

## THE PROXIMAL NORMAL FORMULA IN BANACH SPACE

J. M. BORWEIN AND J. R. GILES

**ABSTRACT.** Approximation by proximal normals to the Clarke generalized subdifferential for a distance function generated by a nonempty closed set and the normal cone to the set generated by the proximal normals are important tools in nonsmooth analysis. We give simple general versions of such formulae in infinite dimensional Banach spaces which satisfy different geometrical conditions. Our first class, of spaces with uniformly Gâteaux differentiable norm includes the Hilbert space case and the formulae is attained through dense subsets. Our second class, of reflexive Kadec smooth spaces is the most general for which such formulae can be obtained for all nonempty closed sets in the space. Our technique also allows us to establish the existence of solutions for a class of optimization problems substantially extending similar work of Ekeland and Lebourg.

**RESUME.** L'approximation par les normales proximales au sous-différentiel généralisé de Clarke pour une fonction de distance produit d'un ensemble non-vidé fermé et le cône normal à l'ensemble produit des normales proximales sont objets d'importance pour l'analyse non-régulière. Nous donnons deux versions simples et générales de telles formules dans les espaces de Banach de dimensions infinies. Premièrement, nous examinons la classe des espaces avec norme uniformément Gâteaux-dérivable qui comprend les espaces de Hilbert. Deuxièmement, nous examinons la classe espaces réflexifs Kadec et lisses. Enfin, notre méthode produit l'existence des solutions pour une classe de problèmes à l'optimisation.

**1. Introduction.** Given a locally Lipschitz function  $\phi$  on an open subset  $G$  of a normed linear space  $X$ , the *generalized subdifferential* of  $\phi$  at  $x$  is defined as

$$\partial\phi(x) \equiv \left\{ f \in X^* : f(y) \leq \limsup_{\substack{z \rightarrow x \\ t \rightarrow 0+}} \frac{\phi(z + ty) - \phi(z)}{t} \right\}.$$

When he first introduced the generalized subdifferential, Clarke gave the following definition in  $\mathbf{R}^n$ :

$\partial\phi(x)$  is the convex hull of the cluster points of the derivatives  $\nabla\phi(x_n)$  for points  $x_n$  where they exist as  $x_n \rightarrow x$  [6, Theorem 2.5.1]

(that there are plenty of such sequences  $x_n$  is guaranteed by Rademacher's Theorem in  $\mathbf{R}^n$ ).

---

Received by the editors June 20, 1986.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 49A52; Secondary 46B20.

Ce rapport a été publié en partie grâce à une subvention du Fonds FCAR pour l'aide et soutien à la recherche.

The authors wish to express their thanks to Professor Frank Clarke for providing them the opportunity to carry out this research at Centre de Recherches Mathématiques, Université de Montréal.

Given a nonempty closed set  $K$  in a normed linear space  $X$ , the *distance function*  $d_K$  generated by  $K$  is defined by

$$d_K(x) = \inf_{y \in K} \|x - y\|.$$

Associated with the generalized subdifferential of the *distance function*  $d_K$  at  $x \in \text{bdy } K$  is a geometrical object called the *normal cone* to  $K$  at  $x$  which is defined as

$$N_K(x) \equiv \left\{ \overline{\bigcup_{\lambda \geq 0} \lambda \partial d_K(x)}^{w^*} \right\},$$

the weak\* closed convex cone generated by  $\partial d_K(x)$ . In  $\mathbf{R}^n$ , the generalized subdifferential of the distance function  $d_K$  at  $x \in \text{bdy } K$  has an interesting geometrical formulation:

$\partial d_K(x)$  is the convex hull of the origin and the cluster points of  $v_n/\|v_n\|$  where  $v_n \perp K$  at points  $x_n \in \text{bdy } K$  as  $x_n \rightarrow x$  and  $\|v_n\| \rightarrow 0$  [6, Theorem 2.5.6].

The corresponding expression for the normal cone  $N_K(x)$  in terms of approximating normals to  $K$  in  $\mathbf{R}^n$  is called the *proximal normal formula*.

An infinite dimensional formulation of the proximal normal formula was given by Borwein and Strojwas in any reflexive Banach space with Kadec norm where the norm is also Fréchet differentiable [4, Theorem 3.1]. A second proof for Hilbert space using the abundant geometrical properties of that space has recently been given by Loewen [11].

Here we present two elementary proofs of the formula in Banach spaces satisfying different geometrical conditions. The first uses uniform Gâteaux differentiability of the norm and this enables the formulae to be expressed in terms of dense subsets. The second replaces this geometrical spatial condition by a Fréchet subdifferentiability condition on the distance function and gives the most general situation in which the formulae can be obtained for all nonempty closed subsets. We conclude by showing that the second approach generalizes to allow us to establish the existence of solutions for a class of optimization problems extending the work of Ekeland and Lebourg [8, Theorem 3.11].

Consider a nonempty closed set  $K$  in a normed linear space  $X$  and the distance function  $d_K$  generated by  $K$ . Denote by  $E(K)$  the set of points in  $X \setminus K$  for which there exists a closest point in  $K$ ; that is,  $x \in E(K)$  if and only if there exists an  $e(x) \in K$  such that  $\|x - e(x)\| = d_K(x)$ .

Throughout, for  $x \in X$ ,  $x \neq 0$ , we denote by  $f_x$  a subgradient of the norm at  $x$ ; that is,  $f_x(x) = \|x\|$  and  $\|f_x\| = 1$ . We say that  $X$  is *smooth* if each  $x \neq 0$  has only one subgradient  $f_x$ . We have the following preliminary property locating certain norm subgradients in the generalized subdifferential of the distance function.

**LEMMA 1.** *For any  $x \in E(K)$  and every  $e(x) \in K$  such that  $\|x - e(x)\| = d_K(x)$ , there exists an  $f_{x-e(x)} \in \partial d_K(x)$ .*

PROOF. Given  $y \in X$ ,

$$\begin{aligned} d_K^0(x)(y) &\equiv \limsup_{\substack{z \rightarrow x \\ t \rightarrow 0+}} \frac{d_K(z + ty) - d_K(z)}{t} \\ &\geq \limsup_{t \rightarrow 0+} \frac{d_K(x) - d_K(x - ty)}{t} \\ &\geq \limsup_{t \rightarrow 0+} \frac{\|x - e(x)\| - \|x - e(x) - ty\|}{t} \end{aligned}$$

So for  $y \equiv x - e(x)$ ,

$$d_K^0(x)(x - e(x)) = \|x - e(x)\| \geq -d_K^0(x)(-(x - e(x))).$$

But then there exists an  $f \in \partial d_K(x)$  such that  $f(x - e(x)) = \|x - e(x)\|$ . Since the distance function  $d_K$  is Lipschitz one,  $\|f\| \leq 1$  for all  $f \in \partial d_K(x)$ . So  $f \in \partial \|x - e(x)\|$  [3, Theorem 5].  $\square$

Of course, if  $X$  is smooth, then for each  $x \in E(K)$ ,  $\partial \|x - e(x)\|$  is singleton and  $\partial \|x - e(x)\| \subseteq \partial d_K(x)$ .

To establish our formulae for points on the boundary of the set, we depend heavily on the following preliminary lemma developed for locally Lipschitz functions on a general normed linear space.

Given any locally Lipschitz function  $\phi$  on an open subset  $G$  of a normed linear space  $X$ , for any  $x \in G$  we denote by

$\phi_x^>$  the set of weak\* cluster points of  $f_n \in \partial \phi(z_n)$  as  $z_n \rightarrow x$  and  $\phi(z_n) > \phi(x)$ .

LEMMA 2. *If  $\phi$  attains a local minimum at  $x \in G$  then*

$$\partial \phi(x) = \overline{\text{co}}^{w^*} (\{0\} \cup \phi_x^>).$$

PROOF. The weak\* upper semicontinuity of the subdifferential mapping  $x \rightarrow \partial \phi(x)$  implies that  $\partial \phi(x)$  contains  $\phi_x^>$ . The fact that  $\phi$  has a local minimum at  $x$  implies that  $0 \in \partial \phi(x)$ . So we only need to prove containment of  $\partial \phi(x)$ . Suppose not. Then we can strongly separate some  $f_0 \in \partial \phi(x)$  from the other set which is weak\* compact and convex. That is, there exists a  $y \in X$ , and an  $r > 0$  such that

$$f_0(y) - r > \sup\{f(y) : f \in \phi_x^> \cup \{0\}\} \equiv s \geq 0.$$

Then

$$f_0(y) - r > s \geq \sup\{f(y) : f \in \phi_x^>\}.$$

Since  $f_0 \in \partial \phi(x)$ ,

$$0 < s + r < f_0(y) \leq \limsup_{\substack{z \rightarrow x \\ t \rightarrow 0+}} \frac{\phi(z + ty) - \phi(z)}{t}.$$

Now there exists  $z_n \rightarrow x$  and  $t_n \rightarrow 0+$  such that

$$0 < s + \frac{r}{2} < f_0(y) - \frac{r}{2} < \frac{\phi(z_n + t_n y) - \phi(z_n)}{t_n}$$

for all  $n$  sufficiently large.

We may suppose that  $\phi(x) = 0$ , the minimum of  $\phi$  on  $G$ . Clearly  $\phi(z_n + t_n y) > 0$  for all such  $n$ . For a fixed  $n$  sufficiently large, consider the interval  $[z_n, z_n + t_n y]$ .

Since  $\phi$  is continuous there exists a largest  $0 \leq r_n < 1$  such that  $\phi(z_n + r_n t_n y) = \phi(z_n)$ . Then on the interval  $(z_n + r_n t_n y, z_n + t_n y]$ ,  $\phi$  is strictly positive. By the Mean Value Theorem [6, Theorem 2.3.7], there exists an  $a_n \in (z_n + r_n t_n y, z_n + t_n y)$  and an  $f_n \in \partial\phi(a_n)$  such that

$$f_n(t_n y) \geq f_n((1 - r_n)t_n y) = \phi(z_n + t_n y) - \phi(z_n + r_n t_n y) > 0.$$

But  $\phi(a_n) > 0$  so  $f_n$  satisfies the defining property in  $\phi_x^>$  and a weak\* cluster point  $f_*$  of  $\{f_n \text{ as } n \rightarrow \infty\}$  satisfies  $f_* \in \phi_x^>$ . But then

$$0 < s + r/2 < f_0(y) - r/2 \leq f_*(y) \leq s$$

which provides a contradiction.  $\square$

**2. The case using uniform Gâteaux differentiability of the norm.** We first establish a proximal normal formula for those normed linear spaces where the norm is uniformly Gâteaux differentiable.

The norm of a normed linear space  $X$  is said to be *uniformly Gâteaux differentiable* if, given  $\varepsilon > 0$  and  $y \in X$  there exists a  $\delta(\varepsilon, y) > 0$  such that for every  $x, \|x\| = 1$  there is a continuous linear functional  $f_x$  on  $X$  and

$$|(\|x + ty\| - \|x\|)/t - f_x(y)| < \varepsilon \quad \text{for all } 0 < |t| < \delta.$$

This property has the following useful characterization.

**LEMMA 3.** *The norm is uniformly Gâteaux differentiable if and only if, for each  $y \in X$  and any  $r > 0$*

$$\lim_{t \rightarrow 0} \sup_{\|x\| > r} (f_{x+ty} - f_x)(y) = 0$$

for  $f_x \in \partial\|x\|$  and  $f_{x+ty} \in \partial\|x + ty\|$  [3, Lemma 7(i)] and [12, p. 299].

We consider a situation where  $E(K)$  is dense in  $X \setminus K$ . For any subset  $P(K)$  of  $E(K)$  which is also dense in  $X \setminus K$ , choose for each  $x \in P(K)$  a point  $p(x) \in K$  such that  $\|x - p(x)\| = d_K(x)$ .

Consider a smooth normed linear space  $X$ . Given  $x \in X \setminus \text{int } K$ , denote by  $D_x(P)$  the set of weak\* cluster points of  $f_{z_n - p(z_n)}$  for  $z_n \in P(K)$  as  $z_n \rightarrow x$ .

**THEOREM 4.** *In a normed linear space  $X$  where the norm is uniformly Gâteaux differentiable, given any nonempty closed set  $K$  where  $E(K)$  is dense in  $X \setminus K$ ,*

(i) *for  $x \in X \setminus K$ ,*

$$\partial d_K(x) = \overline{\text{co}}^{w^*} D_x(P).$$

(ii) *For  $x \in \text{bdy } K$ ,*

$$\partial d_K(x) = \overline{\text{co}}^{w^*} (\{0\} \cup D_x(P))$$

*and  $N_K(x)$  is the weak\* closed convex cone generated by  $\{0\}$  and  $D_x(P)$  (a proximal normal formula in Banach space).*

**PROOF.** (i) From the weak\* upper semicontinuity of the subdifferential mapping  $x \rightarrow \partial d_K(x)$  we see that  $\partial d_K(x)$  contains  $D_x(P)$ . So we need only prove the containment of  $\partial d_K(x)$ . Suppose not. Then we can strongly separate some  $f_0 \in \partial d_K(x)$  from the other set which is weak\* compact and convex. That is, there exists a  $y \in X$  and  $r > 0$  such that

$$f_0(y) - r > \sup\{f(y) : f \in D_x(P)\}.$$

Since  $f_0 \in \partial d_K(x)$

$$f_0(y) \leq \limsup_{\substack{z \rightarrow x \\ t \rightarrow 0+}} \frac{d_K(z + ty) - d_K(z)}{t}.$$

Now there exists  $z_n \rightarrow x$  and  $t_n \rightarrow 0+$  such that

$$f_0(y) - \frac{r}{2} < \frac{d_K(z_n + t_n y) - d_K(z_n)}{t_n}$$

for sufficiently large  $n$ .

Since  $x \in X \setminus K$ ,  $z_n \in X \setminus K$  for  $n$  sufficiently large. Now since  $P(K)$  is dense in  $X \setminus K$ , for each such  $n$  we can choose  $z'_n \in P(K)$  such that  $\|z'_n - z_n\| < t_n^2/2$ . Then

$$\begin{aligned} f_0(y) - \frac{r}{2} &< \frac{d_K(z'_n + t_n y) - d_K(z'_n)}{t_n} + t_n \\ &\leq \frac{\|z'_n - p(z'_n) + t_n y\| - \|z'_n - p(z'_n)\|}{t_n} + t_n \\ &\leq f_{z'_n - p(z'_n) + t_n y}(y) + t_n. \end{aligned}$$

Lemma 3 implies that

$$f_0(y) - \frac{r}{2} \leq \limsup_{z'_n \rightarrow x} f_{z'_n - p(z'_n)}(y)$$

since the norm is uniformly Gâteaux differentiable. But a weak\* cluster point  $f_*$  of  $f_{z'_n - p(z'_n)}$  as  $z'_n \rightarrow x$  satisfies  $f_0(y) - r/2 \leq f_*(y)$  and this contradicts the separation property.

(ii) Applying Lemma 2 to the distance function  $d_K$  at  $x \in \text{bdy } K$ , we have that  $\phi_x^>$  is the set of weak\* cluster points of elements of  $\partial d_K(z_n)$  for  $z_n \in X \setminus K$  as  $z_n \rightarrow x$ . But by (i) we have for each  $z_n \in X \setminus K$ ,

$$\partial d_K(z_n) = \overline{\text{co}}^{w^*} D_{z_n}(P).$$

So for  $x \in \text{bdy } K$ ,

$$\partial d_K(x) = \overline{\text{co}}^{W^*}(\{0\} \cup D_x(P)). \quad \square$$

Note that for  $x \in \text{bdy } K$  and  $z \in P(K)$

$$\|x - p(z)\| \leq \|x - z\| + \|z - p(z)\| \leq 2\|x - z\|$$

so  $p(z) \rightarrow x$  as  $z \rightarrow x$ . But also

$$\|z - x\| \leq \|z - p(z)\| + \|p(z) - x\|$$

so if  $p(z) \rightarrow x$  and  $\|z - p(z)\| \rightarrow 0$  then  $z \rightarrow x$ .

We say that  $v$  is a  $P(K)$ -proximal normal to  $K$  at  $y \in \text{bdy } K$  if there exists a  $z \in P(K)$  and  $y = p(z)$  and  $v = \lambda(z - p(z))$  for some  $\lambda > 0$ .

In a Hilbert space  $X$  which is self-dual, we have for  $x \in \text{bdy } K$ , that

$D_x(P)$  is the set of weak cluster points of  $v_n/\|v_n\|$  where  $v_n$  is a  $P(K)$ -proximal normal to  $K$  at  $y_n \in \text{bdy } K$  and  $y_n \rightarrow x$  and  $\|v_n\| \rightarrow 0$ .

**COROLLARY 5.** *Given any closed nonempty set  $K$  in a Hilbert space  $X$ , for  $x \in \text{bdy } K$*

$$\partial d_K(x) = \overline{\text{co}}(\{0\} \cup D_x(P))$$

and  $N_K(x)$  is the closed convex cone generated by  $\{0\}$  and  $D_x(P)$  (the proximal normal formula in Hilbert space).

PROOF. A Hilbert space has uniformly Gâteaux differentiable norm and is uniformly convex. Edelstein [7] showed that any closed nonempty set  $K$  in a uniformly convex Banach space  $X$  has  $E(K)$  dense  $X \setminus K$ . It follows that Theorem 4 applies to all closed nonempty sets in Hilbert space.  $\square$

We should notice that Theorem 4 and Corollary 5 are an improvement of the existing formulae even in  $\mathbf{R}^n$  in that they only require an approach by a dense subset of proximal normals.

A normed linear space  $X$  is said to have a *Kadec norm* if for any sequence  $x_n$  which converges weakly to  $x$  where also  $\|x_n\| \rightarrow \|x\|$  then  $x_n$  converges in norm to  $x$ .

Borwein has shown [2, Theorem 4.1] that in a Banach space  $X$  which has a Kadec norm and is such that  $X^*$  has an equivalent strictly convex dual norm, any norm closed boundedly relatively weakly compact subset  $K$  has  $E(K)$  generically dense in  $X \setminus K$ .

This extended Lau's result [10, Theorem 4] that in any reflexive Banach space  $X$  with Kadec norm, every closed nonempty subset  $K$  has  $E(K)$  second category. Konjagin has shown [9, Theorem 2] that if  $X$  is a Banach space which does not belong to this class then there exists a closed nonempty set  $K$  where  $E(K)$  is not dense in  $X \setminus K$ . So this is the class of Banach spaces in which the density property holds for all closed nonempty subsets.

We should note that in every finite dimensional space the norm is Kadec and in general every locally uniformly convex norm is Kadec; in particular, the  $L_p$  ( $1 < p < \infty$ ) spaces have Kadec norm. Moreover, every reflexive Banach space can be renormed to have norm simultaneously locally uniformly convex and Fréchet differentiable. Since the normal cone, unlike the distance function, does not change on renorming, this renorming property is often useful in applications such as might follow from the material in the next section.

**3. The case using Fréchet subdifferentiability of the distance function.** The proof of the proximal normal formula in Theorem 4 is particularly simple because of the uniform Gâteaux differentiability of the norm. We now extend the result without this geometrical condition.

In [5, Proposition 6.5], Borwein and Strojwas introduced the notion of a Fréchet subderivative of an extended real function  $\phi$  on a normed linear space  $X$ . A continuous linear functional  $f$  on  $X$  is said to be a *Fréchet subderivative* of  $\phi$  at  $x$  if

$$\liminf_{\substack{\|y\| \rightarrow 0 \\ y \neq 0}} \|y\|^{-1} (\phi(x+y) - \phi(x) - f(y)) \geq 0.$$

The set of Fréchet subderivatives of  $\phi$  at  $x$  is denoted by  $\partial^F \phi(x)$  and  $\phi$  is said to be *Fréchet subdifferentiable* at  $x$  if  $\partial^F \phi(x)$  is not empty. They show [4, Theorem 6.2] that every lower semicontinuous function  $\phi$  on a reflexive Banach space  $X$  is densely Fréchet subdifferentiable on its domain.

Given a nonempty closed set  $K$  in a normed linear space  $X$  we denote by  $F(K)$  the set of points in  $X \setminus K$  where  $d_K$  is Fréchet subdifferentiable. So if  $X$  is reflexive then  $F(K)$  is dense in  $X \setminus K$ .

We establish a proximal normal formula using Fréchet subderivatives through the following lemma.

LEMMA 6. *On a reflexive Banach space  $X$  given any nonempty closed set  $K$ , for  $x \in F(K)$  if  $f \in \partial^F d_K(x)$  then there exists an  $r(x) \in K$  such that  $\|x - r(x)\| = d_K(x)$  and  $f \in \partial\|x - r(x)\|$ . If  $X$  is smooth then  $\partial^F d_K(x)$  is singleton. If the norm has the Kadec property, then any minimizing sequence  $\{x_n\}$  has a subsequence which converges in norm to  $r(x)$  and so  $r(x) \in K$ ; if also the norm is Gâteaux (Fréchet) differentiable at  $x - r(x)$  then  $f$  is the Gâteaux (Fréchet) derivative of  $d_K$  at  $x$ .*

PROOF. Given  $\varepsilon > 0$ , consider a real sequence  $\varepsilon/2 > t_n \rightarrow 0+$  and choose  $\bar{x}_n \in K$  such that  $\|x - \bar{x}_n\| \leq d_K(x) + t_n^2$ . Then since  $f \in \partial^F d_K(x)$ ,

$$f(\bar{x}_n - x) \leq \frac{d_K(x + t_n(\bar{x}_n - x)) - d_K(x)}{t_n} + \frac{\varepsilon}{2}$$

for  $n$  sufficiently large,

$$\begin{aligned} &\leq \frac{\|x + t_n(\bar{x}_n - x) - \bar{x}_n\| - \|x - \bar{x}_n\|}{t_n} + \varepsilon \\ &= \varepsilon - \|x - \bar{x}_n\|. \end{aligned}$$

So

$$\liminf_{n \rightarrow \infty} f(x - \bar{x}_n) \geq d_K(x).$$

But  $\|f\| \leq 1$  so  $\|f\| = 1$  and

$$\lim_{n \rightarrow \infty} f(x - \bar{x}_n) = d_K(x).$$

Since  $X$  is reflexive there exists a subsequence  $\bar{x}_{n(k)}$  weakly convergent to some  $r(x)$ . Then

$$\lim_{n \rightarrow \infty} \|x - \bar{x}_{n(k)}\| = d_K(x) = f(x - r(x)) \leq \|x - r(x)\|.$$

Therefore,  $x - \bar{x}_{n(k)}$  converges weakly to  $x - r(x)$  and  $\|x - \bar{x}_{n(k)}\| \rightarrow \|x - r(x)\|$ . We conclude that  $f(x - r(x)) = \|x - r(x)\|$  and  $\partial^F d_K(x) \subseteq \partial\|x - r(x)\|$ . If  $X$  is smooth then  $\partial^F d_K(x)$  is singleton and  $f = f_{x-r(x)}$ . If the norm has the Kadec property,  $\bar{x}_{n(k)}$  is norm convergent to  $r(x)$  and so  $r(x) \in K$ . But then, for all  $y \in X$ ,  $\|y\| = 1$  there exists a  $\delta > 0$  such that

$$f_{x-r(x)}(y) \leq (d_K(x + ty) - d_K(x))/t + \varepsilon$$

for all  $0 < |t| < \delta$

$$\begin{aligned} &\leq (\|x - r(x) + ty\| - \|x - r(x)\|)/t + \varepsilon \\ &\leq f_{x-r(x)+ty}(y) + \varepsilon \end{aligned}$$

for all  $0 < |t| < \delta$ . We deduce that if the norm is Gâteaux (Fréchet) differentiable at  $x - r(x)$  then  $d_K$  is Gâteaux (Fréchet) differentiable at  $x$ .  $\square$

We are now in a position to establish another general form of the proximal normal formula in Banach space. Consider a smooth reflexive Banach space  $X$  with Kadec norm. Given a nonempty closed set  $K$ , for any  $x \in X$  we denote by  $F_x$  the set of weak cluster points of  $f_{z_n-r(z_n)}$  for  $z_n \in F(K)$  as  $z_n \rightarrow x$ .

**THEOREM 7.** *In a smooth reflexive Banach space  $X$  with Kadec norm, given any nonempty closed set  $K$*

(i) *for  $x \in X \setminus K$ ,*

$$\partial d_K(x) = \overline{\text{co}} F_x = \overline{\text{co}} D_x(E(K)).$$

(ii) *For  $x \in \text{bdy } K$ ,*

$$\partial d_K(x) = \overline{\text{co}}(\{0\} \cup F_x) = \overline{\text{co}}(\{0\} \cup D_x(E(K)))$$

*and  $N_K(x)$  is the weak\* closed convex cone generated by  $\{0\}$  and  $F_x$  (a proximal normal formula in Banach space).*

**PROOF.** Borwein and Strojwas have shown [5, Corollary 7.4] that in a reflexive Banach space  $X$ , for any  $x \in X$ ,  $\partial d_K(x)$  is the closed convex hull of the weak cluster points of  $\partial^F d_K(z_n)$  as  $z_n \rightarrow x$ . Since  $X$  is reflexive,  $F(K)$  is dense in  $X \setminus K$ . But Lemma 6 tells us that in a reflexive Banach space with Kadec norm,  $\partial d_K(x) \subseteq \overline{\text{co}} F_x$ . Therefore,

(i) for  $x \in X \setminus K$ ,  $\partial d_K(x) = \overline{\text{co}} F_x$ ,

the other containment following from Lemma 1 since  $X$  is smooth. Now we apply Lemma 2 to the distance function  $d_K$  at  $x \in \text{bdy } K$ . Using the result for points in  $X \setminus K$ ,

(ii) for  $x \in \text{bdy } K$

$$\partial d_K(x) = \overline{\text{co}}(\{0\} \cup F_x). \quad \square$$

It follows from Lemma 6 and this theorem that in a reflexive Banach space with Kadec norm if the norm is also Gâteaux (Fréchet) differentiable on  $X \setminus \{0\}$  then  $d_K$  is Gâteaux (Fréchet) differentiable at points where it is Fréchet subdifferentiable on  $X \setminus K$ . So for the derivative taken in the appropriate sense we have, for  $X \setminus K$

$\partial d_K(x)$  is the closed convex hull of the weak cluster points of the derivatives  $\nabla d_K(x_n)$  for points  $x_n$  where they exist as  $x_n \rightarrow x$ ,

and for  $x \in \text{bdy } K$

$\partial d_K(x)$  is the closed convex hull of 0 and the weak cluster points of the derivatives  $\nabla d_K(x_n)$  for points  $x_n \in X \setminus K$  where they exist as  $x_n \rightarrow x$ .

These formulae correspond closely to the original definitions given by Clarke for  $\mathbf{R}^n$ .

Smoothness is assumed in Theorem 7 to make the formula workable in practice. After Lemma 1, we noted that if the normed linear space  $X$  is smooth, then for each  $x \in E(K)$  and each  $e(x) \in E(K)$  where  $\|x - e(x)\| = d_K(x)$  we have  $\partial \|x - e(x)\| \subseteq \partial d_K(x)$ . But we should also notice that if  $X$  is not smooth then there may and usually will exist an  $f \in \partial \|x - e(x)\|$  where  $f \notin \partial d_K(x)$ . In this case we will need the added provision that we consider only those elements  $f \in \partial \|x - e(x)\| \cap \partial d_K(x)$ . But this can be difficult to apply in practice. Without such a provision the results may not hold. The following elementary example illustrates the problem.

**EXAMPLE 8.** Consider  $\mathbf{R}^2$  with norm  $\|\cdot\|_1$  and  $K$  to be the closed upper half plane. For  $x \equiv (0, -1)$  we have  $d_K(x) = 1$  and  $e(x) = (0, 0)$ . Also for  $y \equiv (\lambda, \mu)$  we have

$$f_{x-e(x)}(y) = \alpha\lambda - \mu \quad \text{for all } -1 \leq \alpha \leq 1.$$



But  $\partial d_K(x)$  is singleton and is given by  $f(y) = -\mu$ . Now the normal cone to  $K$  at  $(0, 0)$  is clearly the negative  $\mu$  axis generated by  $(0, -1)$  but the closed convex cone generated by all the  $f_{z_n} - e(z_n)$  for  $z_n$  in the open lower half plane as  $z_n \rightarrow (0, 0)$  is the wedge  $\{y \equiv (\lambda, \mu): \mu + |\lambda| \leq 0\}$  generated by  $(\alpha, -1)$  for all  $-1 \leq \alpha \leq 1$ .  $\square$

Borwein, Fitzpatrick and Giles have shown [3, Theorem 12] that in a Banach space  $X$  where  $X^*$  has strictly convex norm, given a nonempty closed set  $K$  if for each  $x \in X \setminus K$

$$\limsup_{\|y\| \rightarrow 0} \frac{d_K(x+y) - d_K(x)}{\|y\|} = 1$$

then  $d_K$  is a convex function.

So the following property of Fréchet subderivatives for distance functions should be contrasted with the property that distance functions are densely Fréchet subdifferentiable on any reflexive Banach space [5, Theorem 2].

**THEOREM 9.** *In a Banach space  $X$  where  $X^*$  has strictly convex norm, if for a given nonempty closed set  $K$ ,  $d_K$  is Fréchet subdifferentiable at each  $x \in X \setminus K$  then  $d_K$  is a convex function.*

**PROOF.** From the proof of Lemma 6 we see that, for  $n$  sufficiently large we have both

$$f(x - \bar{x}_n) \leq \frac{d_K(x + t_n(x - \bar{x}_n)) - d_K(x)}{t_n} + \frac{\varepsilon}{2}$$

and

$$\|x - \bar{x}_n\| - \varepsilon \leq f(x - \bar{x}_n).$$

So

$$\|x - \bar{x}_n\| \leq \frac{d_K(x + t_n(x - \bar{x}_n)) - d_K(x)}{t_n} + \frac{3\varepsilon}{2}.$$

Then

$$\limsup_{\|y\| \rightarrow 0} \frac{d_K(x+y) - d_K(x)}{\|y\|} = 1. \quad \square$$

Konjagin's example [9, Theorem 2] actually provides in any non-Kadec space, a subset  $K$  such that  $d_K$  is linear (Fréchet) on a neighborhood but no point in the neighbourhood has a closest point in  $K$ .

It is of interest to note that, following an argument given by Asplund [1, p. 235], the differentiability properties of a distance function in Hilbert space can be deduced immediately from convex analysis. This enables us to deduce the Hilbert space formula more directly and with additional information.

**THEOREM 10.** *In a Hilbert space  $X$  the distance function  $d_K$  generated by any nonempty closed set  $K$  is generically Fréchet differentiable on  $X \setminus K$ .*

**PROOF.** We may write

$$d_K^2(x) = \|x\|^2 - \phi_K(x) \quad \text{where } \phi_K(x) = \sup_{y \in K} \{\|x\|^2 - \|x - y\|^2\}.$$

Now

$$\phi_K(x) = \sup_{y \in K} \{2(y, x) - \|y\|^2\}.$$

and so as a supremum of affine functionals  $\phi_K$  is a convex function. Since  $X$  is an Asplund space,  $\phi_K$  and therefore  $d_K$  is generically Fréchet differentiable on  $X \setminus K$ .  $\square$

Moreover, for Hilbert space, Theorem 4(i) now follows from the fact that

$$2d_K(x)\partial d_K(x) = 2x - \partial\phi_K(x)$$

and the convex analysis result that

$$\partial\phi_K(x) \text{ is the closed convex hull of all the weak cluster points of } \nabla\phi_K(z_n) \text{ for } z_n \text{ belonging to any dense subset and } z_n \rightarrow x.$$

Then Theorem 4(ii) follows from Lemma 2.

**4. An application to a class of perturbed optimization problems.** In their paper [8, Theorem 3.11], Ekeland and Lebourg consider a family of optimization problems. We now show how the discussion of the proximal normal formula using Fréchet subdifferentiability generalizes to provide an illuminating approach to such problems.

**THEOREM 11.** *Consider a reflexive Banach space  $X$  with Kadec norm,  $Y$  a topological space and  $T$  a continuous proper mapping from  $Y$  into  $X$  with  $T(Y)$  bounded. Consider the function*

$$\phi(x) = \inf_{y \in Y} \{g(y) + h(\|x - Ty\|)\}$$

where  $g$  is lower semicontinuous on  $Y$  and bounded from below, and  $h$  is continuous on  $[0, \infty)$  with derivative  $h'$  continuous and positive on  $(0, \infty)$  and  $h(0) = 0$ .

(i) *For each point  $x_0 \in X$  from the dense subset where  $\phi$  is Fréchet subdifferentiable there exists a  $y_0 \in Y$  such that*

$$\phi(x_0) = g(y_0) + h(\|x_0 - Ty_0\|).$$

(ii) *When  $X$  is smooth, for any  $x \in X \setminus T(Y)$*

*$\partial\phi(x)$  is the closed convex hull of the weak cluster points of  $h'(\|z_n - Tw_n\|)f_{z_n - Tw_n}$  for those  $z_n$  where  $\phi(z_n)$  attains its infimum at  $w_n \in Y$  and  $z_n \rightarrow x$ .*

**PROOF.** (i) Given  $\varepsilon > 0$ , consider a real sequence  $\varepsilon/2 > t_n \rightarrow 0+$  and choose  $y_n \in Y$  such that

$$g(y_n) + h(\|x_0 - Ty_n\|) \leq \phi(x_0) + t_n^2/2.$$

For  $f \in \partial^F\phi(x_0)$ , and  $n$  large

$$f(Ty_n - x_0) \leq \frac{h((1 - t_n)\|x_0 - Ty_n\|) - h(\|x_0 - Ty_n\|)}{t_n} + \varepsilon.$$

Since  $X$  is reflexive, there exists a subsequence  $Ty_{n(k)}$  weak convergent to some  $\bar{y}_0$  and  $\|Ty_{n(k)} - x_0\|$  convergent to some  $d$ . If  $d = 0$  then since  $T$  is proper [8, p. 205] the set  $\{y_n\}$  has a cluster point  $y_0 \in Y$  and  $Ty_0 = x_0$ . If  $d > 0$  then  $f(\bar{y}_0 - x_0) \leq -h'(d)d$ . Similarly, for any  $x \in X$ ,  $\|x\| = 1$

$$f(x) \leq \frac{h(\|x_0 - Ty_{n(k)}\| + t_{n(k)}\|x\|) - h(\|x_0 - Ty_{n(k)}\|)}{t_{n(k)}} + \varepsilon$$

since  $h'$  is strictly positive, and so  $f(x) \leq h'(d)\|x\|$  and  $\|f\| \leq h'(d)$ . Therefore,  $f(x_0 - \bar{y}_0) \geq h'(d)d$  implies that  $\|x_0 - \bar{y}_0\| \geq d$ . Then  $x_0 - Ty_{n(k)}$  converges weakly to  $x_0 - \bar{y}_0$  but also  $\|x_0 - Ty_{n(k)}\| \rightarrow \|x_0 - \bar{y}_0\| = d$ . So  $f(x_0 - \bar{y}_0) = h'(\|x_0 - \bar{y}_0\|)\|x_0 - \bar{y}_0\|$ ; that is,  $f \in h'(\|x_0 - \bar{y}_0\|)\partial\|x_0 - \bar{y}_0\|$ . By the Kadec property of the norm,  $Ty_{n(k)}$  is norm convergent to  $\bar{y}_0$ . But again since  $T$  is proper, the set  $\{y_{n(k)}\}$  has a cluster point  $y_0 \in Y$  and  $Ty_0 = \bar{y}_0$ . We conclude that in both cases this infimum is attained at  $y_0$ .

(ii) Since  $X$  is reflexive and  $\phi$  lower semicontinuous on  $X$ , it follows from [5, Theorem 6.2] that  $\phi$  is Fréchet subdifferentiable on a dense subset of  $X$ . As in [8, p. 213],  $\phi$  is locally Lipschitz on  $X$  so we can apply [5, Corollary 7.4] and our formula for  $\partial\phi(x)$  follows from (i).  $\square$

Again smoothness in Theorem 11(ii) ensures that each  $h'(\|x_0 - Ty_0\|)f_{x_0 - Ty_0}$  is actually a generalized subgradient of  $\phi$  at  $x_0$ .

Like the Ekeland and Lebourg theorem, ours covers a wide class of optimization problems; it includes the cases when  $g = 0$  or  $h = t^r$  where  $r \geq 1$ . But we should point out that their techniques require a Banach space which simultaneously is uniformly rotund and has a uniformly Fréchet differentiable norm.

## REFERENCES

1. E. Asplund, *Chebyshev sets in Hilbert space*, Trans. Amer. Math. Soc. **444** (1969), 235–240.
2. J. M. Borwein, *Weak local supportability and applications to approximation*, Pacific J. Math. **82** (1979), 323–338.
3. J. M. Borwein, S. P. Fitzpatrick and J. R. Giles, *The differentiability of real functions on normed linear spaces using generalized subgradients*, J. Math. Anal. Appl. (to appear).
4. J. M. Borwein and H. M. Strojwas, *Proximal analysis and boundaries of closed sets in Banach space. I: Theory*, J. Canad. Math. Soc. **38** (1986), 431–452.
5. ———, *Proximal analysis and boundaries of closed sets in Banach space. II: Applications*, J. Canad. Math. Soc. (to appear).
6. F. H. Clarke, *Optimization and nonsmooth analysis*, Canad. Math. Soc. Ser. Monographs and Advanced Texts, Wiley, New York, 1983.
7. M. Edelstein, *On nearest points of sets in uniformly convex Banach spaces*, J. London Math. Soc. **43** (1968), 375–377.
8. I. Ekeland and G. Lebourg, *Generic Fréchet-differentiability and perturbed optimization problems in Banach spaces*, Trans. Amer. Math. Soc. **224** (1976), 193–216.
9. S. V. Konjagin, *On approximation properties of closed sets in Banach spaces and the characterization of strongly convex spaces*, Soviet Math. Dokl. **21** (1980), 418–422.
10. K. S. Lau, *Almost Chebyshev subsets in reflexive Banach spaces*, Indiana Univ. Math. J. **2** (1978), 791–795.
11. P. D. Loewen, *The proximal normal formula in Hilbert space*, Nonlinear Anal. (to appear).

DEPARTMENT OF MATHEMATICS, STATISTICS AND COMPUTING SCIENCE, DALHOUSIE UNIVERSITY, HALIFAX, NOVA SCOTIA, CANADA B3H 3J5

DEPARTMENT OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCE, THE UNIVERSITY OF NEWCASTLE, NEWCASTLE, N.S.W., AUSTRALIA