GENERALIZED DIFFERENTIABILITY FOR A CLASS OF NONDIFFERENTIABLE OPERATORS WITH APPLICATIONS TO NONSMOOTH OPTIMIZATION

THOMAS W. REILAND

Communicated by B. Mond

Abstract

A theory of generalized gradients is presented for a class of Lipschitz vector-valued mappings from a Banach space to a locally convex order complete vector lattice. Necessary optimality conditions are obtained for nonconvex programming problems on Banach spaces with vector-valued operator constraints and/or an arbitrary set constraint. Sufficient optimality conditions are also obtained under mild convexity assumptions.

1980 *Mathematics subject classification (Amer. Math. Soc.)* (1985 *Revision*): Primary 49 B 27, Secondary 90 C 48.

1. Introduction and summary

The desire to optimize increasingly sophisticated models of complex man-made and naturally occurring systems has motivated considerable interest over the past quarter century in the area of research now frequently referred to as "nonsmooth analysis and optimization." Briefly put, this area's primary goal is to include broader classes of mappings in a suitably generalized theory of differentiation that is conducive to the analysis of optimization problems. Initial results by Rockafellar [34–38], Moreau [27] and McLinden [26] dealt with convex functions; Valadier [51], Ioffe-Levin [23], Zowe [54–55], Kutateladze [25], Rubinov [44], Borwein [1] and Papageorgiou [28] made important generalizations to convex mappings into ordered vector spaces. Clarke [3–7]

introduced a generalized gradient for lower semicontinuous and locally Lipschitz functions whose analytical virtues were recognized from the outset. A partial list of other significant contributions to the nonconvex case includes the works of Halkin [11], Hiriart-Urruty [14], Ioffe [19–22], Kusraev [24], Papageorgiou [29], Penot [30], Rockafellar [39–41], Thibault [46–50], Warga [53] and many others. For excellent summaries of the theory, motivation, and application of generalized gradients and extensive references we refer the reader to Clarke [7], Hiriart-Urruty [15] and Rockafellar [42].

In this paper we study a mapping $f: X \to V$ that has a Lipschitz-type property at $\overline{x} \in X$, where X is a Banach space and V is a locally convex ordered topological vector space that is an order complete vector lattice with a normal positive cone. We provide results for a generalized directional derivative $f^\circ(\overline{x}; y)$ and generalized gradient $\partial f(\overline{x})$ and derive optimality conditions for programming problems with vector-valued operator constraints and/or an arbitrary set constraint. Our assumptions differ from those in previous works that have dealt with vector-valued mappings. For example, Kusraev [24] uses the absolute value operator in his definition of Lipschitz mapping, whereas we assume that the difference quotient $\lambda^{-1}[f(x + \lambda y) - f(x)]$ has its values in an order interval of V for all positive sufficiently small λ, for all x in a neighborhood of $\overline{x}$ and y in a neighborhood of $\theta \in X$. In Section 2 we show that this property is less stringent than the Lipschitz definition employed by Kusraev. In [46] and [47] Thibault studies compactly Lipschitzian mappings, while in [49] he takes an approach based on a theorem of Sweetser.

The paper is organized as follows. Section 2 contains the main results concerning directional derivatives, generalized gradients and optimality conditions; only a few simple proofs are given in this section. Section 3 provides proofs and examples while Section 4 includes a summary and brief comments concerning future research.

2. Main results

In this section we present the main results concerning directional derivatives, generalized gradients and optimality conditions. Proofs of some results are included in this section; the remaining proofs are given in Section 3.

2.1. DIRECTIONAL DERIVATIVES AND GENERALIZED GRADIENTS. Let U be an open subset of a Banach space X and let V be 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 $V_+ = \{v: v \geq 0\}$ that is closed. We also assume V is an order complete vector lattice (ocvl) with a normal positive cone; that is, V is an ocvl with a neighborhood

basis $\mathcal{W}$ of the origin θ such that, for $W \in \mathcal{W}$, $W = (W + V_+) \cap (W - V_+)$. Equivalently, the normality of V_+ implies the existence of a neighborhood basis of the origin consisting of sets W such that if $x, z \in W$, then any y satisfying $x \leq y \leq z$ is also in W. A consequence of the normality of V_+ that will be used in the sequel is that if $\{x_k\}_{k \in K}$, $\{y_k\}_{k \in K}$ and $\{z_k\}_{k \in K}$ are nets in V satisfying $x_k \leq y_k \leq z_k$ for all $k \in K$, and if $\lim_{k \in K} x_k = \lim_{k \in K} z_k = \theta$, then $\lim_{k \in K} y_k = \theta$ [31, page 62, Corollary 1.3].

DEFINITION 1. The mapping $f: X \to V$ is *order Lipschitz on* U if there exist a neighborhood W of $\theta \in X$, $\varepsilon > 0$, and $m \leq M$ in V such that $\lambda^{-1}[f(x + \lambda y) - f(x)] \in [m, M]$ for all $\lambda \in (0, \varepsilon]$, $x \in U$, and $y \in W$. The mapping f is *order Lipschitz at* $\overline{x} \in X$ if there exists a neighborhood N of $\overline{x}$ such that f is order Lipschitz on N, and *locally order Lipschitz on* U if f is order Lipschitz at $\overline{x}$ for every $\overline{x} \in U$.

Kusraev [24] employs a definition of a Lipschitz mapping using the absolute value operator $|\cdot|: V \to V$, where $|v| = v^+ + v^-$ with $v^+ = \sup(v, \theta)$ and $v^- = \sup(-v, \theta)$: $f: X \to V$ is locally Lipschitz if, for any $\overline{x} \in X$, there exists a neighborhood N of $\overline{x}$ and $k \in V_+$ such that $|f(x') - f(x'')| \leq k\|x' - x''\|$ for all $x', x'' \in N$. If $V = R$, this definition reduces to the definition of a real-valued locally Lipschitz function. If f is locally Lipschitz using the absolute value operator, choose a neighborhood N_0 of $\overline{x}$, a neighborhood $W = \{y \in X: \|y\| < \delta\}$, $\delta > 0$, of $\theta \in X$, and $\varepsilon > 0$ such that $x + \lambda y \in N$ for all $x \in N_0$, $y \in W$ and $\lambda \in (0, \varepsilon]$. Then for all $x \in N_0$, $y \in W$ and $\lambda(0, \varepsilon]$ we have $|f(x + \lambda y) - f(x)| \leq k\|\lambda y\| = \lambda k\|y\|$, which implies $-\delta k \leq -k\|y\| \leq \lambda^{-1}[f(x + \lambda y) - f(x)] \leq k\|y\| \leq \delta k$; thus $\lambda^{-1}[f(x + \lambda y) - f(x)] \in [m, M]$, where $m = -\delta k$ and $M = \delta k$. Hence, if f is locally Lipschitz using the absolute value operator, then f is locally order Lipschitz as defined in Definition 1.

An order Lipschitz mapping need not be continuous, as the following example shows.

EXAMPLE 1. 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 $\mathcal{T}((c), \ell^1)$, that is, the weak topology on (c). Since ℓ^1 is the dual of (c), ℓ^1 is norm-determining for (c), hence by W is bounded relative to the norm topology. In particualr, 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 $\overline{B} = \{x: \|x\| \leq 1\}$ 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 order Lipschitz on (c). However, $f(x)$ is not continuous since the dual of (c) is not the sequence space $\phi = \{x = (x_n): x_n = 0$ for all but a finite number of choices of $n\}$ [31, page 135]. We note that if E is a normed lattice, then the lattice operations are weakly continuous if and only if E is finite dimensional; also, if $o(E, E')$ denotes the topology of uniform convergence on all order intervals in E', then if E is a locally convex vector lattice, $o(E, E')$ is the coarsest translation invariant topology on E finer than the weak topology $\mathcal{T}(E, E')$ and for which the lattice operations are continuous [45, page 254].

When defining a directional derivative our preference is to allow consideration of any "direction" $y \in X$, not just those in the neighborhood W in Definition 1. Thus, we note that W is radial, that is, for each $y \in X$ there exists a scalar $\alpha_y > 0$ such that $\alpha y \in W$ for all $|\alpha| \le \alpha_y$ [31, page 192].

DEFINITION 2. If $f: X \to \mathcal{V}$ is order Lipschitz at $\overline{x}$, the *generalized directional derivative of f at $\overline{x}$ in the direction* $y \in X$, denoted $f^\circ(\overline{x}; y)$, is given by

$$f^\circ(\overline{x}; y) = \inf_{\substack{N \in \mathcal{N}(\overline{x}) \\ \varepsilon > 0}} \sup_{\substack{x \in N \\ 0 < \lambda \le \varepsilon}} \lambda^{-1}[f(x + \lambda y) - f(x)]$$

where $\mathcal{N}(\overline{x})$ is a neighborhood base of $\overline{x}$ in X.

If V is a Daniell space, that is, if every decreasing net $(W_j)_{j \in J}$ that is bounded below converges (in which case $\lim_{j \in J} W_j = \inf_{j \in J} W_j$) [23], then we can write

$$f^\circ(\overline{x}, y) = \lim_{\substack{\varepsilon \to 0^+ \\ N \in \mathcal{N}(\overline{x})}} \left(\sup_{\substack{x \in N \\ 0 < \lambda \le \varepsilon}} \lambda^{-1}[f(x + \lambda) - f(x)] \right).$$

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

We shall now examine several situations in which the generalized directional derivative is a continuous mapping. For $f: X \to V$, we define the epigraph of f, denoted epi f, by

$$\text{epi } f = \{(x, v) \in X \times V | v \ge f(x)\}.$$

PROPOSITION 2. *If int epi $f^\circ(\overline{x}; \cdot)$ is nonempty, then $f^\circ(\overline{x}; \cdot)$ is a continuous mapping from X to V.*

The assumption that int epi $f^\circ(\overline{x}; \cdot)$ is nonempty implies that int V_+ is nonempty, so this assumption is rather stringent.

The *interior pseudo-tangent cone* of epi f at $(\overline{x}, f(\overline{x}))$, denoted

$$I_{\text{epi}\,f}(\overline{x}, f(\overline{x})),$$

is the (open) set of points $(y, v) \in X \times V$ for which there exists a neighborhood E of $(\overline{x}, f(\overline{x}))$, a neighborhood F of (y, v), and a real number $\varepsilon > 0$ such that $E \cap \text{epi}\,f + (0, \varepsilon)F \subset \text{epi}\,f$ [48]. In the language of Rockafellar [41], $I_{\text{epi}\,f}(\overline{x}, f(\overline{x}))$ is the set of points $(y, v) \in X \times V$ with respect to which epi f is *epi-Lipschitzian* at $(\overline{x}, f(\overline{x}))$, that is, there exists a neighborhood G of (y, v), a neighborhood J of $(\overline{x}, f(\overline{x}))$, and $\lambda > 0$ such that

(2.1) $(x', w') + t(y', v') \in \text{epi}\,f$

for all $(x', w') \in \text{epi}\,f \cap J$, $t \in (0, \lambda)$ and $(y', v') \in G$. Each vector (y', v') satisfying (2.1) is said to be in the *hypertangent cone of epi f at $(\overline{x}, f(\overline{x}))$*, denoted $H_{\text{epi}\,f}(\overline{x}, f(\overline{x}))$. Note that

$$I_{\text{epi}\,f}(\overline{x}, f(\overline{x})) \subseteq H_{\text{epi}\,f}(\overline{x}, f(\overline{x})) \subseteq T_{\text{epi}\,f}(\overline{x}, f(\overline{x})),$$

where $T_{\text{epi}\,f}(\overline{x}, f(\overline{x}))$ is the (Clarke) *tangent cone*: if S is a nonempty subset of Y, a locally convex Hausdorff topological vector space, then $y \in Y$ is in the *tangent cone* of S at $\overline{x} \in \text{cl}\,S$, denoted $T(S; \overline{x})$ or $T_S(\overline{x})$, if for every net $(x_j)_{j \in J}$ in $\text{cl}\,S$ converging to $\overline{x}$ and every net $(t_j)_{j \in J}$ of positive real numbers converging to zero there exists a net $(y_j)_{j \in J}$ converging to y such that $x_j + t_j y_j \in S$ for each $j \in J$ [41]. If Y is a Banach space we can replace nets with sequences in the above definition and obtain the formulation given by Hiriart-Urruty [14]. $T(S; \overline{x})$ is always a closed convex cone.

If $I_{\text{epi}\,f}(\overline{x}, f(\overline{x})) \neq \varnothing$, then since $T_{\text{epi}\,f}(\overline{x}, f(\overline{x}))$ is convex and

$$\text{int}\,T_{\text{epi}\,f}(\overline{x}, f(\overline{x})) \subseteq I_{\text{epi}\,f}(\overline{x}, f(\overline{x}))\qquad [41],$$

we have $I_{\text{epi}\,f}(\overline{x}, f(\overline{x})) = \text{int}\,H_{\text{epi}\,f}(\overline{x}, f(\overline{x}))$. In addition, note that the definition of $f^\circ(\overline{x}; \cdot)$ implies $f^\circ(\overline{x}; y) = \inf\{v \in V | (y, v) \in H_{\text{epi}\,f}(\overline{x}, f(\overline{x}))\}$, and if V is Daniell, $H_{\text{epi}\,f}(\overline{x}, f(\overline{x})) \supset \{(y, v) \in X \times V | f^\circ(\overline{x}; y) \in v - \text{int}\,V_+\}$ since

$$\inf_{\substack{N \in \mathcal{N}(\overline{x}) \\ \varepsilon > 0}} \sup_{\substack{x \in N \\ 0 < \lambda \leq \mathcal{E}}} \lambda^{-1}[f(x + \lambda y) - f(x)]$$

$$= \lim_{\substack{N \in \mathcal{N}(\overline{x}) \\ \varepsilon \to 0^+}} \sup_{\substack{x \in N \\ 0 < \lambda \leq \mathcal{E}}} \lambda^{-1}[f(x + \lambda y) - f(x)] \in v - \text{int}\,V_+.$$

Hence, the interior of $H_{\text{epi}\,f}(\overline{x}, f(\overline{x}))$ is the same as the interior of $\text{epi}\,f^\circ(\overline{x}; \cdot)$.

To summarize, if $I_{\text{epi}\,f}(\overline{x}, f(\overline{x})) \neq \varnothing$, then

$$I_{\text{epi}\,f}(\overline{x}, f(\overline{x})) = \text{int}\,H_{\text{epi}\,f}(\overline{x}, f(\overline{x})),$$

and if V is Daniell, then $\text{int}\,H_{\text{epi}\,f}(\overline{x}, f(\overline{x})) = \text{int}\,\text{epi}\,f^\circ(\overline{x}; \cdot)$. We have established the following corollary to Proposition 2.

COROLLARY 1. *If $I_{epi\,f}(\overline{x}, f(\overline{x}))$ is nonempty and V is Daniell, then $f^\circ(\overline{x}; \cdot)$ is a continuous mapping from X to V.*

REMARKS. 1. A sufficient condition that $I_{epi\,f}(\overline{x}, f(\overline{x}))$ be nonempty is that the mapping $f: X \to V$ be locally Lipschitz at $\overline{x}$ with the absolute value operator $|\cdot|: V \to V$ and that int $V_+ \neq \varnothing$. In fact, if there exists $k \in V_+$ and $N \in \mathcal{N}(\overline{x})$ such that

$$|f(x'') - f(x')| \le k\|x'' - x'\|, \quad \forall x', x'' \in N,$$

then there exist a real number $\varepsilon > 0$ and $v \in$ int V_+ such that $f(x' + ty') \le f(x') + tv$ for all $x' \in \overline{x} + \varepsilon B$, $y' \in \varepsilon B$ and $t \in [0, \varepsilon]$, where $B = \{x \in X| \|x\| < 1\}$. Hence $f(x' + ty') \le z + w$ when $(x', z) \in [(\overline{x} + \varepsilon B) \times N_1] \cap \text{epi} f$, $t \in [0, \varepsilon]$, and $(y', w) \in (\varepsilon B) \times N_2$, where $N_1 \in \mathcal{N}(f(\overline{x}))$ and $N_2 = v + \text{int} V_+ \in \mathcal{N}(2v)$. Thus $(\theta, 2v) \in I_{epi\,f}(\overline{x}, f(\overline{x}))$.

2. The Lipschitz property of $f: X \to V$ at $\overline{x}$ with the absolute value operator implies the continuity of $f^\circ(\overline{x}; \cdot)$ in a more direct manner. Indeed, the Lipschitz property implies $|f^\circ(\overline{x}; y| \le k\|y\|$ for any $y \in X$; thus $-k\|y\| \le f^\circ(\overline{x}; y) \le k\|y\|$ and, since V_+ is normal, $\lim_{y \to \theta} f^\circ(\overline{x}; y) = \theta$. Hence $f^\circ(\overline{x}; \cdot)$ is a sublinear mapping continuous at the origin and thus continuous on X by [47, Lemma 2.4].

The continuity of the generalized directional derivative $f^\circ(\overline{x}; \cdot)$ leads to several results concerning the generalized gradient. Hence, we make the following definition.

DEFINITION 3. The mapping $f: X \to V$ is *regular* at $\overline{x}$ if f is order Lipschitz at $\overline{x}$ and if $f^\circ(\overline{x}; \cdot)$ is a continuous mapping from X to V.

Denote by $L(X, V)$ the vector space of continuous linear mappings from X to V; $L_s(X, V)$ denotes this space endowed with the topology of pointwise convergence.

DEFINITION 4. Let $f: X \to V$ be order Lipschitz at $\overline{x}$. The *generalized gradient of f at $\overline{x}$*, denoted $\partial f(\overline{x})$, is the set of al $T \in L(X, V)$ satisfying $T(y) \le f^\circ(\overline{x}; y)$, for all $y \in X$.

PROPOSITION 3. *The generalized gradient $\partial f(\overline{x})$ of f at $\overline{x}$ is convex and satisfies $-\partial f(\overline{x}) = \partial(-f)(\overline{x})$.*

PROOF. The convexity of $\partial f(\overline{x})$ follows directly from the definition;

$$-\partial f(\overline{x}) = \partial(-f)(\overline{x})$$

is a consequence of Proposition 1 which states that $f^\circ(\overline{x}; -y) = (-f)^\circ(\overline{x}, y)$ for all $y \in X$.

PROPOSITION 4. *If the mapping $f: X \to V$ is regular at $\overline{x}$, then the generalized gradient $\partial f(\overline{x})$ is the set of all linear mappings $T: X \to V$ such that $T(y) \leq f^\circ(\overline{x}, y)$ for all $y \in X$.*

THEOREM 1. *If the mapping $f: X \to V$ is regular at $\overline{x}$, then the generalized gradient $\partial f(\overline{x})$ is a nonempty, closed, convex, equicontinuous subset of $L_s(X, V)$ with $f^\circ(\overline{x}; y) = \max\{T(y) | T \in \partial f(\overline{x})\}$. If, in addition, the order intervals in V are compact, then $\partial f(\overline{x})$ is compact in $L_s(X, V)$.*

REMARK. Note that since $\partial f(\overline{x})$ is equicontinuous, it is bounded in $L_s(X, V)$ [45, page 83].

2.2. OPTIMALITY CONDITIONS. In this section we give necessary and sufficient optimality conditions for nondifferentiable mathematical 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 [24] and Thibault [46], 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" [47, Definition 1.1]. In addition, our proof of the Kuhn-Tucker necessary conditions (Theorem 2), which recalls a paper of Guignard [10], does not explicitly use the assumption 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 $\overline{x}$ any closed convex subset $\Gamma_g(\overline{x})$, 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 [20] (where such multifunctions are called *fans*) and lead to necessary conditions which have as special cases the necessary conditions of Clarke [5], Hiriart-Urruty [12, 14] and Ioffe [19]. Indeed, Ioffe [22] has successfully merged this approach with the above-mentioned *fans* to develop more general necessary conditions.

Let X be a Banach space, 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 Lipschitz at $\overline{x} \in S$, that is, there exists a neighborhood N of $\overline{x}$ and a constant $K > 0$ such that $|f(x) - f(y)| \leq K\|x - y\|$ for all $x, y \in N$. Consider the problem

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

$\overline{x}$ is a local minimum of f on S if f is finite at $\overline{x}$ and if there exists a neighborhood N of $\overline{x}$ such that $f(x) \geq f(\overline{x})$ for every $x \in S \cap N$; $\overline{x}$ is a

minimum of f on S if f is finite at $\overline{x}$ and $f(x) \geq f(\overline{x})$ for every $x \in S$. If C is a nonempty subset of Y, a locally convex Hausdorff topological vector space, a vector $y \in Y$ is an *adherent displacement* for C at $\overline{x} \in \operatorname{cl} C$ if there exist a net $(t_j)_{j \in J}$ of positive real numbers converging to zero and a net $(y_j)_{j \in J}$ converging to y such that $\overline{x} + t_j y_j \in C$ for all j. Following Rockafellar [42, page 14], we denote the set of adherent displacements for C at $\overline{x}$ by $K(C; \overline{x})$ and refer to $K(C; \overline{x})$ as the *contingent cone* of C at $\overline{x}$. If Y is a Banach space then, just as for the tangent cone $T(C; \overline{x})$ defined previously, we can replace nets with sequences and obtain the formulation given by Hiriart-Urruty [14]. $K(S; \overline{x})$ is a closed cone with vertex θ and satisfies $T(S; \overline{x}) \subseteq K(S; \overline{x})$. The closure of $\operatorname{co} K(S; \overline{x})$, the convex hull of $K(S; \overline{x})$, is denoted $P(S; \overline{x})$. The *polar cone* of a nonempty set A of X is given by $A^\circ \equiv \{x^* \in X^* | x^*(x) \leq 0 \ \forall x \in A\}$, where X^* denotes the topologial dual of X. If A is empty, $A^0 = X^*$. If A^* is a nonempty set of X^*, the *prepolar* of A^* is $^\circ(A^*) \equiv \{x \in X | x^*(x) \leq 0 \ \forall x^* \in A^*\}$. If A^* is empty, $^\circ(A^*) = X$. $A^\circ(^\circ(A^*))$ is a weak*-closed (weakly closed) convex cone in $X^*(X)$.

We begin our study of optimality conditions with two results, needed in the sequel, that are special cases of necessary conditions given by Hiriart-Urruty.

PROPOSITION 5 [14, Theorem 6]. *If $\overline{x}$ is a local minimum of f on S, then*

$$(2.2) \qquad\qquad f^\circ(\overline{x}; d) \geq 0 \quad \forall d \in K(S; \overline{x}).$$

Necessary conditions are also often expressed in terms of the polar cones of subsets of the adherent displacement set.

PROPOSITION 6 [14, Theorem 8]. *Let $\overline{x}$ be a local minimum of f on S and M a convex cone contained in $K(S; \overline{x})$; then*

$$(2.3) \qquad\qquad 0 \in \partial f(\overline{x}) + M^0.$$

REMARKS 1. Condition (2.3) is sharpest when $K(S; \overline{x})$ is convex, in which case (2.3) becomes

$$(2.4) \qquad\qquad 0 \in \partial f(\overline{x}) + [K(S; \overline{x})]^\circ.$$

If, in addition, f is continuously differentiable at $\overline{x}$, then $\partial f(\overline{x}) = \{\nabla f(\overline{x})\}$ [5, Proposition 4] and (2.4) reduces to $0 \in \nabla f(\overline{x}) + [K(S; \overline{x})]^\circ$, that is, $\nabla f(\overline{x}) \in -[K(S; \overline{x})]^\circ$ which, since $[K(s; \overline{x})]^\circ = [P(S; \overline{x})]^\circ$, is the well-known optimality condition in differentiable programming given by Guignard [10] and Varaiya [52].

2. To establish the optimality condition in differentiable programming noted in remark 1, it is *not* necesary to assume that $K(S; \overline{x})$ is convex. The

convexity requirement is needed in the nondifferentiable case since $f°(\bar{x}; d)$ ≥ 0 for all $d \in K(S; \bar{x})$ cannot be extended to clco $K(S, \bar{x}) = P(S; \bar{x})$. Thus, for nondifferentiable objective functions, relation (2.4) does not hold without the convexity of $K(S; \bar{x})$. To illustrate this fact we include an example due to Hiriart-Urruty [12, page 80]. Let $X = E^2$, $f: E^2 \to R$ is given by $f(x_1, x_2) = 1 - \exp(x_2 - |x_1|)$, $S = \{(x_1, x_2) \in E^2: x_2 - |x_1| \leq 0\}$; $\bar{x} = (0, 0)$ is a minimum of f on S, $[K(S; \bar{x})]° = \{(0, 0)\}$, and $\partial f(\bar{x}) = co\{(1, -1), (-1, -1)\}$.

A statement of sufficient conditions requires the following preliminaries. A locally Lipschitz function $f: X \to R$ on a Banach space X is *locally pseudo-convex at* x_0 (see [15, page 75]) if there exists a neighborhood N of x_0 such that

$$\forall x \in N, \; f°(x_0; x - x_0) \geq 0 \Rightarrow f(x) \geq f(x_0).$$

If this property holds for all $x_0 \in X$, then f is *locally pseudoconvex*. If the relation holds globally, that is, if $N = X$, f is *pseudoconvex*. A subset $A \subset X$ is *pseudoconvex at* $x_0 \in cl\, A$ if $x - x_0 \in P(A; x_0)$ for all $x \in A$, and *strictly pseudoconvex at* x_0 if $x - x_0 \in K(A; x_0)$ for all $x \in A$.

PROPOSITION 7. *Suppose f is pseudoconvex on S and S is pseudoconvex at $\bar{x} \in S$; then $0 \in \partial f(\bar{x}) + [P(S; \bar{x})]°$ is a sufficient condition for $\bar{x}$ to be a minimum of f on S.*

REMARKS. 1. If f is also continuously differentiable, the sufficient condition becomes $\nabla f(\bar{x}) \in -[P(S; \bar{x})]°$, which is the sufficient condition given by Guignard [10, Theorem 1].

2. A "local minimum" analogue of the above result follows directly if f is locally pseudoconvex at $\bar{x}$ and if S is locally pseudoconvex at $\bar{x}$, where the latter means that there exists a neighborhood N of $\bar{x}$ such that $x - \bar{x} \in P(S; \bar{x})$ for all $x \in S \cap N$. Hiriart-Urruty [15, Theorem 15] states that $0 \in \partial f(\bar{x}) + [K(S; \bar{x})]°$ (note that $[K(S; \bar{x})]° = [P(S; \bar{x})]°$) is a sufficient condition for $\bar{x}$ to be a *local* minimum of f under the assumptions that f be locally pseudoconvex at $\bar{x}$ and that S be locally *strictly* pseudoconvex at $\bar{x}$; this latter condition is termed "condition L" by Hiriart-Urruty.

3. A more desirable sufficient condition is possible in Proposition 7, but it is acquired at the expense of strengthening the assumption on S by using the (Clarke) tangent cone $T(S; \bar{x})$. If f is pseudoconvex on S (as in Proposition 7) and if $x - \bar{x} \in T(S; \bar{x})$ for all $x \in S$, then $0 \in \partial f(\bar{x}) + [T(S; \bar{x})]^0$ 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 $T(S; \bar{x}) = K(S; \bar{x}) = P(S; \bar{x})$ [14, page 83] and the sufficient condition immediately above is equivalent to the sufficient condition in Proposition 7.

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 with a normal positive cone. A and B are nonempty subsets in X and V, respectively, and $g: X \to V$ is order Lipschitz at $\overline{x} \in S$ where $S = \{x \in A | g(x) \in B\}$. Let

$$J = \{x \in X | T(x) \in P(B; g(\overline{x})) \text{ for each } T \in \partial g(\overline{x})\}$$

and

$$H^* = \{h \in X^* | h \in \mu \partial g(\overline{x}), \mu \in [P(B; g(\overline{x}))]^\circ\},$$

where $\mu \partial g(\overline{x}) = \{\mu \circ T | T \in \partial g(\overline{x})\}$. 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 $G \cap J = \operatorname{cl} M$ for a convex cone M contained in $K(S; \overline{x})$, where $K(S; \overline{x})$ is the contingent cone for S at $\overline{x} \in S$; suppose also that $G^\circ + J^\circ$ is closed. If $\overline{x}$ is a local minimum of f over S, then there exists $\mu \in [P(B; g(\overline{x}))]^\circ$, where $[P(B; g(\overline{x}))]^\circ$ is the polar cone of the closure of the convex hull of $K(B; g(\overline{x}))$, such that $0 \in \partial f(\overline{x}) + \mu \partial g(\overline{x}) + G^\circ$.*

REMARK. Theorem 2 generalizes to nondifferentiable problems with operator constraints Guignard's well-known Kuhn-Tucker conditions for differentiable programming [10, Theorem 3] which hold for $M = P(S; \overline{x})$. Beyond the generalization to nondifferentiable problems, Theorem 2 is applicable to a wider class of programs, though in general it is not as sharp as the result in the differentiable case. For example, in the extreme case we can always choose $G = M = \{0\}$, but then $G^0 = X^*$. If $G = T(S; \overline{x})$, then $0 \in \partial f(\overline{x}) + \mu \partial g(\overline{x}) + [T(S; \overline{x})]^\circ$ and we have a result consistent with the necessary condition $0 \in \partial f(\overline{x}) + [T(S; \overline{x})]^\circ$ established by Clarke in a slightly different form [5, Lemma 2]. The "larger" G is, the sharper the result in Theorem 2; hence the strongest conclusion in general occurs when $K(S; \overline{x})$ is itself convex and $M = K(S; \overline{x})$; then $\operatorname{cl} M = P(S; \overline{x})$ and our result is identical to the optimality condition in differentiable programming. However, in the differentiable case it is not necessary to assume $K(S; \overline{x})$ is convex (see the remark following Proposition 6).

Sufficient conditions are obtained by imposing mild convexity assumptions.

THEOREM 3. *If G is a closed convex cone in X such that $x - \overline{x} \in G$ for all $x \in S$, if there exists $\mu \in [P(B; g(\overline{x}))]^\circ$ such that $0 \in \partial f(\overline{x}) + \mu \partial g(\overline{x}) + G^\circ$, if S is strictly pseudoconvex at $\overline{x}$ and $T(K(S; \overline{x})) \subset K(B; g(\overline{x}))$ for all $T \in \partial g(\overline{x})$, and if f is pseudoconvex on S, then $\overline{x}$ is optimal for f over S.*

3. Examples and proofs

3.1. EXAMPLES AND PROOFS FOR SECTION 2.1. In this section we discuss situations which give rise to order Lipschitz functions, give examples of generalized gradients, and prove Propositions 1, 2, 4, and Theorem 1.

As before, we consider a mapping $f: X \to V$ where unless indicated otherwise, X is a Banach space and V is a locally convex ordered topological vector space that is an order complete vector lattice with a closed, convex normal positive cone V_+. The mapping f is *convex* if

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

for each $x_1, x_2 \in X$ and each $\lambda \in [0, 1]$; we say f is *closed* if the epigraph of f is closed in the product topology on $X \times V$. If in addition V is a Banach space and $B_V = \{v \in V \mid \|v\| \leq 1\}$, then following Robinson [33, pages 140–141], we set $I_{V_+} \equiv (B_V + V_+) \cap (B_V - V_+)$ and say that the mapping $f: X \to V$ is *locally V_+-Lipschitzian at* $\overline{x} \in X$ if there exists a neighborhood N of $\overline{x}$ and a constant k such that $f(x) - f(y) \in k\|x - y\|I_{V_+}$ for each $x, y \in N$. If I_{V_+} is order bounded (a rather severe restriction) and f *is locally V_+-Lipschitzian at* $\overline{x}$, *then* f is order Lipschitz at $\overline{x}$. If $f: X \to V$ is closed and convex and I_{V_+} is order bounded, then f is locally order Lipschitz on X [33, Theorem 5].

The mapping $f: X \to V$ is *strictly differentiable at* $\overline{x}$ *in the Hadamard sense* if there exists a continuous linear mapping $\nabla f(\overline{x}): X \to V$ such that the mappings

$$\phi_{x,\lambda}(y) = \frac{f(x + \lambda y) - f(x)}{\lambda} - \nabla f(\overline{x})y$$

converge to θ uniformly on all compact y-sets. If X is a Banach lattice and the cone K of positive continuous linear mappings in $L(X, V)$ is generating, that is, $K - K = L(X, V)$, then $L(X, V) \subset L^b(X, V)$, where $L^b(X, V)$ denotes the space of order-bounded linear mappings from X to V [31, page 173]. If, in addition, there is an order-bounded y-neighborhood W of $\theta \in X$ where the above convergence is uniform and a neighborhood Z of $\theta \in V$ such that $Z + [a, b]$ is order-bounded, where $\nabla f(\overline{x})(W) \subseteq [a, b]$, then f is order Lipschitz at $\overline{x}$. We note that if f is *strictly differentiable at* $\overline{x}$ *in the full limit sense*, then the same holds in the Hadamard sense [40, page 346].

We now proceed with the proofs of Propositions 1, 2, and 4 and Theorem 1.

PROOF OF PROPOSITION 1 The proof of the sublinearity of $f^\circ(\overline{x}; \cdot)$ follows that for real-valued Lipschitz functions, while $f^\circ(\overline{x}; y) \leq M$ for all $y \in W$ follows directly from Definitions 1 and 2. For any given $y \in X$, $f^\circ(\overline{x}; \alpha_y y) =$

$\alpha_y f^\circ(\overline{x}, y) \leq M$, so $f^\circ(\overline{x}; y) \leq \alpha_y^{-1} M$ and thus $f^\circ(\overline{x}; y) \in V$. Finally,

$$(-f)^\circ(\overline{x}, y) = \inf_{\substack{N \in \eta(\overline{x}) \\ \varepsilon > 0}} \sup_{\substack{x \in N \\ 0 < \lambda \leq \varepsilon}} \lambda^{-1}[-f(x + \lambda y) + f(x)]$$

$$= \inf_{\substack{N \in \eta(\overline{x}) \\ \varepsilon > 0}} \sup_{\substack{x \in N \\ 0 < \lambda \leq \varepsilon}} \lambda^{-1}[f(x + \lambda y + \lambda(-y)) - f(x + \lambda y)]$$

$$= f^\circ(\overline{x}; -y).$$

PROOF OF PROPOSITION 2. Since V_+ is normal, the order intervals $[v, w] = \{z \in V | v \leq z \leq w\}$ in V are bounded in the topology of V [31, page 62]. Therefore, since $f^\circ(\overline{x}, \cdot)$ is convex, it is continuous on X if it is bounded above in a neighborhood of one point [51, page 71]. But the interior of epi $f^\circ(\overline{x}; \cdot)$ is included in the set of $(y, v) \in X \times V$ such that $f^\circ(\overline{x}; \cdot)$ is bounded above by v in a neighborhood of y.

PROOF OF PROPOSITION 4. Suppose $T: X \to V$ is a linear mapping satisfying $T(y) \leq f^\circ(\overline{x}; y)$ for all $y \in X$. By the linearity of T, $-T(y) = T(-y) \leq f^\circ(\overline{x}, -y)$, thus $-f^\circ(\overline{x}; -y) \leq T(y) \leq f^\circ(\overline{x}, y)$. Since V_+ is normal and $f^\circ(\overline{x}; \cdot)$ is continuous, $\lim_{y \to \theta} T(y) = \theta$ and hence T is continuous on X.

PROOF OF THEOREM 1. Observe that $\partial f(\overline{x})$ is the convex analytic subdifferential of $f^\circ(\overline{x}; \cdot)$ at zero. Then since f is assumed regular at $\overline{x}$, the theorem follows from Théorème 6 and Corollaire 7 in [51].

We conclude this section with two examples of generalized gradients.

EXAMPLE 2. Let $X = V = C[0, 1]$, the Banach space of continuous real-valued functions on the interval $[0, 1]$ with the supremum norm. The cone of nonnegative functions in $C[0, 1]$ is normal for the topology generated by the above norm [31, page 65]. Let $\phi: [0, 1] \times R \to R$ be measurable in t for any x in R and assume there exists a nonnegative function $k(\cdot) \in L^1[0, 1]$ and an $\varepsilon > 0$ such that

$$|\phi(t, y) - \phi(t, z)| \leq k(t)|y - z|$$

whenever $|y - \overline{x}(t)| < \varepsilon$, $|z - \overline{x}(t)| < \varepsilon$, and $0 \leq t \leq 1$. Define the mapping

$$f(x(\cdot)): C[0, 1] \to C[0, 1]$$

by

$$f(x(\cdot))(t) = \int_0^t \phi(\tau, x(\tau)) \, d\tau;$$

then f is Lipschitz at $\overline{x}(\cdot)$ with the absolute value operator. Indeed, for any $t \in [0, 1]$ and $z(\cdot)$, $y(\cdot)$ in an ε-neighborhood of $\overline{x}(\cdot)$,

$$
\begin{aligned}
f(z(\cdot))(t) - f(y(\cdot))(t) &= \int_0^t [\phi(\tau, z(\tau)) - \phi(\tau, y(\tau))] \, d\tau \\
&\leq \int_0^t k(\tau)|z(\tau) - y(\tau)| \, d\tau \\
&\leq \int_0^1 k(\tau)|z(\tau) - y(\tau)| \, d\tau \\
&\leq \|z(\cdot) - y(\cdot)\| \int_0^1 k(\tau) \, dt = \|k(\cdot)\| \, \|z(\cdot) - y(\cdot)\|.
\end{aligned}
$$

Since the argument above is symmetric in $y(\cdot)$ and $z(\cdot)$, we have

$$
|f(z(\cdot) - f(y(\cdot))|(t) \leq K(t)\|z(\cdot) - y(\cdot)\|,
$$

where $K(t) = \|k(\cdot)\|$ for all $t \in [0, 1]$. By the remarks following Definition 1, f is order Lipschitz at $\overline{x}(\cdot)$.

When f is regular at $\overline{x}(\cdot)$ we will use the concept of *fan* introduced by Ioffe [20–22] to show that the expression for the generalized gradient is

$$
(1) \qquad \partial f(\overline{x}(\cdot))(h(\cdot))(t) \subseteq \int_0^t \partial \phi(\tau, \overline{x}(\tau))(h(\tau)) \, d\tau.
$$

The above expression means that for $T \in \partial f(\overline{x}(\cdot))$ and $h(\cdot) \in C[0, 1]$, $T(h(\cdot)(t)$ can be written $\int_0^t \xi_T(\tau)h(\tau) \, d\tau$, where $\tau \to \xi_T(\tau)h(\tau)$ is a measurable selection of the set-valued mapping $\tau \to \partial \phi(\tau, \overline{x}(\tau))h(\tau)$.

Define the set-valued mapping $A: C[0, 1] \to C[0, 1]$ by

$$
A(h(\cdot)) = \{v(\cdot) \in C[0, 1] | v(\cdot) = T(h(\cdot)) \text{ for some } T \in \partial f(\overline{x}(\cdot))\}.
$$

Then by Theorem 1 and [20, page 7; 21, page 96], A is a bounded odd fan generated by $\partial f(\overline{x})$ with support function $s(v^*, h) \equiv \sup\{v^*(T(h)) | T \in \partial f(\overline{x})\}$, where $v^* \in V^*$, and adjoint fan $A^*(v^*) = \{x^* \in X^* | x^* = T^* v^*$ for some $T \in \partial f(\overline{x})\}$. For any $h(\cdot) \in C[0, 1]$ we will show that $A(h(\cdot))$ is a subset of $f'(\overline{x}(\cdot))(h(\cdot))$, where the upper derivative $f'(\overline{x}(\cdot))$ of f at $\overline{x}(\cdot)$ is the fan with support function

$$
f^\circ(\overline{x}; v^*, h) \equiv \limsup_{\substack{x \to \overline{x} \\ \lambda \downarrow 0}} \lambda^{-1} v^*(f(x + \lambda h) - f(x))
$$

[20, page 36; 21, page 108]. The relation $A(h(\cdot)) \subseteq f'(\overline{x}(\cdot))(h(\cdot))$ will verify (1) by application of [22, Example 2.5.3] where it is shown that (1) holds for the upper derivative $f'(\overline{x}(\cdot))$.

The adjoint fan $f'^*(\overline{x}(\cdot)): V^* \to X^*$ of $f'(\overline{x}(\cdot))$ is the set-valued mapping defined by

$$
f'^*(\overline{x}(\cdot))(v^*) = \{x^* \in X^* | f^\circ(\overline{x}(\cdot); v^*, x(\cdot)) \geq x^*(x(\cdot)) \, \forall x(\cdot) \in X\},
$$

that is, $f'^*(\overline{x}(\cdot))(v^*)$ is the generalized gradient of the function $f_{v^*}: X \to R$ given by $f_{v^*}(x) = v^*(f(x))$. Now for any $v^* \in V^*$, $x^* \in A^*(v^*)$ and $h(\cdot) \in X$, we have $x^*(h(\cdot)) \leq f^{\circ*}(\overline{x}(\cdot); v^*)(h(\cdot)) \leq f^\circ(\overline{x}(\cdot); v^*, h(\cdot))$; hence $x^* \in f'^*(\overline{x}(\cdot))(v^*)$, which implies $A^*(v^*) \subseteq f'^*(\overline{x}(\cdot))(v^*)$, and thus $A(h(\cdot)) \subseteq f'(\overline{x}(\cdot))(h(\cdot))$, as desired.

EXAMPLE 3. Let $X = L_1[0, 1]$ and $V = L_2[0, 1]$; for $x(\cdot) \in L_1[0, 1]$ consider the mapping

$$f(x(\cdot))(t) = \int_0^1 \sqrt{s + t}|x(s)| \, ds.$$

Since $\int_0^1 [\int_0^1 \sqrt{s + t}|x(s)| ds]^2 \, dt \leq 2 \int_0^1 [\int_0^1 |x(s)| ds]^2 \, dt < \infty$ and $f(y(\cdot))(t) - f(x(\cdot))(t) \leq \int_0^1 \sqrt{s + t}(|y(s) - x(s)|) \, ds \leq \sqrt{2}\|y(\cdot) - x(\cdot)\|$, f is an order Lipschitz mapping from $L_1[0, 1]$ to $L_2[0, 1]$. The set of functions in $L_2[0, 1]$ that are nonnegative almost everywhere is a normal cone [31, page 65]. To determine the generalized gradient of f at $\theta \in X$ note that, for $h(\cdot) \in X$, $f^\circ(\theta; h(\cdot))(t) \leq \int_0^1 \sqrt{s + t}|h(s)| \, ds$ and hence $\partial f(\theta)$ is a subset of the set of $T \in L(X, V)$ such that $T(h(\cdot))(t) \leq \int_0^1 \sqrt{s + t}|h(s)| \, ds$ for all $h(\cdot) \in X$ and $t \in [0, 1]$. To obtain an integral representation of $\partial f(\theta)$, we note that since $V = L_2[0, 1]$, every $T \in L(X, V)$ is weakly compact [9, page 483, Corollary 3]. Then by [9, page 507, Theorem 10] if $T \in L(X, V)$ there is an essentially unique bounded measurable function $\phi: [0, 1] \to L_2[0, 1]$, called the kernel, such that

$$T(h(\cdot))(t) = \int_0^1 \phi(s)(t)h(s) \, ds.$$

Hence $\partial f(\theta)$ is a subset of the set of $T \in L(X, V)$ that have a corresponding essentially unique kernel $\phi: [0, 1] \to L_2[0, 1]$ such that

$$\int_0^1 \phi(s)(t)h(s) \, ds \leq \int_0^1 \sqrt{s + t}|h(s)| \, ds$$

for all $h(\cdot) \in L_1[0, 1]$ and $t \in [0, 1]$.

3.2 EXAMPLES AND PROOFS FOR SECTION 2.2. In this section we prove Proposition 7 and Theorems 2 and 3. We conclude with sample applications of the Kuhn-Tucker conditions of Theorem 2 for particular choices of A and B in the constraint set $S = \{x \in A | g(x) \in B\}$.

PROOF OF PROPOSITION 7. The condition $0 \in \partial f(\overline{x}) + [P(S; \overline{x})]^\circ$ implies $0 = T + \gamma$, where $T \in \partial f(\overline{x})$ and $\gamma \in [P(S; \overline{x})]^\circ$. therefore, for all $x \in S$, $0 = T(x - \overline{x}) + \gamma(x - \overline{x})$. Since S is pseudoconvex at $\overline{x}, x - \overline{x} \in P(S; \overline{x})$ for all $x \in S$, whish implies $\gamma(x - \overline{x}) \leq 0$. Thus $T(x - \overline{x}) \geq 0$ and, for all $x \in S$, $f^\circ(\overline{x}; x - \overline{x}) \geq T(x - \overline{x}) \geq 0$, which by the pseudoconvexity of f implies $f(x) \geq f(\overline{x})$.

PROOF OF THEOREM 2. Since $\overline{x}$ is a local minimum of f on S, we have by Proposition 7 that $0 \in \partial f(\overline{x}) + M^\circ$ for any convex cone M contained in $K(S; \overline{x})$. Since $M^\circ = (\text{cl } M)^\circ$ (property C2, [10]) and $J^0 + G^0$ is closed, then $M^\circ = J^\circ + G^\circ$ (property C3, [10]) and $0 \in \partial f(\overline{x}) + J^\circ + G^\circ$. Let $\gamma \in {}^\circ(H^*)$; then $\mu(T(\gamma)) \leq 0$ for any $\mu \in [P(B; g(\overline{x}))]^\circ$ and $T \in \partial g(\overline{x})$. Now suppose that $T(\gamma) \notin P(B; g(\overline{x}))$; then since $P(B; g(\overline{x}))$ is a closed convex cone, by the strong separation theorem [9, page 417] there exists $v^* \in V^*$ such that $v^*(T(\gamma)) > 0 \geq v^*(w)$ for any $w \in P(B; g(\overline{x}))$, which implies that $v^* \in [P(B; g(\overline{x}))]^\circ$. Then $v^*(T(\gamma)) \leq 0$ and this contradicts $v^*(T(\gamma)) > 0$. Therefore, $T(\gamma) \in P(B; g(\overline{x}))$, that is, for each $\gamma \in {}^\circ(H^*)$ we have shown $\gamma \in J$. Hence ${}^\circ(H^*) \subset J$ and since H^* is a closed convex cone, $H^* = ({}^\circ(H^*))^\circ \supset J^\circ$, which shows that there exists $\mu \in [P(B; g(\overline{x}))]^\circ$ such that $0 \in \partial f(\overline{x}) + \mu \partial g(\overline{x}) + G^\circ$.

PROOF OF THEOREM 3. There exists $\mathcal{E} \in \partial f(\overline{x}), T \in \partial g(\overline{x})$ and $x^* \in G^\circ$ such that $0 = \mathcal{E} + \mu \circ T + x^*$, hence $0 = \mathcal{E}(x - \overline{x}) + \mu(T(x - \overline{x})) + x^*(x - \overline{x})$. Since S is strictly pseudoconvex we have $T(x - \overline{x}) \in K(B; g(\overline{x}))$ and thus $\mu(T(x - \overline{x})) \leq 0$; also, $x^*(x - \overline{x}) \leq 0$ for al $x \in S$, hence $\mathcal{E}(x - \overline{x}) \geq 0$. Since f is pseudoconvex on S, it follows that $f^\circ(\overline{x}; x - \overline{x}) \geq \mathcal{E}(x - \overline{x}) \geq 0$, which implies $f(x) \geq f(\overline{x})$ for any $x \in S$.

EXAMPLE 4. Let X and V be the Euclidean space 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 $g_i(\overline{x}) = 0$ for all $i \in I$ and $g_j(\overline{x}) < 0$ for all $j \in J$, where $\overline{x} \in S = \{x \in A | g(x) \leq 0\}$. Then $[P(B; g(\overline{x}))]^0 = [P(E^m_-; g(\overline{x}))]^\circ = \{\lambda \in E^m | \lambda \geq 0, \lambda g(\overline{x}) = 0\} = \{\lambda \in E^m | \lambda_i \geq 0, i \in I, \lambda_j = 0, j \in J\}$. If $\overline{x}$ minimizes f over S, the necessary conditions of Theorem 2 imply that there exist scalars $\lambda_i \geq 0$ such that $\lambda_i g_i(\overline{x}) = 0, i = 1, \ldots, m$, and $0 \in \partial f(\overline{x}) + \sum_{i=1}^m \lambda_i \partial g_i(\overline{x}) + G^\circ$. If $A = E^n$ and $G^\circ = [P(E^n; \overline{x})]^\circ = \{0\}$, we have $0 \in \partial f(\overline{x}) + \sum_{i=1}^m \lambda_i \partial g_i(\overline{x})$; moreover, if f and g are continuously differentiable at $\overline{x}$, the latter condition reduces to $0 = \nabla f(\overline{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\overline{x})$. Note that both $K = \{x \in E^n | \mathcal{E}_i x \leq 0 \text{ for each } \mathcal{E}_i \in \partial g_i(\overline{x}), i \in I\}$ and $H^* = \{h \in E^n | h = \sum_{i \in I} \lambda_i \mathcal{E}_i, \lambda_i \geq 0, \mathcal{E}_i \in \partial g_i(\overline{x})\}$ are closed convex cones. If $K(S, \overline{x})$ is convex (which occurs if, for example, S is convex) then to obtain the strongest result G must be a closed convex cone such that $J \cap G = P(S; \overline{x})$ and $J^\circ + G^\circ$ is closed. In the differentiable case (where the convexity of $K(S; \overline{x})$ is not required) Guignard [10] shows that this assumption on G is the weakest constraint qualification for finite-dimensional programming problems.

EXAMPLE 5. Consider the problem $\min\{f(x) | x \in A, g_i(x) = 0, i \in I, g_j(x) \leq 0, j \in J\}$ where $I = \{1, \ldots, m - r\}$ and $J = \{m - r + 1, \ldots, m\}$. Thus $B = \{0\} \times E^r_-$ where $\{0\} \subset E^{m-r}$. The optimality conditions of Theorem 2,

derived as in the above example, state that there exist scalars λ_i, $i \in I$, $\lambda_j \geq 0$, $j \in J$, such that $0 \in \partial f(\overline{x}) + \sum_{k \in I \cup J} \lambda_k \partial g_k(\overline{x}) + G^\circ$.

EXAMPLE 6. 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 with a normal positive cone, and $B = \{0\}$, then $P(B; g(\overline{x})) = \{0\}$ and Theorem 2 says that there exists $\mu \in V^*$ such that $0 \in \partial f(\overline{x}) + \mu \partial g(\overline{x}) + G^\circ$. Multiplier rules for infinite dimensional equality constraints have appeared only recently; Ioffe [19], [22] for example, provides such a rule for V a (not necessarily ordered) Banach space.

EXAMPLE 7. If X and V are as in Example 6, then corresponding to the problem

(P1)   minimize $f(x)$ subject to $x \in A$, $g(x) \in B$,

where B is a closed convex cone, is the saddle-value problem

(P2)   find $\overline{x} \in A$, $\overline{\mu} \in B^\circ$ satisfying $\phi(\overline{x}, \mu) \leq \phi(\overline{x}, \overline{\mu}) \leq \phi(x, \overline{\mu})$
            for all $x \in A$, $\mu \in B^\circ$

where $\phi(x, \mu) \equiv f(x) + \mu(g(x))$. Without going too far into detail, (see [32] for proofs) we will state relationships between solutions to P1 and P2. For fixed $\mu \in B^\circ$, we denote by $\partial_x \phi(\overline{x}, \mu)$ the generalized gradient at $\overline{x}$ of the function $x \to \phi(x, \mu)$ ($x \to \phi(x, \mu)$ is order Lipschitz at $\overline{x}$ since f is Lipschitz at $\overline{x}$, g is order Lipschitz at $\overline{x}$, and μ is an order bounded linear functional). If $(\overline{x}, \overline{\mu})$ solves P2, then (i) $0 \in \partial_x \phi(\overline{x}, \overline{\mu}) + M^\circ$, where M is any convex cone contained in $K(A; \overline{x})$, (ii) $g(\overline{x}) \in B$, and (iii) $\overline{\mu}(g(\overline{x})) = 0$. As a consequence, if $(\overline{x}, \overline{\mu})$ solves P2, then $\overline{x}$ solves P1. Conversely, if A is pseudoconvex at $\overline{x} \in A$, $\phi(x, \overline{\mu})$ is pseudoconvex on A, $0 \in \partial_x \phi(\overline{x}, \overline{\mu}) + [K(A; \overline{x})]^\circ$, and $(\overline{x}, \overline{\mu})$ satisfy conditions (ii) and (iii) above, then $(\overline{x}, \overline{\mu})$ solves P2. In addition, if $\overline{x}$ solves P1, A is pseudoconvex at $\overline{x}$ and the hypotheses of Theorem 2 are satisfied with $G = K(A; \overline{x})$, then $(\overline{x}, \overline{\mu})$ solves P2 if $\phi(x, \overline{\mu})$ is pseudoconvex on A, $\overline{\mu}$ satisfies $0 \leq \overline{\mu}(y) \leq y$ for all $y \in V_+$, f and $\overline{\mu} \circ g$ are regular at $\overline{x}$ in the sense of Clarke [7, page 39] and $\overline{\mu}(g(\overline{x})) = 0$.

4. Summary

For a vector-valued function $f: X \to V$ that is order Lipschitz at $\overline{x}$ we have defined and obtained properties for the generalized directional derivative $f^\circ(\overline{x}, y)$ and the generalized gradient $\partial f(\overline{x})$. In particular, we have discussed conditions under which the sublinear mapping $f^\circ(\overline{x}; \cdot)$ is continuous and have shown that when this is the case, $f(\overline{x})$ is nonempty, convex, closed

and equicontinuous (as a subset of $L(X, V)$ with the topology of pointwise convergence) and $f°(\overline{x}; y) = \max\{T(y) | T \in \partial f(\overline{x})\}$. If the order intervals in V are compact, then $\partial f(\overline{x})$ 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 [19], for example, which in turn could lead to more general optimality conditions.

References

[1] J. M. Borwein, 'Continuity and differentiability properties of convex operators', *Proc. London Math. Soc.* (3) **44** (1982), 420–444.

[2] N. Bourbaki, *Varietes differentielles et analytiques, elements de mathematiques* XXXIII (Hermann, Paris, 1967).

[3] F. H. Clarke, *Necesary conditions for nonsmooth problems in optimal control and the calculus of variations*, (Ph.D. Thesis, University of Washington, Seattle, 1973).

[4] F. H. Clarke, 'Generalized gradients and applications', *Trans. Amer. Math. Soc.* **205** (1975), 247–262.

[5] F. H. Clarke, 'A new approach to Lagrange multipliers', *Math. Oper. Res.* **2** (1976), 165–174.

[6] F. H. Clarke, 'Generalized gradients of Lipschitz functionals', *Adv. in Math.* **40** (1981), 52–66.

[7] F. H. Clarke, *Optimization and nonsmooth analysis*, (Wiley and Sons, New York, 1983).

[8] R. W. Chaney, 'On sufficient conditions in nonsmooth optimization', *Math. Oper. Res.* **7** (1982), 463–475.

[9] N. Dunford and J. T. Schwartz, *Linear operators, Part* I: *General theory*, (Wiley and Sons, New York, 1964).

[10] M. Guignard, 'Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space', *SIAM J. Control Optim.* **7** (1969), 232–241.

[11] H. Halkin, 'Interior mapping theorem with set-valued derivatives', *J. Analyse Math.* **30** (1976), 200–207.

[12] J. B. Hiriart-Urruty, 'On optimality conditions in nondifferentiable programming', *Math. Programming* **14** (1978), 73–86.

[13] J. B. Hiriart-Urruty, 'Refinements of necessary optimality conditions in nondifferentiable programming I', *Appl. Math. Optim.* **5** (1979), 63–82.

[14] J. B. Hiriart-Urruty, 'Tangent cones, generalized gradients and mathematical programming in Banach spaces,' *Math. Oper. Res.* **4** (1979), 79–97.

[15] J. B. Hiriart-Urruty, 'New concepts in nondifferentiable programming,' *Bull. Soc. Math. France* **50** (1979), 57–85.

[16] J. B. Hiriart-Urruty, 'Approximating a second-order directional derivative for nonsmooth convex functions,' *SIAM J. Control Optim.* **20** (1982), 783–807.

[17] J. B. Hiriart-Urruty, 'Limiting behaviour of the approximate first-order and second-order directional derivatives for a convex function', *Nonlinar Anal.* **6** (1982), 1309–1326.

[18] J. B. Hiriart-Urruty, J. J. Strodiot and V. H. Nguyen, 'Generalized hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data', *Appl. Math. Optim.* **11** (1984), 43–56.

[19] A. D. Ioffe, 'Necessary and sufficient conditions for a local minimum. 1: A reduction theorem and first-order conditions,' *SIAM J. Control Optim.* **17** (1979), 245–250.

[20] A. D. Ioffe, 'Nonsmooth analysis: differential calculus of nondifferentiable mappings,' *Trans. Amer. Math. Soc.* **266** (1981), 1–56.

[21] A. D. Ioffe, 'Nonsmooth analysis and the theory of fans,' *Convex analysis and optimization* edited by J. P. Aubin and R. B. Vinter, (Pitman, Marshfield, Mass., 1982), pp. 93–117.

[22] A. D. Ioffe, 'Necessary conditions in nonsmooth optimization,' *Math. Oper. Res.* **9** (1984), 159–189.

[23] A. D. Ioffe and L. Levin, 'Subdifferentials of convex functions', *Trans. Moscow Math. Soc.* **26** (1972), 1–72.

[24] A. G. Kusraev, 'On necessary conditions for an extremum of nonsmooth vector-valued mappings,' *Soviet Math. Dokl.* **19** (1978), 1057–1060.

[25] S. S. Kutateladze, 'Support sets of sublinear operators', *Dokl. Akad. Nauk* SSSR **230** (1976), 1029–1032.

[26] L. McLinden, 'Dual operations on saddle functions,' *Trans. Amer. Math. Soc.* **179** (1973), 363–381.

[27] J. J. Moreau, *Fonctionelles convexes*, (Lecture notes, Seminaire Equations aux Derivees Partielles, College de France, 1966).

[28] N. S. Papageorgiou, 'Nonsmooth analysis on partially ordered vector spaces: part 1: Convex case,' *Pacific J. Math.* **107** (1983), 403–458.

[29] N. S. Papageorgiou, Nonsmooth analysis on partially ordered vector spaces: Part 2: Nonconvex case, Clarke's theory,' *Pacific J. Math.* **109** (1983), 463–495.

[30] J. P. Penot, 'Calcul sous-differentiels et optimisation,' *J. Funct. Anal.* **27** (1978), 248–276.

[31] A. L. Peressini, *Ordered topological vector spaces* (Harper and Row, New York, 1967).

[32] T. W. Reiland and J. H. Chou, *Saddlepoint problems and duality in nondifferential programming*, (Report #205, North Carolina State University Graduate Program in Operations Research, 1984).

[33] S. M. Robinson, 'Regularity and stability theorems for convex multivalued functions,' *Math. Oper. Res.* **1** (1976), 130–142.

[34] R. T. Rockafellar, *Convex analysis* (Princeton Univ. Press, Princeton, N.J., 1970).

[35] R. T. Rockafeller, Conjugate convex functions in optimal control and the calculus of variations', *J. Math. Anal. Appl* **32** (1970), 174–222.

[36] R. T. Rockafeller, 'Existence and duality theorems for convex problems of Bolza', *Trans. Amer. Math. Soc.* **159** (1971), 1–40.

[37] R. T. Rockafeller, 'Integrals which are convex functionals, II', *Pacific J. Math.* **39** (1971), 439–469.

[38] R. T. Rockafeller, '*Conjugate duality and optimization* (Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974).

[39] R. T. Rockafeller, 'Clarke's tangent cones and the boundaries of closed sets in R^n, *Nonlinear Anal.* **3** (1979), 145–154.

[40] R. T. Rockafeller, 'Directionally Lipschitzian functions and subdifferential calculus,' *Proc. London Math. Soc.* **39** (1979), 331–355.

[41] R. T. Rockafeller, 'Generalized directional derivatives and subgradients of nonconvex functions,' *Canad. J. Math.* **22** (1980), 257–280.

[42] R. T. Rockafeller, *The theory of subgradients and its applications to problems of optimization. Convex and nonconvex functions* (Heldermann-Verlag, Berlin, 1981).

[43] R. T. Rockafeller, 'Marginal values and second-order necessary conditions for optimality,' *Math. Programming* **26** (1983), 245–286.

[44] A. M. Rubinov, 'Sublinear operators and their applications,' *Russian Math. Surveys* **32** (1977), 115–175.

[45] H. H. Schaefer, *Topological vector spaces* (Springer-Verlag, New York, 1971).

[46] L. Thibault, 'Fonctions compactement Lipschitziennes et programmation mathematique,' *C. R. Acad. Sci. Paris Ser. A* **287** (1978), 213–216.

[47] L. Thibault, Subdifferentials of compactly Lipschitzian vector-valued functions,' *Ann. Mat. Pura Appl.* **125** (1980), 157–192.

[48] L. Thibault, 'Subdifferentials of nonconvex vector-valued functions,' *J. Math. Anal. Appl.* **86** (1982), 319–344.

[49] L. Thibault, 'On generalized differentials and subdifferentials of Lipschitz vector-valued functions,' *Nonlinear Anal.* **6** (1982), 1037–1053.

[50] L. Thibault, 'Tangent cones and quasi-interiorly tangent cones to multifunctions,' *Trans. Amer. math. Soc.* **277** (1983), 601–62.

[51] M. Valadier, 'Sous-differentiabilite de fonctions convexes a valeurs dans un espace vectoriel ordonne,' *Math. Scand.* **30** (1972), 65–74.

[52] P. P. Varaiya, 'Nonlinear programming in Banach space,' *SIAM J. Appl. Math.* **15** (1967), 284–293.

[53] J. Warga, *Derivative containers, inverse functions and controllability, calculus of variations and control theory*, (D. L. Russell, ed., Academic Press, New York, 1976).

[54] J. Zowe, 'Subdifferentiability of convex functions with values in an ordered vector space,' *Math. Scand.* **34** (1974), 69–83.

[55] J. Zowe, 'A duality theorem for a convex programming problem in order complete vector lattices,' *J. Math. Anal. Appl.* **50** (1975), 273–287.

Department of Statistics and Graduate
Program in Operations Research
North Carolina State University
Box 8203
Raleigh, North Carolina 27695-8203
U.S.A.