Internat. J. Math. & Math. Sci.
VOL. 15 NO. 1 (1992) 65-82

NONSMOOTH ANALYSIS AND OPTIMIZATION ON PARTIALLY ORDERED VECTOR SPACES

THOMAS W. REILAND

Department of Statistics
and
Graduate Program in Operations Research
Box 8203
North Carolina State University
Raleigh, NC 27695-8203

(Received February 21, 1991 and in revised form July 16, 1991)

ABSTRACT. Interval-Lipschitz mappings between topological vector spaces are defined and compared with other Lipschitz-type operators. A theory of generalized gradients is presented when both spaces are locally convex and the range space is an order complete vector lattice. Sample applications to the theory of nonsmooth optimization are given.

KEY WORDS AND PHRASES. Interval-Lipschitz Mapping, Subdifferential, Optimality Conditions.

1980 AMS SUBJECT CLASSIFICATION CODE. Primary: 49B27. Secondary: 90C48.

1. INTRODUCTION.

The purpose of this paper is to introduce a broad class of Lipschitz-type operators and to present new results concerning first-order optimality conditions for nonsmooth nonconvex programs in infinite dimensions.

Significant progress in deriving more general optimality conditions for mathematical programming models has been made in recent years as a result of advances in nonsmooth analysis and optimization. The study of nonsmooth problems is motivated in part by the desire to optimize increasingly sophisticated models of complex man-made and naturally occurring systems that arise in areas ranging from economics, operations research, and engineering design to variational principles that correspond to partial differential equations. Results in nonsmooth optimization have expedited understanding of the salient aspects of the classic smooth theory and identified concepts fundamental to optimality that are not intertwined with differentiability assumptions. We mention as examples in this regard the works of Hiriart-Urruty [1], where the convexity of a tangent cone is required for optimality in the nonsmooth case but not when differentiability is assumed, and Clarke [2] where standard assumptions in optimal control are weakened.

First-order optimality conditions have received the most scrutiny and in general are well-understood. In terms of first principles they require, for example, that two problem-specific sets be nonintersecting or that a certain map not be locally surjective. Smoothness is not a fundamental prerequisite for these properties to hold. Analysis serves as the link between the above mentioned conditions and their equivalent expression in useable and verifiable algebraic forms. Research in nonsmooth analysis is motivated in part by the attitude that the essentials of optimality are sufficiently amenable and extensive to allow their application to nondifferentiable (and nonconvex) problems, provided an appropriate analysis is developed.

This paper makes a contribution to nonsmooth analysis and optimization based on these ideas. Our approach and subsequent results, while new in many respects, continue the work of others in extending the applicability of differential calculus. For example, generalized derivatives are defined in the well-known theory of distributions; however, these derivatives are of little use in optimization since their values are often not well-defined at local extrema.

The systematic development of nonsmooth analysis began in the late 1960's and early 1970's. Initial results by Rockafellar [3-7] , Moreau [8] and McLinden [9] dealt with convex, concave, and convex-concave functions. Valadier [10], Ioffe-Levin [11], Zowe [12, 13], Kutateladze [14], Rubinov [15], Borwein [16] and Papageorgiou [17] made important generalizations to convex mappings into ordered vector spaces. However, there is no general agreement on exactly what to do except in the convex case. The "quasidifferentials" of Pshenichnyi [18], "$\geq$ -gradients" of Bazaara, Goode and Nashed [19], "subdifferentials" of Penot [20] and the "derivative containers" of Warga [21] marked the initial thrusts into the nonconvex, nonsmooth setting. Clarke [2, 22-25] introduced a generalized gradient for nonconvex functions whose analytical virtues were recognized from the outset. His approach, like our approach in this paper, is essentially a "convexifying" process utilizing properties inherent in the function rather than that of assuming the existence of convex and/or linear approximations.

Since the initial contribution of Clarke, the theory and applications of generalized gradients has grown to such an extent that a survey is beyond the scope of this introduction. For excellent summaries of the theory, motivation and applications of generalized gradients and extensive references we refer the reader to Clarke [2], Hiriart-Urruty [1] and Rockafellar [26]; in addition, Borwein and Strojwas [27] provide an insightful comparison of several recent directional derivatives and generalized gradients of the same genre as Clarke's gradient. The excellent papers by Papageorgiou [17, 28] and Ioffe [29, 30] provide many fundamental results in nonsmooth analysis for vector-valued mappings.

We conclude this section with a brief summary of the main results. In Section 2 we introduce interval-Lipschitz mappings and show that several other classes of mappings introduced in the context of nonsmooth analysis and/or optimization, such as strictly differentiable mappings, the Lipschitz operators of Kusraev [31] and Papageorgiou [28], the order-Lipschitz mappings of Reiland [32, 33], convex mappings, and sublinear mappings are special cases of interval-Lipschitz mappings. In Section 3 we define and exhibit properties of a generalized directional derivative and subdifferential and make comparisons with several other directional derivatives and subdifferentials in the literature. We establish optimality conditions in Section 4 and relate these to other optimality conditions involving Lipschitz operators and quasidifferentiable functions. A distinguishing feature of our optimality conditions is that they allow for an infinite-dimensional equality constraint. Ioffe [30] obtains results for problems in Banach spaces with an infinite-dimensional Lipschitz equality constraint operator or finitely many directionally Lipschitzian equality constraint functions.

2. INTERVAL-LIPSCHITZ MAPPINGS.

Unless specified otherwise, in this section X and V denote, respectively, a linear topological space and an ordered topological vector space. We will denote the zero elements of X and V by θ. We will occasionally make the assumption that the positive cone $V_+ := \{v \in V: v \geq \theta\}$ is normal, that is, there is a neighborhood base $\mathcal{W}$ of the origin $\theta \in V$ such that, for $W \in \mathcal{W}$, $W = (W+V_+) \cap (W-V_+)$. Such neighborhoods are said to be full or saturated. Several consequences of normality utilized in the sequel can be found in Peressini [34]. We will always make explicit mention of this assumption when it is being used.

DEFINITION 1. The mapping $f: X \to V$ is interval-Lipschitz at $\bar{x} \in X$ if there exists neighborhoods N of $\bar{x}$ and W of $\theta \in X$, $\epsilon > 0$, two mappings m and M from W into V satisfying $m(y) \leq M(y)$, and a mapping r from $(0,\epsilon] \times X \times X$ into V satisfying $\lim_{\substack{t\downarrow 0\\ x\to\bar{x}}} r(t,x;y) = \theta$ for all $y \in W$, such that

$$t^{-1}[f(x+ty) - f(x)] \in [m(y), M(y)] + r(t,x;y)$$

for all $x \in N$, $y \in W$ and $t \in (0,\epsilon]$. If U is an open subset of X, f is locally interval-Lipschitz on U if f is interval-Lipschitz at $\bar{x}$ for every $\bar{x} \in U$.

If X is a normed space, V=R, and f is Lipschitz at $\bar{x} \in X$ in the usual sense, that is, there exist a neighborhood N_0 of $\bar{x}$ and $k \in R^+$ such that $|f(x) - f(y)| \leq k\|x-y\|$ for all x, $y \in N_0$, then f is interval-Lipschitz at $\bar{x}$. Indeed, select a neighborhood N of $\bar{x}$ and a circled neighborhood W of $\theta \in X$ such that $N + W \subseteq N_0$; then for $x \in N$, $y \in W$ and $t \in (0,1]$, $|f(x+ty)-f(x)| \leq tk\|y\|$ and the choices $m(y) = -k\|y\|$, $M(y) = k\|y\|$, $r\equiv 0$ show that f is interval-Lipschitz at $\bar{x}$. Below we provide additional sample classes of operators that are interval-Lipschitz.

EXAMPLE 1. For X a Banach space and V an order complete Banach lattice, Papageorgiou [28] defines a mapping $f: X \to V$ to be locally o-Lipschitz if for every open bounded subset U of X there is a $k \in V_+ := \{v \in V: v \geq 0\}$, the positive cone of V, such that $|f(x) - f(z)| \leq k\|x-z\|$ for all x, $z \in U$. If f is locally o-Lipschitz and U is an open bounded subset of X, then f is locally interval-Lipschitz on U. Indeed, if $\bar{x} \in U$, choose a neighborhood N of $\bar{x}$ and a circled neighborhood W of $\theta \in X$ such that $N+W \subseteq U$. Then for $x \in N$, $y \in W$, and $t \in (0,1]$, we have $|f(x+ty) - f(x)| \leq kt\|y\|$; the same choices for $m(\cdot)$, $M(\cdot)$, and r as in the preceding paragraph show that f is interval-Lipschitz at $\bar{x}$. Since $\bar{x} \in U$ was arbitrary, f is locally interval-Lipschitz on U.

EXAMPLE 2. If X is a normed vector space, $f: X \to V$ is strictly differentiable at $\bar{x} \in X$ if there exists a continuous linear mapping $\nabla f(\bar{x}): X \to V$ such that

$$\lim_{\substack{x\to\bar{x}\\ z\to\bar{x}\\ x\neq z}} [f(x) - f(z) - \nabla f(\bar{x})(x-z)]/\|x-z\| = 0 .$$

If we choose $m(y) = M(y) = \nabla f(\bar{x})y$ and $r(t,x;y) = t^{-1}[f(x+ty)-f(x) - t\nabla f(\bar{x})y]$, then $\lim_{\substack{t\downarrow 0\\ x\to\bar{x}}} r(t,x;y) = 0$ and f is interval-Lipschitz at $\bar{x}$.

EXAMPLE 3. If $f:X \to V$ is sublinear (i.e., subadditive and positively homogeneous), then f is interval-Lipschitz on X. In fact, if u and z are in X, then by the sublinearity of f, $f(u) - f(z) \leq f(u - z)$ and $-f(z-u) \leq f(u) - f(z)$. Thus, for x and y in X and $t > 0$, $-f(-ty) \leq f(x+ty) - f(x) \leq f(ty)$ and the choices $r\equiv 0$, $m(y) = -f(-y)$, $M(y) = f(y)$ show that f is interval-Lipschitz.

EXAMPLE 4A. If V is a vector lattice, Kusraev [31] defines a mapping $f: X \to V$ to be Lipschitz at $\bar{x}$ in X if there exists a neighborhood N_0 of $\bar{x}$ and a continuous monotone sublinear operator $P: X \to V$ such that $|f(u) - f(v)| \leq P(u-v)$ for all u,v in N_0. Let N be a neighborhood of $\bar{x}$ and W a circled neighborhood of θ in X such that $N + W \subseteq N_0$. Then the sublinearity of P and the choices $m(y) = -P(y)$, $M(y) = P(y)$ and $r\equiv 0$ show that f is interval-Lipschitz at $\bar{x}$.

EXAMPLE 4B. If X is a Banach space, then the inequality in Kusraev's definition of a Lipschitz mapping f at $\bar{x}$ in Example 4a can be stated as $|f(u) - f(v)| \leq k\|u-v\|$ for all $u,v \in N_0$ and for some $k \in V_+$. These Lipschitz mappings are equivalent to the subclass of interval-Lipschitz mappings, called order-Lipschitz mappings, on the Banach space X where $m(y) = v_1$, $M(y) = v_2$, and $r(\cdot,\cdot;y) = 0$ for all $y \in W$. Indeed, if f is Lipschitz at $\bar{x}$ according to Kusraev, then choosing neighborhoods N of $\bar{x}$ and W of θ in X such that $N + W \subseteq N_0$ and selecting $m(y) = -k$, $M(y) = k$, and $r\equiv 0$ shows that f is order-Lipschitz at $\bar{x}$. Conversely, suppose f is order-Lipschitz at $\bar{x}$ with $m(y) = v_1$, $M(y) = v_2$, and $r(\cdot,\cdot;y) = 0$ for all $y \in W$. Let the real number $\rho > 0$ be such that $B(\bar{x},2\rho) := \{x\in X: \|\bar{x}-x\|<2\rho\} \subseteq N$, $B(\theta,2\rho) \subseteq W$ and choose $\sigma > 0$ such that $\rho^{-1}\sigma < \epsilon$. Then for all $x,y \in B(\bar{x},\sigma)$ we have $f(y) - f(x) = f(x+\rho^{-1}\|y-x\|\cdot\rho((y-x)/\|y-x\|) - f(x) \in \rho^{-1}\|y-x\|[v_1,v_2]$, if $x \neq y$; since $\rho^{-1}\|y-x\| < \epsilon$ and $\rho(y-x)/\|y-x\| \in W$, $|f(y) - f(x)| \leq k\|y-x\|$ for all $x,y \in B(\bar{x},\sigma)$, where $k=\rho^{-1}(|v_1| + |v_2|) \in V_+$, and thus f is Lipschitz at $\bar{x}$ according to Kusraev.

REMARK. If X is a Banach space, V is an order complete Banach lattice and $f: X \to V$ is locally o-Lipschitz according to Papageorgiou [28] (see Example 1), then if int $V_+ \neq \emptyset$, f is Lipschitz at $\bar{x}$ according to Kusraev for any $\bar{x} \in X$. Indeed, let v_0 be in the interior of V_+; then $[-v_0, v_0] + \bar{x}$ is a (convex) neighborhood of $\bar{x}$ and is (topologically) bounded since the normality of V_+ implies that order bounded sets are topologically bounded (Peressini [34, p. 62].

The next example shows that an interval-Lipschitz mapping is not necessarily continuous.

EXAMPLE 5. Let (c) be the space of all convergent sequences of real numbers with norm $\|x\|_\infty = \sup\{|x_n|\}$ and let W be an open bounded neighborhood of $\theta \in (c)$ relative to the topology $\sigma((c), \ell^1)$, i.e., the weak topology on (c). Since ℓ^1 is the dual of (c), ℓ^1 is norm-determining for (c) (Taylor [35, p. 202]), hence by Taylor [35, p. 208] W is bounded relative to the norm topology. In particular, W is absorbed by $B = \{x: \|x\| < 1\}$, thus there exists $\lambda_0 > 0 \ni \lambda W \subseteq B$ for all $|\lambda| \leq \lambda_0$. Let $W_0 = \lambda_0 W$; then W_0 is order bounded since $B = \{x: \|x\| \leq 1\}$ coincides with $[-e,e]$ in (c), where $e = (e_n)$, $e_n = 1$ for all n. Therefore, since $f: (c) \to (c)$ given by $f(x) = |x|$ is sublinear, for any $x \in (c)$ and $y \in W_0$ we have

$$\lambda^{-1}[f(x+\lambda y) - f(x)] \leq f(y) = |y| \in [-e,e] ,$$

which shows that $f(x) = |x|$ is interval Lipschitz on (c). However, $f(x)$ is not continuous since the dual of (c) is not the sequence space $\varphi = \{x=(x_n):x_n = 0$ for all but a finite number of choices of $n\}$ (Peressini [34, p. 135]).

The following example shows that convex mappings are interval-Lipschitz.

EXAMPLE 6. Let X and V be as in Example 1. The mapping $f: X \to V$ is convex if $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ for all $\lambda \in [0,1]$ and $x,y \in X$. If f is convex and majorized in a neighborhood of $x_0 \in X$, then by Theorem 3.2 in Papageorgiou [17] and Example 1, f is interval-Lipschitz on X.

We conclude this section with a brief comparison of interval-Lipschitz mappings and two similar Lipschitz-type operators proposed by Thibault [36]. Unless specified otherwise, X and V are linear topological vector spaces. Thibault [36] defines a compactly Lipschitzian mapping at a point as follows: $f:X \to V$ is compactly Lipschitzian at $\bar{x} \in X$ if there is mapping $K:X \to \text{Comp}(V):=$ {nonempty compact subsets of V} and a mapping $r:(0,1] \times X \times X$ into V such that

(i) $\lim\limits_{\substack{t \downarrow 0 \\ x \to \bar{x}}} r(t,x;y) = 0$ for each $y \in X$;

(ii) for each $y \in X$ there is a neighborhood Ω of $\bar{x}$ and $\eta \in (0,1]$ such that

$$t^{-1}[f(x+ty)-f(x)] \in K(y) + r(t,x;y) \text{ for all } x \in \Omega \text{ and } t \in (0,\eta] \ .$$

This definition does not require the range space to be ordered as in Definition 1 and hence in this respect can be considered more general than our definition. However, the approach taken in this paper and in Thibault [36] (and in many other works as well) to derive a theory of generalized gradients requires that the range space be ordered. In this case, Definition 1 takes explicit account of the order structure. In addition, the order interval $[m(y), M(y)]$ is in general not compact. If V is normal, then the order interval $[m(y), M(y)]$ is bounded and hence by Alaoglu's Theorem is ω^*-compact if V is a dual space; however, it is in general not compact for any other stronger topology. From this viewpoint, Definition 1 can be considered somewhat more general than Thibault's definition.

For a mapping $f: X \to V$, V an ordered topological vector space, Thibault [36] defines f to be order-Lipschitz at a point $\bar{x} \in X$ as follows: there exist mappings $\underline{h}$ and $\bar{h}$ of X into V and a mapping $r:(0,1] \times X \times X \to V$ such that

(i) $\underline{h}(x) \leq \bar{h}(x)$ for all $x \in X$ and $\lim\limits_{x \to \theta} \bar{h}(x) = \theta$;

(ii) $\lim\limits_{\substack{t \downarrow 0 \\ x \to \bar{x}}} r(t,x;y) = \theta$ for all $y \in X$;

(iii) for each $y \in X$ there is a neighborhood Ω of $\bar{x}$ and $\eta \in (0,1]$ such that

$$t^{-1}[f(x+ty)-f(x)] \in [\underline{h}(y),\bar{h}(y)] + r(t,x;y) \text{ for all } t \in (0,\eta],\ x \in \Omega \ .$$

There are no implications between the above definition and Definition 1 without additional technical assumptions. For instance, if f is order-Lipschitz at $\bar{x} \in X$ according to Thibault and in addition there is a neighborhood W of $\theta \in X$ with a corresponding neighborhood Ω of $\bar{x}$ and $\eta \in (0,1]$ such that

$$t^{-1}[f(x+ty)-f(x)] \in [\underline{h}(y),\bar{h}(y)] + r(t,x;y) \text{ for all } x \in \Omega,\ t \in (0,\eta),\ y \in W,$$

then f is interval-Lipschitz at $\bar{x}$ according to Definition 1 with $m = \underline{h}$ and $M = \bar{h}$. Conversely, suppose f is interval-Lipschitz at $\bar{x}$ according to Definition 1 with the additional assumptions that $\lim_{y\to\theta} M(y) = \theta$ and $\lim_{\substack{t\downarrow 0\\ x\to\bar{x}}} r(t,x;y) = \theta$ for all $y \in X$ (not just for all $y \in W$). There exists an element W_0 of a neighborhood basis of $\theta \in X$ such that $W_0 \subseteq W$ with W_0 radial (Peressini [34, p. 162]). Thus, for each $y \in X$ there exists $\lambda_y > 0$ such that $\lambda y \in W_0$ for all λ with $|\lambda| \le \lambda_y$. Then f is order-Lipschitz at $\bar{x}$ according to Thibault with $\eta = \min(\epsilon,\lambda_y,1)$.

3. GENERALIZED DIRECTIONAL DERIVATIVES AND SUBDIFFERENTIALS.

Unless specified otherwise, in this section X denotes a locally convex Hausdorff topological vector space and V denotes a locally convex ordered topological vector space, that is, V is a Hausdorff locally convex topological vector space and an ordered vector space with a convex positive cone $V_+ = \{v \in V\colon v \ge 0\}$ that is closed. We also assume V is an order complete vector lattice for its order structure, that is, sup(u,v) exists for all u,v in V and sup B exists for each nonempty subset B of V that is order bounded above.

The subdifferential of an interval-Lipschitz mapping will be defined in terms of a directional derivative which we now introduce.

DEFINITION 2. If $f\colon X \to V$ is interval-Lipschitz at $\bar{x}$, the <u>generalized directional derivative</u> of f at $\bar{x}$ in the direction $y \in X$, denoted $f^0(\bar{x};y)$, is given by

$$f^0(\bar{x};y) = \inf_{\substack{N\in\eta(\bar{x})\\ \epsilon>0}} \ \sup_{\substack{x\in N\\ 0<t\le\epsilon}} t^{-1}[f(x+ty)-f(x)]$$

where $\eta(\bar{x})$ is a neighborhood base of $\bar{x}$ in X.

If X is a Banach space, V=R, and f is Lipschitz at $\bar{x}$ (which implies f is interval-Lipschitz at $\bar{x}$), then $f^0(x;\cdot)$ coincides with <u>Clarke's generalized directional derivative</u> at $\bar{x}$; see Clarke [2, 22-25]. If V is an order complete Banach lattice and f is locally o-Lipschitz (see Example 1) then $f^0(\bar{x};\cdot)$ also coincides with the <u>generalized o-directional derivative of f at $\bar{x}$ in the direction</u> y defined by Papageorgiou [28]. The <u>Clarke derivative</u> of f at $\bar{x}$ defined by <u>Kusraev</u> [31] coincides with $f^0(\bar{x};\cdot)$ if the range space and the filter in Kusraev [31] are, respectively, order complete and limited to the neighborhood filter of $\bar{x}$.

The next two results exhibit properties of $f^0(\bar{x};y)$ as a mapping of $y \in X$.

PROPOSITION 1. The mapping $y \to f^0(\bar{x};y)$ is a sublinear mapping from X to V that satisfies $f^0(\bar{x};y) \le M(y)$ for all $y \in W$ and $f^0(\bar{x};-y) = (-f)^0(\bar{x};y)$ for every $y \in X$.

PROOF. The proof of the sublinearity of $f^0(\bar{x};\cdot)$ follows that for real-valued Lipschitz functions, while $f^0(\bar{x};y) \le M(y)$ for all $y \in W$ follows directly from Definitions 1 and 2. For any given $y \in X$, there exists $\alpha_y > 0$ such that $\alpha y \in W$ for $|\alpha| \le \alpha_y$; hence $f^0(\bar{x};\alpha_y y) = \alpha_y f^0(\bar{x};y) \le M(\alpha_y y)$, so $f^0(\bar{x};y) \le \alpha_y^{-1} M(\alpha_y y)$ and thus $f^0(\bar{x};y) \in V$. Finally

$$(-f)^0(\bar{x};y) = \inf_{\substack{N\in\eta(\bar{x})\\ \epsilon>0}} \ \sup_{\substack{x\in N\\ 0<t\le\epsilon}} t^{-1}[-f(x+ty)+f(x)]$$

$$= \inf_{\substack{N\in\eta(\bar{x})\\ \epsilon>0}} \ \sup_{\substack{x\in N\\ 0<t\le\epsilon}} t^{-1}[f(x+ty+t(-y)) - f(x+ty)]$$

$= f^0(\bar{x};-y)$. ∎

REMARK. Note that since $f^0(\bar{x};\cdot)$ is sublinear, by Example 3 it is interval-Lipschitz on X.

The next result exhibits several sufficient conditions for $f^0(\bar{x};\cdot)$ to be a continuous mapping. For f: $X \to V$ we define the epigraph of f, denoted epi f, by epi $f:=\{(x,v) \in X \times V | v \geq f(x)\}$. Recall that the positive cone V_+ in V is <u>normal</u> if there exists a neighborhood basis $\mathcal{W}$ of $\theta \in V$ such that $W = (W+V_+)\cap(W-V_+)$ for all $W \in \mathcal{W}$ (Peressini [34, p. 61]).

PROPOSITION 2. If the positive cone V_+ of V is normal, then each of the following conditions implies that $f^0(\bar{x};\cdot)$ is continuous:

(i) *int epi* $f^0(\bar{x};\cdot)$ *is nonempty*;

(ii) $\lim_{y\to\theta} M(y) = \theta$, where the convergence is an order convergence;

(iii) $M(\cdot)$ *is continuous* at $\theta \in X$.

PROOF. (i) Since the order intervals in V are bounded in the topology of V and $f^0(\bar{x};\cdot)$ is convex, $f^0(\bar{x};\cdot)$ is continuous on X if it is bounded above in a neighborhood of one point (Valadier [10, p. 71]). But int epi $f^0(\bar{x};\cdot)$ is included in the set of $(y,v) \in X \times V$ such that $f^0(\bar{x};\cdot)$ is bounded above by v in a neighborhood of y.

(ii) If y is a point in W, then by Proposition 1, $0 = f^0(\bar{x};0) = f^0(\bar{x};y-y) \leq f^0(\bar{x};y) + f^0(\bar{x};-y) \leq f^0(\bar{x};y) + M(-y)$, and thus $-M(-y) \leq f^0(\bar{x};y) \leq M(y)$. Since V_+ is normal and $\lim_{y\to\theta} M(y) = \theta$, we conclude $\lim_{y\to\theta} f^0(\bar{x};y) = \theta$ (Peressini [34, p. 62]) which shows that $f^0(\bar{x};\cdot)$ is continuous at the origin. Since $f^0(\bar{x};\cdot)$ is continuous at the origin and sublinear, it is continuous on X (Thibault [36, Lemma 2.4]) or Borwein [16, Cor. 2.4]).

(iii) Since $f^0(\bar{x};y) \leq M(y)$ for each $y \in W$ and $f^0(\bar{x};\cdot)$ is convex, the continuity of $f^0(\bar{x};\cdot)$ at $\theta \in X$ follows directly from Borwein [16, Prop. 2.3] since $M(\cdot)$ is assumed continuous at $\theta \in X$. The continuity of $f^0(\bar{x};\cdot)$ on X follows as in part (ii). ∎

The continuity of $f^0(\bar{x};\cdot)$ leads to several results concerning the subdifferential. Hence we make the following definition.

DEFINITION 3. The mapping f: $X \to V$ is <u>regular</u> at $\bar{x} \in X$ if f is interval-Lipschitz at $\bar{x}$ and if $f^0(\bar{x};\cdot)$ is a continuous mapping from X to V.

Denote by L(X,V) the vector space of linear mappings from X to V. $\mathcal{L}(X,V)$ denotes the space of continuous linear mappings from X to V; $\mathcal{L}_S(X,V)$ denotes the latter space endowed with the topology of pointwise convergence.

DEFINITION 4. Let f: $X\to V$ be interval-Lipschitz at $\bar{x} \in X$. The <u>subdifferential</u> of f at $\bar{x}$, denoted $\partial f(\bar{x})$, is defined as follows:

$$\partial f(\bar{x}):=\{T \in \mathcal{L}(X,V) | T(y) \leq f^0(\bar{x};y) \ \forall \ y \in X\}.$$

If f is Lipschitz at $\bar{x}$ and V=R, the above definition coincides with <u>Clarke's subdifferential</u> [2, 22-25]. If f is locally o-Lipschitz (see Example 1), then Definition 4 is the <u>generalized gradient</u> of f at $\bar{x}$ defined by <u>Papageorgiou</u> [28, Def. 3.2]; finally, if f is Lipschitz at $\bar{x}$ according to Kusraev (see Example 2), then the above definition coincides with Kusraev's subdifferential (Kusraev [31, Def.3]).

If we ignore the topological structure on X and V and deal only with the algebraic structures, then we can define the algebraic subdifferential of f at $\bar{x}$, denoted $\partial_\alpha f(\bar{x})$; thus

$$\partial_\alpha f(\bar{x}) := \{T \in L(X,V) \mid T(y) \le f^o(\bar{x},y)\ \forall\, y \in X\}.$$

REMARK. The subdifferential $\partial f(\bar{x})$ can be empty; indeed, if f is linear and discontinuous, then $\partial f(\bar{x}) = \emptyset$ since $f^o(\bar{x};y) = f(y)$ for all $y \in X$.

PROPOSITION 3. The subdifferential $\partial f(\bar{x})$ of f at $\bar{x}$ is convex and satisfies $-\partial f(\bar{x}) = \partial(-f)(\bar{x})$.

PROOF. The convexity of $\partial f(\bar{x})$ follows directly from the definition; $-\partial f(\bar{x}) = \partial(-f)(\bar{x})$ is a consequence of the relation $f^o(\bar{x};-y) = (-f)^o(\bar{x};y)$, for all $y \in X$, proved in Proposition 1. ∎

PROPOSITION 4. If f is regular at $\bar{x}$ and V_+ is normal, then $\partial f(\bar{x}) = \partial_\alpha f(\bar{x})$, that is, $\partial f(\bar{x})$ is the set of all linear mappings $T: X \to V$ such that $T(y) \le f^o(\bar{x};y)$ for all $y \in X$.

PROOF. Suppose $T: X \to V$ is a linear mapping satisfying $T(y) \le f^o(\bar{x};y)$ for all $y \in X$. By the linearity of T, $-T(y) = T(-y) \le f^o(\bar{x};-y)$, thus $-f^o(\bar{x};-y) \le T(y) \le f^o(\bar{x};y)$. Since V_+ is normal and $f^o(\bar{x};\cdot)$ is continuous, $\lim_{y\to\theta} T(y) = \theta$ and hence T is continuous on X.

THEOREM 1. Under the assumption of Proposition 4, the subdifferential $\partial f(\bar{x})$ is a nonempty, closed, convex, equicontinuous subset of $L_S(X,V)$ with

$$f^o(\bar{x};y) = \max\{T(y) \mid T \in \partial f(\bar{x})\}\ .$$

If, in addition, the order intervals in V are compact, then $\partial f(\bar{x})$ is compact in $L_S(X,V)$.

PROOF. The subdifferential $\partial f(\bar{x})$ is the convex subdifferential of $f^o(\bar{x};\cdot)$ at zero. Then since f is assumed regular at $\bar{x}$, the results follow from Theoreme 6 and Corollaire 7 in Valadier [10].

REMARK. Theorem 1 provides a connection between the subdifferential of Definition 4 and the quasidifferential of Pschenichnyi [18]. A real-valued function defined on a topological vector space E is quasidifferentiable at $\bar{x} \in E$ in the sense of Pschenichnyi if

$$f'(\bar{x};d) := \lim_{\alpha\downarrow 0} \alpha^{-1}[f(\bar{x}+\alpha d) - f(\bar{x})]$$

exists for all $d \in E$ and if $\exists$ a nonempty weak*-closed subset $M_f(\bar{x})$ of $E^* \ni$

$$f'(\bar{x};d) = \mathrm{Max}\{x^*(d) \mid x^* \in M_f(\bar{x})\}.$$

Thus, by Theorem 1, if the real-valued function f defined on X (a locally convex Hausdorff spaced with normal cone) is interval-Lipschitz and regular at $\bar{x}$ with $f'(\bar{x};d) = f^o(\bar{x};d)$, then f is quasidifferentiable at $\bar{x}$.

REMARK. It is natural to consider a comparison of $\partial f(\bar{x})$ and $\partial_c f(\bar{x})$, the convex subdifferential of f at $\bar{x}$, and to compare $\partial f(\bar{x})$ with the Frechet or Gateaux derivative of f at $\bar{x}$. By Theorem 3.2 in Papageorgiou [28] the subclass of interval-Lipschitz mappings known as locally o-Lipschitz mappings (see Example 1) has a subdifferential $\partial f(\bar{x})$ such that $\partial f(\bar{x}) = \partial_c f(\bar{x})$ when f is convex. Similarly, a locally o-Lipschitz mapping $f: X \to Y$ that is continuously Gateaux differentiable for $\|\cdot\|_1$ on Y, where $\|y\|_1 := \inf\{k \mid |y| \le ke\}$ (e is the strong unit on the Banach lattice Y), satisfies $\partial f(\bar{x}) = \{f'(\bar{x})\}$ by Papageorgiou [28, Th. 3.3].

4. OPTIMALITY CONDITIONS

In this section we show that our approach to the local analysis of nonsmooth operators introduced in Sections 2 and 3 has relevance to mathematical programming. In particular, we give necessary and sufficient optimality conditions for nondifferentiable programming problems with real-valued objective functions and constraints consisting of either an arbitrary set or an arbitrary set and a vector-valued operator. While the results are related to those obtained in Kusraev [31] and Thibault [36], where the objective functions are vector-valued, our assumptions and proof techniques are somewhat different. Specifically, Kusraev's vector-valued mappings are Lipschitz with the absolute value operator while Thibault's mappings are "compactly Lipschitzian" [36, Def. 1.1]. In addition, our proof of the Kuhn-Tucker necessary conditions (Theorem 2), which recalls a paper of Guignard [37], does not explicitly use the assumptions that the range space of the constraint operator is an ordered space. This raises the possibility of substituting for the generalized gradient of the constraint operator g at $\bar{x}$ any closed convex subset $\Gamma_g(\bar{x})$, say, of $L_S(X,V)$ that satisfies the conditions we require of the generalized gradient. This approach could generate various closed convex-valued multifunctions as in Ioffe [29] (where such multifunctions are called fans) and lead to necessary conditions which have as special cases the necessary conditions of Clarke [24], Hiriart-Urruty [1, 38, 39] and Ioffe [40]. Ioffe [30] has in fact used the concept of fan to develop more general necessary conditions.

Let X be a Banach space, V as described at the beginning of Section 3, S a nonempty subset of X, and f an extended real-valued function on X which, unless stated otherwise, is assumed to be finite and interval-Lipschitz at $\bar{x} \in S$. Consider the problem:

minimize $f(x)$, subject to $x \in S$;

$\bar{x}$ is a local minimum of f on S if f is finite at $\bar{x}$ and if there exists a neighborhood N of $\bar{x}$ such that $f(x) \ge f(\bar{x})$ for every $x \in S \cap N$; $\bar{x}$ is a minimum of f on S if f is finite at $\bar{x}$ and $f(x) \ge f(\bar{x})$ for every $x \in S$. The *contingent* cone *of S at $x_0 \in clS$* (closure of *S*), denoted *$K(S;x_0)$*, is defined as follows:

$$K(S;x_0) := \{d \in X \mid \exists t_n > 0, \{x_n\} \subseteq S,\ x_n \to x_0 \text{ with } d = \lim_{n\to\infty} t_n(x_n - x_0)\}$$

$$= \{d \in X \mid \exists t_n \downarrow 0,\ d_n \to d \text{ with } x_0 + t_n d_n \in S \text{ for all } n\}\ . \qquad (4.1)$$

The *(Clarke) tangent cone* of S at $x_0 \in clS$, denoted $\mathcal{T}(S;x_0)$, is the following set:

$$\mathcal{T}(S;x_0) := \{d \in X \mid \text{ for every } \{x_n\} \subseteq clS \ni x_n \to x_0 \text{ and for every } \{t_n\} \ni t_n \downarrow 0,\ \exists \{d_n\} \ni d_n \to d \text{ with } x_n + t_n d_n \in S \text{ for all } n\}\ .$$

$K(S:x_0)$ is a closed cone and $\mathcal{T}(S;x_0)$ is a closed convex cone with $\mathcal{T}(S;x_0) \subseteq K(S;x_0)$.

The closure of the convex hull of $K(S;x_0)$ is denoted $P(S;x_0)$. The *polar cone* of a nonempty set $A \subseteq X$ is given by $A^o:=\{x^* \in X^* | x^*(x) \leq 0 \ \forall \ x \in A\}$, where X^* is the topological dual of X; if $A = \emptyset$, $A^o:=X^*$. If $A^* \subseteq X^*$ is nonempty, the *prepolar* of A^* is ${}^o(A^*):=\{x \in X | x^*(x) \leq 0 \ \forall \ x^* \in A^*\}$. If $A^* = \emptyset$, ${}^o(A^*):=X$. $A^o({}^o(A^*))$ is a weak*-closed (weakly closed) convex cone in $X^*(X)$.

We begin our study of optimality with three results that give necessary conditions for a vector $\bar{x}$ to be a local minimum.

PROPOSITION 5. If $\bar{x}$ is a local minimum of f on S=X, then $0 \in \partial f(\bar{x})$.

PROOF. Consider a sequence $\{t_n\} \subseteq (0,1]$ converging to 0 and select neighborhoods N of $\bar{x}$ and W of $\theta \in X$, a constant $\epsilon > o$, and m, M and r satisfying Definition 1. We may assume $f(x) \leq f(\bar{x})$ for all $x \in N$. For each $y \in W$ there exists n_0 such that $t_n^{-1}[f(\bar{x}+t_n y) - f(\bar{x})] - r(t_n, \bar{x};y) \in [m(y), M(y)]$ and $\bar{x} + t_n y \in N$ for all n $\geq n_0$. In addition, there exists a convergent subsequence $\{t_{\alpha(n)}^{-1}[f(\bar{x} + t_{\alpha(n)}y) - f(\bar{x})]\}$ since $[m(y), M(y)]$ is compact. Therefore,

$$f^o(\bar{x};y) = \lim_{\substack{\epsilon \downarrow 0 \\ N \in \eta(\bar{x})}} \sup_{\substack{x \in N \\ 0<t\leq\epsilon}} t^{-1}[f(x + ty) - f(x)]$$

$$\geq \lim_{n\to\infty} t_{\alpha(n)}^{-1}[f(\bar{x} + t_{\alpha(n)}y) - f(\bar{x})] \geq 0 .$$

Since W is radial, we conclude $f^o(\bar{x};y) \geq 0$ for $y \in X$ and hence that $0 \in \partial f(\bar{x})$. ∎

REMARK. Proposition 5 is related to a necessary condition for an unconstrained optimum of a quasidifferentiable function on E^n. A real-valued function f on E^n is quasidifferentiable at x if f is directionally differentiable at x and if there exists convex compact sets $\underline{\partial} f(x)$ and $\bar{\partial} f(x)$ in E^n such that

$$f'(x;d) = \max_{v \in \underline{\partial} f(x)} \langle v,d\rangle + \min_{w \in \bar{\partial} f(x)} \langle w,d\rangle$$

(Demyanov and Rubinov [41]). Polyakova [42] has shown that $-\bar{\partial} f(\bar{x}) \subseteq \underline{\partial} f(\bar{x})$ is a necessary condition for $\bar{x}$ to be a minimum of a quasidifferentiable function f on E^n. By Theorem 1, if the real-valued function f on E^n is order-Lipschitz and regular at $\bar{x}$, then f is quasidifferentiable at $\bar{x}$ with $\bar{\partial} f(\bar{x}) = \{0\}$ and $\underline{\partial} f(\bar{x}) = \partial f(\bar{x})$, thus the optimality condition immediately above reduces to the condition in Proposition 5: $0 \in \partial f(\bar{x})$. However, Proposition 5 is applicable in the broader context of infinite - dimensional spaces. In addition Proposition 5 generalizes several results in the literature obtained for Lipschitz functions on a Banach space, e.g., Clarke [2], Ioffe [30, 40] and Thibault [36].

PROPOSITION 6. If $\bar{x}$ is a local minimum of f on S, m and M in Definition 1 are continuous, and $\bar{x}$ is such that

$$f^o(\bar{x};y) = \lim_{\substack{\epsilon \downarrow 0 \\ N_1 \in \eta(\bar{x}) \\ N_2 \in \eta(y)}} \sup_{\substack{x \in N_1 \\ v \in N_2 \\ 0<t\leq\epsilon}} t^{-1}[f(x + tv) - f(x)]$$

for all $y \in K(S;\bar{x})$, then $f^o(\bar{x};y) \geq 0$ for all $y \in K(S;\bar{x})$.

PROOF. Suppose $y \in K(S;\bar{x})$ and let $\{t_n\}$ and $\{y_n\}$ be the sequences corresponding to (4.1). In addition, choose N, W, ϵ, m, M and r satisfying Definition 1 with m and M continuous. There exists n_1 such that $\bar{x} + t_n y_n \in N$ for all $n \geq n_1$, hence

$$\bar{x} + t_n y_n \in S \cap N \text{ for all } n \geq n_1 \tag{4.2}$$

by (4.1). We may assume $f(\bar{x}) \leq f(x)$ for all $x \in S \cap N$. Since W is radial, corresponding to y and each y_n there exists $\alpha_y > 0$ and $\alpha_n > 0$, respectively, such that $\alpha y \in W$ for $|\alpha| \leq \alpha_y$ and $\alpha y_n \in W$ for $|\alpha| \leq \alpha_n$. Hence there exists n_2 such that

$$t_n^{-1}[f(\bar{x}+t_n\alpha_n y_n) - f(\bar{x})] - r(t_n,\bar{x};\alpha_n y_n) \in [m(\alpha_n y_n), M(\alpha_n y_n)] \tag{4.3}$$

for all $n \geq n_2$,

and thus (4.2) and (4.3) hold for all $n \geq n_0 := \max\{n_1, n_2\}$. Since $\{y_n\}$ converges to y, there exists a sequence of α_n's that converges to α_y. In addition, since each $[m(\alpha_n y_n), M(\alpha_n y_n)\}$ is compact and m and M are continuous, there exists a convergent subsequence $t_{\sigma(n)}^{-1}[f(\bar{x} + t_{\sigma(n)}\alpha_{\sigma(n)}y_{\sigma(n)} - f(\bar{x})]$. Therefore, since $y \in K(S;\bar{x})$, $\alpha_y y \in K(S;\bar{x})$ and

$$\alpha_y f^0(\bar{x};y) = f^0(\bar{x};\alpha_y y) = \lim_{\substack{\epsilon \downarrow 0 \\ N_1 \in \eta(\bar{x}) \\ N_2 \in \eta(\alpha_y y)}} \sup_{\substack{0<t\leq\epsilon \\ x\in N_1 \\ v\in N_2}} t^{-1}[f(x + tv) - f(x)]$$

$$\geq \lim_{n\to\infty} t_{\sigma(n)}^{-1}[f(\bar{x} + t_{\sigma(n)}\, \alpha_{\sigma(n)}\, y_{\sigma(n)}) - f(\bar{x})] \geq 0 ,$$

which implies $f^0(\bar{x};y) \geq 0$. ∎

REMARK. The assumption in Proposition 6 concerning $f^0(\bar{x};y)$ plays a role similar to "condition ($\mathscr{C}$)" imposed by Hiriart-Urruty [38, p. 89] to obtain the same necessary optimality condition.

It is customary to express optimality conditions in terms of the polar cones of the cones of displacement. A result of this type is presented below. Recall first that if C is a nonempty subset of X, the distance function d_C: $X \to R$, defined by $d_C(x) = \inf[\|x-c\| \mid c \in C\}$, is a globally Lipschitz function on X with Lipschitz constant 1.

PROPOSITION 7. Let $\bar{x}$ be a local minimum of f on S. If f is regular at $\bar{x}$, then $0 \in \partial f(x) + (\mathscr{T}(S;\bar{x}))^0$.

PROOF. Since f is interval-Lipschitz at $\bar{x}$, choose neighborhoods N and W, mappings m, M and r, and $\epsilon > 0$ that satisfy Definition 1. We will first show that there exists a neighborhood N_0 of $\bar{x}$ over which $\bar{x}$ minimizes $f(x) + \rho^{-1}(|M(\bar{y}) + r(\bar{t},x;\bar{y})| + |m(\bar{y}) + r(\bar{t},x;\bar{y})|)d_S(x)$ for some $\bar{y} \in W$ and some $\bar{t} \in (0,\epsilon]$, where $\rho > 0$ is such that $B(\theta,2\rho) \leq W$. By way of contradiction, suppose this result is false. Then there exists a sequence $\{x_n\}$ converging to $\bar{x}$ such that $f(x_n) + \rho^{-1}(|M(y) + r(t,x_n;y)| + |m(y) + r(t,x_n;y)|)\, d_S(x_n) < f(\bar{x})$ for $y \in W$ and $t \in (0,\epsilon]$. There exists n_0 such that $d_S(x_n) > 0$ for $n \geq n_0$, since otherwise x_n belongs to S and the above inequality contradicts the local optimality of $\bar{x}$. Since $d_S(x_n)$ converges to 0 as $n \to \infty$, we can choose n sufficiently large so that f is order-Lipschitz at x_n in a neighborhood of radius $2d_S(x_n)$ and with the same neighborhood W, mappings M, m and r, and $\epsilon > 0$ mentioned at the beginning of the proof. There exists $s_n \in S$ such that $\|s_n - x_n\| \leq \min\{\rho\epsilon, (1+\eta)d_S(x_n)\}$, where $\eta \in (0,1)$ satisfies $f(x_n) + \rho^{-1}(|M(y) + r(t,x_n;y)|+|m(y) + r(t,x_n:y)|)(1+\eta)d_S(x_n) < f(\bar{x})$ for $y \in W$ and $t \in (0,\epsilon]$. Since $s_n = x_n + t_0 y_0$ where $t_0 := \rho^{-1}\|s_n-x_n\| \leq \epsilon$ and $y_0 := \rho\|s_n-x_n\|^{-1}(s_n-x_n) \in W$ with $\|s_n-x_n\| < 2d_S(x_n)$, we have

$$f(s_n) \le f(x_n) + t_0(|M(y_0) + r(t_0,x_n;y_0)|+|m(y_0) + r(t_0,x_n;y_0)|)$$

$$\le f(x_n) + \rho^{-1}(|M(y_0) + r(t_0,x_n;y_0)|+|m(y_0) + r(t_0,x_n;y_0)|)(1+\eta)d_S(x_n)$$

$$< f(\bar{x}) ,$$

which contradicts the local optimality of $\bar{x}$. Thus $\bar{x}$ is a local minimum of $f(x) + \rho^{-1}(|M(\bar{y}) + r(\bar{t},x;\bar{y})|+|m(\bar{y}) + r(\bar{t},x;\bar{y})|)d_S(x_n)$ for some $\bar{y} \in W$ and $\bar{t} \in (0,\epsilon]$. Since $\lim_{\substack{t\downarrow 0\\ x\to\bar{x}}} r(t,x;\bar{y}) = 0$, we have $k:=\rho^{-1}(|M(\bar{y})|+|r(\bar{t},\bar{x};\bar{y})|+|m(\bar{y})|+|r(\bar{t},\bar{x};\bar{y})|) \ge \rho^{-1}(|M(\bar{y}) + r(\bar{t},x;\bar{y})|+|m(\bar{y}) + r(\bar{t},x;\bar{y})|)$; therefore $\bar{x}$ is also a local minimum of $f(x) + kd_S(x)$. Finally, since $\partial(f_1+f_2)(\bar{x}) \subseteq \partial f_1(\bar{x}) + \partial f_2(\bar{x})$ where f_1 and f_2 are interval-Lipschitz at $\bar{x}$ and f_1 is regular at $\bar{x}$, by Proposition 5 and Clarke [2, Prop. 2.4, p. 51] we conclude that

$$0 \in \partial(f(\bar{x}) + kd_S(\bar{x})) \subseteq \partial f(\bar{x}) + k\partial d_S(\bar{x}) \subseteq \partial f(\bar{x}) + (\mathcal{T}(S,\bar{x}))^0 . \quad \blacksquare$$

If f is Lipschitz, then a stronger necessary condition than the one in Proposition 7 can be obtained.

PROPOSITION 8. Let $\bar{x}$ be a local minimum of f on S, where f is Lipschitz at $\bar{x}$, and M a convex cone contained in $K(S;\bar{x})$; then

$$0 \in \partial f(\bar{x}) + M^0 . \tag{4.4}$$

PROOF. Since f is assumed Lipschitz at $\bar{x}$, the result follows directly from Theorems 7 and 8 in Hiriart-Urruty [38]. ∎

REMARKS. 1) Condition (4.4) is sharpest when $K(S;\bar{x})$ is convex, in which case (4.4) becomes

$$0 \in \partial f(\bar{x}) + [K(S;\bar{x})]^0 . \tag{4.5}$$

If, in addition, f is continuously differentiable at $\bar{x}$, then $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$ by Rockafellar [4, Proposition 4] and (4.5) reduces to $0 \in \nabla f(\bar{x}) + [K(S;\bar{x})]^0$, i.e., $\nabla f(\bar{x}) \in -[K(S;\bar{x})]^0$ which, since $[K(S;\bar{x})]^0 = [P(S;\bar{x})]^0$, is the well-known optimality condition in differentiable programming given by Guignard [37].

2) To establish the optimality condition in differentiable programming noted in remark 1, it is <u>not</u> necessary to assume that $K(S;\bar{x})$ is convex. The convexity requirement is needed in the nondifferentiable case since

$$f^0(\bar{x};d) \ge 0 \ \forall \ d \in K(S;\bar{x})$$

cannot be extended to $c\ell\{co\ K(S;\bar{x})\} = P(S;\bar{x})$. Thus, for nondifferentiable objective functions, relation (4.5) does not hold without the convexity of $K(S;\bar{x})$. To illustrate this fact we include an example due to Hiriart-Urruty [1, p. 80]. Let $X = E^2$, $f: E^2 \to R$ is given by $f(x_1, x_2) = 1 - \exp(x_2 - |x_1|)$, $S = \{(x_1,x_2) \in E^2: x_2 - |x_1| \le 0\}$; $\bar{x} = (0,0)$ is a minimum of f on S, $[K(S;\bar{x})]^0 = \{(0,0)\}$, and $\partial f(\bar{x}) = co\{(1,-1), (-1,-1)\}$.

A statement of sufficient conditions requires the following preliminaries. A function $f: X \to R$ that is interval-Lipschitz at $\bar{x}$ is <u>pseudoconvex over S at $\bar{x}$</u> if for all $x \in S$, $f^0(\bar{x};x-\bar{x}) \geq 0$ implies $f(x) \geq f(\bar{x})$. A subset $A \subseteq X$ is <u>pseudoconvex at x_0</u> $\in$ cl A if $x - x_0 \in P(A;x_0)$ for all $x \in A$, and <u>strictly pseudoconvex at x_0</u> if $x - x_0 \in K(A;x_0)$ for all $x \in A$.

PROPOSITION 9. Suppose f is pseudoconvex over S at $\bar{x} \in S$ and S is pseudoconvex at $\bar{x}$; then $0 \in \partial f(\bar{x}) + [P(S;\bar{x})]^0$ is a sufficient condition for $\bar{x}$ to be a minimum of f on S.

PROOF. The condition $0 \in \partial f(\bar{x}) + [P(S;\bar{x})]^0$ implies $0 = T + \gamma$, where $T \in \partial f(\bar{x})$ and $\gamma \in [P(S;\bar{x})]^0$. Therefore, for all $x \in S$, $0 = T(x-\bar{x}) + \gamma(x-\bar{x})$. Since S is pseudoconvex at $\bar{x}$, $x - \bar{x} \in P(S;\bar{x})$ for all $x \in S$, which implies $\gamma(x - \bar{x}) \leq 0$. Thus $T(x - \bar{x}) \geq 0$ and, for all $x \in S$, $f^0(\bar{x};x-\bar{x}) \geq T(x-\bar{x}) \geq 0$, which by the pseudoconvexity of f implies $f(x) \geq f(\bar{x})$. ∎

REMARKS. 1) A "local minimum" analogue of the above result follows directly if f is pseudoconvex over $S \cap N_1$ at $\bar{x}$, for some neighborhood N_1 of $\bar{x}$, and if S is locally pseudoconvex at $\bar{x}$, where the latter means that there exists a neighborhood N_2 of $\bar{x}$ such that $x - \bar{x} \in P(S;\bar{x})$ for all $x \in S \cap N_2$. Hiriart-Urruty [39, Th. 5] states (for f Lipschitz at $\bar{x}$) that $0 \in \partial f(\bar{x}) + [K(S;\bar{x})]^0$ (note that $[K(S;\bar{x})]^0 = [P(S;\bar{x})]^0$) is a sufficient condition for $\bar{x}$ to be a <u>local</u> minimum of f on S under the assumptions that f be locally pseudoconvex at $\bar{x}$ and that S be locally <u>strictly</u> pseudoconvex at $\bar{x}$; this latter condition is termed "Condition L" by Hiriart-Urruty.

2) A more desirable sufficient condition is possible in Proposition 9, but it is acquired at the expense of strengthening the assumption on S by using the (Clarke) tangent cone $\mathcal{T}(S;\bar{x})$. If f is pseudoconvex over S at $\bar{x}$ (as in Proposition 9) and if $x - \bar{x} \in \mathcal{T}(S;\bar{x})$ for all $x \in S$, then $0 \in \partial f(\bar{x}) + [\mathcal{T}(S;\bar{x})]^0$ is a sufficient condition for $\bar{x}$ to be a minimum of f on S. If S is locally convex at $\bar{x}$, that is, there is a neighborhood N of $\bar{x}$ such that $S \cap N$ is convex, then $\mathcal{T}(S;\bar{x}) = K(S;\bar{x}) = P(S;\bar{x})$ Hiriart-Urruty [38, p. 83] and the sufficient condition immediately above is equivalent to the sufficient condition in Proposition 9.

To state the problem with an explicit operator constraint, let V be a locally convex ordered topological vector space that is an order complete vector lattice. A and B are nonempty subsets in X and V, respectively, and $g: X \to V$ is interval-Lipschitz at $\bar{x} \in S$ where $S = \{x \in A | g(x) \in B\}$. Let $J = \{x \in X | T(x) \in P(B;g(\bar{x}))$ for each $T \in \partial g(\bar{x})\}$ and $H^* = \{h \in X^* | h \in \mu\partial g(\bar{x}),\ \mu \in \{P(B;g(\bar{x}))\}^0\}$, where $\mu\partial g(\bar{x}) = \{\mu \cdot T | T \in \partial g(\bar{x})\}$. Note that J is a closed convex cone and H^* is a cone.

THEOREM 2 (KUHN-TUCKER CONDITIONS). Suppose H^* is closed and G is a closed convex cone in X such that $G \cap J = \mathcal{T}(S;\bar{x})$ and $G^0 + J^0$ is closed. If $\bar{x}$ is a local minimum of f over S, where f is regular at $\bar{x}$, then there exists $\mu \in [P(B;g(\bar{x}))]^0$ such that $0 \in \partial f(\bar{x}) + \mu\partial g(\bar{x}) + G^0$.

PROOF. Since $\bar{x}$ is a local minimum of f on S, we have by Proposition 7 that $0 \in \partial f(\bar{x}) + (\mathcal{T}(S;\bar{x}))^0$. Since $J^0 + G^0$ is closed, then $(\mathcal{T}(S;\bar{x}))^0 = J^0 + G^0$ (property G3, Guignard [37]) and $0 \in \partial f(\bar{x}) + J^0 + G^0$. Let $\gamma \in {}^0(H^*)$; then $\mu(T(\gamma)) \leq 0$ for any $\mu \in [P(B;g(\bar{x}))]^0$ and $T \in \partial g(\bar{x})$. Now suppose that $T(\gamma) \notin P(B;g(\bar{x}))$; then since $P(B;g(\bar{x}))$ is a closed convex cone, by the strong separation theorem (Dunford and Schwartz [43, p. 417]) there exists $v^* \in V^*$ such that $v^*(T(\gamma)) > 0$

$\geq v^*(w)$ for any $w \in P(B;g(\bar{x}))$, which implies that $v^* \in [P(B;g(\bar{x}))]^o$. Then $v^*(T(\gamma)) \leq 0$ and this contradicts $v^*(T(\gamma)) > 0$. Therefore, $T(\gamma) \in P(B;g(\bar{x}))$, that is, for each $\gamma \in {}^o(H^*)$ we have shown $\gamma \in J$. Hence ${}^o(H^*) \subset J$ and since H^* is a closed convex cone, $H^* = ({}^o(H^*))^o \supset J^o$, which shows that there exists $\mu \in [P(B;g(\bar{x}))]^o$ such that $0 \in \partial f(\bar{x}) + \mu\partial g(\bar{x}) + G^o$. ▌

REMARK. Theorem 2 provides a multiplier rule for an infinite dimensional equality constraint. If X is a Banach space, V is a locally convex ordered topological vector space that is an ocvl, and $B = \{0\}$, then $P(B;g(\bar{x})) = \{0\}$ and Theorem 2 says that there exists $\mu \in V^*$ such that $0 \in \partial f(\bar{x}) + \mu\partial g(\bar{x}) + G^o$. Multiplier rules for infinite dimensional equality constraints have appeared only recently; Ioffe [30, 40], for example, provides such a rule for V a (not necessarily ordered) Banach space.

The optimality condition in Theorem 2 compares favorably with other results in the literature. For example, if $G = \mathcal{I}(S;\bar{x})$, then $0 \in \partial f(\bar{x}) + \mu\partial g(\bar{x}) + [\mathcal{I}(S;\bar{x})]^o$ and we have a result consistent with the necessary condition $0 \in \partial f(\bar{x}) + [\mathcal{I}(S;\bar{x})]^o$ established by Clarke [24, Lemma 2] in a slightly different form. Theorem 2 also generalizes results of Hiriart-Urruty [1, Th. 6] and Demyanov [44, Th. 7] (see REMARK after Prop. 9), for Lipschitz functions on R^n, and is related to a result of Ioffe [40, Prop. 1] for Lipschitz functions on a Banach space.

EXAMPLE 7. The role of the various sets in Theorem 2 is perhaps better understood by considering the finite-dimensional case. Let X and V be the Euclidean spaces E^n and E^m, respectively. If $B = E^m_- = \{y \in E^m | y \leq 0\}$, the problem becomes $\min\{f(x) | x \in A, g(x) \leq 0\}$. Let I and J be such that $g_i(\bar{x}) = 0$ for all $i \in I$ and $g_j(\bar{x}) < 0$ for all $j \in J$, where $\bar{x} \in S = \{x \in A | g(x) \leq 0\}$. Then $[P(B;g(\bar{x}))]^o = [P(E^m_-;g(\bar{x}))]^o = \{\lambda \in E^m | \lambda \geq 0, \lambda g(\bar{x}) = 0\} = \{\lambda \in E^m | \lambda_i \geq 0, i \in I, \lambda_j = 0, j \in J\}$. If $\bar{x}$ minimizes f over S, the necessary conditions of Theorem 1 imply that there exist scalars $\lambda_i \geq 0$ such that $\lambda_i g_i(\bar{x}) = 0$, $i = 1,\ldots, m$ and $0 \in \partial f(\bar{x}) + \Sigma_{i=1}^m \lambda_i \partial g_i(\bar{x}) + G^o$. If $A = E^n$ and $G^o = [P(E^n;\bar{x})]^o = \{0\}$, we have $0 \in \partial f(\bar{x}) + \Sigma_{i=1}^m \lambda_i \partial g_i(\bar{x})$; moreover, if f and g are continuously differentiable at $\bar{x}$, the latter condition reduces to $0 = \nabla f(\bar{x}) + \Sigma_{i=1}^m \lambda_i \nabla g_i(\bar{x})$. Note that both $J = \{x \in E^n | \mathcal{E}_i x \leq 0 \text{ for each } \mathcal{E}_i \in \partial g_i(\bar{x}), i \in I\}$ and $H^* = \{h \in E^n | h = \Sigma_{i \in I} \lambda_i \mathcal{E}_i, \lambda_i \geq 0, \mathcal{E}_i \in \partial g_i(\bar{x})\}$ are closed convex cones.

If f is Lipschitz at $\bar{x} \in S$, then we can obtain necessary conditions that in general are more precise than those in Theorem 2.

THEOREM 3 (KUHN-TUCKER CONDITIONS). Suppose H^* is closed and G is a closed convex cone in X such that $G \cap J = \text{cl}M$, for a convex cone M contained in $K(S;\bar{x})$, and $G^o + J^o$ is closed. If $\bar{x}$ is a local minimum of f over S, where f is Lipschitz at $\bar{x}$, then there exists $\mu \in [P(B;g(\bar{x}))]^o$ such that $0 \in \partial f(\bar{x}) + \mu\partial g(\bar{x}) + G^o$.

PROOF. In the proof of Theorem 2 use Proposition 8 instead of Proposition 7 and the relation $M^o = (\text{cl}M)^o$ (property C2, Guignard [37]). ▌

Sufficient conditions are obtained by imposing mild convexity assumptions.

THEOREM 4. If G is a closed convex cone in X such that $x - \bar{x} \in G$ for all $x \in S$, if there exists $\mu \in [P(B;g(\bar{x}))]^o$ such that $0 \in \partial f(\bar{x}) + \mu\partial g(\bar{x}) + G^o$, if S is strictly pseudoconvex at $\bar{x}$ and $T(K(S;\bar{x})) \subset K(B;g(\bar{x}))$ for all $T \in \partial g(\bar{x})$, and if f is pseudoconvex over S at $\bar{x}$, then $\bar{x}$ is optimal for f over S.

PROOF. There exists $\mathcal{E} \in \partial f(\bar{x})$, $T \in \partial g(\bar{x})$ and $x^* \in G^o$ such that $0 = \mathcal{E} + \mu \cdot T + x^*$, hence $0 = \mathcal{E}(x - \bar{x}) + \mu(T(x - \bar{x})) + x^*(x - \bar{x})$. Since S is strictly pseudoconvex at $\bar{x}$, for all $x \in S$ we have $T(x - \bar{x}) \in K(B;g(\bar{x}))$ and thus $\mu(T(x - \bar{x})) \leq 0$; also, $x^*(x - \bar{x}) \leq 0$ for all $x \in S$, hence $\mathcal{E}(x - \bar{x}) \geq 0$. Hence, for all $x \in S$, $f^o(\bar{x}; x - \bar{x}) \geq \mathcal{E}(x - \bar{x}) \geq 0$ which, since f is pseudoconvex over S at $\bar{x}$, implies $f(x) \geq f(\bar{x})$.

4. SUMMARY

For a vector-valued function $f: X \to V$ that is interval-Lipschitz at $\bar{x}$ we have defined and obtained properties for the generalized directional derivative $f^o(\bar{x};y)$ and the generalized gradient $\partial f(\bar{x})$. In particular, we have discussed conditions under which the sublinear mapping $f^o(\bar{x};\cdot)$ is continuous and have shown that when this is the case, $\partial f(\bar{x})$ is nonempty, convex, closed and equicontinuous (as a subset of $\mathcal{L}(X,V)$ with the topology of pointwise convergence) and $f^o(\bar{x};y) = \max\{T(y) | T \in \partial f(\bar{x})\}$. If the order intervals in V are compact, then $\partial f(\bar{x})$ is also compact. We also have obtained necessary and sufficient optimality conditions for a nondifferentiable mathematical programming problem with a vector-valued operator constraint and/or an arbitrary set constraint. The proof techniques point to future research in the area of convex-valued multifunctions as in Ioffe [30], for example, which in turn could lead to more general optimality conditions.

REFERENCES

1. HIRIART-URRUTY, J. B. "On Optimality Conditions in Nondifferentiable Programming," Mathematical Programming 14, (1978) 73-86.
2. CLARKE, F. H. Optimization and Nonsmooth Analysis, Wiley and Sons, New York, (1983).
3. ROCKAFELLAR, R. T. Convex Analysis, Princeton University Press, Princeton, New Jersey, (1970a).
4. ROCKAFELLAR, R. T. "Conjugate Convex Functions in Optimal Control and the Calculus of Variations," Journal of Mathematical Analysis and Applications 32, (1970b) 174-222.
5. ROCKAFELLAR, R. T. "Existence and Duality Theorems for Convex Problems of Bolza," Transactions of the American Mathematical Society 159, (1971a) 1-40.
6. ROCKAFELLAR, R. T. "Integrals Which are Convex Functionals, II," Pacific Journal of Mathematics 39, (1971b) 439-469.
7. ROCKAFELLAR, R. T. "Conjugate Duality and Optimization," Society for Industrial and Applied Mathematics, Philadelphia, (1971b).
8. MOREAU, J. J. "Fonctionelles Convexes," lecture notes, Seminaire "Equations aux Derivees Partielles," (1966) College de France.
9. MCLINDEN, L. "Dual Operations on Saddle Functions," Transactions of the American Mathematical Society 179, (1973) 363-381.
10. VALADIER, M. "Sous-Differentiabilite de Fonctions Convexes a Valeurs dans un Espace Vectoriel Ordonne," Mathematica Scandinavia 30, (1972) 65-74.
11. IOFFE, A. D., and LEVIN, L. "Subdifferentials of Convex Functions," Transactions of the Moscow Mathematical Society 26, (1972) 1-72.
12. ZOWE, J. "Subdifferentiability of Convex Functions with Values in an Ordered Vector Space," Mathematica Scandinavia 34, (1974) 69-83.
13. ZOWE, J. "A Duality Theorem for a Convex Programming Problem in Order Complete Vector Lattices," Journal of Mathematical Analysis and Applications 50, (1975) 273-287.
14. KUTATELADZE, S. S. "Support Sets of Sublinear Operators," Doklady Akadamie Nauk SSSR 230, (1976) 1029-1032.

15. RUBINOV, A. M. "Sublinear Operators and their Applications," Russian Mathematical Surveys 32, (1977) 115-175.

16. BORWEIN, J. M. "Continuity and Differentiability Properties of Convex Operators," Proceedings of the London Mathematical Society (3) 44, (1982) 420-444.

17. PAPAGEORGIOU, N. S. "Nonsmooth Analysis on Partially Ordered Vector Spaces: Part 1 - Convex Case," Pacific Journal of Mathematics 107, (1983a) 403-458.

18. PSHENICHNYI, B. N. Necessary Conditions for an Extremum, Marcel Dekker, New York, (1971).

19. BAZARAA, M. S., GOODE, J. J. and NASHED, M. A. "On the Cones of Tangents with Applications to Mathematical Programming," Journal of Optimization Theory and Applications 13, (1974) 389-426.

20. PENOT, J. P. "Calcul Sous-differentiels et Optimisation," Journal of Functional Analysis 27, (1978) 248-276.

21. WARGA, J. "Derivative Containers, Inverse Functions and Controllability," in Russell, D. L., ed., Proceedings of the MRC Symposium on the Calculus of Variations and Optimal Control, Academic Press, New York, (1976).

22. CLARKE, F. H. "Necessary Conditions for Nonsmooth Problems in Optimal Control and the Calculus of Variations," Ph.D. Dissertation, University of Washington, Seattle, WA, (1973).

23. CLARKE, F. H. "Generalized Gradients and Applications," Transactions of the American Mathematical Society 205, (1975) 247-262.

24. CLARKE, F. H. "A New Approach to Lagrange Multipliers," Mathematics of Operations Research 2, (1976) 165-174.

25. CLARKE, F. H. "Generalized Gradients of Lipschitz Functionals," Advances in Mathematics 40, (1981) 52-66.

26. ROCKAFELLAR, R. T. The Theory of Subgradients and its Applications to Problems of Optimization. Convex and Nonconvex Functions. Helderman Verlag, Berlin, (1981).

27. BORWEIN, J. M. and STROJWAS, H. M. "Directionally Lipschitzian Mappings on Baire Spaces," Canadian Journal of Mathematics XXXVI, (1984) 95-130.

28. PAPAGEORGIOU, N. S. "Nonsmooth Analysis on Partially Ordered Vector Spaces: Part 2 - Nonconvex Case, Clarke's Theory," Pacific Journal of Mathematics 109, (1983b) 463-495.

29. IOFFE, A. D. "Nonsmooth Analysis: Differential Calculus of Nondifferentiable Mappings," Transactions of the American Mathematical Society 266, (1981) 1-56.

30. IOFFE, A. D. "Necessary Conditions in Nonsmooth Optimization," Mathematics of Operations Research 9, (1984) 159-189.

31. KUSRAEV, A. G. "On Necessary Conditions for an Extremum of Nonsmooth Vector-Valued Mappings," Soviet Mathematics Doklady 19, (1978) 1057-1060.

32. REILAND, T. W. "Nonsmooth Analysis and Optimization for a Class of Nonconvex Mappings," in Anderson, E. J. and Philpot, A. B., eds., Proceedings of the International Conference on Infinite-Dimensional Programming, Cambridge University, September 1984, Springer-Verlag, (1985).

33. REILAND, T. W. "Nonsmooth Analysis of Vector-Valued Mappings with Contributions to Nondifferentiable Programming," Numerical Functional Analysis and Optimization 3-4, (1985-86) 301-323.

34. PERESSINI, A. L. Ordered Topological Vector Spaces. Harper and Row, New York, (1967).

35. TAYLOR, A. E. Functional Analysis. Wiley and Sons, New York, (1958).

36. THIBAULT, L. "Subdifferentials of Compactly Lipschitzian Vector-Valued Functions," Annali di Matematica pura ed applicata CSSV, (1980) 157-192.

37. GUIGNARD, M. "Generalized Kuhn-Tucker Conditions for Mathematical Programming Problems in a Banach Space," SIAM Journal on Control and Optimization 7, (1969) 232-241.

38. HIRIART-URRUTY, J. B. "Tangent Cones, Generalized Gradients and Mathematical Programming in Banach Spaces," Mathematics of Operations Research 4, (1979a) 79-97.

39. HIRIART-URRUTY, J. B. "New Concepts in Nondifferentiable Programming," Bulletin de la Societe Mathematique de France 50, (1979b) 57-85.

40. IOFFE , A. D. "Necessary and Sufficient Conditions for a Local Minimum .1: A Reduction Theorem and First-Order Conditions," SIAM Journal on Control and Optimization 17, (1979) 245-250.

41. DEMYANOV, V. F. and RUBINOV, A. M. "On Quasidifferentiable Functionals," Soviet Mathematics Doklady 21, (1980) 14-17.

42. POLYAKOVA, L. N. "Necessary Conditions for an Extremum of Quasidifferentiable Functions," Vestnik Leningrad University Mathematics 13, (1981) 241-247.

43. DUNFORD, N. and SCHWARTZ, J. T. Linear Operators Part I: General Theory, Wiley and Sons, New York, (1964).

44. DEMYANOV, V. F. "Quasidifferentiable Functions: Necessary Conditions and Descent Directions," in Demyanov, V. F. and Dixon, L. C. W., eds., Quasidifferential Calculus, North Holland, Amsterdam, 20-43, (1986).