

## CLARKE'S GRADIENTS AND EPSILON-SUBGRADIENTS IN BANACH SPACES

BY

JAY S. TREIMAN<sup>1</sup>

**ABSTRACT.** A new characterization of Clarke's normal cone to a closed set in a Banach space is given. The normal cone is characterized in terms of weak-star limits of epsilon normals. A similar characterization of Clarke's generalized gradients is also presented. Restrictions must be placed on the Banach spaces to make the formulas valid.

**1. Introduction.** The main purpose of this paper is to extend the characterizations of Clarke's normal cone in terms of proximal normals and Clarke's generalized gradients in terms of proximal subgradients to Banach spaces. These normals and generalized gradients have found many uses in control theory [1, 3], differential inclusions [3, 4], mathematical programming [9, 16], and other fields. The recent book by F. H. Clarke [4] gives an excellent survey and has an extensive list of references. Much of this work has been done using either Clarke's tangent cone and directional derivatives or the properties of Lipschitz functions.

Many of the sharpest results in finite dimensions require the characterizations of the normal cone and generalized gradients in terms of proximal normals and subgradients. Both of these characterizations require the existence of closest points in a closed set  $C \subset \mathbf{R}^n$  to points in the complement of  $C$ . Since closest points may not exist if  $C$  is in an infinite dimensional space, something different is required. Before defining these normals and subgradients, we review the definitions of Clarke's tangent and normal cones and Clarke's generalized gradients.

Let  $E$  be a Banach space,  $C$  a closed subset of  $E$ , and  $x$  a member of  $C$ . The *Clarke tangent cone*,  $T_C(x)$ , to  $C$  at  $x$  is the set of all  $y$  in  $E$  such that for all  $\varepsilon > 0$  there are  $\lambda, \delta > 0$  with the drops

$$(1.1) \quad x' + (0, \lambda] \cdot B(y, \varepsilon)$$

intersecting  $C$  for all  $x' \in C \cap B(x, \delta)$  and  $\lambda' \in (0, \lambda]$ . Here  $B(y, \varepsilon)$  is the open ball centered at  $y$  with radius  $\varepsilon$  and

$$(0, \lambda] \cdot V := \{tv: t \in (0, \lambda], v \in V\}.$$

This tangent cone is always nonempty, closed and convex.

---

Received by the editors, December 10, 1984 and, in revised form, March 25, 1985.

1980 *Mathematics Subject Classification*. Primary 26B05, 26B25; Secondary 46B20.

*Key words and phrases*. Nondifferentiable functions, generalized gradients, tangent cones, normal cones.

<sup>1</sup> Research supported in part by the United States Air Force Office of Scientific Research, U.S.A.F., under grant number F49620-82-K-0012.

Since  $T_C(x)$  is a closed convex cone, it has a polar, the (Clarke) normal cone,  $N_C(x)$ , such that

$$N_C^\circ(x) = T_C(x) = N_C^{\circ\circ}(x).$$

The normal cone is a nonempty weak\* closed convex cone.

The normal cone and the generalized gradients of a lower semicontinuous (lsc) function  $f$  are related through the *epigraph* of  $f$ ,

$$\text{epi } f := \{(x, \alpha) \in E \times \mathbf{R} : \alpha \geq f(x)\}.$$

The *generalized gradients* of a lsc  $f$  at  $x$  are given by

$$\partial f(x) := \{v^* \in E^* : (v^*, -1) \in N_{\text{epi } f}(x, f(x))\}.$$

From the definition it is clear that  $\partial f(x)$  is a weak\* closed convex set.

As was mentioned earlier, there is a characterization of generalized gradients in terms of proximal subgradients if  $E = \mathbf{R}^n$ . A *proximal subgradient* to  $f$  at  $\bar{x}$  is a  $v^* \in E^*$  such that for some  $r > 0$

$$f(x) - \langle v^*, x - \bar{x} \rangle + r|x - \bar{x}|^2$$

has a local minimum at  $\bar{x}$ . This is equivalent to saying that, for some  $r > 0$ ,  $f$  is supported below by

$$f(\bar{x}) + \langle v^*, x - \bar{x} \rangle - r|x - \bar{x}|^2$$

in a neighborhood of  $x$ . It can be shown [15] that if  $f: \mathbf{R}^n \rightarrow \mathbf{R}$  is lsc, then

$$(1.2) \quad \partial f(x) = \text{cl co}\{\hat{\partial} f(x) + \hat{\partial}^\infty f(x)\},$$

where

$$\hat{\partial} f(x) := \left\{ v^* : \exists v^{*k} \rightarrow v^* \text{ with } v^{*k} \text{ a proximal subgradient to } f \text{ at } x^k \xrightarrow{f} x \right\},$$

and

$$\hat{\partial}^\infty f(x) := \left\{ v^* : \exists v^{*k} \rightarrow v^* \text{ with } v^{*k} \text{ a proximal subgradient to } \tau_k f \right. \\ \left. \text{at } x^k \xrightarrow{f} x \text{ and } \tau_k \searrow 0 \right\}.$$

The set  $\hat{\partial}^\infty f(x)$  can be interpreted as the limit points at infinity of sequences of proximal subgradients.

The definition of the normal cone in terms of proximal normals is somewhat simpler. A *proximal normal* to a closed  $C \subset \mathbf{R}^n$  at  $x$  is a  $v$  in  $\mathbf{R}^n$  such that  $x$  is the closest point in  $C$  to  $x + tv$  for some  $t > 0$ . The normal cone to  $C$  at  $x$  is given by

$$(1.3) \quad N_C(x) = \text{cl co}\left\{ v : \exists v^k \rightarrow v \text{ with } v^k \text{ a proximal normal to } C \text{ at } x^k \xrightarrow{C} x \right\}.$$

The original proofs of (1.2) and (1.3) rely on Rademacher's theorem to construct proximal normals and proximal subgradients. Using the proof of (1.3) in the paper of Borwien and Strojwas [2] or the proof in [18], one can avoid Rademacher's theorem and extend the result to Hilbert spaces. See the work of Thibault [17] for another extension of (1.2).

In many infinite dimensional spaces it appears that (1.2) and (1.3) may not hold because of the structure of the spaces. The work of Borwein and Strojwas [2] depends on the specific norm on the Banach space. In this paper generalizations of (1.2) and (1.3) are given that are independent of the norm on the Banach space. These results depend on the geometric structure of certain spaces and Ekeland's variational principle [7]. First, Clarke's normal cone is characterized in terms of  $\varepsilon$ -normals. Then a characterization of Clarke's generalized gradients in terms of  $\varepsilon$ -subgradients is given. Throughout, the relationships between this work and other recent results are explored. These results were originally in the unpublished work [19].

**2. Local normals.** Let  $E$  be a Banach space with open unit ball  $B$  and dual  $E^*$ . In this section two types of normals will be defined. These normals correspond to the Fréchet derivative and to the Hadamard derivative. Recall that a function  $f$  on  $E$  has Fréchet (Hadamard) derivative  $v^*$  at  $x$  if

$$\lim_{t \searrow 0} \left| \frac{f(x + th) - f(x) - \langle v^*, th \rangle}{t} \right| \rightarrow 0$$

uniformly on  $B$  (uniformly on compacta). These concepts are covered in detail in Flett [8].

In the following  $\langle \cdot, \cdot \rangle$  denotes the inner product on  $\mathbf{R}^n$  and the canonical pairing between  $E^*$  and  $E$  otherwise.

**DEFINITION 2.1.** Let  $C$  be a closed subset of  $E$ ,  $x \in C$ ,  $\varepsilon > 0$  and  $v^* \in E^*$ . Then  $v^*$  is a *Fréchet  $\varepsilon$ -normal* ( $F_\varepsilon$ -normal) to  $C$  at  $x$  if there is a neighborhood  $X$  of  $x$  such that  $0 \geq \langle v^*, x' - x \rangle - \varepsilon |x' - x|$  for all  $x' \in X \cap C$ .

**REMARK.** (1) In both of the definitions of normals it is not assumed that the normals are bounded away from the origin since 0 is in the polar of any subcone of  $E$ , and it is desired to characterize the polar of  $T_C(x)$ .

(2) This definition is almost identical to that of Kruger and Mordukhovich [12] but is more suited to the purposes of this paper.

In order to show that, for any  $\varepsilon > 0$ , the points where there are  $F_\varepsilon$ -normals of norm one is dense in the boundary of  $C$ , the existence of a Fréchet differentiable bump function is required; see Ioffe [11]. Here it will be assumed that there is an equivalent norm that is Fréchet differentiable off 0.

The other type of normal used in this work corresponds to the Hadamard derivative. In Definition 2.1 neighborhoods are used. The concept of a neighborhood is not broad enough for the second type of normal. Thus the following definition is required. In this definition the singular point 0 is ignored.

**DEFINITION 2.2.** A set  $S$  is a *sponge* around  $x$  if for every  $h \in E \setminus \{0\}$ , there is a neighborhood  $H$  of  $h$  and a  $\lambda > 0$  such that  $x + [0, \lambda] \cdot H \subset S$ .

**EXAMPLE 2.3.** Let  $f: E \rightarrow \mathbf{R}$  be Lipschitz around  $x$  with Gateaux derivative  $v^*$  at  $x$ . The set of  $z \in E$  where  $f(z) \geq f(x) + \langle v^*, z - x \rangle - \varepsilon |z - x|$  is a sponge around  $x$  for any  $\varepsilon > 0$ . Similarly, the set of  $z \in E$  where  $f(z) \leq f(x) + \langle v^*, z - x \rangle + \varepsilon |z - x|$  is a sponge around  $x$  for any  $\varepsilon > 0$ . This fact will be used in the proofs of the main theorems in this paper.

EXAMPLE 2.4. In the next section the “typical” sponge that is used can be described as follows. Let  $A_1$  be

$$A_1 := \{z: 0 < \langle v^*, z \rangle - \varepsilon|z|\}.$$

This is the complement of the cone  $\{z: 0 \geq \langle v^*, z \rangle - \varepsilon|z|\}$ . Let  $X$  be the set of all  $x$  in the interior of the cone  $\{z: 0 > \langle v^*, z \rangle - 2\varepsilon|z|\}$  with  $|x| = 1$ , and for each  $x$  let  $N_x$  be a neighborhood of  $x$ , and let  $\lambda_x > 0$  be a scalar. The set  $A_1 \cup \bigcup_{x \in X} (0, \lambda_x] \cdot N_x$  is a sponge around 0.

A sponge around  $x$  is a neighborhood of  $x$  if either  $S$  is convex or  $E$  is finite dimensional. It is easy to find sponges around  $x$  that are not neighborhoods of  $x$ . Dolecki [6] defines a similar object; however, in the usual topology his absorbing sets must be neighborhoods.

With this the second type of normal can be defined.

DEFINITION 2.5. Let  $C$  be a closed subset of  $E$ ,  $x \in C$  and  $\varepsilon > 0$ . A  $v^* \in E^*$  is a *Hadamard  $\varepsilon$ -normal* ( $H_\varepsilon$ -normal) to  $C$  at  $x$  if there is a sponge  $S$  around  $x$  such that  $0 \geq \langle v^*, x' - x \rangle - \varepsilon|x' - x|$  for all  $x' \in S \cap C$ .

The following example shows there are situations where  $H_\varepsilon$ -normals of norm 1 exist at every point on the boundary of  $C$ , but the points where there are  $F_\varepsilon$ -normals of norm 1 are not even dense on  $\text{bdry } C$  for some  $\varepsilon > 0$ .

EXAMPLE 2.6. Let  $E$  be a Banach space with a norm that is Gateaux differentiable on  $E \setminus \{0\}$  but is nowhere Fréchet differentiable; see [13]. Let  $C = E \setminus B$ . For any  $\varepsilon > 0$ , define

$$A_\varepsilon := \{x \in \text{bdry } C: \exists \text{ an } F_\varepsilon\text{-normal to } C \text{ at } x \text{ of norm } 1\}.$$

Then  $A_\varepsilon \subset \text{rel int } A_{2\varepsilon}$ . Here the relative interior is with respect to the boundary of  $C$ . This implies that the set

$$A := \bigcap_{\varepsilon > 0} A_\varepsilon = \{x \in \text{bdry } C: \text{the norm is Fréchet differentiable at } x\}$$

is a  $G_\delta$  relative to the boundary of  $C$ . Since  $A = \emptyset$ , there is an  $\varepsilon_0 > 0$  such that

$$A' = \text{rel int}(\text{bdry } C \setminus A_{\varepsilon_0}) \neq \emptyset.$$

For all  $x \in A'$  there are no  $F_{\varepsilon_0}$ -normals of norm 1 to  $C$  at  $x$ . However, if  $v^*$  is the Gateaux derivative of the norm function at  $x \in \text{bdry } C$ , then  $v^*$  is an  $H_\varepsilon$ -normal to  $C$  at  $x$  for all  $\varepsilon > 0$ .

**3. The normal cones.** In this section three normal cones are defined, using the  $\varepsilon$ -normals of the previous section. They will be related to Clarke's normal cone in the next section. The first two cones have similar definitions involving the following set:

$$\hat{N}(C, x) := \left\{ v^*: \exists v^{*k} \xrightarrow{w^*} v^*, x^k \xrightarrow{C} x \text{ and } \varepsilon_k \searrow 0 \right. \\ \left. \text{with } v^{*k} \text{ an } F_{\varepsilon_k}\text{-normal (} H_{\varepsilon_k}\text{-normal) to } C \text{ at } x^k \right\}.$$

We will denote  $\hat{N}(C, x)$  by  $\hat{N}_F(C, x)$  in the case of  $F_\varepsilon$ -normals and by  $\hat{N}_H(C, x)$  in the Hadamard case.

DEFINITION 3.1. Let  $C$  be a closed subset of a Banach space  $E$  and  $x \in C$ .

(a) The *Fréchet normal cone* to  $C$  at  $x$  is given by

$$N_F(C, x) = \text{cl}^* \text{co } \hat{N}_F(C, x).$$

(b) The *Hadamard normal cone* to  $C$  at  $x$  is given by

$$N_H(C, x) = \text{cl}^* \text{co } \hat{N}_H(C, x).$$

Example 2.6 shows that the cones in Definition 3.1 can be different. It is always true that  $N_F(C, x) \subset N_H(C, x)$ . This follows from the fact that if  $v^*$  is an  $F_\varepsilon$ -normal to  $C$  at  $x$  then  $v^*$  is an  $H_\varepsilon$ -normal to  $C$  at  $x$ .

Counterexample 3.1 in [18] shows that there are sets in any infinite dimensional Banach space such that  $N_H(C, x) \neq N_C(x)$ . In that example  $N_H(C, x) = E$ , whereas  $N_C(x) = \{v^*: \langle v^*, y \rangle \leq 0\}$  for some nonzero  $y \in E$ . The problem is in the lack of uniformity of the sponges used in the definition of  $\hat{N}_H(C, x)$ . The following definition will eliminate this difficulty.

DEFINITION 3.2. Let  $y \in E$ . A sequence  $S_m$  of sponges around  $x^m$  for each  $m$  is *uniform in  $y$*  if there exists an  $M > 0$ , a sequence  $\lambda_m > 0$  and a neighborhood  $Y$  of  $y$  such that for all  $m > M$ ,

$$x^m + [0, \lambda_m] \cdot Y \subset S_m.$$

DEFINITION 3.3. Let  $v^{*k}$  be a sequence of  $H_{\varepsilon_k}$ -normals to  $C$  at  $x^k$  with corresponding sponges  $S_k$ . The sequence  $v^{*k}$  is a *uniform* sequence of  $H_{\varepsilon_k}$ -normals to  $C$  if  $v^{*k} \xrightarrow{w^*} v^*$  for some  $v^* \in E^*$ ,  $\varepsilon_k \searrow 0$ , and the sponges are uniform in  $y$  for all  $y$  with  $\langle v^*, y \rangle > 0$ .

DEFINITION 3.4. Let  $C$  be a closed subset of  $E$ ,  $x \in C$ , and let

$$\hat{N}_U(C, x) := \left\{ v^*: \exists \text{ a uniform sequence } v^{*k} \text{ of } H_{\varepsilon_k}\text{-normals to } C \text{ at } x^k \xrightarrow{C} x \right. \\ \left. \text{with } v^{*k} \xrightarrow{w^*} v^*, \varepsilon_k \searrow 0 \right\}.$$

The *uniform normal cone* to  $C$  at  $x$  is

$$N_U(C, x) := \text{cl}^* \text{co } \hat{N}_U(C, x).$$

We have been calling the objects defined in this section cones. The following lemma is easy to prove.

LEMMA 3.5. *All of the sets  $\hat{N}_F(C, x)$ ,  $\hat{N}_H(C, x)$ ,  $\hat{N}_U(C, x)$ ,  $N_F(C, x)$ ,  $N_H(C, x)$  and  $N_U(C, x)$  are cones.*

**4. Normal cone theorem.** In this section a new characterization of Clarke's normal cone in Banach spaces is provided. The proof follows the lines of the proof of Theorem 4.1 in [18]. In the proof of Theorem 4.2 the following result is used.

THEOREM 4.1 [7]. *Let  $E$  be a Banach space, and let  $f$  be a lsc function from  $E$  to  $(-\infty, +\infty]$  that is bounded from below. Let  $\varepsilon > 0$  and suppose  $z$  satisfies  $f(z) \leq \inf f + \varepsilon$ . Then for all  $\lambda > 0$  there is an  $x$  such that*

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

and

$$\forall x' \neq x, \quad f(x') > f(x) - (\varepsilon/\lambda)|x' - x|.$$

The main result of this section is Theorem 4.2. With a slight modification of terminology, Theorem 4.2 shows that if  $E$  has an equivalent norm that is Fréchet differentiable off 0 (e.g.  $E$  is reflexive), then the weak\* closed convex hull of Kruger and Mordukhovich's normal cone [12] is the Clarke normal cone.

In the following theorem we require that the Banach spaces have equivalent norms with certain differentiability properties. Recall that all reflexive spaces have an equivalent norm that is Fréchet differentiable off 0, and any separable space has an equivalent norm that is Gateaux differentiable off 0. A discussion of these facts is contained in §5.5 of Diestel [5].

**THEOREM 4.2.** *If  $E$  is a Banach space with an equivalent norm that is Gateaux differentiable off 0 and  $C$  is a closed subset of  $E$ , then for all  $x$  in  $C$ ,*

$$(4.1) \quad N_U(C, x) = N_C(x) \subset N_H(C, x).$$

*If, in addition,  $E$  has an equivalent norm that is Fréchet differentiable off 0, then  $N_F(C, x) = N_C(x)$ . If  $E$  is any Banach space, then  $N_F(C, x) \subset N_U(C, x) \subset N_C(x)$ .*

The next lemma is required to prove Theorem 4.2. It deals with a very special case.

**LEMMA 4.3.** *Let  $E$  be a Banach space of dimension greater than 1 with a norm that is Gateaux (Fréchet) differentiable off 0. Let  $x, y \in E$ ,  $|y| = 1$ ,  $\rho \in (0, 1)$  and  $\beta' > \beta > 0$ . If  $w$  is such that  $|w - (x + \beta y)| = \beta\rho$  and  $w \notin D = x + (\beta, \beta') \cdot B(y, \rho)$ , then  $\langle -w^*, y \rangle \geq \rho$ , where  $w^*$  is the Gateaux (Fréchet) derivative of the norm at  $w - (x + \beta y)$ .*

**PROOF.** Take  $\beta = 1$  and  $x = 0$ . If  $\xi = \langle -w^*, y \rangle$ , the hyperplane  $H := \{z: \langle w^*, z - w \rangle = 0\}$  intersects the line  $\{y - ty: t \in \mathbf{R}\}$  at  $y - \rho y/\xi$ . The choice of  $w^*$  as the Gateaux (Fréchet) derivative of the norm at  $w - y$  guarantees that  $H$  is the unique supporting hyperplane to  $B(y, \rho)$  at  $w$  and that  $\langle w^*, w - y \rangle = |w - y|$ . Since  $D$  is a convex set containing  $B(y, \rho)$ ,  $H$  is tangent to  $D$ . From these facts it is elementary to show that  $1 - \rho/\xi \geq 0$  and, hence, that  $\langle -w^*, y \rangle \geq \rho$ . (See Figure 1.)

□

**PROOF OF THEOREM 4.2.** First, the inclusion  $N_C(x) \subset N_U(C, x)$  is proven by showing that  $N_U^\circ(C, x) \subset N_C^\circ(x) = T_C(x)$ . If  $T_C(x) = E$ , we are done. Thus we assume that  $T_C(x) \neq E$ . The case of  $\dim E = 1$  is left to the reader. Assume that the dimension of  $E$  is greater than 1.

Let  $x \in C$  and let  $y \in E \setminus T_C(x)$  be of norm 1. We will construct a  $v^*$  in  $N_U(C, x)$  such that  $\langle v^*, y \rangle > 0$ . The contrapositive of (1.1) gives a  $\rho_0 > 0$ , a sequence  $z^k \xrightarrow{C} x$  and a  $\beta_k > 0$  such that

$$C \cap \{z^k + (0, \beta_k) \cdot B(y, \rho_0)\} = \emptyset.$$

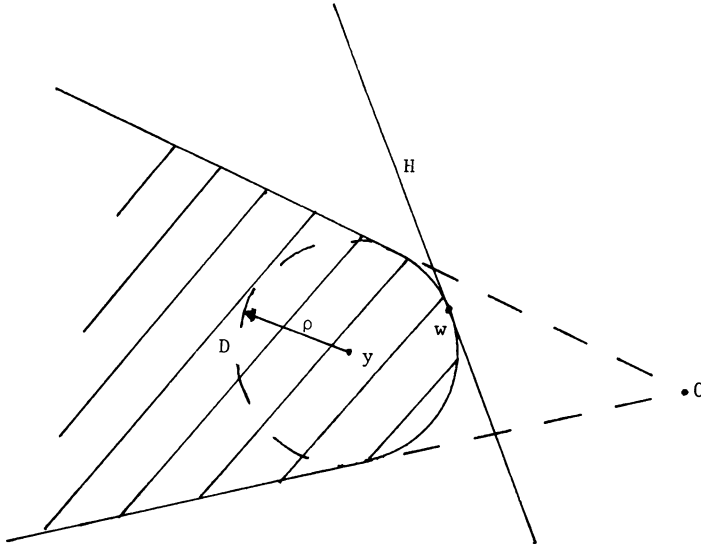


FIGURE 1

Fix  $k$  and let

$$\rho_1 = \sup \{ \rho > 0 : \exists \beta' > 0 \text{ with } C \cap [z^k + (0, \beta') \cdot B(y, \rho)] = \emptyset \} \geq \rho_0,$$

and let  $\varepsilon_k \in (0, 1/k)$ .

If  $\rho_1 = 1$  and  $v^{*k}$  is the derivative of the norm at  $-y$ , then it is elementary to show that  $v^{*k}$  is an  $H_{\varepsilon_k}$ -normal to  $C$  at  $x^k = z^k$  with  $\langle v^{*k}, y \rangle = 1 \geq \rho_0$ .

Assume that  $\rho_1 < 1$  and take  $\rho_2$  and  $\rho_3$  arbitrarily close to  $\rho_1$ ,  $\rho_0 \leq \rho_2 \leq \rho_1 < \rho_3 < 1$ , and such that for some  $\beta_2 > 0$ ,  $A := z^k + (0, \beta_2] \cdot B(y, \rho_2)$  does not intersect  $C$ . The proper choice of  $\rho_3$  guarantees that there is a  $\beta_3 \leq \beta_2$  such that

$$B = C \cap \{ z^k + \beta_3 B(y, \rho_3) \} \neq \emptyset.$$

Take  $\beta_3$  so that  $\beta_3 \leq \beta_2$  and  $\delta = (1/\beta_3)(\sup_{z' \in B} \langle z^k + \beta_3 y, z' \rangle) - \rho_2$  is arbitrarily close to 0. This can be done by adjusting  $\rho_2$  and  $\rho_3$ .

Define

$$\phi(z) := \min_{\tau \in [\beta_3, \beta_2]} \frac{1}{\tau} |z - (z^k + \tau y)|.$$

The level sets of  $\phi$  are the boundaries of sets similar to  $D$  in Lemma 4.3. This function is continuous on  $C$  and bounded below by  $\rho_2$ . If  $z' \in B$  then

$$\phi(z') \leq \inf_{z \in C} \phi(z) + \delta.$$

Theorem 4.1 can be applied to any such  $z'$  and any  $\lambda > 0$  to yield an  $x^k \in C$  such that

$$\phi(x^k) \leq \phi(z'), \quad |x^k - z'| \leq \lambda$$

and

$$(4.2) \quad \forall z \in C \quad \phi(z) \geq \phi(x^k) - (\delta/\lambda)|z - x^k|.$$

Take  $\tau_0$  with

$$(4.3) \quad \phi(x^k) = (1/\tau_0) |x^k - (z^k + \tau_0 y)|.$$

Then (4.2) and (4.3) combined give

$$\frac{1}{\tau_0} |z - (z^k + \tau_0 y)| \geq \frac{1}{\tau_0} |x^k - (z^k + \tau_0 y)| - \frac{\delta}{\lambda} |z - x^k|$$

or

$$(4.4) \quad |z - (z^k + \tau_0 y)| \geq |x^k - (z^k + \tau_0 y)| - (\delta\tau_0/\lambda) |z - x^k|.$$

If  $-v^{*k}$  is the Gateaux derivative of  $z \mapsto |z - (z^k + \tau_0 y)|$  at  $x^k$ , then for any  $\theta > 0$  there is a sponge  $N_k$  around  $x^k$  so that, if  $z \in N_k$ ,

$$(4.5) \quad |z - (z^k + \tau_0 y)| \leq |x^k - (z^k + \tau_0 y)| + \langle -v^{*k}, z - x^k \rangle + \theta |z - x^k|.$$

Combining (4.4) and (4.5) yields

$$0 > \langle v^{*k}, z - x^k \rangle - (\theta + \delta\tau_0/\lambda) |z - x^k|$$

on  $N_k \cap C$ . By choosing  $\theta$ ,  $\delta$  and  $\lambda$  appropriately, one has

$$|x^k - z^k| < 1/k, \quad \theta + \delta\tau_0/\lambda < \varepsilon_k \quad \text{and} \quad \tau_0 < \beta_2,$$

so that  $v^{*k}$  satisfies the conclusions of Lemma 4.3. Thus  $x^k \xrightarrow{C} x$ ,  $\varepsilon_k \searrow 0$ , and by the theorem of Hagler and Sullivan [10] there is a subsequence  $\{v^{*j}\}$  of  $\{v^{*k}\}$  converging weak\* to some  $v^*$ ,  $|v^*| \leq 1$  with  $\langle v^*, y \rangle \geq \rho_0$ .

The uniformity of the sequence of sponges follows easily from the facts that  $v^{*j} \xrightarrow{w^*} v^*$ , and the norm is a Lipschitz function (see Lemma 6.2). Thus  $v^* \in N_U(C, x)$ .

This demonstrates that  $N_U^\circ(C, x) \subset T_C(x)$  or  $N_U(C, x) \supset N_C(x)$ .

We now show that if  $E$  is any Banach space, then  $N_U(C, x) \subset N_C(x)$ . Because  $N_C(x)$  is a weak\* closed convex set, it is only necessary to show that  $\hat{N}_U(C, x) \subset N_C(x)$ .

For all  $v^* \in \hat{N}_U(C, x)$  we will show that  $y \notin T_C(x)$  for any  $y$  with  $\langle y, v^* \rangle > 0$ . This demonstrates that  $v^* \in N_C(x)$  and, hence, that  $\hat{N}_U(C, x) \subset N_C(x)$ .

Let  $v^* \in \hat{N}_U(C, x)$ , and let  $v^{*k} \xrightarrow{w^*} v^*$  be a uniform sequence of  $H_{\varepsilon_k}$ -normals to  $C$  at  $x^k \xrightarrow{C} x$  with  $\varepsilon_k \searrow 0$ . Let  $y \in E \setminus \{0\}$  be such that  $\langle v^*, y \rangle > \nu > 0$ .

The definition of a uniform sequence of  $H_{\varepsilon_k}$ -normals guarantees that there is an  $\eta$ ,  $0 < \eta < \nu$ , such that  $x^k + [0, \tau_k] \cdot B(y, \eta) \subset S_k$  for all  $k$  greater than some  $K$ . There is a  $K_1 > K$  such that if  $k > K_1$  then  $\langle v^{*k}, y \rangle > 0$  and  $\eta + \varepsilon_k < \langle v^{*k}, y \rangle$ .

From the definition of a uniform sequence of  $H_{\varepsilon_k}$ -normals and the above inequality there are a sequence  $\tau_k > 0$  and an  $M$  such that if  $k > M$  then

$$[x^k + (0, \tau_k) \cdot B(y, \eta)] \cap C = \emptyset.$$

The contrapositive of (1.1) shows that  $y \notin T_C(x)$ . Hence  $v^* \in N_C(x)$ , and thus  $\hat{N}_U(C, x) \subset N_C(x)$ . This complete the proof.  $\square$

**5. Subgradients, from finite to infinite dimensions.** In the previous sections a formula for the normal cone in terms of  $\varepsilon$ -normals was derived. In the finite



dimensional case Rockafellar [15] translated proximal normals into proximal subgradients. This is not as easy in Banach spaces. Instead we follow the lines of the proof of Theorem 4.2.

Recall that the generalized gradient of  $f$  at  $x$  is given by

$$\partial f(x) := \{v^*: (v^*, -1) \in N_{\text{epi } f}(x, f(x))\}.$$

What will be done is to find for each  $(y, \alpha)$  not in  $T_{\text{epi } f}(x, f(x))$  a sequence of  $\varepsilon$ -subgradients  $v^{*k}$  to  $f$  at  $x^k \xrightarrow{f} x$  such that  $\langle (v^{*k}, -1), (y, \alpha) \rangle$  is bounded away from zero. Taking the weak\* closed convex hull of weak\* limits of these sequences will yield Clarke's generalized gradients.

As with the  $\varepsilon$ -normals, there are two types of  $\varepsilon$ -subgradients to consider; those corresponding to the Fréchet derivative and those corresponding to the Hadamard derivative.

**DEFINITION 5.1.** Let  $f$  be an lsc function on a Banach space  $E$ . A  $v^* \in E^*$  is an  $F_\varepsilon$ -subgradient ( $H_\varepsilon$ -subgradient) to  $f$  at  $x$  if  $f(w) \geq f(x) + \langle v^*, w - x \rangle - \varepsilon|w - x|$  on a neighborhood (sponge) around  $x$ .

These definitions of  $\varepsilon$ -subgradients are related to those of Ioffe [11]. He says that  $v^*$  is a Fréchet  $\gamma$ -subgradient to  $f$  at  $x$  if

$$\liminf_{|h| \rightarrow 0} \frac{f(z+h) - f(z) - \langle v^*, h \rangle}{|h|} \geq -\gamma,$$

and  $u^*$  is a Dini  $\gamma$ -subdifferential of  $f$  at  $x$  if

$$\langle u^*, h \rangle \leq d^-f(x; h) + \gamma|h|, \quad \forall h$$

where  $d^-f(x; h)$  is the lower Hadamard derivative (variation) of  $f$  at  $x$  in the direction  $h$ :

$$d^-f(x; h) := \liminf_{\substack{t \searrow 0 \\ h' \rightarrow h}} \frac{f(x + th') - f(x)}{t}.$$

It is simple to show that if  $v^*$  is a Fréchet (Dini)  $\gamma$ -subdifferential of  $f$  at  $x$ , then for any  $\varepsilon > \gamma$ ,  $v^*$  is an  $F_\varepsilon$ -subgradient ( $H_\varepsilon$ -subgradient) of  $f$  at  $x$ , and that any  $F_\varepsilon$ -subgradient ( $H_\varepsilon$ -subgradient) is also a Fréchet (Dini)  $\varepsilon$ -subdifferential of  $f$  at  $x$ .

As in the characterization of the normal cone, weak\* limits of  $H_\varepsilon$ -subgradients give a set of subgradients that may be too large. To capture the uniformity required, we make the following definition.

**DEFINITION 5.2.** For each  $k$  let  $v^{*k}$  be an  $H_{\varepsilon_k}$ -subgradient to  $\tau_k f$  at  $x^k$ . The sequence  $v^{*k}$  is *uniform* for  $\{\tau_k\}_{k=1}^\infty$  if for any  $\gamma > 0$  there is a sequence  $S_k$  of sponges around  $x^k$  that is uniform in all  $y \in E \not\equiv \{0\}$  such that

$$\tau_k f(x) \geq \tau_k f(x^k) + \langle v^{*k}, w - x^k \rangle - \gamma_k |w - x^k|$$

on  $S_k$ . Here  $\gamma_k = \max\{\gamma, \varepsilon_k\}$  and  $\tau_k > 0$ .

The only cases considered here are when  $\tau_k \rightarrow 1$  or 0 and  $\varepsilon_k \searrow 0$ .

This definition guarantees that if  $v^{*k} \xrightarrow{w^*} v^*$  is a uniform sequence of  $H_{\varepsilon}$ -subgradients to  $\tau_k f$  at  $x^k \xrightarrow{f} x$  with  $\tau_k = 1$ , then Rockafellar's upper subderivative [14] is greater than or equal to  $\langle v^*, y \rangle$  for all  $y \in E$ . A similar statement holds for  $\tau_k \searrow 0$ .

The definitions of our generalized gradients can now be stated.

DEFINITION 5.3. Let

$$\hat{\partial}_F f(x) := \left\{ v^*: \exists F_{\varepsilon_k}\text{-subgradients } v^{*k} \xrightarrow{w^*} v^*, \varepsilon_k \searrow 0 \text{ to } f \text{ at } x^k \xrightarrow{f} x \right\}$$

and

$$\hat{\partial}_F^\infty f(x) := \left\{ v^*: \exists F_{\varepsilon_k}\text{-subgradients } v^{*k} \xrightarrow{w^*} v^*, \varepsilon_k \searrow 0, \right. \\ \left. \text{to } \tau_k f \text{ at } x^k \xrightarrow{f} x, \text{ where } \tau_k \searrow 0 \right\}.$$

The Fréchet generalized gradient of  $f$  at  $x$  is

$$\partial_F f(x) := \text{cl}^* \text{co} [\hat{\partial}_F f(x) + \hat{\partial}_F^\infty f(x)].$$

If  $\hat{\partial}_F f(x) = \emptyset$ , we say that  $\partial_F f(x) = \emptyset$ .

DEFINITION 5.4. Let

$$\hat{\partial}_U f(x) := \left\{ v^*: \exists \text{ a uniform sequence of } H_{\varepsilon_k}\text{-subgradients } v^{*k} \xrightarrow{w^*} v^* \text{ to } f \right. \\ \left. \text{at } x^k \xrightarrow{f} x, \text{ where } \varepsilon_k \searrow 0 \right\}$$

and

$$\hat{\partial}_U^\infty f(x) := \left\{ v^*: \exists \text{ a uniform sequence of } H_{\varepsilon_k}\text{-subgradients } v^{*k} \xrightarrow{w^*} v^* \text{ to } \tau_k f \right. \\ \left. \text{at } x^k \xrightarrow{f} x, \text{ where } \varepsilon_k \searrow 0 \text{ and } \tau_k \searrow 0 \right\}.$$

The uniform generalized gradient of  $f$  at  $x$  is

$$\partial_U f(x) := \text{cl}^* \text{co} \{ \hat{\partial}_U f(x) + \hat{\partial}_U^\infty f(x) \}.$$

If  $\hat{\partial}_U f(x) = \emptyset$  we say that  $\partial_U f(x) = \emptyset$ .

The sets  $\hat{\partial}_F^\infty f(x)$  and  $\hat{\partial}_U^\infty f(x)$  are always cones, and if the points where there are  $F_\varepsilon$ -subgradients are dense in the domain of  $f$ , then  $\hat{\partial}_F^\infty f(x)$  is nonempty. This is true if  $E$  has an equivalent norm that is Fréchet differentiable. The above definitions are norm independent, since changing the norm only multiplies the  $\varepsilon$ 's by a constant.

As with the subdifferential definitions, Ioffe [11] has defined corresponding generalized gradients. His definitions involve the use of families of subspaces and

use only finite limits. Because of this, unless the function  $f$  is Lipschitz, Ioffe's generalized gradients cannot capture Clarke's generalized gradients. There are cases where Ioffe's gradients are much larger than  $\partial f$  (let  $f$  be the indicator function of the set in Counterexample 3.2 in [18]) or much smaller than  $\partial f$  (let  $f: \mathbf{R} \rightarrow \mathbf{R}$  be 0 if  $x \leq 0$  and  $-x^{1/2}$  if  $x \geq 0$ ).

**6. The gradient theorem.** In this section the main theorem concerning generalized gradients is stated and proven. Recall that all reflexive Banach spaces have an equivalent norm that is Fréchet differentiable off 0, and all separable Banach spaces have an equivalent norm that is Gateaux differentiable off 0. §5.5 of Diestel [5] includes a discussion of these facts.

**THEOREM 6.1.** *Let  $E$  be a Banach space with an equivalent norm that is Gateaux differentiable off 0. If  $f$  is a lower semicontinuous function on  $E$ , then  $\partial_U f(x) = \partial f(x)$ . If, in addition,  $E$  has an equivalent norm that is Fréchet differentiable off 0, then  $\partial_F f(x) = \partial f(x)$ . If  $E$  is any Banach space, then  $\partial_F f(x) \subset \partial_U f(x) \subset \partial f(x)$ .*

*Note.* If the uniformity condition on the sequence of  $H_{\varepsilon_k}$ -subgradients is removed and  $E$  has an equivalent norm that is Gateaux differentiable off 0, then  $\partial f(x)$  will be contained in the resulting object. This follows from the inclusion  $\partial f(x) \subset \partial_U f(x)$ .

The case of  $\partial_F f(x)$  is easier and follows the same lines as the case of  $\partial_U f(x)$ . Because of this, the theorem is proven only for  $\partial_U f(x)$ .

**PROOF.** The inclusion  $\partial f(x) \supset \partial_U f(x)$  follows from the definition of  $\partial_U f(x)$  and the contrapositive of (1.1). The proof of this follows the lines of the proof that  $N_C(x) \supset N_U(C, x)$  in §4.

The proof that  $\partial f(x) \subset \text{cl}^* \text{co} \{ \hat{\partial} f(x) + \hat{\partial}_U^\infty f(x) \}$  is more difficult. Assume that the norm on  $E$  is Gateaux differentiable off 0 by renorming if necessary.

If  $T_{\text{epi } f}(x, f(x)) = E \times \mathbf{R}$ , then  $\partial f(x) = \emptyset$ . Because  $(v^*, -1)$  is an  $H_\varepsilon$ -normal to  $\text{epi } f$  whenever  $v^*$  is an  $H_\varepsilon$ -subgradient to  $f$ , Theorem 4.2 shows that  $\partial_U f(x) = \emptyset$ . Therefore assume that  $T_{\text{epi } f}(x, f(x)) \neq E \times \mathbf{R}$ .

Let  $\bar{x}$  be a point where  $f$  is finite and  $(y, \alpha) \notin T_{\text{epi } f}(x, f(x)) \subset E \times \mathbf{R}$ . Here  $(y, \alpha) \neq (0, 0)$ . The norm on  $E \times \mathbf{R}$  used here is  $\|(z, \xi)\| = (|z|^2 + \xi^2)^{1/2}$ . This norm is Gateaux differentiable off 0 since the norm on  $E$  is.

The idea is to find a  $v^*$  in either  $\hat{\partial}_U f(x)$  or  $\hat{\partial}_U^\infty f(x)$  such that  $\langle (v^*, -1), (y, \alpha) \rangle > 0$  or  $\langle v^*, y \rangle > 0$ , respectively. We deal with the case when  $y \neq 0$  first.

By adding an appropriate linear function to  $f$ ,  $\alpha$  may be set to 0. The contrapositive of (1.1) gives a  $\rho_0 > 0$ , a sequence  $z^k \xrightarrow{f} \bar{x}$  and a sequence  $\lambda_k > 0$  such that

$$\text{epi } f \cap \{ (z^k, f(z^k)) + (0, \lambda_k) \cdot B((y, 0), \rho_0) \} = \emptyset$$

for each  $k$ .

Fix  $k$  and by adding appropriate constants take  $z^k = 0$  and  $f(z^k) = 0$ . A differentiable penalty function is added to  $f$  to restrict the possible locations of the  $\varepsilon_k$ -subgradients; its derivatives satisfy

$$(6.1) \quad \langle v^*, y \rangle \geq \beta_1 \rho_0 |v^*|$$

for some  $\beta_1 > 0$  with  $\beta_1$  independent of  $k$ . Let

$$F(z) := -\kappa \langle w^*, z \rangle + \inf_{\mu \in (\omega, \mu_0)} \begin{cases} 0 & \text{if } |z - (\mu - \omega)y| \leq \beta_2 \mu \rho_2, \\ K \left( \frac{2|z - (\mu - \omega)y| - \mu \rho_2}{2} \right)^2 & \text{if } |z - (\mu - \omega)y| \in [\beta_2 \mu \rho_2, \beta_3 \mu \rho_3], \\ +\infty & \text{otherwise.} \end{cases}$$

Here  $\rho_2 < \rho_3 < \rho_0$ ,  $\mu_0 < \omega + \lambda_k$ ,  $\omega$  is arbitrarily small and  $w^*$  is chosen so that  $|w^*| = 1$ ,  $\langle w^*, y \rangle = 1$  and  $\langle w^*, y \rangle > \sigma$  for some  $\sigma > 0$  and all  $y' \in B(y, \rho_0)$ . These constants, except  $\omega$  and  $\mu_0$ , are independent of  $k$ .

The constants can be chosen so that outside  $A := \omega y - [\bigcup_{\mu \in (\omega, \mu_2)} \mu \bar{B}(y, \rho_3)]$ ,  $f(x) + F(x) > 0$  for some  $\mu_2 \in (\omega, \mu_1)$ . Since  $f(x)$  is lsc, it is assumed that  $f$  is bounded below by  $-\delta$  on some neighborhood of  $x$ , and all of the points we deal with are in that neighborhood. Choose  $K$  so that  $F(x) > \delta$  on the boundary of  $\omega y + (\omega, \mu_1) \cdot B(y, \rho_3)$ . The derivative of  $F$  at any point in the interior of  $A$  is  $-\kappa w^* + s^*$ , where  $s^* = 0$  or  $\langle s^*/|s^*|, y \rangle < -\rho_2$ . This yields

$$\langle -F', y \rangle \geq \kappa + \rho_2 |s^*|$$

on the interior of  $A$ .

Ekeland's variational principle (Theorem 4.1) can be applied to the lsc function  $G(z) = f(z) + F(z)$ , since this function is bounded from below. Take a  $z^{k,1}$  such that

$$G(z^{k,1}) \leq \min\{G(0) = 0, \inf G(w) + 1/2k\}.$$

Theorem 4.1 yields an  $x^k$  such that  $G(x^k) \leq G(z^{k,1})$ ,

$$(6.2) \quad \forall z \quad G(z) \geq G(x^k) - (1/2k)|z - x^k|,$$

$x^k$  is in the interior of the domain of  $F$  and  $x^k \in A$ .

Since  $F$  is Gateaux differentiable at  $x^k$ ,

$$(6.3) \quad F(z) \leq F(x^k) + \langle F'(x^k), z - x^k \rangle + (1/2k)|z - x^k|$$

on some sponge  $S_k$  around  $x^k$ .

Combining (6.2), (6.3) and the bound on  $\langle F', y \rangle$  yields

$$f(x) \geq f(x^k) + \langle -F'(x^k), z - x^k \rangle - (1/k)|z - x^k|$$

on  $S_k$ , where

$$(6.4) \quad |F'(x^k)| > \kappa.$$

The result of Hagler and Sullivan [10] shows that the sequence  $v^{*k} = -F'(x^k)/|F'(x^k)|$  has a weak\* convergent subsequence  $v^{*j}$  with  $v^{*j} \xrightarrow{w^*} v^*$ . Equation

(6.4) implies that  $\langle v^*, y \rangle \geq \frac{1}{2} \min\{\kappa, \rho_2\}$ . Since  $x^k \xrightarrow{f} \bar{x}$ , this completes the construction of the  $H_{\epsilon_k}$ -subgradients.

The only thing left to show is that  $v^{*j}$  is a uniform sequence of  $H_{\epsilon_k}$ -subgradients. This will give a  $v^*$  in either  $\hat{\partial}_U f(\bar{x})$  or  $\hat{\partial}_U^\infty f(\bar{x})$  such that  $\langle (v^*, -1), (y, \alpha) \rangle > 0$  or  $\langle v^*, y \rangle > 0$ , demonstrating that  $\partial f(\bar{x}) \subset \text{cl}^* \text{co}\{\hat{\partial}_U f(\bar{x}) + \hat{\partial}_U^\infty f(\bar{x})\}$ .

Assume that  $|F'(x^j)| \rightarrow \infty$ . The functions  $F$  are Lipschitz around  $x^j$  with Lipschitz constants  $L_j = 2|F'(x^j)|$ . Lemma 6.2 shows that the sequence  $v^{*j}$  is a uniform sequence of  $H_{\epsilon_j}$ -subgradients to  $f/L_j$  at  $x^j \xrightarrow{f} x$ . Thus  $v^* \in \hat{\partial}_U^\infty f(\bar{x})$ .

Similarly, if  $|F'(x^j)|$  is bounded, there is a  $v^* \in \hat{\partial} f(\bar{x})$  with  $\langle (v^*, -1), (y, \alpha) \rangle > 0$ .

The case of  $(0, -1) \notin T_{\text{epi}f}(\bar{x}, f(\bar{x}))$  is similar to the preceding case. Instead of the  $F$  used above, one sees the function described below.

Since  $(0, -1) \notin T_{\text{epi}f}(x, f(x))$ , there exist sequences  $x^k \xrightarrow{f} x$  and  $\lambda_k > 0$  and a  $\rho_0 > 0$  such that, for each  $k$ ,

$$\text{epi} f \cap \{(x^k, f(x)) + (0, \lambda_k] \cdot \bar{B}((0, -1), \rho_0)\} = \emptyset.$$

Fix  $k$  and  $\rho_1, \rho_0 > \rho_1 > 0$ . Let  $\epsilon_k = 1/k$ . Set  $x^k = 0$  and  $f(x^k) = 0$ . Take  $\omega$  small and less than  $\lambda_k$ . Let

$$F(z) := \begin{cases} \left( (\rho_1 \omega)^2 - |z|^2 \right)^{1/2} & \text{if } |z| \leq \eta \omega, \\ \left( (\rho_1)^2 - \omega^2 \eta^2 \right)^{1/2} + \frac{\rho_1}{(1 - \rho_1^2)^{1/2}} (|z| - \omega \eta) & \text{if } |z| \in (\eta \omega, \lambda_k \rho_1], \\ +\infty, & \text{otherwise.} \end{cases}$$

Here  $\eta$  is chosen to make this function differentiable.

Adding  $f$  and  $F$  gives an lsc function that is 0 at 0 and is greater than 0 on the boundary of  $B(0, \lambda_k \rho_1)$ . In addition, the derivatives of the  $F$  are bounded. The proof now follows the same lines as the case of  $y \neq 0$ .  $\square$

The following lemma was used in the proof of Theorem 6.1. It shows that the sequences of  $H_{\epsilon_k}$ -subgradients constructed in the proof of Theorem 6.1 are uniform sequences.

**LEMMA 6.2.** *For each  $k > 0$  let  $v^{*k}$  be the Gateaux derivative of a Lipschitz function  $p_k(z)$  at a point  $x^k$ . Assume that  $v^{*k} \xrightarrow{w^*} v^*$ . If all of the  $p_k$ 's have the same Lipschitz constant  $L$ , then for all  $\delta > 0$  there is a sequence of sponges  $S_k$  around  $x^k$  that is uniform in all  $y \in E$  and, for  $z \in S_k$ ,*

$$p_k(z) \geq p_k(x^k) + \langle v^{*k}, z - x^k \rangle - \delta |z - x^k|.$$

**PROOF.** Let  $y \in E \setminus \{0\}$ . By the hypotheses there is a  $K$  so that when  $k \geq K$ ,  $|v^{*k}| \leq L$  and  $|\langle v^{*k} - v^*, y \rangle| \leq (\delta/4)|y|$ . Let  $k \geq K$ . Since  $v^{*k}$  is the Gateaux derivative of  $p_k(z)$  at  $x^k$ , there is a  $\lambda_k > 0$  so that if  $t \in [0, \lambda_k]$  then

$$p_k(x^k + ty) \geq p_k(x^k) + \langle v^{*k}, ty \rangle - \delta |ty|/4.$$

Hence

$$p_k(x^k + ty) \geq p_k(x^k) + \langle v^*, ty \rangle - \delta|ty|/2.$$

There exists a  $\gamma > 0$  such that if  $y'$  is in  $B(y, \gamma)$  then

$$p_k(x^k + ty') \geq p_k(x^k) + \langle v^*, ty' \rangle - \delta|ty'|.$$

Thus the  $S_k$ 's can be taken to absorb  $B(y, \gamma)$  for all  $k \geq K$ .  $\square$

## REFERENCES

1. J. P. Aubin and F. H. Clarke, *Shadow prices and duality for a class of optimal control problems*, SIAM J. Control Optim. **17** (1979) 567–586.
2. J. M. Borwein and H. M. Strojwas, *Proximal analysis and boundaries of closed sets in Banach spaces. Part 1: Theory*, Canad. J. Math. (to appear).
3. F. H. Clarke, *The Erdmann condition and Hamiltonian inclusions in optimal control and the calculus of variations*, Canad. J. Math. **32** (1980), 494–507.
4. ———, *Optimization an nonsmooth analysis*, Wiley, New York, 1983.
5. J. Diestel, *Geometry of Banach spaces: selected topics*, Springer-Verlag, New York, 1975.
6. S. Dolecki, *Tangency and differentiation: some applications of convergence*, Ann. Mat. Pura Appl. (4) **80** (1982), 223–255.
7. I. Ekeland, *On the variational principle*, J. Math. Anal. Appl. **47** (1974), 324–353.
8. T. M. Flett, *Differential analysis*, Cambridge Univ. Press, Cambridge, 1980.
9. A. A. Goldstein, *Optimization of Lipschitz continuous functions*, Math. Programming **13** (1977), 14–22.
10. J. Hagler and H. Sullivan, *Smoothness and weak\* sequential compactness*, Proc. Amer. Math. Soc. **78** (1980), 497–503.
11. A. D. Ioffe, *Approximate subdifferentials of nonconvex functions*, Cahiers Math Decision, no. 8120, Univ. Paris-Dauphine, 1981.
12. A. Ja. Kruger and B. Sh. Mordukhovich, *Extreme points and Euler equations in nondifferentiable optimization problems*, Dokl. Akad. Nauk SSSR **24** (1980), 684–687.
13. R. R. Phelps, *Differentiability of convex functions on Banach spaces*, lecture notes (unpublished), University College, London, 1978.
14. R. T. Rockafellar, *Generalized directional derivatives and subgradients of nonconvex functions*, Canad. J. Math. **32** (1980), 257–280.
15. ———, *Proximal subgradients, marginal values and augmented Lagrangians in nonconvex optimization*, Math. Oper. Res. **6** (1981), 424–436.
16. ———, *The theory of subgradients and its applications to problems of optimization*, Heldermann, Berlin, 1981.
17. L. Thibault, *Quelques proprietes des sous-differentiels de fonctions reelles localement Lipschitziennes definies sur un espace de Banach separable*, C. R. Acad. Sci. Paris Ser. A-B **282** (1976), A507–A510.
18. J. S. Treiman, *Characterization of Clarke's tangent and normal cones in finite and infinite dimensions*, Nonlinear Anal. **7** (1983), 771–783.
19. ———, *A new characterization of Clarke's tangent cone and its applications to subgradient analysis and optimization*, Ph.D. Thesis, Univ. of Washington, 1983.

DEPARTMENT OF MATHEMATICS, LEHIGH UNIVERSITY, BETHLEHEM, PENNSYLVANIA 18015

Current address: Department of Mathematics, Western Michigan University, Kalamazoo, Michigan 49008