The Clarke Subdifferential and Operator Integral

Liqun Qi
Center for Operations Research & Econometrics
Universite Catholique Louvain
1348 Louvain-la-Neuve, Belgium

John R. Birge
Department of Industrial & Operations Engineering
University of Michigan
Ann Arbor, MI 48109

Technical Report 88-14

September 1988

THE CLARKE SUBDIFFERENTIAL

AND

OPERATOR INTEGRAL

Liqun Qi*

Center for Operations Research & Econometrics

Universite Catholique Louvain

1348 Louvain-la-Neuve, Belgium

John R. Birge**

Department of Industrial and Operations Engineering

University of Michigan

Ann Arbor, MI 48109

*The work of this author was sponsored by the National Science Foundation under Grant CCR-8801489, and by the Air Force Systems Command, USAF, under Grant No. AFOSR-88-0090. The U.S. Government has certain rights in this material, and is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

**The work of this author was supported in part by office of Naval Research grant N00014-86-K-0628.

Abstract

A locally Lipschitz function in an open set in a Banach space is called a primal function if its subdifferential is single-valued everywhere except a first-category set and the projection of its subdifferential at any line segment is single-valued a.e. on that segment. Maximal monotone operator in separable or reflexive Banach spaces and subdifferentials of primal functions are examples of maximal normal operators. A maximal normal operator is the subdifferential of a primal function if and only if it is cyclically normal. Through the definition of operator integral, we have equality formulas for exchanging integral and subdifferential operations.

Keywords: Subdifferentials, integrals, first-category sets.

1980 Mathematical Subject Classification (1985 Revision):

Primary 26E25; Secondary 28B20.

I. Introduction

The Clarke subdifferential theory [3] [4] [24] is a milestone of nonsmooth analysis. The subdifferential of a nonsmooth function at a point is a set instead of a single point. This becomes the main idea of nonsmooth analysis. However, under the usual set operations, the Clarke subdifferential calculus has only inclusion relationship in general. For example, generally, we have only

$$\partial(f + g)(x) \subseteq \partial f(x) + \partial g(x). \tag{1}$$

Analyses, optimal control, the calculus of variations and stochastic programming, consider the Clarke subdifferential of the integral functional f on a Banach space X given by

$$f(x) = \int_T f_t(x)\, \mu(dt), \tag{2}$$

where f_t is a family of functions on X. For example, in stochastic programming (2) may be the expectation of a recourse function [1] [26]. If we apply subgradient-based methods such as the bundle method [12], stochasticc quasigradient method [9] and other generalized gradient methods [25], we hope to compute the subdifferential of f through the integral of subdifferentials of f_t. In the finite-horizon case, X is finite-dimensional. In infinite-horizon cases, optimal control and the calculus of varations, X may be infinite-dimensional [4] [5] [6] [7] [8]. However, in general, we only have [4],

$$\partial\int_T f_t(x)\, \mu(dt) \subseteq \int_T \partial f_t(x)\mu(dt). \tag{3}$$

Our point of view for these inclusion relationships is that, if we deal with the operators among operators ∂f, ∂g and ∂f_t, instead of the operations among sets $\partial f(x)$, $\partial g(x)$ and $\partial f_t(x)$, we will get equality relations. In [17] [18] [19] [20], we developed in finite-dimensional spaces a maximal normal operator theory to deal with these operator operations. In a certain sense, this theory is parallel to the maximal monotone operator theory [2] in the nonconvex case. In this development, we consider locally Lipschitz functions with almost everywhere single-valued subdifferentials, called primal functions. Subdifferentials of primal functions and maximal monotone operators are examples of maximal normal operators. A maximal normal operator is the subdifferential of a primal function if and only if it is cyclically normal. If the domain is connected, the primal function is determined up to a constant. Various functions are shown to be primal functions. For example, continuously differentiable functions, convex functions, concave functions, and differences of convex functions are primal functions. More generally, semismooth functions [13] are primal functions [19]. Besides, the Euclidean distance function of a set C is a primal function if and only if bd Cl C has zero measure [21].

We defined function addition $\oplus$ for two maximal normal operators. In this way, the Clarke subdifferential calculus has equalities. For example, for two primal functions f and g defined on an open set D, we have

$$\partial(f + g) = \partial f \oplus \partial g. \tag{4}$$

Compare (4) with (1). Similarly we have equalities for $\partial(f \bullet g)$ and $\partial(f/g)$. All the maximal normal operators in an open set form a linear space with scaler multiplication and function addition $\oplus$. The Clarke subdifferential operator is thus a linear operator from the primal function space to the maximal normal operator space.

In this paper, we generalize this idea to Banach spaces and develop an operator equality form of (3).

In Banach spaces, we use "single-valued everywhere except a first-category set" to replace the "single-valued almost everywhere" term. Almost all the above discussions on maximal normal operators and primal functions can be thus generalized to general Banach spaces without involving a measure. In Sections 2-4, we discuss maximal normal operators, cyclical normality, and operator addition $\oplus$ in Banach spaces. We call it operator addition $\oplus$ now to emphasize operator operations.

In Sections 5-6, we discuss the operator integral op-$\int$, based upon our definitions of maximal normal operators on finite-dimensional spaces and general Banach spaces respectively. We establish the operator form of (3) as

$$\partial \textstyle\int_T f_t \, \mu(dt) = \text{op-}\textstyle\int_T \partial f_t \, \mu(dt). \qquad (5).$$

2. Maximal Normal Operators in Banach Spaces

Let X be a Banach space and X^* be its dual. For x in X and u in X^*, we adopt the convention of using $\langle u, x\rangle$ or $\langle x, u\rangle$ for u(x). We denote by $||x||$ the norm in X and by $||u||^*$ the norm in X^*:

$$||u||^* := \{\langle u, x\rangle : x \in B\}$$

where B is the closed unit ball in X.

Let Y be an open subset of X. Consider a set-valued operator $F: Y \rightarrow X^*$. For any h in X, define $F_h: Y \rightarrow \mathbf{R}$ by

$$F_h(x) = \langle F(x), h\rangle$$

for any x in Y. Let sing F: = $\{x \varepsilon Y: F(x)$ is single-valued$\}$.

Definition 2.1 We call a set-valued operator $F: Y \Rightarrow X^*$ a maximal normal operator if

(i) F is locally bounded on Y;

(ii) Y\sing F is a first-category set in X;

(iii) for any $x \varepsilon Y$,

$$F(x) = w^*\text{-cl conv } \{u \varepsilon X^*: u = w^*\text{-lim } F(x_i),\ x_i \varepsilon \text{ sing } F,\ x_i \Rightarrow x\};$$

(iv) for any line segment $[x, x + h]$ in Y, F_h is single-valued almost everywhere in this segment in the sense of one-dimensional Lebesque measure. ■

It is easy to see that

Proposition 2.2 If $F:Y \Rightarrow X^*$ is a maximal normal operator, then F is w^*-closed and for any x in Y, F(x) is a convex, w^*-compact subset of X^*. ■

The w^*-compactness follows from Alaoglu's theorem.

Example 2.3 If $F: Y \Rightarrow X^*$ is single-valued and continuous (with the norm topology) everywhere in Y, then we call F a continuous operator. A continuous operator is clearly a maximal normal operator. By Corollary 2.2.1 of [4], if $f: Y \Rightarrow \mathbf{R}$ is continuously differentiable at x, then the subdifferential of f at x is $\{f'(x)\}$. Thus the subdifferential of a continuously differentiable function $f: Y \Rightarrow \mathbf{R}$ is a continuous operator. ■

Proposition 2.4 If X is either separable of reflexive, then any maximal monotone operator $F: Y \to X^*$ is a maximal normal operator on Y.

Proof The local boundedness is a well-known property of the maximal monotone operator [23]. According to Kenderov [11], a maximal monotone operator in a reflexive or separable Banach space is single-valued everywhere except a first-category set. Condition (iii) of Definition 2.1 follows Robert [22]. Also see [16]. Since F_h is monotone in $[x, x + h]$, we have condition (iv) of Definition 2.1. ■

Note that we use first-category sets to replace sets of measure zero in Banach spaces because of the single value of maximal monotone operators outside these sets. In finite dimensions maximal monotone operators are also single-valued except for a set of measure zero [14].

Definition 2.5 Suppose that $f: Y \to \mathbf{R}$ is a locally Lipschitz function. If $F = \partial f$ is a maximal normal operator, then we call f a primal function. ■

Example 2.6 If F is a maximal normal operator and a is a real number, then aF is also a maximal normal operator. Certainly, a continuously differentiable function is a primal function. If X is either separable or reflexive, then convex functions and concave function are also primal functions. Later we will give more examples of primal functions, such as the differences of two convex functions. ■

3. Cyclical Normality

A maximal monotone operator is the subdifferential of a convex function if and only if it is cyclically monotone [2]. We now extend this to the general case.

We call an open set Y linearly connected if for any two points x and y in Y, there is a finite number of line segments $[x_0 = x, x_1]$, $[x_1, x_2]$,..., $[x_{k-1}, x_k = y]$ in Y to connect these two points.

<u>Definition 3.1</u> A maximal normal operator F, defined on a linearly connected open set Y, F: Y $\Rightarrow$ X*, is called cyclically normal on Y if for any cycle of line segments $[x_0, x_1]$, $[x_1, x_2]$,..., $[x_{k-1}, x_k]$, $[x_k, x_0]$ in Y,

$$\sum_{i=0}^{k} \int_0^1 \langle F(x_i + t(x_{i+1} - x_i)), x_{i+1} - x_i \rangle dt = 0,$$

where $x_{k+1} = x_0$ and the integral is meaningful since the integrand is single-valued almost everywhere in the sense of one-dimensional Lebesque's measure. ■

<u>Theorem 3.2</u> A maximal normal operator F, defined on a linearly connected open set Y, is the subdifferential of a primal function f, if and only if F is cyclically normal. In this case, f is determined up to a constant. Suppose that x_0 is a fixed point in Y, and x is any point in Y such that there line segments $[x_0, x_1]$, $[x_1, x_2]$,..., $[x_{k-1}, x_k = x]$ in Y to connect x_0 and x. Then

$$f(x) = f(x_0) + \sum_{i=1}^{k} \int_0^1 \langle F(x_{i-1} + t(x_i - x_{i-1})), x_i - x_{i-1} \rangle dt. \qquad (6)$$

<u>Proof</u> If F = ∂f, then

$$\langle F(x_i + t(x_{i+1} - x_i), x_{i+1} - x_i \rangle = \partial Q(t),$$

where $Q(t) = f(x_i + t(x_{i+1} - x_i))$ for $t \in (0, 1)$, and $\partial Q(t) = Q'(t)$ whenever it is single-valued. Thus,

$$\sum_{i=0}^{k} \int_0^1 \langle F(x_i + t(x_{i+1} - x_i)), x_{i+1} - x_i \rangle dt$$

$$= \sum_{i=0}^{k} \int_0^1 \frac{d}{dt} (f(x_i + t(x_{i+1} - x_i))dt$$

$$= \sum_{i=0}^{k} (x_{i+1} - x_i) = 0,$$

where $x_{k+1} = x_0$, for any $[x_0, x_1], x_1, x_2], \ldots, [x_k, x_0]$ in Y, i.e., F is cyclically normal. If F is cyclically normal, we may define a function f by (6), where $f(x_0)$ is a constant. From the lcoal boundedness of F, we see that f is locally Lipschitz. For any y in Y and any h in X, [y, y + th] is in Y for a small t. By the cyclical normality of F and the definition of f,

$$f(y + th) - f(y) = t \int_0^1 \langle F(y + sth), h \rangle ds = t \int_0^1 F_h(y + sth)ds.$$

Thus, for any x in Y and any h in X,

$$\sup\{\langle u, h\rangle : u \in \partial f(x)\} = f^\circ(x; h)$$

$$= \limsup_{\substack{y \to x \\ t \downarrow 0}} [f(y + th) - f(y)]/t = \limsup_{\substack{y \to x \\ t \downarrow 0}} \int_0^1 F_h(y + sth)ds$$

$$\leq \limsup_{Z \to x} \{F_h(Z): F_h(Z) \text{ is single-valued}\}$$

$$\leq \sup \{\langle u, h\rangle : u \in F(x)\}.$$

Therefore,

$$F(x) \supseteq \partial f(x).$$

Especially, if $x \in$ sing F, then $F(x) = \partial f(x)$. By Propositions 2.1.5 and 2.1.2 of

[4], ∂f is w*-closed and $\partial f(x)$ is convex for any x. According to (iii) of Definition 2.1, for any x in Y,

$$F(x) = w^*\text{-cl conv } \{u \in x^*: u = w^*\text{-lim } \partial f(x_i),\ x_i \in \text{sing } F,\ x_i \to x\}$$

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

Hence $F(x) = \partial f(x)$ for all x in Y, i.e., $F = \partial f$. Clearly, such f is determined up to a constant $f(x_0)$. ■

4. Operator Sum

Suppose that F and G are two maximal normal operators on Y.

Definition 4.1 The operator sum $H = F \oplus G$ is defined by

$$H(x) = F(x) + G(x), \text{ if } x \in \text{sing } F \cap \text{sing } G.$$

$$H(x) = w^*\text{-cl conv } \{u \in x^*: u = w^*\text{-lim } H(x_i),\ x_i \in \text{sing } F \cap \text{sing } G,\ x_i \to x\},$$

otherwise. ■

Theorem 4.2 If F and G are two maximal normal operators on Y, then their operator sum $H = F \oplus G$ is also a maximal normal operator on Y.

Proof The local boundedness follows the definition of the operator sum and the uniform boundedness theorem. Clearly,

$$\text{sing } H \supseteq (\text{sing } F \cap \text{sing } G),$$

i.e.,

$$Y \setminus \text{sing } H \subseteq (Y \setminus \text{sing } G) \cup (Y \setminus \text{sing } G)$$

is also a first-category set. For any $u = w^*\text{-}\lim H(x_i)$, $x_i \in \operatorname{sing} H$, $x_i \to x$, we may choose y_i very close to x_i such that $y_i \in \operatorname{sing} F \cap \operatorname{sing} G$, $y_i \to x$ and $u = w^*\text{-}\lim H(y_i)$. Thus,

$$H(x) = w^*\text{-cl conv } \{u \in x^*: u = w^*\text{-}\lim H(x_i),\ x_i \in \operatorname{sing} H,\ x_i \to x\},$$

for x not in $\operatorname{sing} F \cap \operatorname{sing} G$. However, this is also clearly true for u in $\operatorname{sing} F \cap \operatorname{sing} G$. Finally, for any x in Y and any h in X such that $F_h(x)$ and $G_h(x)$ are single-valued, obviously $H_h(x) = F_h(x) + G_h(x)$. This proves (iv) of Definition 2.1. Therefore, H is also a maximal normal operator. ■

<u>Corollary 4.3</u> All the maximal normal operators on Y form a linear space with the operator addition $\oplus$ and the scalar multiplication. ■

We call this space the maximal normal operator space on Y. We may also define the operator difference of two maximal normal operators F and G by

$$F \ominus G: = F \oplus (-G).$$

It is not difficult to see that all the continuous operators on Y, all the cyclically maximal normal operators on Y, all the differences of two maximal monotone operators in the case that X is either separable or reflexive, form three distinct subspaces of the maximal normal operator space on Y. They do not contain each other. See discussions on the finite-dimensional case [17] [18] [19]. We may also define semicontinuous operators as in the finite-dimensional case [19].

As in the finite-dimensional case, the introduction of the operator sum of maximal normal operators makes the basic calculus of the Clarke subdifferential have equalities.

<u>Theorem 4.4</u> Suppose that f,g: $Y \to \mathbf{R}$ are two primal functions. Then their sum and product are also primal functions and

(a) $\partial(f + g) = \partial f \oplus \partial g$;

(b) $\partial(f \bullet g) = f\partial g \oplus g\partial f$.

Furthermore, if $g(x) \neq 0$ for any x in Y, then f/g is also a primal function and

(c) $\partial(f/g) = [g\partial f \oplus (-f\partial g)]/g^2$.

<u>Proof</u> Because of similarity, we only prove (a). If x is in sing $\partial f \cap$ sing ∂g, then

$$\partial(f + g)(x) = \partial f(x) + \partial g(x) = (\partial f \oplus \partial g)(x).$$

Suppose that x is not in sing $\partial f \cap$ sing ∂g. Since $\partial(f + g)$ is w*-closed and $\partial(f + g)(x)$ is convex,

$$(\partial f \oplus \partial g)(x)$$

$$= \quad \text{w*-cl conv } \{u \in X^*: u = \text{w*-lim } (\partial f \oplus \partial g)(x_i),\ x_i \in \text{sing } \partial f \cap \text{sing } \partial g,\ x_i \to x\},$$

$$\subseteq \quad \text{w*-cl conv } \{u \in X^*: u = \text{w*-lim } \partial(f + g)(x_i),\ x_i \in \text{sing } \partial f \cap \text{sing } \partial g,\ x_i \to x\}$$

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

On the other hand, for any h in X,

$$\sup \{\langle u, h\rangle : u \in \partial(f + g)(x)\} = (f + g)^{\circ}(x)$$

$$= \limsup_{\substack{y \to x \\ t \downarrow 0}} [(f+g)(y+th) - (f+g)(y)]/t$$

$$= \limsup_{\substack{y \to x \\ t \downarrow 0}} \int_0^1 (\partial(f+g))_h \, (y + sth)ds$$

$$= \limsup_{\substack{y \to x \\ t \downarrow 0}} \int_0^1 (\partial f \oplus \partial g)_h \, (y + sth)ds$$

$$\leq \limsup_{z \to x} \{\partial f \oplus \partial g)_h(z): \text{ it is single-valued}\}$$

$$\leq \sup \{<u, h>: u \varepsilon (\partial f \oplus \partial g)(x)\}.$$

Thus, $\partial(f + g)(x) \subseteq (\partial f \oplus \partial g)(x)$. Therefore, the equality holds, i.e.,

$\partial(f + g) = \partial f \oplus \partial g$. ■

Corollary 4.5 All the primal functions on Y form a linear space. The Clarke subdifferential operator ∂ is a linear operator from the primal function space onto the cyclically maximal normal operator space. Especially, the difference of two convex functions is also a primal function. ■

A function is called a d.c. function if it can be locally expressed as the difference of two convex functions [10]. Since the maximal normal operator is locally defined, a d.c. function is a primal function too.

In the finite-dimensional space, we proved in the sense of the Lebesgue measure that a semismooth function is a primal function [19]. In [21], Ralph proved that the Euclidean distance function of a set C is a primal function if and only if the boundary of the closure of C has zero measure. An open questions is whether these results can be extended into infinite-dimensional cases in the context of first-category sets. Notice that the boundary of the closure of a set is a first-category set.

5. Operator Integral in Finite-Dimensional Spaces

In the last section, we studied the operator sum of two maximal normal operators, which makes the basic calculus of the Clarke subdifferential have equalities. We now study the operator integral of a family of maximal normal operators, which will be taken over a positive σ-finite measure space $(T, \mathcal{S}, \mu)$:

$$\text{op-}\int_T F_t \, \mu(dt),$$

where for each t, $F_t: Y \twoheadrightarrow X^*$ is a maximal normal operator and,

Assumption 5.1 For each x in Y, the set $D=\{t:F_t(x) \text{ is single-valued}\}$ is measurable and the map $t \twoheadrightarrow F_t(x)$ from D to X is measurable relative to the restriction of μ to D. ■

This study aims to establish the following formula:

$$\partial \int_T f_t \, \mu(dt) = \text{op-}\int_T \partial f_t \, \mu(dt), \qquad (7).$$

where for each t, $f_t: Y \twoheadrightarrow \mathbf{R}$ is a primal function.

We first discuss in the case that X is finite-dimensional. As said in the introduction, the left hand side may be the subdifferential of the expectation functional term in the objective function of stochastic programming with recourse. Finding this subdifferential is necessary for subgradient-based methods, such as the bundle method, stochastic quasigradient method and other generalized gradient methods.

The definition of maximal normal operators in finite-dimensional spaces is based on the Lebesgue measure [17] [18]. Suppose that Y is an open set in $\mathbf{R}^n$.

Definition 5.2 A set-valued operator $F: Y \to \mathbf{R}^n$ is called a maximal normal operator if

(i) F is locally bounded on Y;

(ii) F is single-valued almost everywhere in the sense of the Lebesque measure on Y;

(iii) for any x in Y,

$$F(x) = \text{cl conv} \{u: u = \lim F(x_i),\ x_i \in \text{sing } F,\ x_i \to x\}. \qquad ■$$

All the results in Sections 2-4 can be derived from this definition in the finite-dimensional spaces and in fact, we have results on semismooth functions, quasidifferentials and Euclidean distance functions. See [17] [18] [19] [20] [21]. Similarly, a locally Lipschitz function is called a primal function if its subdifferential is a maximal normal operator. We now discuss the operator integral based upon Definition 5.2 and given Assumption 5.1.

Definition 5.3 The operator integral $H = \text{op-}\int_T F_t\, \mu(dt)$ is defined by

$$H(x) = \int_T F_t(x)\, \mu(dt) \tag{8}$$

if $F_t(x)$ is single-valued almost everywhere in the sense of the measure μ in T. Denote the set of such x as W. Then for other x in Y,

$$H(x) = \text{cl conv}\,\{u: u = \lim H(x_i),\ x_i \in W,\ x_i \to x\}. \qquad \blacksquare$$

Lemma 5.4 The set $Y \backslash W$ has measure zero.

Proof Let $C = \{(t, x): F_t(x)$ is not single-valued$\}$. Then for each t, $C(t) = \{x \in Y: (t, x) \in C\}$ has measure zero. Thus, by the Fubini Theorem, for almost every x, $C(x) := \{t \in T: (t, x) \in C\}$ has measure zero, i.e., almost every x is in W. This establishes the lemma. $\blacksquare$

Theorem 5.5 Suppose that

(a) for each $t \in T$, $f_t: Y \to \mathbf{R}$ is a primal function;

(b) for each x in Y, the map $t \to f_t(x)$ is measurable;

(c) for some $k(\cdot) \in L^1(T, \mathbf{R})$ (the space of integrable functions from T to $\mathbf{R}$), for all x and y in Y and t in T, one has

$$|f_t(x) - f_t(y)| \leq k(t)\, \|x - y\|.$$

Then the integral functional f on Y given by

$$f(x) = \int_T f_t(x)\mu(dt)$$

is also a primal function and

$$\partial f = \partial \int_T f_t\, \mu(dt) = \text{op-}\int_T \partial f_t\, \mu(dt).$$

Proof By Theorem 2.7.2 of [4], f is a Lipschitz function on Y and for each x in Y,

$$\partial f(x) \subseteq \int_T \partial f_t(x)\mu(dt).$$

Thus, when $w \in W$, $\partial f(x) = H(x)$, where $H = \text{op-}\int_T \partial f_t\, \mu(dt)$.

By Lemma 5.3, ∂f is single-valued almost everywhere on Y. By Proposition 5 of [18], f is a primal function, i.e., H is a maximal normal operator. Thus $\partial f = H$. This proves the theorem. ∎

6. Operator Integral in Banach Spaces

We now study the operator integral of a family of maximal normal operators in Banach spaces based upon Definition 2.1. There suppose that $(T, \mathfrak{S}, \mu)$ is a positive complete σ-finite measure space, that for each t, $F_t: Y \rightarrow X^*$ is a maximal normal operator under Definition 2.1, where X is a Banach space, and that Assumption 5.1 holds.

Definition 6.1 The operator integral $H = \text{op-}\int_T F_t\, \mu(dt)$ is defined by

$$H(x) = \int_T F_t(x)\, \mu(dt)$$

if $F_t(x)$ is single-valued almost everywhere in the sense of the measure μ in T. Denote the set of such x as W. Then for other x in Y,

$$H(x) = w^*\text{-cl conv } \{u \in X^*;\ u = w^*\text{-lim } H(x_i),\ x_i \in W,\ x_i \rightarrow x\}.$$ ∎

A difference from the finite-dimensional case is that the Fubini theorem cannot be applied directly to ensure that Y/W is a first-category set. If a homemorphism exists from X onto a complete measure space that maps the first category sets into sets of measure zero, then the Fubini Theorem may again be applied. Unfortunately, such mappings are often impossible to find. An alternative is to define a topology on T in which sets of measure zero and first-category sets agree. This is possible given our assumptions on $(T, \mathfrak{I}, \mu)$ (See Chapter 22 of [15]). The corresponding topology $\mathfrak{I} = \{\phi(A)\backslash N : A \in \mathfrak{I}, \mu(N) = 0\}$, where a is a lower density mapping.

<u>Lemma 6.2</u> If X is separable and, there exists a sequence, $\{C_n\}$ of nonempty open sets in $\mathfrak{I}$ such that every nonempty open set of $\mathfrak{I}$ contains some C_n, then $Y\backslash W$ is a first-category set.

<u>Proof</u> Define C,C(t) and C(x) as in the proof of Lemma 5.4. For each t, C(t) is a first-category set, and, for each x, C(x) is measurable. By Theorems 22.4, 22.5 and 22.6 of [15], we can introduce a topology $\mathfrak{I}$ in T, in which, the nowhere dense sets and sets of measure zero agree. By Theorem 22.7 of [15], C(x) has the property of Baire with respect to $\mathfrak{I}$. Hence, C has the property of Baire in the product topology.

By Theorem 15.4 in [15], if C(t) is a first-category set and X is separable, then C is a first-category set. The Kuratowski-Ulam Theorem (15.1 in [15]) under the assumptions on $\mathfrak{I}$, implies that C(x) is of first-category for all x except a set of first-category. All first category sets in $\mathfrak{I}$ are countable unions of nullsets. Hence, C(x) has measure zero except for a set of first-category. ■

Theorem 6.3 Suppose that conditions (a), (b) and (c) of Theorem 5.4 are satisfied. Denote $F_t = \partial f_t$. Let W be the set as defined in Definition 6.1. If either (d) T is countable,

or (e) the conditions in Lemma 6.2 hold

or (f) $Y \setminus W$ is a first-category set, T is a separable metric space, μ is a regular measure and the mapping $t \rightarrow \partial f_t(x)$ is upper semicontinuous (w*-) for each x in Y,

then the conclusions of Theorem 5.5 hold. ∎

The proof is similar to the proof of Theorem 5.5. Condition (d) or (e) or (f) is used to invoke Theorem 2.7.2 of [4].

Acknowledgment

The authors would like to thank Professor Frank Forelli for his useful comments and suggestions.

REFERENCES

[1] J.R. Birge and L. Qi, Computing block-angular karmarkar projections with applications to stochastic programming, Management Science (to appear).

[2] H. Brézis, Operateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert, North-Holland, Amsterdam, 1973.

[3] F.H. Clarke, Generalized gradients and applications, Trans. Amer. Math. Soc. 205 (1975), 247-262.

[4] F.H. Clarke, Optimization and Nonsmooth Analysis, Wiley, New York, 1983.

[5] F.H. Clarke, Periodic solutions of Hamilton's equations and local minima of the dual action, Trans. Amer. Math. Soc. 287 (1985), 239-254.

[6] F.H. Clarke, The dual action, optimal control, and generalized gradients, Math. Control Theory 14 (1985), 109-119.

[7] F.H. Clarke, Hamiltonian analysis of the generalized problem of Bolza, Trans. Amer. Math. Soc. 301 (1987), 385-400.

[8] F.H. Clarke and R.B. Vinter, Regularity properties of solutions to the basic problem in the calculus of variations. Trans. Amer. Math. Soc. 289 (1985), 73-98.

[9] Y. Ermoliev, On the stochastic quasi-gradient method and stochastic quasi-Feyer sequences, Cibernetika 2 (1969), 208-220.

[10] J.B. Hiriart-Urruty and H. Tuy, Essays on Nonconvex Optimization, North-Holland, Amsterdam, 1988.

[11] P.S. Kenderov, The set-valued monotone mappings are almost everywhere single-valued, C. R. Acad. Bulgare Sci. 27 (1974), 1173-1175.

[12] C. Lemarechal, Bundle methods in nonsmooth optimization, Nonsmooth Optimization (C. Lemarechal and R. Mifflin eds.) Pergamon Press, New York, 1978, pp. 79-102.

[13] R. Mifflin, Semismooth and semiconvex functions in constrained optimization, SIAM J. Control. & Optimization 15 (1977), 959-972.

[14] F. Mignot, Contrôl dans les inéquations variationelles elliptiques, J. Functional Analysis 22 (1976), 130-185.

[15] J. C. Oxtoby, Measure and Category, Springer-Verlag, Berlin, 1980 .

[16] L. Qi, Uniqueness of the maximal extension of a monotone operator, Nonlinear Analysis 7 (1983), 325-332.

[17] L. Qi, The maximal normal operator space and integration of subdifferentials of nonconvex functions, Nonlinear Analysis (to appear).

[18] L. Qi, Maximal normal operators and the second-order derivatives of nonsmooth functions, Industrial Engineering Technical Report 88-3, University of Wisconsin-Madison, Madison, 1988.

[19] L. Qi, Semismoothness and decomposition of maximal normal operators, Journal of Math. Anal. & Applications (to appear).

[20] L. Qi, Quasidifferentials and maximal normal operators. Industrial Engineering Technical Report, 88-9, University of Wisconsin-Madison, Madison, 1988.

[21] D. Ralph, Strict differentiability and primality of the Euclidean distance function, manuscript.

[22] R. Robert, Contributions a l'analyse non lineaire, These d'Etat, Grenoble, 1976.

[23] R.T. Rockafellar, Local boundedness of nonlinear, monotone operators, Mich. Math. 5 16 (1969), 397-407.

[24] R.T. Rockafellar, The Theory of Subgradients and Its Applications to Problems of Optimization. Convex and Nonconvex Functions, Holderman-Verlag, West Berlin, 1981.

[25] N.Z. Shor, Generalize gradient methods of non-differentiable optimization employing space dilation operations, Mathematical Programming: State of the Art 1982 (A. Bachem, M. Grötschel and B. Korte eds.) Springer-Verlag, Berlin, 1983, pp. 501-529.

[26] R.J.-B. Wets, Stochastic programming: solution techniques and approximation scheme, Mathematical Programming: State of the Art 1982 (A. Bachem, M. Grötschel and B. Korte eds.) Springer-Verlag, Berlin, 1983, pp. 566-603.