdot) = E^1 .$$

Then (2.3.2) and (2.3.3) hold. Equality is ensured in (2.3.2) and (2.3.3) if the hypotheses of Proposition 2.4.5 are satisfied.

Remark 2.5.6: Although Proposition 2.5.5 is not as strong as Theorem 2.1.2, the special case of Proposition 2.5.5, in which $f_1 :\equiv 0$, is just as strong as [Ro3, Theorem 3], the corresponding corollary of Theorem 2.1.2.

CHAPTER III

Generalized Subdifferential Calculus: The Vector-Valued Case

3.1. Directional derivatives and subgradients for vector-valued functions.

In recent years, a number of different concepts of generalized derivative for vector-valued functions (e.g. [Cl2], [Cl4], [Ro5], [Th1], [Hi4], [A2], [Bo4], [Pa1], [Pa2]) and set-valued mappings ([Pe2], [Th2], [Ga1], [A1]) have been proposed and investigated. Calculus formulae for these various generalized derivatives have been proven by Clarke, Hiriart-Urruty, Penot, Thibault, and others. Of particular relevance to the present work are the papers of Thibault ([Th1], [Th2]), which generalize the main calculus formulae of [Ro3] to the vector-valued and set-valued settings. Thibault's approach, like ours, centers around directional derivatives and subgradients defined in association with tangent cones to epigraphs.

The cornerstone of our development of a generalized subdifferential calculus for vector-valued functions is a tangent cone inclusion proven in Proposition 4.4 of [Th2]. As a direct corollary of this inclusion, we will prove a vector-valued analogue of (2.3.29), our inequality involving the upper subderivative of the composition of an isotone function and n l.s.c. functions. Combining this result with subgradient inclusions for vector-valued convex functions will enable us to prove vector-valued counterparts of some of the formulae in section 2.3. Specializing to the scalar-valued case, we will then prove an infinite-dimensional version of Theorem 2.3.15 which almost subsumes the results of [Ro3] (We assume the functions involved are "l.s.c. at x_0", an assumption not made in [Ro3].) and allows us to prove new product and quotient rules for functions with infinite-dimensional domains. We also give a variant of Theorem 2.3.15 in which the complicated conditions (2.3.28) are "decoupled". Finally, we wrap up our study of generalized subdifferential calculus by investigating whether any additional chain rule formulations not already considered might be possible.

We begin by describing the setting for the discussion in this chapter. Let E be a l.c.s., and let E_1 (or (E_1,S)) denote an ordered l.c.s. (abbreviated o.t.v.s.) with an ordering $\leq_S$ induced by a closed convex cone $S \subset E_1$; i.e., for $x,y \in E_1$, $x \leq_S y$ if and only if $y-x \in S$. Such an ordering is automatically transitive. It is reflexive if $0 \in S$ and is antisymmetric if in addition S is pointed — i.e., $S \cap -S = \{0\}$. Call $\dot{E}_1 := E_1 \cup \{+\infty\}$ the space E_1 with adjunction of a supremum $+\infty$, and similarly define $\dot{S} := S \cup \{+\infty\}$. Denote by $\overline{E}_1 := E_1 \cup \{\pm\infty\}$ the space E_1 with the adjunction of a supremum $+\infty$ and an infimum $-\infty$. Addition and scalar multiplication are extended in the usual way to $\dot{E}_1$ and $\overline{E}_1$ with $(+\infty) + (-\infty) = +\infty$. We will assume throughout that $0 \in S$ and that (E_1,S) is <u>order complete</u>; i.e., every nonempty subset of E_1 which is bounded above (below) in E_1 has a supremum (infimum) in E_1.

Denote by $L(E,E_1)$ the space of all continuous linear mappings from E into E_1. For a mapping $f: E \to (\bar{E}_1,S)$, define the (effective) <u>domain</u> of f to be the set

$$\text{(3.1.1)} \qquad \operatorname{dom} f := \{x \in E \mid f(x) \neq +\infty\}$$

and the S-epigraph of f to be the set

$$\text{(3.1.2)} \qquad \operatorname{epi}_S f := \{(x,y) \in E \times E_1 \mid f(x) \leq_S y\} .$$

If $\operatorname{epi}_S f$ is a convex set, f is said to be <u>S-convex.</u> If $f: E \to (\bar{E}_1,S)$ has nonempty domain, it is called <u>proper</u>. As in the scalar-valued case, if $f: E \to (\bar{E}_1,S)$ is S-convex, $\operatorname{epi}_S f$ is closed, and there exists $x_0 \in E$ with $f(x_0) \in E_1$, then f is proper. We demonstrate this below:

<u>Proposition 3.1.1</u>: Let $f: E \to (\bar{E}_1,S)$ be an S-convex function such that $\operatorname{epi}_S f$ is closed.

If there exists $x_0 \in E$ with $f(x_0) = -\infty$, then f is equal to $-\infty$ throughout its domain.

<u>Proof</u>: Suppose $f(x_0) = -\infty$. Then $(x_0,y) \in \operatorname{epi}_S f$ for all $y \in E_1$. Let $\hat{x} \in \operatorname{dom} f$, and let $\hat{y} \in E_1$ be such that $(\hat{x},\hat{y}) \in \operatorname{epi}_S f$. Let $y \in E_1$ and $t \in [0,1]$ be given. Then

there exists $y_0 \in E_1$ such that $ty_0 + (1-t)y \leq_S y$. Since f is S-convex, $(tx_0 + (1-t)\hat{x}, y) \in \mathrm{epi}_S f$. Now let $t \downarrow 0$. Since $\mathrm{epi}_S f$ is closed, we conclude that $(\hat{x}, y) \in \mathrm{epi}_S f$. Because y is an arbitrary element of E, $f(\hat{x})$ must equal $-\infty$. □

For a proper S-convex function $f: E \to (E_1, S)$ and $x_0 \in \mathrm{dom}\, f$, define the <u>S directional derivative</u> of f at x_0 with respect to $h \in E$ by

$$(3.1.3) \qquad f'_S(x_0;h) := \inf_{t>0} {}_S \frac{f(x_0+th)-f(x_0)}{t}$$

Define the S subgradient of f at x_0 to be the set

$$(3.1.4) \qquad \partial_S f(x_0) := \{T \in L(E,E_1) \mid T(x-x_0) \leq_S f(x)-f(x_0) \text{ for all } x \in E\} .$$

Notice that $\partial_S f$ may be written equivalently as

$$\partial_S f(x_0) := \{T \in L(E, E_1) \mid T(h) \leq f'_S(x_0;h) \text{ for all } h \in E\} .$$

For information on the calculus of S subgradients, see [Zo1], [Pa1], [Ku1], [Ku2], and [Bo9].

Now we extend the definitions of section 1.5 to the above setting.

Definition 3.1.2: Let R be a q-cone, and let $f: E \to (\overline{E}_1, S)$ and $x_0 \in f^{-1}(E_1)$. Define

$R_S(f,x_0) := \{(y,d) \in E \times E_1 \mid \forall\ V \in N(0)$ in E_1, $*\ U \in N(0)$ in E, $\#\ X \times Y \in N((x_0,f(x_0)))$, $\$\ \lambda > 0$, $\exists\ W \in K(X \times Y)$, $\exists\ Z \in M(U)$, $\#'\ (x,r) \in W \cap \mathrm{epi}_S f$, $\$'\ t \in (0,\lambda)$, $*'\ u' \in Z$, $\exists\ v' \in V$, $(x,r) + t(y+u', d+v') \in \mathrm{epi}_S f\}$

Remark 3.1.3: (a) The expression "$(x,r) + t(y+u',d+v') \in \mathrm{epi}_S f$" in Definition 3.1.2 can also be written "$\frac{f(x+t(y+u'))-r}{t} \leq_S d + v'$". As in section 1.5, it follows that if $(y,d) \in \hat{R}_S(f,x_0)$, then $(y,d') \in R_S(f,x_0)$ whenever $d \leq_S d'$.

(b) If $* := \forall$ and $M(Y) := Y$ in Definition 3.1.2, then $\hat{R}_S(f,x_0) = R(\mathrm{epi}_S f, (x_0, f(x_0)))$. For such tangent cones, we will denote $R(\mathrm{epi}_S f, (x_0, f(x_0)))$ simply as $R_f(x_0)$ - e.g., if R is the Clarke tangent cone, we will use "$T_f(x_0)$" in place of "$T_{\mathrm{epi}_S f}((x_0, f(x_0)))$".

Definition 3.1.4: Let R be a q-cone, and let $f: E \to (\overline{E}_1, S)$ and $x_0 \in f^{-1}(E_1)$. The R-S directional derivative of f at x_0 with respect to $y \in E$ is defined by

$$(3.1.5) \qquad f_S^R(x_0;y) := \inf\{d \mid (y,d) \in \hat{R}_S(f,x_0)\}.$$

The R-S subgradient of f at x_0 is the set

$$(3.1.6)\qquad \partial_S^R f(x_0) := \{T\in L(E,E_1) \mid T(y) \le_S f_S^R(x_0;y) \text{ for all } y\in E\}.$$

Remark 3.1.5: (a) As in the case of extended real valued functions, if R' and R'' are two q-cones such that $\hat{R}'_S(f,x_0) \subset \hat{R}''_S(f,x_0)$, it follows that $f_S^{R''}(x_0;\cdot) \le_S f_S^{R'}(x_0;\cdot)$ and $\partial_S^{R''}f(x_0) \subset \partial_S^{R'} f(x_0)$.

(b) Thibault introduces special cases of these definitions in [Th1]. He denotes $f_S^I(x_0;\cdot)$ by "$f^{\square}(x_0;\cdot)$", $f_S^T(x_0;\cdot)$ by "$f^{\dagger}(x_0;\cdot)$", and $\hat{I}(f,x_0)$ by "$Q(f,x_0)$". We will adopt the latter two notations and suppress the S in "$f_S^R(x_0;\cdot)$", writing it as "$f^R(x_0;\cdot)$".

We saw in section 1.5 that $\hat{R}(f,x_0) = \text{epi } f^R(x_0;\cdot)$ for $f\colon E \to \overline{\mathbb{R}}$. The analogous relationship will not generally hold in this vector-valued setting without additional assumptions about both R and S. The space (E_1,S) (alternately, the cone S) is said to be Daniell if for every decreasing net $(y_j)_{j\in J}$ in E_1 which is bounded below, $\lim_J y_j = \inf_J y_j$. If in addition the function $(y_1,y_2) \mapsto \sup(y_1,y_2)$ is continuous, (E_1,S) is called a Daniell topological lattice (D.t.l.).

Proposition 3.1.6: Let R be a q-cone, $f\colon E \to (\overline{E}_1,S)$, and $x_0 \in E$ such that $f(x_0) \in E_1$. Then

$$(3.1.7)\qquad \hat{R}_S(f,x_0) \subset \text{epi}_S f^R(x_0;\cdot).$$

If in addition (E_1,S) is a D.t.l. and R is such that $* := \forall$, $\# := \forall$, and $* := \forall$, then

$$(3.1.8) \qquad \hat{R}_S(f,x_0) = \mathrm{epi}_S f^R(x_0;\cdot) \ .$$

<u>Proof:</u> Let $(y,d) \in \hat{R}_S(f,x_0)$. By (3.1.5), $f_S^R(x_0;y) \leq_S d$; i.e., $(y,d) \in \mathrm{epi}_S f^R(x_0;\cdot)$. Conversely, suppose (E_1,S) is a D.t.l., and let $(y,d) \in \mathrm{epi}_S f^R(x_0;\cdot)$. If $f^R(x_0;\cdot) = -\infty$, then $(y,d) \in \hat{R}_S(f,x_0)$ by Remark 3.1.3(a). Otherwise, define $D := \{d \in E_1 \mid (y,d) \in \hat{R}_S(f,x_0)\}$, and suppose $d_1,d_2 \in D$. Let $V \in N(0)$ in E_1 . Since $(y_1,y_2) \mapsto \sup(y_1,y_2)$ is continuous, there exists $V_1 \subset V$ with $V_1 \in N(0)$ and $\sup(v_1,v_2) \in V$ whenever $v_1,v_2 \in V_1$. There exist $U \in N(0)$ in E , $X \times Y \in N((x_0,f(x_0)))$, $\lambda > 0$, $W \in K(X\times Y)$, and $Z \in M(0)$ such that for all $(x,r) \in W \cap \mathrm{epi}_S f$, $t \in (0,\lambda)$, $u' \in Z$, there exist $v_1',v_2' \in V_1$ with $\frac{f(x+t(y+u'))-r}{t} \leq d_i + v_i'$ for $i = 1,2$. It follows that

$$\frac{f(x+t(y+u'))-r}{t} \leq d_i + \sup(v_1',v_2') \quad \text{for} \quad i = 1,2 ,$$

and so

$$\frac{f(x+t(y+u'))-r}{t} \leq \inf(d_1,d_2) + \sup(v_1',v_2') \ .$$

Since $\sup(v_1',v_2') \in V$, we conclude that $\inf(d_1,d_2) \in D$. The rest of the proof is outlined in [Th1, Remark (4), page 330]. Consider the decreasing net of infima of finite subsets of D. The set D is easily seen to be closed by an argument similar to that of Theorem 1.3.4. Since (E_1, S) is Daniell, it follows then that $d \in D$. Thus $(y,d) \in \hat{R}_S(f,x_0)$ and (3.1.8) holds. □

<u>Definition 3.1.7:</u> The function $f\colon E \to (\bar{E}_1, S)$ is said to be <u>R epigraph regular at x_0</u> if (3.1.8) is satisfied.

Proposition 3.1.6 shows that f is R epigraph regular at each $x_0 \in f^{-1}(E_1)$ if (E_1,S) is a D.t.l. and $R = I,H,A,E,$ or L. The reason we isolate this property and assign it a name is that in many of the subdifferential calculus theorems in this chapter, it is important that the functions involved be T epigraph regular. At this point we pause to list some classes of functions that have this property.

(i) If $f\colon E \to (\bar{E}_1,S)$ is strictly compactly Lipschitzian at x_0 (see [Th1], [Th2], or [Th3] for a definition), then $T_f(x_0) = Q(f,x_0)$ ([Th2, Corollary 2.12], [Th1, Proposition 3.10]). Thus if (E_1,S) is a D.t.l., f is T epigraph regular at x_0. Thibault proves subdifferential calculus rules for such functions in [Th3].

(ii) If $f\colon E \to (\dot{E}_1,S)$ is S-convex and $x_0 \in \operatorname{int} \operatorname{dom} f$, then [Th1, Proposition 2.7] and [Bo10, Theorem 1.1] combine to show that $f^{\dagger}(x_0;\cdot) = f'(x_0;\cdot)$. Hence if (E_1,S) is a D.t.l., Proposition 3.1.6 with $R := A$ shows that f is T epigraph regular.

(iii) If $f: E \to (\overline{E}_1, S)$ is strictly differentiable at x_0, then $T_f(x_0) = Q(f,x_0) = \nabla f(x_0) + S$ and $f^{\dagger}(x_0;\cdot) = f^I(x_0;\cdot) = \nabla f(x_0)$ (see [Th1], [Th2, Corollaries 1.7 and 2.11]). Thus f is T epigraph regular and I epigraph regular regardless of whether (E_1,S) is a D.t.l.

(iv) We say $F: (E_1,K) \to (\overline{E}_2,S)$ is $(K-S)$ isotone on $D \subset E_1$ if $F(x) \leq_S F(y)$ whenever $x,y \in D$ and $x \leq_K y$. An examination of the proof of Proposition 3.1.6 shows that if $(y_1,y_2) \mapsto \inf(y_1,y_2)$ is continuous in (E_1,K), (E_2,S) is a D.t.l., and F is isotone on a neighbourhood of x_0, then F is R epigraph regular at x_0 for $R = T$, k as well as the other cones mentioned above.

(v) Products of R epigraph regular functions.

<u>Lemma 3.1.8</u>: Let $f_i: E \to (\overline{E}_i,K_i)$, $i = 1,\ldots,n$, and define $f: E \to \pi(\overline{E}_i,K_i)$ by $f(x) := (f_1(x),\ldots,f_n(x))$. Suppose that each f_i is R epigraph regular at $x_0 \in E$, and that $(y,d_1,\ldots,d_n) \in \hat{R}(f,x_0)$ if and only if $(y,d_i) \in \hat{R}(f_i,x_0)$, $i = 1,\ldots,n$. Then $f^R(x_0;y) = (f_1^R(x_0;y),\ldots,f_n^R(x_0;y))$ and f is R epigraph regular at x_0.

<u>Proof</u>: The first assertion follows immediately from (3.1.5). To prove that f is R epigraph regular at x_0, suppose $f^R(x_0;y) \leq (d_1,\ldots,d_n)$. Then $f_i^R(x_0;y) \leq d_i$, and since each f_i is R epigraph regular, $(y,d_i) \in \hat{R}(f_i,x_0)$.

By hypothesis, $(y, d_1, \ldots, d_n) \in \hat{R}(f, x_0)$. Therefore (3.1.8) holds and f is R epigraph regular at x_0. □

We will apply Lemma 3.1.8 several times in sections 3.3 and 3.4.

It is hoped that further research will lengthen the above list. Of course any $f: E \to \overline{\mathbb{R}}$ is R epigraph regular for any R and any x_0 for which f is finite (section 1.5).

Observe that if $f: E \to (\overline{E}_1, S)$ is T epigraph regular at x_0, then $\mathrm{epi}_S f^{\dagger}(x_0; \cdot)$ is closed, and we have by Proposition 3.1.1 that either

(3.1.9) (a) $f^{\dagger}(x_0; 0) = 0$, in which case $f^{\dagger}(x_0; \cdot)$ is proper, or

(b) $f^{\dagger}(x_0; 0) = -\infty$, in which case $f^{\dagger}(x_0; \cdot)$ is equal to $-\infty$ throughout dom f and $\partial_S^T f(x_0) = \phi$,

just as in the scalar-valued case. We will make occasional use of this fact later. We will also find use for a vector-valued version of Lemma 2.2.10, which we prove below.

Lemma 3.1.9: Let $F: (E_1, K) \to (\overline{E}_2, S)$ be K-S isotone on a neighbourhood of $x_0 \in F^{-1}(E_2)$ in E_1. Let R be a q-cone with [illegible] $= \Xi$ which satisfies the following conditions:

$W := (W_1 \cap W_3) \times W_2 \in K((X \cap X_2) \times Y)$, there exists $Z \in M(U)$ such that for all $(x,r) \in W \cap \mathrm{epi}_S F$, $\forall$ $t \in (0, \min(\lambda, \lambda_0))$, $\forall$ $h \in Z$,

$$\frac{F(x+t(y_2+h))-r}{t} \leq_S d + v$$

for some $v \in V$. For any such x , t , and h , $x + t(y_1+h) \leq_K x + t(y_2+h)$, and both $x + t(y_2+h)$ and $x + t(y_1+h)$ are in X_1 . By isotonicity of F on X_1 ,

$$\frac{F(x+t(y_1+h))-r}{t} \leq_S \frac{F(x+t(y_2+h))-r}{t}$$

and hence $(y_1,d) \in \hat{R}_S(F,x_0)$, as required. □

Among the q-cones which satisfy the hypotheses of Lemma 3.1.8 are $K_C(x_0)$, $k_C(x_0)$, $T_C(x_0)$, and $I_C(x_0)$.

As in [Ro3] and [Th1], we will be primarily concerned here with $f^{\uparrow}(x_0;\cdot)$ and $f^I(x_0;\cdot)$. The strategy of the proofs in this chapter will be similar to that pursued in [Th1] - directional derivative inequalities involving $f^{\uparrow}(x_0;\cdot)$ and $f^I(x_0;\cdot)$ combine with subgradient inclusions for S-convex functions to produce subgradient inclusions involving $\partial_S^T f$ and $\partial_S^I f$. This strategy is again based on that fact that $f^{\uparrow}(x_0;\cdot)$ and $f^I(x_0;\cdot)$ are S-convex functions and $\partial_S^T f(x_0) = \partial_S f^{\uparrow}(x_0;\cdot)(0)$, $\partial_S^I f(x_0) = \partial_S f^I(x_0;\cdot)(0)$.

(3.1.10) If $X_1, X_2 \in N(x_0)$, then
$K(X_1) \cap K(X_2) \subset K(X_1 \cap X_2)$

(3.1.11) If $X \in N(x_0)$, $Y \in N(F(x_0))$, then
$K(X \times Y) = K(X) \times K(Y)$.

Then $F^R(x_0;\cdot)$ is K-S isotone on all of E_1.

Proof: Let $(y_2,d) \in \hat{R}_S(F,x_0)$, and let $y_1 \in E_1$ be such that $y_1 \leq_K y_2$. It suffices to show that (y_1,d) is also in $\hat{R}_S(F,x_0)$. To this end, let $V \in N(0)$ in E_2 be given. There exists $X_1 \in N(x_0)$ on which F is isotone. For a given $U \in N(0)$ in E_1, there exists $X_2 \in N(x_0)$ and $\lambda_0 > 0$ such that $X_2 + [0,\lambda_0)(y_i+U) \subset X_1$ for $i = 1,2$. Since $(y_2,d) \in \hat{R}_S(F,x_0)$, $\forall\, U \in N(0)$, there exists $X \times Y \in N((x_0,F(x_0)))$, $\exists\, \lambda > 0$, there exist $W \in K(X \times Y)$ and $Z \in M(U)$ such that for all $(x,r) \in W \cap \mathrm{epi}_S F$, $\exists'\, t \in (0,\lambda)$, $\forall'\, h \in Z$,

$$\frac{F(x+t(y_2+h))-r}{t} \leq_S d + v$$

for some $v \in V$. By (3.1.11), any $W \in K(X \times Y)$ is equal to $W_1 \times W_2$ with $W_1 \in K(X)$ and $W_2 \in K(Y)$. By (3.1.10), if $W_1 \in K(X)$ and $W_3 \in K(X_2)$, $W_1 \cap W_3 \in K(X \cap X_2)$. Putting these facts together, we see that $\forall\, U \in N(0)$, there exists $(X \cap X_2) \times Y \in N((x_0,F(x_0)))$, $\exists\, \lambda > 0$; there exists $W_1 \cap W_3 \in K(X \cap X_2)$ and $W_2 \in K(Y)$ with

In chapter 2, $f_+(x_0;\cdot)$ and $f_{\square}(x_0;\cdot)$ play an important subsidiary role in formulating conditions for equality. This aspect of the theory does not seem to carry over in general to the vector-valued case, since the proofs of Propositions 2.4.1 and 2.4.2 depend heavily on the fact that the usual ordering on $\mathbb{R}$ is <u>total</u>; i.e., $\mathbb{R} = \mathbb{R}_+ \cup -\mathbb{R}_+$.

We close this introductory section with a note on notation: Since $\partial_S^T f = \partial_S f$ if f is S-convex [Th1, Proposition 2.7] and $\partial_S^T f = \partial_S^I f$ if $\partial_S^I f$ is nonempty ([Th2, Corollary 2.4], [Th1, Lemma 3.7]), we will hereafter write $\partial_S^T f$ and $\partial_S^I f$ simply as $\partial_S f$.

3.2. An important tangent cone inclusion

In this section, we show that an inclusion involving the Clarke tangent cone to the graph of the composition of two relations [Th2, Proposition 4.4] can be employed quite effectively in the derivation of subdifferential calculus formulae. We begin by introducing some necessary terminology involving relations.

<u>Definition 3.2.1</u>: Let E_1, E_2 be l.c.s., and let $M: E_1 \to 2^{E_2}$ be a relation (alternately "set-valued mapping" or "multifunction").

(a) The graph of M is the set

$$\mathrm{Gr}\, M := \{(x,y) \in E_1 \times E_2 \mid y \in M(x)\}.$$

(b) The <u>inverse</u> of M is the relation $M^{-1}: E_2 \to 2^{E_1}$ defined by $M^{-1}(y) := \{x \in E_1 \mid y \in M(x)\}$.

(c) The relation M is said to be lower semicontinuous (l.s.c.) at $(x_0,y_0) \in \mathrm{Gr}\, M$ relative to a set $D \subset E$ containing x_0 if for each $Y \in N(y_0)$, there exists $X \in N(x_0)$ such that $Y \cap M(x) \neq \phi$ for all $x \in D \cap X$.

(d) Let E be a l.c.s., and let $N: E \to 2^{E_1}$ be a relation. The <u>composition</u> of M and N is the relation $M \circ N: E \to 2^{E_2}$ defined by $(M \circ N)(x) := \{z \in E_2 \mid \exists\, y \in N(x) \cap M^{-1}(z)\}$.

<u>Definition 3.2.2</u> [Th2]: Let E, E_1 be l.c.s., and let $M: E \to 2^{E_1}$ be a relation. The <u>quasi-interiorly tangent cone</u> to Gr M at (x_0,y_0) is the set

$$(3.2.1) \qquad \begin{aligned} Q_M(x_0,y_0) := \{(u,v) \in E \times E_1 \mid\ & \forall\, V \in N(v),\ \exists\, U \in N(u), \\ & \exists\, X \in N(x_0),\ \exists\, Y \in N(y_0),\ \exists\, \lambda > 0 \text{ such that} \\ & \forall\, (x',y') \in (X \times Y) \cap \mathrm{Gr}\, M,\ \forall\, t \in (0,\lambda),\ \forall\, u' \in U, \\ & \exists\, v' \in V \text{ with } (x',y') + t(u',v') \in \mathrm{Gr}\, M\} . \end{aligned}$$

<u>Remark 3.2.3</u>: (a) We will be especially concerned here with the <u>S-epigraph relation</u>

$$\mathrm{epi}_S f(x) := \{y \in E_1 \mid f(x) \leq_S y\}$$

corresponding to a given $f: E \to (\overline{E}_1, S)$. Observe that when $M := \mathrm{epi}_S f$, $Q(f,x_0) := Q_M(x_0, f(x_0)) = \hat{I}(f,x_0)$ for all $x_0 \in f^{-1}(E_1)$.

(b) It is shown in [Th1] that $Q(f,x_0)$ is convex under certain assumptions on f. This is extended in [Th2, Corollary 2.3], where it is shown that in fact $Q_M(\cdot)$ is <u>always</u> convex. The proof is similar to that of Theorem 1.3.21 or Theorem 1.3.23.

(c) In this section, we will use "$T_M(x_0,y_0)$" as a shorthand notation for "$T_{grM}((x_0,y_0))$", following [Th2].

(d) When $Q_M(x_0,y_0)$ is nonempty, $T_M(x_0,y_0) = \mathrm{cl}\, Q_M(x_0,y_0)$ [Th2, Corollary 2.4]. This means in particular that for $f\colon E \to (\bar{E}_1, S)$, $\partial_S^T f(x_0) := \partial_S^I f(x_0)$ whenever the latter is nonempty, as we mentioned at the end of section 3.1.

Having made these definitions, we can now state the aforementioned tangent cone inclusion.

<u>Theorem 3.2.4</u> [Th2, Propositions 4.2 and 4.4]: Let E, E_1, E_2 be l.c.s., let $M\colon E_1 \to 2^{E_2}$ and $N\colon E \to 2^{E_1}$ be relations, and let $y_0 \in N(x_0)$ and $z_0 \in M(y_0)$. If the relation from $E \times E_2$ into E_1 defined by $(x,z) \longrightarrow N(x) \cap M^{-1}(z)$ is l.s.c. at $((x_0,z_0),y_0)$ relative to $\mathrm{Gr}(M \circ N)$, then for all $u \in E$,

$$(3.2.2) \qquad [Q_M(y_0,z_0) \circ T_N(x_0,y_0)](u) \subset T_{M\circ N}(x_0,z_0)(u)$$

and

$$(3.2.3) \qquad [Q_M(y_0,z_0) \circ Q_N(x_0,y_0)](u) \subset Q_{M\circ N}(x_0,z_0)(u) .$$

We will apply both (3.2.2) and (3.2.3) in proving subdifferential calculus rules. Thibault applies Theorem 3.2.4 in the case in which N is $\mathrm{Gr}\, F$ for a function $F\colon E \to E_1$ which is strictly differentiable at $x_0 \in E$. If we in addition set $M := \mathrm{epi}\, f$ for some $f\colon E_1 \to \overline{\mathbb{R}}$, (3.2.2) can be applied to prove [Ro3, Theorem 3]. In [Th1], Thibault proves a vector-valued generalization of [Ro3, Theorem 3], although not by this method.

If we let $M := \mathrm{epi}_S F$ and $N := \mathrm{epi}_K f$ be epigraph relations for $f\colon E \to (\overline{E}_1,K)$ and $F\colon (E_1,K) \to (\overline{E}_2,S)$, Theorem 3.2.4 yields directional derivative chain rules for $(F\circ f)^{\dagger}(x_0;\cdot)$ and $(F\circ f)^{I}(x_0;\cdot)$ <u>if</u> $\mathrm{epi}_S(F\circ f) = \mathrm{epi}_S F \circ \mathrm{epi}_K f$. This condition is satisfied in the important special case in which F is K-S isotone, as we now demonstrate.

<u>Lemma 3.2.5</u>: Let $f\colon E \to (\overline{E}_1,K)$ and $F\colon (E_1,K) \to (\overline{E}_2,S)$ with $F(+\infty) = +\infty$. Assume either f is proper or $F(-\infty) = -\infty$ and $\inf_{S,E_1} F(y) = -\infty$. Then

$$(3.2.4) \qquad \mathrm{epi}_S(F\circ f) \subset \mathrm{Gr}(\mathrm{epi}_S F \circ \mathrm{epi}_K f) .$$

If in addition F is K-S isotone on the set $D := \{y \in E_1 \mid f(x) \leq_K y \text{ for some } x \in E\}$, then

$$(3.2.5) \qquad \mathrm{epi}_S(F\circ f) = \mathrm{Gr}(\mathrm{epi}_S F \circ \mathrm{epi}_K f) .$$

<u>Proof</u>: Let $(x,z) \in \mathrm{epi}_S(F\circ f)$. Since $F(+\infty) = +\infty$, $f(x) \neq +\infty$. If $f(x) \in E_1$, then $(f(x),z) \in \mathrm{epi}_S F$, and $(x,f(x)) \in \mathrm{epi}_K f$ since $0 \in K$. Hence (3.2.4) holds. If $f(x) = -\infty$, let $y \in E_1$ be such that $F(y) \leq_S z$. Since $f(x) \leq_K y$, we again get (3.2.4). Conversely, suppose F is isotone on D , and let $(x,z) \in \mathrm{Gr}(\mathrm{epi}_S F \circ \mathrm{epi}_K f)$. Then there exists $y \in E_1$ such that $f(x) \leq_K y$ and $F(y) \leq_S z$. By the isotonicity of F , $F(f(x)) \leq_S F(y) \leq_S z$, and so $(x,z) \in \mathrm{epi}_S(F\circ f)$. Therefore (3.2.5) holds.

We can now prove our main directional derivative chain rule.

<u>Theorem 3.2.6</u>: Let $f: E \to (\overline{E}_1,K)$ and $F: (E_1,K) \to (\overline{E}_2,S)$ be such that $F(+\infty) = +\infty$ and either

(i) f is proper, or

(ii) $F(-\infty) = -\infty$ and $\inf\limits_{E_1}{}_S\, F(y) = -\infty$.

Let $x_0 \in f^{-1}(E_1)$ with $F(f(x_0)) \in E_2$, and assume F is K-S isotone on the union of some neighbourhood of $f(x_0)$ and the set $\{y \in E_1 \mid f(x) \leq_K y \text{ for some } x \in E\}$. Suppose $F^I(f(x_0);+\infty) = +\infty$ and that either

(iii) $f^{\dagger}(x_0;\cdot)$ is proper, or

(iv) $F^I(f(x_0);-\infty) = -\infty$ and $\inf\limits_{E_1}{}_S\, F^I(f(x_0);y) = -\infty$.

Make the following additional assumption:

(3.2.6) For each $Y \in N(f(x_0))$, there exist $X \in N(x_0)$ and $Z \in N(F(f(x_0)))$ such that for all $(x,z) \in X \times Z$ with $(F \circ f)(x) \leq_S z$, there exists $y \in Y$ satisfying $f(x) \leq_K y$ and $F(y) \leq_S z$.

If f is T epigraph regular at x_0, then for all $y \in E$,

$$(3.2.7) \quad (F \circ f)^{\dagger}(x_0;y) \leq_S F^{I}(f(x_0);\, f^{\dagger}(x_0;y)) .$$

If f is I epigraph regular at x_0, then for all $y \in E$,

$$(3.2.8) \quad (F \circ f)^{I}(x_0;y) \leq_S F^{I}(f(x_0);\, f^{I}(x_0;y)) .$$

<u>Proof:</u> Define $N\colon E \to 2^{E_1}$ and $M\colon E_1 \to 2^{E_2}$ by $N(x) := \mathrm{epi}_K f(x) := \{y \in E_1 \mid f(x) \leq_K y\}$ and $M(y) := \mathrm{epi}_S F(y)$. By Lemma 3.2.5, $\mathrm{Gr}(M \circ N) = \mathrm{epi}_S(F \circ f)$. Now $F^{I}(f(x_0);\cdot)$ is isotone by Lemma 3.1.9, so

$$\begin{aligned}
& F^{I}(f(x_0); f^{\dagger}(x_0;y)) \\
&= \inf_S\{F^{I}(f(x_0);d) \mid f^{\dagger}(x_0;y) \leq_K d\} \\
&= \inf_S\{F^{I}(f(x_0);d) \mid (y,d) \in T_f(x_0)\} \quad \text{if } f \text{ is } T \text{ epigraph} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{regular at } x_0 , \\
&= \inf_S\{r \mid (d,r) \in Q(F,f(x_0)),\, (y,d) \in T_f(x_0)\} \\
&= \inf_S\{r \mid r \in [Q_M(f(x_0),F(f(x_0))) \circ T_N(x_0,f(x_0))](y)\} .
\end{aligned}$$

By definition,

$$(F\circ f)^{\dagger}(x_0;y) = \inf_S\{r \mid r \in T_{M\circ N}(x_0,F(f(x_0)))(y)\} .$$

Thus to prove (3.2.7), it suffices to show that

$$[Q_M(f(x_0);F(f(x_0))) \circ T_N(x_0,f(x_0))](y)$$
$$\subset T_{M\circ N}(x_0,F(f(x_0)))(y) .$$

Observe that assumption (3.2.6) is simply the statement that the relation

$$(x,z) \mapsto \mathrm{epi}_K f(x) \cap (\mathrm{epi}_S F)^{-1}(z)$$

is l.s.c. at $((x_0,F(f(x_0))),f(x_0))$ relative to $\mathrm{epi}_S F\circ f$. Thus we may apply Theorem 3.2.4, and (3.2.2) gives the desired inclusion. The proof of (3.2.8) parallels that of (3.2.7).

<u>Corollary 3.2.7</u> (cf. [Thl, Proposition 4.4(i)]): Let $f: E \to E_1$ be strictly differentiable at $x_0 \in E$, and let $F: E_1 \to (\overline{E}_2,S)$ be such that $F(f(x_0)) \in E_2$. Then for all $y \in E$,

$$(3.2.9) \qquad (F\circ f)^{I}(x_0;y) \leq F^{I}(f(x_0);\nabla f(x_0)y) .$$

<u>Proof:</u> In Theorem 3.2.6, let $K = \{0\}$. As we saw in section 3.1, f is I epigraph regular at x_0 and $f^{I}(x_0;y) = \nabla f(x_0)y$. With $K = \{0\}$, F is automatically

K-S isotone since $0 \in S$. Since f is continuous at x_0, (3.2.6) is automatically satisfied by Proposition 3.2.8 below. Hence (3.2.8) reduces to (3.2.9) in this case. □

In chapter 2, we studied two seemingly distinct chain-rule formulations. Corollary 3.2.7 shows that Theorem 3.2.6 encompasses both of them.

Theorem 3.2.6 promises to have many more applications as long as we can find easily verifiable conditions guaranteeing that (3.2.6) holds. We now give two such conditions, beginning with the one just applied in Corollary 3.2.7.

<u>Proposition 3.2.8</u>: Suppose $f: E \to (\bar{E}_1, K)$ in Theorem 3.2.6 is continuous at x_0. Then (3.2.6) holds.

<u>Proof</u>: Let $Y \in N(f(x_0))$. If f is continuous at x_0, there exists $X \in N(x_0)$ such that $f(X) \subset Y$. Let Z be any neighbourhood of $F(f(x_0))$, and suppose $(x,z) \in X \times Z \cap \mathrm{epi}_S(F \circ f)$. Then $f(x) \in Y$, so $y := f(x)$ satisfies $f(x) \leq_K y$, $F(y) \leq_S z$, and $y \in Y$. Therefore (3.2.6) holds. □

<u>Definition 3.2.9</u>: Suppose (E_1, K) is an o.t.v.s. The cone K is said to be <u>normal</u> if there exists a base of neighbourhoods of the origin in E_1 such that $V = (V+K) \cap (V-K)$ for each V in the neighbourhood base. Such neighbourhoods are called "normal" or "full".

<u>Definition 3.2.10</u>: A function $f\colon E \to (\overline{E}_1, K)$ is <u>lower semicontinuous</u> (l.s.c.) at $x_0 \in f^{-1}(E_1)$ if for each $Y \in N(f(x_0))$, there exists $X \in N(x_0)$ such that $f(X) \subset Y + K$. It is <u>upper semicontinuous</u> (u.s.c.) at x_0 if for each $Y \in N(f(x_0))$, there exists $X \in N(x_0)$ such that $f(X) \subset Y - K$.

<u>Remark 3.2.11</u>: (a) Notice that Definition 3.2.10 reduces to Definition 2.2.1 (c) in the case where $E_1 := \mathbb{R}$.

(b) The cone K is normal if and only if for any $y \in E_1$, there exists a base of neighbourhoods $Y \in N(y)$ such that

$$Y = (Y+K) \cap (Y-K)$$

for each such Y. Notice that if K is normal, the function $f\colon E \to (\overline{E}_1, K)$ is continuous at x_0 if and only if it is l.s.c. and u.s.c. at x_0.

(c) We will often assume normality in the sequel. This assumption provides an important link between the topology and order structure of (E_1, K). There are a number of o.t.v.s.'s which have the combination of properties usually needed to apply Theorem 3.2.6 - D.t.l. plus normality - including many Banach lattices (see [Perl], [Th3], [Bo4], [Bo10] for further details).

<u>Proposition 3.2.12</u>: In Theorem 3.2.6, assume that K is normal and $f\colon E \to (\overline{E}_1, K)$ is l.s.c. at x_0. Also assume that the following condition holds:

(3.2.10) For each $Y \in N(f(x_0))$, there exist $Y_1 \in N(f(x_0))$ and $Z \in N(F(f(x_0)))$ such that $F^{-1}(Z-S) \cap (Y_1+K) \subset Y-K$. Then (3.2.6) holds.

<u>Proof:</u> Let $Y \in N(f(x_0))$ be a full neighbourhood, and choose $Z \in N(F(f(x_0)))$ and $Y_1 \in N(f(x_0)))$ as postulated in (3.2.10). Since f is l.s.c. at x_0, there exists $X \in N(x_0)$ such that $f(X) \subset (Y \cap Y_1) + K$. Now suppose $(x,z) \in X \times Z$ satisfies $F(f(x)) \leq_S z$. Then $f(x) \in Y_1 + K$ and $F(f(x)) \in Z-S$, and so $f(x) \in Y-K$ by (3.2.10). Hence $f(x) \in (Y+K) \cap (Y-K) = Y$, and (3.2.6) is satisfied with $y := f(x)$. □

In the case in which $E := \mathbb{R}^m$, $E_1 := \mathbb{R}^n$, $E_2 := \mathbb{R}$, and $K := \mathbb{R}^n_+$, the assumption that F is l.s.c. and strictly isotone in the ith coordinate at $f(x_0) = (f_1(x_0),\ldots,f_n(x_0))$ and f_i is l.s.c. at x_0 for each i will guarantee that (3.2.10) is satisfied, as an examination of the proof of Lemma 2.2.8 reveals. Moreover, it is also clear from the proof of Lemma 2.2.8 that assumptions (1) and (2) of Lemma 2.2.8 will ensure that (3.2.6) holds.

Using Propositions 3.2.8 and 3.2.12, we can derive the following chain rule for extended real-valued functions from Theorem 3.2.6:

<u>Theorem 3.2.13:</u> Let E be a l.c.s. and let $f: E \to \overline{\mathbb{R}}$ and $F: \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ be extended real-valued functions with $F(+\infty) = +\infty$.

Assume that either f is proper or that $F(-\infty) = -\infty$ and $\inf_{y\in\mathbb{R}} F(y) = -\infty$. Suppose $f(x_0)$ and $F(f(x_0))$ are finite. Assume that either $f^{\uparrow}(x_0;0) = 0$ or that $F^{\uparrow}(f(x_0);-\infty) = -\infty$ and $\inf_{y\in\mathbb{R}} F^{\uparrow}(f(x_0);y) = -\infty$. Suppose F is isotone on $\{y\,|\,f(x) \le y$ for some $x \in E\} \cup B_{\varepsilon}(f(x_0))$ for some $\varepsilon > 0$.

Assume either

(1) f is continuous at x_0, or

(2) f is l.s.c. at x_0 and F is strictly isotone at $f(x_0)$.

Assume further that

$$(3.2.11) \qquad \text{Range } f^{\uparrow}(x_0;\cdot) \cap \text{int dom } F^{\uparrow}(f(x_0);\cdot) \neq \phi .$$

Then for all $y \in E$,

$$(3.2.12) \qquad (F\circ f)^{\uparrow}(x_0;y) \le F^{\uparrow}(f(x_0);f^{\uparrow}(x_0;y)) .$$

Moreover

$$(3.2.13) \qquad \partial(F\circ f)(x_0) \subset \{\lambda\partial f(x_0)\,|\,\lambda\in\partial F(f(x_0)),\lambda\ge 0\},$$

where by convention $\lambda\partial f(x_0) = \phi$ if $\partial f(x_0) = \phi$. Equality holds in (3.2.13) if in addition F is subdifferentiably regular at $f(x_0)$ and f is subdifferentiably weakly regular at x_0. Equality holds in (3.2.12) if these regularity conditions hold and $f^{\uparrow}(x_0;0) = 0$.

Proof: By an argument similar to that in Theorem 2.3.15, (3.2.13) follows from (3.2.12) by Theorem 2.2.12. To establish (3.2.12) we need to verify that (3.2.6) holds under the hypotheses above. If (1) holds, (3.2.6) is satisfied by Proposition 3.2.8. Suppose (2) holds. We will verify that (3.2.10) is satisfied. To do so, let $\delta_0 > 0$ be given, and call $\delta := \min(\varepsilon, \delta_0)$ and $Y := B_\delta(f(x_0))$. Set $\lambda := F(f(x_0)+\delta) - F(f(x_0))$. Since F is strictly isotone at $f(x_0)$, $\lambda > 0$. Now set $Z := B_{\lambda/2}(F(f(x_0)))$. If $F(y) \le F(f(x_0)) + \frac{\lambda}{2} < F(f(x_0) + \delta)$, then $y \le f(x_0) + \delta$ by the isotonicity of F. Hence (3.2.10) holds for any $Y_1 \in N(f(x_0))$, Z as above, and $K = S = \mathbb{R}_+$. By Proposition 3.2.12, for all $y \in \mathbb{R}$,

$$(F\circ f)^\uparrow(x_0;y) \le F^I(f(x_0);f^\uparrow(x_0;y)) .$$

But since (3.2.11) holds, we conclude that in fact $(F\circ f)^\uparrow(x_0;y) \le F^\uparrow(f(x_0);f^\uparrow(x_0;y))$ for all $y \in \mathbb{R}$ by an argument similar to that used in [Ro3, Theorem 2]. The assertions about equality follow as in Proposition 2.4.4. □

The proof that (3.2.6) holds under hypotheses (1) and (2) of Theorem 3.2.13 is a much simplified version of the argument used to verify Lemma 2.2.8. Notice that the function F in Theorem 3.2.13 does not have to be l.s.c. at $f(x_0)$, in contrast to the case of $m > 1$ in Lemma 2.2.8.

If we can derive a vector-valued version of Theorem 2.2.12, we will be able to use Theorem 3.2.6 to establish subgradient inclusions for nonconvex vector-valued functions. In deriving such a theorem, we will make use of the following Lagrange multiplier theorem which is a special case of [Bo9, Theorem 3.1]:

<u>Theorem 3.2.14</u>: Let $f: E_1 \times E_2 \to (\dot{E}_3, S)$ be S-convex and let $g: E_1 \times E_2 \to (\dot{E}_4, K)$ be K-convex. Consider the convex program

$$p(0) := \inf_S \{f(x_1,x_2) \mid g(x_1,x_2) \leq_K 0\} .$$

Assume that there exists $(x_1,x_2) \in E_1 \times E_2$ such that $g(x_1,x_2) \lneq_K 0$ and such that

(3.2.14) $\quad f(\cdot,x_2)$ is continuous at x_1

and

(3.2.15) $\quad 0 \in \operatorname{int}(g(x+V,x_2)+K)$ for any $x \in E$ with $g(x,x_2) \leq_K 0$ and any $V \in N(0)$ in E.

Then there exists $\lambda \in L(E_4,E_3)$ such that $\lambda(K) \subset S$ and $f(x,y) + \lambda(g(x,y)) \geq p(0)$ for all $(x,y) \in E_1 \times E_2$.

We now give a vector-valued counterpart to Theorem 2.2.12.

<u>Theorem 3.2.15</u>: Let $f: E \to (\dot{E}_1, K)$ be K-convex, and let $F: (E_1,K) \to (\dot{E}_2,S)$ be S-convex and K-S isotone on E_1.

Suppose $x_0 \in \operatorname{dom} f \cap f^{-1}(\operatorname{dom} F)$. Assume there exists a point in the range of f at which F is continuous. Then

$$(3.2.16) \quad \partial_S(F\circ f)(x_0) = \{\partial_S(\lambda f)(x_0) \mid \lambda \in \partial_S F(f(x_0)), \lambda(K) \subset S\},$$

where by convention $\lambda f(x) = +\infty$ if $f(x) = +\infty$. If in addition K is Daniell, $E_2 := \mathbb{R}$, $S := \mathbb{R}_+$, and f is continuous at x_0, then

$$(3.2.17) \quad \partial_S(F\circ f)(x_0) = \{\lambda \partial_K f(x_0) \mid \lambda \in \partial_S F(f(x_0)), \lambda(K) \subset S\},$$

where by convention the right hand side of (3.2.17) is considered to be empty if $\partial_K f(x_0) = \phi$.

<u>Proof</u>: Let $T \in \partial_K f(x_0)$ and $\lambda \in \partial_S F(f(x_0))$ such that $\lambda(K) \subset S$. For all $x \in E$, $T(x-x_0) \le_K f(x)-f(x_0)$, so since $\lambda(K) \subset S$, $\lambda T(x-x_0) \le_S \lambda(f(x)-f(x_0))$. Since $\lambda \in \partial_S F(f(x_0))$,

$$\lambda(f(x)-f(x_0)) \le_S F(f(x))-F(f(x_0))$$

for all $x \in E$. Therefore $\lambda T \in \partial_S(F\circ f)(x_0)$, and

$$\begin{aligned}\{\lambda\partial_K f(x_0) \mid \lambda \in \partial_S F(f(x_0)), \lambda(K) \subset S\} \\ \subset \{\partial_S(\lambda f)(x_0) \mid \lambda \in \partial_S F(f(x_0)), \lambda(K) \subset S\} \\ \subset \partial_S(F\circ f)(x_0).\end{aligned}$$

Conversely, suppose $T \in \partial_S(F\circ f)(x_0)$. Since F is K-S isotone,

$$F(f(x_0))-T(x_0) = \inf_S\{F(y)-T(x) \mid f(x)-y \leq_K 0\} .$$

Since F is continuous at some element of the range of f, (3.2.14) and (3.2.15) hold for this convex program, and we may apply Theorem 3.2.14. There exists $\lambda \in L(E_1,E_2)$ with $\lambda(K) \subset S$ and

$$(3.2.18) \qquad F(y)-T(x) + \lambda(f(x)-y) \geq_S F(f(x_0))-T(x_0)$$

for all $x \in E$, $y \in E_1$. Setting $x := x_0$ in (3.2.18), we obtain

$$\lambda(y-f(x_0)) \leq_S F(y) - F(f(x_0))$$

for all $y \in E_1$, and so $\lambda \in \partial_S F(f(x_0))$. Setting $y := f(x_0)$ in (3.2.18), we get that

$$T(x-x_0) \leq_S \lambda(f(x)-f(x_0))$$

for all $x \in E$, and so $T \in \partial_S(\lambda f)(x_0)$. Thus (3.2.16) holds.

Now suppose K is Daniell, $E_2 := \mathbb{R}$, $S := \mathbb{R}_+$, and f is continuous at x_0. Then for $\lambda \in \partial_S F(f(x_0))$ with $\lambda(K) \subset S$, the set $\lambda\partial_K f(x_0)$ is compact in the weak operator topology [Bo4]. Suppose $\partial_S(\lambda f)(x_0) \neq \lambda\partial_K f(x_0)$. Since K is Daniell, there exists $h \in E$ such that

$$\lambda f'_K(x_0;h) = \sup_{T\in\lambda\partial_K f(x_0)} T(h)$$

$$\leq \sup_{T\in\partial_S(\lambda f)(x_0)} T(h) \leq (\lambda f)'_S(x_0;h)$$

contradicting the fact [Bo4, Proposition 3.8] that

$$\lambda f'_K(x_0;h) = (\lambda f)'_S(x_0;h) .$$

Therefore $\partial_S(\lambda f)(x_0) = \lambda\partial_K f(x_0)$ and (3.2.16) holds. □

By combining Theorems 3.2.6 and 3.2.15, we can prove the following chain rule for generalized subgradients:

Theorem 3.2.16: Under the hypotheses of Theorem 3.2.6 assume, in addition that if $f^{\uparrow}(x_0;0) = 0$, then $F^I(f(x_0);\cdot)$ is proper and continuous at some element of range $f^{\uparrow}(x_0;\cdot)$. Then

$$(3.2.19) \qquad \partial_S(F\circ f)(x_0) \subset \{\partial_S(\lambda f)(x_0) \mid \lambda \in \partial_S F(f(x_0)), \lambda(K) \subseteq S\} .$$

In the special case in which $E_2 := \mathbb{R}$, $S := \mathbb{R}_+$, and $f^{\uparrow}(x_0;\cdot)$ is continuous at 0, then

$$(3.2.20) \qquad \partial_S(F\circ f)(x_0) \subset \{\lambda\partial_K f(x_0) \mid \lambda \in \partial_S F(f(x_0)), \lambda(K) \geq 0\} .$$

Proof: If $f^{\uparrow}(x_0;0) = -\infty$, then by the hypotheses of Theorem 3.2.6 and (3.2.7), $(F\circ f)^{\uparrow}(x_0;0) = -\infty$, so that (3.2.19) and (3.2.20) hold vacuously. If $f^{\uparrow}(x_0;0) = 0$

and $F^I(f(x_0);\cdot)$ is proper and continuous at some element of range $f^\dagger(x_0;\cdot)$, we may combine (3.2.7) with (3.2.16) and (3.2.17) to obtain (3.2.19) and (3.2.20). □

<u>Corollary 3.2.17</u>: Let $f: E \to (E_1, K)$ be continuous and T epigraph regular at $x_0 \in E$. Assume that K is Daniell. Let $T \in L(E_1, \mathbb{R})$ be in K^+. Suppose $f^\dagger(x_0;0) = 0$ and $f^\dagger(x_0;\cdot)$ is continuous at 0. Then

$$\partial(Tf)(x_0) \subset T\partial_K f(x_0) . \tag{3.2.20}$$

<u>Proof</u>: Let $F := T$, $E_2 := \mathbb{R}$, $S := \mathbb{R}_+$ in Theorem 3.2.16. □

<u>Remark 3.2.18</u>: We mention here some conditions sufficient to guarantee that $f^\dagger(x_0;\cdot)$ is continuous at y. Suppose E and E_1 are complete metric spaces and $f: E \to (\overline{E}_1, S)$ is T epigraph regular at x_0. Then by [Bo9, Proposition 2.3], $f^\dagger(x_0;\cdot)$ is continuous at any point in the core of dom $f^\dagger(x_0;\cdot)$.

3.3. Formulae involving n functions

In order to derive vector-valued analogues of some of the results of section 2.3, we now will examine $f^\dagger(x_0;\cdot)$ in the case in which $f: E \to (\overline{E}_1, K)$ and

$$(E_1, K) = \prod_{i=1}^{n} (E_{1i}, K_i) . \tag{3.3.1}$$

Such a function f naturally breaks up into $(f_1,\ldots,f_n)$, with $f_i: E \to (\overline{E}_{1i},K_i)$. To derive directional derivative inequalities involving $f_i^{\dagger}(x_0;\cdot)$, we need more information about the relationship between $f^{\dagger}(x_0;\cdot)$ and $(f_1^{\dagger}(x_0;\cdot),\ldots,f_n^{\dagger}(x_0;\cdot))$. This relationship is investigated in Propositions 3.3.1 and 3.3.2.

<u>Proposition 3.3.1</u>: Let $f: E \to (\overline{E}_1,K)$ with (E_1,K) as in (3.3.1). Suppose each $f_i: E \to (\overline{E}_{1i},K_i)$ is u.s.c. at $x_0 \in f^{-1}(E_1)$. Then for all $y \in E$,

$$(3.3.2) \qquad f^{\dagger}(x_0;y) \geq_K (f_1^{\dagger}(x_0;y),\ldots,f_n^{\dagger}(x_0;y)) .$$

<u>Proof</u>: Let $(y,d_1,\ldots,d_n) \in T_f(x_0)$, and let $Y \in N(x_0)$ and $V_i \in N(d_i)$, $i = 1,\ldots,n$. There exist $X \in N(x_0)$, $Z_i \in N(f_i(x_0))$, and $\lambda > 0$ such that for all $(x,z_1,\ldots,z_n) \in X \times \prod_{i=1}^{n} Z_i$ with $f_i(x) \leq_{K_i} z_i$, $i=1,\ldots,n$, there exists $(y',d_1',\ldots,d_n') \in Y \times \prod_{i=1}^{n} V_i$ such that $\frac{f_i(x+ty')-z_i}{t} \leq_{K_i} d_i'$. Since each f_i is u.s.c. at x_0, there exists $X' \subset X$ such that $X' \in N(x_0)$ and $f_i(X') \subset Z_i - \dot{K}_i$. Now for all $(x,z_1) \in X' \times Z_1$ with $f_1(x) \leq_{K_1} z_1$, we have $f_i(x) \in Z_i - \dot{K}$ for each i and so there exist $z_i \in Z_i$, $i = 2,\ldots,n$ with $f_i(x) \leq_{K_i} z_i$. Thus for all $(x,z_1) \in X \times Z_1$ with $f_1(x) \leq_{K_1} z_1$ and for all $t \in (0,\lambda)$, there exists $(y',d_1') \in Y \times V_1$ such that $\frac{f_1(x+ty')-z_1}{t} \leq_{K_1} d_1'$. Hence $(y,d_1) \in T_{f_1}(x_0)$. By a

similar argument, $(y,d_i) \in T_{f_i}(x_0)$, $i=2,\dots,n$, and we obtain (3.3.2). □

Proposition 3.3.2: Let $f: E \to (\bar{E}_1,K)$ with (E_1,K) as in (3.3.1). Let $x_0 \in f^{-1}(E_1)$. Then for all $y \in E$,

$$(3.3.3) \qquad f^{\uparrow}(x_0;y) \leq_K (f_1^{\uparrow}(x_0;y), f_2^I(x_0;y), \dots, f_n^I(x_0;y))$$

and

$$(3.3.4) \qquad f^I(x_0;y) \leq_K (f_1^I(x_0;y), \dots, f_n^I(x_0;y)) .$$

Proof: Let $y \in E$ be such that $(y,d_1) \in T_{f_1}(x_0)$, $(y,d_i) \in Q_{\mathrm{epi}\, f_i}(x_0,f_i(x_0))$, $i = 2,\dots,n$ for some $d_i \in E_{1i}$, $i = 1,\dots,n$. It suffices to show that $(y,d_1,\dots,d_n) \in T_f(x_0)$.

To this end, let $Y_1 \in N(y)$ and $V_i \in N(d_i)$, $i=1,\dots,n$. There exist $X_i \times Z_i \in N((x_0,f_i(x_0)))$, $\lambda_i > 0$, and $Y_i \in N(y)$, $i = 2,\dots,n$ such that for all $(x_i,z_i) \in \mathrm{epi}_{K_i} f_i \cap (X_i \times Z_i)$, $t_i \in (0,\lambda_i)$, and $y_i' \in Y_i$, there exist $d_i' \in V_i$ with $\dfrac{f_i(x_i+t_iy_i')-z_i'}{t} \leq d_i'$. In addition, there exist $X_1 \times Z_1 \in N((x_0,f_1(x_0)))$ and $\lambda_1 > 0$ such that for all $(x,z) \in \mathrm{epi}_{K_1} f_1 \cap (X_1 \times Z_1)$ and all $t \in (0,\lambda_1)$, there exist $d_1' \in V_1$ and $y_1' \in Y_1 \cap \dots \cap Y_n$ with $\dfrac{f_1(x+ty_1')-z}{t} \leq d_1'$. Thus for all $(x,z_1,\dots,z_n) \in (\bigcap_{i=1}^{n} X_i) \times Z_1 \times \dots \times Z_n \cap \mathrm{epi}_K f$ and $t \in (0,\min(\lambda_1,\dots,\lambda_n))$; there exist $y' \in Y_1 \cap \ldots \cap Y_n$ and $(d_1', \ldots, d_n') \in \prod_{i=1}^{n} V_i$ with

$$\frac{f(x+ty') - (z_1, \ldots, z_n)}{t} \leq_K (d_1', \ldots, d_n') .$$

Therefore $(y, d_1, \ldots, d_n) \in T_f(x_0)$, and (3.3.3) holds. The proof of (3.3.4) is easier, and we leave it to the reader. □

<u>Corollary 3.3.3</u>: Let $E_{1i} := \mathbb{R}$ and $K_i := \mathbb{R}_+$ in Proposition 3.3.2. Assume that f_i, $i = 2, \ldots, n$ are directionally Lipschitzian at x_0 and

$$(3.3.5) \qquad \operatorname{dom} f_1^{\uparrow}(x_0;\cdot) \cap \bigcap_{i=2}^{n} \operatorname{int} \operatorname{dom} f_i^{\uparrow}(x_0;\cdot) \neq \phi .$$

Then for all $y \in E$,

$$(3.3.6) \qquad f^{\uparrow}(x_0;y) \leq_K (f_1^{\uparrow}(x_0;y), \ldots, f_n^{\uparrow}(x_0;y)) .$$

<u>Proof</u>: Call $D_i := \operatorname{dom} f_i^{\uparrow}(x_0;\cdot)$, $i = 1, \ldots, n$. The object is to show that for all $y \in E$, if $(y, d_i) \in T_{f_i}(x_0)$, $i = 1, \ldots, n$, then $(y, d_1, \ldots, d_n) \in T_f(x_0)$. If $y \notin D_i$ for some i, this implication holds vacuously. If $y \in D_1 \cap \bigcap_{j=2}^{n} \operatorname{int} D_j$, then $f_i^{\uparrow}(x_0;y) = f_i^{I}(x_0;y)$ by [Ro4, Theorem 3], and the inequality (3.3.6) holds by Proposition 3.3.2. If $y \in \bigcap_{i=1}^{n} D_i$, argue as in the proof of [Ro3, Theorem 2]: There exists $h \in D_1 \cap \bigcap_{i=2}^{n} \operatorname{int} D_i$, so

$(1-\varepsilon)y + \varepsilon h \in D_1 \cap \bigcap_{i=2}^{n} \operatorname{int} D_i$ for any $\varepsilon \in (0,1)$. Now $f^{\dagger}(x_0;\cdot)$ is $\mathbb{R}^n_+$-convex and l.s.c. and each $f_i^{\dagger}(x_0;\cdot)$ is convex and l.s.c., so letting $\varepsilon \downarrow 0$, we obtain the inequality (3.3.6) for y. □

We remark here that (3.3.6) will hold for infinite-dimensional (E_1,K) if the functions $f_2,\ldots,f_n$ in Proposition 3.2.6 are strictly compactly Lipschitzian at x_0 (defined in [Th1], [Th2]). For such functions, $f_j^{\dagger}(x_0;\cdot) = f_j^I(x_0;\cdot)$ ([Th1, Proposition 3.10]).

<u>Proposition 3.3.4:</u> Under the hypotheses of Theorem 3.2.6, let (E_1,K) be of the form (3.3.1). Then for all $y \in E$,

$$(3.3.7)\qquad (F\circ f)^{\dagger}(x_0;y) \leq_S F^I(f(x_0);f_1^{\dagger}(x_0;y),f_2^I(x_0;y),\ldots,f_n^I(x_0;y))$$

and

$$(3.3.8)\qquad (F\circ f)^I(x_0;y) \leq_S F^I(f(x_0);f_1^I(x_0;y),\ldots,f_n^I(x_0;y)).$$

<u>Proof:</u> The inequalities (3.3.7) and (3.3.8) follow immediately from (3.3.3) and (3.3.4) and the isotonicity of $F^I(f(x_0);\cdot)$ (Lemma 3.1.9). □

We next consider an important special case of Proposition 3.3.4 - that in which $F(z_1,\ldots,z_n) = \sum_{i=1}^{n} z_i$. Verifying that condition (3.2.6) holds in this special case is the first step toward establishing a subgradient sum formula as in [Th1].

Lemma 3.3.5: Let $f\colon E \to \prod_{i=1}^{n} (\overline{E}_1, K)$ be l.s.c. at $x_0 \in f^{-1}(\prod_{i=1}^{n} E_1)$, and let $F\colon \prod_{i=1}^{n} (\overline{E}_1, K) \to (\overline{E}_1, K)$ be defined by $F(z_1,\ldots,z_n) := \sum_{i=1}^{n} z_i$, where $\sum_{i=1}^{n} z_i = +\infty$ if any $z_i = +\infty$. Assume K is normal. Then (3.2.6) holds.

Proof: Let V be a full neighbourhood of 0 in E_1, and let V_i, $i = 1,\ldots,n$ be symmetric full neighbourhoods of 0 in E_1 such that $\sum_{i=1}^{n} V_i \subset V$. Call $Y := f(x_0) + \prod_{i=1}^{n} V$ $\in N(f(x_0))$. Since f is l.s.c. at x_0, there exists $X \in N(x_0)$ such that

$$
\begin{aligned}
f_i(X) &\subset f_i(x_0) + V_i + K \\
&\subset f_i(x_0) + V + K \quad \text{for each } i .
\end{aligned}
$$

Let $V' = \bigcap_{i=1}^{n} V_i \in N(0)$, and suppose

$$(x,z) \in X \times \left(\sum_{i=1}^{n} f_i(x_0) + V'\right)$$

with $\sum_{i=1}^{n} f_i(x) \leq_K z$. Then

$$f_i(x) \in \sum_{i=1}^{n} f_i(x_0) + V' - K . \tag{3.3.9}$$

Since $-f_i(x) \in -f_i(x_0) - V_i - K$ and each V_i is symmetric, it follows from (3.3.9) that for each i,

$$f_i(x) \subset \sum_{i=1}^{n} f_i(x_0) + V' - K - \sum_{j\neq i} f_j(x)$$

$$\subset \sum_{i=1}^{n} f_i(x_0) + V_i - K - \sum_{j\neq i} (f_j(x_0)+V_j + K)$$

$$\subset f_i(x_0) + V - K .$$

Since V is a full neighbourhood, the above inclusions show that $f_i(x) \subset f_i(x_0) + V$. Letting $y := f(x)$, we have $f(x) \leq_{\pi K} y$, $F(y) \leq_K z$, and $y \in Y$, as required to prove condition (3.2.6). □

<u>Proposition 3.3.6</u>: Let K be normal, and let $f_i : E \to (\overline{E}_1, K)$ $i = 1,\ldots,n$ be l.s.c. at $x_0 \in \bigcap_{i=1}^{n} f_i^{-1}(E_1)$. If $f := (f_1,\ldots,f_n)$ is T epigraph regular at x_0, then for all $y \in E$,

$$(\sum_{i=1}^{n} f_i)^{\uparrow}(x_0;y) \leq_K f_1^{\uparrow}(x_0;y) + \sum_{i=2}^{n} f_i^{I}(x_0;y) . \tag{3.3.10}$$

If f is I epigraph regular, then for all $y \in E$,

$$(\sum_{i=1}^{n} f_i)^{I}(x_0;y) \leq_K \sum_{i=1}^{n} f_i^{I}(x_0;y) . \tag{3.3.11}$$

<u>Proof</u>: Since (3.2.6) holds under these assumptions, we may apply Proposition 3.3.4. Then (3.3.7) reduces to (3.3.10) and (3.3.8) becomes (3.3.11). □

One way to guarantee that f in Proposition 3.3.6 is I epigraph regular is to assume and f_i is continuous at x_0 and (E_1,K) is a D.t.l. Then (3.3.2) and (3.3.4) combine to give $f^I(x_0;y) = (f_1^I(x_0;y),\ldots,f_n^I(x_0;y))$, and f is I epigraph regular by Lemma 3.1.8. If in addition $f_i^{\uparrow}(x_0;\cdot) = f_i^I(x_0;\cdot)$, $i = 2,\ldots,n$, f is also T epigraph regular.

Unfortunately, Theorem 3.2.15 is not strong enough to enable us to derive a subgradient sum formula; in fact, all (3.2.16) tells us in this case is that $\partial_K(\sum_{i=1}^n f_i)(x_0) = \partial_K(\sum_{i=1}^n f_i)(x_0)$. There is, however, a subgradient sum formula for convex vector-valued functions (see [Zo1], [Bo9], [Pa1]). We state here the n-function version of this result.

<u>Theorem 3.3.7</u>: Let $f_i: E \to (E_1,K)$, $i = 1,\ldots,n$ be K-convex functions, and let $x_0 \in \bigcap_{i=1}^n \operatorname{dom} f_i$. If there exists a point $z \in \bigcap_{i=1}^n \operatorname{dom} f_i$ such that f_i, $i = 2,\ldots,n$ are continuous at z, then

$$\partial_K(\sum_{i=1}^n f_i)(x_0) = \sum_{i=1}^n \partial_K f_i(x_0) . \tag{3.3.12}$$

We may now combine Proposition 3.3.6 and Theorem 3.3.7 to obtain a subgradient sum formula. Such a formula has also been derived by Thibault [Th1]. His result requires fewer assumptions than the one we give here; by deriving (3.3.10) directly rather than via Theorem 3.2.6, one can avoid the hypotheses of epigraph regularity [Th1, Lemma 3.6].

Theorem 3.3.8 (cf. [Th1, Proposition 3.9]): Let S be normal, and let $f_i: E \to (\overline{E}_1, S)$, $i = 1,\ldots,n$, be l.s.c. at $x_0 \in \bigcap_{i=1}^{n} f_i^{-1}(E_1)$. Assume that $f := (f_1,\ldots,f_n)$ is T epigraph regular at x_0, that $f_1^{\uparrow}(x_0;\cdot)$ and $f_i^{I}(x_0;\cdot)$, $i = 2,\ldots,n$ are proper, and that there is a point $z \in \operatorname{dom} f_1^{\uparrow}(x_0;\ ) \cap \bigcap_{i=2}^{n} \operatorname{dom} f_i^{I}(x_0;\ )$ at which each $f_i^{I}(x_0;\cdot)$, $i = 2,\ldots,n$ is continuous. Then

$$\partial_S\left(\sum_{i=1}^{n} f_i\right)(x_0) \subset \sum_{i=1}^{n} \partial_S f_i(x_0) \, . \tag{3.3.13}$$

Proof: Apply Theorem 3.3.7 to $f_1^{\uparrow}(x_0;\cdot)$, $f_i^{I}(x_0;\cdot)$, $i = 2,\ldots,n$. Then (3.3.10) and (3.3.12) combine to give (3.3.13). □

We next consider the case of $F(y_1,\ldots,y_n) := \sup_{1\le i\le n} y_i$, beginning with a demonstration that (3.2.6) is satisfied in this case.

Lemma 3.3.9: Suppose S is normal and induces a lattice ordering. Let $f: E \to \prod_{i=1}^{n} (\overline{E}_1, S)$ be l.s.c. at $x_0 \in f^{-1}\left(\prod_{i=1}^{n} E_1\right)$, and let $F: \prod_{i=1}^{n} (\overline{E}_1, S) \to (\overline{E}_1, S)$ be defined by $F(z_1,\ldots,z_n) := \sup_{1\le i\le n} z_i$. Call

$$I(x_0) := \{i \in \{1,\dots,n\} \mid f_i(x_0) = F(f(x_0))\} .$$

Assume f_i is continuous at x_0 for each $i \notin I(x_0)$. Then (3.2.6) holds.

<u>Proof</u>: Let V_i, $i = 1,\dots,n$ be full neighbourhoods of 0 in E_1, and call $Y := f(x_0) + \prod_{i=1}^{n} V_i \in N(f(x_0))$. By hypothesis, there exists $X \in N(x_0)$ such that

$$f_i(X) \subset f_i(x_0) + V_i + \dot{S} \quad \text{for each} \quad i \in I(x_0) ,$$

and $$f_i(X) \subset f_i(x_0) + V_i \quad \text{for each} \quad i \notin I(x_0) .$$

Set $Z := F(f(x_0)) + \bigcap_{i=1}^{n} V_i$. Now suppose $(x,z) \in X \times Z$ with $F(f(x)) \leq_S z$. Then for each $i \in I(x_0)$, $f_i(x) \in f_i(x_0) + V_i - S$, so $f_i(x) \in f_i(x_0) + V_i$ since V_i is full. Set $y := f(x)$. Then $f(x) \leq_{\pi S} y$, $F(y) \leq_S z$, and $y \in Y$. Hence (3.2.6) is satisfied. □

<u>Proposition 3.3.10</u>: Suppose S induces a lattice ordering on E_1 and is normal. Let $f_i : E \to (\dot{E}_1,S)$, $i = 1,\dots,n$ be l.s.c. at x_0. Define
$I(x_0) := \{i \in \{1,\dots,n\} \mid f_i(x_0) = \sup_{1\leq i\leq n} f_i(x)\}$. Assume f_i is continuous at x_0 for each $i \notin I(x_0)$. If $f := (f_1,\dots,f_n)$ is I epigraph regular at x_0, then for all $y \in E$,

$$(3.3.14) \qquad (\sup_{1\le i\le n} f_i)^{I}(x_0;y) \le_S \sup_{I(x_0)} f_i^{I}(x_0;y) .$$

If f is T epigraph regular at x_0, then for all $y \in E$,

$$(3.3.15) \qquad (\sup f_i)^{\dagger}(x_0;y) \le_S \sup_{I(x_0)} \{f_{j_1}^{\dagger}(x_0;y), f_{j_2}^{I}(x_0;y), \ldots, f_{j_m}^{I}(x_0;y)\}$$

where $I(x_0) = \{j_1,\ldots,j_m\}$.

<u>Proof</u>: Let $F : \prod_{i=1}^{n} (\dot{E}_1,S) \to (\dot{E}_1,S)$ be defined by $F(y_1,\ldots,y_n) := \sup_{1\le i\le n} y_i$. By Lemma 3.3.9, condition (3.2.6) holds, and we may apply Proposition 3.3.4. In this case, (3.3.8) and (3.3.7) become (3.3.14) and (3.3.15). □

Again in this case, Theorem 3.2.15 will not give the appropriate formula for convex functions. However, we can derive such a formula by applying Theorem 3.2.14, as we now demonstrate.

<u>Theorem 3.3.11</u>: Let $f_i : E \to (\dot{E}_1,S)$, $i = 1,\ldots,m$ be proper S-convex functions, and suppose $x_0 \in \bigcap_{i=1}^{m} \operatorname{dom} f_i$ is such that $f_1(x_0) = \ldots = f_m(x_0)$. Assume that S is pointed and induces a lattice ordering, and that there exists $\hat{x} \in \bigcap_{i=1}^{n} \operatorname{dom} f_i$ such that f_i, $i = 2,\ldots,n$ are continuous at $\hat{x}$. Then

$$(3.3.16)\qquad \partial_S(\sup_{1\le i\le m} f)(x_0) = \{\sum_{i=1}^{m} \partial_S(\lambda_i f_i)(x_0) \mid \lambda_i \in L(E_1,E_1),\ \lambda_i(S)\subset S,\ \sum_{i=1}^{m} \lambda_i = I_{E_1}\} .$$

where $I_{E_1}: E_1 \to E_1$ is the identity mapping on E_1. (Note: If $\lambda_j \equiv 0$ in (3.3.16), $\lambda_j f_j = i_{\text{dom } f_j}$ by convention.)

Proof: Call $h(x) := \sup_{1\le i\le n} f_i(x)$, and suppose T is an element of the right hand side of (3.3.16). Then there exist $\lambda_i \in L(E_1,E_1)$, with $\sum_{i=1}^{m} \lambda_i = I_{E_1}$ and $\lambda_i(S) \subset S$ such that for all $x \in E$,

$$T(x-x_0) \le_S \sum_{i=1}^{m} \lambda_i f_i(x) - \sum_{i=1}^{m} \lambda_i f_i(x_0)$$

$$\le_S \sum_{i=1}^{m} \lambda_i h(x) - \sum_{i=1}^{m} \lambda_i h(x_0)$$

$$= h(x) - h(x_0) .$$

Therefore $T \in \partial_S h(x_0)$. Conversely, suppose $T \in \partial_S h(x_0)$. Since $T \in \partial_S h(x_0)$ if and only if $0 \in \partial_S(\sup_{1\le i\le m}(f_i - T)(\cdot))(x_0)$, we may assume without loss of generality that $T = 0$. Then $h(x_0) = \inf_S\{z \mid f_1(x)-z \le_S 0,\dots,f_m(x)-z \le_S 0\}$. For this program, (3.2.14) and (3.2.15) hold as long as $\bigcap_{i=1}^{n} \text{dom } f_i \ne \phi$, so we may apply Theorem 3.2.14. There then exist $\lambda_i \in L(E_1,E_1)$ with $\lambda_i(S) \subset S$ and

$$(3.3.17) \qquad h(x_0) \le_S z + \sum_{i=1}^{m} \lambda_i (f_i(x) - z)$$

for all $x \in E$ and $z \in E_1$. Setting $x := x_0$ in (3.3.17), we see that

$$0 \le_S (I_{E_1} - \sum_{i=1}^{m} \lambda_i) z \quad \text{for all} \quad z \in E_1 .$$

Since S is pointed, we then have $\sum_{i=1}^{m} \lambda_i = I_{E_1}$. By (3.3.17), this means that

$$\sum_{i=1}^{m} \lambda_i f_i(x_0) = h(x_0) \le_S \sum_{i=1}^{m} \lambda_i f_i(x) ,$$

or

$$0 \le_S \sum_{i=1}^{m} \lambda_i f_i(x) - \sum_{i=1}^{m} \lambda_i f_i(x_0) .$$

Hence $0 \in \partial(\sum_{i=1}^{m} \lambda_i f_i)(x_0)$. Finally, since there exists $\hat{x} \in \bigcap_{i=1}^{m} \operatorname{dom} f_i$ with $f_2,\ldots,f_m$ continuous at $\hat{x}$, $\lambda_2 f_2,\ldots,\lambda_m f_m$ are also continuous at $\hat{x}$, and we may apply Theorem 3.3.7 to the $\lambda_i f_i$ to obtain $0 \in \sum_{i=1}^{m} \partial(\lambda_i f_i)(x_0)$. Therefore (3.3.16) holds. □

Proposition 3.3.10 and Theorem 3.3.11 combine to give the following result:

Theorem 3.3.12: Suppose that S is pointed, normal, and induces a lattice ordering on E_1. Let $f_i: E \to (\dot{E}_1, S)$, $i = 1,\dots,n$ be l.s.c. at $x_0 \in \bigcap_{i=1}^{m} \operatorname{dom} f_i$, and assume f_i is continuous at x_0 for each $i \notin I(x_0)$ and $f:=(f_1,\dots,f_n)$ is T epigraph regular at x_0. Assume $f_{j_i}^{I}(x_0;\cdot)$, $i = 2,\dots,m$ and $f_{j_1}^{\dagger}(x_0;\cdot)$ are proper, where $I(x_0) = \{j_1,\dots,j_m\}$, and that there is a point $z \in \operatorname{dom} f_{j_1}^{\dagger}(x_0;\cdot) \cap \bigcap_{i=2}^{m} f_{j_i}^{I}(x_0;\cdot)$ at which each $f_{j_i}^{I}(x_0;\cdot)$ $i = 2,\dots,m$ is continuous. Then

$$(3.3.18) \qquad \partial_S\Big(\sup_{1\le i\le m} f_i\Big)(x_0) \subset \Big\{ \sum_{I(x_0)} \partial_S(\lambda_i f_i) \,\Big|\, \lambda_i \in L(E_1,E_1),\ \lambda_i(S) \subset S \text{ and } \sum_{I(x_0)} \lambda_i = I_{E_1} \Big\}$$

Proof: Apply Theorem 3.3.11 to the functions $f_{j_1}^{\dagger}(x_0;\cdot)$, $f_{j_i}^{I}(x_0;\cdot)$, $i = 2,\dots,m$. Then (3.3.15) and (3.3.16) combine to give (3.3.18). □

Remark 3.3.13: Certain questions not addressed here need to be dealt with more thoroughly in future work. For example, what conditions guarantee that $f^{I}(x_0;\cdot)$ is continuous (an assumption made often in sections 3.2 and 3.3)? We know from section 3.1 that if E_1 is a complete metric space, (E_1,S) is a D.t.l. and $f: E \to (\overline{E}_1,S)$ is strictly compactly Lipschitzian at $x_0 \in E$, then $f^{I}(x_0;\ ) = f^{\dagger}(x_0;\cdot)$ is continuous on the core of its domain. However, less restrictive conditions for continuity of $f^{I}(x_0;\cdot)$ are probably provable. Another question: What are conditions for equality in our chain rules?

3.4. The scalar-valued case

Among the consequences of our results of sections 3.2 and 3.3 are new subdifferential calculus results for extended real valued functions. We have already seen one such formula in Theorem 3.2.13. In this section we give two more applications to the scalar-valued case. The first is an n-function version of Theorem 3.2.13:

<u>Theorem 3.4.1</u>: Let E be a l.c.s., and let $f_i\colon E \to \overline{\mathbb{R}}$, $i = 1,\dots,n$ and $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be extended real valued functions with $F(y_1,\dots,y_n) = +\infty$ if some $y_i = +\infty$. Assume either that each f_i is proper or that $\inf_{y\in\mathbb{R}^n} F(y) = -\infty$ and $F(y_1,\dots,y_n) = -\infty$ if some $y_i = -\infty$ and each $y_i < +\infty$. Call $f := (f_1,\dots,f_n)$. Suppose $f(x_0)$ and $F(f(x_0))$ are finite. Assume either that each $f_i^{\uparrow}(x_0;0) = 0$ or that $\inf_{y\in\mathbb{R}^n} F^{\uparrow}(f(x_0);y) = -\infty$ and $F^{\uparrow}(f(x_0);f_1^{\uparrow}(x_0;0),\dots,f_n^{\uparrow}(x_0;0)) = -\infty$. Suppose F is isotone on $(\text{Range } f + \mathbb{R}^n_+) \cup B_\varepsilon(f(x_0))$ for some $\varepsilon > 0$. Assume that each f_i is continuous at x_0, that f_i, $i = 2,\dots,n$ are directionally Lipschitzian at x_0, and that

$$(3.4.1)\qquad \{(r_1,\ldots,r_n)\in \mathbb{R}^n \mid \exists\ y\in E,\ f_i^{\uparrow}(x_0;y)\le r_i,\ i=1,\ldots,n\} \cap \operatorname{int}\operatorname{dom} F^{\uparrow}(f(x_0);\cdot)\neq \phi$$

$$(3.4.2)\qquad \operatorname{dom} f_1^{\uparrow}(x_0;\cdot)\cap \bigcap_{i=2}^{n} \operatorname{int}\operatorname{dom} f_i^{\uparrow}(x_0;\cdot)\neq\phi\ .$$

Then for all $y \in E$,

$$(3.4.3)\qquad (F\circ f)^{\uparrow}(x_0;y)\le F^{\uparrow}(f(x_0);f_1^{\uparrow}(x_0;y),\ldots,f_n^{\uparrow}(x_0;y))\ .$$

Moreover,

$$(3.4.4)\qquad \partial(F\circ f)(x_0)\subset\{\lambda(\partial f_1(x_0),\ldots,\partial f_n(x_0))\mid \lambda\in\mathbb{R}^n_+\ ,\ \lambda\in\partial F(f(x_0))\}\ ;$$

Equality holds in (3.4.4) if in addition F is subdifferentiably regular at $f(x_0)$, $f_2,\ldots,f_n$ are subdifferentiably regular at x_0 , and f_1 is subdifferentiably weakly regular at x_0 . Equality holds in (3.4.3) if these regularity conditions hold and $f_i^{\uparrow}(x_0;0) = 0$, $i = 1,\ldots,n$.

<u>Proof</u>: Consider Theorem 3.2.6 with $(E_1,K) = (\mathbb{R}^n,\mathbb{R}^n_+)$ and $(E_2,S) = (\mathbb{R},\mathbb{R}_+)$. Continuity of each f_i guarantees that (3.2.6) is satisfied. Since (3.4.2) holds, Proposition 3.3.1 and Corollary 3.3.3 combine to give $f^{\uparrow}(x_0;\cdot) = (f_1^{\uparrow}(x_0,\cdot),\ldots,f_n^{\uparrow}(x_0;\cdot))$. By Lemma 3.1.8, f is T epigraph regular at x_0 . Then by Theorem 3.2.6, we have for all $y \in E$ that

$$(F \circ f)^{\dagger}(x_0;y) \leq F^{I}(f(x_0);f^{\dagger}(x_0;y))$$
$$= F^{I}(f(x_0);f_1^{\dagger}(x_0;y),\ldots,f_n^{\dagger}(x_0;y)) .$$

Now since $F^{I}(f(x_0;\cdot)$ is isotone and thus directionally Lipschitzian, assumption (3.4.1) implies that in fact

$$(F \circ f)^{\dagger}(x_0;y) \leq F^{\dagger}(f(x_0);f_1^{\dagger}(x_0;y),\ldots,f_n^{\dagger}(x_0;y)) ,$$

by an argument similar to that used in Corollary 3.3.3.

As usual, we use our interiority assumptions twice. Assumptions (3.4.1) and (3.4.2) enable us to apply Theorem 2.2.12. By a now familiar argument, Theorem 2.2.12 and the inequality (3.4.3) combine to give (3.4.4). The assertions about equality follow as in Proposition 2.4.4. □

Working directly rather than via Theorem 3.2.6, it is possible to prove a variant of Theorem 3.4.1 in which the f_i's do not necessarily have to be continuous at x_0. We state this result below. The proof is straightforward and involves no new techniques, so we leave it to the reader.

<u>Theorem 3.4.2</u>: In Theorem 3.4.1, assume in addition that F is l.s.c. Replace the hypothesis that each f_i is continuous at x_0 with the hypothesis that for each $i \in \{1,\ldots,n\}$, either

(i) f_i is l.s.c. at x_0 and F is strictly isotone in the i^{th} coordinate at $f(x_0)$, or

(ii) f_i is continuous at x_0 .

Replace conditions (3.4.1) and (3.4.2) with the following condition:

$$(3.4.5) \qquad \{y \in E \mid (f_1^{\uparrow}(x_0;y),\ldots,f_n^{\uparrow}(x_0;y)) \in \operatorname{int} \operatorname{dom} F^{\uparrow}(f(x_0);\cdot)\}$$
$$\cap \operatorname{dom} f_1^{\uparrow}(x_0;\cdot) \cap \bigcap_{i=2}^{n} \operatorname{int} \operatorname{dom} f_i^{\uparrow}(x_0;\cdot) \neq \phi$$

Then the conclusions of Theorem 3.4.1 remain valid.

From Theorem 3.4.2 we may derive several calculus rules in the setting of [Ro3], including the following product and quotient rules not given in [Ro3].

<u>Corollary 3.4.3</u>: Let $f_i: E \to \mathbb{R}$, $i = 1,\ldots,n$ be nonnegative on E and positive and l.s.c. at $x_0 \in \bigcap_{i=1}^{n} \operatorname{dom} f_i$. Assume (3.4.2) holds. Then for all $y \in E$,

$$(3.4.6) \qquad \left(\prod_{i=1}^{n} f_i\right)^{\uparrow}(x_0;y) \leq \sum_{i=1}^{n} \left(\prod_{j\neq i} f_j(x_0)\right) f_i^{\uparrow}(x_0;y) .$$

Moreover,

$$(3.4.7) \qquad \partial\left(\prod_{i=1}^{n} f_i\right)(x_0) \subset \sum_{i=1}^{n} \left(\prod_{j\neq i} f_j(x_0)\right) \partial f_i(x_0) .$$

Equality holds in (3.4.7) if f_1 is subdifferentiably weakly regular at x_0 and $f_2,\ldots,f_n$ are subdifferentiably regular at x_0 . Equality holds in (3.4.6) if in addition each $f_i^{\uparrow}(x_0;0) = 0$.

Proof: In Theorem 3.4.2, consider $F: \mathbb{R}^n \to \mathbb{R}$ defined by $F(y_1,\ldots,y_n) = \prod_{i=1}^{n} y_i$. Since each f_i is nonnegative on E and positive at x_0, F is isotone on the union of range $f + \mathbb{R}^n_+$ and $B_\varepsilon(f(x_0))$ for some $\varepsilon \geq 0$. Condition (3.4.5) is automatically satisfied since $\operatorname{dom} F^\dagger(f(x_0);\cdot) = \mathbb{R}^n$ and since we have assumed (3.4.2). Thus all the hypotheses of Theorem 3.4.2 are satisfied. In this case, (3.4.3) reduces to (3.4.6). If $f_i^\dagger(x_0;0) = -\infty$ for some i, then the right hand side of (3.4.6) is $-\infty$ and $(\Pi f_i)^\dagger(x_0;0) = -\infty$, so $\partial(\prod_{i=1}^{n} f_i)(x_0) = \phi$ and equality holds in (3.4.7). Otherwise each $f_i^\dagger(x_0;\cdot)$ is proper, and (3.4.4) becomes (3.4.7) in this case. Equality under the conditions given follows from Theorem 3.4.2 since F is subdifferentiably regular at $f(x_0)$. □

Proposition 3.4.4: Let $f: E \to \overline{\mathbb{R}}$ be nonnegative on E and positive and l.s.c. at x_0, and let $g: E \to \overline{\mathbb{R}}$ be positive on E and continuous at x_0. Assume

$$(3.4.8) \qquad \operatorname{dom} f^\dagger(x_0;\cdot) \cap \operatorname{int} \operatorname{dom}(-g)^\dagger(x_0;\cdot) \neq \phi .$$

Then for all $y \in E$,

$$(3.4.9) \qquad \left(\frac{f}{g}\right)^\dagger(x_0;y) \leq \frac{f(x_0)(-g)^\dagger(x_0;y)+g(x_0)f^\dagger(x_0;y)}{(g(x_0))^2}$$

Moreover,

$$(3.4.10) \qquad \partial(\tfrac{f}{g})(x_0) \subset \frac{f(x_0)\partial(-g)(x_0) + g(x_0)\partial f(x_0)}{(g(x_0))^2}$$

Equality holds in (3.4.10) if f is subdifferentiably weakly regular at x_0 and $\frac{1}{g}$ is subdifferentiably regular at x_0. Equality holds in (3.4.9) if in addition $f^{\uparrow}(x_0;0) = 0$ and $(-g)^{\uparrow}(x_0;0) = 0$.

Proof: Apply Corollary 3.4.3 with $n = 2$, $f_1 := f$, and $f_2 := \frac{1}{g}$. By Lemma 2.3.19, $\mathrm{dom}(\frac{1}{g})^{\uparrow}(x_0;\cdot) = \mathrm{dom}(-g)^{\uparrow}(x_0;\cdot)$, so (3.4.8) is equivalent to (3.4.2) in this case. Then (3.4.9) and (3.4.10) follow from (3.4.6), (3.4.7), and Lemma 2.3.19, as in the proof of Proposition 2.3.20. □

For functions with finite-dimensional domains, the conditions (3.4.1) and (3.4.2) can be weakened, as we saw in Theorem 2.3.15. However, the hypothesis used there ((2.3.28)) has the disadvantage of being rather complicated in general. We now show that by employing the methods of chapter 2, we can weaken condition (3.4.2) in the case in which E is finite-dimensional. As a result, we can obtain a variant of Theorem 2.3.15 in which parts (i) and (ii) of (2.3.28) are decoupled at the price of an extra continuity assumption. The first step is to establish a technical lemma along the lines of Lemmas 2.2.6 and 2.2.8.

Lemma 3.4.5: Let $f_i: \mathbb{R}^m \to \overline{\mathbb{R}}$, $i = 1,\ldots,n$ be extended real valued functions which are finite at $x_0 \in \mathbb{R}^m$. Define $A: (\mathbb{R}^m \times \mathbb{R})^n \to \mathbb{R}^m \times \mathbb{R}^n$, by $A(x_1,y_1,\ldots,x_n,y_n) :=$ $(x_1,y_1,y_2,\ldots,y_n)$ and $G: (\mathbb{R}^m \times \mathbb{R})^n \to \mathbb{R}^{m(n-1)}$ by $G(x_1,y_1,\ldots,x_n,y_n) := (x_1-x_2,\ldots,x_1-x_n)$. Then condition (2.2.3) is satisfied with A as above,

$C := \text{epi } f_1 \times \ldots \times \text{epi } f_n \cap G^{-1}(0)$, and

$z_0 := (x_0,f_1(x_0),\ldots,x_0,f_n(x_0))$.

Proof: Let $\varepsilon > 0$ be given, and let $X := B_\varepsilon(z_0)$. Then $Z := A(X) \in N(Az_0)$. Suppose $(\hat{x}_1,\hat{y}_1,\hat{y}_2,\ldots,\hat{y}_n) \in Z \cap A(C)$. Since $(\hat{x}_1,\hat{y}_1,\ldots,\hat{y}_n) \in Z$, there exists $(x_1,y_1,\ldots,x_n,y_n) \in X$ with $\hat{x}_1 = x_1 \in B_\varepsilon(x_0)$ and $\hat{y}_i = y_i \in B_\varepsilon(f_i(x_0))$ for $i = 1,\ldots,n$. Since $(\hat{x}_1,\hat{y}_1,\ldots,\hat{y}_n) \in A(C)$, there exists $(x_1',y_1',\ldots,x_n',y_n')$ with $x_j' = \hat{x}_1$, $y_j' = \hat{y}_j$ and $f_j(x_j') \le y_j'$ for $j = 1,\ldots,n$. Hence $(x_1',y_1',\ldots,y_n') \in C \cap X$, and so $(\hat{x}_1,\hat{y}_1,\hat{y}_2,\ldots,\hat{y}_n) \in A(C \cap X)$. Therefore (2.2.3) holds. □

Having proven Lemma 3.4.5, we can now establish a finite-dimensional version of Corollary 3.3.3.

Proposition 3.4.6: Let $f_i: \mathbb{R}^m \to \overline{\mathbb{R}}$ be finite and strictly l.s.c. at x_0, and assume the sets $\text{dom } f_i^{\uparrow}(x_0;\cdot)$, $i = 1,\ldots,n$ are in strong general position. Call $f := (f_1,\ldots,f_n)$ and $S := \mathbb{R}^n_+$. Then for all $y \in \mathbb{R}^m$,

$$f^{\uparrow}(x_0;y) \leq_S (f_1{}^{\uparrow}(x_0;y),\ldots,f_n{}^{\uparrow}(x_0;y)) . \tag{3.4.11}$$

Proof: Define A, G, and z_0 as in Lemma 3.4.5, and define $D := \text{epi } f_1 \times\ldots\times \text{epi } f_n$. Then

$$\begin{aligned} \text{epi}_S f^{\uparrow}(x_0;\cdot) &\supset T_f(x_0) \quad \text{by Proposition 3.1.6} \\ &= T_{A(D\cap G^{-1}(0))}(x_0,f(x_0)) \\ &\supset A(T_{D\cap G^{-1}(0)}(z_0)) \end{aligned}$$

by Lemma 3.4.5 and Proposition 2.2.4

$$\supset A(T_D(z_0) \cap \nabla G(z_0)^{-1}(0)) ,$$

by our strong general position hypothesis and Corollary 2.2.3,

$$= A(\{(x_1,y_1,\ldots,x_n,y_n) \mid f_j^{\uparrow}(x_0;x_j) \leq y_j,\ x_j = x_1,\ j = 1,\ldots,n\})$$

since the Clarke tangent cone is product-preserving (section 1.4),

$$= \{(x,y_1,\ldots,y_n) \mid f_i{}^{\uparrow}(x_0;x) \leq y_i ,\ i = 1,\ldots,n\}$$

$$= \text{epi}_S(f_1{}^{\uparrow}(x_0;\cdot),\ldots,f_n{}^{\uparrow}(x_0;\cdot))$$

Therefore $\text{epi}_S f^{\uparrow}(x_0;\cdot) \supset \text{epi}_S(f_1{}^{\uparrow}(x_0;\ ),\ldots,f_n{}^{\uparrow}(x_0;\cdot))$ and (3.4.11) holds. □

Remark 3.4.7: (a) The inequality (3.4.11) does not hold in general. For example, define $f_1 : \mathbb{R} \to \overline{\mathbb{R}}$ and $f_2 : \mathbb{R} \to \overline{\mathbb{R}}$ by $f_1 := i_{\mathbb{Q}}$ and $f_2 := i_{(\mathbb{R}\setminus\mathbb{Q})\cup\{0\}}$, and let $x_0 := 0$. Call

$S := \mathbb{R}^2_+$. Then $\mathrm{epi}_S f = \{(0,y,z)\,|\,y \geq 0,\ z \geq 0\}$ and

$$f^{\dagger}(0;y) = \begin{cases} (0,0) & \text{if } y=0 \\ +\infty & \text{else} \end{cases}, \text{ while } f_1^{\dagger}(0;\cdot) = f_2^{\dagger}(0;\cdot) = i_{\mathbb{R}}(\cdot) .$$

In this example the strong general position assumption holds, but the functions f_1 and f_2 are not l.s.c. at 0 .

Another example: Define $C_1 := \{(x,y) \in \mathbb{R}^2\,|\,y = x^2\}$, $C_2 := \{(x,y) \in \mathbb{R}^2\,|\,y = -x^2\}$, and $f_1: \mathbb{R}^2 \to \overline{\mathbb{R}}$ and $f_2: \mathbb{R}^2 \to \overline{\mathbb{R}}$ by $f_1 := i_{C_1}$ and $f_2 := i_{C_2}$. Again let $S := \mathbb{R}^2_+$. Then $\mathrm{epi}_S f = \{(0,0,y,z)\,|\,y \geq 0,\ z \geq 0\}$ and

$$f^{\dagger}((0,0);y) = \begin{cases} (0,0) & \text{if } y = (0,0) \\ +\infty & \text{else} \end{cases} .$$

On the other hand, $f_1{}^{\dagger}((0,0);\cdot) = f_2{}^{\dagger}((0,0);\cdot) = i_C(\cdot)$, where $C = \{(x,y) \in \mathbb{R}^2\,|\,y = 0\}$. In this example f_1 and f_2 are l.s.c. but the strong general position assumption is not satisfied.

(b) Comparing Proposition 3.3.1 with Proposition 3.4.6, we see that for $f := (f_1,\ldots,f_n)$, $f^{\dagger}(x_0;\cdot) = (f_1{}^{\dagger}(x_0;\cdot),\ldots, f_n{}^{\dagger}(x_0;\cdot))$ whenever $f_i: \mathbb{R}^m \to \overline{\mathbb{R}}$, $i = 1,\ldots,n$ are continuous at x_0 and the sets $\mathrm{dom}\, f_i{}^{\dagger}(x_0;\cdot)$ are in strong general position. In particular, if each f_i is locally Lipschitzian near x_0 , then $f^{\dagger}(x_0;\cdot) = (f_1^0(x_0;\cdot),\ldots, f_n^0(x_0;\cdot))$ and $\partial^T_{\mathbb{R}^n_+} f(x_0) = (\partial f_1(x_0),\ldots,\partial f_n(x_0))$. This means that for locally Lipschitzian functions $f: \mathbb{R}^m \to \mathbb{R}^n$, $\partial^T_{\mathbb{R}^n_+} f$ is a "rougher" local approximation than the generalized Jacobian of Clarke ([Cl2], [Cl4]).

For further discussion, see [Hi4].

Theorem 3.4.8: In Theorem 3.4.1, let $E := \mathbb{R}^m$ and assume each $f_i : \mathbb{R}^m \to \overline{\mathbb{R}}$ is continuous at x_0. Replace assumption (3.4.2) with the assumption that $\text{dom } f_i^{\uparrow}(x_0;\cdot)$, $i = 1,\ldots,n$ are in strong general position. Then (3.4.3) and (3.4.4) hold. Equality holds in (3.4.3) and (3.4.4) under the assumptions given in Theorem 3.4.1.

Proof: The proof parallels that of Theorem 3.4.1, with Proposition 3.4.6 used in place of Corollary 3.3.3. □

Remark 3.4.9: In Theorem 3.4.7, $\text{int dom } F^{\uparrow}(f(x_0);\cdot)$ is nonempty since $\text{int } \mathbb{R}_+^n \neq \phi$ and $F^{\uparrow}(f(x_0);\cdot)$ is isotone on $\mathbb{R}^n$ [Ro4, Theorem 3 and Proposition 4]. Thus if $E := \mathbb{R}^m$, condition (3.4.1) is equivalent to the condition

$$\text{(3.4.12)} \quad \{(r_1,\ldots,r_n) \in \mathbb{R}^n \mid \exists\, y \in E,\ f_i^{\uparrow}(x_0;y) \le r_i\ ,\ i = 1,\ldots,n\} - \text{dom } F^{\uparrow}(f(x_0);\cdot) = \mathbb{R}^n .$$

Since condition (2.3.28) implies that (3.4.12) holds and that $\text{dom } f_i^{\uparrow}(x_0;\cdot)$, $i = 1,\ldots,n$ are in strong general position, the combination of (3.4.1) and the strong general position hypothesis is potentially weaker than (2.3.28). This weakening is achieved at the price of extra continuity requirements on the f_i's.

3.5. Limitations of the generalized subdifferential calculus

In sections 2.3, 2.4, and 3.4, we have obtained "chain rules" for $(F\circ f)^{\uparrow}(x_0;\cdot)$ and $\partial(F\circ f)(x_0)$ with F and f of the following types:

(a) $F\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ strictly l.s.c. at $f(x_0)$, $f\colon \mathbb{R}^n \to \mathbb{R}^n$ strictly differentiable at x_0.

(b) $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ isotone, $f = (f_1,\ldots,f_n)$ with each $f_i\colon E \to \overline{\mathbb{R}}$ l.s.c. and either (2.3.26) or (2.3.27) satisfied.

(In fact, as we have pointed out previously, (a) may be derived as a special case of (b).) Are other more general chain rules possible? For example, is there a chain rule for $F\colon \mathbb{R}^n \to \mathbb{R}$ merely directionally Lipschitzian or for F strictly differentiable and $f_i\colon E \to \overline{\mathbb{R}}$, $i = 1,\ldots,n$ l.s.c.? We investigate these questions in this section.

One way to approach this problem is to examine the proofs of known chain rules and see if similar arguments might be used to prove these results for broader classes of functions. A good place to start is with the calculus of Clarke subgradients for locally Lipschitzian functions ([C14], [Hi1], [Hi2]) developed by Clarke and Hiriart-Urruty. For example, the following is a chain rule given by Hiriart-Urruty in [Hi1, page VIII.14]:

<u>Theorem 3.5.1</u>: Let E be a Banach space, and let $f\colon E \to \mathbb{R}^m$ be locally Lipschitzian near x_0 and $F\colon \mathbb{R}^m \to \mathbb{R}$

be locally Lipschitzian near $f(x_0)$. Then

$$(3.5.1) \qquad \partial(F \circ f)(x_0) \subset \operatorname{conv}\{\sum_{i=1}^{m} u_i x_i^* \mid (u_1,\ldots,u_m) \in \partial F(f(x_0)),\ (x_1^*,\ldots,x_m^*) \in \prod_{i=1}^{m} \partial f_i(x_0)\}.$$

The proof of Theorem 3.5.1 involves two applications of the mean value theorem of Lebourg for locally Lipschitzian functions ([Le1]; [Hi3], [Cl4, Theorem 2.3.7]) and uses the fact that $\partial f(x_0)$ is weak star compact for locally Lipschitzian functions. This line of argument is not possible for broader classes of functions, for two reasons:

(i) $\partial f(x_0)$ is not weak star compact if f is not locally Lipschitzian.

(ii) No analogous mean-value theorem is possible for functions which are not at least directionally Lipschitzian (see [Hi3]).

It is clear, then, that there are definite difficulties involved in an attempt to generalize such a chain rule. We will be able to understand these difficulties more clearly by considering in detail the following special case of Theorem 3.5.1, given in [Cl1, Proposition 10] and [Hi1, Chapter 8]. We include a proof along the lines of the one sketched in [Cl1].

Theorem 3.5.2: Let E be a Banach space, and let $f\colon E \to \mathbb{R}$ be locally Lipschitzian near $x_0 \in E$ and $F\colon \mathbb{R} \to \mathbb{R}$ continuously differentiable near $f(x_0)$. Then for all $y \in E$,

$$(3.5.2) \qquad (F \circ f)^0(x_0;y) = F'(f(x_0))f^0(x_0;y) .$$

Moreover,

$$(3.5.3) \qquad \partial(F \circ f)(x_0) = F'(f(x_0))\ f(x_0) .$$

Proof: For all $y \in E$,

$$(F \circ f)^0(x_0;y) = \limsup_{\substack{x \to x_0 \\ t \downarrow 0}} \frac{(F \circ f)(x+ty) - (F \circ f)(x)}{t}$$

$$= \limsup_{\substack{x \to x_0 \\ t \downarrow 0}} F'(z(t,x,y)) \frac{f(x+ty) - f(x)}{t}$$

for some $z(t,x,y)$ between $f(x)$ and $f(x+ty)$, by the classical mean-value theorem,

$$= \lim_{\substack{x \to x_0 \\ t \downarrow 0}} F'(z(t,x,y)) \limsup_{\substack{x \to x_0 \\ t \downarrow 0}} \frac{f(x+ty) - f(x)}{t}$$

$$= F'(f(x_0))f^0(x_0;y)$$

since f is continuous at x_0 and F is continuously differentiable near $f(x_0)$. Hence (3.5.2) holds. To prove (3.5.3), first assume that $F'(f(x_0)) \geq 0$. Then

$$\begin{aligned}\partial(F\circ f)(x_0) &= \partial((F\circ f)^0(x_0;\cdot))(0) \\ &= \partial(F'(f(x_0))f^0(x_0;\cdot))(0) \quad \text{by (3.5.2)} \\ &= F'(f(x_0))\partial f^0(x_0;\cdot)(0) \\ &\quad \text{since } F'(f(x_0)) \geq 0 \\ &= F'(f(x_0))\partial f(x_0) \quad .\end{aligned}$$

Now $-\partial(F\circ f)(x_0) = \partial(-(F\circ f))(x_0)$, so if $F'(f(x_0)) < 0$, we have

$$\begin{aligned}-\partial(F\circ f)(x_0) &= \partial(-(F\circ f))(x_0) \\ &= \partial((-F)\circ f)(x_0) \\ &= -F'(f(x_0))\partial f(x_0) \\ &\quad \text{since } -F'(f(x_0)) > 0 \quad .\end{aligned}$$

Multiplying both sides of the equation above by -1, we obtain (3.5.3) for this case also. □

The proof of (3.5.2) relies on the following fact, whose proof we leave to the reader.

<u>Lemma 3.5.3</u>: Let E be a l.c.s., and let $g: E \to \overline{\mathbb{R}}$ and $h: E \to \overline{\mathbb{R}}$ be extended real valued functions. Suppose $\lim_{x\to x_0} h(x) = A \in \mathbb{R}$ and $\limsup_{x\to x_0} g(x) = B \in \overline{\mathbb{R}}$. Then

(i) $\limsup_{x\to x_0} h(x)g(x) = AB$ if $0 < A < +\infty$

(ii) $\limsup_{x \to x_0} h(x)g(x) = AB$ if $-\infty < A \leq 0$

and g is bounded in a neighbourhood of x_0.

Lemma 3.5.3 indicates that in order to prove (3.5.2) for more general f, we will have to at least assume that $F'(f(x_0)) > 0$. The assumption that $F'(f(x_0)) \geq 0$ will be needed to prove (3.5.3) for more general f, since the relationship $\partial(-f)(x_0) = -\partial f(x_0)$ does not necessarily hold outside the locally Lipschitzian case.

In this section, we prove what seems to be the best possible generalization of Theorem 3.5.2. We begin by introducing some notation and proving a technical lemma.

Definition 3.5.4 [Ro3], [Ro4]: Let E, E_1 be l.c.s., and let $g: E \times E_1 \to \bar{\mathbb{R}}$ be an extended real valued function. For $(x_0, y_0) \in E \times E_1$, define

$$\limsup_{x \to x_0} \inf_{y \to y_0} g(x,y) = \sup_{Y \in N(y_0)} \inf_{X \in N(x_0)} \sup_{x \in X} \inf_{y \in Y} g(x,y) .$$

We have previously used the following lemma in the proof of Lemma 2.3.19.

Lemma 3.5.5: Let E, E_1 be l.c.s., and let $g: E \times E_1 \to \bar{\mathbb{R}}$ and $h: E \times E_1 \to \bar{\mathbb{R}}$ be extended real valued functions. Suppose $\lim_{\substack{x \to x_0 \\ y \to y_0}} h(x,y) = A$ and $\limsup_{x \to x_0} \inf_{y \to y_0} g(x,y) = B$. Then

(i) $\limsup_{x\to x_0} \inf_{y\to y_0} h(x,y)g(x,y) = AB$ if $0 < A < +\infty$.

(ii) The same equation holds for $-\infty < A \le 0$ if g is bounded in a neighbourhood of (x_0,y_0) .

<u>Proof of (i)</u>: For any given $\lambda \in (0,A)$, there exist $X_1 \in N(x_0)$, $Y_1 \in N(y_0)$ such that $A - \lambda \le h(x,y) \le A + \lambda$ for all $x \in X_1$, $y \in Y_1$. We first establish that $\limsup_{x\to x_0} \inf_{y\to y_0} h(x,y)g(x,y) \le AB$.

<u>Case 1</u>: $0 \le B < +\infty$. Let $Y \in N(y)$ and $\varepsilon > 0$ be given. Choose $\lambda \in (0,A)$ small enough so that $(A+B)\lambda + \lambda^2 < \varepsilon$ and X_1 , Y_1 as above. Then there exists $X_2 \in N(x_0)$ such that for all $x \in X_2$, there exists $y \in Y \cap Y_1$ with $g(x,y) \le B + \lambda$. Thus for all $x \in X_1 \cap X_2$, there exists $y \in Y \cap Y_1$ with $h(x,y)g(x,y) \le (A+\lambda)(B+\lambda) = AB + (A+B)\lambda + \lambda^2 \le AB + \varepsilon$. Since ε was arbitrary, $\limsup_{x\to x_0} \inf_{y\to y_0} h(x,y)g(x,y) \le AB$.

<u>Case 2</u>: $-\infty < B < 0$. Let $Y \in N(y)$ and $\varepsilon > 0$ be given. Let $\lambda \in (0,\min(A,-B))$ be such that $(A-B)\lambda \le \varepsilon$, and choose X_1 , Y_1 as above. There exists $X_2 \in N(x_0)$ such that for each $x \in X_2$, there exists $y \in Y \cap Y_1$ with $g(x,y) \le B + \lambda$. So for each $x \in X_1 \cap X_2$, there exists $y \in Y \cap Y_1$ with $h(x,y)g(x,y) \le (A-\lambda)(B+\lambda) = AB + (A-B)\lambda - \lambda^2 \le AB + \varepsilon$. Since ε was arbitrary, $\limsup_{x\to x_0} \inf_{y\to y_0} h(x,y)g(x,y) \le AB$.

Case 3: $B = -\infty$. Let $m < 0$ and $Y \in N(y)$ be given, and choose $\lambda \in (0,A)$ and X_1, Y_1 as above. There exists $X_2 \in N(x_0)$ such that for all $x \in X_2$, there exists $y \in Y \cap Y_1$ with $g(x,y) \le \frac{m}{A-\lambda}$. Then for all $x \in X_1 \cap X_2$, there exists $y \in Y \cap Y_1$ such that $h(x,y)g(x,y) \le (A-\lambda)\frac{m}{(A-\lambda)} = m$. Since m was arbitrary, $\limsup_{x\to x_0} \inf_{y\to y_0} = -\infty = AB$.

Case 4: $B = +\infty$. Choose $\lambda \in (0,A)$, and let $X \in N(x_0)$ and $M > 0$ be given. Choose X_1, Y_1 as above. There exists $Y \in N(y_0)$ such that for some $x \in X \cap X_1$, $g(x,y) \ge M$ for all $y \in Y$. So there exists $x \in X \cap X_1$ such that for all $y \in Y \cap Y_1$, $h(x,y)g(x,y) \ge M(A-\lambda)$. Since M and X were arbitrary, $\limsup_{x\to x_0} \inf_{y\to y_0} g(x,y)h(x,y) = +\infty = AB$. We have now established that $\limsup_{x\to x_0} \inf_{y\to y_0} g(x,y)h(x,y) \le AB$. It follows, then, that

$$
\begin{aligned}
B &= \limsup_{x\to x_0} \inf_{y\to y_0} g(x,y) \\
&\le \lim_{x\to x_0} \frac{1}{h(x_0)} \limsup_{x\to x_0} \inf_{y\to y_0} g(x,y)h(x,y) \\
&= \frac{1}{A} \limsup_{x\to x_0} \inf_{y\to y_0} g(x,y)h(x,y) \quad .
\end{aligned}
$$

Hence $\limsup_{x\to x_0} \inf_{y\to y_0} g(x,y)h(x,y) \ge AB$, and (i) holds.

We leave the proof of (ii), which is similar, to the reader. We make the important observation that if $A \le 0$, local boundedness of g is required in both the proof of

$\limsup_{x\to x_0} \inf_{y\to y_0} h(x,y)g(x,y) \le AB$ and that of

$\limsup_{x\to x_0} \inf_{y\to y_0} h(x,y)g(x,y) \ge AB$. □

We now prove what seems to be the best possible version of Theorem 3.5.2 with f not necessarily locally Lipschitzian.

<u>Theorem 3.5.6</u>: Let E be a l.c.s. and let $f: E \to \bar{\mathbb{R}}$ be continuous at $x_0 \in E$, and let $F: \mathbb{R} \to \mathbb{R}$ be continuously differentiable near $f(x_0)$. Assume $F'(f(x_0)) > 0$. Then for all $y \in E$,

$$(3.5.4) \qquad (F\circ f)^{\uparrow}(x_0;y) = F'(f(x_0))f^{\uparrow}(x_0;y) .$$

Moreover,

$$(3.5.5) \qquad \partial(F\circ f)(x_0) = F'(f(x_0))\partial f(x_0) .$$

<u>Proof</u>: The proof of (3.5.4) parallels that of (3.5.2) with Lemma 3.5.5 used in place of Lemma 3.5.3. Now if $f^{\uparrow}(x_0;0) = -\infty$, $(F\circ f)^{\uparrow}(x_0;0)$ will also $= -\infty$ by (3.5.4), and both sides of (3.5.5) will be empty. Otherwise, $f^{\uparrow}(x_0;\cdot)$ is proper, $\partial f(x_0)$ is nonempty, and the proof of (3.5.5) parallels that of (3.5.3). □

Remark 3.5.7: (a) If $F'(f(x_0)) < 0$, neither inclusion of (3.5.5) will hold in general. For example, define $F: \mathbb{R} \to \mathbb{R}$ by $F(x) := -x$ and $f: \mathbb{R} \to \mathbb{R}$ by $f(x) := -|x|^{1/2}$. Then $\partial(F \circ f)(0) = \mathbb{R}$, but $\partial f(0) = \phi$. On the other hand, if $F(x) := -x$ but $f(x) := |x|^{1/2}$, $\partial(F \circ f)(0) = \phi$ while $F'(f(x_0))\partial f(x_0) = \mathbb{R}$.

(b) There are also counterexamples if $F'(f(x_0)) = 0$. If we define $F: \mathbb{R} \to \mathbb{R}$ by $F(x) := x^2$ and $f: \mathbb{R} \to \mathbb{R}$ by $f(x) = -|x|^{1/2}$, $\partial(F \circ f)(0) = [-1,1]$ while $\partial f(0) = \phi$. On the other hand, if $F(x) := -x^{4/3}$ and $f(x) := |x|^{1/2}$, then $\partial(F \circ f)(0) = \phi$ while $F'(f(0)) = 0$ and $\partial f(0) = \mathbb{R}$.

We now have at least a partial answer to our original questions. Even if F is continuously differentiable and f is continuous, the requirement that F be locally isotone (as it is if $F'(f(x_0)) > 0$) or that f be locally Lipschitzian seems to be needed in any chain rule for $(F \circ f)^{\uparrow}(x_0; \ )$ or $\partial(F \circ f)$. This requirement is not surprising in view of the inherent "one-sidedness" of our epigraph-centered approach.

CHAPTER IV

Lim inf Formulae and Approximate Subdifferentials

4.1. Lim infs of tangent cones

In this section we survey some significant recent results concerning the relationship between $T_C(x_0)$ and various tangent cones $R_C(x)$ at points x near x_0.

Definition 4.1.1: Let E, E_1 be l.c.s., and let $C \subset E$ and $x_0 \in C$. Let $H: E \to 2^{E_1}$ be a relation. Define

$$\liminf_{\substack{x \to x_0 \\ C}} H(x) := \bigcap_{V \in N(0)} \bigcup_{X \in N(x_0)} \bigcap_{x \in X \cap C} H(x) + V \tag{4.1.1}$$

and

$$\limsup_{\substack{x \to x_0 \\ C}} H(x) := \bigcap_{V \in N(0)} \bigcap_{X \in N(x_0)} \bigcup_{x \in X \cap C} H(x) + V \tag{4.1.2}$$

Observe that if $H_1, H_2: E \to 2^{E_1}$ are such that $\mathrm{Gr}\, H_1 \subset \mathrm{Gr}\, H_2$, then $\liminf_{\substack{x \to x_0 \\ C}} H_1(x) \subset \liminf_{\substack{x \to x_0 \\ C}} H_2(x)$ and $\limsup_{\substack{x \to x_0 \\ C}} H_1(x) \subset \limsup_{\substack{x \to x_0 \\ C}} H_2(x)$.

Penot has obtained the following result relating $T_C(x_0)$, $\liminf\limits_{x \xrightarrow{C} x_0} K_C(x)$, and $\liminf\limits_{x \xrightarrow{C} x_0} P_C(x)$ in finite dimensions.

<u>Theorem 4.1.2</u> [Pe1]: Let $C \subset \mathbb{R}^n$ be closed near $x_0 \in C$. Then

$$(4.1.3) \qquad \liminf_{x \xrightarrow{C} x_0} K_C(x) = \liminf_{x \xrightarrow{C} x_0} P_C(x) = T_C(x_0) .$$

Neither inclusion of (4.1.3) remains true in infinite dimensions without additional hypotheses (see [Pe3], [Tr1], [Bo6], [Bo8] for examples).

We next discuss the infinite-dimensional extensions of Theorem 4.1.2 which have been proved by Treiman, Penot, Borwein, and Strojwas.

<u>Theorem 4.1.3</u> [Tr1]: Let E be a l.c.s., and let $x_0 \in C \subset E$. Then

$$(4.1.4) \qquad \liminf_{x \xrightarrow{C} x_0} K_C(x) \subset T_C^*(x_0) .$$

If E is a Banach space and C is closed near x_0, then

$$(4.1.5) \qquad \liminf_{x \xrightarrow{C} x_0} K_C(x) \subset T_C(x_0) .$$

If in addition C is epi-Lipschitzian at x_0, then equality holds in (4.1.5).

Treiman obtains (4.1.5) as a consequence of (4.1.4) and the fact that $T_C^*(x_0) = T_C(x_0)$ for locally closed subsets of a Banach space (Theorem 1.7.1). Penot [Pe3] gives a simpler proof of (4.1.5) via the "drop theorem" of Danes [Da1].

Of course (4.1.5) still holds if the contingent cone is replaced by any tangency operator which is always contained in the contingent cone. If C is epi-Lipschitzian, equality in (4.1.5) will still be achieved if K is replaced by k, as we now demonstrate.

<u>Proposition 4.1.4</u>: Let E be a Banach space, and let $C \subset E$ be closed near $x_0 \in C$. Then

$$\liminf_{\substack{x \to x_0 \\ C}} k_C(x) \subset T_C(x_0) . \tag{4.1.6}$$

Equality holds in (4.1.6) if C is in addition epi-Lipschitzian at x_0.

<u>Proof</u>: As was mentioned above, (4.1.6) is an immediate consequence of (4.1.5). Suppose C is epi-Lipschitzian at x_0, and let $y \in H_C(x_0)$. Then there exist $X \in N(x_0)$ and $\lambda > 0$ such that for all $x \in X \cap C$ and for all $t \in (0,\lambda)$, $x + ty \in C$. Thus $y \in k_C(x_0)$, and it follows that

$$\begin{aligned} H_C(x_0) &\subset \bigcup_{X \in N(x_0)} \bigcap_{x \in X} k_C(x) \\ &\subset \liminf_{\substack{x \to x_0 \\ C}} k_C(x) \end{aligned}$$

Now $\liminf_{x \to_C x_0} k_C(x)$ is closed, and since C is epi-Lipschitzian at x_0, $T_C(x_0) = \mathrm{cl}\, H_C(x_0)$. Therefore

$T_C(x_0) \subset \liminf_{x \to_C x_0} k_C(x)$. □

Notice that our proof of Proposition 4.1.4 provides an easier proof of the last assertion of Theorem 4.1.3 than that given in [Tr1].

The following extension of Theorem 4.1.3 has been established by Borwein.

<u>Theorem 4.1.5</u> [Bo6]: Let E be a reflexive Banach space, and let $C \subset E$ be closed near $x_0 \in C$. Then

$$\liminf_{x \to_C x_0} P_C(x) \subset T_C(x_0) \ . \tag{4.1.7}$$

Equality holds in (4.1.7) if in addition C is epi-Lipschitzian at x_0.

For the weak contingent and pseudotangent cones even more satisfactory results, with no epi-Lipschitzian requirement, are possible in a reflexive Banach space.

<u>Theorem 4.1.6</u> (Strojwas and Borwein): Let E be a reflexive Banach space, and let $C \subset E$ be closed near $x_0 \in C$. Then

$$\liminf_{x \to_C x_0} K_C^W(x) = T_C(x_0) \tag{4.1.8}$$

and

$$\liminf_{\substack{x \to x_0 \\ C}} P_C^w(x) = T_C(x_0) \tag{4.1.9}$$

where "w" denotes the weak topology.

Borwein has also shown that (4.1.8) characterizes reflexive space and has exhibited counterexamples to (4.1.7), (4.1.8), and (4.1.9) for closed sets in nonreflexive Banach spaces. All these results will appear in future work of Borwein and Strojwas.

These "lim inf" results have already proven to be a powerful tool in nonsmooth analysis. We now describe some of their applications.

(i) Tangent cone inclusions: Watkins [W1] has employed Theorem 4.1.2 in proving inclusion (2.3.20) of Proposition 2.3.11.

(ii) Subdifferential calculus: As was mentioned in section 2.1, Ioffe [I2] has used Theorem 4.1.2 to derive Corollaries 2.3.5 and 2.3.7 and Proposition 2.3.14 from corresponding results for approximate subdifferentials.

(iii) R-quasiconvexity and convexity: Sufficient conditions for optimality in differentiable non-linear programming ([Gu1], [Bo3]) often involve the assumption of R-quasiconvexity, defined below:

<u>Definition 4.1.7</u> [Bo6]: Let E be a l.c.s. and R a tangent cone. The set $C \subset E$ is <u>R-quasiconvex at x_0</u> $\in C$ if

$$(4.1.10) \qquad C - x_0 \subset R(C,x_0) .$$

If C is R-quasiconvex at all of its points, it is said to be R-quasiconvex.

If R is the pseudotangent cone, the term "pseudoconvex" is used in place of "P-quasiconvex" ([Bo3], [Gu1]).

<u>Definition 4.1.8</u>: Let E be a l.c.s.

(a) For $x,y \in E$, define $[x,y] := \mathrm{conv}\{x,y\}$.

(b) A set $C \subset E$ is said to be starshaped if the set

$$(4.1.11) \qquad \text{star } C := \{x \in C \mid [x,y] \subseteq C \text{ for all } y \in C\}$$

is nonempty.

One consequence of Theorems 4.1.3, 4.1.5, and 4.1.6 is the following result. For the proof, see [Bo6].

<u>Proposition 4.1.9</u>: Let C be a closed subset of a Banach space E . Then

$$(4.1.12) \qquad \text{star } C = \bigcap_{x \in C} (T_C(x) + x) = \bigcap_{x \in C} (K_C(x) + x) = \bigcap_{x \in C} (K_C(x) + x) .$$

If E is reflexive, then in addition

$$\text{(4.1.13)} \qquad \operatorname{star} C = \bigcap_{x \in C} (K_C^w(x) + x) = \bigcap_{x \in C} (P_C^w(x) + x) = \bigcap_{x \in C} (P_C(x) + x) .$$

An immediate corollary of Proposition 4.1.9 is a characterization of R-quasiconvex sets in Banach spaces (or reflexive Banach spaces) for various tangent cones R.

Corollary 4.1.10: A closed subset C of a Banach space E is T-quasiconvex (K-quasiconvex, k-quasiconvex) if and only if it is convex. If E is reflexive, C is pseudoconvex (P^w-quasiconvex, K^w-quasiconvex) if and only if it is convex.

Proof: The "if" part follows from the fact that all of these tangent cones have property (1) of chapter 1. Suppose C is T-quasiconvex. Then

$$C \subset \bigcap_{x \in C} (T_C(x) + x) = \operatorname{star} C \text{ by (4.1.12)},$$

and hence C is convex. The proof of "R-quasiconvexity implies convexity" for the other tangent cones is exactly the same. □

(iv) Properness of tangent cones and surjectivity of relations:

Definition 4.1.11: Let E be a l.c.s. and R a tangency operator. We say that $x_0 \in C \subseteq E$ is a R-proper point of C if $R(C,x_0) \neq E$, and that x_0 is a R-improper point of C if $R(C,x_0) = E$.

Another consequence of Theorem 4.1.3 and 4.1.6 is the following result on properness of tangent cones.

Proposition 4.1.12 [Bo6]: (a) Let E be a Banach space and R a tangency operator such that $R(D,x_0) \subset K_D(x_0)$ for all $(D,x_0) \in 2^E \times E$. If C is a closed subset of E, the R-proper points of C are dense in the boundary of C.
(b) The same conclusion is true if E is reflexive and $R(D,x_0) \subset P_D^W(x_0)$ for all $(D,x_0) \in 2^E \times E$.

Proof: If the assertion in (a) is true for the contingent cone, it certainly holds for any smaller tangency operator. The proof for $R := K$ is given in [Bo6]. The proof of (b) is similar. □

As an immediate corollary of either Proposition 4.1.9 or Proposition 4.1.12, we can deduce a result about tangency operators which are everywhere improper.

<u>Corollary 4.1.13</u>: (a) Let E be a Banach space, and let R be a tangency operator satisfying $R(D,x_0) \subset K_D(x_0)$ for all $(D,x_0) \in 2^E \times E$. Then any closed set $C \subset E$ which is R-improper at each of its points must be equal to E.
(b) The same conclusion holds if E is reflexive and R satisfies $R(D,x_0) \subset P_D^W(x_0)$ for all $(D,x_0) \in 2^E \times E$.

Part (a) of Corollary 4.1.13 is also proved in [Gal] via the drop theorem. It is then used to derive a number of conditions sufficient for the surjectivity of a relation. In the case in which E is reflexive, part (b) of Corollary 4.1.13 may be used to derive extensions of the results of [Gal].

4.2. Applications to the Study of Directional Derivatives and Subgradients

We now explore some of the implications of Theorems 4.1.3 and 4.1.5 in the case in which C is the epigraph of an extended real valued function. In particular, we extend to a reflexive Banach space setting some finite-dimensional results given in [I2].

Let E be a Banach space, and let $f: E \to \bar{\mathbb{R}}$ be a function which is finite and strictly l.s.c. at x_0. If $C := \operatorname{epi} f$ in Theorem 4.1.3, (4.1.5) becomes

$$(4.2.1) \qquad \liminf_{(x,r) \xrightarrow[\operatorname{epi} f]{} (x_0,f(x_0))} K_{\operatorname{epi} f}(x,r) \subset T_{\operatorname{epi} f}(x_0,f(x_0)).$$

In fact, since f is strictly l.s.c. at x_0, (4.2.1) can be written (see [I2, proof of Proposition 6]) more simply as

$$(4.2.2) \qquad \liminf_{(x,f(x))\to(x_0,f(x_0))} K_{\text{epi } f}(x,f(x)) \subset T_{\text{epi } f}(x_0,f(x_0)).$$

By results of section 1.5, an equivalent way to write (4.2.2) is

$$(4.2.3) \qquad \liminf_{x \to_f x_0} \text{epi } f_+(x;\cdot) \subset \text{epi } f^{\dagger}(x_0;\cdot) ,$$

where "$x \to_f x_0$" is a convenient shorthand notation for " $x \to x_0$ with $f(x) \to f(x_0)$".

If E is a reflexive space, we have in addition that

$$(4.2.4) \qquad \liminf_{x \to_f x_0} \text{epi } \overline{\text{co}}\, f_+(x;\cdot) \subset \text{epi } f^{\dagger}(x_0;\cdot)$$

by Theorem 4.1.5. Equality holds in (4.2.3) and (4.2.4) if f is directionally Lipschitzian at x_0 or if E is finite-dimensional. Of course if f is continuous at x_0, (4.2.3) and (4.2.4) can be written with "$x \to_f x_0$" replaced by "$x \to x_0$".

We next demonstrate that inclusions (4.2.3) and (4.2.4) translate into more explicit relationships between $f^{\dagger}(x_0;\cdot)$ and $f'_+(x_0;\cdot)$.

Proposition 4.2.1 (cf. [I2, Proposition 6): Let E be a Banach space, and let f be finite and strictly l.s.c. at x_0. Then for all $h_0 \in E$,

$$\limsup_{x \to_f x_0} \inf_{h \to h_0} f_+(x;h) \geq f^\dagger(x_0;h_0) . \tag{4.2.5}$$

If E is reflexive, we have in addition that

$$\limsup_{x \to_f x_0} \inf_{h \to h_0} \overline{\mathrm{co}}\, f_+(x;h) \geq f^\dagger(x_0;h_0) . \tag{4.2.6}$$

Equality holds in (4.2.5) and (4.2.6) if E is finite-dimensional or if f is directionally Lipschitzian at x_0.

Proof: We prove (4.2.5) and leave the proof of (4.2.6), which is analogous, to the reader. If $f^\dagger(x_0;h_0) = -\infty$, (4.2.5) is automatically satisfied. Suppose $f^\dagger(x_0;h_0) = +\infty$, and let $r \in \mathbb{R}$. Then $f^\dagger(x_0;h_0) > r$, and by (4.2.3), $(h_0,r) \notin \liminf_{x \to_f x_0} \operatorname{epi} f_+(x;\cdot)$. Thus there exists $\delta > 0$ such that for all $X \in \mathcal{N}(x_0)$ and $\lambda > 0$, there exists $x \in X$ with $f(x) \in B_\lambda(f(x_0))$ such that $f_+(x;h) \geq r + \delta$ for all $h \in B_\delta(h_0) := \{h \in E \,\big|\, \|h-h_0\| < \delta\}$. Since r was arbitrarily chosen,

$$\limsup_{x \to_f x_0} \inf_{h \to h_0} f_+(x;h) = +\infty .$$

Now let $\varepsilon > 0$ be given, and suppose $r \in \mathbb{R}$ is such that

$f^{\uparrow}(x_0;h_0) = r + \varepsilon$. Again by (4.2.3) there exists $\delta > 0$ such that for all $X \in N(x_0)$ and $\lambda > 0$, there exists $x \in X$ with $f(x) \in B_\lambda(f(x_0))$ such that for all $h \in B_\delta(h_0)$,

$$\begin{aligned} f_+(x;h) &\geq r + \delta \\ &= f^{\uparrow}(x_0;h_0) - \varepsilon + \delta \\ &\geq f^{\uparrow}(x_0;h_0) - \varepsilon . \end{aligned}$$

Since $\varepsilon > 0$ was arbitrarily chosen,

$$\limsup_{x \to_f x_0} \inf_{h \to h_0} f_+(x;h) \geq f^{\uparrow}(x_0;h_0) .$$

Hence (4.2.5) holds. Now suppose that either E is finite-dimensional or f is directionally Lipschitzian at x_0. Then equality holds in (4.2.3). Suppose $f^{\uparrow}(x_0;h_0) \leq r$, and let $\delta > 0$ be given. Then there exist $X \in N(x_0)$ and $\lambda > 0$ such that for all $x \in X$ with $f(x) \in B_\lambda(f(x_0))$, there exists $h \in B_\delta(h_0)$ such that $f_+(x;h) \leq r + \delta$. Since δ is arbitrary, $\limsup_{x \to_f x_0} \inf_{h \to h_0} f_+(x;h) \leq r$, and it follows that $\limsup_{x \to_f x_0} \inf_{h \to h_0} f_+(x;h) \leq f^{\uparrow}(x_0;h_0)$. Therefore we have equality in (4.2.5). □

In the special case in which $f: E \to \overline{\mathbb{R}}$ is locally Lipschitzian near x_0, $f_+(x_0;\cdot) = f^A(x_0;\cdot)$ and $f^{\uparrow}(x_0;\cdot) = f^0(x_0;\cdot)$ by Corollary 1.5.19. In this case, equality holds in (4.2.3), and (4.2.3) becomes

$$\liminf_{x \to x_0} \text{epi } f^A(x;\cdot) = \text{epi } f^0(x_0;\cdot) \ . \tag{4.2.7}$$

If E is reflexive, we also have by (4.2.4) that

$$\liminf_{x \to x_0} \text{epi } \overline{\text{co}}\ f^A(x;\cdot) = \text{epi } f^0(x_0;\cdot) \ . \tag{4.2.8}$$

We may apply (4.2.7) and (4.2.8) as in Proposition 4.2.1 to prove the following result:

<u>Proposition 4.2.2</u>: Let E be a Banach space, and let $f: E \to \overline{\mathbb{R}}$ be locally Lipschitzian near $x_0 \in E$. Then for all $h_0 \in E$,

$$\limsup_{x \to x_0} f^A(x;h_0) = f^0(x_0;h_0) \ . \tag{4.2.9}$$

Moreover, if E is reflexive, then

$$\limsup_{x \to x_0} \overline{\text{co}}\ f^A(x;h_0) = f^0(x_0;h_0) \ . \tag{4.2.10}$$

<u>Proof</u>: Let $\varepsilon > 0$ and $h_0 \in E$ be given, and suppose $f^0(x_0;h_0) = r + \varepsilon$. By (4.2.7), $(h_0,r_0) \notin \liminf_{x \to x_0} \text{epi } f^A(x;\cdot)$. As a result, there exists $\delta > 0$ such that for all $X \in N(x_0)$, there exists $x \in X$ such that $f^A(x;h) \geq r + \delta$ for all $h \in B_\delta(h_0)$. In particular,

$$\begin{aligned} f^A(x;h_0) &\geq r + \delta \\ &= f^0(x_0;h_0) - \varepsilon + \delta \\ &\geq f^0(x_0;h_0) - \varepsilon . \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, it follows that $\limsup_{x \to x_0} f^A(x;h_0) \geq f^0(x_0;h_0)$. Conversely, let $\varepsilon > 0$ and $h_0 \in E$ be given, and suppose $f^0(x_0;h_0) = r$. Let $X_1 \in N(x_0)$ be a neighbourhood on which f is locally Lipschitzian with Lipschitz constant $C > 0$. Choose $\delta \in (0, \frac{\varepsilon}{2C}]$, and choose $X_2 \in N(x_0)$ and $\lambda > 0$ such that $X_2 + (0,\lambda)B_\delta(h_0) \subset X_1$. Now by (4.2.7), there exists $X \in N(x_0)$ with $X \subset X_2$ such that for all $x \in X$,

$$f^A(x;h') \leq f^0(x_0;h_0) + \frac{\varepsilon}{2}$$

for some $h' \in B_\delta(h_0)$. Now for any $x \in X$, $t \in (0,\lambda)$, and $h \in B_\delta(h_0)$,

$$\begin{aligned} \frac{f(x+th_0)-f(x)}{t} &\leq \frac{f(x+th)-f(x)}{t} + C||h_0-h|| \\ &\leq \frac{f(x+th)-f(x)}{t} + \frac{\varepsilon}{2} . \end{aligned}$$

Thus for all $x \in X$ and some $h' \in B_\delta(h_0)$,

$$\begin{aligned} f^A(x;h_0) &\leq f^A(x;h') + \frac{\varepsilon}{2} \\ &\leq f^0(x_0;h_0) + \varepsilon . \end{aligned}$$

Since ε was arbitrary, this means that

$$\limsup_{x \to x_0} f^A(x;h_0) \leq f^0(x_0;h_0) \ .$$

Therefore (4.2.9) holds. The proof of (4.2.10) parallels that of (4.2.9). □

<u>Definition 4.2.3</u>: Let E be a l.c.s. with dual space E', and let $f: E \to \overline{\mathbb{R}}$ be finite at $x_0 \in E$. The <u>approximate subdifferential</u> of f at x_0 is the set

(4.2.11) $$\partial_a f(x_0) := w^* \limsup_{x \to_f x_0} \partial^K f(x) \ ,$$

where by "w* lim sup" we mean that the neighbourhoods V in definition (4.1.2) are taken to be weak star neighbourhoods.

This definition is an extension of the one made by Ioffe [I2] for the case in which E is finite dimensional.

Propositions 4.2.1 and 4.2.2 can be used to derive information about the relationship between $\partial_a f(x_0)$ and $\partial^T f(x_0)$ when E is a Banach space, as we now demonstrate. (Analogous finite dimensional results are given in [I2]).

<u>Proposition 4.2.4</u>: Let E be a Banach space, and let $f: E \to \overline{\mathbb{R}}$ be locally Lipschitzian near $x_0 \in E$. Then

(4.2.12) $$\partial_a f(x_0) \subset \partial f(x_0) \ .$$

<u>Proof</u>: Suppose $x^* \in E' \backslash \partial f(x_0)$. Then there exist $h_0 \in E$ and $\varepsilon > 0$ such that

$$\langle x^*, h_0 \rangle - 2\varepsilon = f^0(x_0; h_0) .$$

Define $V := \{x' \in E' \,\big|\, |\langle x', h_0 \rangle| < \varepsilon\}$. Then for all $x' \in x^* + V$,

$$\langle x', h_0 \rangle - \varepsilon \geq f^0(x_0; h_0) .$$

By Proposition 4.2.2,

$$\langle x', h_0 \rangle - \varepsilon \geq \limsup_{x \to_f x_0} f_+(x; h_0)$$

for all $x' \in x^* + V$. Thus there exists $\delta > 0$ such that for all $x \in B_\delta(x_0)$ with $f(x) \in B_\delta(f(x_0))$, $f_+(x; h_0) < \langle x', h_0 \rangle$ for all $x' \in x^* + V$. Therefore $x^* \notin w^* \limsup_{x \to_f x_0} \partial^K f(x)$, and so (4.2.12) holds. □

<u>Corollary 4.2.5</u>: Under the hypotheses of Proposition 4.2.4,

$$(4.2.13) \qquad \overline{\text{conv}}\, \partial_a f(x_0) \subset \partial f(x_0) .$$

<u>Proof</u>: This follows immediately from (4.2.12) since $\partial f(x_0)$ is closed and convex. □

If f is not locally Lipschitzian, (4.2.12) does not seem to be quite obtainable, but we can derive the following result:

Proposition 4.2.6: Let E be a Banach space, and let $f: E \to \overline{\mathbb{R}}$ be finite and strictly l.s.c. at x_0. Assume either E is finite-dimensional or f is directionally Lipschitzian at x_0. Then

$$(4.2.14) \qquad \limsup_{x \to_f x_0} \partial^K f(x) \subset \partial^T f(x_0) .$$

Proof: Suppose $x^* \notin \partial^T f(x_0)$. Then there exists $h_0 \in E$ with $\langle x^*,h_0\rangle - 2\varepsilon = f^{\uparrow}(x_0;h_0)$ for some $\varepsilon > 0$. Define $V := \{x'\in E' \mid \|x'\| \le \varepsilon\}$. Then for all $x' \in x^* + V$, $\langle x',h_0\rangle - \varepsilon \ge f^{\uparrow}(x_0;h_0)$. By Proposition 4.2.1,

$$\langle x',h_0\rangle - \varepsilon \ge \limsup_{x \to_f x_0} \inf_{h \to h_0} f_+(x;h)$$

for all $x' \in x^* + V$. Let $Y \in N(h_0)$ be such that $|\langle x',h\rangle - \langle x',h_0\rangle| < \frac{\varepsilon}{3}$ for all $h \in Y$ and for all $x'' \in x^* + V$. There exists $\delta > 0$ such that for all $x \in B_\delta(x_0)$ with $f(x) \in B_\delta(f(x_0))$, there exists $h \in Y$ with

$$\begin{aligned} f_+(x;h) &\le \langle x',h_0\rangle - \frac{2\varepsilon}{3} \\ &< \langle x',h_0\rangle - \frac{\varepsilon}{3} \\ &< \langle x',h\rangle \quad \text{for all } x' \in x^* + V . \end{aligned}$$

Thus $x' \notin \partial^K f(x)$ for all $x \in B_\delta(x_0)$ and $f(x) \in B_\delta(f(x_0))$. Therefore $x^* \notin \limsup_{x \to_f x_0} \partial^K f(x)$, and (4.2.14) holds. □

Our next proposition generalizes a result of [I2].

<u>Proposition 4.2.7</u>: Let E be a reflexive Banach space, and let f be locally Lipschitzian near $x_0 \in E$. Then

$$(4.2.15) \qquad \overline{\text{conv}}\ \partial_a f(x_0) = \partial f(x_0) .$$

<u>Proof</u>: Let $h \in E$ and $\varepsilon > 0$ be given. Then $f^0(x_0;h) = r + \varepsilon > r$ for some $r \in \mathbb{R}$. By (4.2.10), there exists a sequence $x_n \to x_0$ with

$$\overline{\text{co}}\ f^A(x_n;h) \geq r + \frac{\varepsilon}{2} > r .$$

Now by a standard theorem of convex analysis (see for example [Hol, section 14]), $\sup_{T \in \partial(\overline{\text{co}}\ f^A(x_n;\cdot))(0)} \langle T,h\rangle = \overline{\text{co}}\ f^A(x_n;h)$.

Thus there exists a sequence $T_n \in \partial(\overline{\text{co}}\ f^A(x_n;\cdot))(0) \subset \partial^K f(x_n)$ with $\langle T_n,h\rangle > r$. Since f is locally Lipschitzian near x_0, the sets $\partial(\overline{\text{co}}\ f^A(x_n;\cdot))(0)$ are equi-w* compact, so we may assume that T_n converges in the weak star topology to some $T_0 \in E'$. Hence $\langle T_0,h\rangle \geq r$, and so

$$T_0 \in \partial_a f(x_0) .$$

Therefore $\sup_{T \in \partial_a f(x_0)} T(h) \geq f^0(x_0;h)$, and it follows that

(4.2.16) $$\overline{\text{conv}}\ \partial_a f(x_0) \supset \partial f(x_0) .$$

We have already proven the opposite inclusion in Corollary 4.2.5. □

<u>Corollary 4.2.8</u>: Let E be a reflexive Banach space, and let $f\colon E \to \overline{\mathbb{R}}$ be locally Lipschitzian near $x_0 \in E$. Then $\partial^K f(x) \neq \phi$ densely near x_0.

<u>Proof</u>: Since f is locally Lipschitzian near x_0, $\partial f(x_0) \neq \phi$. By (4.2.16), $\partial_a f(x_0)$ must be nonempty also. Suppose $x^* \in \partial_a f(x_0)$. Let $X \in N(x_0)$ be given, along with a weak star neighbourhood V of 0 in E'. Then there exists $x' \in x^* + V$ and $x \in X$ with $x' \in \partial^K f(x)$. Since X was arbitrary, $\partial^K f(x) \neq \phi$ densely near x_0. □

<u>Example 4.2.9</u>: The following example, given in [I2], shows that "$\overline{\text{conv}}$" is needed in (4.2.15). Define $f\colon \mathbb{R} \to \mathbb{R}$ by $f(x) := -|x|$, and let $x_0 := 0$. Then $T_{\text{epi } f}(0,0) = \text{epi}(|x|)$, so $\partial f(0) = [-1,1]$, while $K_{\text{epi } f}(0,0) = \text{epi } f$ and

$$K_{\text{epi } f}(x,-|x|) = \begin{cases} \{(x,y) \in \mathbb{R}^2 \mid y \geq x\} & \text{if } x < 0 \\ \{(x,y) \in \mathbb{R}^2 \mid y \geq -x\} & \text{if } x > 0 . \end{cases}$$

Thus $\partial^K f(0) = \phi$, $\partial^K f(x) = \{1\}$ if $x < 0$, and $\partial^K f(x) = \{-1\}$ if $x > 0$. Hence $\partial_a f(0) = \{-1,1\}$, showing that even for Lipschitz functions of one variable, $\partial_a f(x_0)$

is not necessarily a convex set and "$\overline{\text{conv}}$" is needed in (4.2.15).

Example 4.2.10: We now give two functions which are not locally Lipschitzian and for which (4.2.15) is not true.

(a) Define $f\colon \mathbb{R} \to \overline{\mathbb{R}}$ by

$$f(x) := \begin{cases} x & \text{if } x \leq 0 \\ \sqrt{x} & \text{if } x > 0 \end{cases},$$

and let $x_0 := 0$. Then $T_{\text{epi } f}(0,0) = \{(x,y) \in \mathbb{R}^2 \mid x \leq 0, y \geq 0\}$, so $\partial f(0) = [0,\infty)$. On the other hand, $\partial^K f(0) = [1,\infty)$, while

$$\partial^K f(x) = \begin{cases} \{1\} & \text{if } x < 0 \\ \{\frac{1}{2\sqrt{x}}\} & \text{if } x > 0 \end{cases}.$$

Thus $\partial_a f(0) = [1,\infty)$, and (4.2.15) does not hold.

(b) Define $f\colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) := \begin{cases} 0 & \text{if } x \leq 0 \\ -\sqrt{x} & \text{if } x > 0 \end{cases},$$

and let $x_0 := 0$. Then $T_{\text{epi } f}(0,0) = \{(x,y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0\}$, so $\partial f(0) = (-\infty,0]$, while

$$\partial^K f(x) = \begin{cases} \{0\} & \text{if } x < 0 \\ \phi & \text{if } x = 0 \\ \{-\frac{1}{2\sqrt{x}}\} & \text{if } x > 0 \end{cases} .$$

Therefore $\partial_a f(0) = \{0\}$, and (4.2.15) does not hold. This may be the example intended in [I2, page 398]. For the example given there, $f\colon \mathbb{R} \to \mathbb{R}$ defined by

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

it is correctly stated that $\partial f(0) = (-\infty, 0]$, but $\partial_a f(0)$ is incorrectly given as $\{0\}$. In fact,

$$\partial^K f(x) = \begin{cases} \{-\frac{1}{2\sqrt{-x}}\} & \text{if } x < 0 \\ (-\infty, 0] & \text{if } x = 0 \\ \{0\} & \text{if } x > 0 \end{cases} ,$$

so $\partial_a f(0) = (-\infty, 0]$.

By applying Theorems 4.1.4 and 4.1.6, a development parallel to that given here may be carried out, replacing $f_+(x_0;\cdot)$ by the appropriate directional derivatives. We also mention in passing that Theorem 4.1.6 and some related recent results of Borwein can be used to prove significant new generic subdifferentiability results, as well as extend the theory of [Ro7] to a reflexive Banach space setting.

Details will appear in future work by Borwein.

4.3. A product rule for approximate subdifferentials

In [I2], Ioffe develops a calculus of approximate sub-differentials for extended real valued functions with finite-dimensional domain. Even though approximate sub-differentials are not in general convex, they do admit an extensive calculus, as Ioffe demonstrates. In particular, analogues of Corollary 2.3.5 and Proposition 2.3.14 exist for approximate subdifferentials under the same hypotheses. To illustrate the method employed in [I2], we apply Ioffe's method in this section to prove a product rule for continuous, locally positive functions with finite-dimensional domain. The first step is closely patterned after [I2, Lemma 3].

<u>Lemma 4.3.1</u>: Let $f_1, f_2: \mathbb{R}^m \to \overline{\mathbb{R}}$ be continuous and positive at $z \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$. Then for any $\delta > 0$,

$$(4.3.1) \qquad \partial^K(f_1 f_2)(z) \subset \bigcup_{x_i \in U(f_i, z, \delta)} \left(f_1(x_1)\partial^K f_2(x_2) + f_2(x_2)\partial^K f_1(x_1) + B_\delta(0)\right)$$

where

$$U(f,z,\delta) := \{x \in \mathbb{R}^m \,|\, \|x-z\| < \delta, |f(x)-f(z)| < \delta\} \quad .$$

<u>Proof</u>: If $\partial^K(f_1 f_2)(z) = \phi$, (4.3.1) holds automatically. Assume, then, that $x^* \in \partial(f_1 f_2)(z)$. Replacing $(f_1 f_2)(x)$

by $(f_1 f_2)(x) - \langle x^*, x\rangle$ if necessary, we can assume without loss of generality that $x^* = 0$. We are then required to show that

$$\bigcup_{x_i \in U(f_i,z,\delta)} (f_1(x_1)\partial^K f_2(x_2) + f_2(x_2)\partial^K f_1(x_1)) \cap B_\delta(0) \neq \phi .$$

Now by [I2, Lemma 1], the function $g_\delta(x) := f_1(x)f_2(x) + \delta||x-z||$ attains a strict local minimum at z. Choose $\varepsilon \in (0,\delta)$ such that $g_\delta(x_i) > g_\delta(z)$, $f_i(x_i) \geq f_i(z) - \alpha$ with $\alpha^2 + \alpha(f_1(x_1)+f_2(x_2)) \leq 1$ whenever $x_i \neq z$, $||x_i - z|| \leq \varepsilon$. This is possible since the f_i's are continuous at z. Next consider

$$P_r(u,v,x) := f_1(u)f_2(v) + \frac{r}{2}(||u-x||^2 + ||v-x||^2) + \delta||x-z|| .$$

Let u_r, v_r, x_r be such that $||u_r - z|| \leq \varepsilon$, $||v_r - z|| \leq \varepsilon$, $||x_r - z|| \leq \varepsilon$, and

$$P_r(u_r,v_r,x_r) = \min\{P_r(u,v,x) \,\big|\, ||u-z|| \leq \varepsilon,\ ||v-z|| \leq \varepsilon,\ ||x-z|| \leq \varepsilon\} .$$

Then

$$\begin{aligned} f_1(z)f_2(z) &+ \frac{r}{2}(||u_r - x_r||^2 + ||v_r - x_r||)^2 + \delta||x_r - z|| \\ &\leq (f_1(u_r)+\alpha)(f_2(v_r)+\alpha) + \frac{r}{2}(||u_r-x_r||^2 + ||v_r - x_r||^2) \\ &\qquad + \delta||x_r - z|| \end{aligned}$$

$$\leq f_1(u_r)f_2(v_r) + 1 + \frac{r}{2}(||u_r-x_r||^2 + ||v_r-x_r||^2) + \delta||x_r-z|| .$$

$$= P_r(u_r,v_r,x_r) + 1$$

$$\leq P_r(z,z,z) + 1 = f_1(z)f_2(z) + 1 .$$

Thus $\frac{r}{2}(||u_r-x_r||^2+||v_r-x_r||^2) + \delta||x_r-z|| \leq 1$, and so $\{u_r\}$, $\{v_r\}$, and $\{x_r\}$ are bounded. We may assume that they converge. Since r can be made arbitrarily large, $\{u_r\}$, $\{v_r\}$, and $\{x_r\}$ all converge to the same $\bar{x}$. Therefore

$$g_\delta(\bar{x}) = f_1(\bar{x})f_2(\bar{x}) + \delta||\bar{x}-z||$$
$$\leq f_1(z)f_2(z) = g_\delta(z) ,$$

which can only be true if $\bar{x} = z$.
Now assume r is so large that $||x_r-z|| < \varepsilon$, $||u_r-z|| < \varepsilon$, $||v_r-z|| < \varepsilon$, $|f_1(u_r)-f_1(z)| < \delta$, and $|f_2(v_r)-f_2(z)| < \delta$. Then P_r attains an unconditional local minimum at (u_r,v_r,x_r) . This also means that $G_1(u) := P_r(u,v_r,x_r)$ attains an unconditional local minimum at u_r , so by Theorem 1.5.21, $(f_1)_+(u_r;u)f_2(v_r) - \langle u_r^*,u\rangle \leq 0$ for all $u \in \mathbb{R}^m$, where $u_r^* = -r(u_r-x_r)$. Similarly, $G_2(v) := P_r(u_r,v,x_r)$ attains an unconditional local minimum at v_r , so

$$f_1(u_r)(f_2)_+(v_r;v) - \langle v_r^*,v\rangle \geq 0$$

for all $v \in \mathbb{R}^m$, where $v_r^* = -r(v_r - x_r)$. Moreover, $G_3(x) := P_r(u_r, v_r, x)$ attains an unconditional local minimum at x_r, so

$$\langle u_r^*, x\rangle + \langle v_r^*, x\rangle + \delta||x|| \geq 0$$

for all $x \in \mathbb{R}^m$. Therefore

$$u_r^* \in f_2(v_r)\partial^K f_1(u_r),$$

$$v_r^* \in f_1(u_r)\partial^K f_2(v_r),$$

and $u_r^* + v_r^* \in B_\delta(0)$, completing the proof. □

<u>Theorem 4.3.2</u>: Suppose $f_1, f_2: \mathbb{R}^m \to \overline{\mathbb{R}}$ are continuous and positive at $z \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$, and suppose

$$(4.3.2) \qquad \operatorname{dom} f_1^{\uparrow}(z;\cdot) - \operatorname{dom} f_2^{\uparrow}(z;\cdot) = \mathbb{R}^m .$$

Then

$$(4.3.3) \qquad \partial_a(f_1 f_2)(z) \subset f_1(z)\partial_a f_2(z) + \partial_a f_1(z) f_2(z) .$$

<u>Proof</u>: If $\partial_a(f_1 f_2)(z) = \phi$, (4.3.3) holds automatically. Let $x^* \in \partial_a(f_1 f_2)(z)$. Then there exist sequences $\{x_n\}$ and $\{x_n^*\}$ such that $x_n \to z$, $x_n^* \to x^*$, $(f_1 f_2)(x_n) \to (f_1 f_2)(z)$, and $x_n^* \in \partial^K(f_1 f_2)(x_n)$. By Lemma 4.3.1, there exist u_n, v_n, u_n^*, and v_n^* such that $||x_n - u_n|| \to 0$, $||x_n - v_n|| \to 0$,

$f_1(u_n)-f_1(x_n) \to 0$, $f_2(v_n)-f_2(x_n) \to 0$, $u_n^* \in \partial^K f_1(u_n) f_2(v_n)$, $v_n^* \in \partial^K f_2(v_n) f_1(u_n)$, and $||x_n^*-u_n^*-v_n^*|| \to 0$. As in [I2, Theorem 4], it suffices to show that either u_n^* or v_n^* is bounded, or equivalently, that either $\frac{v_n^*}{f_1(u_n)}$ or $\frac{u_n^*}{f_2(v_n)}$ is bounded. (Here we use the fact that f_1 and f_2 are positive near z and continuous at z .) The proof that either $\frac{v_n^*}{f_1(u_n)}$ or $\frac{u_n^*}{f_2(v_n)}$ is bounded uses Proposition 4.2.1 ([I2, Proposition 6]) and (4.3.2) and is identical to the proof of [I2, Theorem 4]. □

It is of course possible to prove versions of Lemma 4.3.1 and Theorem 4.3.2 involving n functions. In the n-function version of Theorem 4.3.2, assumption (4.3.2) would be replaced by the assumption that dom $f_i^{\uparrow}(z;\cdot)$, $i=1,\ldots,n$ be in strong general position.

An analogue of Theorem 2.3.15 for approximate subdifferentials has not yet been established, and it seems difficult to do so by the method used here. Such a result may well be true, however.

Observe that all of the subdifferential calculus formulae of chapter 2, as well as Theorem 4.3.2 and the results of [I2] for approximate subdifferentials, require assumptions involving the upper subderivative. For example, Corollary 2.3.5, Corollary 2.4.14, and [I2, Theorem 4] are sum rules involving different directional derivatives and subgradients, but all require assumption (4.3.2). The lim inf formulae of section 4.1 help explain the Clarke tangent cone's essential role in constraint qualifications. They show that $T_C(x_0)$ contains information about $K_C(x)$, $k_C(x)$ and $P_C(x)$ at points x near x_0, exactly what is needed in subdifferential calculus results for $f_+(x_0;\cdot)$, $f_\Box(x_0;)$, $\partial^K f$, $\partial^k f$, or $\partial_a f$.

CHAPTER V

Optimization

5.1. A basic necessary condition for optimality

The theory developed in chapters 1 through 4 can be readily applied to obtain necessary conditions for optimality in mathematical programs. Such applications are, in fact, one of the main motivations for studying tangent cones and subdifferential calculus.

To prove rudimentary necessary conditions, the basic ingredients are Theorem 1.5.21 and a subgradient sum formula. For example, let $f: E \to \overline{\mathbb{R}}$ and $C \subset E$, and suppose x_0 is a local minimizer for the mathematical program

$$\text{(P)} \qquad \min\{f(x) \mid x \in C\} .$$

Then an equivalent statement is that x_0 is a local minimizer of the function $f + \iota_C$. Suppose that R is a tangent cone satisfying two conditions:

(5.1.1) (i) $R(C,x) \subset K(C,x)$ for all $(C,x) \in 2^E \times E$.

(ii) A subgradient sum formula $\partial^R(f_1+f_2)(x) \subset \partial^R f_1(x) + \partial^R f_2(x)$ holds under certain conditions involving f_1, f_2, and x.

Then if $f + i_C$ satisfies the hypotheses guaranteeing (5.1.1)(ii) at x_0,

$$\begin{aligned} 0 &\in \partial^R(f+i_C)(x_0) \\ &\subset \partial^R f(x_0) + \partial^R i_C(x_0) . \end{aligned}$$

Thus we obtain the abstract Kuhn-Tucker type condition

(5.1.2) $$0 \in \partial^R f(x_0) + \partial^R i_C(x_0) .$$

In this chapter, we apply our knowledge of tangent cones and subdifferential calculus to give conditions under which (5.1.2) holds for various subgradients ∂^R, functions f, and sets C. In this section, we focus on (5.1.2) for various subgradients ∂^R.

Rockafellar has obtained the following result in the case $R := \uparrow$.

<u>Theorem 5.1.1</u>: Let E be a l.c.s., $C \subset E$, and $f: E \to \bar{\mathbb{R}}$, and suppose x_0 is a local minimizer for (P). Assume either

(5.1.3) (a) $T_C(x_0) \cap \operatorname{int} \operatorname{dom} f^{\uparrow}(x_0;\cdot) \neq \phi$ and f is directionally Lipschitzian at x_0, or

(b) $\operatorname{dom} f^{\uparrow}(x_0;\cdot) \cap \operatorname{int} T_C(x_0) \neq \phi$ and C is epi-Lipschitzian at x_0.

Then

(5.1.4) $$0 \in \partial f(x_0) + N_C(x_0) .$$

Proof: Condition (5.1.3) is assumption (2.1.5) of Theorem 2.1.2 with $f_1 := i_C$ and $f_2 := f$ or vice versa. Condition (5.1.1)(i) is true for the Clarke tangent cone, so (5.1.2) (which in this case is (5.1.4)) holds. $\square$

In the case in which f and C are convex in Theorem 5.1.1, (5.1.3) reduces to the assumption that either $C \cap \operatorname{int} \operatorname{dom} f \neq \phi$ or $\operatorname{int} C \cap \operatorname{dom} f \neq \phi$, and (5.1.4) is the well-known "Pshenichnyi's condition" (see [Hol, section 14])

(5.1.5) $$\partial f(x_0) \cap (C - x_0)^+ \neq \phi .$$

Notice that (5.1.3) is automatically satisfied when f is locally Lipschitzian near x_0.

A finite-dimensional version of Theorem 5.1.1 with weaker hypotheses is possible via Corollary 2.3.5. We state it below.

Theorem 5.1.2: Let $f: \mathbb{R}^m \to \overline{\mathbb{R}}$ be strictly l.s.c. at $x_0 \in C \subset \mathbb{R}^m$, where x_0 is a local minimizer for (P) and C is closed near x_0. Assume

(5.1.6) $$\operatorname{dom} f^{\uparrow}(x_0;\cdot) - T_C(x_0) = \mathbb{R}^m .$$

Then (5.1.4) holds.

<u>Proof</u>: Let $f_1 := f$, $f_2 := i_C$ in Corollary 2.3.5. In this case (2.3.8) is exactly (5.1.6). □

There is an analogous result for approximate subdifferentials.

<u>Theorem 5.1.3</u> [I2, Proposition 9]: Under the hypotheses of Theorem 5.1.2,

$$(5.1.7) \qquad 0 \in \partial_a f(x_0) + N_a(C,x_0) \ ,$$

where $\qquad N_a(C,x_0) := \partial_a i_C(x_0)$.

<u>Proof</u>: By Theorem 1.5.21, $0 \in \partial^K(f+i_C)(x_0)$, and by definition,

$$\partial^K(f+i_C)(x_0) \subset \partial_a(f+i_C)(x_0) \ .$$

Hence $\qquad 0 \in \partial_a(f+i_C)(x_0)$

$$\subset \partial_a f(x_0) + N_a(C,x_0)$$

by [I2, Theorem 4]. □

If R is not in general convex, it will be difficult to verify (5.1.1)(ii). However, if the directional derivative inequality

$$(5.1.8) \qquad (f_1+f_2)^R(x_0;\cdot) \leq f_1^{R_1}(x_0;\cdot) + f_2^{R_2}(x_0;\cdot)$$

is satisfied, we can derive optimality conditions in terms of "upper convex approximates", mentioned in section 2.4. We now generalize the definition given in Definition 2.4.17 for the case $R = R_1 = R_2 := k$.

Definition 5.1.4: (a) The function $h\colon E \to \overline{\mathbb{R}}$ is an R upper convex approximate to the function $f\colon E \to \overline{\mathbb{R}}$ at $x_0 \in E$ if

(5.1.9) $h(y) \geq f^R(x_0;y)$ for all $y \in E$.

(5.1.10) $h(\cdot)$ is proper, convex, l.s.c., and positively homogeneous.

(b) For a given R upper convex approximate to f at x_0, the set

$$\partial_h^R f(x_0) := \{x' \in E' \mid h(y) \geq \langle y, x' \rangle \text{ for all } y \in E\}$$

is called the h R-subgradient of f at x_0.

Now suppose h is an R upper convex approximate to f at x_0, D is a closed convex subcone of $R(C,x_0)$, and (5.1.8) is satisfied. Then if x_0 is a local minimizer for (P), we have for all $y \in \mathbb{R}^m$ that

$$\begin{aligned} 0 &\leq (f+i_C)^R(x_0;y) \\ &\leq f^R(x_0;y) + i_C^R(x_0;y) \quad \text{by (5.1.8)} \\ &= f^R(x_0;y) + i_{R(C,x_0)}(y) \\ &\leq h(y) + i_D(y) . \end{aligned}$$

Hence $0 \in \partial(h+i_D)(0)$.

If the hypotheses for the convex subgradient sum formula are satisfied for $h + i_D$ (ri dom h $\cap$ ri D $\neq \phi$ if $E := \mathbb{R}^m$, dom h $\cap$ int D $\neq \phi$ or int dom h $\cap$ D $\neq \phi$ in general), we obtain

$$0 \in \partial h(0) + \partial i_D(x_0) = \partial_h^R f(x_0) + D^0 .$$

Previous work with this concept has usually been carried out in the case $R := L$ ([Psl], [Il]), since (5.1.8) holds for $R := L$ without any conditions on f_1 and f_2 . For the same reason, another natural candidate is $R := E$. The results of section 2.4 open up some new possibilities. We give one example here.

Theorem 5.1.5: Under the hypotheses of Theorem 5.1.2, suppose $h: \mathbb{R}^m \to \overline{\mathbb{R}}$ is a k upper convex approximate to f at x_0 satisfying $h(\cdot) \leq f^{\dagger}(x_0;\cdot)$, and let D be a closed convex subcone of $k_C(x_0)$ such that $T_C(x_0) \subset D$. Then

$$(5.1.11) \qquad 0 \in \partial_h^k f(x_0) + D^0 .$$

Proof: Condition (5.1.6) ensures that we may apply (2.4.24) with $f_1 := f$, $f_2 := i_C$. It also implies that dom h $- D = \mathbb{R}^m$, so that we can apply [Rol, Theorem 23.8] with $n := 2$, $f_1 := h$, and $f_2 := i_D$. The inclusion (5.1.11) then follows from the discussion preceding the statement of the theorem. □

Results analogous to Theorem 5.1.5 can be derived by applying (2.4.25) in place of (2.4.24). Notice that Theorem 5.1.2 is the special case of Theorem 5.1.5 obtained by setting $h(\cdot) = f^{\uparrow}(x_0;\cdot)$ and $D := T_C(x_0)$. Also note that (5.1.11) can also be written

$$(5.1.12) \qquad 0 \in \partial_h^k f(x_0) + N_D(0)$$

since $\qquad D = T_D(0)$.

5.2. Some special cases, and a Fritz John type necessary condition

We discussed a Kuhn-Tucker type necessary condition for optimality in (P) in section 5.1. In this section, we study some important special cases of the set C in (P) and combine our results and some recent work of Watkins [W1] to prove a new Fritz John type necessary condition for (P).

We begin by considering the constraint set $C := \{x \in \mathbb{R}^m | G(x) = 0\}$, where $G: \mathbb{R}^m \to \mathbb{R}^p$ is strictly differentiable at $x_0 \in G^{-1}(0)$ and $\nabla G(x_0)\mathbb{R}^m = \mathbb{R}^p$. As we saw in Corollary 1.3.18,

$$K_C(x_0) = k_C(x_0) = T_C(x_0) = \nabla G(x_0)^{-1}(0) .$$

Hence $\qquad K_C^0(x_0) = k_C^0(x_0) = N_C(x_0) = \bigcup_{\lambda \in \mathbb{R}^p} \nabla G(x_0)$.

We next consider the set

$$C := \{x \in \mathbb{R}^m | g(x) \le 0\} ,$$

where $g: \mathbb{R}^m \to \overline{\mathbb{R}}$ is strictly l.s.c. at $x_0 \in g^{-1}(0)$. For this constraint set, we can derive the following result.

Proposition 5.2.1 (cf. [Ro3, Theorem 5]): Let $g: \mathbb{R}^m \to \overline{\mathbb{R}}$ be strictly l.s.c. at $x_0 \in g^{-1}(0)$, and let $C := \{x \in \mathbb{R}^m | g(x) \le 0\}$. Assume $0 \notin \partial g(x_0)$. Then

$$(5.2.1) \qquad T_C(x_0) \supset \{y \in \mathbb{R}^m | g^{\uparrow}(x_0;y) \le 0\} .$$

If in addition $\partial g(x_0) \neq \phi$, then

$$(5.2.2) \qquad N_C(x_0) \subset \bigcup_{\lambda \ge 0^+} \lambda \partial g(x_0) .$$

Proof: In Proposition 2.3.11, let $n := 2$, $D_1 := \{(z,u) \in \mathbb{R}^m \times \mathbb{R} | u = 0\}$, and $D_2 := \text{epi } g$. To apply Proposition 2.3.11, we must verify that $T_{D_1}(x_0,0) - T_{D_2}(x_0,0) = \mathbb{R}^{m+1}$, i.e., that $\{(z,u) | u = 0\} - \text{epi } g^{\uparrow}(x_0;\cdot) = \mathbb{R}^{m+1}$. To this end, let $(x',r') \in \mathbb{R}^m \times \mathbb{R}$. Since $0 \notin \partial g(x_0)$, there exists $y \in \mathbb{R}^m$ with $g^{\uparrow}(x_0;y) < 0$. Since $g^{\uparrow}(x_0;\cdot)$ is positively homogeneous, this means that for any $r \in \mathbb{R}$, there exists $\bar{y} \in \mathbb{R}^m$ with $g^{\uparrow}(x_0;\bar{y}) \le r$. We may then choose $y' \in \mathbb{R}^m$ such that $(y',-r') \in \text{epi } g^{\uparrow}(x_0;\cdot)$, and so

$$(x',r') = (x'+y',0) - (y',-r')$$
$$\in \{(z,u)\,|\,u = 0\} - \text{epi } g^{\dagger}(x_0;\cdot) .$$

Now by (2.3.20),

$$T_{C\times 0}(x_0,0) \supset (\mathbb{R}^m\times 0) \cap \text{epi } g^{\dagger}(x_0;\cdot) .$$

Hence $T_C(x_0) \supset \{y \in \mathbb{R}^m | g^{\dagger}(x_0;y) \le 0\}$. By (2.3.21),

$$N_{C\times 0}(x_0,0) \subset (0 \times \mathbb{R}) + N_{D_2}(x_0,0) .$$

If $\partial g(x_0) \neq \phi$, we have by (2.3.23) that

$$N_{D_2}(x_0) \subset \bigcup_{\lambda\ge 0^+} \lambda(\partial g(x_0),-1) .$$

Hence $N_C(x_0) \subset \bigcup_{\lambda\ge 0^+} \lambda\partial g(x_0)$. □

The infinite-dimensional analogue of Proposition 5.2.1 requires that C be epi-Lipschitzian at x_0 ([Ro3]).

By an argument paralleling that of Proposition 5.2.1, we can apply Corollary 2.4.13 to prove a related result.

<u>Proposition 5.2.2</u>: Under the hypotheses of Proposition 5.2.1,

(5.2.3) $$k_C(x_0) = \{y \in \mathbb{R}^m | g_{\Box}(x_0;y) \le 0\}$$

and

(5.2.4) $$K_C(x_0) \supset \{y \in \mathbb{R}^m | g_{+}(x_0;y) \le 0\} .$$

<u>Proof</u>: By the proof of Proposition 5.2.1, the hypotheses of Corollary 2.4.13 are satisfied. Then (5.2.3) and (5.2.4) follow from (2.4.21) and (2.4.22). □

We can readily apply Propositions 5.2.1 and 5.2.2 to prove similar results for constraint sets of the form

$$C := \{x \mid g_i(x) \leq 0 \ , \ i = 1,\ldots,n\} \ .$$

<u>Proposition 5.2.3</u>: Let $g_i : \mathbb{R}^m \to \overline{\mathbb{R}}$, $i = 1,\ldots,n$ be strictly l.s.c. at $x_0 \in \mathbb{R}^m$, and define $C := \{x \in \mathbb{R}^m \mid g_i(x) \leq 0, \ i = 1,\ldots,n\}$. Call $I(x) := \{i \in \{1,\ldots,n\} \mid g_i(x) = 0\}$. Assume that g_i is continuous at x_0 for each $i \notin I(x_0)$, that $0 \notin \partial g_i(x_0)$ for each $i \in I(x_0)$, and that the sets $\{y \mid g_i^{\uparrow}(x_0;y) \leq 0\}$, $i \in I(x_0)$, are in strong general position. Then

$$(5.2.5) \qquad T_C(x_0) \supset \{y \mid g_i^{\uparrow}(x_0;y) \leq 0, \ i = 1,\ldots,n\}$$

$$(5.2.6) \qquad k_C(x_0) = \{y \mid (g_i)_{\square}(x_0;y) \leq 0, \ i = 1,\ldots,n\}$$

$$(5.2.7) \qquad K_C(x_0) \supset \{y \mid (g_1)_{+}(x_0;y) \leq 0, \ (g_i)_{\square}(x_0;y) \leq 0, \ i = 2,\ldots,n\}$$

If in addition $\partial g_i(x_0) \neq \phi$, $i \in I(x_0)$, then

$$(5.2.8) \qquad N_C(x_0) \subset \{\sum_{I(x_0)} \lambda_i \partial g_i(x_0) \mid \lambda_i \geq 0^+\} \ .$$

Proof: Let $D_i := \{x \in \mathbb{R}^m | g_i(x) \le 0\}$, $i = 1,\ldots,n$ in Proposition 2.3.11. For $i \notin I(x_0)$, $T_{D_i}(x_0) = \mathbb{R}^m$. Thus $\bigcap_{i=1}^{n} T_{D_i}(x_0) = \bigcap_{I(x_0)} T_{D_i}(x_0)$, and $T_{D_i}(x_0)$, $i=1,\ldots,n$ are in strong general position if and only if $T_{D_i}(x_0)$, $i \in I(x_0)$ are in strong general position. By Proposition 5.2.1, $T_{D_i}(x_0) \supset \{y | g_i^{\uparrow}(x_0;y) \le 0\}$, for each $i \in I(x_0)$. Since $\{y | g_i^{\uparrow}(x_0;y) \le 0\}$, $i \in I(x_0)$ are in strong general position, so are $T_{D_i}(x_0)$, $i \in I(x_0)$, and we may apply Proposition 2.3.11. Combining (2.3.20) and (5.2.1), we obtain (5.2.5), while (2.3.21) and (5.2.2) combine to give (5.2.8). The proofs of (5.2.6) and (5.2.7) are similar to that of (5.2.5), with the extensions of (2.4.21) and (2.4.22) to n sets applied in place of (2.3.20). □

Now we can employ Propositions 5.2.1 and 5.2.2 to derive an important special case of Theorem 5.1.2 and a result along the lines of Theorem 5.1.5.

Theorem 5.2.4: Let $D \subset \mathbb{R}^m$ be closed near x_0 , $f: \mathbb{R}^m \to \overline{\mathbb{R}}$ and $g_i: \mathbb{R}^m \to \overline{\mathbb{R}}$, $i = 1,\ldots,n$ strictly l.s.c. at x_0 , and $G: \mathbb{R}^m \to \mathbb{R}^p$ strictly differentiable at x_0 , where x_0 is a local minimizer of

$$(5.2.9) \quad \min\{f(x) | g_i(x) \le 0,\ i=1,\ldots,n,\ G(x)=0,\ x \in D\} .$$

Call $I(x) := \{i \in \{1,\ldots,n\} | g_i(x) \le 0\}$. Assume that $0 \notin \partial g_i(x_0)$ and $\partial g_i(x_0) \ne \phi$ for each $i \in I(x_0)$, that g_i is continuous at x_0 for each $i \notin I(x_0)$, and that $\nabla G(x_0)\mathbb{R}^m = \mathbb{R}^p$. In addition, assume that $\{y | g_i^{\uparrow}(x_0;y) \le 0\}$, $i \in I(x_0)$, $T_D(x_0)$, and $\nabla G(x_0)^{-1}(0)$ are in strong general position, and that

$$(5.2.10) \qquad \operatorname{dom} f^{\uparrow}(x_0;\cdot) - \{y | g_i^{\uparrow}(x_0;y) \le 0,\ i=1,\ldots,n,\ \nabla G(x_0)y=0,\ y \in T_D(x_0)\} = \mathbb{R}^m .$$

Then

$$(5.2.11) \qquad 0 \in \partial f(x_0) + \sum_{I(x_0)} \lambda_i \partial g_i(x_0) + \lambda \nabla G(x_0) + N_D(x_0),$$
$$\text{for some } \lambda_i \ge 0^+ \text{ and } \lambda \in \mathbb{R}^p .$$

<u>Proof</u>: Call $C := \{x \in \mathbb{R}^m | x \in D,\ g_i(x) \le 0,\ i=1,\ldots,n,\ G(x) = 0\}$.

Since $\nabla G(x_0)$ is of full rank, $T_{G^{-1}(0)}(x_0) = \nabla G(x_0)^{-1}(0)$. By our strong general position assumption, Propositions 2.3.11 and 5.2.1 imply that $\{y | g_i^{\uparrow}(x_0;y) \le 0,\ i=1,\ldots,n,\ y \in T_D(x_0),\ \nabla G(x_0)y=0\} \subset T_C(x_0)$. Thus assumption (5.2.10) guarantees that (5.1.6) holds. By Theorem 5.1.2,

$$0 \in \partial f(x_0) + N_C(x_0) \subset \partial f(x_0) + \{\sum_{I(x_0)} \lambda_i \partial g_i(x_0) | \lambda_i \ge 0^+\}$$
$$+ \{\lambda \nabla G(x_0) | \lambda \in \mathbb{R}^p\} + N_D(x_0)$$

by Propositions 2.3.11 and 5.2.1. Therefore

$0 \in \partial f(x_0) + \sum_{I(x_0)} \lambda_i \partial g_i(x_0) + \lambda \nabla G(x_0) + N_D(x_0)$ for some $\lambda_i \geq 0^+$ and $\lambda \in \mathbb{R}^p$. □

<u>Theorem 5.2.5</u>: Under the hypotheses of Theorem 5.2.4, suppose in addition that $h: \mathbb{R}^m \to \overline{\mathbb{R}}$ and $h_i: \mathbb{R}^m \to \overline{\mathbb{R}}$, $i \in I(x_0)$, are k upper convex approximates to f and g_i, $i \in I(x_0)$, at x_0 such that $h(\cdot) \leq f^{\uparrow}(x_0;\cdot)$ and $h_i(\cdot) \leq g_i^{\uparrow}(x_0;\cdot)$. Let S be a closed convex subcone of $k_D(x_0)$ such that $T_D(x_0) \subset S$. Then

$$0 \in \partial_h^k f(x_0) + \sum_{I(x_0)} \lambda_i \partial_{h_i}^k g_i(x_0) + \lambda \nabla G(x_0) + S^0 . \tag{5.2.12}$$

for some $\lambda_i \geq 0^+$ and $\lambda \in \mathbb{R}^p$.

<u>Proof</u>: Define the set C as in Theorem 5.2.4, and define $C_i := \{x \in \mathbb{R}^m | g_i(x) \leq 0\}$, $i = 1,\ldots,n$, $C_{n+1} := G^{-1}(0)$, and $C_{n+2} := D$. By hypothesis,

$$\operatorname{dom} f^{\uparrow}(x_0;\cdot) - \operatorname{dom} i_C^{\uparrow}(x_0;\cdot) = \mathbb{R}^m ,$$

so by Corollary 2.4.14, for all $y \in \mathbb{R}^m$,

$$0 \leq f_{\Box}(x_0;y) + (i_C)_{\Box}(x_0;y) .$$

Now since $T_{C_i}(x_0)$, $i = 1,\ldots,n+2$ are in strong general position, the n-function version of Corollary 2.4.14 implies that for all $y \in \mathbb{R}^m$,

$$0 \le f_{\Box}(x_0;y) + \sum_{j=1}^{n+2} i_{k_{C_j}(x_0)}(y) .$$

Now call $D_i := k_{C_i}(x_0)$ for $i \notin I(x_0)$, $i \le n+1$; $D_i := \{y | h_i(y) \le 0\}$ for $i \in I(x_0)$, and $D_{n+2} := S$. By Proposition 5.2.2, each $D_j \subset k_{C_j}(x_0)$. Thus

$$0 \le h(y) + \sum_{j=1}^{n+2} i_{D_j}(y) \quad \text{for all } y \in \mathbb{R}^m ,$$

and so $0 \in \partial(h + \sum_{j=1}^{n+2} i_{D_j})(0)$. Assumption 5.2.10 implies that $\text{dom}\, h - \text{dom}(\sum_{j=1}^{n+2} i_{D_j}) = \mathbb{R}^m$, so

$$0 \in \partial h(0) + \partial(\sum_{j=1}^{n+2} i_{D_j}(0))$$

$$\subset \partial h(0) + \sum_{j=1}^{n+2} \partial i_{D_j}(0)$$

by our strong general position assumption and [Ro3, Theorem 23.8]. In other words,

$$0 \in \partial_h^k f(x_0) + \{\sum_{I(x_0)} \lambda_i \partial_{h_i}^k g_i(x_0) | \lambda_i \ge 0\}$$

$$+ \{\lambda \nabla G(x_0) | \lambda \in \mathbb{R}^p\} + S^0 ,$$

and (5.2.12) holds. □

Observe that Theorem 5.2.4 is a special case of Theorem 5.2.5 with $h(\cdot) := f^{\dagger}(x_0;\cdot)$, $h_i(\cdot) \doteq g_i^{\dagger}(x_0;\cdot)$, $i \in I(x_0)$ and $S := T_D(x_0)$. We also note that (5.2.10) and the general position assumption in Theorems 5.2.4 and 5.2.5 are equivalent to the single assumption that $f^{\dagger}(x_0;\cdot)$, $\{y | g_i^{\dagger}(x_0;y) \leq 0\}$, $T_D(x_0)$, and $\nabla G(x_0)^{-1}(0)$ are in strong general position.

We can also derive necessary conditions for optimality in (5.2.9) via the following "Dubovitskii-Milyutin" result due to Watkins [W1]:

<u>Theorem 5.2.6</u>: Let $C_i \subset \mathbb{R}^m$, $i = 1,\ldots,p$ be sets which are closed near x_0 and satisfy $\bigcap_{i=1}^{p} C_i = \{x_0\}$, and suppose that at least one of the sets $T_{C_i}(x_0)$ is not a subspace. Then there exist $a_i \in N_{C_i}(x_0)$, not all equal to 0, such that $\sum_{i=1}^{p} a_i = 0$.

The proof of Theorem 5.2.6 relies on Proposition 2.3.11 and can be found in [W1]. Watkins applies Theorem 5.2.6 to prove a Fritz John type theorem for a constrained optimization problem involving locally Lipschitzian functions. By employing Proposition 5.2.1, we can establish a generalization of this result in which the functions involved are not necessarily locally Lipschitzian. We first need a technical lemma. It's proof is straightforward, and we leave it to the reader.

<u>Lemma 5.2.7</u>: Let $g: E \times E \to \overline{\mathbb{R}}$ and $h: E \times E \to \overline{\mathbb{R}}$ be extended real-valued functions, and let $x_0, y_0 \in E$. If $\lim_{\substack{x \to x_0 \\ y \to y_0}} h(x,y)$ is finite, then

$$\limsup_{x \to x_0} \inf_{y \to y_0} (g+h)(x,y) = \limsup_{x \to x_0} \inf_{y \to y_0} g(x,y) + \lim_{\substack{x \to x_0 \\ y \to y_0}} h(x,y) .$$

<u>Theorem 5.2.8</u>: Let $D \subset \mathbb{R}^m$ be closed near x_0, $f: \mathbb{R}^m \to \overline{\mathbb{R}}$ and $g_i: \mathbb{R}^m \to \overline{\mathbb{R}}$, $i = 1,\ldots,n$ strictly l.s.c. at x_0, and $G: \mathbb{R}^m \to \mathbb{R}^p$ strictly differentiable at x_0, where x_0 is a local minimizer of (5.2.9). Assume that $0 \notin \partial g_i(x_0) \neq \phi$ for each $i \in I(x_0)$, that g_i is continuous at x_0 for each $i \notin I(x_0)$, and that $\nabla G(x_0)\mathbb{R}^m = \mathbb{R}^p$. Then

$$(5.2.13) \qquad 0 \in \lambda_0 \partial f(x_0) + \sum_{I(x_0)} \lambda_i \partial g_i(x_0) + \lambda \nabla G(x_0) + N_D(x_0)$$

where $\lambda_0, \lambda_1 \geq 0^+$ and $\lambda \in \mathbb{R}^p$. Moreover, λ and the particular elements of $\lambda_0 \partial f(x_0)$, $\lambda_i \partial g_i(x_0)$ for $i \in I(x_0)$, and $N_D(x_0)$ in (5.2.13) are not all equal to 0.

<u>Proof</u>: Let $C_0 := \{x \in \mathbb{R}^m | f(x) \leq f(x_0) - (x-x_0)^2\}$,
$C_i := \{x \in \mathbb{R}^m | g_i(x) \leq 0\}$, $i = 1,\ldots,n$,
$C_{n+1} := \{x \in \mathbb{R}^m | G(x) = 0\}$, and $C_{n+2} := D$. Since x_0 is a local minimizer for (P), $\bigcap_{i=0}^{n+2} C_i = \{x_0\}$. Now $T_{C_0}(x_0)$ is not a subspace, so by Theorem 5.2.6, there exist

$a_i \in N_{C_i}(x_0)$, $i = 0,\ldots,n+2$, not all equal to 0 , such that $\sum_{i=0}^{n+2} a_i = 0$. For $i \in I(x_0)$, $N_{C_i}(x_0) \subset \bigcup_{\lambda \geq 0^+} \lambda \partial g_i(x_0)$ by (5.2.2). For $i \in \{1,\ldots,n\} \backslash I(x_0)$, $N_{C_i}(x_0) = \{0\}$ since each such g_i is continuous at x_0 and $g_i(x_0) < 0$. Since $\nabla G(x_0)$ is of full rank, $N_{C_{n+1}}(x_0) \subset \bigcup_{\lambda \in \mathbb{R}^p} \lambda \nabla G(x_0)$. Call $\bar{f}(x) := f(x) - f(x_0) + (x-x_0)^2$. By Lemma 5.2.7, $\bar{f}^{\uparrow}(x_0;\cdot) \leq f^{\uparrow}(x_0;\cdot)$, so $\partial \bar{f}(x_0) = \partial f(x_0)$ and

$$N_{C_0}(x_0) \subset \bigcup_{\lambda \geq 0^+} \partial \bar{f}(x_0) = \bigcup_{\lambda \geq 0^+} \lambda \partial f(x_0) .$$

Putting these facts together, we obtain (5.2.13). □

No Fritz John type analogue of Theorem 5.2.5 has yet been established.

5.3 The objective function in problem (P): some examples from the literature.

The generalized subdifferential calculus developed in Chapters 2 and 3 enables us to work with a variety of objective functions in (P) . In this section, we give some examples of problems from the recent optimization literature which are tractable with this subdifferential calculus.

(a) Weak Pareto minima

Definition 5.3.1: Let E be a l.c.s. and $C \subset E$, and let $f_i : E \to \bar{\mathbb{R}}$, $i = 1,\ldots,n$ be extended real valued functions. We say x_0 is weakly $(\mathbb{R}^n_+)$ Pareto minimal for $f_1,\ldots,f_n$ with respect to C if there is no $x \in C$ such that $f_i(x) < f_i(x_0)$ for each i.

In [Mil], necessary conditions for weak Pareto minimality are developed in the case in which E is a Banach space and $C := \{x \in E \mid g_j(x) \le 0,\ j \in J,\ h_k(x) = 0,\ k \in K,\ x \in D\}$ where J and K are finite index sets and $f_i, g_j : E \to \bar{\mathbb{R}}$, and $h_k = E \to \bar{\mathbb{R}}$ are locally Lipschitzian. Since x_0 is weakly Pareto minimal for $f_1,\ldots,f_n$ with respect to C if and only if

$$0 = \min\{\max_{1\le i\le n} (f_i(x) - f_i(x_0)) \mid x \in C\},$$

the problem Minami considers is actually a special case of that studied in [Cl 1], and Minami derives his necessary conditions by mimicking the arguments of [Cl 1].

The results of [Mi1] can be generalized to functions ~~which are~~ not necessarily locally Lipschitzian via Theorem 5.1.1 (for infinite-dimensional E) or Theorem 5.1.2 and 5.1.3 (for finite-dimensional E). We consider the finite-dimensional case here.

<u>Proposition 5.3.2</u>: Let $f_i : \mathbb{R}^m \to \bar{\mathbb{R}}$, $i = 1,\dots,n$ be strictly l.s.c. at x_0 and let $C \subset \mathbb{R}^m$ be closed near x_0, where x_0 is weakly Pareto minimal for $f_1,\dots,f_n$ with respect to C. Assume that $\operatorname{dom} f_1^{\uparrow}(x_0;\cdot),\dots,\operatorname{dom} f_n^{\uparrow}(x_0;\cdot)$ and $T_C(x_0)$ are in strong general position, and that $\partial f_i(x_0)$, $i = 1,\dots,n$ are non-empty. Then

$$(5.3.1) \qquad 0 \in \sum_{i=1}^{n} \lambda_i \partial f_i(x_0) + N_C(x_0)$$

and

$$(5.3.2) \qquad 0 \in \sum_{i=1}^{n} \alpha_i \partial_a f_i(x_0) + N_a(C, x_0)$$

where each $\lambda_i \geq 0^+$, $\alpha_i \geq 0^+$, and $\sum_{i=1}^{n} \lambda_i = 1$, $\sum_{i=1}^{n} \alpha_i = 1$.

<u>Proof</u>: Define $f : \mathbb{R}^m \to \bar{\mathbb{R}}$ by $f(x) := \max_{1 \leq i \leq n} (f_i(x) - f_i(x_0))$. Then $0 = f(x_0) = \min \{f(x) \mid x \in C\}$. Our strong general position assumption ensures that the hypotheses of Theorems 5.1.2 and 5.1.3 are satisfied. Thus

$$0 \in \partial f(x_0) + N_C(x_0) \quad \text{by Theorem 5.1.2,}$$

$$\subset \{\sum_{i=1}^{n} \lambda_i \partial f_i(x_0) \mid \lambda_i \geq 0^+, \sum_{i=1}^{n} \lambda_i = 1\} + N_C(x_0)$$

by Proposition 2.3.14.

Therefore (5.3.1) is satisfied. The proof of (5.3.2) follows similarly from Theorem 5.1.3 and the n-function version of [I 2, Corollary 4.6]. □

Special cases of Proposition 5.3.1 for specific constraint sets C can be proven as in section 5.2. We leave the details to the reader.

We can also treat this problem as in Theorems 5.1.5 and 5.2.5.

<u>Proposition 5.3.3</u>: In addition to the hypotheses of Proposition 5.3.2, suppose $h_i : \mathbb{R}^m \to \bar{\mathbb{R}}$, $i = 1,\ldots,n$ are k upper convex approximates of f_i at x_0 such that $h_i(\cdot) \leq f_i^{\uparrow}(x_0;\cdot)$ for each i. Let $D \subset \mathbb{R}^m$ be a closed convex cone satisfying $T_C(x_0) \subset D \subset N_C(x_0)$. Then

(5.3.3) $$0 \in \sum_{i=1}^{n} \lambda_i \partial^k_{h_i} f_i(x_0) + D^0 ,$$

where each $\lambda_i \geq 0^+$ and $\sum_{i=1}^{n} \lambda_i = 1$.

<u>Proof</u>: The proof parallels the arguments used in proving Theorems 5.1.5 and 5.2.5. □

(b) Fractional programming.

One immediate application of the quotient rules derived in chapters 2 and 3 is in establishing necessary optimality conditions in nonsmooth fractional programming. Specifically, we consider the problem

(5.3.4) $\quad \min \{f(x)/g(x) \mid x \in C\}$.

where at a local minimizer x_0 , $f : E \to \bar{\mathbb{R}}$ is positive and l.s.c. (or strictly l.s.c. if $E := \mathbb{R}^m$) and $g : E \to \bar{\mathbb{R}}$ is positive and continuous. It is a usual assumption in fractional programming (see [Cr1]) that the numerator be nonnegative and the denominator positive, so our positivity assumptions on both the numerator and denominator are not much more restrictive than what is normally required. On the other hand, we require less smoothness on the part of f and g than has previously been assumed. For previous work on quotient rules for nonsmooth functions, see [Hi 1] (or [Cl4]) and [Bo1] .

We can readily apply Theorem 5.1.1 and Proposition 3.4.4 to prove necessary optimality conditions for (5.3.4) in the setting of [Ro3] , as we now demonstrate.

<u>Proposition 5.3.4</u>: Let $f : E \to \bar{\mathbb{R}}$ be nonnegative on E and positive and l.s.c. at x_0 , and let $g : E \to \bar{\mathbb{R}}$ be positive on E and continuous at x_0 , where x_0 is a

local minimizer for (5.3.4). Assume that C is epi-Lipschitzian at x_0 and

(5.3.5) $\quad \operatorname{dom} f^{\uparrow}(x_0;\cdot) \cap \operatorname{dom}(-g)^{\uparrow}(x_0;\cdot) \cap \operatorname{int} T_C(x_0) \neq \phi$,

and that $-g$ is directionally Lipschitzian at x_0 and

(5.3.6) $\quad \operatorname{dom} f^{\uparrow}(x_0;\cdot) \cap \operatorname{int} \operatorname{dom}(-g)^{\uparrow}(x_0;\cdot) \neq \phi$.

Then

(5.3.7) $$0 \in \frac{f(x_0)\partial(-g)(x_0) + g(x_0)\partial f(x_0)}{(g(x_0))^2} + N_C(x_0)$$

<u>Proof</u>: By (3.4.9), assumption (5.3.5) implies that $\operatorname{dom}(f/g)^{\uparrow}(x_0;\cdot) \cap \operatorname{int} T_C(x_0) \neq \phi$. Thus we may apply Theorem 5.1.1 to obtain

$$\begin{aligned} 0 &\in \partial(f/g)(x_0) + N_C(x_0) \\ &\subset \frac{f(x_0)\partial(-g)(x_0) + g(x_0)\partial f(x_0)}{(g(x_0))^2} + N_C(x_0) \end{aligned}$$

by (5.3.6) and Proposition 3.4.4. □

Of course (5.3.7) is true under weaker assumptions if E is finite-dimensional.

<u>Proposition 5.3.5</u>: Let $f : \mathbb{R}^m \to \bar{\mathbb{R}}$ be nonnegative on $\mathbb{R}^m$ and positive and strictly l.s.c. at x_0 , and let $g : \mathbb{R}^m \to \bar{\mathbb{R}}$ be positive on $\mathbb{R}^m$ and continuous at x_0 , where x_0 is a local minimizer for (5.3.4). Assume that $\operatorname{dom} f^{\uparrow}(x_0;\cdot)$, $\operatorname{dom}(-g)^{\uparrow}(x_0;\cdot)$, and $T_C(x_0)$ are in strong

general position. Then (5.3.7) holds.

Proof: By Theorem 5.1.2,

$$0 \in \partial(f/g)(x_0) + N_C(x_0)$$

$$\subset \frac{f(x_0)\partial(-g)(x_0) + g(x_0)\partial f(x_0)}{g(x_0))^2} + N_C(x_0)$$

by Proposition 2.3.20. □

We can prove a similar result involving approximate subdifferentials with the help of Theorem 4.3.2 and the following technical lemma.

Lemma 5.3.6: Suppose $g : E \to \bar{\mathbb{R}}$ is continuous and positive at $x_0 \in \operatorname{dom} g$. Then for all $y \in E$,

$$(5.3.8) \qquad (\tfrac{1}{g})_+(x_0;y) = \frac{(-g)_+(x_0;y)}{(g(x_0))^2} .$$

Moreover,

$$(5.3.9) \qquad \partial_a(\tfrac{1}{g})(x_0) = \frac{1}{(g(x_0))^2}\partial_a(-g)(x_0) .$$

Proof: The proof of (5.3.8) parallels that of Lemma 2.3.19 and uses Lemma 3.5.3. It follows from (5.3.8) that

$$(5.3.10) \qquad \partial^K(\tfrac{1}{g})(x_0) = \frac{1}{(g(x_0))^2}\partial^K(-g)(x_0) ,$$

and (5.3.9) is a direct consequence of (5.3.10). □

Proposition 5.3.7: Let $f : \mathbb{R}^m \to \bar{\mathbb{R}}$ and $g : \mathbb{R}^m \to \bar{\mathbb{R}}$ be continuous and positive at x_0, a local minimizer for (5.3.4). Assume that $\operatorname{dom} f^{\uparrow}(x_0;\cdot)$, $\operatorname{dom}(-g)^{\uparrow}(x_0;\cdot)$, and $T_C(x_0)$ are in strong general position. Then

$$(5.3.11) \qquad 0 \in \frac{f(x_0)\,\partial_a(-g)(x_0) + g(x_0)\,\partial_a f(x_0)}{(g(x_0))^2} + N_a(C,x_0) .$$

Proof: Since $\operatorname{dom}(f/g)^{\uparrow}(x_0;\cdot) - T_C(x_0) = \mathbb{R}^m$ and $\operatorname{dom} f^{\uparrow}(x_0;\cdot) - \operatorname{dom}(-g)^{\uparrow}(x_0;\cdot) = \mathbb{R}^m$, we may apply Theorems 4.3.2 and 5.1.3. Thus

$$0 \in f(x_0)\,\partial_a(\tfrac{1}{g})(x_0) + \frac{1}{g(x_0)}\,\partial_a f(x_0) + N_a(C,x_0)$$

$$\subset \frac{f(x_0)\,\partial_a(-g)(x_0) + g(x_0)\,\partial_a f(x_0)}{(g(x_0))^2} + N_a(C,x_0)$$

by Lemma 5.3.6. □

Among the consequences of Theorem 2.4.11 are product and quotient rules for $f_{\square}(x_0;\cdot)$, although we did not explicitly state them in chapter 2. As a result, we can give an optimality condition for (5.3.4) involving upper convex approximates.

Proposition 5.3.8: In addition to the hypotheses of Proposition 5.3.5, suppose that $h_i : \mathbb{R}^m \to \bar{\mathbb{R}}$, $i = 1,2$ are k upper convex approximates of f and $-g$, respectively, with $h_1(\cdot) \le f^{\uparrow}(x_0;\cdot)$ and $h_2(\cdot) \le (-g)^{\uparrow}(x_0;\cdot)$. Let D be a closed convex cone such that $T_C(x_0) \subset D \subset k_C(x_0)$. Then

$$(5.3.12) \qquad 0 \in \frac{f(x_0)\,\partial^k_{h_2}(-g)(x_0) + g(x_0)\,\partial^k_{h_1} f(x_0)}{(g(x_0))^2} + D^0 .$$

We can also use our subdifferential calculus results to formulate optimality conditions for the problem of generalized fractional programming [Cr 1]:

$$(5.3.13) \qquad \min\{ \max_{1\le i\le n} \frac{f_i(x)}{g_i(x)} \mid x \in C\} .$$

We state below some necessary optimality conditions for (5.3.13) which can be derived by applying Propositions 2.3.14 and 2.3.20.

Proposition 5.3.9: Let $f_i : \mathbb{R}^m \to \bar{\mathbb{R}}$, $i = 1,\dots,n$ be nonnegative on $\mathbb{R}^m$ and positive and strictly l.s.c. at x_0, and let $g_i : \mathbb{R}^m \to \mathbb{R}$, $i = 1,\dots,n$ be positive on $\mathbb{R}^m$ and continuous at x_0 , where x_0 is a local minimizer for (5.3.13). Call $I(x) := \{i \mid \frac{f_i(x)}{g_i(x)} = \max_{1\le j\le n} \frac{f_j(x)}{g_j(x)}\}$. Suppose f_j is continuous at x_0 for each $j \notin I(x_0)$. Assume that $\partial(f_j/g_j)(x_0) \neq \phi$ for each $j \in I(x_0)$, that

$$(5.3.14) \qquad \operatorname{dom} f_i^{\uparrow}(x_0;\cdot) - \operatorname{dom}(-g_i)^{\uparrow}(x_0;\cdot) = \mathbb{R}^m,\ i \in I(x_0),$$

and that the sets $\hat{T}_C(x_0)$ and

$$\operatorname{dom} f_i^{\uparrow}(x_0;\cdot) \cap \operatorname{dom}(-g_i)^{\uparrow}(x_0;\cdot),\ i \in I(x_0)$$

are in strong general position. Then

$$(5.3.15) \qquad 0 \in \sum_{I(x_0)} \lambda_i \left(\frac{g_i(x_0)\,\partial f_i(x_0) + f_i(x_0)\,\partial(-g_i)(x_0)}{(g_i(x_0))^2} \right) + N_C(x_0)$$

where each $\lambda_i \geq 0^+$ and $\sum_{I(x_0)} \lambda_i = 1$.

(c) Penalty-type objective functions.

Consider problem (P) in the case in which $C := \{x \in E \mid g(x) \leq 0 \ , \ h(x) = 0\}$ where $g : E \to \bar{\mathbb{R}}$, $h : E \to \bar{\mathbb{R}}$. In connection with this program, unconstrained problems like the following are often studied.

(P1) $\quad \min\limits_{x \in E} f(x) + c_1 g^+(x) + c_2 |h(x)|$,

where $c_1, c_2 > 0$ and $g^+(x) := \max(0, g(x))$.

(P2) $\quad \min\limits_{x \in E} f(x) + \exp(g^+(x) - 1) + \exp(|h(x)| - 1)$

Charalambous [Ch 1] derives necessary optimality conditions for (P1), (P2), and other similar penalty-type functions in the case in which $E := \mathbb{R}^m$ and $f, g,$ and h are continuously differentiable.

Employing the subdifferential calculus results of [Ro3] and chapter 3, one can easily derive optimality conditions for (P1) and (P2) where E is a l.c.s. and for much more general $f, g,$ and h . Corresponding finite-dimensional results under weaker constraint qualifications are obtainable by the results of chapter 2.

First of all we formulate the following general result, which follows easily from the subgradient sum formula and Theorems 3.2.13 and 2.3.15:

Proposition 5.3.10: Let $f : E \to \bar{\mathbb{R}}$, and let $g : E \to \bar{\mathbb{R}}$ and $h : E \to \bar{\mathbb{R}}$ be continuous at x_0, a local minimizer of (P1). Assume that g and $|h|$ are directionally Lipschitzian at x_0, and that

$$\text{(5.3.16)} \qquad \operatorname{dom} f^{\uparrow}(x_0;\cdot) \cap \operatorname{int} \operatorname{dom} g^{\uparrow}(x_0;\cdot) \cap \operatorname{int} \operatorname{dom} |h|^{\uparrow}(x_0;\cdot) \neq \phi$$

Then

$$\text{(5.3.17)} \qquad 0 \in \partial f(x_0) + c_1 \partial g^{+}(x_0) + c_2 \partial |h|(x_0).$$

If instead x_0 is a local minimizer for (P2), then

$$\text{(5.3.18)} \qquad 0 \in \partial f(x_0) + \exp(g^{+}(x_0) - 1)\partial g^{+}(x_0) + \exp(|h(x_0)| - 1)\partial |h|(x_0).$$

If f is strictly l.s.c. at x_0 and $E := \mathbb{R}^m$, then (5.3.16) and the directional Lipschitz hypotheses may be replaced by the assumption that $\operatorname{dom} f^{\uparrow}(x_0;\cdot)$, $\operatorname{dom} g^{\uparrow}(x_0;\cdot)$, and $\operatorname{dom} |h|^{\uparrow}(x_0;\cdot)$ are in strong general position.

We have not said anything in Proposition 5.3.10 about how to calculate $\partial g^+(x_0)$ and $\partial|h|(x_0)$. We can deduce more information about these subgradients by using [Ro 3, Theorem 4] or Proposition 2.3.14. First we examine $\partial g^+(x_0)$. There are three cases:

(i) If $g(x_0) > 0$ and g is l.s.c. at x_0, then $\partial g^+(x_0) = \partial g(x_0)$

(ii) If $g(x_0) < 0$ and g is u.s.c. at x_0, then $\partial g^+(x_0) = \{0\}$.

(iii) If $g(x_0) = 0$ and $\partial g(x_0) \neq \phi$, then $\partial g^+(x_0) \subset \{\lambda \partial g(x_0) \mid 0^+ \leq \lambda \leq 1\}$.

We can use the fact that $|h(x)| = \max\,(h(x), -h(x))$ to calculate information about $\partial|h|(x_0)$. Again there are three cases:

(i) If $h(x_0) > 0$ and h is l.s.c. at x_0, then $\partial|h|(x_0) = \partial h(x_0)$.

(ii) If $h(x_0) < 0$ and h is u.s.c. at x_0, then $\partial|h|(x_0) = \partial(-h)(x_0)$.

(iii) If $h(x_0) = 0$, h is directionally Lipschitzian at x_0, int dom $h^\dagger(x_0;\cdot) \cap \text{dom}(-h)^\dagger(x_0;\cdot) = \phi$, and $\partial h(x_0) \neq \phi$, then

$$\partial|h|(x_0) \subset \{\lambda_1 \partial h(x_0) + \lambda_2 \partial(-h)(x_0) \mid \lambda_1, \lambda_2 \geq 0^+, \lambda_1 + \lambda_2 = 1\}.$$

A combination of Proposition 5.3.10 and the above information gives more detailed necessary conditions for optimality in (P1) and (P2). One can also obtain corresponding results involving approximate subdifferentials or upper convex approximates. We leave the details to the reader.

We also leave aside the important questions of when a penalty function like (P1) or (P2) attains a local minimum at the same point as the associated constrained problem. Such penalty functions are called <u>exact</u>. If a penalty function is exact, one can derive optimality conditions for the constrained problem by studying the penalty function (see for example the first part of [I1]).

In [I1], Ioffe studies exact penalty functions of the form $g(G(x))$, where $G : E \to E_1$ is continuously differentiable and $g : E_1 \to \mathbb{R}$ is convex and positively homogeneous. Such penalty functions can be associated with smooth optimization problems with finitely many equality and inequality constraints. We conclude this section by giving necessary optimality conditions for a more general problem.

<u>Proposition 5.3.11</u>: Let $G : E \to E_1$ be strictly differentiable at x_0, and let $g : E_1 \to \bar{\mathbb{R}}$ be finite at $G(x_0)$, where x_0 is a local minimizer for $g \circ G$. Assume g is directionally Lipschitzian at $G(x_0)$ and

(5.3.19) range $\nabla G(x_0) \cap \text{int dom } g^{\uparrow}(G(x_0);\cdot) \neq \phi$.

Then there exists $T \in \partial g(G(x_0))$ such that $T\nabla G(x_0) = 0$. If $E := \mathbb{R}^n$, $E_1 := \mathbb{R}^m$, and g is strictly l.s.c. at x_0 , the same conclusion holds if (5.3.19) and the directionally Lipschitzian hypothesis are replaced by

(5.3.20) $\nabla G(x_0)\mathbb{R}^n - \text{dom } g^{\uparrow}(x_0;\cdot) = \mathbb{R}^m$.

Proof: This follows from Theorem 1.5.21, [Ro 3, Theorem 3], and Corollary 2.3.7. □

5.4 Isotone tangent cones and necessary conditions for optimality.

A brief reference was made in section 1.1 to a quick method of writing optimality conditions for (P) involving tangency operators which are isotone with respect to set inclusion. In this short section we develop this idea further, beginning with a special case.

Proposition 5.4.1: Let $f : E \to \bar{\mathbb{R}}$ be finite at $x_0 \in E$, where x_0 is a local minimizer for (P). Then

(5.4.1) $f^{+}(x_0;y) \geq 0$ for all $y \in K_C(x_0)$.

Proof: Let $y \in K_C(x_0)$. Since x_0 is a local minimizer for (P) , there exists $X \in N(x_0)$ such that $C \cap X \subset \{x \in E | f(x) \geq f(x_0)\}$. By the localization property

(Theorem 1.4.6), $y \in K_{C \cap X}(x_0)$, and since the contingent cone has property (3) (isotonicity), $y \in K_{\{x \in E | f(x) \geq f(x_0)\}}(x_0)$.

Now let $Y \in N(y)$ and $\lambda > 0$ be given. Since $y \in K_{\{x \in E | f(x) \geq f(x_0)\}}(x_0)$, there exist $y' \in Y$ and $t \in (0,\lambda)$ such that $\dfrac{f(x_0 + ty') - f(x_0)}{t} \geq 0$. Hence,

$$f^+(x_0;y) := \limsup_{\substack{y' \to y \\ t \downarrow 0}} \frac{f(x_0 + ty') - f(x_0)}{t} \geq 0 . \qquad \square$$

<u>Corollary 5.4.2</u>: Let $f : E \to \bar{\mathbb{R}}$ and $g : E \to \bar{\mathbb{R}}$ be finite at $x_0 \in E$, a local minimizer of $\min\{f(x) | g(x) \leq 0\}$. Assume that g is strictly l.s.c. at x_0, that $g(x_0) = 0$, and that $0 \notin \partial g(x_0)$. Then $f^+(x_0;y) \geq 0$ for all y such that $g_+(x_0;y) \leq 0$.

<u>Proof</u>: Let $C := \{x \in E | g(x) \leq 0\}$ in (P), and let $y \in E$ be such that $g_+(x_0;y) \leq 0$. By (5.2.4), $y \in K_C(x_0)$, so it follows from Proposition 5.4.1 that $f^+(x_0;y) \geq 0$. $\square$

Proposition 5.4.1 is a special case of the following more general principle:

<u>Theorem 5.4.3</u>: Suppose R_1 and R_2 are q-cones such that R_1 is isotone with respect to set inclusion and $R_2(C,x) = E\backslash R_1(E\backslash C,x)$ for all $(C,x) \in 2^E \times E$. Then if x_0 is a local minimizer for (P),

$$f^{R_2}(x_0;y) \geq 0 \quad \text{for all} \quad y \in R_1(C,x_0) .$$

Proof: Let $y \in R_1(C,x_0)$. Since x_0 is a local minimizer for (P), there exists $X \in N(x_0)$ such that $(C \cap X) \cap \{x \in E | f(x) < f(x_0)\} = \phi$. Since R_1 satisfies property (3c) of section 1.1,

$$R_1(C \cap X,x_0) \cap R_2(\{x \in E | f(x) < f(x_0)\},x_0) = \phi .$$

Now by Theorem 1.4.6, $y \in R_1(C \cap X,x_0)$, so $y \notin R_2(\{x \in E | f(x) < f(x_0)\},x_0)$. It is a straightforward matter to verify that this means that $f^{R_2}(x_0;y) \geq 0$, as asserted. □

Corollary 5.4.4: Suppose $f : E \to \bar{\mathbb{R}}$ is finite at x_0 , a local minimizer for (P) . Then

(5.4.2) $\quad f_+(x_0;y) \geq 0 \quad$ for all $\quad y \in L_C(x_0)$.

(5.4.3) $\quad f_\square(x_0;y) \geq 0 \quad$ for all $\quad y \in \ell_C(x_0)$.

(5.4.4) $\quad f^\ell(x_0;y) \geq 0 \quad$ for all $\quad y \in k_C(x_0)$.

(5.4.5) $\quad f^A(x_0;y) \geq 0 \quad$ for all $\quad y \in E_C(x_0)$.

(5.4.6) $\quad f^E(x_0;y) \geq 0 \quad$ for all $\quad y \in A_C(x_0)$.

Corollary 5.4.5: Under the hypotheses of Corollary 5.4.2, $f^\ell(x_0;y) \geq 0$ for all y such that $g_\square(x_0;y) < 0$.

Proof: This follows immediately from (5.4.4) and (5.2.3). □

For a given isotone q-cone R_1, there does not necessarily exist a complementary q-cone R_2 satisfying the hypotheses of Theorem 5.4.3. There will exist such an R_2 if $K(X) := x_0$ and $M(Y) := Y$ or $M(Y) := y$. These cases are listed in (5.4.1) through (5.4.6).

One interesting consequence of Theorem 5.4.3 is an optimality condition involving upper convex approximates.

<u>Theorem 5.4.6</u>: In addition to the hypotheses of Theorem 5.4.3, suppose that h is a R_2 upper convex approximate to f at x_0 and D is a closed convex subcone of $R_1(C,x_0)$. If either $\operatorname{dom} h \cap \operatorname{int} D \neq \phi$ or $\operatorname{int} \operatorname{dom} h \cap D \neq \phi$, then

$$(5.4.7) \qquad 0 \in \partial_h^{R_2} f(x_0) + D^0 .$$

<u>Proof</u>: By Theorem 5.4.3, we have for all $y \in E$ that

$$\begin{aligned} 0 &\leq f^{R_2}(x_0;y) + i_{R_1(C,x_0)}(y) \\ &\leq h(y) + i_D(y) . \end{aligned}$$

Hence $\qquad 0 \in \partial(h + i_D)(0)$

$\qquad\qquad \subset \partial h(0) + D^0$ by Theorem 2.1.1.

Therefore (5.4.7) holds. $\square$

Note that in order to apply Theorem 5.4.6, the function f and set C need to be such that the upper convex approximate h and convex cone D postulated in the theorem actually exist. This will not be the case if either $R_1(C, x_0) = \phi$ or $f^{R_2}(x_0;\cdot)$ is identically equal to $+\infty$. We leave aside here the question of conditions guaranteeing the existence of the h and D in Theorem 5.4.6 for various choices of R_1.

5.5. Vector optimization

In this section we study optimization problems in the vector-valued setting of Chapter 3. Let $f : E \to (\bar{E}_1, S)$, where again S is a closed convex cone which contains 0 and induces an order complete ordering on E_1, and let $C \subset E$. We consider the problem

$$\text{(PV)} \qquad \inf_S\{f(x) \mid x \in C\} .$$

A point at which the infimum in (PV) is attained is said to be <u>efficient</u> or <u>Pareto minimal</u>. We will begin our study of (PV) by examining a special case which can be handled as in section 5.1.

<u>Definition 5.5.1</u>: The point $x_0 \in C$ is a strong local minimizer for f with respect to C if there exists $X \in N(x_0)$ such that for all $x \in X \cap C$, $f(x_0) \le_S f(x)$.

For strong minimality we can prove an analogue of Theorem 1.5.21 and develop necessary conditions as in section 5.1.

Theorem 5.5.2: Suppose x_0 is a strong local minimizer for f with respect to E. Then

$$0 \in \partial_S^K f(x_0) .$$

Proof: Let $y \in E$ be given, and suppose $(y,d) \in K_f(x_0)$. There exists $X \in N(x_0)$ such that $f(x) \geq_S f(x_0)$ for all $x \in X$. Choose $Y \in N(y)$ and $\lambda > 0$ such that

$$x_0 + [0,\lambda_0)Y \subset X , \quad \text{and let} \quad U \in N(0)$$

in E_1 be given. Since $(y,d) \in K_f(x_0)$, there exist $t \in (0,\lambda)$, $y' \in Y$, and $u \in U$ such that

$$0 \leq_S \frac{f(x_0 + ty') - f(x_0)}{t} \leq_S d + u .$$

Thus for every $U \in N(0)$, $(d + U) \cap S \neq \phi$, i.e. $d \in \mathrm{cl}S = S$. Since $d \geq_S 0$ for every d such that $(y,d) \in K_f(x_0)$, we conclude that $f^K(x_0;y) \geq_S 0$. Therefore $0 \in \partial_S^K f(x_0)$. □

Corollary 5.5.3: Under the hypotheses of Theorem 5.5.2,

$$(5.5.1) \qquad 0 \in \partial_S^R f(x_0)$$

for any q-cone R which is always contained in the contingent cone.

In analogy with the scalar-valued case, we can define the indicator function of a set $C \subset E$ with respect to S by

$$i_C(x) := \begin{cases} 0 & \text{if} \quad x \in C \\ +\infty & \text{if} \quad x \in E\backslash C \end{cases}$$

Then if x_0 is a strong local minimizer for f with respect to C, x_0 is a strong local minimizer for $f + i_C$ with respect to E. By Corollary 5.5.3, we then have

$$0 \in \partial_S^R(f + i_C)(x_0)$$

for any R smaller than the contingent cone. If in addition a subgradient sum formula for ∂_S^R is applicable,

(5.5.2) $\quad 0 \in \partial_S^R f(x_0) + \partial_S^R i_C(x_0)$.

We now formulate conditions under which (5.5.2) holds at a strong local minimizer with $R := T$.

<u>Proposition 5.5.4</u>: Let S be normal, let $f : E \to (\bar{E}_1, S)$, and let $C \subset E$ be closed near x_0, a strong local minimizer for f with respect to C. Assume that f is l.s.c. at x_0, that $f^L(x_0;\cdot)$ is proper, and that there is a point $z \in \text{dom}\, f^I(x_0;\cdot) \cap T_C(x_0)$ such that $f^I(x_0;\cdot)$ is continuous at z. Then

(5.5.3) $\quad 0 \in \partial_S f(x_0) + N_C(x_0)$.

<u>Proof</u>: As in the scalar-valued case, $\text{dom}\, i_C{}^{\uparrow}(x_0;\cdot) = T_C(x_0)$ and $i_C{}^{\uparrow}(x_0;\cdot) = i_{T_C(x_0)}(\cdot)$, so $i_C{}^{\uparrow}(x_0;\cdot)$ is proper and $\partial_S i_C(x_0) = N_C(x_0)$. An application of [Th1, Proposition 3.9] yields (5.5.3). □

Specific cases of (5.5.3) may be calculated by means of subdifferential calculus formulae from Chapter 3. We will not go into further detail here.

The inclusion (5.5.1) will not hold in general for an efficient point that is not strongly minimal. An alternate strategy for tackling (PV) is to obtain efficient points for (PV) by finding minima of a <u>scalarization</u> - an associated scalar optimization problem. With this in mind, we next examine the concept of <u>proper efficiency</u>, introduced in [Bo 2].

<u>Definition 5.5.5</u>: Let R be a q-cone. A point $x_0 \in C$ is a (local) R proper efficient point for (PV) if it is efficient and

$$(5.5.4) \qquad R(f(C) + S, f(x_0)) \cap -S = \{0\} .$$

A thorough discussion of K proper efficiency is given in [Bo 2], with K defined sequentially as in (1.8.5). We will now study this concept for other tangent cones, in particular $k_C(x_0)$ and $T_C(x_0)$.

Observe first of all that if x_0 is a R proper efficient point, it is also a R' proper efficient point for any R' which contains $\{0\}$ and is always contained in R. For example, any K proper efficient point is also k proper efficient and T proper efficient. Not all T proper efficient or k proper efficient points are K proper efficient, however.

Example 5.5.6: (a) Let $E = E_1 := \mathbb{R}^2$, $S := \mathbb{R}^2_+$, $f := I$, and $C := \{(x,y) \in \mathbb{R}^2 | y \geq x^2 \text{ or } x \geq y^2\}$ in (PV). Then $K_{C+S}(0,0) = k_{C+S}(0,0) = \{(x,y) \in \mathbb{R}^2 | y \geq 0 \text{ or } x \geq 0\}$, so $(0,0)$ is not K or k proper efficient. However, $T_{C+S}(0,0) = \{(x,y) \in \mathbb{R}^2 | x \geq 0, y \geq 0\}$, so $(0,0)$ is T proper efficient.

(b) Let $E = E_1 := \mathbb{R}^2$, $S := \{(x,y) | y \geq x/2, x \leq 0\}$, $f := I$, and $C := \text{epi } g$, where $g : \mathbb{R} \to \mathbb{R}$ is defined by

$$g(x) := \begin{cases} 0 & \text{if } x \leq 0 \\ \frac{1}{2^{n+1}} & \text{if } \frac{1}{2^{n+1}} \leq x < \frac{1}{2^n}, \quad n = 0, \pm 1, \pm 2, \ldots \end{cases}$$

Then $T_{C+S}(0,0) = \{(x,y) | x \leq 0, y \geq 0\}$ and $k_{C+S}(0,0) = \{(x,y) | y \geq \max(x, \frac{x}{2})\}$, so $(0,0)$ is T and k proper efficient. On the other hand, $K_{C+S}(0,0) = \{(x,y) | y \geq \frac{x}{2}\}$, so $(0,0)$ is not K proper efficient.

Following [Bo 2], we can readily formulate a scalarization result.

Theorem 5.5.7: (cf. [Bo 2, Theorem 1]): Suppose S is such that there exists $s^+ \in E_1'$ satisfying

$$(5.5.5) \quad s^+(s) > 0 \text{ for all } s \in S \backslash \{0\},$$

and suppose that x_0 is a minimizer for

$$(P(s^+)) \quad \min\{s^+ f(x) | x \in C\}.$$

Then if R is a q-cone which always contains 0 and is always contained in the contingent cone, x_0 is R proper efficient for (PV).

<u>Proof</u>: It suffices to show that x_0 is K proper efficient. Let $y \in K_{f(C)+S}(f(x_0))$, and let $Y \in N(y)$ be given. Then for all $\lambda > 0$, there exist $t \in (0,\lambda)$ and $y' \in Y$ with $f(x_0) + ty' \in f(C) + S$, so $ty' \in f(C) - f(x_0) + S$. Then for any s^+ satisfying (5.5.5), $s^+(y') \geq 0$. Now if y were in $-S\setminus\{0\}$, there would exist $Y' \in N(y)$ with $s^+(Y') < 0$. Therefore

$K_{f(C)+S}(f(x_0)) \cap -S = \{0\}$. □

In [Bo 2], two equivalent criteria are given which guarantee that (5.5.5) holds:

(5.5.6) (i) $S \cap -S = \{0\}$ and S is locally compact.

(ii) The dual cone S^+ has nonempty-interior in some topology which gives E_1 as the dual of E_1'.

Further conditions on R, f, and C are needed for a converse of Theorem 5.5.7 to hold.

<u>Theorem 5.5.8</u>. (cf. [Bo 2, Theorem 2]). Suppose f is S-convex, C is convex, R is closed and has property (1), and either (5.5.6) (i) or (5.5.6) (ii) is satisfied. Then if x_0 is R proper efficient for (PV), x_0 is

optimal for $(P(s^+))$ for some s^+ satisfying (5.5.5).

<u>Proof:</u> The set $f(C) + S$ is convex, so since R has property (1) and is closed.

$$(5.5.7) \quad f(C) + S - f(x_0) \subset R(f(C) + S, f(x_0)) =: N .$$

The set N is closed and convex, and since x_0 is properly efficient, $N \cap -S = \{0\}$. It is shown in [Bo 2] that if either (5.5.6) (i) or (ii) holds, there exists s^+ satisfying (5.5.5) and $s^+(N) \geq 0$. In particular, we have by (5.5.7) that $s^+(f(x) + S - f(x_0)) \geq 0$ for all $x \in C$. Since $0 \in S$, it follows that $s^+f(x) \geq s^+f(x_0)$ for all $x \in C$. Therefore x_0 is optimal for $(P(s^+))$. □

Tangent cones R which satisfy the hypotheses of Theorems 5.5.7 and 5.5.8 include $K_C(x_0)$, $k_C(x_0)$, and $T_C(x_0)$.

Theorem 5.5.7 shows we can find K, k, or T proper efficient points for (PV) by solving $(P(s^+))$. We now apply the results of Section 5.1 to give a necessary optimality condition for $(P(s^+))$.

<u>Theorem 5.5.9:</u> Let (E_1, S) be such that there exists $s^+ \in E_1'$ satisfying (5.5.5). Let $f : E \to (\bar{E}_1, S)$ and $C \subset E$, and suppose x_0 is a minimizer for $(P(s^+))$. Assume C is epi – Lipschitzian at x_0 and

$$(5.5.8) \quad \operatorname{dom}(s^+f)^{\uparrow}(x_0;\cdot) \cap \operatorname{int} T_C(x_0) \neq \phi .$$

Then

$$(5.5.9) \qquad 0 \in \partial(s^+f)(x_0) + N_C(x_0) .$$

If $E := \mathbb{R}^m$, (5.5.9) holds if (5.5.8) and the epi-Lipschitzian hypothesis are replaced by the assumption that s^+f is strictly l.s.c. at x_0, C is closed near x_0, and

$$(5.5.10) \qquad \operatorname{dom}(s^+f)^\uparrow(x_0;\cdot) - T_C(x_0) = \mathbb{R}^m .$$

<u>Proof:</u> Apply Theorems 5.1.1 and 5.1.2. Inclusion (5.1.4) in this case is (5.5.9). □

<u>Theorem 5.5.10:</u> Suppose S is Daniell and (E_1,S) admits a functional s^+ satisfying (5.5.5). Let $f : E \to (E_1,S)$ be continuous and T epigraph regular at x_0, a minimizer of $(P(s^+))$. Assume that $C \subset E$ is epi-Lipschitzian at x_0, that $f^\uparrow(x_0;0) = 0$ and $f^\uparrow(x_0;\cdot)$ is continuous at 0, and that

$$(5.5.11) \qquad \operatorname{dom} f^\uparrow(x_0;\cdot) \cap \operatorname{int} T_C(x_0) \neq \phi .$$

Then

$$(5.5.12) \qquad 0 \in s^+\partial_S f(x_0) + N_C(x_0) .$$

If $E := \mathbb{R}^m$, (5.5.12) holds if (5.5.11) and the epi-Lipschitzian hypothesis are replaced by the assumption that C is closed near x_0 and (5.1.6) is satisfied.

Proof: Our hypotheses are sufficient to guarantee that $(s^+f)^\uparrow(x_0;\cdot) \leq s^+f^\uparrow(x_0;\cdot)$ by Theorem 3.2.6, so (5.5.8) can be replaced by (5.5.11) and (5.5.10) can be replaced by (5.1.6). By Corollary 3.2.17, $\partial(s^+f)(x_0) \subset s^+\partial_S f(x_0)$, so (5.5.12) holds. □

We close this section with a comment on differentiable Pareto optimization. In [Bo 2], Borwein examines the special case of (PV) in which E and E_1 are normed spaces, $f : E \to (E_1,S)$ is Frechet differentiable, and $C := \{x \in E | g(x) \in B, x \in D\}$, where $g : E \to (E_1,S)$ is Frechet differentiable, $B \subset E_1$, and $D \subset E$. He develops Guignard-type necessary conditions for K proper efficient points of (PV) in this case. We will not study those conditions here, but we do remark in passing that section 6 of [Bo 2] gives an analogous result for k proper efficient points if the contingent cone is simply replaced by k throughout the section. A similar result for T proper efficient points will not hold without the additional assumptions that f is strictly differentiable at x_0 and (1.4.17) holds.

5.6 Directions for further research.

In this chapter, we have surveyed some of the applications of subdifferential calculus in optimization theory. There are additional important topics that have not been considered here, including

(1) specific vector optimization problems, like the least element problem and order complementarity problem (see [Bo 4] for definitions and references).

(2) stability of perturbed optimization problems, including results on directional derivatives and subgradients of marginal functions of mathematical programs (e.g. [Bo 5], [Ro 6]).

(3) higher-order optimality conditions (e.g. [Bo 5], [Il]).

It is hoped that the research in this thesis can lend new insights to research in these areas.

Much more can and should be done to elucidate the meaning, capabilities, and limitations of the results presented in these five chapters. The study of specific nonsmooth functions and optimization problems should prove to be fruitful in this regard.

[A1] J.-P. Aubin, "Contingent derivatives of set-valued maps and existence of solutions to non-linear inclusions and differential inclusions," Math. Anal. Appl., Advances in Math., Suppl. Studies 7A (1981) 159-229.

[A2] J.-P. Aubin, "Ioffe's fans and generalized derivatives of vector-valued maps", in: J.-P. Aubin and R.B. Vinter, eds., Convex Analysis and Optimization, Pitman, London, 1982, pp. 2-17.

[A3] J.-P. Aubin, "Lipschitz behavior of solutions to convex minimization problems", Mathematics of Operations Research 9 (1984) 87-111.

[Ba1] M.S. Bazaraa, J.J. Goode, M.F. Nashed and C.M. Shetty, "Nonlinear programming without differentiability in Banach Spaces. Necessary and sufficient constraint qualifications", Applicable Analysis 5 (1976) 165-173.

[Bo1] J.M. Borwein, "Fractional programming without differentiability", Mathematical Programming 11 (1976) 283-290.

[Bo2] J.M. Borwein, "Proper efficient points for maximizations with respect to cones", SIAM J. Cont. Opt. 15 (1977) 57-63.

[Bo3] J.M. Borwein, "Weak tangent cones and optimization in a Banach space", SIAM J. Cont. Opt. 16 (1978) 512-522.

[Bo4] J.M. Borwein, "Continuity and differentiability properties of convex operators", Proc. London Math. Society 44 (1982) 420-444.

[Bo5] J.M. Borwein, "Stability and regular points of inequality systems", to appear in J.O.T.A.

[Bo6] J.M. Borwein, "A tangent cone separation principle", Dalhousie University research report 84-03.

[Bo7] J.M. Borwein and H.M. Strojwas, "Directionally Lipschitzian mappings on Baire spaces", Can. J. Math. 36 (1984) 95-130.

[Bo8] J.M. Borwein and H.M. Strojwas, "Tangential approximations", preprint, 1984.

[Bo9] J.M. Borwein, "A Lagrange multiplier theorem and a sandwich theorem for convex relations", Math. Scand. 48 (1981) 189-204.

[Bo10] J.M. Borwein, "Subgradients of convex operators", Dalhousie University preprint, 1983.

[Ch1] C. Charalambous, "On conditions for optimality of a class of nondifferentiable functions", J.O.T.A. 43 (1984) 135-142.

[Cl1] F.H. Clarke, "A new approach to Lagrange multipliers", Math. of O.R. 1 (1976) 165-174.

[Cl2] F.H. Clarke, "On the inverse function theorem", Pac. J. Math. 64 (1976) 97-102.

[Cl3] F.H. Clarke, "Generalized gradients of Lipschitz functionals", Adv. Math. 40 (1981) 52-67.

[Cl4] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.

[Cr1] J.-P. Crouzeix, Jacques A. Ferland, and Siegfried Schaible, "Duality in generalized linear fractional programming", Mathematical Programming 27 (1983) 342-354.

[Da1] J. Danes, "A geometric theorem useful in nonlinear functional analysis", Boll. Un. Mat. Ital. 6 (1982) 369-375.

[Do1] S. Dolecki, "Tangency and differentiation: some applications of convergence theory", Annali di Matematica pura ed applicata 130 (1982) 223-255.

[E1] I. Ekeland, "On the variational principle", J. Math. Anal. Appl. 47 (1974) 324-353.

[E2] I. Ekeland and R. Temam, Convex Analysis and Variational Problems, North Holland, Amsterdam, 1976.

[Ga1] S. Gautier, G. Isac, and J.-P. Penot, "Surjectivity of multifunctions under generalized differentiability assumptions", Bull. Australian Math. Society 28 (1983) 13-21.

[Gu1] M. Guignard, "Generalized Kuhn-Tucker conditions for mathematical programming in a Banach space", SIAM J. Cont. 7 (1969) 232-241.

[Hi1] J.-B. Hiriart-Urruty, Contributions à la Programmation Mathématique: Deterministe et Stochastique, These Université de Clermont-Ferrand II, 1977.

[Hi2] J.-B. Hiriart-Urruty, "Tangent cones, generalized gradients and mathematical programming in Banach spaces", Mathematics of Operations Research 4 (1979) 79-97.

[Hi3] J.-B Hiriart-Urruty, "Mean value theorems in nonsmooth analysis", Numerical Functional Analysis and Optimization 2 (1980) 1-30.

[Hi4] J.-B. Hiriart-Urruty, "Characterizations of the plenary hull of the generalized Jacobian matrix", Mathematical Programming Study 17 (1982) 1-12.

[Ho1] R.B. Holmes, Geometric Functional Analysis and its Applications, Springer-Verlag, New York, 1975.

[I1] A.D. Ioffé, "Necessary and sufficient conditions for a local minimum", SIAM J. Cont. Opt. 17 (1979) 245-288.

[I2] A.D. Ioffe, "Approximate subdifferentials and applications I: the finite-dimensional theory", Trans. Am. Math. Soc. 281 (1984) 389-416.

[Ku1] S.S. Kutateladze, "Formulas for computing subdifferentials", Soviet Math. Dokl. 18 (1977) 146-148.

[Ku2] S.S. Kutateladze, "Convex operators", Russian Math. Surveys 34 (1979) 181-214.

[Le1] G. Lebourg, "Generic differentiability of Lipschitzian functions", Trans. Am. Math. Soc. 256 (1979) 125-144.

[Ma1] D.H. Martin, R.J. Gardner, and G.G. Watkins, "Indicating cones and the intersection principle for tangential approximants in abstract multiplier rules", J.O.T.A. 33 (1981) 515-537.

[Ma2] D.H. Martin and G.G. Watkins, "Cores of tangent cones and Clarke's tangent cone", NRIMS Technical Report TWISK 304, 1983.

[Pa1] N.S. Papageorgiou, "Nonsmooth Analysis on partially ordered vector spaces: Part 1 - convex case", Pac. J. Math. 107 (1983) 403-458.

[Pa2] N.S. Papageorgiou, "Nonsmooth Analysis on partially ordered vector spaces: Part 2 - nonconvex case, Clarke's theory", Pac. J. Math. 109 (1983) 463-499.

[Pe1] J.-P. Penot, "A characterization of tangential regularity", Nonlinear Analysis: Theory, Methods, and Applications 5 (1981) 625-641.

[Pe2] J.-P. Penot, "Differentiability of relations and differential stability of perturbed optimization problems", SIAM J. Cont. Opt. 22 (1984) 529-551.

[Pe3] J.-P. Penot, "The Clarke's tangent cone and limits of tangent cones", preprint, 1984.

[Per1] A.L. Peressini, Ordered topological vector spaces, Harper and Row, New York, 1967.

[Ps1] B.N. Pshenichnyi and R.A. Hacatrjan, "Constraints of equality type in nonsmooth optimization problems", Soviet. Math. Dokl. 26(1982) 659-662.

[Ro1] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, N.J., 1970.

[Ro2] R.T. Rockafellar, "Clarke's tangent cones and the boundaries of closed sets in $\mathbb{R}^n$", Nonlinear Analysis: Theory, Methods and Applications 3 (1979) 145-154.

[Ro3] R.T. Rockafellar, "Directionally Lipschitzian functions and subdifferential calculus", Proc. London Math. Soc. 39 (1979) 331-355.

[Ro4] R.T. Rockafellar, "Generalized directional derivatives and subgradients of nonconvex functions", Can. J. Math. 32 (1980) 157-180.

[Ro5] R.T. Rockafellar, The Theory of Subgradients and its Applications to Problems of Optimization: Convex and Nonconvex Functions, Heldermann Verlag, Berlin, 1981.

[Ro6] R.T. Rockafellar, "Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming", Mathematical Programming Study 17 (1982) 28-66.

[Ro7] R.T. Rockafellar, "Extensions of subgradient calculus with applications to optimization", to appear in Nonlinear Analysis: Theory, Methods and Applications.

[Th1] L. Thibault, "Subdifferentials of nonconvex vector-valued functions", J. Math. Anal. Appl. 86 (1982) 319-344.

[Th2] L. Thibault, "Tangent cones and quasi-interiorly tangent cones to multifunctions", Trans. Am. Math. Soc. 277 (1983) 601-621.

[Th3] L. Thibault, "Subdifferentials of compactly Lipschitzian vector-valued functions, Annali di Matematica pura ed applicata 125 (1980) 157-192.

[Tr1] J.S. Treiman, "Characterization of Clarke's tangent and normal cones in finite abd infinite dimensions", Nonlinear Analysis: Theory, Methods, and Applications 7 (1983) 771-783.

[U1] C. Ursescu, "Tangent sets' calculus and necessary conditions for extremality", SIAM J. Cont. Opt. 20 (1982) 563-574.

[V1] M. Vlach, "On necessary conditions of optimality in linear spaces", Commentationes Mathematicae Universitatis Carolinae 11 (1970) 501-513.

[V2] M. Vlach, "On the cones of tangents", Methods of Operations Research 37 (1980) 251-256.

[V3] M. Vlach, "Approximation operators in optimization theory", Zeitschrift für O.R. 25 (1981) 15-23.

[V4] M. Vlach, "Closures and neighbourhoods properties induced by tangential approximations", in: G. Hammer and D. Pallaschke, eds. Selected Topics in Operations Research and Mathematical Economics, Springer-Verlag, Berlin, 1984, pp. 119-127.

[W1] G.G. Watkins, "Nonsmooth Milyutin-Dubovitskii theory and Clarke's tangent cone", submitted to Math. of O.R.

[Za1] C. Zalinescu, "On convex sets in general position", to appear in Linear Algebra and its Applications.

[Zo1] J. Zowe, "A duality theorem for a convex programming problem in order complete vector lattices", J. Math Anal. Appl. 50 (1975) 273-287.