ANNALES DE L'I. H. P., SECTION C

R. T. ROCKAFELLAR

Maximal monotone relations and the second derivatives of nonsmooth functions

Annales de l'I. H. P., section C, tome 2, nº 3 (1985), p. 167-184

<http://www.numdam.org/item?id=AIHPC_1985__2_3_167_0>

© Gauthier-Villars, 1985, tous droits réservés.

L'accès aux archives de la revue « Annales de l'I. H. P., section C » (http://www.elsevier.com/locate/anihpc) implique l'accord avec les conditions générales d'utilisation (http://www.numdam.org/conditions). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/

Ann. Inst. Henri Poincaré,
Vol. 2, n° 3, 1985, p. 167-184. *Analyse non linéaire*

Maximal monotone relations and the second derivatives of nonsmooth functions

by

R. T. ROCKAFELLAR (*)

ABSTRACT. — Maximal monotone relations serve as a prototype from which properties can be derived for the subdifferential relations associated with convex functions, saddle functions, and other important classes of functions in nonsmooth analysis. It is shown that the Clarke tangent cone at any point of the graph of a maximal monotone relation is actually a linear subspace. This fact clarifies a number of issues concerning the generalized second derivatives of nonsmooth functions.

Key Words: Nonsmooth analysis, convex functions, maximal monotone relations, generalized second derivatives, Lipschitzian manifolds.

RÉSUMÉ. — Nous déduisons, grâce à l'étude de multi-applications monotones maximales, diverses propriétes relatives aux multi-applications sous-gradients de fonctions convexes, fonctions de selle et autres types de fonctions importants en analyse sous-différentielle. Nous montrons, qu'en tout point du graphe d'une multi-application monotone maximale, le cône tangent, au sens de Clarke, est en réalité un sous-espace vectoriel. Ce fait éclaircit certaines questions concernant les dérivées généralisées du second ordre de fonctions non-différentiables.

Mots-clés : Analyse sous-différentielle, fonctions convexes, multi-applications monotones maximales, dérivées généralisées de second ordre, variétés lipschitziennes.

(*) Supported in part by a grant from the National Science Foundation at the University of Washington, Seattle.

1. INTRODUCTION

Generalized theories of differentiation in convex analysis [*20*] and more recently the nonsmooth analysis of Clarke [*8*] associate with an extended-real-valued function f on R^n a multifunction (set-valued mapping) ∂f with graph in $R^n \times R^n$. The elements of $\partial f(x)$ are called the subgradients or generalized gradients of f at x, and they are used in characterizing first-order derivative properties of f such as are important especially in the analysis of problems of optimization. Since second-order properties could be useful in such analysis too, there have been various attempts to extend the operation of subdifferentiation from f to ∂f. No simple approach has seemed entirely satisfying, however, so this area of research is still in a state of flux. The purpose of the present article is to establish a number of facts that should help to clarify the situation and shed light on the limits of the possible.

Convex functions have been the main focus for work on generalized second derivatives. The classical theorem of Alexandrov [*1*] says that a finite convex function on an open convex set is twice differentiable almost everywhere in the sense of having a second-order Taylor's expansion. Alexandrov's proof is couched in a geometric language that is nowadays hard to follow, but the same thing has been proved in terms of the theory of distributions by Reshetniak [*20*]. It is closely connected with a result of Mignot [*16*, Theorem 1.3] according to which a maximal monotone relation is once differentiable almost everywhere on the interior of its effective domain. Indeed, when $f: R^n \to R \cup \{+\infty\}$ is convex, lower semicontinuous and proper (not identically $+\infty$), the subdifferential relation ∂f is a maximal monotone relation whose effective domain includes the interior of the convex set $\operatorname{dom} f = \{x \mid f(x) < \infty\}$ (see [*20*, § 24]).

The drawback with twice differentiability in this classical sense, of course, is that it tells us nothing about the behavior of f at boundary points of dom f or interior points where f has a « kink ». Such are just the kinds of points where the minimum of f may occur, so efforts have been made to include them in some generalized definition of second derivative. One approach, followed by Lemarechal and Nurminski [*15*], Auslender [*6*], and Hiriart-Urruty [*13*] [*14*] has been to exploit certain properties of the support function of the ε-subdifferential $\partial_\varepsilon f(x)$ of f at x as $\varepsilon \downarrow 0$. This idea is motivated by computational considerations, but it only leads in general to « approximate » second derivatives, and it is limited in concept to the case of f convex.

Another approach has been to consider tangent cones of various kinds to the graph of ∂f and view these as the graphs of derivative relations. This approach has been pioneered by Aubin [*5*], who has observed in particular that the Clarke tangent cone at a point of the graph of ∂f is the graph of a closed convex process which, if f is convex, is also a monotone relation. Aubin has used this concept along with surjectivity conditions to derive Lipschitz stability for the optimal solutions to a parameterized class of convex optimization problems in the Fenchel duality format.

The results in this paper will show that the kind of derivative multifunction that occurs in applications of this second approach is actually the inverse of a linear transformation. One conclusion to be drawn is that the cases where this approach works are more special than has been realized. On the other hand, it will be seen that the properties in such cases are also much stronger than reported. Furthermore the proposed derivatives for ∂f can be characterized in terms of limits of second-order difference quotients for f that can make sense even at points where f is not smooth or continuous.

The plan of the paper is to treat first the graphs of maximal monotone relations and related multifunctions which can be regarded as Lipschitzian manifolds of a certain sort. The results are then applied to the subdifferentials of convex functions and tied to second derivative properties of the functions themselves. So-called lower-C^2 (strongly subsmooth) functions are covered at the same time. Extensions to saddle functions, and other functions that occur as the Lagrangians in optimization problems with constraints or perturbations, would be possible, but we do not pursue them here, due to lack of space. For simplicity we limit attention to R^n, although most of the results have some infinite-dimensional analogue, at least in a separable Hilbert space.

2. LIPSCHITZIAN MANIFOLDS

A function $F : U \to R^m$, where U is open in R^n, is said to be *Lipschitzian* (*with modulus* γ) if $|F(u') - F(u)| \leqq \gamma |u' - u|$ for all u and u' in U. The classical theorem of Rademacher [*26*] asserts that such a function is differentiable almost everywhere: for almost every $u \in U$ there is a linear transformation $A = \nabla F(u)$ such that

$$F(u') = F(u) + A(u' - u) + o(|u' - u|).$$

This obviously says something about the geometry of the graph set

$$\text{gph } F = \{ (u, v) \mid u \in U, v = F(u) \} \subset R^{m+n}.$$

Our aim is to utilize such geometry in the study of certain important classes of sets that may not at first appear to be the graphs of Lipschitzian functions but can be interpreted as such through a change in coordinates.

The following concept will be useful. A subset M of R^N will be called a a *Lipschitzian manifold* if it is locally representable as the graph of a Lipschitzian function in the sense that: for every $\bar{x} \in M$ there is an open neighborhood X of $\bar{x}$ in R^N and a one-to-one mapping Φ of X onto an open set in $R^n \times R^m$ (where $n + m = N$) with Φ and Φ^{-1} continuously differentiable, such that $\Phi(M \cap X)$ is the graph of some Lipschitzian function $F : U \to R^m$, where U is some open set in R^n. Clearly the Φ and F in this definition are not uniquely determined by M and $\bar{x}$, but the integer n is. It is the dimension of M around $\bar{x}$, and it must in fact be the same for all $\bar{x} \in M$ if M is connected, in which case one can appropriately speak of M as a Lipschitzian manifold of *dimension n* in R^N. An immediate example is the following.

PROPOSITION 2.1. — *If* $F : R^n \to R^m$ *is locally Lipschitzian, then the set* $M = \text{gph } F$ *is a Lipschitzian manifold of dimension n in* $R^n \times R^m$.

For a less obvious example that will be of great interest to us later, we recall the notion of maximal monotonicity. A relation or multifunction $D : R^n \rightrightarrows R^n$ (assigning to each $x \in R^n$ a subset $D(x) \subset R^n$ that might be empty) is said to be *monotone* (in the sense of Minty [*17*])if

$$(x_1 - x_2) \cdot (y_1 - y_2) \geqq 0 \quad \text{for all} \quad x_1, x_2, \quad \text{and} \quad y_1 \in D(x_1), \quad y_2 \in D(x_2).$$

It is *maximal* monotone if, in addition, its graph

$$\text{gph } D = \{ (x, y) \mid x \in R^n, \ y \in D(x) \}$$

is maximal, or in other words, if there does not exist another monotone relation $E : R^n \rightrightarrows R^n$ having $\text{gph } E \supset \text{gph } D$, $\text{gph } E \neq \text{gph } D$. (Every monotone relation can be extended to one which is maximal in this sense.)

The study of maximal monotone relations is closely connected with the study of subdifferentials of convex functions [*20*, § *24*], saddle functions [*20*, § *35*], lower-C^2 (strongly subsmooth) functions [*21*] [*22*] and other topics of importance in nonsmooth analysis and variational theory. Right now we need only mention the fact (see Minty [*17*]) that if D is maximal monotone, then the relation $P = (I + D)^{-1}$ is actually a single-valued mapping of all of R^n into R^n which is nonexpansive, i. e. globally Lipschitzian with modulus $\gamma = 1$. The following is essentially well known, although it has not previously been expressed in the language of Lipschitzian manifolds.

PROPOSITION 2.2. — *If* $D : R^n \rightrightarrows R^n$ *is maximal monotone, then the set* $M = \text{gph } D$ *is a Lipschitzian manifold of dimension n in* $R^n \times R^n$.

Proof. — Let $P = (I + D)^{-1}$ as above and also $Q = (I + D^{-1})^{-1}$. Then P and Q are both Lipschitzian (globally), since D^{-1} as well as D is maximal monotone. Furthermore one has

$$(x, y) = (P(u), Q(u)) \Leftrightarrow y \in D(x) \quad \text{and} \quad x + y = u\,.$$

(Indeed, for any u the vector $P(u)$ is the unique x such that $u \in (I + D)(x)$, i. e. $u - x \in D(x)$. Then $u - x$ must correspondingly be the unique y such that $u = y \in D^{-1}(y)$, i. e. $u - x$ must be $Q(u)$. In particular, $P + Q = I$.) Consider now the linear transformation

$$\Phi(x, y) = (x + y, x - y),$$

which is one-to-one from $R^n \times R^n$ onto $R^n \times R^n$. Trivially, Φ and Φ^{-1} are differentiable. The image of $M = \operatorname{gph} D$ under Φ is evidently the set of all pairs (u, v) such that $v = P(u) - Q(u)$. Thus it is the graph of the Lipschitzian function $F : R^n \to R^n$, where $F = P - Q = 2P - I$. This demonstrates that M fits the definition of Lipschitzian manifold. □

COROLLARY 2.3. — *Let* $F : R^n \to \overline{R}$ *be a closed proper convex function, and let* ∂f *be the subdifferential of* f *in the sense of convex analysis. Then the set* $M = \operatorname{gph} \partial f$ *is a Lipschitzian manifold of dimension* n *in* $R^n \times R^n$.

Proof. — The relation $D = \partial f$ is maximal monotone, as is well known; cf. [*20*, § 24]. □

COROLLARY 2.4. — *Let* $L : R^n \times R^m \to \overline{R}$ *be a closed proper saddle function (with* $L(x, y)$ *convex in* x*, concave in* y*), and let* ∂L *be the subdifferential of* L *in the sense of convex analysis. Then the set* $M = \operatorname{gph} \partial L$ *is a Lipschitzian manifold of dimension* $n + m$ *in* $(R^n \times R^m) \times (R^n \times R^m)$.

Proof. — The definition of ∂L is explained in [*20*, § 35] along with the concepts of « closed » and « proper » that are required here. It is shown in [*24*] that the linear transformation

$$(x, y, w, z) \to (x, y, w, -z)$$

transforms the graph of ∂L into a maximal monotone relation. □

One other case deserves explicit mention. Recall that a function $f : X \to R$ where X is open in R^n, is called lower-C^2 [*21*] (strongly subsmooth [*22*]) if each $\bar{x} \in X$ has a neighborhood on which f can be represented as a max of C^2 functions:

$$f(x) = \max_{s \in S} g_s(x),$$

where S is a compact topological space, each g_s is twice differentiable on the neighborhood in question, and the values of g_s and its first and second partial derivatives are continuous jointly in x and s. (Here S could in particular be a finite set in the discrete topology.)

COROLLARY 2.5. — *Let* $f : X \to R$ *be lower-*C^2*, where* X *is open in* R^n*, and let* ∂f *be the Clarke subdifferential of* f*. Then the set* $M = \operatorname{gph} \partial f$ *is a Lipschitzian manifold of dimension* n *in* R^n.

Proof. — In a local « max representation » of f such as above, the Hessian matrices $\nabla^2 g_s(x)$ are continuous in s and x, and their eigenvalues are therefore uniformly bounded below as long as x remains in a compact set. Thus it is possible locally to find a value $\mu > 0$ such that when μI is added to all these matrices, one gets only positive definite matrices. The corresponding functions $g_s(x) + (\mu/2)|x|^2$ are then convex locally, and so is $f(x) + (\mu/2)|x|^2$. In other words, for any $\bar{x} \in X$ there is a $\mu > 0$ such that the function $f(x) + (\mu/2)|x|^2$ is convex on some neighborhood of $\bar{x}$. Choose a compact convex neighborhood C of $\bar{x}$ that is contained in a neighborhood of the type just mentioned. The function

$$h(x) = f(x) + (\mu/2)|x|^2 + \psi_C(x), \tag{2.1}$$

where ψ_C is the indicator of C, is then closed proper convex with

$$\partial h(x) = \partial f(x) + \mu x \quad \text{for all} \quad x \in \operatorname{int} C. \tag{2.2}$$

We know from Corollary 1 that gph ∂h is a Lipschitzian manifold of dimension n. The linear transformation $(x, y) \to (x, y + \mu x)$, which obviously is invertible, identifies the portion of gph ∂f lying over int C with the corresponding portion of gph ∂h. This shows that the portion of gph ∂f lying over int C is a Lipschitzian manifold of dimension n too. □

3. TANGENT SPACES

For a general closed set $M \subset R^n$ and a point $\bar{x} \in M$, there are two concepts of tangent cone that have received much attention in nonsmooth analysis. The *contingent cone* (or Bouligand tangent cone) is

$$K_M(\bar{x}) = \limsup_{t \downarrow 0} t^{-1}[M - \bar{x}],$$

whereas the *Clarke tangent cone* is

$$T_M(\bar{x}) = \liminf_{\substack{t \downarrow 0 \\ x(\in X) \to \bar{x}}} t^{-1}[M - x].$$

Both cones are always closed, but the Clarke tangent cone is also *convex* [*9*] [*23*]; moreover

$$T_M(\bar{x}) = \liminf_{x(\in M) \to \bar{x}} K_M(x)$$

(see Cornet [*11*] and Penot [*18*]).

The concepts of set limits that are employed here are the usual ones (see Salinetti and Wets [*27*], for instance): for a sequence of nonempty sets S_ν in R^N, one has

$$\limsup_{\nu \to \infty} S_\nu = \{ w \mid \exists w_\nu \in S_\nu \text{ such that } w \text{ is a cluster point of } \{ w_\nu \} \}, \tag{3.1}$$

$$\liminf_{\nu \to \infty} S_\nu = \{ w \mid \exists w_\nu \in S_\nu \text{ such that } w \text{ is the limit point of } \{ w_\nu \} \}. \tag{3.2}$$

(3.3) $$S = \lim_{\nu \to \infty} S_\nu \quad \text{if} \quad S = \limsup_{\nu \to \infty} S_\nu = \liminf_{\nu \to \infty} S_\nu .$$

The classical concept of a (linear) *tangent space* to M at $\bar{x}$ refers to a subspace S of R^N such that actually

(3.4) $$S = \lim_{t \downarrow 0} t^{-1}[M - \bar{x}]$$

(in which case $S = K_M(\bar{x})$ in particular). When such a subspace exists, we say M is *smooth* at $\bar{x}$. If actually

(3.5) $$S = \lim_{\substack{t \downarrow 0 \\ x(\in M) \to \bar{x}}} t^{-1}[M - x]$$

(in which case also $S = T_M(\bar{x})$), we say M is *strictly smooth* at $\bar{x}$.

In applying these concepts to a Lipschitzian manifold, we shall need to relate them to directional differentiability properties of a function $F : U \to R^m$, where U is open in R^n. A vector $\bar{k} \in R^m$ is the (one-sided) *directional derivative* (in the « Hadamard sense ») of F at a point $\bar{u} \in U$ with respect to a vector $\bar{h} \in R^n$ if

(3.6) $$\lim_{\substack{t \downarrow 0 \\ h \to \bar{h}}} \frac{F(\bar{u} + th) - F(\bar{u})}{t} = \bar{k} .$$

One then writes $F'(\bar{u}; \bar{h}) = \bar{k}$. It is said to be a *strict* directional derivative if actually

(3.7) $$\lim_{\substack{t \downarrow 0 \\ (u,h) \to (\bar{u},\bar{h})}} \frac{F(u + th) - F(u)}{t} = \bar{k} .$$

When F is Lipschitzian, the limits $h \to \bar{h}$ are superfluous in these formulas; one can then just take $h \equiv h$ without changing the limit values.

Differentiability of F at $\bar{u}$ means that $F'(\bar{u}; \bar{h})$ exists for all $\bar{h} \in R^n$ and is linear as a function of $\bar{h}$; the linear transformation $\bar{h} \to F'(\bar{u}; \bar{h})$ is what we denote by $\nabla F(\bar{u})$. We call F *strictly differentiable* at $\bar{u}$ if the same holds but $F'(\bar{u}; \bar{h})$ is a strict directional derivative for all $\bar{h}$.

PROPOSITION 3.1. — *Let* $F : U \to R^m$, *where* $U \subset R^n$ *is open. Let* $\bar{u} \in U$ *and* $\bar{v} = F(\bar{u})$, *so that* $(\bar{u}, \bar{v}) \in M := \text{gph}\, F$. *Then*

a) M *is smooth at* $(\bar{u}, \bar{v})$ *if and only if* F *is differentiable at* $\bar{u}$, *in which case the tangent space to* M *at* $(\bar{u}, \bar{v})$ *is* $S = \text{gph}\, \nabla F(\bar{u})$;

b) M *is strictly smooth at* $(\bar{u}, \bar{v})$ *if and only if* F *is strictly differentiable at* $\bar{u}$.

Proof. — These facts are classical in nature. Their proof, which is left to the reader, is just a matter of expressing classical notions in the language of set convergence (for which the article of Salinetti and Wets [27] provides appropriate tools). □

THEOREM 3.2. — *Let* $F : U \to R^m$ *be Lipschitzian, where* $U \subset R^n$ *is open. Let* $\bar{u} \in U$ *and* $\bar{v} = F(\bar{u})$, *so that* $(\bar{u}, \bar{v}) \in M := \text{gph } F$. *Then the Clarke tangent cone* $T_M(\bar{u}, \bar{v})$ *is not just a cone but a (linear) subspace of* $R^n \times R^m$.

One has $(\bar{h}, \bar{k}) \in T_M(\bar{u}, \bar{v})$ *if and only if* $F'(\bar{u}; \bar{h}) = \bar{k}$ *as a strict directional derivative. Moreover this is true if and only if*

$$\lim_{u(\in U') \to \bar{u}} \nabla F(u)\bar{h} = \bar{k}, \tag{3.8}$$

where $U' = \{ u \in U \mid F$ *is differentiable at* $u \}$.

Proof. — According to definition (3.2) one has $(\bar{h}, \bar{k}) \in T_M(\bar{u}, \bar{v})$ if and only if for every sequence $(u_\nu, v_\nu) \to (\bar{u}, \bar{v})$ in M and every sequence $t_\nu \downarrow 0$, there is a sequence $(h_\nu, k_\nu) \to (\bar{h}, \bar{k})$ with $(u_\nu, v_\nu) + t_\nu(h_\nu, k_\nu) \in M$ for all ν. Since $M = \text{gph } F$ with F continuous, this condition reduces to the following: For every sequence $u_\nu \to \bar{u}$ in U and every sequence $t_\nu \downarrow 0$, there is a sequence $h_\nu \to \bar{h}$ with

$$\frac{F(u_\nu + t_\nu h_\nu) - F(u_\nu)}{t_\nu} \to \bar{k}. \tag{3.9}$$

But F is actually Lipschitzian, so

$$| F(u_\nu + t_\nu h_\nu) - F(u_\nu + t_\nu \bar{h}) | \leqq \lambda t_\nu | h_\nu - \bar{h} |$$

for a certain modulus λ. The limit (3.9) is therefore unaffected if h_ν is replaced simply by $\bar{h}$. Thus the condition is: for every sequence $u_\nu \to \bar{u}$ in U and every sequence $t_\nu \downarrow 0$, one has

$$\frac{F(u_\nu + t_\nu \bar{h}) - F(u_\nu)}{t_\nu} = \bar{k}.$$

In other words, $(\bar{h}, \bar{k}) \in T_M(\bar{u}, \bar{v})$ if and only if

$$\lim_{\substack{t \downarrow 0 \\ u \to \bar{u}}} \frac{F(u + t\bar{h}) - F(u)}{t} = \bar{k}. \tag{3.10}$$

which again by the Lipschitz property is equivalent to (3.7), the defining condition for a strict directional derivative. Clearly too, (3.10) can be written as

$$\lim_{\substack{t \downarrow 0 \\ u' \to \bar{u}}} \frac{F(u') - F(u' - t\bar{h})}{t} = \bar{k}.$$

where $u' = u + t\bar{h}$, $u = u' - t\bar{h}$, so if $(\bar{h}, \bar{k}) \in T_M(\bar{u}, \bar{v})$ we must also have $(-\bar{h}, -\bar{k}) \in T_M(\bar{u}, \bar{v})$. We already know that $T_M(\bar{u}, \bar{v})$ is a convex cone, such being true always of the Clarke tangent cone, so we can conclude from this property that $T_M(\bar{u}, \bar{v})$ is actually a subspace.

Now we need to verify (3.8) as an alternative criterion for (3.10). One direction is easy: if (3.10) holds, then in particular (3.8) holds, inasmuch as

$$\nabla F(u)\bar{h} = F'(u;\bar{h}) = \lim_{t\downarrow 0} \frac{F(u+t\bar{h}) - F(u)}{t}$$

for $u \in U'$. The opposite direction of argument relies on Rademacher's theorem, i. e. the fact that U' differs from U by only a set of measure zero. We can take $\bar{h} \neq 0$ and assume for simplicity that U is a bounded open cylinder whose axis is in the direction of $\bar{h}$: There is an open interval I containing 0, such that every $u \in U$ is uniquely of the form $w + t\bar{h}$ for some $t \in I$ and w in the disk $D = \{ w \in U \mid w . \bar{h} = 0 \}$. The set of pairs $(w, t) \in D \times I$ such that $w + t\bar{h} \notin U'$ is then of measure zero, and so too must be its I cross-section for almost every $w \in D$ (as follows from Fubini's theorem when applied to the integral expressing the volume of the set in question). In particular, therefore, there is a dense subset D_0 of D such that for every $w \in D_0$ one has $w + t\bar{h} \in U'$ for almost every $t \in I$. Then the set $U_0 = \{ w + th \mid w \in D_0, t \in I \}$ is dense in U and has the property that for every $u \in U_0$, the set of $t \in R$ with $u + t\bar{h} \in U$ but $u + t\bar{h} \notin U'$ is of measure zero.

Invoking now the assumption that (3.8) holds, we consider arbitrary $\varepsilon > 0$ and choose a corresponding $\delta > 0$ such that $u' \in U$ when $|u' - \bar{u}| < \delta$ and

$$|\bar{k} - \nabla F(u')\bar{h}| \leqq \varepsilon \quad \text{when} \quad u' \in U', \qquad |u' - \bar{u}| < \delta .$$

Next we choose $\alpha > 0$ small enough that

$$|(u + \tau\bar{h}) - \bar{u}| < \delta \quad \text{when} \quad |u - \bar{u}| < \alpha \quad \text{and} \quad \tau \in [0, \alpha] .$$

Then the set $U_0^\alpha = \{ u \in U_0 \mid |u - \bar{u}| < \alpha \}$ is dense in a neighborhood of $\bar{u}$ and has the property that for all $u \in U_0^\alpha$, one has for almost every $\tau \in [0, \alpha]$ that $\nabla F(u + \tau\bar{h})$ exists and $|\bar{k} - \nabla F(u + \tau\bar{h})\bar{h}| \leqq \varepsilon$. The function $\tau \to F(u + \tau\bar{h})$ is itself Lipschitzian and therefore is the integral of its derivative, which is $\nabla F(u + \tau\bar{h})\bar{h})$ almost everywhere when $u \in U_0^\alpha$ and $t \in [0, \alpha]$, as just seen. Thus when $u \in U_0^\alpha$ and $\tau \in [0, \alpha]$ we have

$$F(u + t\bar{h}) = F(u) + \int_0^t \nabla F(u + \tau\bar{h})\bar{h} d\tau ,$$

so that

$$\bar{k} - \frac{F(u + t\bar{h}) - F(u)}{t} = \frac{1}{t}\int_0^t [\bar{k} - \nabla F(u + \tau\bar{h})\bar{h}] d\tau$$

with $|\bar{k} - \nabla F(u + \tau\bar{h})\bar{h}| \leqq \varepsilon$ in the integrand. It follows that

$$\left| \bar{k} - \frac{F(u + t\bar{h}) - F(u)}{t} \right| \leqq \varepsilon$$

for all $t \in [0, \alpha]$ and $u \in U_0^\alpha$, hence by continuity of F also for u in the closure of U_0^α, i. e. whenever $|u - \bar{u}| \leqq \alpha$. Since ε was arbitrary, this demonstrates that (3.10) must hold. Thus (3.8) does imply (3.10). □

COROLLARY 3.3. — *Let* $F : U \to R^m$, *where* $U \subset R^n$ *is open. If* $\bar{u} \in U$ *is such that* $F'(\bar{u}; \bar{h})$ *exists as a strict directional derivative for every* $\bar{h} \in R^n$, *then* F *is in fact strictly differentiable at* $\bar{u}$.

Proof. — An elementary compactness argument demonstrates that if the limit in (3.7) exists for every $\bar{h}$, then the difference quotients in (3.7) must be uniformly bounded in norm when $\bar{h}$ belongs to the unit ball and t belongs to an interval $(0, \varepsilon)$. This implies F is locally Lipschitzian. When F is Lipschitzian around $\bar{u}$, however, Theorem 3.2 is applicable and says that the set

$$\{ (\bar{h}, \bar{k}) \mid F'(\bar{u}; \bar{h}) = \bar{k} \text{ strictly} \}$$

is a subspace of $R^n \times R^m$. A subspace of such special type is the graph of a linear transformation from R^n to R^m if and only if its projection in the first argument is all of R^n. □

COROLLARY 3.4. (Clarke [9]). — *Let* $F : U \to R^m$ *be Lipschitzian, where* $U \subset R^n$ *is open. Let* $\bar{u} \in U$ *and let* $U' = \{ u \in U \mid F$ *is differentiable at* $u \}$. *Then* F *is strictly differentiable at* $\bar{u}$ *if and only if* $\bar{u} \in U'$ *and the mapping* $\nabla F : u \to \nabla F(u)$ *is continuous at* $\bar{u}$ *relative to* U'.

Proof. — This follows from Corollary 3.3 and the equivalences in Theorem 3.2. □

The main consequences of these results for Lipschitzian manifolds will now be stated.

THEOREM 3.5. — *Let* M *be a Lipschitzian manifold of dimension n in* R^N, *and let* M′ *be the set of points* $x \in M$ *where* M *is smooth, i. e. actually has a tangent space* $S_M(x)$. *Then*

a) M′ *differs from* M *by only a set of measure zero (with respect to n-dimensional Hausdorff measure), and for every* $x \in M'$ *the tangent space* $S_M(x)$ *is of dimension n.*

b) At every $\bar{x} \in M$ *the Clarke tangent cone* $T_M(\bar{x})$ *is actually a subspace of dimension no greater than n, namely*

$$T_M(\bar{x}) = \liminf_{\bar{x}(\in M') \to x} S_M(x).$$

c) M *is strictly smooth at* $\bar{x}$ *if and only if* $T_M(\bar{x})$ *has dimension n. This is true if and only if* $\bar{x} \in M'$ *and the mapping* $S_M : x \to S_M(x)$ *is continuous at* $\bar{x}$ *relative to* M′.

Proof. — Representing M locally as the graph of a Lipschitzian function on an open set in R^n, as is possible by definition, we get (*a*) as a consequence of Proposition 3.1 (*a*). Then (*b*) follows from Theorem 3.2, while (*c*) follows from Corollary 3.4 in combination with Proposition 3.1 (*b*). (For a sequence of n-dimensional subspaces of R^{N}, the « lim inf » cannot be n-dimensional unless it is actually a « lim ».) □

COROLLARY 3.6. — *Let* $\mathrm{M} = \mathrm{gph}\, \mathrm{D}$, *where* $\mathrm{D} : \mathrm{R}^n \to \mathrm{R}^n$ *is a maximal monotone relation or one of the subdifferential multifunctions considered in Corollaries 2.3, 2.4 or 2.5. Then the assertions* (*a*), (*b*) *and* (*c*) *of Theorem 3.5 hold for* M.

In the case of subdifferentials, Corollary 3.6 has important implications for the theory of second derivatives of nonsmooth functions. These will be traced in the next section. Another consequence can be stated immediately, however. To do this we introduce for a multifunction $\mathrm{D} : \mathrm{R}^n \to \mathrm{R}^n$ the notation

$$(3.11)\quad \mathrm{dom}\, \mathrm{D} = \{ x \mid \mathrm{D}(x) \neq \phi \}, \qquad \mathrm{rge}\, \mathrm{D} = \{ y \mid \exists x \quad \text{with} \quad y \in \mathrm{D}(x) \}.$$

THEOREM 3.7. — *Let* $\mathrm{D} : \mathrm{R}^n \to \mathrm{R}^n$ *be as in Corollary 3.6* (*or indeed, any multifunction whose graph is a Lipschitzian manifold of dimension n*). *Let* $\bar{x} \in \mathrm{dom}\, \mathrm{D}$ *and* $\bar{y} \in \mathrm{D}(\bar{x})$. *Let* $\mathrm{A} : \mathrm{R}^n \to \mathrm{R}^n$ *be the multifunction whose graph is the Clarke tangent cone* $\mathrm{T}_{\mathrm{M}}(\bar{x}, \bar{y})$, *where* $\mathrm{M} = \mathrm{graph}\, \mathrm{D}$.

a) *If* $0 \in \mathrm{int}\,(\mathrm{dom}\, \mathrm{A})$, *then* $\bar{x} \in \mathrm{int}\,(\mathrm{dom}\, \mathrm{D})$ *and* D *is single-valued and Lipschitzian on a neighborhood of* $\bar{x}$. *In fact* A *is a linear transformation, and* D *is strictly differentiable at* $\bar{x}$ *with* $\nabla \mathrm{D}(\bar{x}) = \mathrm{A}$.

b) *If* $0 \in \mathrm{int}\,(\mathrm{rge}\, \mathrm{A})$, *then* $\bar{y} \in \mathrm{int}\,(\mathrm{rge}\, \mathrm{D})$, *and* D^{-1} *is single-valued and Lipschitzian on a neighborhood of* $\bar{y}$. *In fact* A^{-1} *is a linear transformation, and* D^{-1} *is strictly differentiable at* $\bar{y}$ *with* $\nabla(\mathrm{D}^{-1})(\bar{y}) = \mathrm{A}^{-1}$.

Proof. — Obviously (*b*) is just the application of (*a*) to D^{-1}. To prove (*a*) we invoke the fact that M is an n-dimensional Lipschitzian manifold (cf. Proposition 2.2 and Corollaries 2.3, 2.4, and 2.5). According to Theorem 3.5 (*b*), $\mathrm{T}_{\mathrm{M}}(\bar{x}, \bar{y})$ is a subspace of $\mathrm{R}^n \times \mathrm{R}^n$ having dimension at most n. Since dom A is by definition the projection of $\mathrm{T}_{\mathrm{M}}(\bar{x}, \bar{y})$ in the first component, the condition $0 \in \mathrm{int}\,(\mathrm{dom}\, \mathrm{A})$ implies that the dimension of $\mathrm{T}_{\mathrm{M}}(\bar{x}, \bar{y})$ equals n. Then by Theorem 3.5 (*c*), M is strictly smooth at $(\bar{x}, \bar{y})$.

Consider now, as in the definition of « Lipschitzian manifold » at the beginning of § 2, a coordinate transformation Φ that represents M locally around $(\bar{x}, \bar{y})$ as the graph of a Lipschitzian function $\mathrm{F} : \mathrm{U} \to \mathrm{R}^n$ (with U open in R^n). Since M is strictly smooth at $(\bar{x}, \bar{y})$, so is the graph of F at the point $(\bar{u}, \mathrm{F}(\bar{u}))$ which corresponds to $(\bar{x}, \bar{y})$; thus by Proposition 3.1 (*b*), F is strictly differentiable at $\bar{u}$. Let $\varphi(u) = (\xi(u), \eta(u))$ denote the point $(x, y) \in \mathrm{M}$ that corresponds to $(u, \mathrm{F}(u))$. Then φ, ξ, and η are Lipschitzian on U and strictly differentiable at $\bar{u}$, Moreover the range of the linear transformation $\nabla\varphi(\bar{u})$ is the image, under the derivative $\nabla\Phi^{-1}(\bar{u}, \mathrm{F}(\bar{u}))$ of the inverse coordinate transformation, of the tangent space to gph F at $(\bar{u}, F(\bar{u}))$, which is the graph of $\nabla F(\bar{u})$; thus it is $T_M(\bar{x}, \bar{y})$. The range of the linear transformation $\nabla\xi(\bar{u})$ is therefore the image of $T_M(\bar{x}, \bar{y})$ under the projection in the first component, and we know this to be all of R^n. Thus $\nabla\xi(\bar{u})$ is nonsingular. By the inverse function theorem (in the Lipschitzian version of Clarke [*10*], for instance, since the Clarke generalized Jacobian reduces to $\nabla\xi(\bar{u})$ in the present case) the inverse ξ^{-1} exists as a Lipschitzian function in a neighborhood of $\bar{x} = \xi(\bar{u})$. Then locally around $(\bar{x}, \bar{y})$ we have

$$(x, y) \in M \Leftrightarrow [x = \xi(u),\ y = \eta(u)] \Leftrightarrow y = \eta(\xi^{-1}(x)).$$

But $M = \text{gph } D$. Therefore D reduces in a neighborhood of $\bar{x}$ to a single-valued Lipschitzian mapping, namely $\eta \circ \xi^{-1}$. Utilizing again the fact that $T_M(\bar{x}, \bar{y})$ is a subspace of dimension n, and applying Proposition 3.1 (*b*) to D at $\bar{x}$, we see that D must be strictly differentiable at $\bar{x}$ with $\nabla D(\bar{x}) = A$. □

4. SECOND DERIVATIVES

In applying Theorem 3.5 to the graphs of the subdifferentials of convex functions, saddle functions, or lower-C^2 functions, as is permissible by Corollaries 2.3, 2.4, and 2.5, we gain insight into second derivative properties of such functions. We shall not attempt here to develop any general theory of second derivatives that goes beyond the bounds of the conclusions which can immediately be drawn in this manner. Nevertheless it will be necessary to consider certain generalized limits of second-order difference quotients in order to formulate our results.

The limit concept we need is that of *epi-convergence*, which corresponds to set convergence of the epigraphs of functions. The theory of such convergence can be found in Dolecki, Salinetti and Wets [*12*] (see also Wets [*28*], Rockafellar and Wets [*25*]). The basic notions are as follows.

Suppose $\{ g_\nu \}$ is a sequence of lower semicontinuous functions from R^n to the extended reals $\bar{R}$. (The epigraphs

$$\text{epi } g_\nu = \{ (x, \alpha) \in R^{n+1} \mid \alpha \geqq g_\nu(x) \}$$

are closed sets that determine the functions g_ν completely.) One says that

$$g = \operatorname*{epi\text{-}lim\,sup}_{\nu\to\infty} g_\nu$$

if g is given by

$$(4.1) \qquad g(\bar{x}) = \lim_{\varepsilon \downarrow 0} \, [\limsup_{\nu\to\infty} \, [\inf_{|x-\bar{x}| \leqq \varepsilon} g_\nu(x)]] \quad \text{for all} \quad \bar{x},$$

and one says that

$$g = \operatorname*{epi\text{-}lim\,inf}_{\nu\to\infty} g_\nu$$

if g is given by

(4.2) $$g(\bar{x}) = \lim_{\varepsilon \downarrow 0} [\liminf_{\nu \to \infty} [\inf_{|x-\bar{x}| \leq \varepsilon} g_\nu(x)]] \quad \text{for all} \quad \bar{x}.$$

The first case corresponds to

$$\text{epi } g = \liminf_{\nu \to \infty} [\text{epi } g_\nu]$$

in the sense of (3.2), whereas the second corresponds to

$$\text{epi } g = \limsup_{\nu \to \infty} [\text{epi } g_\nu]$$

in the sense of (3.1). One says that

$$g = \operatorname*{epi-lim}_{\nu \to \infty} g_\nu$$

if both (4.1) and (4.2) are true, i. e. if

$$\text{epi } g = \lim_{\nu \to \infty} [\text{epi } g_\nu]$$

in the sense of (3.3).

We shall also need the notion of a *generalized purely quadratic* convex function on R^n. By this we mean a function expressible in the form

(4.3) $$q(x) = \begin{cases} (1/2)x \cdot \mathrm{Q}x & \text{if} \quad x \in \mathrm{N}, \\ +\infty & \text{if} \quad x \notin \mathrm{N}, \end{cases}$$

where N is a subspace of R^n (possibly all of R^n) and Q is symmetric and positive semidefinite. Our motivation for this concept is the following fact.

PROPOSITION 4.1. — *Let q be a closed proper convex function on R^n. Then for the graph of the subdifferential ∂q to be a subspace in $\mathrm{R}^n \times \mathrm{R}^n$, it is necessary and sufficient that q be a generalized purely quadratic convex function (up to an additive constant).*

Proof. — If q does have the form in (4.3), then

$$\partial q(x) = \mathrm{Q}x + \partial\psi_\mathrm{N}(x) \quad \text{for all} \quad x$$

by [*20*, Theorem 23.8], where ψ_N is the indicator of N and has $\partial\psi_\mathrm{N}(x) = \mathrm{N}^\perp$ for $x \in \mathrm{N}$, $\partial\psi_\mathrm{N}(x) = \phi$ for $x \notin \mathrm{N}$. This means that

$$y \in \partial q(x) \Leftrightarrow [x \in \mathrm{N} \quad \text{and} \quad y - \mathrm{Q}x \in \mathrm{N}^\perp].$$

Then gph ∂q is indeed a subspace $\mathrm{S} \subset \mathrm{R}^n \times \mathrm{R}^n$.

Conversely if gph ∂q is a subspace S, which by Corollary 2.3 must be of dimension n, let N denote the projection of S in the first component, i. e. $\mathrm{N} = \text{dom } \partial q$. Then N too is a subspace (in particular a relatively open convex set), and N must then be the effective domain of q:

$$q(x) < \infty \Leftrightarrow x \in \mathrm{N}$$

(apply [20, Theorem 23.4]). This implies that

$$\partial q(x) + N^{\perp} = \partial q(x) \quad \text{for all} \quad x, \tag{4.4}$$

inasmuch as $N^{\perp}$ is a normal cone to N at every $x \in N$. Therefore

$$S = S \cap [N \times N] + (0, N^{\perp}).$$

Since $\dim S = n \dim N + \dim N^{\perp}$, it follows that the subspace $S_0 = S \cap [N \times N]$ has the same dimension as N. Thus it is the graph of a linear transformation $Q_0 : N \to N$, and one has

$$\partial q(x) = \begin{cases} Q_0 x + N^{\perp} & \text{if} \quad x \in N, \\ \phi \quad \text{if} \quad x \notin N. \end{cases}$$

Let q_0 be the restriction of q to N. In terms of N rather than the larger space R^n, we have $\partial q_0(x) = \{ Q_0 x \}$, so q_0 is a convex function that is actually differentiable everywhere with $\nabla q_0 = Q_0$. Then the function $q_0 - c$, where $c = q_0(0)$, has to be purely quadratic:

$$q_0(x) - c = (1/2) x \cdot Q_0 x \quad \text{for all} \quad x \in N,$$

and Q_0 has to be positive semidefinite. We can extend Q_0 to a positive semidefinite linear transformation $Q : R^n \to R^n$. Then (4.3) holds for $q - c$ in place of q. □

COROLLARY 4.2. — *If q is a generalized purely quadratic convex function on R^n, then so is the conjugate function q^*.*

Proof. — $\partial q^* = \partial q^{-1}$ by [20] Corollary 23.5.1]. Furthermore, when $0 \in \partial q(0)$ one has $q(0) = 0$ if and only if $q^*(0) = 0$, as follows immediately from the formulas for conjugacy [20, Theorem 23.5]. □

In the sequel we shall be concerned with the second-order difference quotients

$$\Delta_{x,y,t}(h) = \frac{f(x + th) - f(x) - ty \cdot h}{t^2}. \tag{4.5}$$

THEOREM 4.3. — *Let $f : R^n \to \overline{R}$ be a closed proper convex function, and let $M = \operatorname{gph} \partial f$. For $(\bar{x}, \bar{y}) \in M$ to be a smooth point of M, it is necessary and sufficient that there exist a generalized purely quadratic convex function $q_{\bar{x},\bar{y}} : R^n \to \overline{R}$ such that*

$$q_{\bar{x},\bar{y}} = \operatorname*{epi-lim}_{t \downarrow 0} \Delta_{\bar{x},\bar{y},t}. \tag{4.6}$$

The stronger condition

$$q_{\bar{x},\bar{y}} = \operatorname*{epi-lim}_{\substack{t \downarrow 0 \\ (x,y)(\in M) \to (\bar{x},\bar{y})}} \Delta_{x,y,t}. \tag{4.7}$$

characterizes $(\bar{x}, \bar{y})$ as a strictly smooth point of M.

Proof. — The function $\Delta_{x,y,t} : \mathbf{R}^n \to \overline{\mathbf{R}}$ is closed proper convex with $\Delta_{x,y,t}(0) = 0$ and

$$\partial\Delta_{x,y,t}(h) = t^{-1}[\partial f(x + th) - y].$$

Thus

$$\text{gph}\,\partial\Delta_{x,y,t} = t^{-1}[\mathrm{M} - (x, y)]. \tag{4.8}$$

The defining property for $(\bar{x}, \bar{y})$ to be a smooth point of M, namely

$$\lim_{t \downarrow 0} t^{-1}[\mathrm{M} - (\bar{x}, \bar{y})] = \mathrm{S}$$

for some subspace $\mathrm{S} \subset \mathbf{R}^n \times \mathbf{R}^n$, can be written by virtue of (4.8) as

$$\lim_{t \downarrow 0} [\text{gph}\,\partial\Delta_{\bar{x},\bar{y},t}] = \mathrm{S}. \tag{4.9}$$

Since each of the functions $\Delta_{\bar{x},\bar{y},t}$ for $t > 0$ is closed proper convex with $\Delta_{\bar{x},\bar{y},t}(0) = 0$, the existence of the limit (4.9) is equivalent to the existence of a closed proper convex function $q_{\bar{x},\bar{y}}$ with $q_{\bar{x},\bar{y}}(0) = 0$ such that (4.6) holds (see Attouch [*2*]). Then $q_{\bar{x},\bar{y}}$ must be a generalized purely quadratic convex function by Proposition 4.1.

The strictly smooth case (4.7) falls out in the same way. □

A generalized purely quadratic convex function $q_{\bar{x},\bar{y}}$ satisfying (4.6). when it exists, obviously serves as a kind of *second derivative* function for f. We shall then say that f *is twice differentiable in the generalized sense* at $\bar{x}$ relative to the subgradient $\bar{y} \in \partial f(\bar{x})$. In the case of (4.7) we shall speak of *strict* twice differentiability in the generalized sense.

Note that $q_{\bar{x},\bar{y}}$ need not be finite everywhere when $\bar{x} \in \text{int}\,(\text{dom}\,f)$. It is easy to verify, for instance, that dom $q_{\bar{x},\bar{y}}$ must be contained in the normal cone to the convex set $\partial f(\bar{x})$ at $\bar{y}$. When $q_{\bar{x},\bar{y}}$ does happen to be finite everywhere, the epi-limits in (4.6) and (4.7) are equivalent to *pointwise* convergence of the functions in question (cf. Dolecki, Salinetti and Wets [*12*]); these difference quotient functions, being convex, must then actually converge uniformly on bounded sets [*20*, Theorem 10.9].

COROLLARY 4.4. — *Let f be a closed proper convex function on* $\mathbf{R}^n$, *and suppose f is twice differentiable in the generalized sense at $\bar{x}$ relative to the subgradient $\bar{y} \in \partial f(\bar{x})$, with $q_{\bar{x},\bar{y}}$ as the second derivative function. Then the conjugate f^* is twice differentiable in the generalized sense at $\bar{y}$ relative to the subgradient $\bar{x} \in \partial f^*(\bar{y})$, and the corresponding second derivative function is the conjugate $q_{\bar{x},\bar{y}}^*$.*

The same holds also for strict twice differentiability.

Proof. — Recall that $\partial f^* = \partial f^{-1}$ and $\partial q_{\bar{x},\bar{y}}^* = \partial q_{\bar{x},\bar{y}}^{-1}$. □

COROLLARY 4.5. — *Let f be a closed proper convex function on* $\mathbf{R}^n$, *and let* M′ *be the set of all $(\bar{x}, \bar{y})$ such that $\bar{y} \in \partial f(\bar{x})$ and f is twice differentiable* *in the generalized sense at* $\bar{x}$ *relative to* $\bar{y}$. *For* $(x, \bar{y}) \in M'$ *to be such that the twice differentiability is strict, it is necessary and sufficient that the corresponding second derivative function satisfy*

$$q_{\bar{x},\bar{y}} = \operatorname*{epi\text{-}lim}_{(x,y)(\in M') \to (\bar{x},\bar{y})} q_{x,y}\,. \tag{4.10}$$

Proof. — The functions $q_{x,y}$ all vanish at 0, so (4.10) is equivalent to

$$\text{gph}\ \partial q_{\bar{x},\bar{y}} = \lim_{(x,y)(\in M') \to (\bar{x},\bar{y})} \text{gph}\ q_{x,y}$$

(see Wets [*28*]). Recognizing gph $\partial q_{x,y}$ as the tangent space $S_M(x, y)$ to $M = \text{gph}\ \partial f$ at (x, y), we need only invoke Theorem 3.5 to obtain the desired conclusion. □

COROLLARY 4.6. — *Let* f *be a closed proper convex function on* R^n, *and let* $\bar{y} \in \partial f(\bar{x})$. *If the limit on the right side of* (*4.7*) (*where* $M = \text{gph}\ \partial f$) *exists at all, then in fact* f *must be strictly twice differentiable in the generalized sense at* $\bar{x}$ *relative to* $\bar{y}$.

Proof. — We know that the limit function g on the right side of (4.7) exists if and only if the set limit

$$\lim_{\substack{t \downarrow 0 \\ (x,y)(\in M) \to (\bar{x},\bar{y})}} \text{gph}\ \partial \Delta_{x,y,t} \tag{4.11}$$

exists, in which event the latter is gph ∂g (see Wets [*28*]; g is a certain closed proper convex function). In view of (4.8), however, the limit (4.11) is the Clarke tangent cone $T_M(\bar{x}, \bar{y})$, which is a subspace by Theorem 3.5. Then g must be a generalized purely quadratic convex function by Proposition 4.1. □

COROLLARY 4.7. — *Let* f *be a closed proper convex function on* R^n, *and let* $\bar{y} \in \partial f(\bar{x})$. *Suppose that* f *is strictly twice differentiable in the general sense at* $\bar{x}$ *relative to* $\bar{y}$, *and that the corresponding second derivative function* $q_{\bar{x},\bar{y}}$ *is finite everywhere. Then there is actually a neighborhood of* $\bar{x}$ *on which* f *is finite and continuously differentiable, and the gradient mapping* ∇f *is Lipschitzian; at* $\bar{x}$ *one has* $\nabla f(\bar{x}) = \bar{y}$ *and* ∇f *strictly differentiable.*

Proof. — This follows in the context of the preceding results by Theorem 3.7 (*a*) as applied to $D = \partial f$. □

COROLLARY 4.8. (*Alexandrov's Theorem*). — *Let* f *be a closed proper convex function on* R^n. *Then at almost every* $\bar{x} \in \text{int}\,(\text{dom}\, f)$ *there is a quadratic* (*finite but not necessarily purely quadratic*) *function* $q_{\bar{x}}$ *such that*

$$f(x) = q_{\bar{x}}(x) + o(|x - \bar{x}|^2)\,. \tag{4.12}$$

Proof. — Let $D = \partial f$. Then D is a maximal monotone relation (Corollary 2.3) and

$$\operatorname{int}(\operatorname{dom} D) = \operatorname{int}(\operatorname{dom} f)$$

(see [*20*, Theorem 23.4]). By the theorem of Mignot [*16*, Theorem 1.3], D is differentiable at almost every $\bar{x} \in \operatorname{int}(\operatorname{dom} f)$, the graph of $\nabla D(\bar{x})$ being then, of course, the tangent space $S_M(\bar{x}, \bar{y})$ to $M = \operatorname{gph} D$ at $(\bar{x}, \bar{y})$, where $\bar{y}$ is the unique element of $D(\bar{x}) = \partial f(\bar{x})$ (and consequently $\bar{y} = \nabla f(\bar{x})$, cf. [*20*, Theorem 25.1]). Theorem 4.3 identifies this as the case where (4.6) holds and the set $\operatorname{gph} \partial q_{\bar{x},\bar{y}} = S_M(\bar{x}, \bar{y})$ projects in the first component onto all of R^n, i. e. one has $\operatorname{dom} q_{\bar{x},\bar{y}} = R^n$. Then the epi-convergence in (4.6) can be expressed as pointwise convergence

$$q_{\bar{x},\bar{y}}(h) = \lim_{t \downarrow 0} \Delta_{\bar{x},\bar{y},t}(h) \tag{4.13}$$

that is uniform on bounded sets (cf. the observation that precedes Corollary 4.4). Condition (4.12) is just another way of writing (4.13) with this uniformity taken into account. □

Lower-C^2 functions such as appear in Corollary 2.5 enjoy almost the same generalized second derivative properties as convex functions.

THEOREM 4.10. — *Let $f: X \to R$ be lower-C^2, where X is open in R^n, and let $M = \operatorname{gph} \partial f$, where ∂f is the Clarke subdifferential of f. Then the conclusions in Theorem 4.3 and Corollaries 4.5, 4.6, 4.7, and 4.8 are valid.*

Proof. — As demonstrated in the proof of Corollary 2.5, there is for every $\bar{x} \in X$ a compact convex neighborhood C of $\bar{x}$ and a number $\mu > 0$ such that the function h in (2.1) is closed proper convex, and ∂h satisfies (2.2). One need only apply the results in question to h. □

For saddle functions as in Corollary 2.4, there are complete analogues of all the results in this section, but the type of convergence that describes the limits of the second-order difference quotients is somewhat more complicated. This type of convergence has been developed by Attouch and Wets [*3*] [*4*]. The details will not be given here.

REFERENCES

[*1*] A. D. ALEXANDROV, The existence almost everywhere of the second differential of a convex function and some associated properties of convex surfaces, *Ucenye Zapiski Leningr. Gos. Univ. Ser. Mat.*, t. **37**, 1939, p. 3-35 (in Russian).

[*2*] H. ATTOUCH, Familles d'opérateurs maximaux monotones et mesurabilité, *Annali di Matematica Pura et Applicata*, t. **120**, 1979, p. 35-111.

[*3*] H. ATTOUCH and R.J.-B. WETS, A convergence for bivariate functions aimed at the convergence of saddle values, in *Mathematical Theory of Optimization* (P. Cecconi and T. Zolezzi, eds.), Springer-Verlag Lecture Notes in Math., t. **997**, 1983.

[4] H. ATTOUCH and R. J.-B. WETS, A convergence theory for saddle functions, *Trans. Amer. Math. Soc.*, t. **226**, 1983.
[5] J. P. AUBIN, Lipschitz behavior of solutions to convex minimization problems, *Math. of Op. Research*, t. **9**, 1984, p. 87-111.
[6] A. AUSLENDER, *Differential properties of the support function of the ε-subdifferential of a convex function*, Math. Programming, forthcoming.
[7] A. AUSLENDER, Stability in mathematical programming with nondifferentiable data, *SIAM J. Control Opt.*, t. **22**, 1984, p. 239-254.
[8] F. H. CLARKE, *Nonsmooth Analysis and Optimization*, Wiley-Interscience, 1983.
[9] F. H. CLARKE, Generalized gradients and applications, *Trans. Amer. Math. Soc.*, t. **205**, 1975, p. 247-262.
[10] F. H. CLARKE, On the inverse function theorem, *Pacific J. Math.*, t. **67**, 1976, p. 97-102.
[11] B. CORNET, *Regularity properties of normal and tangent cones*, forthcoming.
[12] S. DOLECKI, G. SALINETTI and R. J.-B. WETS, Convergence of functions: equi-semi-continuity, *Trans. Amer. Math. Soc.*, t. **276**, 1983, p. 409-429.
[13] J.-B. HIRIART-URRUTY, Approximating a second-order directional derivative for nonsmooth convex functions, *SIAM J. Control Opt.*, t. **20**, 1982, p. 783-807.
[14] J.-B. HIRIART-URRUTY, Calculus rules on the approximate second-order directional derivative of a convex function, *SIAM J. Control Opt.*, t. **22**, 1984, p. 381-404.
[15] C. LEMARECHAL and E. NURMINSKI, Sur la différentiabilité de la fonction d'appui du sous-différentiel approaché, *C. R. Acad. Sci. Paris*, t. **290** Sér. A, 1980, p. 855-858.
[16] F. MIGNOT, Contrôle dans les inéquations variationnelles elliptiques, *J. Functional Anal.*, t. **22**, 1976, p. 130-185.
[17] G. J. MINTY, Monotone (nonlinear) operations in Hilbert space, *Duke Math. J.*, t. **29**, 1962, p. 341-346.
[18] J.-P. PENOT, A characterization of tangential regularity, *Nonlinear Anal.*, t. **5**, 1981, p. 625-643.
[19] Yu. G. RESHETNIAK, Generalized derivatives and differentiability almost everywhere, *Mat. Sbornik*, t. **75**, 1968, p. 323-334.
[20] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[21] R. T. ROCKAFELLAR, Monotone operators associated with saddle functions and minimax problems, in *Nonlinear Functional Analysis*, Part I (F. E. Browder, ed.), Proc. of Symposia in Pure Math., *Amer. Math. Soc.*, t. **18**, 1970, p. 241-250.
[22] R. T. ROCKAFELLAR, Generalized subgradients in mathematical programming, in *Mathematical Programming Bonn 1982: The State of the Art* (A. Bachem et al., eds.), Springer-Verlag, Berlin, 1983, p. 368-380.
[24] R. T. ROCKAFELLAR and R.J.-B. WETS, Variational systems: An introduction, in *Multifunctions and Integrands* (G. Salinetti, ed.), Springer-Verlag Lecture Notes in Math. (1984).
[25] S. SAKS, *Theory of the Integral*, second revised edition, Hafner Publishing Co., New York, 1937.
[26] G. SALINETTI and R. J.-B. WETS, On the convergence of closed-valued measurable multifunctions, *Trans. Amer. Math. Soc.*, t. **266**, 1981, p. 275-289.
[27] R. J.-B. WETS, Convergence of convex function, variational inequalities and convex optimization problems, in *Variational Inequalities and Complementarity Problems* (R. Cottle, F. Giannessi and J.-L. Lions, eds.), Wiley, 1980, p. 405-419.

(Manuscrit reçu le 25 juin 1984)