Canad. Math. Bull. Vol. **31** (3), 1988

# THE PROXIMAL SUBGRADIENT FORMULA IN BANACH SPACE

BY
PHILIP D. LOEWEN

ABSTRACT. The proximal subgradient formula is a refinement due to Rockafellar of Clarke's fundamental proximal normal formula. It expresses Clarke's generalized gradient of a lower semicontinuous function in terms of analytically simpler proximal subgradients. We use the infinite-dimensional proximal normal formula recently given by Borwein and Strojwas to derive a new version of the proximal subgradient formula in a reflexive Banach space $X$ with Frechet differentiable and locally uniformly convex norm. Our result improves on the one given by Borwein and Strojwas by referring only to the given norm on $X$.

Proximal analysis is an indispensable part of Clarke's nonsmooth calculus. Its importance derives partly from the elegance with which it combines both geometric and analytic viewpoints, but primarily from its exceptional versatility as a calculating tool. The basic result is Clarke's "proximal normal formula" [4], which explains how the (Clarke) normal cone to an arbitrary closed set can be described in terms of simpler objects called proximal normals. The formula has many applications, especially in the analysis of various value functions arising in optimization theory. In these cases the closed set of interest is the value function's epigraph, and the special shape of an epigraph set can be used to simplify the proximal normal formula. The refined result is the "proximal subgradient formula" first proven by Rockafellar [14] for the finite-dimensional case. Roughly speaking, it explains how the (Clarke) generalized gradient of a lower semicontinuous function $V:\mathbf{R}^n \to \mathbf{R} \cup \{\pm\infty\}$ can be described in terms of simpler objects called proximal subgradients. This paper offers an infinite-dimensional formulation of the proximal subgradient formula, based on the infinite-dimensional version of the proximal normal formula recently given by Borwein and Strojwas [2]. Although these authors have also considered the infinite-dimensional proximal subgradient formula [3], our result improves upon theirs by eliminating the need to consider all equivalent smooth norms

Received by the editors February 24, 1987.
Key words: proximal normals, proximal subgradients, generalized gradients, Clarke's normal cone.
AMS Subject Classification (1980): 26B05, 46B20.




https://doi.org/10.4153/CMB-1988-051-9 Published online by Cambridge University Press

on the base space when defining the set of singular limiting proximal subgradients.

This paper has only two sections. The first of these reviews the basics of proximal analysis in the finite-dimensional case, while introducing definitions and notation of more general scope. The second section presents the proximal subgradient formula in any reflexive Banach space with locally uniformly convex and Frechet differentiable norm.

## 1. **Introduction.**

DEFINITION 1.1. *Let $(X, \|\cdot\|)$ be a Banach space, and let $C$ be a nonempty closed subset of $X$. The distance function $d_C$ and the metric projection $\pi_C$ are given by*

$$d_C(x) = \inf\{ \|x - c\| : c \in C\}, \tag{1.1}$$

$$\pi_C(x) = \{c \in C : \|x - c\| = d_C(x)\}. \tag{1.2}$$

Given a point $c \in C$, we say that a nonzero vector $v \in X$ is *perpendicular* to $C$ at $c$ (written $v \perp C$ at $c$) if $v \in \pi_C^{-1}(c) - c$, i.e. $c \in \pi_C(c + v)$.

Now let $X = \mathbf{R}^n$, and let $C \subset \mathbf{R}^n$ contain a point $c$. Clarke's proximal normal formula asserts that the normal cone to $C$ at $c$ is given by

$$N_C(c) = \text{co}\{\lim \lambda_i v_i : \lambda_i \geqq 0,\ v_i \perp C \text{ at } c_i,\ c_i \in C,\ c_i \to c\}. \tag{1.3}$$

In this finite-dimensional case, the scalar multiples $\lambda_i v_i$ of perpendicular vectors appearing in (1.3) are the proximal normals. When $X$ is infinite-dimensional the normal cone is a subset of $X^*$ and a more general definition of proximal normals is required – see Definition 2.1. The utility of (1.3) for evaluating $N_C(c)$ springs from the fact that when $X = \mathbf{R}^n$ (or when $X$ is a Hilbert space),

$$v \perp C \text{ at } c \Leftrightarrow \langle v, c' - c\rangle \leqq (1/2)\|c' - c\|^2 \quad \forall c' \in C. \tag{1.4}$$

The "proximal normal inequality" appearing on the right side of (1.4) can be reformulated as the assertion that the point $c$ minimizes over $C$ the functional $\langle -v, c'\rangle + (1/2)\|c' - c\|^2$. An analysis of this auxiliary optimization problem can lead to a complete understanding of $N_C(c)$ via (1.3).

The program sketched above takes on particular interest when the set $C$ in question arises as the epigraph of some lower semicontinuous function. Indeed, let $X$ be a Banach space on which some $V : X \to \mathbf{R} \cup \{\pm\infty\}$ is defined. Assume that $V$ is finite-valued at some $x \in X$, and that the epigraph set epi $V = \{ (x, r) : r \geqq V(x) \}$ is locally closed near $(x, V(x))$. Then Clarke's *generalized gradient* and *asymptotic generalized gradient* of $V$ at $x$ may be expressed as

$$\partial V(x) = \{\zeta \in X^* : (\zeta, -1) \in N_{\text{epi}\, V}(x, V(x))\}, \tag{1.5}$$

$$\partial^\infty V(x) = \{\zeta \in X^* : (\zeta, 0) \in N_{\text{epi}\, V}(x, V(x))\}. \tag{1.6}$$

When $X = \mathbf{R}^n$, one can evaluate these measures of the first-order behaviour of $V$ by using (1.3) to compute the normal cone to epi $V$. Indeed, this is the method of choice when $V$ is the value function of some perturbed optimization problem: sensitivity studies and proofs of necessary conditions based on this approach may be found, for example, in [11, 14, 15, 16, 5, 7].

Thanks to Rockafellar [14], we now know that when $C$ = epi $V$ for some lower semicontinuous $V:\mathbf{R}^n \to \mathbf{R} \cup \{\pm\infty\}$, the special shape of $C$ allows us to retain equality in (1.3) while imposing additional conditions on the sequences of perpendiculars described on the right-hand side. The resulting identity is called the proximal subgradient formula; its extra conditions, to be studied in detail below, exclude some of the technical problems which come up if (1.3) is applied directly to (1.5) and (1.6). The proofs in [16] bear witness to this assertion.

In spite of their many uses, the proximal subgradient and proximal normal formulas have been limited to finite-dimensional spaces until recently. Borwein and Strojwas [2] gave the first complete generalization of the proximal normal formula to an infinite-dimensional setting. (Treiman's earlier results [17, 18] use Frechet epsilon-normals instead of proximal normals, and are slightly weaker.) Their proof has been illustrated for Hilbert spaces in [13], and simplified considerably in [1]. Infinite-dimensional applications along the lines described above have also appeared: see [6] and [7], for example. In [3], Borwein and Strojwas also present a version of the proximal subgradient formula. Their result suffers from the need to involve all equivalent smooth norms on the Banach space in question. The main result of this paper is a sharper version of the proximal subgradient formula in which only the given norm must be considered. We prove it in the next section by combining the infinite-dimensional proximal normal formula of [2] with an elementary argument involving nearest points.

2. **The proximal subgradient formula.** Throughout this section, we consider a reflexive Banach space $(X, \|\cdot\|)$ whose norm is both Frechet differentiable at all nonzero points of $X$ and locally uniformly convex. Any reflexive Banach space can be given an equivalent norm with these two properties ("Trojanski's theorem"); also, the usual norms of the classical $L^p$-spaces $(1 < p < \infty)$ and all Hilbert spaces have both properties. See Diestel [10] for precise definitions and proofs. For convenience of notation we will sometimes write $n(x) = \|x\|$. Thus the subgradient set $\partial n(x)$ consists of all functionals $\xi \in X^*$ for which

$$(2.1) \qquad \langle \xi, y - x \rangle \leqq \|y\| - \|x\| \quad \forall y \in X.$$

In particular, if $x \neq 0$ then $\partial n(x)$ consists of a single functional $\xi = \nabla n(x)$ obeying $\|\xi\|_* = \xi(x) = 1$; and moreover the mapping $y \to \nabla n(y)$ is norm-to-norm continuous at $x$.

https://doi.org/10.4153/CMB-1988-051-9 Published online by Cambridge University Press

DEFINITION 2.1. *Let C be a closed subset of X containing a point c. If $\pi_C^{-1}(c)$ contains points distinct from c, then the set of proximal normal functionals to C at c is the subset of $X^*$ given by*

$$PN_C(c) = \{\lambda\nabla n(v):\lambda \geqq 0,\ v \perp C \text{ at } c\}; \tag{2.2}$$

*if $\pi_C^{-1}(c) = \{c\}$ then $PN_C = \{0\}$.*

Here is the infinite-dimensional version of the proximal normal formula, from which the proximal subgradient formula will be derived.

THEOREM 2.2 [2]. *Let C be a closed subset of X containing a point c. Then Clarke's normal cone to C at c is given by*

$$N_C(c) = \overline{\text{co}}\{w^* - \lim \zeta_i:\zeta_i \in PN_C(c_i),\ c_i \in C,\ c_i \to c\}. \tag{2.3}$$

We will use the following technical lemma.

LEMMA 2.3. *For any $x \in X$ with $\|x\| = 1$, and any $\delta > 0$, $r \in (0, 1)$, let*

$$M_r(x, \delta) = \{y \in X:\|y\| \geqq 1,\ \|y - rx\| \leqq 1 - r + \delta\}. \tag{2.4}$$

*Then one has*

$$\lim_{\delta\to 0+} \sup\{\, \|y - x\|:y \in M_r(x, \delta)\,\} = 0. \tag{2.5}$$

PROOF. If the conclusion is false then there must be a constant $\epsilon > 0$ and sequences $y_i \subseteq X$, $\delta_i \to 0^+$ such that $y_i \in M_r(x, \delta_i)$ and

$$\|y_i - x\| \geqq 2\epsilon \quad \forall i.$$

Defining $x_i = y_i/\|y_i\|$, the inclusion $y_i \in M_r(x, \delta_i)$ readily implies $1 \leqq \|y_i\| \leqq 1 + \delta_i$, so that $\|x_i - y_i\| \leqq \delta_i$. For all $i$ sufficiently large we have $\delta_i < \epsilon$, whence $\|x_i - x\| \geqq \epsilon$. The local uniform convexity of $X$ implies that, for some $\eta > 0$, $\|x_i + x\| \leqq 2(1 - \eta)$. We now estimate

$$\begin{aligned}\|y_i - rx\| &= \|(1 + r)x_i - (x_i - y_i) - r(x + x_i)\|\\ &\geqq (1 + r) - \delta_i - 2r(1 - \eta)\\ &= 1 - r + \delta_i + 2(r\eta - \delta_i).\end{aligned}$$

For all $i$ sufficiently large, the term $r\eta - \delta_i$ is positive, and the inclusion $y_i \in M_r(x, \delta_i)$ is contradicted. The lemma holds.

The remainder of this section deals with the case $C = \text{epi } V$ for a function $V:X^n \to \mathbf{R} \cup \{\pm\infty\}$. For convenience we assume that $V$ is lower semi-continuous, although the proofs remain correct if we postulate only that epi $V$ is locally closed near the point of interest. Note that epi $V$ is a subset of the product space $X \times \mathbf{R}$, whose norm we take to be

$$\|(x, r)\| = (\|x\|^2 + |r|^2)^{1/2}. \tag{2.6}$$

Clearly the unit ball of $X \times \mathbf{R}$ inherits the smoothness and convexity properties of the ball in $X$, so Theorem 2.2 remains applicable.

Let $(x, r) \in \text{epi } V$ be given. Then elements of $PN_{\text{epi } V}(x, r)$ have the form $(\zeta, -\epsilon) \in X^* \times \mathbf{R}$, and clearly $\epsilon \geqq 0$. Indeed, if $\epsilon > 0$ we must have $r = V(x)$, while if $\epsilon = 0$ and $(\zeta, 0) \in PN_{\text{epi } V}(x, r)$ then it follows that $(\zeta, 0) \in PN_{\text{epi } V}(x, V(x))$ also. These observations allow us to rewrite (2.3) as follows:

$$(2.7)\ N_{\text{epi } V}(x, V(x)) = \overline{\text{co}}\{w^* - \lim(\zeta_i, -\epsilon_i):(\zeta_i, -\epsilon_i) \in PN_{\text{epi } V}(x_i, V(x_i)),\ x_i \to x,\ V(x_i) \to V(x)\}.$$

As Rockafellar [14] observed in the finite-dimensional case, there appear to be three possibilities for the weak*-convergent sequences appearing on the right side of (2.7): either

(a) $\epsilon_i \to \epsilon > 0$, in which case $\epsilon_i > 0$ for all $i$ sufficiently large; or

(b) $\epsilon_i \to 0$ but $\epsilon_i > 0$ holds for arbitrarily large values of $i$ – in which case we may obtain the same weak*-limit by considering only the subsequences along which $\epsilon_i > 0$ for all $i$; or

(c) $\epsilon_i = 0$ for all $i$ sufficiently large.

The key assertion of the proximal subgradient formula is that the weak*-limit points of sequences of type (c) are already accounted for by sequences of type (b). Thus the extra condition $\epsilon_i > 0\ \forall i$ can be added to the right side of (2.7) while preserving the identity. This is the content of Theorem 2.4 below.

THEOREM 2.4. *Let $V:X \to \mathbf{R} \cup \{\pm\infty\}$ be lower semicontinuous, with $V(x) \in \mathbf{R}$. If a nonzero functional $(\zeta, 0) \in PN_{\text{epi } V}(x, V(x))$ and any positive tolerance $\rho$ are given, then there exist a point $x' \in X$ with $\|(x', V(x')) - (x, V(x))\| < \rho$ and a functional $(\zeta', -\epsilon') \in PN_{\text{epi } V}(x', V(x'))$ such that $\epsilon' > 0$ and $\|(\zeta', -\epsilon') - (\zeta, 0)\|_* < \rho$.*

PROOF. By definition, $(\zeta, 0) = \lambda\nabla n(z, 0)$ for some $\lambda > 0$ and some vector $(z, 0) \perp \text{epi } V$ at $(x, V(x))$. Without loss of generality we may assume $x = 0$, $V(0) = 0$, and $\|z\| = 1$. Let $\eta > 0$ be given. To prove the theorem, it suffices to produce points $x \in X$ and $(y, -\epsilon) \in (1/2)(z, 0) + 4\eta B$ such that $\|(x, V(x))\| < \eta$, $\epsilon > 0$, and $(y, -\epsilon) \perp \text{epi } V$ at $(x, V(x))$. For in this case the vector $(\zeta', -\epsilon') = \lambda\nabla n(2(y, -\epsilon))$ is in $PN_{\text{epi } V}(x, V(x))$ and can be made arbitrarily close to $(\zeta, 0)$ in the norm of $X^*$ by simply reducing $\eta$. (The last statement holds because $\nabla n:X \to X^*$ is norm-to-norm continuous [10, p. 29].)

Throughout the proof we denote the open unit ball of $X \times \mathbf{R}$ by $B$, and write $E = \text{epi } V$. Observe that since $d_E(z, 0) = \|z\| = 1$, we have $[(z, 0) + B] \cap E = \emptyset$. In fact, the special shape of $E$ implies

$$(2.8)\qquad [(z, -t) + B] \cap E = \emptyset \quad \forall t \geqq 0.$$

Let us begin by applying Lemma 2.3 to find some $\delta \in (0, \eta)$ so small that

$$\sup\{ \|(x, v)\| : \|(x, v) - (z, 0)\| \geqq 1, \\ \|(x, v) - (z/2, 0)\| \leqq 1/2 + 4\delta\} < \eta. \tag{2.9}$$

Then we define the cylinder

$$C = \{ (z/2, 0) + (u, -t) : \|u\| \leqq \delta, t \in [0, \delta] \}. \tag{2.10}$$

Now the nearest point theorem of Lau [12] states that dom $\pi_E$ contains a dense $G_\delta$ subset of $X \times \mathbf{R}$. In particular, int $C \neq \emptyset$, so $C \cap$ dom $\pi_E$ is dense in $C$. Consider the following two mutually exclusive statements about these points: either

(a) for all $(y, v) \in C \cap$ dom $\pi_E$ one has $(x, r) \in \pi_E(y, v) \Rightarrow r = v$; or

(b) there exists $(y, v) \in C \cap$ dom $\pi_E$ such that $r > v$ for some $(x, r) \in \pi_E(y, v)$. We will show that (a) is impossible and deduce the desired results from (b).

An important ingredient in our analysis of (a) is the function

$$d(t) := d_E(z/2, -t), \quad t \in [0, \delta].$$

Note that $d$ is Lipschitz of rank 1 on $[0, \delta]$, and obeys $d(0) = 1/2$. Let us prove that $d(t) > 1/2$ for $t \in (0, \delta]$. Indeed, if this were not the case then there would be some fixed $t \in (0, \delta]$ for which $d(t) \leqq 1/2$. In particular, there would be a sequence $(x_i, r_i) \in E$ for which

$$\|(x_i, r_i) - (z/2, -t)\| < d(t) + 1/i \quad \forall i.$$

Note, however, that $\|(x_i, r_i) - (z, -t)\| \geqq 1$ $\forall i$ by (2.8). Thus $(x_i, r_i) \to (0, -t)$ by Lemma 2.3, and this contradicts the lower semicontinuity of $V$. Thus $d(t) > 1/2$ for all $t \in (0, \delta]$, as claimed.

Now suppose (a) holds. Then fix any $s \in [0, \delta]$ and consider any sequence $(y_i, -s_i)$ in $C \cap$ dom $\pi_E$ converging to $(z/2, -s)$. Let $(x_i, r_i)$ be any sequence obeying $(x_i, r_i) \in \pi_E(y_i, -s_i)$. Then we have $r_i = -s_i$ by (a), and

$$\begin{aligned} d(s) &= d_E(z/2, -s) \\ &= \lim d_E(y_i, -s_i) \\ &= \lim\|(y_i, -s_i) - (x_i, r_i)\| \\ &= \lim\|z/2 - x_i\|. \end{aligned}$$

Thus $d(s)^2 = \lim\|z/2 - x_i\|^2$; for any $t \in [0, \delta]$ we deduce

$$\begin{aligned} d(t)^2 &\leqq \lim\|(z/2, -t) - (x_i, -s_i)\|^2 \\ &= \lim( \|z/2 - x_i\|^2 + |t - s_i|^2) \\ &= d(s)^2 + (t - s)^2. \end{aligned}$$

This implies that the Lipschitz function $d$ obeys

$$\frac{d(t) - d(s)}{t - s} \leqq \frac{t - s}{d(t) + d(s)} \quad 0 \leqq s < t \leqq \delta,$$

whence $d'(t) \leqq 0 \; \forall t \in [0, \delta]$. This forces $d(\delta) \leqq d(0) = 1/2$, a contradiction. So alternative (a) is impossible.

We are forced to accept condition (b), which provides points $(y_0, v_0) \in C \cap \text{dom}\, \pi_E$ and $(x, r) \in \pi_E(y_0, v_0)$ such that $r > v_0$. Setting $(y, -\epsilon) = (y_0, v_0) - (x, r)$ gives $(y, -\epsilon) \perp E$ at $(x, r)$ and $\epsilon > 0$. It follows immediately that $r = V(x)$. Moreover, $\| (y_0, v_0) - (z/2, 0) \| \leqq \sqrt{2\delta}$ implies $\| (x, V(x)) \| < \eta$ by (2.9). Finally, we have

$$\begin{aligned} \| (y, -\epsilon) - (z/2, 0) \|^2 &= \|y - z/2\|^2 + \epsilon^2 \\ &\leqq ( \|y - y_0\| + \|y_0 - z/2\| )^2 + (v_0 - r)^2 \\ &\leqq 2(\eta^2 + \delta^2) + 2(\eta^2 + \delta^2) \\ &< 8\eta^2. \end{aligned}$$

This shows that $(y, -\epsilon) \in (1/2)(z, 0) + 2\sqrt{2}\eta B$ and completes the proof.

REMARK. When $X = \mathbf{R}^n$ one has dom $\pi_E = \mathbf{R}^n$. In this special case, the cylinder $C$ in the proof above may be replaced by the line segment $(z/2, -t)$ for $t \in [0, \delta]$, and there is no need to consider sequences in the analysis of $d(t)$ or the study of (a). The resulting argument rivals Rockafellar's original proof for simplicity.

Theorem 2.4 clearly implies that adding the constraint $\epsilon_i > 0 \; \forall i$ to the right side of (2.7) leaves that identity intact. To recover the usual form of the proximal subgradient formula now requires little more than new terminology and notation.

DEFINITION 2.5. *Let the lower semicontinuous function $V: X \to \mathbf{R} \cup \{\pm\infty\}$ be finite at $x$. We call the vector $\zeta \in X^*$ a proximal subgradient of $V$ at $x$ if $(\zeta, -1) \in PN_{\text{epi}\, V}(x, V(x))$. The vector $\zeta$ is a limiting proximal subgradient of $V$ at $x$, written $\zeta \in \hat{\partial} V(x)$, if $\zeta = w^* - \lim \zeta_i$ for a sequence of proximal subgradients $\zeta_i$ of $V$ at points $x_i \to x$ for which $V(x_i) \to V(x)$. Finally, $\zeta$ is a singular limiting proximal subgradient of $V$ at $x$, written $\zeta \in \hat{\partial}^\infty V(x)$, if $\zeta = w^* - \lim \lambda_i \zeta_i$ for some sequence of proximal subgradients $\zeta_i$ as described above together with some sequence $\lambda_i \to 0^+$.*

Let us denote by $\hat{N}_{\text{epi}\, V}(x, V(x))$ the set of limiting proximal normals to epi $V$ whose convex hull is computed on the right side of (2.7). As we have observed earlier, the nonzero elements of $\hat{N}_{\text{epi}\, V}(x, V(x))$ have two distinct forms. Either the limiting proximal normal is $(\zeta, -\epsilon)$ for some $\epsilon > 0$, in which case it can be written as $\epsilon(\hat{\zeta}, -1)$ for some $\hat{\zeta} \in \hat{\partial} V(x)$, or else it is $(\zeta, 0)$, in which case $\zeta \in \hat{\partial}^{\infty} V(x)$ by Theorem 2.4. In either case, we have

$$\hat{N}_{\text{epi}\, V}(x, V(x)) = N \cup N^{\infty}, \text{ where} \tag{2.11}$$

$$N = \{\epsilon(\zeta, -1) : \epsilon \geqq 0, \zeta \in \hat{\partial} V(x)\},$$

$$N^{\infty} = \{(\zeta, 0) : \zeta \in \hat{\partial}^{\infty} V(x)\}.$$

In view of the definitions (1.5) and (1.6), we obtain from (2.11) the following result.

THEOREM 2.6 (proximal subgradient formula). *Let $X$ be a reflexive Banach space whose norm is both Frechet differentiable and locally uniformly convex. Let $V: X \to \mathbf{R} \cup \{\pm\infty\}$ be lower semicontinuous everywhere, and finite-valued at $x$. Then*

$$\partial V(x) = \overline{\text{co}}[\hat{\partial} V(x) + \hat{\partial}^{\infty} V(x)].$$

Moreover, if co $\hat{N}_{\text{epi}\, V}(x, V(x))$ happens to be closed, then one has

$$\partial V(x) = \text{co}[\hat{\partial} V(x) + \hat{\partial}^{\infty} V(x)],$$

$$\partial^{\infty} V(x) = \text{co}\, \hat{\partial}^{\infty} V(x).$$

PROOF. Apply to (2.11) the geometrical proposition of Rockafellar [15] which is presented for arbitrary normed linear spaces in [13, Proposition 4.2].

## REFERENCES

1. J. M. Borwein and J. R. Giles, *The proximal normal formula in Banach space*, Trans. Amer. Math. Soc. **302** (1987), pp. 371-381.

2. J. M. Borwein and H. M. Strojwas, *Proximal analysis and boundaries of closed sets in Banach space — Part I: Theory*, Can. J. Math. **38** (1986), pp. 431-452.

3. ———, *Proximal analysis and boundaries of closed sets in Banach space — Part II: Applications*, Can. J. Math., to appear.

4. F. H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), pp. 247-262.

5. ———, *Optimization and nonsmooth analysis*, New York, John Wiley, 1983.

6. ———, *Perturbed optimal control problems*, IEEE Trans. Auto. Control AC-31 (1986), pp. 535-542.

7. F. H. Clarke and P. D. Loewen, *The value function in optimal control: sensitivity, controllability, and time-optimality*, SIAM J. Control Optim. **24** (1986), pp. 243-263.

8.———, *State constraints in optimal control: a case study in proximal normal analysis*, SIAM J. Control Optim. **25** (1987), pp. 1440-1456.

9. M. M. Day, *Normed linear spaces*, 3d ed. Berlin: Springer-Verlag, 1973.

10. J. Diestel, *Geometry of Banach spaces — Selected topics*, Berlin: Springer-Verlag, 1975. Lecture notes in mathematics, p. 485.

11. J. Gauvin, *The generalized gradient of a marginal function in mathematical programming*, Math. of Oper. Res. **4** (1979), pp. 458-463.

12 K. S. Lau, *Almost Chebyshev subsets in reflexive Banach spaces*, Indiana Univ. Math. J. **27** (1978), pp. 791-795.

13. P. D. Loewen, *The proximal normal formula in Hilbert space*, Nonlinear Analysis **11** (1987), pp. 979-995.

14. R. T. Rockafellar, *Proximal subgradients, marginal values, and augmented Lagrangians in nonconvex optimization*, Math. of Oper. Res. **6** (1981), pp. 424-436.

15. ———, *Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming*, Math. Prog. Study **17** (1982), pp. 28-66.

16. ———, *Extensions of subgradient calculus with applications to optimization*, Nonlinear Analysis **9** (1985), pp. 665-698.

17. J. S. Treiman, *Characterization of Clarke's tangent and normal cones in finite and infinite dimensions*, Nonlinear Analysis **7** (1983), pp. 771-783.

18. ———, *Clarke's gradients and epsilon-subgradients in Banach spaces*, Trans. Amer. Math. Soc. **294** (1986), pp. 65-78.

University of British Columbia
Vancouver, B.C. V6T 1Y4