PACIFIC JOURNAL OF MATHEMATICS
Vol. 109, No. 2, 1983

NONSMOOTH ANALYSIS ON PARTIALLY ORDERED VECTOR SPACES: PART 2–NONCONVEX CASE, CLARKE'S THEORY

NIKOLAOS S. PAPAGEORGIOU

The purpose of this paper is to extend the recently developed Clarke theory of generalized gradients to vector valued mappings. For that we introduce the notion of locally o-Lipschitz mappings and develop a subdifferential calculus for them. In this process, we have the opportunity for comparison with analogous results obtained in the convex case.

1. Introduction. We know (see [**14**]) that a convex mapping majorized in a neighborhood of a point is locally o-Lipschitz in the interior of its domain. So it is natural to go a step further and ask whether we can have an analogous subdifferential theory for locally o-Lipschitz mappings.

In the real valued case, this problem was first introduced and successfully solved by Clarke [**2**]—[**7**]. After Clarke, others have also contributed in this or parallel directions, e.g., Aubin [**1**], Halkin [**9**], Hiriart-Urruty [**11**]—[**14**], Rockafellar [**22**], and Warga [**27**].

In this paper, we construct a similar theory for vector valued mappings. Having as our starting point the locally o-Lipschitz mappings, we define the generalized gradient for such mappings, and using that we develop a complete subdifferential theory. Although we face serious analytical difficulties working with vector valued mappings (lack of functional separability results), nevertheless introducing the notion of generalized o-directional derivatives and using the results obtained in [**19**], we are able to obtain several new results that will be potentially useful in solving nonsmooth, nonconvex vector valued extremal problems. Similar work was done very recently by Ioffe [**16**], [**17**] and Thibault [**15**]. In the last section of this paper, we will compare our results with those obtained in the above-mentioned works.

All through this paper, X will be a Banach space and Y an order complete Banach lattice. Any additional assumptions will be mentioned explicitly. The definitions and notational conventions are the same as those introduced in §2 of [**19**].

In the next section we introduce the locally o-Lipschitz mappings, which play an important role in this theory, and we examine several of their properties.

In §3 we introduce the notion of the generalized o-directional derivative and, using that, define the generalized gradient (or generalized subdifferential) of locally o-Lipschitz mappings and study its properties in detail. We use the results obtained in [**19**].

In §4 we develop the subdifferential calculus related to generalized gradients. As we shall see, several of the results of the convex theory (see [**19**]) have their analogs in the new theory.

In §5 we give a brief "preview" of the possible applications of this theory in vector valued optimization problems. In fact, this is the topic of a forthcoming paper by the author [**20**].

Finally, we conclude with a discussion of related work.

We should mention that we could have been more general in our presentation by assuming that X is a locally convex t.v.s. and Y a locally convex lattice and define $f\colon X \to Y$ to be locally o-Lipschitz if for every bounded open set U there is a $y \in K_Y^+$ and a seminorm p such that $|f(x) - f(z)| \leq yp(x - z)$ for all $x, z \in U$. Then most of the results presented here hold. But we have decided not to go into such generality in order to avoid unnecessary technical complications.

In the sequence by ∂cf, we will denote the subdifferential in the sense of convex analysis (here we deviate slightly from the notation used in [**19**].)

Needless to say that beside these ideas is always the ingenuity of F. Clarke, whose theory of generalized gradients opened new perspectives in nonsmooth analysis and optimization theory.

2. Locally o-Lipschitz mappings.

The class of locally o-Lipschitz mappings was introduced in [**19**]. Let us recall the definition.

DEFINITION 2.1. A mapping $f\colon X \to Y$ is said to be *locally o-Lipschitz* if and only if for every U open and bounded subset of X there is a $y \in K_Y^+$ such that $|f(x) - f(z)| \leq y\|x - z\|$ for all $x, z \in U$. We denote this class of mappings by $L_{\mathrm{ip}}^{\mathrm{oc}}(X, Y)$.

First observe that since Y is a Banach lattice, $|f(x) - f(z)| \leq y\|x - z\|$ implies $\|f(x) - f(z)\| \leq \|y\|\,\|x - z\|$ for all $x, z \in U$, which implies that a locally o-Lipschitz mapping is, in fact, locally norm Lipschitz and therefore strongly continuous.

PROPOSITION 2.1. *If $f, g \in L_{\mathrm{ip}}^{\mathrm{oc}}(X, Y)$ then $f + g$ and λf ($\lambda \in R$) also belong to $L_{\mathrm{ip}}^{\mathrm{oc}}(X, Y)$.*

Proof. Let U be a bounded open subset of X and $x, z \in U$. Then

$$|f(x)+g(x)-f(z)-g(z)| \leq |f(x)-f(z)| + |g(x)-g(z)|$$
$$\leq y_f\|x-z\| + y_g\|x-z\| = (y_f+y_g)\|x-z\|.$$

So $f+g \in L_{\mathrm{ip}}^{\mathrm{oc}}(X,Y)$. Also

$$|\lambda f(x) - \lambda f(z)| = |\lambda|\,|f(x)-f(z)| \leq |\lambda| y\|x-z\|,$$

which implies $\lambda f \in L_{\mathrm{ip}}^{\mathrm{oc}}(X,Y)$. □

PROPOSITION 2.2. *If $f \in L_{\mathrm{ip}}^{\mathrm{oc}}(X,Y)$ and $g\colon X \to X$ is locally norm Lipschitz then $f \circ g \in L_{\mathrm{ip}}^{\mathrm{oc}}(X,Y)$.*

Proof. Let U be an open bounded subset of X. For $x, z \in U$, we have

$$(1) \quad |(f\circ g)(x) - (f \circ g)(z)| = |f(g(x)) - f(g(z))| \leq y\|g(x)-g(z)\|.$$

But since g is norm Lipschitz,

$$(2) \qquad \|g(x)-g(z)\| \leq k\|x-z\|, \qquad k \in R^{+}.$$

Substituting (2) in (1) we get

$$|(f\circ g)(x) - (f\circ g)(z)| \leq ky\|x-z\|.$$

Hence $f \circ g$ is locally o-Lipschitz. □

COROLLARY. *If $A \in \mathfrak{L}(X,Z)$ and $f \in L_{\mathrm{ip}}^{\mathrm{oc}}(Z,Y)$ then $f \circ A \in L_{\mathrm{ip}}^{\mathrm{oc}}(X,Y)$.*

DEFINITION 2.2. A sequence of mappings $f_n\colon X \to Y$, $n \in N$, is said to *converge o-pointwise to* f if and only if for all $x \in X$, $f_n(x) \overset{o}{\to} f(x)$.

PROPOSITION 2.3. *If $\{f_n\}_{n\in N} \subseteq L_{\mathrm{ip}}^{\mathrm{oc}}(X,Y)$ with the same Lipschitz constant and $f_n \to f$ o-pointwise, then $f \in L_{\mathrm{ip}}^{\mathrm{oc}}(X,Y)$ with the same Lipschitz constant.*

Proof. For all $n \in N$ and $x, z \in U$, we have

$$|f_n(x) - f_n(z)| \leq \|x-z\|.$$

We know

$$\begin{aligned}|f(x)-f(z)| &= |f(x)-f_n(x)+f_n(x)-f_n(z)+f_n(z)-f(z)| \\ &\leq |f(x)-f_n(x)| + |f_n(x)-f_n(z)| + |f_n(z)-f(z)| \\ &\leq |f(x)-f_n(x)| + y\|x-z\| + |f_n(z)-f(z)|\end{aligned}$$

since for all $n \in N$, $f_n \in L_{\text{ip}}^{\text{oc}}(X, Y)$. Also, since $f_n \to f$ o-pointwise, we know that

$$|f(x) - f_n(x)| \xrightarrow{o} 0 \quad \text{and} \quad |f(z) - f_n(z)| \xrightarrow{o} 0.$$

So

$$\begin{aligned}|f(x) - f(z)| &\leq o - \lim_{n\to\infty} |f(x) - f_n(x)| + y\|x - z\| \\ &\quad + o - \lim_{n\to\infty} |f_n(z) - f(z)| \\ &\to |f(x) - f(z)| \leq y\|x - z\| \quad \text{for all } x, z \in U.\end{aligned}$$

Therefore, $f \in L_{\text{ip}}^{\text{oc}}(X, Y)$ with the same Lipschitz constant. □

COROLLARY. *If $\{f_n\}_{n\in N} \subseteq L_{\text{ip}}^{\text{oc}}(X, Y)$ with the same Lipschitz constant and $f_n \to f$ then $f \in L_{\text{ip}}^{\text{oc}}(X, Y)$ with the same Lipschitz constant.*

By $(f \vee g)(\cdot)$ we denote the mapping ϕ: $X \to Y$ defined by $\phi(x) = (f \vee g)(x) = f(x) \vee g(x)$. Similarly for $(f \vee g)(\cdot)$.

Then we have the following result.

PROPOSITION 2.4. *If $f, g \in L_{\text{ip}}^{\text{oc}}(X, Y)$ then $f \vee g$ and $f \wedge g$ also belong to $L_{\text{ip}}^{\text{oc}}(X, Y)$.*

Proof. Again let $x, z \in U$. Then we know from **[24]** that

$$\begin{aligned}|(f \vee g)(x) - f(\vee g)(z)| &\leq |f(x) - f(z)| + |g(x) - g(z)| \\ &\leq (y_f + y_g)\|x - z\|.\end{aligned}$$

So $f \vee g \in L_{\text{ip}}^{\text{oc}}(X, Y)$. The fact that $f \wedge g \in L_{\text{ip}}^{\text{oc}}(X, Y)$ follows from

$$|(f \wedge g)(x) - (f \wedge g)(z)| \leq |f(x) - f(z)| + |g(x) - g(z)|$$

(see **[24]**). □

DEFINITION 2.3. We will call a linear operator A: $X \to Y$ (so)-bounded if it maps norm bounded subsets of X into order bounded subsets of Y.

Note. Observe that since Y is a Banach lattice, then it is normal and so every o-bounded subset of Y is norm bounded. Hence, every (so)-bounded operator A is bounded, and so by linearity $A \in \mathfrak{L}(X, Y)$, i.e. it is continuous.

PROPOSITION 2.5. *If $A \in L(X, Y)$ is (so)-bounded then $A \in \mathrm{Lip}(X, Y)$. (Clearly the Lipschitzian property is global.)*

Proof. By linearity, we can take $z \in X$ to be the origin. Then define $\langle A \rangle \in Y$ by $\langle A \rangle = \sup_{\|x\| \leq 1} A(x)$. This exists from the fact that A is (so)-bounded and from the order completeness of the Branch lattice Y. Then we have that for all $x \in X$,

$$|Ax| \leq \langle A \rangle \|x\| \to A \in \mathrm{Lip}(X, Y),$$

i.e. A is global Lipschitz with Lipschitz constant $\langle A \rangle \in K_Y^+$. □

REMARK. We saw that since Y is a Banach lattice and therefore normal, then $A \in L(X, Y)$ and A (so)-bounded imply that $A \in \mathfrak{L}(X, Y)$. The converse however is not necessarily true because in general a bounded set is not order bounded.

However, if A is (so)-bounded, then A is (so)-continuous. So if $x_n \to x$, then $A(x_n) \xrightarrow{o} A(x)$.

Now if Y has the diagonal property and $(\mathring{K}_Y^+) \neq \varnothing$, then $A(x_n) \xrightarrow{s} A(x)$, which means $A \in \mathfrak{L}(X, Y)$. So in that case an (so)-bounded operator is continuous. Finally, if Y has the boundedness property (see [**21**]), then Proposition 2.5, we can assume simply that A is (so)-continuous.

In the next section, using the class $L_{\mathrm{ip}}^{\mathrm{oc}}(X, Y)$, we will start building the subdifferential theory for the nonconvex case.

3. Generalized gradients for locally *o*-Lipschitz mappings. Let $f = L_{\mathrm{ip}}^{\mathrm{oc}}(X, Y)$. For such a mapping we introduce the following notion (everything takes place locally).

DEFINITION 3.1. For any $w_n = (z_n, \lambda_n) \to (x, 0)$ in $X \times R^+$, we define

$$f^0(w_n)(x; d) = \overline{\lim_{n \to \infty}} \frac{f(z_n + \lambda_n d) - f(z_n)}{\lambda_n}.$$

Then we call the mapping $f^0(x; d)$ defined by

$$f^0(x; d) = \sup\{f^0(w_n)(x; d) \colon w_n \to (x, 0) \text{ in } X \times R\}$$

the *generalized o-directional derivative of f at x in the direction d*. For simplicity, we usually write

$$f^0(x; d) = \overline{\lim_{\substack{h \to 0 \\ \lambda \downarrow 0}}} \frac{f(x+h+\lambda d) - f(x+h)}{\lambda}.$$

Let us make some remarks about this definition.

First, it is easy to see that if f is real valued, then $f^0(x; \cdot)$ coincides with Clarke's generalized directional derivative at the point x (see [**2**]).

Also observe that for any $w_n = (z_n, \lambda_n) \to (x, 0)$ in $X \times R^+$, we have, by the Lipschitz property of f,

$$\frac{f(z_n + \lambda_n d) - f(z_n)}{\lambda_n} \leq \frac{\lambda_n y\|d\|}{\lambda_n} = y\|d\|.$$

So $|f(z_n + \lambda_n d) - f(z_n)|/\lambda_n$ is o-bounded (for all $n \in N$) and so $(f(z_n + \lambda_n d) - f(z_n))/\lambda_n$ is o-bounded too. Hence

$$\overline{\lim_{n \to \infty}} \frac{f(z_n + \lambda_n d) - f(z_n)}{\lambda_n}$$

exists (i.e. it is finite).

Furthermore, since this is true for all $w_n = (z_n, \lambda_n) \to (x, 0)$ in $X \times R^+$, we conclude that $f^0(x; d)$ exists and, in fact, we have the estimate $|f^0(x; d)| \leq y\|d\|$.

Finally, another useful way to express $f^0(x; d)$ is

$$f^0(x; d) = \inf_{\varepsilon, \delta > 0} \sup_{\substack{\|x - z\| < \varepsilon \\ 0 < \lambda < \delta}} \frac{f(z + \lambda d) - f(z)}{\lambda}.$$

In the sequel, we will investigate $f^0(x; d)$ as a mapping of $d \in X$.

LEMMA 3.1. *The mapping $f^0(x; \cdot)\colon X \to Y$ is sublinear and we have*

$$|f^0(x; d)| \leq y\|d\| \quad \textit{for all } x \in U.$$

Proof. The positive homogeneity of $f^0(x; \cdot)$ results directly from Definition 3.1. So in order to show sublinearity, we have to show subadditivity, i.e.

$$f^0(x; d_1 + d_2) \leq f^0(x; d_1) + f^0(x; d_2)$$

for any $d_1, d_2 \in X$.

From the definition we know that

$$\begin{aligned}
f^0(x; d_1 + d_2) &= \overline{\lim_{\substack{h \to 0 \\ \lambda \downarrow 0}}} \frac{f(x + h + \lambda(d_1 + d_2)) - f(x + h)}{\lambda} \\
&= \overline{\lim_{\substack{h \to 0 \\ \lambda \downarrow 0}}} \left\{ \frac{f(x + h + \lambda d_1 + \lambda d_2) - f(x + h + \lambda d_1)}{\lambda} \right. \\
&\qquad \left. + \frac{f(x + h + \lambda d_1) - f(x + h)}{\lambda} \right\} \\
&\leq \overline{\lim_{\substack{h \to 0 \\ \lambda \downarrow 0}}} \frac{f(x + h + \lambda d_1 + \lambda_2) - f(x + h + \lambda d_1)}{\lambda} \\
&\quad + \overline{\lim_{\substack{h \to 0 \\ \lambda \downarrow 0}}} \frac{f(x + h + \lambda d_1) - f(x + h)}{\lambda} \\
&= f^0(x; d_1) + f^0(x; d_2).
\end{aligned}$$

So, indeed, $f^0(x; \cdot)$ is subadditive and, hence, sublinear for every $x \in U$. As we have already seen, the estimate $|f^0(x; d)| \leq y\|d\|$ is a consequence of the Lipschitz property of f. $\square$

Using this lemma, we obtain the following important result.

PROPOSITION 3.1. *$f^0(x; \cdot) \in \mathrm{Lip}(X, Y)$ for all $x \in U$.*

Proof. First using Lemma 3.1, we have for $d_1, d_2 \in X$:

$$\begin{aligned}
(1) \quad f^0(x; d_1) &= f^0(x; d_1 + d_2 - d_2) \leq f^0(x; d_2) + f^0(x; d_1 - d_2) \\
&\to f^0(x; d_1) - f^0(x; d_2) \leq f^0(x; d_1 - d_2).
\end{aligned}$$

Similarly, interchanging the roles of d_1 and d_2 we get

$$(2) \qquad f^0(x; d_2) - f^0(x; d_1) \leq f^0(x; d_2 - d_1).$$

Then using (1) and (2) above, we have

$$\begin{aligned}|f^0(x; d_1) - f^0(x; d_2)| &= \big(f^0(x; d_1) - f^0(x; d_2)\big) \\ &\qquad \vee \big(-\big(f^0(x; d_1) - f^0(x; d_2)\big)\big) \\ &= \big(f^0(x; d_1) - f^0(x; d_2)\big) \vee \big(f^0(x; d_2) - f^0(x; d_1)\big) \\ &\leq f^0(x; d_1 - d_2) \vee f^0(x; d_2 - d_1) \\ &\leq \big(y\|d_1 - d_2\|\big) \vee \big(y\|d_2 - d_1\|\big) \\ &= y\|d_1 - d_2\|.\end{aligned}$$

Therefore $f^0(x; \cdot) \in \operatorname{Lip}(X, Y)$ for all $x \in U$. □

Note. Since from the above proposition, we have $f^0(x; \cdot)$ is o-Lipschitz, we deduce that it is norm Lipschitz and therefore strongly continuous.

Another way to deduce the continuity of $f^0(x; \cdot)$ is the following. From Lemma 3.1, we know that

$$f^0(x; d) \leq y\|d\| \leq yr \quad \text{for } d \in \mathring{B}_r(0).$$

So $f^0(x; \cdot)$ is majorized in a neighborhood of the origin, and since it is sublinear, by Theorem 3.1 of [**19**] (see also [**26**]), we deduce that $f^0(x; \cdot)$ is continuous on $\mathring{B}_r(0)$. Now we claim that by homogeneity, it is continuous on all of X. Indeed, if d is any point in X, then there is a $\lambda_0 > 0$ such that for $0 \leq \lambda \leq \lambda_0$, $\lambda d \in \mathring{B}_r(0)$. Now at λd, $f^0(x; \cdot)$ is continuous. Again by Lemma 3.1, $f^0(x; \lambda d) = \lambda f^0(x; d)$. So $f^0(x; \cdot)$ is continuous on all of X.

Having the notion of the generalized o-directional derivative, we can proceed and define the notion of the generalized gradient (or generalized subdifferential).

DEFINITION 3.2. We define the generalized gradient of f at x to be the set

$$\partial f(x) = \{A \in \mathfrak{L}(X, Y) \colon A(d) \leq f^0(x; d)\ \forall d \in X\}.$$

Note again, that if f is real valued, then the above set valued mapping is the celebrated Clarke subdifferential (see [**2**], [**3**], [**7**]).

If we ignore the topological structures on X and Y, and consider only their algebraic ones, we get the algebraic generalized gradient of f at x, which we denote by $\partial^a f(x)$. So we have

$$\partial^a f(x) = \{A \in L(X, Y) \colon A(d) \leq f^0(x; d)\ \forall d \in X\}.$$

Observe that from the homogeneity of $f^0(x;\cdot)$, we have $f^0(x;0)=0$. Then, for $x \in U$,

$$(f^0)'(x;0,d) = o-\lim_{\lambda\downarrow 0}\frac{f^0(x;\lambda d)-f^0(x;0)}{\lambda}$$
$$= o-\lim_{\lambda\downarrow 0}\frac{\lambda f^0(x;d)}{\lambda} = f^0(x;d)\quad \forall d\in X.$$

Then by Valadier's **[26]** Proposition 4, we have

$$\partial^\alpha f^0(x;0) = \{A\in L(X,Y)\colon A(d)\le f^0(x;d) = (f^0)'(x;0,d)\}$$
$$\to \partial_c^\alpha f^0(x;0) = \partial^\alpha f(x).$$

But, as we saw previously, $f^0(x;\cdot)$ is strongly continuous, so again from Valadier (Theorem 6 in **[26]**), we have

$$(1)\qquad \partial_c^\alpha f^0(x;0) = \partial_c f^0(x;0) = \partial^\alpha f(x)$$

and

$$(2)\qquad \partial_c f^0(x;0) = \partial f(x).$$

So from (1) and (2), we conclude that $\partial f(x) = \partial^\alpha f(x)$.

So in the sequence, we will no longer distinguish between the generalized gradient and the algebraic generalized gradient (as we did in the convex case (see **[19]**)), since we see that these two coincide. In fact, we get the following chain of identifications:

$$(*)\qquad \partial^\alpha f(x) = \partial f(x) = \partial_c f^0(x;0) = \partial_c^\alpha f^0(x;0).$$

From $(*)$ and Valadier's Theorem 6 and Corollary 7 in **[26]**, we get the following result.

THEOREM 3.1. (1) $\partial_c f^0(x;0) = \partial f(x) = \partial^\alpha f(x) \neq \varnothing$ $\forall x\in U$.

(2) $f^0(x;d) = \max\{A(d)\colon A\in f(x)\}$.

(3) $\partial f(x) \subseteq \mathcal{L}(X,Y)$ *is a convex, equicontinuous and compact subset of* $\mathcal{L}_s(X,Y)$, *if the o-intervals* $[x,y]$ *are w-compact.*

Note. $\mathcal{L}_s(X,Y)$ denotes the space $\mathcal{L}(X,Y)$ with the topology of simple convergence (strong operator topology).

In the sequel we will examine the properties of the multioperator $\partial f(\cdot)$. For similar results in the real valued case, we refer to the two excellent survey articles by Clarke **[9]** (it contains also some new results), and Aubin [1]. Both will be our main references for the remainder of this section.

The generalized gradient $\partial f(\cdot)$ defines a multioperator from X into $2^{\mathfrak{L}_s(X, Y)}$.

We recall that a multioperator is closed if and only if its graph is a closed subset of $X \times \mathfrak{L}_s(X, Y)$.

Viewing $\partial f(\cdot)$ as a multioperator, the following result is true.

PROPOSITION 3.2. *If Y is weakly sequentially complete then the multioperator $\partial f(\cdot)$ is closed.*

Proof. The graph of f is the set

$$\operatorname{Gr} \partial f = \{(x, A) \in X \times \mathfrak{L}_s(X, Y) \colon A(d) \leq f^0(x; d)\ \forall d \in X\}.$$

But

$$\begin{aligned} &A(d) \leq f^0(x; d) \quad \forall d \in X \\ &\quad \leftrightarrow 0 \leq f^0(x; d) - A(d) \leftrightarrow f^0(x; d) - A(d) \in K_Y^+ \\ &\quad \leftrightarrow (f^0(x; d) - A(d), p) \geq 0 \quad \forall p \in (K_Y^+)^*, \forall d \in X. \end{aligned}$$

So we have

$$\begin{aligned} \operatorname{Gr} \partial f = \{(x, a) \in X \times \mathfrak{L}_s(X, Y) \colon (f^0(x; d) - A(d), p) \geq 0, \\ p \in (K_Y^+)^*, d \in X\}. \end{aligned}$$

Hence, $\operatorname{Gr} \partial f$ is the upper level set at zero of the real valued function $(x, A) \to (f^0(x; d) - A(d), p)$ for every $d \in X$ and $p \in (K_Y^+)^*$. Since $A \to (A(d), p)$ is $\mathfrak{L}_s(X, Y)$ continuous, it suffices to show that $x \to (f^0(x; d), p)$ is upper semicontinuous $\forall p \in (K_Y^+)^*$, $d \in X$. But for every $\varepsilon > 0$, there is a $\hat{\delta} > 0$ such that

$$\left(\sup_{\substack{\|z - x\| < \hat{\delta} \\ \lambda < \hat{\delta}/2}} \frac{f(z + \lambda d) - f(z)}{\lambda}, p \right) \leq (f^0(x; d), p) + \frac{\varepsilon}{2}.$$

(This is from the definition of $f^0(x; \cdot)$ and Lemma 8 of [26].) Also from the fact that $f \in L_{\text{ip}}^{\text{oc}}(X, Y)$, we have

$$\frac{f(z + \lambda d_1) - f(z)}{\lambda} \leq \frac{f(z + \lambda d_2) - f(z)}{\lambda} + y\|d_1 - d_2\|.$$

Now let $v \in U$ be such that $\|x - v\| < \hat{\delta}/2$ and $\mu < \hat{\delta}$, $\nu < \hat{\delta}/2$. Then

$$\sup_{\substack{\|z-v\|<\mu \\ \lambda<\nu}} \frac{f(z+\lambda d_1) - f(z)}{\lambda}$$

$$\leq \sup_{\substack{\|z-x\|<\hat{\delta} \\ \lambda<\hat{\delta}/2}} \frac{f(z+\lambda d_2) - f(z)}{\lambda} + y\|d_1 - d_2\|$$

$$\to \left(\sup_{\substack{\|z-v\|<\mu \\ \lambda<\nu}} \frac{f(z+\lambda d_1) - f(z)}{\lambda}, p \right)$$

$$\leq \left(\sup_{\substack{\|z-x\|<\hat{\delta} \\ \lambda<\hat{\delta}/2}} \frac{f(z+\lambda d_2) - f(z)}{\lambda}, p \right) + \big(y\|d_1 - d_2\|, p \big)$$

$$\to \left(\sup_{\substack{\|z-v\|<\mu \\ \lambda<\nu}} \frac{f(z+\lambda d_1) - f(z)}{\lambda}, p \right) \leq \big(f^0(x; d), p \big) + \varepsilon$$

by letting $\|d_1 - d_2\| \leq \varepsilon/(2\|y\| \cdot \|p\|)$. So

$$\big(f^0(v; d), p \big) \leq \big(f^0(x; d), p \big) + \varepsilon \quad \text{for } \|x - v\| < \hat{\delta}/2.$$

Hence $(f^0(\cdot\,; d), p)$ is u.s.c. $\forall d \in X, p \in (K_Y^+)^*$.

This proves that $\partial f(\cdot)$ is a closed multioperator. □

Note. If Y is reflexive then Y is weakly sequentially complete.

Next we have

$$-f^0(x; -d) = -\overline{\lim_{\substack{h\to 0 \\ \lambda \downarrow 0}}} \frac{f(x+h+\lambda(-d)) - f(x+h)}{\lambda}$$

$$= \underline{\lim_{\substack{h\to 0 \\ \lambda \downarrow 0}}} - \frac{f(x+h+\lambda(-d)) - f(x+h)}{\lambda}$$

$$= \underline{\lim_{\substack{h\to 0 \\ \lambda \downarrow 0}}} \frac{f(x+h+(-\lambda)d) - f(x+h)}{-\lambda}$$

$$\leq f^0(x; d).$$

Also,

$$\begin{aligned}(-f)^0(x;d) &= \overline{\lim_{\substack{h\to 0\\ \lambda\downarrow 0}}}\,\frac{-f(x+h+\lambda d)+f(x+h)}{\lambda}\\ &= \overline{\lim_{\substack{h\to 0\\ \lambda\downarrow 0}}}\,-\frac{f(x+h+\lambda d)-f(x+h)}{\lambda}\\ &= -\underline{\lim_{\substack{h\to 0\\ \lambda\downarrow 0}}}\,\frac{f(x+h+\lambda d)-f(x+h)}{\lambda}\\ &= f^0(x;-d).\end{aligned}$$

So we conclude

$$(-f)^0(x;d) = f^0(x;-d) \quad \text{for all } d \in X.$$

Hence, if

$$\begin{aligned}A \in \partial(-f)(x) &\to A(d) \le (-f)^0(x;d) = f^0(x;-d) \quad \forall d \in X\\ &\to -A(d) \le f^0(x;d) \quad \forall d\in X\\ &\to -A \in \partial f(x).\end{aligned}$$

Also if $A \in -\partial f(x)$ then

$$\begin{aligned}&A(-d) \le f^0(x;d) \quad \forall d \in X\\ &\quad\to A(d) \le f^0(x;-d) = (-f)^0(x;d) \quad \text{for all } d\in X.\end{aligned}$$

So $A \in \partial(-f)(x)$. So from the above we deduce the following result.

PROPOSITION 3.3. $\partial(-f)(x) = -\partial f(x)$ *for all* $x \in X$.

The next natural step is to compare $\partial f(x)$ with $\partial_c(x)$ in the case where f is a convex mapping (we know from [**19**], Theorem 3.2, that such an f is locally o-Lipschitz on $\operatorname{int}\operatorname{dom} f$).

THEOREM 3.2. *If* $f \in \operatorname{Conv}(U, Y)$ *where* $U \subseteq X$ *is open and bounded, then* $\partial f(x) = \partial_c f(x)$.

Note. $\operatorname{Conv}(U, X)$ denotes the family of convex mappings from U to Y.

Proof. It is clear from the definitions that

$$f'(x;d) \le f^0(x;d) \quad \forall d \in X$$

(for the definition of $f'(x; d)$, see **[19]** or **[26]**). If we show that in fact we have $f'(x; d) = f^0(x; d)$ for all $d \in X$, then we can see from the definitions of $\partial_c f(x)$ and $\partial f(x)$ that these two multioperators are equal.

Fixing x and $\lambda_0 > 0$, we have

$$\begin{aligned}
&\frac{f(z+\lambda d)}{\lambda} - \frac{f(x+\lambda_0 d)}{\lambda_0} \\
&= \frac{f(z+\lambda d)}{\lambda} - \frac{f(x+\lambda_0 d)}{\lambda} + \frac{f(x+\lambda_0 d)}{\lambda} - \frac{f(x+\lambda_0 d)}{\lambda_0} \qquad (1)\\
&\le \frac{y\|z - x + d(\lambda - \lambda_0)\|}{\lambda} + f(x+\lambda_0 d)\left[\frac{1}{\lambda} - \frac{1}{\lambda_0}\right].
\end{aligned}$$

Also

$$\begin{aligned}
(2)\quad &\frac{f(z)}{\lambda} - \frac{f(x)}{\lambda} + \frac{f(x)}{\lambda} - \frac{f(x)}{\lambda_0} \\
&= \frac{f(z) - f(x)}{\lambda} + f(x)\left[\frac{1}{\lambda} - \frac{1}{\lambda_0}\right] \le \frac{y\|z - x\|}{\lambda} + f(x)\left[\frac{1}{\lambda} - \frac{1}{\lambda_0}\right].
\end{aligned}$$

Adding (2) to (1), we get

$$\begin{aligned}
&\frac{f(z+\lambda d) - f(x)}{\lambda} - \frac{f(x+\lambda_0 d) - f(x)}{\lambda} \\
&\le \frac{y\{2\|z - x\| + \|d(\lambda - \lambda_0)\|\}}{\lambda} \\
&\quad + \left(\frac{1}{\lambda} - \frac{1}{\lambda_0}\right)[f(x+\lambda_0 d) - f(x)] \\
&= \frac{2y\|z - x\|}{\lambda} + \frac{y|\lambda - \lambda_0|\,\|d\|}{\lambda} \\
&\quad + \left(\frac{1}{\lambda} - \frac{1}{\lambda_0}\right)[f(x+\lambda_0 d) - f(x)] \\
&\le \frac{|\lambda - \lambda_0|\,\|d\|y}{\lambda} + \frac{\lambda_0 - \lambda}{\lambda\lambda_0}\lambda_0 y\|d\| + \frac{2y\|z - x\|}{\lambda} \\
&= \frac{2|\lambda - \lambda_0|y\|d\|}{\lambda} + \frac{2y\|z - x\|}{\lambda}.
\end{aligned}$$

So for $\lambda > 0$ there is a $\delta > 0$ such that if $|\lambda - \lambda_0| < \delta$ and $\|x - z\| < \delta$ we have

$$\frac{f(z+\lambda d) - f(z)}{\lambda} \le \frac{f(x+\lambda_0 d) - f(x)}{\lambda_0} + \varepsilon y,$$

$$\sup_{\|x - z\| < \delta} \frac{f(z + (\lambda_0 + \lambda)d) - f(z)}{\lambda_0 + \delta} \le \frac{f(x+\lambda_0 d) - f(x)}{\lambda_0} + \varepsilon y.$$

Taking the infimum with respect to both $\lambda_0, \delta > 0$ we have

$$\inf_{\substack{\lambda_0>0\\ \delta>0}} \sup_{\|x-z\|<\delta} \frac{f(z+(\lambda_0+\lambda)d)-f(z)}{\lambda_0+\delta}$$
$$\leq \inf_{\lambda_0>0} \frac{f(x+\lambda_0 d)-f(x)}{\lambda} + \varepsilon y,$$
$$f^0(x; d) \leq f'(x; d) + \varepsilon y.$$

Let $\varepsilon \downarrow 0$. Then $f^0(x; d) \leq f'(x; d)$. Hence $f^0(x; d) = f'(x; d)$ for all $d \in X$. So $\partial f(x) = \partial_c f(x)$ for all $x \in X$. □

Recall the following definition.

DEFINITION 3.3. A mapping f: $X \to Y$ is said to be *Gateaux differentiable* if and only if

$$\lim_{\lambda \to 0} \frac{f(x+\lambda d)-f(x)}{\lambda} = f'_G(x; d)$$

exists and $f'_G(x, \cdot) \in \mathfrak{L}(X, Y)$. The mapping f is called *continuously Gateaux differentiable* if and only if:

(1) f is Gateaux differentiable.

(2) $x \to f'_G(x; \cdot)$ $(X \to \mathfrak{L}_s(X, Y))$ is continuous from X with the norm topology to $\mathfrak{L}(X, Y)$ with the strong operator topology.

Before going on to our next theorem, we need some auxiliary material.

Let Y have a strong unit e (this is the case when $(\mathring{K}_Y^+) \neq \varnothing$). Define $\|y\|_1 = \inf\{k: |y| \leq ke\}$. Then we have

LEMMA 3.2. *$\|\cdot\|_1$ is a norm on Y and $(Y, \|\cdot\|_1)$ is a Banach space. Furthermore, if τ_1 is the topology induced by $\|\cdot\|_1$ then it is stronger than the topology τ_0 induced by the original norm $\|\cdot\|$.*

Proof. Clearly

$$\| \, \|y\|_1 = |\lambda| \inf\{k': |y| \leq k'e\} = \inf\{|\lambda| k': |y| \leq k'e\}$$
$$= \inf\{k: |y| \leq (k/|\lambda|)e\}$$
$$= \inf\{k: |\lambda| |y| \leq ke\} = \|\lambda y\|_1.$$

Also

$$\begin{aligned}\|w\|_1 + \|y\|_1 &= \inf\{k_1 : |w| \le k_1 e\} + \inf\{k_2 : |y| \le k_2 e\}\\ &= \inf\{k_1 + k_2 : |w| \le k_1 e, |y| \le k_2 e\}\\ &\ge \inf\{k_1 + k_2 : |w| + |y| \le (k_1 + k_2)e\}\\ &\ge \inf\{k_1 + k_2 : |w + y| \le (k_1 + k_2)e\}\\ &= \|w + y\|_1.\end{aligned}$$

Finally,

$$0 = \|y\|_1 = \inf\{k : |y| \le ke\} \leftrightarrow |y| = 0 \leftrightarrow y = 0.$$

So indeed $\|\cdot\|_1$ is a norm on Y.

Now let $\{y_n\}_{n\in N} \subseteq Y$ be a Cauchy sequence in the norm $\|\cdot\|_1$. Then for some $N_0 > 0$ and $n, m \ge N_0$, we have

$$\|y_n - y_m\| < \varepsilon \to |y_n - y_m| < e\varepsilon,$$

which implies $\{y_n\}_{n\in N}$ order converges to some $y \in Y$. So $|y_n - y| < \varepsilon e$ for all $n \ge N_0$, which implies $\|y_n - y\|_1 < \varepsilon$ for all $n \ge N_0$. Hence

$$y_n \overset{\|\cdot\|_1}{\to} y.$$

So indeed Y is a Banach space with respect to the norm $\|\cdot\|_1$. Finally, from the definition of $\|\cdot\|_1$, we see there is $\{\lambda_n\}_{n\in N} \subseteq R^+$ such that $\lambda_n \downarrow \|y\|_1$.

Now

$$|y| \le \lambda_n e \to \|y\| \le \lambda_n \|e\| \to \|y\| \le \|y\|_1 \|e\|.$$

Since e is a strong unit, $\|e\| \le 1$, which implies $\|y\| \le \|y\|_1$. So indeed τ_1 is stronger than τ_0 as claimed by the lemma. □

REMARK. From the above lemma, we deduce the following implications:

(1) $\|\cdot\|_1$-convergence implies $\|\cdot\|$-convergence.

(2) $\|\cdot\|_1$-continuity implies $\|\cdot\|$-continuity.

The converse implications actually also hold (for details see **[24]**).

Also it is easy to see that $(Y, \|\cdot\|_1)$ is still a Banach lattice.

Then we have the following theorem which relates the generalized gradient and the Gateaux differential of f and which is an extension to vector valued mappings of a result of Clarke (**[3]**, **[7]**).

THEOREM 3.3. *If* $f \in L^{\text{oc}}_{\text{ip}}(X, Y)$ *and it continuously Gateaux differentiable for* $\|\cdot\|_1$, *then* $\partial f(x) = \{f'_G(x)\}$

Proof. By the $\|\cdot\|_1$-continuity of the Gateaux differential, we have that for every $\varepsilon > 0$ there is a $\delta > 0$ such that "if $\|x - z\| < \delta$ the $\|f'_G(x; d) - f'_G(z; d)\|_1 < \varepsilon$ for all $d \in F \subseteq Y$ finite." (Continuity in the strong operator topology.)

Now let $v \in X$ such that $\|x - v\| < \delta/2$ and consider $\lambda \in (0, \delta/2\|d\|)$.

Define $\phi(\lambda) = f(v + \lambda d)$. Since f is Gateaux differentiable, it is easy to see that ϕ is also differentiable. Then we have

$$\begin{aligned}\phi'(\lambda) &= \lim_{\mu\to\lambda} \frac{\phi(\mu) - \phi(\lambda)}{\mu - \lambda} = \lim_{\mu\to\lambda} \frac{f(v + \mu d) - f(v + \lambda d)}{\mu - \lambda} \\ &= \lim_{\mu\to\lambda} \frac{f(v + \lambda d + (\mu - \lambda)d) - f(v + \lambda d)}{\mu - \lambda} = f'_G(v + \lambda d; d).\end{aligned}$$

Let $r = \mu - \lambda$ such that $r < \delta/2\|d\|$.

Using the Mean Value Theorem for Bochner integrals, we have

$$\begin{aligned}&\frac{\phi(r) - \phi(0)}{r} - f'_G(x; d) = \frac{1}{r}\int_0^r \phi'(t)\, dt - f'_G(x; d) \\ &\to \frac{f(v + rd) - f(v)}{r} - f'_G(x; d) \\ &= \frac{1}{r}\int_0^r f'_G(v + td; d)\, dt - f'_G(x; d) \\ &= \frac{1}{r}\int_0^r [f'_G(v + td; d) - f'_G(x; d)]\, dt.\end{aligned}$$

Now observe that

$$\|v + td - x\| \le \|v - x\| + t\|d\| < \delta/2 + (\delta/2\|d\|)\|d\| = \delta.$$

So from the $\|\cdot\|_1$-continuity of f'_G and the properties of the Bochner integral (see **[10]**), we have

$$\begin{aligned}&\left\| \frac{f(v + rd) - f(v)}{r} - f'_G(x; d) \right\|_1 \\ &\le \frac{1}{r}\int_0^r \|f'_G(v + td; d) - f'_G(x; d)\|_1\, dt \\ &\le \frac{1}{r}\int_0^r \varepsilon\, dr = \varepsilon.\end{aligned}$$

This follows directly from the definition of the $\|\cdot\|_1$-norm, and means that

$$\left|\frac{f(v+rd)-f(v)}{r} - f'_G(x;d)\right| \le \varepsilon e \to \frac{f(v+rd)-f(v)}{r} \le f'_G(x;d) + \varepsilon e.$$

Then

$$\bigvee_{\substack{\|v-x\|<\delta/2\\ r<\delta/2\|d\|}} \frac{f(v+rd)-f(v)}{r} \le f'_G(x;d)+\varepsilon e$$
$$\to \bigwedge \bigvee_{\substack{\|v-x\|<\delta/2\\ r<\delta/2\|d\|}} \frac{f(v+rd)-f(v)}{r} \le f'_G(x;d)+\varepsilon e$$
$$\to f^0(x;d) \le f'_G(x;d)+\varepsilon e.$$

Let $\varepsilon \to 0$; then we get that $f^0(x;d) \le f'_G(x;d)$. But we know, in general,

$$f'_G(x;d) \le f^0(x;d).$$

So we conclude $f^0(x;d) = f'_G(x;d)$.

Now we know

$$\partial f(x) = \{A \in \mathfrak{L}(X,Y): A(d) \le f^0(x;d)\forall d \in X\}.$$

Since $f'_G(x;\cdot) \in \mathfrak{L}(X,Y)$, we can see we always have $\{f'_G(x;\cdot)\} \subseteq \partial f(x)$ when the Gateaux differential actually exists. Let $A \in \partial f(x)$. Then $A(d) \le f^0(x;d)$ for all $d \in X$ ((1)). Also for every $d \in X$, $-A(d) = A(-d) \le f^0(x;-d) = -f^0(x;d)$ (since we have $f^0(x;\cdot) = f'(x;\cdot) \in \mathfrak{L}(X,Y)$). So we get $A(d) \ge f^0(x;d)$ for all $d \in X$ ((2)).

From (1) and (2) we finally get $A(d) = f^0(x;d)$ for all $d \in X$ and all $A \in \partial f(x)$ so $\partial f(x) = \{f'_G(x;\cdot)\}$. □

Next we examine the generalized gradient of the sum of two locally o-Lipschitz mappings. In the real valued case, Clarke [**2**] proved $\partial(f+g) \subseteq \partial f + \partial g$ holds—and with some additional regularity assumptions, he also proved equality. In turns out that the same is true for our class of mappings.

First we need a definition which extends a notion of Clarke [**7**, Definition 3] to vector valued mappings.

DEFINITION 3.4. We call a mapping $f \in L^{\mathrm{oc}}_{\mathrm{ip}}(X,Y)$ *o-regular* if and only if $f'(x;d) = f^0(x;d)$ for all $d \in X$, where $f'(x;d)$ is just the upper o-directional derivative of f at x in the direction d.

Then from Theorems 3.2 and 3.3, we see that convex and Gateaux differentiable mappings are o-regular.

Now we can state and prove our next theorem.

THEOREM 3.4. *If $f, g \in L_{\text{ip}}^{\text{oc}}(X, Y)$ then*

$$\partial(\alpha f)(x) = \alpha \partial f(x) \quad \text{for } \alpha > 0$$

and

$$\partial(\alpha f + \beta g)(x) \subseteq \alpha \partial f(x) + \beta \partial g(x) \quad \text{for } \alpha, \beta > 0;$$

if f, g are, in addition, o-regular, then equality holds.

Proof. Let $\alpha > 0$.
Then

$$\partial(\alpha f)^0(x; d) = \overline{\lim_{\substack{z \to x \\ \lambda \downarrow 0}}} \frac{\alpha f(z + \lambda d) - \alpha f(z)}{\lambda} = \overline{\lim_{\substack{z \to x \\ \lambda \downarrow 0}}} \frac{\alpha[f(z + \lambda d) - f(z)]}{\lambda}$$

$$= \alpha \overline{\lim_{\substack{z \to x \\ \lambda \downarrow 0}}} \frac{f(z + \lambda d) - f(z)}{\lambda} = \alpha f^0(x; d).$$

Then

$$\partial(\alpha f)(x) = \{A \in \mathfrak{L}(X, Y): A(d) \leq (\alpha f)^0(x; d)\ \forall d \in X\}$$
$$= \{A \in \mathfrak{L}(X, Y): A(d) \leq \alpha f^0(x; d)\ \forall d \in X\}.$$

So, if $A \in \partial(\alpha f)(x)$ then

(1) $\quad A/\alpha \in \partial f(x) \to A \in \alpha \partial f(x) \to \partial(\alpha f)(x) \subseteq \alpha \partial f(x).$

Also, if $B \in \alpha \partial f(x)$ then

(2) $\quad B/\alpha \in \partial f(x) \to B(d)/\alpha \leq f^0(x; d), \qquad d \in X,$
$$\to B(d) \leq \alpha f^0(x; d) \to B(d) \leq (\alpha f)^0(x; d) \quad \forall d \in X.$$

So $B \in \partial(\alpha f)(x)$ and therefore $\alpha \partial f(x) \subseteq \partial(\alpha f)(x)$. From (1) and (2), we conclude

$$\partial(\alpha f)(x) = \alpha \partial f(x) \quad \text{for all } x \in X.$$

To prove the other half of the first part of the theorem, we simply need to show that

$$\partial(f + g)(x) \subseteq \partial f(x) + \partial g(x) \quad \text{for all } x \in X.$$

Recall that $f^0(x;\cdot)$ is sublinear and continuous (in fact o-Lipschitz by Proposition 3.1). So using our generalization of the Moreau-Rockafellar Theorem (see [**19**], Theorem 4.1), the following is true:

(3) $\partial_c(f^0+g^0)(x;0)=\partial_c f^0(x;0)+\partial_c g^0(x;0)=\partial f(x)+\partial g(x).$

On the other hand it is easy to see that

$$f^0(x;\cdot)+g^0(x;\cdot)\geq(f+g)^0(x;\cdot)$$

(to check that, just us the definition of f^0)

(4) $\rightarrow\partial_c(f^0+g^0)(x;0)\supseteq\partial_c(f+g)^0(x;0)=\partial(f+g)(x).$

From (3) and (4) we deduce that

$$\partial(f+g)(x)\subseteq\partial f(x)+\partial g(x).$$

This completes the proof of the first part of the theorem. For the second part, we proceed as follows: If f, g are o-regular, then

$$(f+g)'(x;d)=(f+g)^0(x;d)\quad\forall d\in X$$

and

$$f'(x;d)+g'(x;d)=f^0(x;d)+g^0(x;d)\quad\forall d\in X.$$

But

$$\begin{aligned}&(f+g)'(x;d)=f'(x;d)+g'(x;d)\quad\forall d\in X\\&\quad\rightarrow(f+g)^0(x;d)=f^0(x;d)+g^0(x;d)\quad\forall d\in X.\end{aligned}$$

So we have

(5) $\partial_c(f+g)^0(x;0)=\partial_c f^0(x;0)+\partial_x g^0(x;0)=\partial f(x)+\partial g(x)$

and

(6) $\partial_c(f^0+g^0)(x;0)=\partial_c(f+g)^0(x;0)=\partial(f+g)(x).$

From (5) and (6) we conclude

$$\partial(f+g)(x)=\partial f(x)+\partial g(x).\qquad\square$$

REMARK. It is worth noting an important difference from the convex case. There we saw that, in general, $\partial f(x)+\partial g(x)\subseteq\partial(f+g)(x)$ (see [**19**], §4), while here, as we saw in the previous theorem, in general, it is true that $\partial(f+g)(x)\subseteq\partial f(x)+\partial g(x)$.

4. Generalized subdifferential calculus. Here we develop a calculus for the generalized gradients. The sources of inspiration are two. The first and basic one is the existing calculus for real valued functions developed primarily by Clarke [**2**], [**3**], [**7**], and also Aubin [**1**] and Hiriart-Urruty [**11**], [**12**], [**13**]. The second is the calculus developed in [**19**] for the convex subdifferential. We would like to know to what extent we can have a similar theory for the generalized gradients.

We start with the following result.

Let Z be a Banach space and assume Y has the diagonal property.

PROPOSITION 4.1. *If* $f \in L_{\text{ip}}^{\text{oc}}(Z, Y)$ *and* $A \in \mathfrak{L}(X, Z)$, *then*

$$\partial(f \circ A)(x) \supseteq \partial f(Ax) \circ A \quad \forall x \in X;$$

if, in addition, f is o-regular at Ax_0 and $R(A) = Z$, then

$$\partial(f \circ A)(x_0) = \partial f(Ax_0) \circ A.$$

Proof. We already know from the corollary to Proposition 2.2 that $f \circ A \in L_{\text{ip}}^{\text{oc}}(X, Y)$. Now

$$\begin{aligned}(f \circ A)^0(x; d) &= \overline{\lim_{\substack{h \to 0 \\ \lambda \downarrow 0}}} \frac{(f \circ A)(x + h + \lambda d) - (f \circ A)(x + h)}{\lambda} \\ &= \overline{\lim_{\substack{h \to 0 \\ \lambda \downarrow 0}}} \frac{f(Ax + Ah + \lambda Ad) - f(Ax + Ah)}{\lambda} \\ &= f^0(Ax; Ad)\end{aligned}$$

(since $A \in \mathfrak{L}(X, Z)$, $h \to 0$ implies $Ah \to 0$).

Next let $\Gamma \in \partial f(Ax)$. By the definition of $\partial f(Ax)$, we have

$$\begin{aligned}&\Gamma(h) \le f^0(Ax; h) \quad \forall h \in Z \\ &\quad \to (\Gamma \circ A)(d) = \Gamma(Ad) \le f^0(Ax; Ad) \quad \forall d \in X \\ &\quad \to (\Gamma \circ A)(d) \le (f \circ A)^0(z; d) \quad \forall d \in X \\ &\quad \to (\Gamma \circ A) \in \partial(f \circ A)(x).\end{aligned}$$

So we conclude that $\partial f(Ax) \circ A \subseteq \partial(f \circ A)(x)$.

Now if f is o-regular at x_0, then by the chain rule for o-directional derivatives obtained in [**19**], we have

$$(1) \qquad (f \circ A)^0(x_0; d) = (f \circ A)'(x_0; d) = f'(Ax_0; Ad).$$

(Actually there the chain rule was proved using Lemma 8 of Valadier [**26**] for f convex. But since by diagonal property o and ru convergence coincide and Y is a Banach lattice, then we can see, using Proposition 2.4, p. 162 of [**21**], that the chain rule extends to the present case also.)

Let $B \in \partial(f \circ A)(x_0)$. It is easy to see that $\ker A \subseteq \ker B$. So we have the following situation:

$$\begin{array}{ccc} & A & \\ X & \rightarrow & Z \\ & \searrow & \\ & B & Y \end{array}$$

Since $\operatorname{Im} A = Z$, by the factorization theorem, there is $\Gamma \in \mathfrak{L}(Z, Y)$ such that $B = \Gamma \circ A$. Then by (1) we have

$$B(d) = (\Gamma \circ A)(d) = \Gamma(Ad) \leq f^0(Ax_0; Ad) \quad \forall d \in X.$$

Since A is surjective, we conclude that $\Gamma \in \partial f(Ax_0)$. Therefore, we have $\partial f(Ax_0) \circ A \supseteq \partial(f \circ A(x_0))$. This fact combined with the first part of the proposition implies

$$\partial(f \circ A)(x_0) = \partial f(Ax_0) \circ A. \qquad \square$$

PROPOSITION 4.2. *If $f \in L_{\mathrm{lp}}^{\mathrm{oc}}(X, Y)$, $A \in [\mathfrak{L}(Y)]^+$ and is also invertible, and Y has the diagonal property, then*

$$\partial(A \circ f)(x) = A \circ \partial f(x) \quad \textit{for all } x \in X.$$

Proof. First let us check that the composite map is in fact locally o-Lipschitz. So we have for $x, z \in U$,

$$\begin{aligned} |(A \circ f)(x) - (A \circ f)(z)| &= |A(f(x)) - A(f(z))| \\ &= |A(f(x) - f(z))|. \end{aligned}$$

Using the positivity of A we have

$$|A(f(x) - f(z))| \leq A\big(|f(x) - f(z)|\big).$$

Now since $f \in L_{\mathrm{lp}}^{\mathrm{oc}}(X, Y)$, we have $|f(x) - f(z)| \leq y\|x - z\|$ for some $y \in K_Y^+$ and for all $x, z \in U =$ open and bounded subset of X. Another use of the positivity of A gives up

$$\begin{aligned} &A\big(|f(x) - f(z)|\big) \leq A(y)\|x - z\| \\ &\quad \to |A(f(x) - f(z))| \leq A(y)\|x - z\| \\ &\quad \to |(A \circ f)(x) - (A \circ f)(z)| \leq A(y)\|x - z\| \quad \text{for all } x, z \in U. \end{aligned}$$

So indeed, $A \circ f$ is locally o-Lipschitz with Lipschitz constant $A(y)$.

We now see that

$$\begin{aligned}
&\frac{(A\circ f)(x+h+\lambda d)-(A\circ f)(x+h)}{\lambda}\\
&\quad = A\left(\frac{f(x+h+\lambda d)-f(x+h)}{\lambda}\right)\to (A\circ f)^0(x;d)\\
&\quad = \overline{\lim_{\substack{h\to 0\\ \lambda\downarrow 0}}}\ \frac{(A\circ f)(x+h+\lambda d)-(A\circ f)(x+h)}{\lambda}\\
&\quad = \overline{\lim_{\substack{h\to 0\\ \lambda\downarrow 0}}}\ A\left(\frac{f(x+h+\lambda d)-f(x+h)}{\lambda}\right)\\
&\quad \le \lim_{\varepsilon\downarrow 0} A(r_\varepsilon)
\end{aligned}$$

where

$$r_\varepsilon = \bigvee_{\substack{\|h\|<\varepsilon\\ \lambda<\varepsilon}} \frac{f(x+h+\lambda d)-f(x+h)}{\lambda}.$$

But since Y has the diagonal property, we know (see **[21]**)

$$r_\varepsilon \downarrow f^0(x;d) \to r_\varepsilon \xrightarrow{ru} f^0(x;d).$$

This, by Lemma 2.2 of **[19]**, means $r \xrightarrow{s} f^0(x;d)$. Hence since $A\in\mathfrak{L}(X,Y)$, we deduce that

$$(A\circ f)^0(x;d)\le A(f^0(x;d)).$$

Now let

$$\begin{aligned}
B\in\partial(A\circ f)(x) &\to B(d)\le (A\circ f)^0(x;d)\quad \forall d\in X\\
&\to B(d)\le A(f^0(x;d))\to B\in A\circ\partial f(x)\\
&\to \partial(A\circ f)(x)\subseteq A\circ\partial f(x).
\end{aligned}$$

It is easy to see that the opposite inclusion also holds. So we conclude that

$$\partial(A\circ f)(x) = A\circ\partial f(x)\quad \text{for all } x\in X. \qquad \square$$

Our next goal is to obtain a formula for the generalized gradient of an integral operator.

The final theorem will be based on an analogous result which has been proved for the convex case by Saint-Pierre (see **[23]**).

The proof of the theorem will follow the lines of the proof of the corresponding result for real valued functions due to Clarke (see **[7]**, Theorem 1). However, to achieve that, we first need to develop some auxiliary analytical background, which includes an interesting, in its own sake, generalization of Fatou's Lemma.

Let (Ω, Σ, μ) be a positive measure space, X a separable Banach space and Y a separable, order complete Banach lattice which has the R-N property and is w-seq. complete. By that we mean every weak Cauchy sequence has a weak limit. Recall that all reflexive Banach spaces are weakly sequentially complete (see [**8**]).

In the sequel all vector integrals are defined in the sense of Bochner (see [**10**]).

LEMMA 4.1. *If* $\phi \in L_1(\Omega, Y)$ *then*

$$\left|\int_\Omega \phi(\omega)\, d\mu(\omega)\right| \leq \int_\Omega |\phi(\omega)| d\mu(\omega).$$

Proof. Assume $\int_\Omega |\phi(\omega)| d\mu(\omega) < +\infty$ (where $+\infty$ is a greatest element adjoined to Y) or the above inequality is obvious.

By definition, $\phi(\omega) \leq |\phi(\omega)|$. So by Lemma 3.3 of [**19**], we have

$$\int_\Omega \phi(\omega)\, d\mu(\omega) \leq \int_\Omega |\phi(\omega)|\, d\mu(\omega) \tag{1}$$

and, similarly,

$$-\int_\Omega \phi(\omega)\, d\mu(\omega) \leq \int_\Omega |\phi(\omega)|\, d\mu(\omega). \tag{2}$$

From (1) and (2) we have

$$\left(\int_\Omega \phi(\omega)\, d\mu(\omega)\right) \vee \left(-\int_\Omega \phi(\omega)\, d\mu(\omega)\right) \leq \int_\Omega |\phi(\omega)|\, d\mu(\omega)$$
$$\to \left|\int_\Omega \phi(\omega)\, d\mu(\omega)\right| \leq \int_\Omega |\phi(\omega)|\, d\mu(\omega). \qquad \square$$

The next result is an interesting generalization, for mappings taking values in a Banach lattice, of the well-known Fatou's Lemma.

PROPOSITION 4.3 (*Generalized Fatou's Lemma*) *If* $|f_n| \leq \phi \in L_1(\Omega, Y)\, \forall n \in N$ *and* $\limsup_{n\to\infty} f_n(\omega)$ *exists* μ-*a.e. and is in* $L_1(\Omega, Y)$, *then*

$$\limsup_{n\to\infty} \int_\Omega f_n(\omega)\, d\mu(\omega) \leq \int_\Omega \limsup_{n\to\infty} f_n(\omega)\, d\mu(\omega).$$

Proof. Let $g_m(\omega) = \bigvee_{n \geq m} f_n(\omega)$. Then $g_m(\omega) \downarrow \limsup_{n \to \infty} f_n(\omega)$ μ-a.e. By Lemma 3.3 of [**19**], we know that

$$\text{(1)} \qquad \int_\Omega g_m(\omega)\, d\mu(\omega) \geq \int_\Omega f_n(\omega)\, d\mu(\omega) \quad \forall n \geq m$$

$$\to \int_\Omega g_m(\omega)\, d\mu(\omega) \geq \bigvee_{n \geq m} \int_\Omega f_n(\omega)\, d\mu(\omega)$$

$$\to \bigwedge_{m \in N} \int_\Omega g_m(\omega)\, d\mu(\omega) \geq \bigwedge_{m \in N} \bigvee_{n \geq m} \int_\Omega f_n(\omega)\, d\mu(\omega)$$

$$\to \bigwedge_{m \in N} \int_\Omega g_m(\omega)\, d\mu(\omega) \geq \limsup_{n \to \infty} \int_\Omega f_n(\omega)\, d\mu(\omega).$$

We will show that the left-hand side of the above inequality is just $\int_\Omega g(\omega)\, d\mu(\omega)$, where $g(\omega) = \limsup_{n \to \infty} f_n(\omega)$ μ-a.e. We have

$$\bigwedge_{m \in N} (g_m(\omega) - g(\omega)) = \bigwedge_{m \in N} g_m(\omega) - g(\omega) = 0.$$

Since a separable, order complete Banach lattice is order continuous (see [**24**]), this implies

$$\|g_m(\omega) - g(\omega)\| \to 0 \quad \mu\text{-a.e. as } m \to \infty \to g_m \xrightarrow{s} g \quad \mu\text{-a.e.}$$

From the Dominated Convergence Theorem for Bochner integrals (see [**10**]), we get

$$\int_\Omega g_m(\omega)\, d\mu(\omega) \xrightarrow{s} \int_\Omega g(\omega)\, d\mu(\omega).$$

But a weakly sequential Banach lattice is a *KB*-space in which strong and o^*-convergence are equivalent. Hence we have

$$\int_\Omega g_m(\omega)\, d\mu(\omega) \xrightarrow{o^*} \int_\Omega g(\omega)\, d\mu(\omega).$$

But, since $\{\int_\Omega g_m(\omega)\, d\mu(\omega)\}_{n \to N}$ is a monotone decreasing sequence, order and order* convergence coincide. So

$$\int_\Omega g_n(\omega)\, d\mu(\omega) \downarrow \int_\Omega g(\omega)\, d\mu(\omega).$$

Going back to (1), we get

$$\int_\Omega g(\omega)\, d\mu(\omega) \geq \limsup_{n \to \infty} \int_\Omega g_n(\omega)\, d\mu(\omega) \to \int_\Omega \limsup_n f_n(\omega)\, d\mu(\omega)$$

$$\geq \limsup_n \int_\Omega f_n(\omega)\, d\mu(\omega).$$

So we get the Generalized Fatou Lemma. $\square$

Now we are already to formulate and prove the theorem about the generalized gradient of integral operators. As mentioned earlier, out proofs follows the proof of Theorem 1 of Clarke [7].

Let $f_\omega \in L_{\text{ip}}^{\text{oc}}(X, Y)$ μ-a.e. such that $\omega \to f_\omega$ is weakly measurable and $|f_\omega(x)| \leq \phi(\omega) \in L(\Omega, Y)$ for all $x \in X$.

Then we have the following theorem relating the subdifferential of $f(x) = \int_\Omega f_\omega(x)\, d\mu(\omega)$ with those of $f_\omega(x) \omega \in \Omega$.

THEOREM 4.1. *If the above assumption holds then*

$$\partial f(x) \subseteq \int_\Omega \partial f_\omega(x)\, d\mu(\omega);$$

if, in addition, f_ω is o-regular for μ-almost all $\omega \in \Omega$, then

$$\partial f(x) = \int_\Omega \partial f_\omega(x)\, d\mu(\omega).$$

Proof. We have $f(x) = \int_\Omega f_\omega(x)\, d\mu(\omega)$. Let us calculate its generalized o-directional derivative. We have

$$\begin{aligned} f^0(x; d) &= \overline{\lim_{\substack{h \to 0 \\ \lambda \downarrow 0}}} \frac{f(x + h + \lambda d) - f(x + h)}{\lambda} \\ &= \overline{\lim_{\substack{h \to 0 \\ \lambda \downarrow 0}}} \frac{\int_\Omega f_\omega(x + h + \lambda d)\, d\mu(\omega) - \int_\Omega f_\omega(x + h)\, d\mu(\omega)}{\lambda} \\ &= \overline{\lim_{\substack{h \to 0 \\ \lambda \downarrow 0}}} \int_\Omega \frac{f_\omega(x + h + \lambda d) - f_\omega(x + h)}{\lambda}\, d\mu(\omega). \end{aligned}$$

Using the Generalized Fatou Lemma (see Proposition 4.3), we have

$$\begin{aligned} f^0(x; d) &\leq \int_\Omega \overline{\lim_{\substack{h \to 0 \\ \lambda \downarrow 0}}} \frac{f_\omega(x + h + \lambda d) - f_\omega(x + h)}{\lambda}\, d\mu(\omega) \\ &= \int_\Omega f_\omega^0(x; d)\, d\mu(\omega). \end{aligned}$$

So

$$f_\omega^0(x; d) \leq \int_\Omega f_\omega^0(x; d)\, d\mu(\omega). \tag{1}$$

At this point, note that since X is separable, there is a countable set $D \subseteq X$ such that $\overline{D} = X$. Then

$$\bigvee_{\substack{h \in D \\ \|x-h\|<\varepsilon \\ \lambda \text{ rational}, \lambda<\delta}} \frac{f(x+h+\lambda d)-f(x+h)}{\lambda}$$

is strongly measurable. So let $\varepsilon_n \downarrow 0$ and $\delta_n \downarrow 0$. We conclude $f_\omega^0(x; d)$ is also strongly measurable.

Now consider the mapping.

$$f^N(d) = \int_\Omega f_\omega^0(x; d)\, d\mu(\omega).$$

Because $f_\omega^0(x; \cdot)$ is convex, continuous and majorized in a neighborhood of the origin, then, by Saint-Pierre's result [**23**], we have

$$\partial_c f^N(0) = \int_\Omega \partial_c f_\omega^0(x; 0)\, d\mu(\omega).$$

But recall that

$$\partial_c f_\omega^0(x; 0) = \partial f_\omega(x).$$

So from (1)

$$\partial f(x) \subseteq \int_\Omega \partial f_\omega(x)\, d\mu(\omega).$$

This proves the first part of the theorem.

Now assume f_ω is o-regular for μ-almost all $\omega \in \Omega$. Then

$$\begin{aligned} & f_\omega^0(x; d) = f_\omega'(x; d) \quad \mu\text{-a.e.} \\ (2) \qquad & \rightarrow \int_\Omega f_\omega^0(x; d)\, d\mu(\omega) = \int_\Omega f_\omega'(x; d)\, d\mu(\omega) \\ & \rightarrow f^0(x; d) \leq f'(x; d) \quad (\text{using Proposition 3.1 of } [\mathbf{19}]). \end{aligned}$$

But we know that in general $f'(x; d) \leq f^0(x; d)$. So it follows from (2) that

$$f^0(x; d) = f'(x; d) = \int_\Omega f_\omega^0(x; d)\, d_m(\omega).$$

Applying Saint-Pierre's result (see [23]) to the convex continuous mappings f^0 and f^0_ω we get

$$\partial_c f^0(x;0) = \int_\Omega \partial_c f^0_\omega(x;0)\, d\mu(\omega) \to \partial f(x) = \int_\Omega \partial f(x)\, d\mu(\omega). \qquad \square$$

Again we go back to X a Banach space and Y a Banach lattice.

It is obvious that for mappings defined on product spaces, we can have partial generalized o-directional derivatives and therefore partial generalized gradients.

For such mappings we have the following result.

PROPOSITION 4.4. *If Z is a Banach space, $f \in L^{\mathrm{oc}}_{\mathrm{ip}}(X \times Z, Y)$ and is o-regular at (x, z), then*

$$\partial f(x, z) \subseteq \partial f_x(x, z) \otimes \partial f_z(x, z).$$

Proof. Let $(A, B) \in \partial f(x, z) \subseteq \mathfrak{L}(X, Y) \otimes \mathfrak{L}(Z, Y)$. We have, for $w = (x, z)$,

$$o\text{-}\lim_{\lambda \downarrow 0} \frac{f(x + \lambda h_1, z) - f(x, z)}{\lambda} = f'_x(w; h_1)$$
$$= f'(w; (h_1, 0)) = f^0(w; (h_1, 0)).$$

But

$$(A \otimes B)(h_1, 0) = A(h_1) \leq f'(w; (h_1, 0)) = f^0(w; (h_1, 0)).$$

Hence $A \in \partial_X f(x, z)$. Similarly we deduce that $B \in \partial_z f(x, z)$. So the Proposition follows. $\square$

We conclude this section with some additional useful observations about generalized gradients. In the sequel let X be a Banach lattice, too.

PROPOSITION 4.5. *If $f \in L^{\mathrm{oc}}_{\mathrm{ip}}(X, Y)$ is monotone increasing then $\partial f(x) \subseteq [\mathfrak{L}(X, Y)]^+$.*

Proof. By definition we have

$$f^0(x; d) = \overline{\lim_{\substack{h \to 0 \\ \lambda \to 0}}} \frac{f(x + h + \lambda d) - f(x + h)}{\lambda}.$$

Then using the monotonicity of f, we have, for $d \in K_X^+$,

$$
\begin{aligned}
f(x+h) &\le f(x+h+\lambda d) \qquad (\lambda > 0)\\
&\to f(x+h) - f(x+h+\lambda d) \le 0\\
&\to \frac{f(x+h) - f(x+h+\lambda d)}{\lambda} \le 0\\
&\to \frac{f(x+h+\lambda d + \lambda(-d)) - f(x+h+\lambda d)}{\lambda} \le 0\\
&\to \varlimsup_{\substack{h\to 0\\ \lambda\downarrow 0}} \frac{f(x+h+\lambda d + \lambda(-d)) - f(x+h+\lambda d)}{\lambda} \le 0\\
&\to f^0(x; -d) < 0 \to -f^0(x; -d) \ge 0.
\end{aligned}
$$

We know that $\partial f(x) = \partial_c f^0(x; 0)$. So if $A \in \partial f(x)$, then $-f^0(x; -d) \le A(d) \le f^0(x; d)$. Hence for $d \in K_X^+$, $0 \le A(d)$. Therefore

$$\partial f(x) \subseteq [\mathfrak{L}(X, Y)]^+ \quad \forall x \in X. \qquad \square$$

COROLLARY. *If $f \in L_{\text{ip}}^{\text{oc}}(X, Y)$ is monotone decreasing then $\partial f(x) \subseteq [\mathfrak{L}(X, Y)]^- \ \forall x \in X$.*

This concludes the study of the generalized subdifferential calculus.

In the next section we use the theory developed up to now to get a necessary condition for the existence of a local minimum. More results in that direction together with some new interesting properties of locally o-Lipschitz functions and additional results on the subdifferential calculus appear in **[20]**.

5. A necessary condition for the existence of local minima. Let X be a Banach space and Y an order complete Banach lattice. Let S be a subset of X and f a mapping from S to Y.

In this section we are concerned with the following problem:

[P] "Find $x \in X$, for which there is a neighborhood U such that for all $z \in U \cap S$, we have $f(x) \le f(z)$".

To get necessary condition for the existence of such a point, we need to recall some basic geometric objects.

DEFINITION 5.1. Let $x \in \text{cl}\, S$. The tangent cone of S at x is the set

$$T_S(x) = \Big\{ d \in X\colon d = \lim_{n\to\infty} \lambda_n(x_n - x_0) \text{ where } \lambda_n > 0,\\ x_n \in S \text{ and } x_n \to x_0 \Big\}.$$

The major disadvantage of this cone is that it is not convex and so is not appropriate for using duality techniques, so people were interested in convex subcones of $T_S(x)$. Then in that direction the best definition was introduced by Clarke [2], [3].

DEFINITION 5.2. Let $x \in \mathrm{cl}\, S$. The "Clarke tangent cone of S at x" is the set

$$\begin{aligned}\tau_S(x) = \{d \in X\colon \forall \{x_n\}_{n\in N} \subseteq \mathrm{cl}\, S\, \forall \lambda_n \downarrow 0\\ \exists d_n \text{ s.t. } d_n \to d \text{ and } x_n + \lambda_n d_n \in S \text{ for all } n \in N\}.\end{aligned}$$

This cone is always convex and, of course, $\tau_S(x) \subseteq T_S(x)$.

REMARK. Definition 5.2 is not actually the original definition of Clarke. However, as Rockafellar points out in [22], the two are equivalent by Proposition 3.7 of [2]. (See also Hiriart-Urruty [14].)

Let M be any closed convex subcone of $T_S(x)$.

Then $M_Y^* = \{A \in \mathfrak{L}(X, Y)\colon A(d) \geq 0\, \forall d \in M\}$. We have the following necessary condition for the existence of a local mimimum.

THEOREM 5.1. *If $f \in L_{\mathrm{ip}}^{\mathrm{oc}}(X, Y)$ and x solves* [P] *then* $\partial f(x) \cap M_Y^* \neq \varnothing$.

Proof. We start by showing $f^0(x; d) \leq 0$ for every $d \in T_S(x)$. From Definition 5.1, $d = \lim_{n\to\infty} \lambda_n(x_n - x)$, where $\lambda_n > 0$, $x_n \in S$ and $x_n \to x$. Clearly this means $\lambda_n \uparrow \infty$ so $1/\lambda_n \downarrow 0$. Consider the quotient

$$q_n = \frac{f(x + (1/\lambda_n)d) - f(x)}{(1/\lambda_n)}.$$

Then

$$(*) \qquad q_n = \frac{f(x + (1/\lambda_n)d) - f(x_n) - f(x)}{1/\lambda_n}$$

$$(\text{where } x_n = x + (1/\lambda_n)\lambda_n(x_n - x))$$

$$= \frac{f(x + (1/\lambda_n)d) - f(x_n)}{1/\lambda_n} + \frac{f(x_n) - f(x)}{1/\lambda_n}.$$

Observe that since x is a local minimum (i.e. solves [P]) and eventually $x_n \in U \cap S$, we have $(f(x_n) - f(x))/(1/\lambda_n) \geq 0$. Also, since f is locally o-Lipschitz, we have

$$\frac{f(x + (1/\lambda_n)d) - f(x_n)}{1/\lambda_n} \leq \frac{\|x + (1/\lambda_n)d - x - (x_n - x)\|}{1/\lambda_n}$$

$$= \|d - \lambda_n(x_n - x)\|.$$

The last quantity tends to zero as $n \to \infty$. Using these observations in $(*)$, we get

$$\limsup_{n \to \infty} q_n \geq 0.$$

But $f^0(x; d) \geq \limsup_{n \to \infty} q_n > 0$. Hence $f^0(x; d) \geq 0 \, \forall d \in T_S(x)$. Now, by a result of Kutateladze [**18**], we have

$$\partial_c f^0_{X|M} = \partial_c f^0_X + \partial_c \delta_M \quad \text{where } \delta_M(x) = \begin{cases} 0 & \text{if } x \in M, \\ + & \text{if } x \notin M. \end{cases}$$

Since $0 \in M$ and $f^0_X(M) \geq 0$, we have

$$\begin{aligned} 0 \in \partial_c f^0_{X|M}(0) &\to 0 \in \partial_c f^0_X(0) + \partial_c \delta_M(0) \\ &\to 0 \in \partial f(x) + (-M^*_Y). \end{aligned}$$

Therefore, $\partial f(x) \cap M^*_Y \neq \varnothing$. □

REMARK 1. Since $M^*_Y = \{A \in \mathfrak{L}(X, Y)\colon A(d) \geq 0 \text{ for all } d \in M\}$, it is very easy to see that $\partial S_M(0) = -M^*_Y$.

REMARK 2. Since we always have $\tau_S(x) \subseteq T_S(x)$ and $\tau_S(x)$ is closed and convex, we can let $M = \tau_S(x)$.

6. Discussion of related work. Similar work was recently done by Hiriart-Urruty and Thibault [**15**], Ioffe [**16**], [**17**] and finally by Thibault [**25**]. The one that is closest to our approach is [**25**]. Thibault introduces a new class of vector valued mappings which he calls "compactly Lipschitzian mappings". This class includes the locally Lipschitz functions when the range space is R. For those functions he defines a subdifferential and develops the corresponding calculus. His definition of a compactly Lipschitzian mapping at a point is the following.

"$f\colon X \to Y$ is said to be compactly Lipschitzian at $x \in X$, if there is a mapping $K\colon X \to \operatorname{Comp} Y = \{\text{compact subsets of } Y\}$ and a mapping r: $(0, q] \times X \times Y$ which have the following properties:

(1) $\lim_{t \downarrow 0} r(t, x; v) = 0 \, \forall v \in X$.

(2) For each $v \in X$ there is a neighborhood Ω of $x \in X$ and $\eta \in [0, 1]$ such that

$$(f(x + tv) - f(x))/t \in K(x) + r(x, t; v) \quad \text{for all } x \in \Omega,\ t \in (0, \eta]."$$

Let us see how our definition compares with the above.

For all $x, z \in U \subseteq X$ bounded and open

$$|f(x) - f(z)| \leq y\|x - z\| \quad \text{for some } y \in K_Y^+ .$$

Then

$$\frac{f(x+tv) - f(x)}{t} \in \left[-y\frac{\|x + tv - x\|}{t}, y\frac{\|x + tv - x\|}{t} \right]$$
$$= [-y\|v\|, y\|v\|] \quad \text{and} \quad r \equiv 0.$$

But $[-y|v|, y|v|]$ is not in general compact. Because Y is normal, we can only say that the order interval $[-y\|v\|, y\|v\|]$ is bounded and so by Alaoglu's Theorem is w^*-compact, if Y is a dual space. But it is far from being compact for any other stronger topology and in particular for the norm topology which is of special interest to us. After all, compactness in the norm topology is a quite restrictive requirement because it implies the set under consideration has empty interior.

Therefore our definition, although more naive, appears to be somewhat more general than that of Thibault. However, we should point out that our definition translated in the language of Thibault's definition requires that the perturbation function r be identically zero.

Furthermore, our definition ties better with the convex case, because as we proved in [**19**] (Theorem 3.2), every convex mapping is locally o-Lipschitz in the interior of its domain.

Also the generalized directional derivative that Thibault defines is not, in general, continuous, while ours is even more regular since it is Lipschitz (see Proposition 3.1). Then using his generalized directional derivative, Thibault introduces a subdifferential multioperator, which, however, is not guaranteed to be everywhere nonempty. On the contrary, our subdifferential operator $\partial(x)$ is nonempty for all $x \in X$ (see Theorem 3.1). In addition, Thibault's $\partial f(x)$ is compact in $\mathfrak{L}_s(X, Y)$ under more restrictive hypotheses. Finally, our theorem about the subdifferential of an integral appears to be stronger because it does not require Y to have w compact order intervals (unless Y is reflexive, which is not necessarily the case in our result), and because we give conditions under which equality of two subdifferential multioperators holds.

In [**15**] Hirriart-Urruty and Thibault work with locally norm Lipschitz mappings and characterize the plenary hull of Clarke's generalized Jacobian. Their approach is quite different from ours. However, in connection with this paper, we should point out that it wouldn't have been a good idea to work with the family of locally norm Lipschitz mappings because their definition ignores the lattice structure of the range space. So we wouldn't have been able to give a direct definition of the generalized directional derivative, but only a weak one (i.e. one using the weak topology) as in [**15**]. This, however, is not appropriate for developing a general subdifferential theory close to the real-valued prototype.

Finally, in closing, we should mention Ioffe's very recent work [**17**]. Although considering vector valued mappings, he uses a quite different approach and emphasis. He presents a new outlook on nonsmooth analysis that goes outside Clarke's Theory. He is only considering locally norm Lipschitz mappings, and all analytical objects that he defines are with respect to the weak topology.

Acknowledgement. I would like to thank the anonymous referee for his valuable remarks that improved the presentation of the material in this paper considerably.

REFERENCES

[1] J. P. Aubin, *Gradients généralisés de Clarke*, Ann. Sci. Math. Québec, **2** (1978), 197–252.

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

[3] _____, *A new approach to Lagrange multipliers*, Math. Oper. Res., **1** (1976), 167–174.

[4] _____, *Nonsmooth Analysis and Optimization*, Proceedings of the International Congress of Mathematicians, Helsinki, 1978.

[5] _____, *Euler-Lagrange differential inclusion*, J. Differential Equations, **19** (1975), 80–90.

[6] _____, *On the inverse function theorem*, Pacific J. Math., **64** (1976), 97–101.

[7] _____, *Generalized gradients of Lipschitz functionals*, Adv. in Math., **40** (1981), 52–67.

[8] N. Dunford and J. Schwartz, *Linear Operators*, Vol. 1, Wiley Interscience, New York, 1957.

[9] H. Halkin, *Optimization Without Differentiability*, Lectures Notes in Math., Vol. 680, Springer 1977.

[10] F. Hille and R. Phillips, *Functional Analysis and Semigroups*, Colloq. Publ. Amer. Math. Soc. 1957.

[11] J. B. Hiriart-Urruty. *Conditions necessaires d'optimalité en programmation nondifferentiable*, C.R.A.S., **t284** (1976), 843–845.

[12] _____, *Sur les cônes tangents et leurs applications en programmation nondifferentiable*, C.R.A.S., **t285** (1977), 1381–1384.

[13] J. B. Hiriart-Urruty. *Gradients géneralisés de fonctions composées applications*, C.R.A.S., **t285** (1977), 781–784.

[14] _____, *Tangent cones, generalized gradients and mathematical programming in Banach spaces*, Math. Oper. Res., **4** (1979), 63–82.

[15] J. B. Hiriart-Urruty and L. Thibault, *Existence et caractérisation de differantielles généralisées des applications localement Lipschitziennes d'un espace de Banach séparable dans un espace de Banach réflexif séparable*, C.R.A.S., **t290** (1980), 1081–1094.

[16] A. Ioffe, *Gradients généralisées d'une application localement Lipschitzienne d'un espace de Banach dans un autre*, C.R.A.S., **t289** (1979), 637–640.
[17] _____, *Nonsmooth analysis: differential calculus of nondifferentiable mappings*, Trans. Amer. Math. Soc., **266** (1976), 1428–1431.
[19] N. S. Papageorgiou, *Nonsmooth analysis on partially ordered vector spaces. Part 1: Convex case*, Pacific J. Math., **107** (2) (1983), 403–458.
[20] _____, *On the subdifferential calculus of vector valued mappings with applications in optimization*, Presented in the 11th Inter. Sympos. on Math. Progr. (1982).
[21] A. Peressini, *Ordered Topological Vector Spaces*, Harper and Row, New York 1967.
[22] R. T. Rockafellar, *Clarke's tangent cones and the boundaries of closed sets in R^n*, Nonl. Anal. th. Meth. Appl., **3** (1979), 145–154.
[23] J. Saint-Pierre, *Une extension du théorem de Strassen*, C.R.A.S., **t279** (1974), 5–8.
[24] H. Schaefer, *Banach Lattices and Positive Operators*, Springer, 1974.
[25] L. Thibault, *Subdifferentials of compactly Lipschitzian vector valued functions*, Ann. Math. Pur. Appl.
[26] M. Valadier, *Sous differentiabilité de fonctions convexes a valeurs dans un espace vectoriel ordonne*, Math. Scand., **30** (1972), 65–74.
[27] J. Warga, *Derivate containers, inverse functions, and controllability*, Calculus of Variations and Control Theory (ed. J. Russel) M.R.C. Univ. of Wisconsin, Madison, Academic Press, New York, 1976.

Received October 30, 1981 and in revised form March 19, 1982. Research partially supported by the U.S. Air Force Office of Scientific Research, Grant 77-3177.

HARVARD UNIVERSITY
CAMBRIDGE, MA 02138