DIRECTIONALLY LIPSCHITZIAN FUNCTIONS AND SUBDIFFERENTIAL CALCULUS

By R. T. ROCKAFELLAR†

[Received 17 May 1978]

Abstract

The theory of subgradients of convex functions is recognized for its many applications to optimization and differential equations (for example, Hamiltonian systems, monotone operators). F. H. Clarke has extended the theory to non-convex functions that are merely lower semicontinuous and used it to derive necessary conditions for non-smooth, non-convex problems in optimal control and mathematical programming. For locally Lipschitzian functions, he has proved a number of rules for subgradient calculation that generalize the ones previously known for convex functions.

This paper extends such rules to non-convex functions that are not necessarily locally Lipschitzian. The two main operations considered are the addition of functions and the composition of a function with a differentiable mapping. The theorems are strong enough to cover the main results known in the convex case.

1. Introduction

Let E be a linear topological space (with a locally convex Hausdorff topology), and let f be an extended-real-valued function on E. At each point x where f is finite, there is a weak*-closed convex (possibly empty) subset $\partial f(x)$ of the dual space E^* whose elements are called *subgradients* (or *generalized gradients*) of f at x.

If f is convex, $\partial f(x)$ consists of all $z \in E^*$ such that

$$f(x') \geqslant f(x) + \langle x' - x, z\rangle \quad \text{for all } x' \in E, \tag{1.1}$$

or, in other words,

$$\partial f(x) = \{z \mid f - \langle \cdot, z\rangle \text{ has } x \text{ as a global minimum point}\}. \tag{1.2}$$

This is the case for which the notion of subgradient was originally developed. Rockafellar [**11**] gave the definition for $E = \mathbf{R}^n$ and proved a number of rules for calculating $\partial f(x)$ when f is expressed in terms of other functions. In particular, he showed that

$$\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x) \tag{1.3}$$

† Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under grant number 77-3204 at the University of Washington, Seattle.

if f_1 and f_2 are both finite at x and there exists $\tilde{x}$ such that $f_1(\tilde{x}) < \infty$ anc f_2 is bounded above in a neighbourhood of $\tilde{x}$. He also showed for a conve: function g and linear transformation A (with adjoint A^*) that

$$\partial(g \circ A)(x) = A^*[\partial g(A(s))] \tag{1.4}$$

if g is finite at $A(x) < \infty$ and there exists $\tilde{x}$ such that g is bounded above o a neighbourhood of $A(\tilde{x})$. The proofs of these results in the infinite dimensional case, where A must be continuous, were given in [**12**] an [**13**], respectively; see also [**14**] for a general exposition.

For saddle functions, generalizations of (1.3) and (1.4) have bee obtained by McLinden [**10**].

The definition of $\partial f(x)$ was extended by Clarke [**3**, **4**] to arbitrary low semicontinuous (l.s.c.) functions on $E = \mathbf{R}^n$ by a three-stage metho First Clarke defined $\partial f(x)$ when f is *Lipschitzian* (finite and Lipschi continuous) in a neighbourhood of x and expressed it by means of certa generalized directional derivatives (described below). Next he used th to define normal vectors to closed sets. Finally he defined $\partial f(x)$ in th general case in terms of normal vectors to the epigraph of f, but withou furnishing a directional derivative characterization. Later he used th directional derivatives in the Lipschitzian case as the basis of the definitio of $\partial f(x)$ when E is any normed space [**5**].

We have recently [**15**] supplied an alternative development of Clarke ideas that includes a direct definition of $\partial f(x)$ in terms of a still mor general directional derivative function. This definition covers all th cases at once and does not require E to be normed. It can be express in the following manner, which brings out the natural relationship wi the situation for convex functions and explains why the concept especially relevant for optimization theory.

Let x be any point where f is finite (but not necessarily lower sem continuous). We shall denote the set of all neighbourhoods of x by $\mathcal{N}(x$ Let us say that $y \in E$ is a *vector of approximately uniform descent* for f at (at a rate $\rho > 0$) if for every $Y \in \mathcal{N}(y)$ there exist $X \in \mathcal{N}(x)$, $\delta > 0$, $\lambda >$ such that

(1.5) for all $t \in (0, \lambda)$ and $x' \in X$ with $f(x') \leqslant f(x) + \delta$,
there exists $y' \in Y$ with $f(x' + ty') \leqslant \max\{f(x'), f(x) - \delta\} - t\rho$.

(If f is l.s.c. at x, the 'max' term in (1.5) can be replaced simply by $f(x$ while if f is u.s.c. at x the condition $f(x') \leqslant f(x) + \delta$ is superfluous.) Call x *substationary point* of f if no such y and ($\rho > 0$) exists. Then

$$\partial f(x) \triangleq \{z \in E^* \mid f - \langle \cdot, z \rangle \text{ has } x \text{ as a substationary point}\}. \tag{1.6}$$

When f is convex, any substationary point must actually be a local (and hence global) minimum point, so that (1.6) reduces to (1.2). In the general case, the set $\partial f(x)$ is non-empty if and only if $y = 0$ is not a vector of approximately uniform descent for f at x.

If f has a local minimum at x, then certainly x is a substationary point. However, other extrema also lead to this condition. For example, if f is 'directionally Lipschitzian' at x (a property developed in [**15**] that will play an important role below), it is known that $\partial f(x) = -\partial(-f)(x)$, so if f has a local *maximum* at x, then again x must be a substationary point. Saddle points of finite saddle functions also fit this criterion. Notationally, of course, (1.6) says that

(1.7) $$x \text{ is substationary } \Leftrightarrow 0 \in \partial f(x).$$

We would like to be able to 'calculate' this condition into more explicit forms when f has particular structure, and this is one source of motivation for rules like (1.3) and (1.4).

Such rules have heretofore not been extended to this general framework, but besides the convex and saddle function cases already mentioned, and the obvious case of differentiable functions, some results are known for locally Lipschitzian functions on normed spaces. Clarke [**5**, Proposition 8] has shown for such functions that the inclusion

(1.8) $$\partial(f_1+f_2)(x) \subset \partial f_1(x)+\partial f_2(x)$$

holds, and moreover with equality if f_1 and f_2 are both 'regular' at x. A locally Lipschitzian function f is regular at x in Clarke's sense if for all y the one-sided directional derivative $f'(x;y)$ exists and

(1.9) $$\limsup_{\substack{x'\to x\\ t\downarrow 0}} \frac{f(x'+ty)-f(x')}{t} = f'(x;y).$$

(In particular, a convex function which is finite at x and bounded above in a neighbourhood of x is Lipschitzian and regular at x in this sense.)

As for a generalization of (1.4), Clarke has proved in [**6**, § 13] that

(1.10) $$\partial(g\circ F)(x) \subset A^*[\partial g(F(x))]$$

if F is a continuously Gâteaux differentiable mapping from a normed space E to another normed space E_1 having derivative A at x, and g is a locally Lipschitzian function on E_1. Equality holds if g is 'regular' at $A(x)$.

The purpose of this paper is to extend these results to functions that are not necessarily locally Lipschitzian or even locally finite, and to spaces that are not necessarily normed. Furthermore, this will be accomplished in such a way that the theorems quoted earlier for convex functions will be corollaries. The role of Clarke's 'regularity' will be taken by a notion of 'subdifferential regularity'. The interiority conditions that appear in the convex case will be supplanted by the 'directionally Lipschitzian' property.

Since extended-real-valued functions are covered, the results can be applied to the indicator functions of subsets of E in order to obtain formulas for normal cones. The case of a set defined by a system of inequalities is given explicit treatment.

In the final section, an application is made to the characterization of a relative minimum point, and a conclusion about subgradients is drawn from Ekeland's variational principle.

2. Subderivatives and directionally Lipschitzian functions

The subgradient set $\partial f(x)$ (at a point x where f is finite) corresponds to a kind of directional derivative that for general functions has a complicated description but can be reduced to simpler formulas in many important cases (as explained in [**15**]). In expressing this for general functions, it is expedient to use the notation

$$(2.1)\qquad (x',\alpha')\downarrow_f x \Leftrightarrow (x',\alpha')\to(x,f(x)) \quad \text{in epi}\, f,$$

where

$$(2.2)\qquad \text{epi}\, f = \{(x',\alpha')\in E\times R \mid \alpha' \geqslant f(x')\}.$$

The (*upper*) *subderivative* of f at x with respect to y is

$$(2.3)\qquad f^{\uparrow}(x;y) = \limsup_{(x',\alpha')\downarrow_f x} \inf_{y'\to y} \frac{f(x'+ty')-\alpha'}{t}$$

$$\triangleq \sup_{Y\in\mathcal{N}(y)} \left[\limsup_{\substack{(x',\alpha')\downarrow_f x\\ t\downarrow 0}} \left[\inf_{y'\in Y} \frac{f(x'+ty')-\alpha'}{t} \right]\right].$$

Note that y is a vector of approximately uniform descent at x, as defined in § 1, if and only if $f^{\uparrow}(x;y) < 0$. Definition (1.6) is equivalent to

$$(2.4)\qquad f(x) = \{z\in E^* \mid f^{\uparrow}(x;y) \geqslant \langle y,z\rangle \text{ for all } y\in E\}.$$

It has been demonstrated [**15**, Theorems 2 and 4] that the function $f^{\uparrow}(x;y)$ is lower semicontinuous and sublinear (i.e. convex and positively homogeneous, not $+\infty$ at the origin). One has $\partial f(x) = \emptyset$ if and only if 0 is a vector of approximately uniform descent at x, in which case $f^{\uparrow}(x;0) = -\infty$. Otherwise $f^{\uparrow}(x;0) = 0$ and

$$(2.5)\qquad \sup\{\langle y,z\rangle \mid z\in\partial f(x)\} = f^{\uparrow}(x;y).$$

As already mentioned, the limit (2.3) defining $f^{\uparrow}(x; y)$ can often be reduced to something simpler (see [**15**]). But the facts just cited explain why these subderivatives, whatever their description, must naturally enter into any general study of subgradients. They also have a certain geometric interpretation in terms of 'tangent vectors', and this will be useful in several respects below.

For any set $C \subset E$ and any point $x \in C$, the (Clarke) *tangent cone* to C at x is defined by

$$(2.6) \qquad T_C(x) = \liminf_{\substack{x'(\in C)\to x \\ t\downarrow 0}} t^{-1}(C - x')$$
$$\triangleq \bigcap_{V\in\mathcal{N}(0)} \bigcup_{\substack{X\in\mathcal{N}(x)\\ \lambda>0}} \bigcap_{\substack{x'\in C\cap X\\ t\in(0,\lambda)}} [t^{-1}(C-x')+V].$$

This is always a closed convex cone containing the origin (see [**16, 15**]). Its polar

$$(2.7) \qquad N_C(x) = \{z \in E^* \mid \langle y, z\rangle \leqslant 0 \text{ for all } y \in T_C(x)\}$$

is called the *normal cone* to C at x. For the indicator function

$$(2.8) \qquad \psi_C(x') = 0 \quad \text{if } x' \in C, \qquad \psi_C(x') = +\infty \quad \text{if } x' \notin C,$$

one has

$$(2.9) \quad \psi_C^{\uparrow}(x; y) = 0 \quad \text{if } y \in T_C(x), \qquad \psi_C^{\uparrow}(x; y) = +\infty \quad \text{if } y \notin T_C(x),$$

$$(2.10) \qquad \partial\psi_C(x) = N_C(x).$$

The fact that sets in E can be identified with their indicator functions, and results about subgradients can thereby be specialized in some respects to assertions about normal vectors, is crucial to our approach, especially for applications to constrained optimization.

The fundamental relationship between tangent cones and subderivatives is the following (see [**15**]). The epigraph

$$(2.11) \qquad \operatorname{epi} f^{\uparrow}(x;\, \cdot) = \{(y, \beta) \in E \times R \mid \beta \geqslant f^{\uparrow}(x; y)\}$$

coincides with the tangent cone $T_{\operatorname{epi} f}(x, f(x))$, and consequently one has

$$(2.12) \qquad \partial f(x) = \{z \in E^* \mid (z, -1) \in N_{\operatorname{epi} f}(x, f(x))\}.$$

One of the situations where a considerable simplification is possible in analysing subderivatives is the case where there exists at least one $y \in E$ such that

$$(2.13) \qquad \limsup_{\substack{(x',\alpha')\downarrow_f x\\ y'\to y,\, t\downarrow 0}} \frac{f(x'+ty')-\alpha'}{t} < \infty.$$

Then we say f is *directionally Lipschitzian* at x. In this event, as shown in [**15**, Theorem 3], the set of all y satisfying (2.13) coincides with the open convex cone

$$\text{int}\{y \in E \mid f^{\uparrow}(x; y) < \infty\}, \tag{2.16}$$

and for each y in this cone the limit (2.13) agrees with $f^{\uparrow}(x; y)$. For convenience in applying the main results below, we list some of the criteria for this case that have been established in [**15**].

THEOREM 1. *Each of the following implies that f is directionally Lipschitzian at x (a point where f is finite):*

(a) *f is Lipschitzian on a neighbourhood of x;*

(b) *f is convex and bounded above on a neighbourhood of some point (not necessarily x itself);*

(c) *f is concave and bounded below on a neighbourhood of some point (not necessarily x itself);*

(d) *f is non-decreasing with respect to the partial ordering of E induced by some closed convex cone K with non-empty interior;*

(e) *f is the indicator of a set C that is epi-Lipschitzian at x;*

(f) *$E = \mathbf{R}^n$, f is lower semicontinuous on a neighbourhood of x, and the cone $\{y \mid f^{\uparrow}(x; y) < \infty\}$ is not included in any subspace of lower dimension;*

(g) *$E = \mathbf{R}^n$, f is lower semicontinuous on a neighbourhood of x, and $\partial f(x)$ is non-empty and does not include an entire line.*

Case (a) implies that $\partial f(x)$ is non-empty and the cone (2.16) is all of E. It means that $f(x')$ is finite for all x' in some neighbourhood of x, and

$$\limsup_{\substack{x' \to x \\ y' \to 0 \\ t \downarrow 0}} \frac{f(x' + ty') - f(x')}{t} < \infty. \tag{2.17}$$

In Case (e), C is said to be *epi-Lipschitzian* at x if there exist $X \in \mathcal{N}(x)$, $\lambda > 0$, and a non-empty open set Y such that

$$x' + ty' \in C \quad \text{for all } x' \in C \cap X,\ y' \in Y,\ t \in (0, \lambda) \tag{2.18}$$

(see [**15, 16**]).

3. Subdifferential regularity

The function f will be called *subdifferentially regular* at x if f is finite at x and

$$\liminf_{\substack{y' \to y \\ t \downarrow 0}} \frac{f(x + ty') - f(x)}{t} = f^{\uparrow}(x; y) \quad \text{for all } y. \tag{3.1}$$

When f is the indicator of a set C (containing x), this means C is *tangentially* *regular* at x in the sense that the *contingent cone*

$$K_C(x) = \limsup_{t \downarrow 0} t^{-1}(C-x) \tag{3.2}$$

$$= \bigcap_{V \in \mathcal{N}(0)} \bigcap_{\lambda > 0} \bigcup_{t \in (0,\lambda)} [t^{-1}(C-x)+V]$$

coincides with the tangent cone $T_C(x)$ in (2.6). More generally, f is subdifferentially regular at x if and only if $\operatorname{epi} f$ is tangentially regular at $(x, f(x))$. Indeed, the left-hand side of (3.1) defines the function of y whose epigraph is $K_{\operatorname{epi} f}(x, f(x))$ while, as already noted, for the right-hand side the epigraph is $T_{\operatorname{epi} f}(x, f(x))$.

Recall that z is said to be the gradient of f at z in the *Hadamard sense* if the functions

$$\varphi_t(y) = \frac{f(x+ty)-f(x)}{t} - \langle y, z\rangle \tag{3.3}$$

converge to 0 uniformly on all *compact* y-sets as $t \downarrow 0$. (If $E = \mathbf{R}^n$, this is plain 'differentiability'; it differs from Fréchet differentiability when E is infinite-dimensional in that the latter refers to bounded sets, rather than compact sets.) Let us say that z is a *lower semigradient of f at x (Hadamard sense)* if merely the functions $\min\{\varphi_t, 0\}$ converge in the manner prescribed. When such a z exists, f may be said to be *lower semidifferentiable at x (Hadamard sense)*. This condition can be written

$$f(x') \geqslant f(x) + \langle x'-x, z\rangle + o(x'-x). \tag{3.4}$$

The 'Hadamard sense' specifies for the infinite-dimensional case which of several possible interpretations is to be given to (3.4). We shall also be interested in lower semidifferentiability and lower semigradients z at x in the *full limit sense*, by which we mean that the following condition is satisfied:

$$\liminf_{\substack{y' \to y \\ t \downarrow 0}} \frac{f(x+ty')-f(x)}{t} \geqslant \langle y, z\rangle \quad \text{for all } y \in E. \tag{3.5}$$

PROPOSITION 1. *If z is a lower semigradient of f at x in the full limit sense, then the same holds in the Hadamard sense. Moreover, the converse is true if E is normable (and hence in particular when $E = \mathbf{R}^n$).*

At all events, if z is a lower semigradient of f at x in the full limit sense, then f is lower semicontinuous at x and $z \in \partial f(x)$.

Proof. Suppose (3.5) holds. Then for every $y \in E$ and $\varepsilon > 0$ there exist $Y \in \mathcal{N}(y)$ and $\lambda > 0$ such that

$$[f(x+ty')-f(x)]/t \geqslant \langle y, z\rangle - \varepsilon \quad \text{for all } y' \in Y,\ t \in (0,\lambda). \tag{3.6}$$

Applying this to $y = 0$, we see that f is l.s.c. at x. Since for all $y \in E$,

$$(3.7) \qquad f^{\uparrow}(x; y) \geqslant \limsup_{t \downarrow 0} \inf_{y' \to y} \frac{f(x+ty')-f(x)}{t}$$

$$\geqslant \liminf_{\substack{y' \to y \\ t \downarrow 0}} \frac{f(x+ty')-f(x)}{t} \geqslant \langle y, z \rangle,$$

we conclude from (2.4) that $z \in \partial f(x)$.

Given any compact set $D \subset E$ and $\varepsilon > 0$, there is for each $y \in D$ an open neighbourhood $Y \in \mathcal{N}(y)$ satisfying not only (3.6) but

$$\langle y', z \rangle \leqslant \langle y, z \rangle + \varepsilon \quad \text{for all } y' \in Y.$$

This collection of neighbourhoods is an open covering of D from which a finite subcover can be extracted. Thus there exist points $y_i \in D$, where $i = 1, \ldots, m$, and corresponding neighbourhoods $Y_i \in \mathcal{N}(y_i)$ covering D, along with numbers $\lambda_i > 0$ such that

$$(3.8) \quad y' \in Y_i \quad \text{and} \quad t \in (0, \lambda_i) \quad \Rightarrow$$

$$\langle y', z \rangle \leqslant \langle y_i, z \rangle + \varepsilon \quad \text{and} \quad \frac{f(x+ty')-f(x)}{t} \geqslant \langle y_i, z \rangle - \varepsilon.$$

Taking λ to be the least of the numbers λ_i, we see that for any $y' \in D$ and $t \in (0, \lambda)$ there exists an index i such that (3.8) is applicable; then

$$[f(x+ty')-f(x)]/t - \langle y', z \rangle \geqslant -2\varepsilon.$$

Thus for any $\varepsilon > 0$ there exists $\lambda > 0$ such that

$$\min\{\varphi_t(y'), 0\} \geqslant -2\varepsilon \quad \text{for all } y' \in D,\ t \in (0, \lambda).$$

Since D was any compact subset of E, this establishes that z is a lower semigradient in the Hadamard sense.

For the converse part under the assumption that E is normable, suppose z is a lower semigradient in the Hadamard sense, but (3.5) is false for a certain y. Then there exist sequences $y_k \to y$ and $t_k \downarrow 0$ such that (for some α)

$$[f(x+t_k y_k)-f(x)]/t_k \leqslant \alpha < \langle y, z \rangle \quad \text{for all } k.$$

Let D be the compact set consisting of y and the points y_k. By assumption the functions $\min\{\varphi_t, 0\}$ converge to 0 uniformly on D, so for any $\lambda > 0$ there exists $\varepsilon > 0$ such that

$$[f(x+ty_k)-f(x)]/t \geqslant \langle y_k, z \rangle - \varepsilon \quad \text{for all } k \text{ when } t \in (0, \lambda).$$

This implies for arbitrary $\varepsilon > 0$ that

$$[f(x+t_k y_k)-f(x)]/t_k > \langle y, z \rangle - 2\varepsilon \quad \text{for } k \text{ sufficiently large,}$$

and (3.9) is thereby contradicted. Hence (3.5) cannot fail for any y, and Proposition 1 is proved.

COROLLARY 1. *Suppose that E is normable and $\partial f(x) \neq \emptyset$. Then f is subdifferentially regular at x if and only if every subgradient $z \in \partial f(x)$ is actually a lower semigradient (Hadamard sense).*

COROLLARY 2. *Suppose E is normable. Then f is differentiable at x with gradient z (Hadamard sense) if and only if $f(x)$ is finite and*

$$\lim_{\substack{y' \to y \\ t \downarrow 0}} \frac{f(x+ty')-f(x)}{t} = \langle y, z \rangle \quad \textit{for all } y.$$

PROPOSITION 2. *Suppose f is subdifferentially regular at x and also directionally Lipschitzian at x. Then for all y belonging to*

$$\operatorname{int}\{y \mid f^{\uparrow}(x; y) < \infty\},$$

the one-sided directional derivative $f'(x; y)$ exists and

$$f'(x; y) = f^{\uparrow}(x; y) = \lim_{\substack{y' \to y \\ t \downarrow 0}} \frac{f(x+ty')-f(x)}{t}. \tag{3.11}$$

Proof. As mentioned at the beginning of §2, since f is directionally Lipschitzian we know from [**15**, Theorem 3] that for y of this type the limit in (2.13) agrees with $f^{\uparrow}(x; y)$. Therefore

$$\limsup_{\substack{y' \to y \\ t \downarrow 0}} \frac{f(x+ty')-f(x)}{t} \leqslant f^{\uparrow}(x; y).$$

By assumption, however, (3.1) also holds, so we may conclude (3.11).

COROLLARY. *Suppose f is Lipschitzian in a neighbourhood of x. Then f is subdifferentially regular at x if and only if f is regular in Clarke's sense (that is,* (1.9) *holds).*

Proof. The left-hand side of (1.9) is known to coincide with $f^{\uparrow}(x; y)$ in the Lipschitzian case [**15**, Theorem 3]. On the other hand, when f is Lipschitzian on a neighbourhood of x and $f'(x; y)$ exists, one has

$$\begin{aligned} f'(x; y) &= \lim_{t \downarrow 0} \frac{f(x+ty)-f(x)}{t} + \lim_{\substack{y' \to y \\ t \downarrow 0}} \frac{f(x+ty')-f(x+ty)}{t} \\ &= \lim_{\substack{y' \to y \\ t \downarrow 0}} \frac{f(x+ty')-f(x)}{t}. \end{aligned}$$

Thus (1.9) and (3.1) are equivalent in this case.

Finally, we give two examples (besides the indicator function case mentioned at the start of this section) where f is subdifferentially regular. Other examples will be generated from these by the theorems in §§ 4 and 5.

PROPOSITION 3. *If f is convex and finite at x, then f is subdifferentially regular at x.*

Proof. According to [**15**, Theorem 1], one has $T_C(x) = K_C(x)$ when C is convex. Thus convex sets are everywhere tangentially regular. Applying this to $\operatorname{epi} f$, we deduce that f is subdifferentially regular when $\operatorname{epi} f$ is convex.

Generalizing the terminology of Bourbaki [**1**], we say that f is *strictly differentiable at x in the Hadamard sense* (with gradient z) if it is finite on a neighbourhood of x and the functions

$$\varphi_{x',t}(y) = \frac{f(x'+ty)-f(x')}{t} - \langle y, z\rangle \tag{3.12}$$

converge to 0 uniformly on all *compact* y-sets as $t \downarrow 0$ and $x' \to x$. (The Bourbaki definition corresponds to the Fréchet sense: bounded sets instead of compact sets.) On the other hand, we say f is *strictly differentiable at x in the full limit sense* if it is finite on a neighbourhood of x and

$$\lim_{\substack{x'\to x,\, y'\to y\\ t\downarrow 0}} \frac{f(x'+ty')-f(x')}{t} = \langle y, z\rangle \quad \text{for all } y. \tag{3.13}$$

If f is known to be Lipschitzian in a neighbourhood of x, the limit $y' \to y$ in (3.13) is superfluous, because

$$\lim_{\substack{x'\to x,\, y'\to y\\ t\downarrow 0}} \frac{f(x'+ty')-f(x'+ty)}{t} = 0.$$

PROPOSITION 4. *If f is strictly differentiable at x in the full limit sense, then the same holds in the Hadamard sense. Moreover, the converse is true if E is normable (and hence in particular for $E = \mathbf{R}^n$).*

At all events, if f is strictly differentiable at x in the full limit sense, then f is Lipschitzian on a neighbourhood of x and subdifferentially regular at x, and $\partial f(x)$ reduces to the single element $z = \nabla f(x)$, which is in particular the gradient of f at x in the Hadamard sense.

Proof. The argument for the first part is closely parallel to the one for Proposition 1 and therefore need not be repeated. Suppose now that (3.13) holds. Then (2.17) certainly holds, so f is Lipschitzian in a neighbourhood of x. Also, (3.13) implies Clarke's regularity property (1.9) and hence subdifferential regularity (Corollary to Proposition 2), in fact with

$$\langle y, z\rangle = f'(x; y) = f^{\uparrow}(x; y) = \liminf_{\substack{y' \to y \\ t \downarrow 0}} \frac{f(x+ty')-f(x)}{t} \quad \text{for all } y.$$

Thus z is the unique subgradient and the unique lower semigradient in the full limit sense (hence also in the Hadamard sense, in view of Proposition 1). The same holds for $-z$ and $-f$, so z is actually the gradient of f at x in the Hadamard sense. This finishes the proof.

Proposition 4 yields the following generalization of Clarke's result [**5**, Proposition 4] that $\partial f(x) = \{\nabla f(x)\}$ when E is a Banach space and f is continuously Gâteaux differentiable around x. (See Lebourg [**8**, Theorem 2.1] for another generalization.)

COROLLARY 1. *Suppose that for all x' in some neighbourhood of x, $f(x')$ is finite, $f'(x'; y)$ exists for all $y \in E$ and is continuous in x', y. Then f is subdifferentially regular at x, strictly differentiable at x in the Hadamard sense, and $\partial f(x) = \{\nabla f(x)\}$.*

Proof. Under this hypothesis (3.13) holds, as proved in [**15**, Proposition 5].

In the finite-dimensional case, we can speak simply of 'strict differentiability', since the 'full limit sense' coincides then with the 'Hadamard sense' (and the 'Fréchet sense' used by Bourbaki). The next corollary extends a result of Clarke [**3**, **4**], for locally Lipschitzian functions.

COROLLARY 2. *Suppose that $E = \mathbf{R}^n$ and f is lower semicontinuous on a neighbourhood of x. Then the following conditions are equivalent and imply in particular that f is subdifferentially regular at x:*

(a) *f is strictly differentiable at x;*
(b) *$\partial f(x)$ consists of a single vector;*
(c) *$f^{\uparrow}(x; y) = -f^{\uparrow}(x; -y)$ for all y;*
(d) *f is Lipschitzian in a neighbourhood of x, differentiable at x, and the gradient mapping $\nabla f(x)$ is continuous at x relative to the set of points where f is differentiable.*

Proof. Clarke [**4**] has shown the equivalence of (a), (b), and (d) (in the finite-dimensional case) under the assumption that f is Lipschitzian in a neighbourhood of x. We have established in [**16**, Theorem 4] that the latter property is a consequence of (b) when f is l.s.c. in a neighbourhood of x. The equivalence of (b) and (c) follows from (2.4) and (2.5). (In particular (c) implies $f^{\uparrow}(x; 0) = -f^{\uparrow}(x; 0)$, so $f^{\uparrow}(x; 0)$ cannot be $-\infty$, and $\partial f(x)$ must therefore be non-empty.)

REMARK. Clarke has demonstrated in [4, 6] that certain functions o: the form

(3.14) $$f(x) = \max_{t \in T} g(t, x)$$

are locally Lipschitzian and 'regular' in his sense when the functions $g(t, \cdot)$ have these properties. These too are important examples of functions that are subdifferentially regular.

4. The sum of two functions

We are ready to prove the first of our main theorems, which concerns the subgradients of $f_1 + f_2$. For a number of reasons connected with the applications that are intended, it is important to allow the functions to have $-\infty$ as well as $+\infty$ as values, and this could cause ambiguity in interpreting the sum. We therefore adopt in this connection the convention that $\alpha_1 + \alpha_2 = +\infty$ if either α_1 or α_2 is $+\infty$ (even if the other is $-\infty$). This corresponds to our emphasis on epigraphs and gives universal validity to the relation

(4.1) $$\alpha_1 + \alpha_2 = \inf\{\alpha_1' + \alpha_2' \mid \alpha_1' > \alpha_1,\ \alpha_2' > \alpha_2\}.$$

THEOREM 2. *Let f_1 and f_2 be extended-real-valued functions on E that are finite at x. Suppose that f_2 is directionally Lipschitzian at x and*

(4.2) $$\{y \mid f_1^{\uparrow}(x; y) < \infty\} \cap \operatorname{int}\{y \mid f_2^{\uparrow}(x; y) < \infty\} \neq \emptyset.$$

Then

(4.3) $$(f_1 + f_2)^{\uparrow}(x; y) \leqslant f_1^{\uparrow}(x; y) + f_2^{\uparrow}(x; y) \quad \textit{for all } y,$$

(4.4) $$\partial(f_1 + f_2)(x) \subset \partial f_1(x) + \partial f_2(x),$$

where the set on the right in (4.4) *is also weak*-closed.*

Equality holds in (4.4) *if f_1 and f_2 are also subdifferentially regular. It also holds in* (4.3) *if in addition $f_1^{\uparrow}(x; 0)$ and $f_2^{\uparrow}(x; 0)$ are not $-\infty$ (that is, $\partial f_1(x)$ and $\partial f_2(x)$ are non-empty), and in that event $f_1 + f_2$ is likewise subdifferentially regular.*

Proof. Let $f_0 = f_1 + f_2$ and $l_i(y) = f_i(x; y)$, for $i = 0, 1, 2$. As explained at the beginning of § 2, l_i is lower semicontinuous and sublinear. If $l_i(0) = -\infty$, we have $\partial f_i(x) \neq \emptyset$, but otherwise $l_i(0) = 0$ and $\partial f_i(x) \neq \emptyset$. In the latter case $\partial l_i(0) = \partial f_i(x)$ by (2.4), since l_i is convex and consequently has its subgradients describable by formula (1.2).

We argue first that (4.4) and the weak*-compactness assertion follow from $l_0 \leqslant l_1 + l_2$ (that is (4.3)), while equality in (4.4) can be obtained from $l_0 = l_1 + l_2$. (In these sums, the convention $\infty - \infty = \infty$ is in force.) Suppose indeed that $l_0 \leqslant l_1 + l_2$. If either $l_1(0) = -\infty$ or $l_2(0) = -\infty$, this inequality implies $l_0(0) = -\infty$. Then (4.4) is trivial, because both sides are empty. We can assume therefore that $l_1(0) = l_2(0) = 0$. Then

$$\begin{aligned} \partial f_0(x) &= \{z \mid l_0(y) \geqslant \langle y, z\rangle \text{ for all } y\} \\ &\subset \{z \mid (l_1+l_2)(y) \geqslant \langle y, z\rangle \text{ for all } y\} = \partial(l_1+l_2)(0), \end{aligned} \tag{4.5}$$

where the final equality is valid because l_1+l_2 is a convex function satisfying $(l_1+l_2)(0) = 0$. We can calculate $\partial(l_1+l_2)(0)$ by means of the theorem for convex functions quoted in § 1. The directionally Lipschitzian property of f_2 implies that $f_2^{\uparrow}(x; y)$ is continuous in y on the interior of the set $\{y \mid f_2^{\uparrow}(x; y) < \infty\}$ [**15**, Theorem 3]. Our assumption (4.2) thus provides the existence of a point $\tilde{y}$ such that $l_1(\tilde{y}) < \infty$ and l_2 is bounded above on a neighbourhood of $\tilde{y}$. Hence

$$\partial(l_1+l_2)(0) = \partial l_1(0) + \partial l_2(0) = \partial f_1(x) + \partial f_2(x). \tag{4.6}$$

The combination of (4.5) and (4.6) yields (4.4). Note that (4.6) also establishes the weak*-closedness of $\partial f_1(x) + \partial f_2(x)$. If also $l_0 \geqslant l_1 + l_2$, the inclusion in (4.5) can be reversed to obtain equality in (4.4).

We proceed now to prove (4.3), that is,

$$l_0(y) \leqslant l_1(y) + l_2(y). \tag{4.7}$$

To start with, we consider y in the set (4.2). Let $\beta > l_2(y) = f_2^{\uparrow}(x; y)$. Since f_2 is directionally Lipschitzian at x, and y belongs to (4.2), the limit (2.13) for f_2 is $f_2^{\uparrow}(x; y)$ [**15**, Theorem 3], so it is less than β. Hence there exist $Y_0 \in \mathcal{N}(y)$, $X_0 \in \mathcal{N}(x)$, $\delta_0 > 0$, and $\lambda_0 > 0$, such that

$$\frac{f_2(x'+ty') - \alpha_2'}{t} < \beta \tag{4.8}$$

whenever

$$y' \in Y_0, \quad t \in (0, \lambda_0), \quad x' \in X_0, \quad \alpha_2' \geqslant f_2(x'), \quad |\alpha_2' - f(x')| \leqslant \delta_0.$$

On the other hand, we have by definition

$$l_0(y) = \sup_{Y \in \mathcal{N}(y)} \left[\limsup_{\substack{t \downarrow 0 \\ (x',\alpha') \to (x, f_0(x)) \\ \text{with } \alpha' \geqslant f_0(x')}} \left[\inf_{y \in Y} \frac{f_0(x'+ty') - \alpha'}{t} \right] \right].$$

Here the difference quotient can be expressed as

$$\frac{f_1(x'+ty') - \alpha_1'}{t} + \frac{f_2(x'+ty') - \alpha_2'}{t}, \tag{4.9}$$

where $\alpha_1' + \alpha_2' = \alpha'$, and the 'lim sup' can then be taken equivalently subject to $t \downarrow 0$, $x' \to x$, $\alpha_1' \to f_1(x)$, and $\alpha_2' \to f_2(x)$ with $\alpha_1' \geqslant f_1(x')$ and

$\alpha_2' \geqslant f_2(x')$. Invoking (4.8) in (4.9), we obtain

$$l_0(y) \leqslant \sup_{Y \in \mathcal{N}(y)} \left[\limsup \left[\inf_{y' \in Y} \left(\frac{f_1(x'+ty')-\alpha_1'}{t}+\beta\right)\right]\right]$$
$$= f_1^{\uparrow}(x;\, y)+\beta = l_1(y)+\beta,$$

where the 'lim sup' is as just described. Since this is true for all $\beta > l_2($ we conclude (4.7).

Having established (4.7) for y in the set (4.2), we now consider general If either $l_1(y) = +\infty$ or $l_2(y) = +\infty$, the right-hand side of (4.7) is $+\infty$ a the inequality is trivial. Suppose therefore that y belongs to $D_1 \cap D_2$, whe

$$D_i = \{y \mid l_i(y) < \infty\} \quad \text{(convex)}.$$

By hypothesis there exists $\tilde{y} \in D_1 \cap \operatorname{int} D_2$ (this is the assertion of (4.2 Then by convexity we have

$$(1-\varepsilon)y+\varepsilon\tilde{y} \in D_1 \cap \operatorname{int} D_2 \quad \text{for } \varepsilon \in (0, 1).$$

Points of this kind therefore fall within the case already treated, so th

$$l_0((1-\varepsilon)y+\varepsilon\tilde{y}) \leqslant l_1((1-\varepsilon)y+\varepsilon\tilde{y})+l_2((1-\varepsilon)y+\varepsilon\tilde{y}). \tag{4.10}$$

The functions l_i are convex and lower semicontinuous, so

$$\lim_{\varepsilon \downarrow 0} l_i((1-\varepsilon)y+\varepsilon\tilde{y}) = l_i(y)$$

(cf. [**17**, Corollary 7.5.1]). Taking the limit as $\varepsilon \downarrow 0$ on both sides of (4.1 we obtain (4.7) as desired.

Last on the agenda are the assertions of Theorem 2 about subdifferenti regularity. We have already seen that if either $l_1(0) = -\infty$ or $l_2(0) = -$ the inequality $l_0 \leqslant l_1+l_2$ already established implies equality in (4.4) the trivial sense that both sides represent the empty set. Suppose ther fore that $l_1(0) = 0$ and $l_2(0) = 0$ (i.e. that $\partial f_1(x)$ and $\partial f_2(x)$ are non-empt so in fact $l_1(y) > -\infty$ and $l_2(y) > -\infty$ for all y). If f_1 and f_2 are su differentially regular, then

$$l_i(y) = \liminf_{\substack{y' \to y \\ t \downarrow 0}} \frac{f_i(x+ty')-f_i(x)}{t} \quad \text{for } i = 1, 2. \tag{4.11}$$

Since $l_i(y) > -\infty$, we do not have to worry about the conventi $\infty - \infty = \infty$ in the expression

$$\frac{f_1(x+ty')-f_1(x)}{t}+\frac{f_2(x+ty')-f_2(x)}{t} = \frac{f_0(x+ty')-f_0(x)}{t}$$

(at least in the limit), and the inequality

$$l_1(y)+l_2(y) \leqslant \liminf_{\substack{y' \to y \\ t \downarrow 0}} \frac{f_0(x+ty')-f_0(x)}{t} \leqslant l_0(y) \tag{4.12}$$

is therefore valid. In conjunction with $l_0 \leqslant l_1 + l_2$, we conclude that equality holds in (4.12). Thus equality holds in (4.3) and (4.4), and f_0 is subdifferentially regular at x.

The consequences of Theorem 2 are many and diverse, because of all the criteria furnished in Theorem 1 and the special formulas for $f_i^{\uparrow}(x; y)$ provided in [**15**] for various cases. In fact, the remainder of this paper is essentially an exploration of corollaries of Theorem 2. The ones we now list demonstrate how Theorem 2 covers the main previous results of similar nature.

COROLLARY 1 [**12**]. *Let f_1 and f_2 be convex functions that are finite at x. Suppose there exists $\tilde{x}$ such that $f_1(\tilde{x}) < \infty$ and f_2 is bounded above on a neighbourhood of $\tilde{x}$. Then $\partial(f_1+f_2)(x) = \partial f_1(x) + \partial f_2(x)$.*

Proof. Convexity implies that $f_i^{\uparrow}(x; y) \leqslant f_i'(x; y) < \infty$ for every y such that $f_i(x+\lambda y) < \infty$ for some $\lambda > 0$ [**15**, Theorem 2]. Hence the vector $\tilde{y} = \tilde{x} - x$ belongs to the set (4.2). Theorem 1(b) asserts that f_2 is directionally Lipschitzian at x, and Proposition 3 confirms that f_1 and f_2 are subdifferentially regular.

COROLLARY 2. *Suppose that f_1 is finite at x and f_2 is Lipschitzian on a neighbourhood of x. Then $\partial(f_1+f_2)(x) \subset \partial f_1(x) + \partial f_2(x)$, and there is equality if f_1 and f_2 are also subdifferentially regular at x.*

Proof. Apply Theorem 1(a). The zero vector belongs to (4.2).

Note that Clarke's theorem [**5**, Proposition 8] required both f_1 and f_2 to be Lipschitzian around x (and E to be normed). We have seen in the Corollary of Proposition 1 that 'subdifferential regularity' reduces to Clarke's 'regularity' for Lipschitzian functions.

COROLLARY 3. *Let C_1 and C_2 be subsets of E, and let $x \in C_1 \cap C_2$. Suppose that*

(4.12) $$T_{C_1}(x) \cap \operatorname{int} T_{C_2}(x) \neq \emptyset,$$

and that C_2 is epi-Lipschitzian at x. Then

(4.13) $$T_{C_1 \cap C_2}(x) \supset T_{C_1}(x) \cap T_{C_2}(x),$$

(4.14) $$N_{C_1 \cap C_2}(x) \subset N_{C_1}(x) + N_{C_2}(x),$$

where the set on the right in (4.14) *is weak*-closed. Equality holds in* (4.13) *and* (4.14) *if C_1 and C_2 are tangentially regular at x, and then $C_1 \cap C_2$ is likewise tangentially regular.*

Proof. Apply Theorem 2 to the indicators $f_1 = \psi_{C_1}$, $f_2 = \psi_{C_2}$, usi
condition (e) of Theorem 1.

The finite-dimensional case of Corollary 3 was proved by the author
[16, Theorem 5] in connection with the following simplification.

COROLLARY 4. *Suppose $E = \mathbf{R}^n$. Then the hypothesis in Theorem 2 th f_2 is directionally Lipschitzian can be replaced by the assumption that f_2 lower semicontinuous in a neighbourhood of x. Likewise, the hypothesis Corollary 3 that C_2 is epi-Lipschitzian at x can be replaced by the assumpti that C_2 is closed relative to a neighbourhood of x.*

Proof. Invoke case (f) of Theorem 1.

5. Composition of a function and a differentiable mapping

Theorem 2 and its Corollary 3 can be used to derive formulas for number of other situations where it may not seem, at first sight, that sum of functions is involved. We demonstrate this first for the operati of composition.

A mapping $F: E \to E_1$, where E_1 is another linear topological spa (locally convex, Hausdorff), will be called *strictly differentiable at x in t full limit sense* if there is a continuous linear mapping $A: E \to E_1$ such th

$$\lim_{\substack{x' \to x,\, y' \to y \\ t \downarrow 0}} \frac{F(x'+ty') - F(x')}{t} = A(y) \quad \text{for all } y \in E. \tag{5.1}$$

Just as in the case of $E_1 = \mathbf{R}$ considered in § 3, this property implies stri differentiability in the *Hadamard sense* (defined in the obvious way), a the two are equivalent when E and E_1 are normable. For finite-dime sional spaces, they are equivalent also to strict differentiability in t Fréchet sense.

THEOREM 3. *Let $f = g \circ F$, where g is an extended-real-valued function E_1 and F is a mapping from E to E_1. Suppose that F is strictly differentia at x (in the full limit sense) with derivative $A: E \to E_1$, and that g is finite a directionally Lipschitzian at $F(x)$ with*

$$(\text{range}\, A) \cap \text{int}\{v \mid g^{\uparrow}(F(x); v) < \infty\} \neq \emptyset. \tag{5.2}$$

Then

$$f^{\uparrow}(x; y) \leqslant g^{\uparrow}(F(x); A(y)) \quad \textit{for all } y, \tag{5.3}$$

$$\partial f(x) \subset A^*[\partial g(F(x))]. \tag{5.4}$$

Equality holds in (5.4) if g is subdifferentially regular at x. It also holds (5.3) if, in addition, $g^{\uparrow}(F(x); 0) \neq -\infty$ (that is, $\partial g(F(x)) \neq \emptyset$).

Proof. Let $u = F(x)$, and define h on $E \times E_1$ by

$$h(x', u') = \begin{cases} f(x') & \text{if } u' = F(x'), \\ +\infty & \text{otherwise.} \end{cases} \tag{5.5}$$

Note that $h = f_1 + f_2$, where f_1 is the indicator of the graph of F and $f_2(x', u') = g(u')$. We claim

$$h^{\uparrow}(x, u; y, v) = \begin{cases} f^{\uparrow}(x; y) & \text{if } v = A(y), \\ +\infty & \text{if } v \neq A(y). \end{cases} \tag{5.6}$$

Indeed, from the definitions we have

$$h^{\uparrow}(x, u; y, v) = \limsup_{\substack{(x', u', \alpha') \downarrow_h (x, u) \\ t \downarrow 0}} \inf_{(y', v') \to (y, v)} \frac{h(x' + ty', u' + tv') - \alpha'}{t}, \tag{5.7}$$

where the condition $\alpha' \geqslant h(x', u')$ implicit in the first limit is equivalent to $\alpha' \geqslant f(x')$ and $u' = F(x')$, and where

$$h(x' + ty', u' + tv') = \begin{cases} f(x' + ty') & \text{if } v' = [F(x' + ty') - F(x')]/t, \\ +\infty & \text{otherwise.} \end{cases}$$

Since (5.1) holds, the mixed limit (5.7) is $+\infty$ if $v \neq A(x)$, while otherwise it is

$$h^{\uparrow}(x, u; y, v) = \limsup_{\substack{(x', \alpha') \downarrow_f x \\ t \downarrow 0}} \inf_{y' \to y} \frac{f(x' + ty') - \alpha'}{t} = f^{\uparrow}(x; y).$$

Thus (5.6) is valid, and it follows that

$$\begin{aligned} \partial h(x, u) &= \{(z, w) \mid h^{\uparrow}(x, u; y, v) \geqslant \langle y, z \rangle + \langle v, w \rangle \text{ for all } y, v\} \\ &= \{(z, w) \mid f^{\uparrow}(x; y) \geqslant \langle y, z \rangle + \langle A(y), w \rangle \text{ for all } y\} \\ &= \{(z, w) \mid f^{\uparrow}(x; y) \geqslant \langle y, z + A^*(w) \rangle \text{ for all } y\} \\ &= \{(z, w) \mid z + A^*(w) \in \partial f(x)\}, \end{aligned}$$

and consequently

$$\partial f(x) = \{z \mid (z, 0) \in \partial h(x, u)\}. \tag{5.8}$$

The next step is to apply Theorem 2 to the representation $h = f_1 + f_2$ noted above. The strict differentiability property (5.1) implies for the set $G = \operatorname{graph} F$ that

$$T_G(x, u) = K_G(x, u) = \operatorname{graph} A. \tag{5.9}$$

Thus f_1 is subdifferentially regular at (x, u) with

$$f_1^{\uparrow}(x, u; y, v) = \begin{cases} 0 & \text{if } v = A(y), \\ +\infty & \text{if } v \neq A(y), \end{cases} \tag{5.10}$$

$$\partial f_1(x, u) = N_G(x, u) = (\operatorname{graph} A)^{\circ} = \{(z, w) \mid z = -A^*(w)\}. \tag{5.11}$$

On the other hand, f_2 obviously inherits the directionally Lipschit property from g and has

(5.12) $$f_2^{\uparrow}(x, u; y, v) = g^{\uparrow}(u; v),$$

(5.13) $$\partial f_2(x, u) = \{(0, w) \mid w \in \partial g(u)\}.$$

In particular,

$$\{(y, v) \mid f_1^{\uparrow}(x, u; y, v) < \infty\} \cap \operatorname{int}\{(y, v) \mid f_2^{\uparrow}(x, u; y, v) < \infty\}$$
$$= \{(y, A(y)) \mid A(y) \in \operatorname{int}\{v \mid g^{\uparrow}(u; v) < \infty\}\},$$

and this set is non-empty by assumption (5.2). Therefore the hypoth of Theorem 2 is satisfied, and we have

(5.14) $$h^{\uparrow}(x, u; y, v) \leqslant f_1^{\uparrow}(x, u; y, v) + f_2^{\uparrow}(x, u; y, v),$$

(5.15) $$\partial h(x, u) \subset \partial f_1(x, u) + \partial f_2(x, u).$$

Combining (5.14) with (5.6), (5.10), and (5.12) we get (5.3). Simila (5.4) follows from (5.15), (5.8), (5.11), and (5.13).

For equality in these relations we need only have equality in (5 and (5.15). Since f_1 is already subdifferentially regular at (x, u) $\partial f_1(x, u) \neq \emptyset$, we can conclude the equality from Theorem 2 when subdifferentially regular with $\partial f_2(x, u) \neq \emptyset$. The latter properties equivalent to the corresponding ones for g.

COROLLARY 1 [**13**]. *Let $f = g \circ A$, where A is a continuous linear tr formation from E to E_1 and g is a convex function on E_1. Suppose finite at $A(x)$ and, for some $\tilde{x}$, bounded above in a neighbourhood of Then $\partial f(x) = A^*[\partial g(A(x))]$.*

Proof. Convexity implies that $g^{\uparrow}(u; v) \leqslant g'(u; v) < \infty$ for any u a such that $g(u)$ is finite and $g(u + \lambda v) < \infty$ for some $\lambda > 0$ [**15**, Theorem Hence the vector $\tilde{v} = A(\tilde{x}) - A(x)$ belongs to the set in (5.2). By Theo 1(b), g is directionally Lipschitzian at $A(x)$, and by Proposition 3 subdifferentially regular there.

COROLLARY 2 (Clarke [**6**, § 13]). *Let $f = g \circ F$, where F is a map from E (normed) to E_1 (normed) and g is Lipschitzian on a neighbourhoc $F(x)$. Suppose F is continuously Gâteaux differentiable in the sense tha limit*

$$F'(x'; y) = \lim_{t \downarrow 0} \frac{F(x' + ty) - F(x')}{t}$$

exists for all x', y, and the operators $F'(x; \cdot)$ are linear and continuous depend continuously on x' (in the norm topology for linear operators).

Then $\partial f(x) \subset A^*[\partial g(F(x))]$ (*where* $A = F'(x;\,\cdot)$), *and equality holds if* g *is* *subdifferentially regular at* $F(x)$.

Proof. The continuous Gâteaux differentiability of F implies by the generalized mean value theorem (cf. McLeod [**9**]) that F is strictly differentiable in the full limit sense. As noted in Theorem 1(a), g is directionally Lipschitzian at $F(x)$ if it is Lipschitzian in a neighbourhood of x; then $\partial g(x) \neq \emptyset$.

COROLLARY 3. *Let* $x \in C = F^{-1}(D)$, *where* $F\colon E \to E_1$ *and* $D \subset E_1$. *Suppose* F *is strictly differentiable at* x (*in the full limit sense*) *with derivative* A, *and* D *is epi-Lipschitzian at* $F(x)$ *with*

$$(\operatorname{range} A) \cap \operatorname{int} T_D(F(x)) \neq \emptyset.$$

Then

$$T_C(x) \supset A^{-1}[T_D(F(x))], \tag{5.16}$$

$$N_C(x) \subset A^*[N_D(F(x))]. \tag{5.17}$$

Equality holds in (5.16) *and* (5.17) *if* D *is also tangentially regular at* $F(x)$.

Proof. Apply Theorem 3 to $g = \psi_D$.

The finite-dimensional case of Corollary 3 was established by the author in [**16**, Theorem 5].

COROLLARY 4. *Suppose* $E = \mathbf{R}^n$. *Then the hypothesis that* g *is directionally Lipschitzian in Theorem* 3 *can be replaced by the condition that* g *is lower semicontinuous on a neighbourhood of* $F(x)$. *The assumption in Corollary* 3 *that* D *is epi-Lipschitzian can be replaced by the condition that* D *is closed in a neighbourhood of* $F(x)$.

Proof. This follows from criterion (f) of Theorem 1.

Theorem 3 can also be applied to the calculation of 'partial subgradients'. For a function g on a product space $E^1 \times E^2$ one may consider besides $\partial g(u_1, u_2)$ in $E^{1*} \times E^{2*}$ the sets

$$\partial_1 g(u_1, u_2) = \text{set of subgradients of } g(\cdot\,, u_2) \text{ at } u_1,$$

$$\partial_2 g(u_1, u_2) = \text{set of subgradients of } g(u_1, \cdot\,) \text{ at } u_2.$$

In general

$$\partial g(u_1, u_2) \not\subset \partial g(u_1, u_2) \times \partial_2 g(u_1, u_2) \not\subset \partial g(u_1, u_2).$$

Furthermore, there is no universal relationship between $\partial_1 g(u_1, u_2)$ and

$$\operatorname{proj}_1 \partial g(u_1, u_2) = \{w_1 \in E^{1*} \mid \text{there exists } w_2\colon (w_1, w_2) \in \partial g(u_1, u_2)\}$$

or between $\partial g_2(u_1, u_2)$ and $\operatorname{proj}_2 \partial g(u_1, u_2)$). However, the following corollary shows that when g is Lipschitzian around (u_1, u_2) one does have

$$\partial_1 g(u_1, u_2) \times \partial_2(u_1, u_2) \subset \operatorname{proj}_1 \partial g(u_1, u_2) \times \operatorname{proj}_2 \partial g(u_1, u_2),$$

and that equality holds if g is subdifferentially regular at (u_1, u_2). T
sharpens slightly an observation of Clarke [**6**, §1].

Corollary 5. *Let g be extended-real-valued on $E^1 \times E^2$ and finite directionally Lipschitzian at (u_1, u_2). Suppose*

$$[E^1 \times \{0\}] \cap \operatorname{int}\{(v_1, v_2) \mid g^{\uparrow}(u_1, u_2; v_1, v_2) < \infty\} \neq \emptyset$$

(as is true when g is Lipschitzian around (u_1, u_2)). Then

$$\partial_1 g(u_1, u_2) \subset \operatorname{proj}_1 \partial g(u_1, u_2),$$

and equality holds if g is also subdifferentially regular at (u_1, u_2).

Proof. Take $E = E^1$, $E_1 = E^1 \times E^2$, $F(u_1') = (u_1', u_2)$.

6. Other formulas

We turn now to conclusions that can be drawn from Theorem 2, more precisely from its Corollary 3 in §4, when one or both of the C_1, C_2 is an epigraph. The key to these is a sharper formula than (2 relating the normal cone of epi f at $(x, f(x))$ to $\partial f(x)$ in the case wh $\partial f(x) \neq \emptyset$.

Since $\partial f(x)$ is a non-empty weak*-closed convex set in E^*, it has associated *recession cone* (asymptotic cone):

$$0^+ \partial f(x) = \{w \mid \text{for all } z \in \partial f(x), \text{ for all } t \geqslant 0: z + tw \in \partial f(x)\}. \tag{6.1}$$

Formula (2.12) and the fundamental properties of recession cones [**17, 18**]) imply (because $\partial f(x) \neq \emptyset$) that

$$(w, 0) \in N_{\operatorname{epi} f}(x, f(x)) \Leftrightarrow w \in 0^+ \partial f(x). \tag{6.2}$$

Since $N_{\operatorname{epi} f}(x, f(x))$ is a weak*-closed convex cone contained in the h space $\{(z, \mu) \mid \mu \leqslant 0\}$ (because it is polar to $T_{\operatorname{epi} f}(x, f(x))$, which conta $(0, 1)$), we have

$$N_{\operatorname{epi} f}(x, f(x)) = \bigcup_{\lambda \geqslant 0^+} \lambda(\partial f(x), -1) \tag{6.3}$$

(when $\partial f(x) \neq \emptyset$), where the notation $\lambda \geqslant 0^+$ refers to all the cases wh $\lambda > 0$ and also the case $\lambda = 0^+$. Of course

$$0^+ \partial f(x) = \{0\} \quad \text{if } \partial f(x) \text{ is bounded}, \tag{6.4}$$

which is true when f is Lipschitzian in a neighbourhood of x. In this ev one could just as well write $\lambda \geqslant 0$ in (6.3).

Theorem 4. *Let $f = \max\{f_1, f_2\}$, and let x be a point where $f_1(x) = f$ and this value (which is $f(x)$) is finite. Suppose that*

$$\{y \mid f_1^{\uparrow}(x; y) < \infty\} \cap \operatorname{int}\{y \mid f_2^{\uparrow}(x; y) < \infty\} \neq \emptyset \tag{6.5}$$

and f_2 is directionally Lipschitzian at x. Then

(6.6) $$f^{\uparrow}(x;y) \leqslant \max\{f_1^{\uparrow}(x;y), f_2^{\uparrow}(x;y)\} \quad \text{for all } y.$$

If $\partial f_1(x)$ and $\partial f_2(x)$ are non-empty, then also

(6.7) $$\partial f(x) \subset \bigcup_{\substack{\lambda_1 \geqslant 0^+, \lambda_2 \geqslant 0^+ \\ \lambda_1 + \lambda_2 = 1}} [\lambda_1 \partial f_1(x) + \lambda_2 \partial f_2(x)],$$

where the set on the right-hand side of (6.7) *coincides with the weak*-closed convex hull of $\partial f_1(x)$ and $\partial f_2(x)$. Equality holds in* (6.6) *and* (6.7) *if f_1 and f_2 are also subdifferentially regular at x.*

Proof. All the conclusions are obtained by applying Corollary 3 of Theorem 2 to the epigraphs $C_i = \operatorname{epi} f_i$ at the point (x, α), where $\alpha = f_1(x) = f_2(x) = f(x)$, using (2.12) and (6.3). To see that the required assumptions are satisfied, note first that

(6.8) $$T_{C_1}(x,\alpha) \cap \operatorname{int} T_{C_2}(x,\alpha) = \operatorname{epi} f_1^{\uparrow}(x;\,\cdot\,) \cap \operatorname{int} \operatorname{epi} f_2^{\uparrow}(x;\,\cdot\,).$$

Let $D_i = \{y \mid f_i^{\uparrow}(x;y) < \infty\}$ for $i = 1, 2$. The convex function $f_2^{\uparrow}(x;\,\cdot\,)$ is continuous on $\operatorname{int} D_2$, because f_2 is directionally Lipschitzian at x [**15**, Theorem 3]. Therefore

$$\operatorname{int} \operatorname{epi} f_2^{\uparrow}(x;\,\cdot\,) = \{(y,\beta) \in E \times \mathbf{R} \mid y \in \operatorname{int} D_2, \beta > f_2^{\uparrow}(x;y)\}.$$

The condition that (6.8) be non-empty is thus equivalent to

$$D_1 \cap \operatorname{int} D_2 \neq \varnothing,$$

that is, to assumption (6.5).

COROLLARY (Clarke [**6**, Theorem 2]). *Let $f = \max\{f_1, \ldots, f_m\}$, where the functions f_i are Lipschitzian on a neighbourhood of x. Let*

$$I(x) = \{i \mid f_i(x) = f(x)\}.$$

Then

(6.9) $$f^{\uparrow}(x;y) \leqslant \max_{i \in I(x)} f_i^{\uparrow}(x;y) \quad \textit{for all } y,$$

(6.10) $$\partial f(x) \subset \bigcup \{\sum_{i \in I(x)} \lambda_i \partial f_i(x) \mid \lambda_i \geqslant 0, \sum_{i \in I(x)} \lambda_i = 1\}.$$

If f_i is subdifferentially regular at x for each $i \in I(x)$, then equality holds in (6.9) *and* (6.10).

Proof. Apply Theorem 4 inductively, using (6.4). For $i \notin I(x)$ we have $f_i(x') \leqslant f(x) - \varepsilon$ for all x' in some neighbourhood of x (for some $\varepsilon > 0$) by continuity (as implied by the Lipschitzian property), so $(x, f(x)) \in \operatorname{int} \operatorname{epi} f_i$, and $f^{\uparrow}(x;\,\cdot\,)$ and $\partial f(x)$ are independent of f_i.

REMARK. The cited result of Clarke is slightly more restrictive than tl corollary, in that it assumes the space is normed. However, it is mo general in an important respect: it covers certain kinds of infinite inde sets besides $\{1, \ldots, m\}$.

THEOREM 5. *Let* $C = \{x' \mid f(x') \leqslant 0\}$, *and let* x *be a point satisfyi* $f(x) = 0$. *Suppose that* f *is directionally Lipschitzian at* x *with* $0 \notin \partial f(x) \neq$ *and let* $D = \{y \mid f^{\uparrow}(x;\, y) < \infty\}$. *Then* C *is epi-Lipschitzian at* x *with*

$$T_C(x) \supset \{y \mid f^{\uparrow}(x;\, y) \leqslant 0\}, \tag{6.11}$$

$$\operatorname{int} T_C(x) \supset \{y \mid y \in \operatorname{int} D,\, f^{\uparrow}(x;\, y) > 0\} \neq \varnothing, \tag{6.12}$$

$$N_C(x) \subset \bigcup_{\lambda \geqslant 0^+} \lambda \partial f(x), \tag{6.13}$$

where the set on the right of (6.13) *is weak*-closed. If* f *is also subdifferential regular at* x, *then equality holds in* (6.11), (6.12), *and* (6.13), *and* C *is ta gentially regular at* x.

Proof. The conclusions are obtained by applying Corollary 3 of Theore 2 to $C_1 = \{(z, \mu) \in E \times \mathbf{R} \mid \mu = 0\}$ and $C_2 = \operatorname{epi} f$ at the point $(x, 0)$. One h

$$T_{C_1}(x, 0) = \{(z, \mu) \mid \mu = 0\}, \quad T_{C_2}(x, 0) = \operatorname{epi} f^{\uparrow}(x;\, \cdot), \tag{6.14}$$

and by polarity

$$N_{C_1}(x, 0) = \{(z, \mu) \mid z = 0\}, \quad N_{C_2}(x, 0) = \text{the cone in } (6.3). \tag{6.15}$$

Since the convex function $f^{\uparrow}(x;\, \cdot)$ is continuous on $\operatorname{int} D$ due to f bei directionally Lipschitzian at x [**15**, Theorem 3], one also has

$$\operatorname{int} \operatorname{epi} f^{\uparrow}(x;\, \cdot) = \{(y, \beta) \in E \times \mathbf{R} \mid y \in \operatorname{int} D,\, f^{\uparrow}(x;\, y) < \beta\}, \tag{6.16}$$

and this set is non-empty. Therefore by (6.14),

$$T_{C_1}(x, 0) \cap \operatorname{int} T_{C_2}(x, 0) = \{(y, 0) \mid y \in \operatorname{int} D,\, f^{\uparrow}(x;\, y) < 0\}. \tag{6.17}$$

The latter is also non-empty, for otherwise the set (6.16) would be co tained in the half-space $\{(y, \beta) \mid \beta \geqslant 0\}$, and the same would then be tr of its closure, which includes $\operatorname{epi} f^{\uparrow}(x;\, \cdot)$. That would imply $f^{\uparrow}(x;\, y) \geqslant$ for all y in contradiction with the hypothesis that $0 \notin \partial f(x)$. The assum tions of Corollary 3 of Theorem 2 are therefore satisfied. The assert relations follow at once from the ones in this result and (6.14), (6.15), (6.1

COROLLARY 1. *Let* $C = \{x' \mid f(x') \leqslant 0\}$, *and let* x *be a point satisfyi* $f(x) = 0$. *Suppose that* f *is Lipschitzian in a neighbourhood of* x, *a* $0 \notin \partial f(x)$. *Then* C *is epi-Lipschitzian at* x *with*

$$T_C(x) \supset \{y \mid \langle y, z\rangle \leqslant 0 \text{ for all } z \in \partial f(x)\}, \tag{6.18}$$

$$\operatorname{int} T_C(x) \supset \{y \mid \langle y, z\rangle < 0 \text{ for all } z \in \partial f(x)\}, \tag{6.19}$$

$$N_C(x) \subset \{\lambda z \mid \lambda \geqslant 0,\, z \in \partial f(x)\}, \tag{6.20}$$

where the set on the right in (6.20) *is weak*-closed. If f is also subdifferentially regular at z, then equality holds in* (6.18), (6.19), *and* (6.20), *and C is tangentially regular at x.*

Proof. Since f is Lipschitzian, we have $\partial f(x)$ non-empty and weak*-compact. Thus $0^+\partial f(x) = \{0\}$ and

$$f^\uparrow(x; y) = \max\{\langle y, z\rangle \mid z \in \partial f(x)\} \quad \text{(finite)}$$

for all $y \in E$.

COROLLARY 2. *Let $C = \{x' \mid f_i(x') \leqslant 0,\ i = 1, \ldots, m\}$. For $x \in C$, let $I(x) = \{i \mid f_i(x) = 0\}$. Suppose the functions f_i are Lipschitzian in a neighbourhood of x and*

(6.21) $$0 \notin \operatorname{co}\{\partial f_i(x) \mid i \in I(x)\}.$$

Then C is epi-Lipschitzian at x with

(6.22) $$T_C(x) \supset \{y \mid \langle y, z_i\rangle \leqslant 0 \text{ for all } z_i \in \partial f_i(x),\ i \in I(x)\},$$

(6.23) $$\operatorname{int} T_C(x) \subset \{y \mid \langle y, z_i\rangle < 0 \text{ for all } z_i \in \partial f_i(x),\ i \in I(x)\},$$

(6.24) $$N_C(x) \subset \{\sum_{i \in I(x)} \lambda_i z_i \mid \lambda_i \geqslant 0,\ z_i \in \partial f_i(x)\},$$

where the set on the right in (6.24) *is weak*-closed. If each f_i is also subdifferentially regular at x, then equality holds in* (6.22), (6.23), *and* (6.24), *and C is tangentially regular at x.*

Proof. Apply Corollary 1 with f as in the corollary of the preceding theorem.

Of course, Corollary 2 in turn implies Corollary 1.

7. Application to minimization

Only a very general sort of application can be discussed here, but it illustrates one of the fundamental possibilities of the theory.

THEOREM 6. *Suppose f has a finite local minimum at x relative to a set C (where f is an extended-real-valued function on E and $C \subset E$). Assume that either of the following two conditions is satisfied:*

(a) *$T_C(x) \cap \operatorname{int}\{y \mid f^\uparrow(x; y) < \infty\} \neq \emptyset$ and f is directionally Lipschitzian at x (if $E = \mathbf{R}^n$, the latter can be replaced by: f is l.s.c. on a neighbourhood of x), or*

(b) *$\{y \mid f^\uparrow(x; y) < \infty\} \cap \operatorname{int} T_C(x) \neq \emptyset$ and C is epi-Lipschitzian at x (if $E = \mathbf{R}^n$, the latter can be replaced by: C is closed relative to a neighbourhood of x).*

Then $\partial f(x)$ meets $-N_C(x)$.

Proof. The function $f+\psi_C$ has a local minimum at x, and hence x is substationary point. Either (a) or (b) is sufficient for Theorem 2 to b applicable and yield the inclusion

$$0 \in \partial(f+\psi_C)(x) \subset \partial f(x)+\partial\psi_C(x) = \partial f(x)+N_C(x).$$

This says there exists $z \in \partial f(x)$ with $-z \in N_C(x)$.

REMARK. The set C in Theorem 6 can in particular have the form i Corollary 2 of Theorem 5, and a Lagrange multiplier rule then follows b way of condition (b). At the same time, f can be of the form $f_0+\psi_D$, wher D is a set defined by constraints of some other type perhaps, and ∂f ca then be analysed by another application of Theorem 2. In this way on recovers by 'calculation' the Lagrange multiplier rule proved by Clark [5, Theorem 1] (except to the extent that the latter applies also to *equalit* constraints expressed by locally Lipschitzian, non-smooth functions; suc constraints involve additional considerations that go beyond the presen framework).

Finally, we describe an application to Ekeland's variational princip [7, Theorem 1.1].

THEOREM 7. *Let f be an extended-real-valued lower semicontinuou function on a Banach space E, and let x be a point where f is finite and has ρ-local ε-minimum, in the sense that*

$$f(x) \leqslant \inf\{f(x') \mid \|x'-x\| \leqslant \rho\}+\varepsilon \tag{7.1}$$

(where $0 < \rho \leqslant \infty$, $0 < \varepsilon < \infty$). Choose any $\lambda \in (0,\rho)$. Then there exist x *and $z_\lambda \in \partial f(x_\lambda)$ such that*

$$f(x_\lambda) \leqslant f(x), \quad \|x_\lambda - x\| \leqslant \lambda, \quad \|z_\lambda\| \leqslant \varepsilon/\lambda. \tag{7.2}$$

Proof. Let B and B^* denote the unit balls of E and E^*, and le $g = f+\psi_C$, where $C = x+\rho B$. The function g is lower semicontinuous an has a global ε-minimum at x.

Ekeland's variational principle [7] asserts for any $\lambda > 0$ the existence o x_λ such that $g(x_\lambda) \leqslant g(x)$ (hence $f(x_\lambda) \leqslant f(x)$), $\|x_\lambda - x\| \leqslant \lambda$, and

$$g(x') \geqslant g(x_\lambda)-(\varepsilon/\lambda)\|x'-x_\lambda\| \quad \text{for all } x' \in E. \tag{7.3}$$

Let $h(x') = (\varepsilon/\lambda)\|x'-x_\lambda\|$. We can interpret (7.3) as saying that $g+h$ ha its global minimum at x_λ. Assume that $\lambda < \rho$ so that $x_\lambda \in \operatorname{int} C$. Then coincides with f around x_λ, and $f+h$ therefore has a local minimum at x Hence $0 \in \partial(f+h)(x_\lambda)$. Theorem 2 is applicable, because h is Lipschitzia

$$\partial(f+h)(x_\lambda) \subset \partial f(x_\lambda)+\partial h(x_\lambda) = \partial f(x_\lambda)+(\varepsilon/\lambda)B^*.$$

Thus we may conclude that $0 \in \partial f(x_\lambda) + (\varepsilon/\lambda)B^*$, or, in other words, that there exists $z_\lambda \in \partial f(x_\lambda)$ with $\|z_\lambda\| \leqslant \varepsilon/\lambda$. This completes the proof of Theorem 7.

REMARK. When f is convex and $\rho = \infty$, Theorem 7 reduces to a result of Brøndsted and the author [2, Lemma] that has been used in the global study of the multifunctions ∂f in convex analysis.

REFERENCES

1. N. BOURBAKI, *Variétés différentielles et analytiques*, Eléments de Mathématiques XXXIII (Hermann, Paris, 1967).
2. A. BRØNSTED and R. T. ROCKAFELLAR, 'On the subdifferentiability of convex functions', *Proc. Amer. Math. Soc.* 16 (1965) 605–11.
3. F. H. CLARKE, 'Necessary conditions for nonsmooth problems in optimal control and the calculus of variations', Thesis, University of Washington, Seattle, 1973.
4. —— 'Generalized gradients and applications', *Trans. Amer. Math. Soc.* 205 (1975) 247–62.
5. —— 'A new approach to Lagrange multipliers', *Math. of Operations Research* 1 (1976) 165–74.
6. —— 'Generalized gradients of Lipschitz functionals', *Advances in Math.*, to appear.
7. I. EKELAND, 'On the variational principle', *J. Math. Anal. Appl.* 47 (1974) 324–53.
8. G. LEBOURG, 'Generic smoothness properties of locally Lipschitzian real-valued functions defined on open subsets of infinite-dimensional topological vector spaces', to appear.
9. R. M. MCLEOD, 'Mean value theorems for vector-valued functions', *Proc. Edinburgh Math. Soc.* (2) 14 (1965) 197–209.
10. L. MCLINDEN, 'Dual operations on saddle functions', *Trans. Amer. Math. Soc.* 179 (1973) 363–81.
11. R. T. ROCKAFELLAR, 'Convex functions and dual extremum problems', Thesis, Harvard University, 1963.
12. —— 'Extension of Fenchel's duality theorem for convex functions', *Duke Math. J.* 33 (1966) 81–90.
13. —— 'Integrals which are continuous functionals, II', *Pacific J. Math.* 39 (1971) 439–69.
14. —— *Conjugate duality and optimization*, Conference Board of Mathematical Sciences series 16 (SIAM Publications, Philadelphia, 1974).
15. —— 'Generalized directional derivatives and subgradients of nonconvex functions', *Canad. J. Math.*, to appear.
16. —— 'Clarke's tangent cones and the boundaries of closed sets in R^n', *Nonlin. Analysis*, 2, no. 6 (1978).
17. —— *Convex analysis* (Princeton University Press, 1970).
18. —— 'Level sets and continuity of conjugate convex functions', *Trans. Amer. Math. Soc.* 123 (1966) 46–63.

Department of Mathematics
University of Washington
Seattle, Washington 98195