Nonlinear Analysis, Theory, Methods & Applications, Vol. 9, No. 7, pp. 665–698, 1985.
Printed in Great Britain.
0362-546X/85 $3.00 + .00

EXTENSIONS OF SUBGRADIENT CALCULUS WITH APPLICATIONS TO OPTIMIZATION

R. T. ROCKAFELLAR*
Department of Mathematics, University of Washington, Seattle, WA 98195, U.S.A.

(*Received* 13 *February* 1984; *received for publication* 9 *October* 1984)

Key words and phrases: Subgradient calculus, optimization Lagrange multiplier rules.

1. INTRODUCTION

IN 1973, CLARKE [2] introduced a concept of subgradient for nonconvex, extended-real-valued functions which made possible a far-reaching generalization of the subgradient theory of convex functions [8]. Subgradients in Clarke's sense have subsequently been studied by many authors in both finite and infinite-dimensional spaces; see [3] and [9] for expositions. They have especially turned out to be useful in analyzing problems of optimization, for instance in characterizing solutions and in obtaining conditions for stability under perturbations of data. Central to this purpose are the calculus rules that have been developed for expressing the subgradients of a given function in terms of the subgradients of other functions from which it is constructed. For instance, a great many optimization problems can be formulated in terms of minimizing an extended-real-valued function f over R^n, and the subgradient condition $0 \in \partial f(\bar{x})$ is then necessary for f to have a local minimum at $\bar{x}$ [13, p. 333]. The usefulness of this condition obviously depends, however, on the strength of the rules that are available for calculating $\partial f(\bar{x})$.

This paper is devoted to proving sharper or broader versions of a number of such basic calculus rules. New theorems on Lagrange multipliers in problems of constrained minimization are obtained along the way. For technical reasons connected with the nature of our approach only finite-dimensional spaces are considered here, although advances in underlying theory such as the recent results of Treiman [15] may eventually make possible some extensions to a Banach space setting.

A brief review of basic notions will help to fix notation and terminology. The geometric route to defining subgradients, which was followed by Clarke, depends on first defining cones of normals to an arbitrary closed set $C \subset R^n$. Let us say that a vector $y \in R^n$ is a *proximal normal* to C at a point $\bar{x} \in C$ if for $t > 0$ sufficiently small, the unique point of C nearest to $\bar{x} + ty$ (in the Euclidean norm) is $\bar{x}$. It is a *limiting proximal normal* if there exist points $x^k \in C, x^k \to \bar{x}$, and proximal normals y^k to C at x^k, such that $y^k \to y$. Let

$$\hat{N}_C(\bar{x}) := \{y \mid y \text{ is a limiting proximal normal to } C \text{ at } \bar{x}\}. \tag{1.1}$$

Then Clarke's *normal* cone to C at $\bar{x}$ is

$$N_C(\bar{x}) := \text{cl co } \hat{N}_C(\bar{x}), \tag{1.2}$$

where cl stands for closure and co stands for convex hull. Thus $N_C(\bar{x})$ is always a closed convex

* Research supported by the Air Force Office of Scientific Research, U.S.A.F., under grant number F49620-82-K-0012.

cone containing 0. It is known that $N_C(\bar{x})$ contains some $y \neq 0$ if and only if $\bar{x}$ is a boundary point of C (cf. [10]).

Consider now a function $f: R^n \to \bar{R} = R \cup \{\pm\infty\}$ and a point $\bar{x}$ where f is finite and *strictly lower semicontinuous*; we mean by the latter that for some $\bar{\alpha} \geqq f(\bar{x})$ the function $\min\{f, \bar{\alpha}\}$ is lower semicontinuous on a neighborhood of $\bar{x}$, or in other words, that the epigraph set

$$\operatorname{epi} f = \{(x, \alpha) \in R^n \times R \mid \alpha \geqq f(x)\} \tag{1.3}$$

is closed relative to some neighborhood of $(\bar{x}, f(\bar{x}))$. The normal cone $N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))$ is then well defined and consists of certain vectors $(y, \eta) \in R^n \times R$ such that $\eta \leqq 0$. Clarke's set of subgradients of f at $\bar{x}$ is defined geometrically as

$$\partial f(\bar{x}) := \{y \in R^n \mid (y, -1) \in N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))\}. \tag{1.4}$$

If f is convex, this is the usual subgradient set of convex analysis, whereas if f is strictly differentiable at $\bar{x}$ it reduces to $\{\nabla f(\bar{x})\}$. (In particular f is strictly differentiable at $\bar{x}$ if f is continuously differentiable on a neighborhood of $\bar{x}$; see Clarke [3, p. 30] for more on this concept.)

It is useful sometimes also to consider the set of *singular subgradients* of f at $\bar{x}$, which is

$$\partial^\infty f(\bar{x}) := \{y \in R^n \mid (y, 0) \in N_{\operatorname{epi} f}(\bar{x}, f(\bar{x}))\}. \tag{1.5}$$

Clearly $\partial f(\bar{x})$ is a closed convex set (possibly empty), and $\partial^\infty f(\bar{x})$ is a closed convex cone containing 0. The directions of the rays comprising $\partial^\infty f(\bar{x})$ may be interpreted as the "elements of $\partial f(\bar{x})$ which lie at ∞" (cf. [2, Section 8]: $\partial^\infty f(\bar{x})$ is the recession cone of $\partial f(\bar{x})$ when $\partial f(\bar{x}) \neq \phi$).

One has $\partial^\infty f(\bar{x}) = \{0\}$ if and only if $\partial f(\bar{x})$ is nonempty and bounded, which in turn holds if and only if f is Lipschitz continuous on a neighborhood of $\bar{x}$ (cf. [6]). In fact the Lipschitz modulus.

$$\operatorname{Lip}_f(\bar{x}) := \limsup_{\substack{x \to \bar{x} \\ x' \to \bar{x} \\ x \neq x'}} \frac{|f(x') - f(x)|}{|x' - x|} \tag{1.6}$$

then satisfies

$$\operatorname{Lip}_f(\bar{x}) = \sup\{|y| \mid y \in \partial f(\bar{x})\}. \tag{1.7}$$

If $f = \delta_C$ (the *indicator* of a set C, $\delta_C(x) := 0$ if $x \in C$, ∞ if $x \notin C$), then

$$\partial \delta_C(\bar{x}) = \partial^\infty \delta_C(\bar{x}) = N_C(\bar{x}). \tag{1.8}$$

The sets $\partial f(\bar{x})$ and $\partial^\infty f(\bar{x})$ can also be defined in a dual manner in terms of a certain kind of directional derivative expression for f at $\bar{x}$. This corresponds geometrically to an expression for $N_C(\bar{x})$ as the polar of a certain kind of tangent cone to C at $\bar{x}$; see [3] and [11]. We shall not need to go into this here, but the reader should bear in mind that all our results have a dual statement along such lines.

The chief tool in our approach is the extended limit characterization of $\partial f(\bar{x})$ and $\partial^\infty f(\bar{x})$ which was derived in [12] and makes it possible to reduce questions about these sets to questions about local minima. A vector $y \in R^n$ is said to be a *proximal subgradient* of f at $\bar{x}$ if for some $\tau > 0$

$$f(x) \geqq f(\bar{x}) + y \cdot (x - \bar{x}) - r|x - \bar{x}|^2 \text{ in a neighborhood of } \bar{x},$$

or equivalently

$$f(x) - y \cdot (x - \bar{x}) + r|x - \bar{x}|^2 \text{ has a local minimum at } x = \bar{x}.$$

It is a *limiting proximal subgradient* of f at $\bar{x}$ if there exist points $x^k \to \bar{x}$ with $f(x^k) \to f(\bar{x})$ and proximal subgradients y^k at x^k such that $y^k \to y$. It is a *singular limiting proximal subgradient* of f at $\bar{x}$ if the same holds, except that instead of $y^k \to y$ one has $t_k y^k \to y$ for some sequence of numbers $t_k > 0$, $t_k \to 0$. (Except in the trivial case of $y = 0$, the latter is equivalent to $y^k/|y^k| \to y/|y|$ and means that the sequence y^k converges to the "point at infinity" in the direction of the ray $\{ty \mid t \geqq 0\}$.) Let

$$\hat{\partial} f(\bar{x}) := \{y \mid y \text{ is a limiting proximal subgradient of } f \text{ at } \bar{x}\}, \tag{1.9}$$

$$\hat{\partial}^\infty f(\bar{x}) := \{y \mid y \text{ is a singular limiting proximal subgradient of } f \text{ at } \bar{x}\}. \tag{1.10}$$

Obviously $\hat{\partial} f(\bar{x})$ is a closed set and $\partial^\infty f(\bar{x})$ is a closed cone (a *cone* in R^n being a set K such that $ty \in K$ whenever $y \in K, t > 0$). We proved in [12] that

$$\hat{\partial} f(\bar{x}) = \{y \mid (y, -1) \in \hat{N}_{\mathrm{epi} f}(\bar{x}, f(\bar{x}))\}, \tag{1.11}$$

$$\hat{\partial}^\infty f(\bar{x}) = \{y \mid (y, 0) \in \hat{N}_{\mathrm{epi} f}(\bar{x}, f(\bar{x}))\}, \tag{1.12}$$

and consequently that $0 \in \partial^\infty f(\bar{x})$ and

$$\partial f(\bar{x}) = \mathrm{cl\,co}[\hat{\partial} f(\bar{x}) + \hat{\partial}^\infty f(\bar{x})], \tag{1.13}$$

a formula which can be interpreted as saying that $\partial f(\bar{x})$ is the closed convex hull in the extended sense of [8, Section 17] of the points in $\hat{\partial} f(\bar{x})$ and the "points at infinity" represented by the rays in the cone $\hat{\partial}^\infty f(\bar{x})$. One also has

$$\partial^\infty f(\bar{x}) \supset \mathrm{cl\,co}\, \hat{\partial}^\infty f(\bar{x}) \tag{1.14}$$

and

$$\partial^\infty f(x) = \mathrm{cl\,co}\, \hat{\partial}^\infty f(x) \quad \text{when} \quad \partial f(\bar{x}) = \phi. \tag{1.15}$$

One actually has (cf. [11, proposition 15]):

$$\partial f(\bar{x}) = \mathrm{co}[\hat{\partial} f(\bar{x}) + \hat{\partial}^\infty f(\bar{x})] \quad \text{and} \quad \hat{\partial} f(\bar{x}) = \mathrm{co}\, \hat{\partial}^\infty f(\bar{x}) \quad \text{if the cone } \hat{\partial}^\infty f(\bar{x}) \text{ is pointed.} \tag{1.16}$$

(Recall that a cone K, containing 0 but not necessarily convex, is said to be *pointed* in the equation $y_1 + \ldots + y_m = 0$ for elements $y_i \in K$ is possible only when $y_1 = \ldots = y_m = 0$.)

2. TECHNICAL PRELIMINARIES

For use in subsequent arguments, we need to state several results that are already known or easily follow from results already known.

PROPOSITION 2.1. Let $p: R^m \to \bar{R}$ be finite and strictly lower semicontinuous at $\bar{u}$. Suppose $M(\bar{u})$ and $M^\infty(\bar{u})$ are sets in R^m such that $M^\infty(\bar{u})$ is a cone and $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$, $\hat{\partial}^\infty p(\bar{u}) \subset$ $\subset M^\infty(\bar{u})$. Then

$$\begin{aligned}\partial p(\bar{u}) &= \mathrm{cl\,co}[M(\bar{u}) \cap \partial p(u) + M^\infty(\bar{u}) \cap \partial^\infty p(\bar{u})] \\ &\subset \mathrm{cl\,co}[M(\bar{u}) + M^\infty(u)],\end{aligned} \tag{2.1}$$

and

$$\partial^\infty p(\bar{u}) = \mathrm{cl\,co}[M^\infty(\bar{u}) \cap \partial^\infty p(\bar{u})] \subset \mathrm{cl\,co}\, M^\infty(\bar{u}) \qquad \text{if} \quad \partial p(\bar{u}) = \varnothing. \tag{2.2}$$

Thus

$$M(\bar{u}) \cap \partial p(\bar{u}) \neq \phi \qquad \text{if} \quad \partial p(\bar{u}) \neq \varnothing. \tag{2.3}$$

If $M^\infty(\bar{u})$ is pointed, then $\partial^\infty p(\bar{u})$ is pointed and one actually has

$$\partial p(\bar{u}) = \mathrm{co}[M(\bar{u}) \cap \partial p(\bar{u}) + M^\infty(\bar{u}) \cap \partial^\infty p(\bar{u})] \subset \mathrm{co}[M(\bar{u}) + M^\infty(\bar{u})], \tag{2.5}$$

$$\partial^\infty p(\bar{u}) = \mathrm{co}[M^\infty(\bar{u}) \cap \partial^\infty p(\bar{u}))] \subset \mathrm{co}\, M^\infty(\bar{u}). \tag{2.6}$$

In particular, if $M^\infty(\bar{u}) = \{0\}$ then $\partial p(\bar{u})$ is nonempty and bounded with $\partial p(\bar{u}) \subset \mathrm{co}\, M(\bar{u})$, and p is Lipschitz continuous on a neighborhood of $\bar{u}$ with

$$\mathrm{Lip}_p(\bar{u}) \leqq \sup_{z \in M(\bar{u})} |z|. \tag{2.7}$$

Proof. All these conclusions are apparent from the basic facts about $\partial p(\bar{u})$, $\partial^\infty p(\bar{u})$, $\hat{\partial} p(\bar{u})$ and $\hat{\partial}^\infty p(\bar{u})$ that we reviewed in Section 1; cf. Rockafellar [12, theorem 1] and [11, proposition 15]. ∎

COROLLARY 2.1.1. Let $E \subset R^m$ be closed relative to some neighborhood of the point $\bar{u} \in E$. Suppose $M(\bar{u}) \in R^m$ is a cone such that $\hat{N}_E(\bar{u}) \subset M(\bar{u})$. Then

$$N_E(\bar{u}) = \mathrm{cl\,co}[M(\bar{u}) \cap N_E(\bar{u})] \subset \mathrm{cl\,co}\, M(\bar{u}). \tag{2.8}$$

Thus

$$M(\bar{u}) \cap N_E(\bar{u}) \neq \{0\} \qquad \text{if} \quad N_E(\bar{u}) \neq \{0\}. \tag{2.9}$$

If $M(\bar{u})$ is pointed, then $N_E(\bar{u})$ is pointed and one actually has

$$N_E(\bar{u}) = \mathrm{co}[M(\bar{u}) \cap N_E(\bar{u})] \subset \mathrm{co}\, M(\bar{u}). \tag{2.10}$$

In particular, if $M(\bar{u}) = \{0\}$ then $\bar{u} \in \mathrm{int}\, E$.

Proof. Let $p = \delta_E$ in proposition 2.1. ∎

Our mode of operation on several occasions will be to define appropriate sets $M(\bar{u})$ and $M^\infty(\bar{u})$ for a given function p, verify that $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$ and $\hat{\partial}^\infty p(\bar{u}) \subset M^\infty(\bar{u})$ and then refer to the conclusions of proposition 2.1.

Two results about the behavior of subgradients under limits are recorded next.

PROPOSITION 2.2. Let $f: R^n \to \bar{R}$ be finite and strictly lower semicontinuous at $\bar{x}$. Let $x^k \to \bar{x}$ with $f(x^k) \to f(\bar{x})$, and suppose $y^k \in \partial f(x^k)$, $y^k \to y$. If $\hat{\partial}^\infty f(\bar{x})$ is pointed (as is true in particular when $\hat{\partial}^\infty f(\bar{x}) = \{0\}$, i.e. when f is Lipschitz continuous in a neighborhood of $\bar{x}$), then $y \in \partial f(\bar{x})$.

Proof. If $\hat{\partial}^\infty f(\bar{x})$ is pointed, then so is $\partial f^\infty(\bar{x})$ by proposition 2.1. Then the cone $N_{\mathrm{epi}\, f}(\bar{x}, f(\bar{x}))$ must also be pointed by virtue of (1.4), (1.5), and the fact that all the elements (y, η) of

$N_{\mathrm{epi}f}(\bar{x}, f(\bar{x}))$ have $\eta \leqq 0$. According to [10, corollary 2 on p. 149], the graph of the multifunction $N_{\mathrm{epi}f}$ is then closed at $(\bar{x}, f(\bar{x}))$. Thus from having $(y^k, -1) \in N_{\mathrm{epi}f}(x^k, f(x^k))$, $(x^k, f(x^k)) \to (\bar{x}, f(\bar{x}))$, $(y^k, -1) \to (y, -1)$, we may conclude that $(y, -1) \in N_{\mathrm{epi}f}(\bar{x}, f(\bar{x}))$, i.e. $y \in \partial f(\bar{x})$. ■

For f Lipschitz continuous on a neighborhood of $\bar{x}$, the conclusion $y \in \hat{\partial} f(\bar{x})$ in proposition 2.2 was established originally by Clarke [4]. A generalization for f not Lipschitzian was given by Rockafellar [10, corollary 3 to theorem 2]. The assumptions there were, in effect, that $\partial^\infty f(\bar{x})$ is pointed and $\partial f(\bar{x}) \neq \emptyset$.

PROPOSITION 2.3. (Cf. [11, Proposition 14].) Suppose for $k = 1, 2, \ldots$, that x^k furnishes a finite local minimum of $g + h_k$, where $g: R^n \to \bar{R}$ and $h_k: R^n \to \bar{R}$ are lower semicontinuous and nowhere $-\infty$. If $x^k \to \bar{x}$ with $g(x^k) \to g(\bar{x})$ (finite) and h_k is Lipschitz continuous on a neighborhood of x^k with

$$\mathrm{Lip}_{h_k}(x^k) \to 0, \tag{2.11}$$

then $0 \in \partial g(\bar{x})$.

Proof. This differs from [11, proposition 14] only in the substitution of (2.11) for the condition that $\partial h_k(x^k)$ is nonempty and $\sup\{|z| \mid z \in \partial h_k(x^k)\} \to 0$. The two conditions are equivalent by (1.7); recall that h_k is Lipschitz continuous on a neighborhood of x^k if and only if $\partial h_k(x^k)$ is nonempty and bounded. ■

Rules for calculating $\partial(g + h)$ in terms of ∂g and ∂h are important in a number of situations, and we shall prove a new one in Section 7. As a stepping stone the following known special case will be needed.

PROPOSITION 2.4. (Cf. [13].) Let $g: R^n \to \bar{R}$ be lower semicontinuous, and let $h: R^n \to \bar{R}$ be Lipschitz continuous in a neighborhood of $\bar{x}$, a point where both g and h are finite. Then

$$\partial(g + h)(\bar{x}) \subset \partial g(\bar{x}) + \partial h(\bar{x}), \quad \partial^\infty(g + h)(\bar{x}) = \partial^\infty g(\bar{x}). \tag{2.12}$$

Proof. The inclusion in (2.12) is asserted directly by [13, corollary 2 of theorem 2], but the equation for singular subgradients requires putting some separate facts together. The cited theorem also gives us in [13, (4.3)] the subderivative inequality

$$(g + h)^\uparrow(\bar{x}; w) \leqq g^\uparrow(\bar{x}; w) + h^\uparrow(\bar{x}; w) \qquad \text{for all} \quad w \in R^n$$

(see [13] or [14] for the definition of these expressions), where $h^\uparrow(\bar{x}; w)$ is finite for all w because h is locally Lipschitzian. This implies

$$\{w \in R^n \mid (g + h)^\uparrow(\bar{x}; w) < \infty\} \supset \{w \in R^n \mid g^\uparrow(\bar{x}; w) < \infty\}.$$

The two sets in this inclusion are convex cones whose polars are $\partial^\infty(g + h)(\bar{x})$ and $\partial^\infty g(\bar{x})$, respectively. (In more detail: the epigraph of $g^\uparrow(\bar{x}; \cdot)$ is the closed convex tangent cone $T_{\mathrm{epi}g}(\bar{x}, g(\bar{x}))$ in $R^n \times R$ by [14, theorem 2], and the set $\{w \mid g^\uparrow(\bar{x}; w) < \infty\}$ is the projection of this cone on R^n. The polar projection is then the set of vectors y such that $(y, 0)$ belongs to the polar cone $N_{\mathrm{epi}g}(\bar{x}, g(\bar{x}))$, which is $\partial^\infty g(\bar{x})$ by definition.) Taking polars on both sides of the inclusion we obtain $\partial^\infty(g+h)(\bar{x}) \subset \partial^\infty g(\bar{x})$. But the opposite inequality then follows from this fact as applied to $g+h$ in place of g:

$$\partial^\infty g(\bar{x}) = \partial^\infty(g+h-h)(\bar{x}) \subset \partial^\infty(g+h)(\bar{x}),$$

since $-h$ too is locally Lipschitzian. Hence the equation in (2.12) is correct. ∎

COROLLARY 2.4.1. Let $h: R^n \to \bar{R}$ be Lipschitz continuous in a neighborhood of $\bar{x} \in C$, where $C \subset R^n$ is a closed set. Then

$$\partial(h+\delta_C)(\bar{x}) \subset \partial h(\bar{x}) + N_C(\bar{x}), \qquad \partial^\infty(h+\delta_C)(\bar{x}) = N_C(\bar{x}). \tag{2.13}$$

Proof. Apply (1.8). ∎

COROLLARY 2.4.2. (Cf. [11, proposition 5].) Let $g: R^n \to \bar{R}$ be finite and strictly lower semicontinuous at $\bar{x}$, and let h be continuously differentiable on a neighborhood of $\bar{x}$. Then

$$\partial(g+h)(\bar{x}) = \partial g(\bar{x}) + \{\nabla h(\bar{x})\}, \qquad \partial^\infty(g+h)(\bar{x}) = \partial^\infty g(\bar{x}). \tag{2.14}$$

Proof. One has h Lipschitz continuous at $\bar{x}$ with $\partial h(\bar{x}) = \{\nabla h(\bar{x})\}$. Apply (2.12) to $g+h$ and $g = (g+h) - h$. ∎

Other cases besides corollary 2.4.2 where equality holds for the inclusion in (2.12) are described in [13, theorem 2] but will not be required here. One case where equality holds that has not previously been noted is the following.

PROPOSITION 2.5. Let $f(x, w) = g(x) + h(w)$, where $g: R^n \to \bar{R}$ and $h: R^m \to \bar{R}$ nowhere have the value $-\infty$. If $\bar{x}$ is a point where g is finite and strictly lower semicontinuous, and $\bar{w}$ is a point where h is finite and strictly lower semicontinuous, then

$$\partial f(\bar{x}, \bar{w}) = (\partial g(\bar{x}), \partial h(\bar{w})), \qquad \partial^\infty f(\bar{x}, \bar{w}) = (\partial^\infty g(\bar{x}), \partial^\infty h(\bar{w})). \tag{2.15}$$

Proof. We shall base ourselves on formula (1.13). We demonstrate first that

$$\hat{\partial} f(\bar{x}, \bar{w}) = (\hat{\partial} g(\bar{x}), \hat{\partial} h(\bar{w})). \tag{2.16}$$

If $(y, v) \in \hat{\partial} f(\bar{x}, \bar{w})$, we have $(y^k, v^k) \to (y, v)$, where (y^k, v^k) is a proximal subgradient. If f at (x^k, w^k) with $(x^k, w^k) \to (\bar{x}, \bar{w})$ and $f(x^k, w^k) \to f(\bar{x}, \bar{w})$. Then for certain $r_k > 0$ the function

$$(x, w) \to g(x) + h(w) - y^k[x - x^k] - z^k[w - w^k] + r_k(|x - x^k|^2 + |w - w^k|^2)$$

has a local minimum at (x^k, w^k). But this is the same as saying that the function

$$x \to g(x) - y^k[x - x^k] + r_k|x - x^k|^2$$

has a local minimum at x^k, while the function

$$w \to h(w) - z^k[w - w^k] + r_k|w - w^k|^2$$

has a local minimum at w^k. Thus y^k is a proximal subgradient of g at x^k, and z^k is a proximal subgradient of h at w^k. Since $f(x^k, w^k) \to f(\bar{x}, \bar{w})$ if and only if $g(x^k) \to g(\bar{x})$ and $h(w^k) \to h(\bar{w})$ (due to the lower semicontinuity assumption), we see that $y \in \partial g(\bar{x})$ and $z \in \partial h(\bar{w})$. Thus the inclusion $\subset$ holds in (2.16). The proof of the inclusion $\supset$ is essentially a reversal of this argument.

One verifies similarly that

$$\hat{\partial}^\infty f(\bar{x}, \bar{w}) \subset (\hat{\partial}^\infty g(\bar{x}), \hat{\partial}^\infty h(\bar{w})) = (\hat{\partial}^\infty g(\bar{x}), 0) + (0, \hat{\partial}^\infty h(\bar{w})),$$
$$\hat{\partial}^\infty f(\bar{x}, \bar{w}) \supset (\hat{\partial}^\infty g(\bar{x}), 0) \cup (0, \hat{\partial}^\infty h(\bar{w})),$$

which implies (since the sets in question are cones containing 0) that

$$\operatorname{co} \hat{\partial}^\infty f(\bar{x}, \bar{w}) = \operatorname{co}(\partial^\infty g(\bar{x}), \partial^\infty h(\bar{w})). \tag{2.17}$$

From (2.16) and (2.17) we obtain

$$\begin{aligned}\operatorname{co}[\hat{\partial} f(\bar{x}, \bar{w}) + \hat{\partial}^\infty f(\bar{x}, \bar{w})] &= \operatorname{co}[\hat{\partial} f(\bar{x}, \bar{w}) + \operatorname{co} \hat{\partial}^\infty f(\bar{x}, \bar{w})] \\ &= [(\hat{\partial} g(\bar{x}), \hat{\partial} h(\bar{w})) + \operatorname{co}(\hat{\partial}^\infty g(\bar{x}), \hat{\partial}^\infty h(\bar{w}))] \\ &= \operatorname{co}[(\hat{\partial} g(\bar{x}). \hat{\partial} h(\bar{w})) + (\hat{\partial}^\infty g(\bar{x}), \hat{\partial}^\infty h(\bar{w}))] \\ &= \operatorname{co}(\hat{\partial} g(\bar{x}) + \hat{\partial}^\infty g(\bar{x}), \hat{\partial} h(\bar{w}) + \hat{\partial}^\infty h(\bar{w})) \\ &= (\operatorname{co}[\hat{\partial} g(\bar{x}) + \hat{\partial}^\infty g(\bar{x})], \operatorname{co}[\hat{\partial} h(\bar{w}) + \hat{\partial} h^\infty(\bar{w})])\end{aligned}$$

and consequently by (1.13) that

$$\begin{aligned}\partial f(\bar{x}, \bar{w}) &= (\operatorname{cl}\operatorname{co}[\hat{\partial} g(\bar{x}) + \hat{\partial}^\infty g(\bar{x})], \operatorname{cl}\operatorname{co}[\hat{\partial} h(\bar{w}) + \hat{\partial}^\infty h(\bar{w})]) \\ &= (\partial g(\bar{x}), \partial h(\bar{w}))\end{aligned}$$

as claimed in the first part of (2.15).

If $\partial f(\bar{x}, \bar{w}) = \emptyset$, the second part of (2.15) follows from (2.17) via the general formula (1.15). If $\partial f(\bar{x}, \bar{w}) \neq \emptyset$ and $\partial h(\bar{w}) \neq \emptyset$, and the recession cones in the closed convex sets $\partial f(\bar{x}, \bar{w})$, $\partial g(\bar{x})$, and $\partial h(\bar{w})$, are $\partial^\infty f(\bar{x}, \bar{w})$, $\partial^\infty g(\bar{x})$ and $\partial^\infty h(\bar{w})$, respectively (cf. (1.4) and (1.5); for the theory of recession cones, see [8, Section 8]). Then the recession cone of $(\partial g(\bar{x}), \partial h(\bar{w}))$ is $(\partial^\infty g(\bar{x}), \partial^\infty h(\bar{w}))$, so the second part of (2.15) is implied by the first. ∎

COROLLARY 2.5.1. Let $C \subset R^n$ and $D \subset R^m$ be closed sets, and let $\bar{x} \in C$, $\bar{w} \in D$. Then

$$N_{C\times D}(\bar{x}, \bar{w}) = (N_C(\bar{x}), N_D(\bar{w})). \tag{2.18}$$

Proof. Take $g = \delta_C$, $h = \delta_D$ in proposition 2.5. ∎

Finally we need to recall Clarke's concept of the "generalized Jacobian" $\partial F(\bar{x})$ of a locally Lipschitzian mapping $F: R^n \to R^m$ at a point $\bar{x}$. This is defined as the convex hull of the set of all $m \times n$ matrices of the form $A = \lim_k \nabla F(x^k)$ where $\nabla F(x^k)$ is the Jacobian of F at a point x^k at which F happens to be differentiable and $x^k \to \bar{x}$. The set $\partial F(x^k)$ is nonempty and compact, as well as convex (cf. [3, 2.6]). In fact the Lipschitz modulus

$$\operatorname{Lip}_F(\bar{x}) := \limsup_{\substack{x \to \bar{x} \\ x' \to \bar{x} \\ x \neq x'}} \frac{|F(x') - F(x)|}{|x' - x|} \tag{2.19}$$

satisfies

$$\operatorname{Lip}_F(\bar{x}) = \max_{A \in \partial F(\bar{x})} |A|. \tag{2.20}$$

For notational convenience, we shall write

$$\partial F(\bar{x})v = \{Av \mid A \in \partial F(\bar{x})\}$$
$$y\partial F(\bar{x}) = \{yA \mid A \in \partial F(\bar{x})\}. \tag{2.21}$$

PROPOSITION 2.6 [3]. Suppose $h(x) = y \cdot F(x)$ where $F: R^n \to R^m$ is locally Lipschitzian and $y \in R^m$. Then

$$\partial h(\bar{x}) = y\partial F(\bar{x}). \tag{2.22}$$

Proof. Let S be the set of points x where the Jacobian $\nabla F(x)$ does not exist. By Rademacher's theorem, S is of measure 0 because F is locally Lipschitzian. For $x \notin S$, we have $\nabla h(x) = y\nabla F(\bar{x})$. The function h is itself locally Lipschitzian, so $\partial h(\bar{x})$ is the closed convex hull of all vectors of the form $\lim \nabla h(x^k)$ for $x^k \notin S$, $x^k \to \bar{x}$ (see [8, theorem 25.1]). Thus it is the closed convex hull of all vectors of the form $y \lim \nabla F(x^k)$ for $x^k \notin S, x^k \to \bar{x}$. This is the same as $y\partial F(\bar{x})$. ■

Proposition 2.6 can be viewed as an elementary case of a chain rule of Clarke [3, theorem 2.6.6] where F is composed with the mapping $z \to y \cdot z$ from R^m to R.

3. BASIC THEOREMS ON PERTURBATIONS

We now prove a result which will be the key to a number of new subgradient formulas. It crystallizes the basic principle used by the author in deriving Lagrange multiplier rules in [11] and [12].

THEOREM 3.1. Let $P: R^n \times R^m \to \bar{R}$ be lower semicontinuous, and consider

$$p(u) := \inf_x P(x, u), \qquad X(u) := \operatorname{argmin}_x P(x, u). \tag{3.1}$$

Let $\bar{u}$ be a point where p is finite and the following holds:

$$\exists \eta > 0 \text{ and } \bar{\alpha} > p(\bar{u}) \text{ such that the set}$$
$$\{x \mid \exists u \text{ with } |u - \bar{u}| \leqq \eta, P(x, u) \leqq \bar{\alpha}\} \text{ is bounded.} \tag{3.2}$$

Then p is strictly lower semicontinuous at $\bar{u}$, and for all u satisfying $|u - \bar{u}| \leqq \eta, p(u) \leqq \bar{\alpha}$, the set $X(u)$ is nonempty and compact. When $x^k \in X(u^k)$ with $u^k \to \bar{u}$ and $p(u^k) \to p(\bar{u})$, then the sequence $\{x^k\}$ is bounded and all of its cluster points belong to $X(\bar{u})$. Moreover, for

$$M(\bar{u}) := \bigcup_{\bar{x} \in X(\bar{u})} \{y \mid (0, y) \in \hat{\partial} P(\bar{x}, \bar{u})\}, \tag{3.3}$$
$$M^\infty(\bar{u}) := \bigcup_{\bar{x} \in X(\bar{u})} \{y \mid (0, y) \in \hat{\partial}^\infty P(\bar{x}, \bar{u})\},$$

one has $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$ and $\hat{\partial}^\infty p(\bar{u}) \subset M^\infty(\bar{u})$, so all the assertions in proposition 2.1 are valid.

Remark 3.1. The first person to develop inf-compactness conditions like (3.2) in order to conclude the lower semicontinuity of p in such an abstract setting was Wets [16].

Proof. For arbitrary $\alpha < \bar{\alpha}$, the set

$$\{(x,u) \in R^n \times R^m \mid |u - \bar{u}| \leqq \eta,\ P(x,u) \leqq \bar{\alpha}\} \tag{3.4}$$

is bounded because the set in (3.2) is bounded, and it is closed because P is lower semicontinuous. Hence it is compact. In particular, for fixed u and α satisfying $|u - \bar{u}| \leqq \eta$ and $p(u) \leqq \alpha \leqq \bar{\alpha}$ (e.g. $u = \bar{u}$ and $\alpha = \bar{\alpha}$), the set of points x such that (x, u) belongs to (3.4) is compact, so $X(u)$ is nonempty and compact. Thus when $\alpha \leqq \bar{\alpha}$, the set $\{u \in R^m \mid |u - \bar{u}| \leqq \alpha\}$ is the image of the compact set (3.4) under the projection $(x,u) \to u$ and is itself compact, hence closed. This tells us that the function $u \to \min\{p(u), \alpha\}$ is lower semicontinuous on a neighborhood of $\bar{u}$. In other words, p is strictly lower semicontinuous at $\bar{u}$.

Consider now a sequence $u^k \to \bar{u}$ such that $p(u^k) \to p(\bar{u})$ (finite). Suppose $x^k \in X(u^k)$. For k large enough that $|u^k - \bar{u}| \leqq \eta$ and $p(u^k) \leqq \bar{\alpha}$, the set $X(u^k)$ is included in the set in (3.2) (which is bounded), so $\{x^k\}$ is bounded. Let $\bar{x}$ be a cluster point of $\{x^k\}$; for simplicity of notation, we can suppose $x^k \to \bar{x}$. We have $P(x^k, u^k) = p(u^k) \to p(\bar{u})$, and since P is lower semicontinuous this implies $P(\bar{x}, \bar{u}) \leqq p(\bar{u})$. But also $p(\bar{u}) \leqq P(\bar{x}, \bar{u})$ by the definition of p. Therefore $P(\bar{x}, \bar{u}) = p(\bar{u})$, so that $\bar{x} \in X(\bar{u})$.

Proof of the last statement of the theorem requires showing that the sets defined in (3.3) satisfy $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$ and $\hat{\partial}^\infty p(\bar{u}) \subset M^\infty(\bar{u})$. Let $y \in \hat{\partial} p(\bar{u})$: for a certain sequence $u^k \to \bar{u}$ with $p(u^k) \to p(\bar{u})$ one has $y^k \to y$ with y^k a proximal subgradient of p at u^k, i.e. such that for certain numbers $r_k > 0$ the function

$$u \to p(u) - y^k \cdot u + r_k |u - u^k|^2 \tag{3.5}$$

has a local minimum at u^k. When k is large enough that $|u^k - \bar{u}| \leqq \eta$ and $p(u^k) \leqq \alpha$, there exists $x^k \in X(u^k)$. The function

$$(x,u) \to P(x,u) - u^k \cdot u + r_k |u - u^k|^2 \tag{3.6}$$

then has a local minimum at (x^k, u^k), and so also does the function

$$(x,u) \to P(x,u) - (0, y^k) \cdot (x,u) + r_k |(x,u) - (x^k, u^k)|^2.$$

Thus $(0, y^k)$ is a proximal subgradient of P at (x^k, u^k). As seen above, the sequence $\{x^k\}$ is bounded and has all its cluster points in $X(\bar{u})$. Passing to subsequences, we can arrange that $x^k \to \bar{x} \in X(\bar{u})$. Then $P(x^k, u^k) = p(u^k) \to p(\bar{u}) = P(\bar{x}, \bar{u})$, and since $(0, y^k)$ is a proximal subgradient of P at (x^k, u^k) with $(0, y^k) \to (0, \bar{y})$, we conclude that $(0, \bar{y}) \in \hat{\partial} P(\bar{x}, \bar{u})$.

This proves the inclusion $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$. The proof of $\hat{\partial}^\infty p(\bar{u}) \subset M^\infty(\bar{u})$ is the same, except that instead of $y^k \to y$ one has $t_k y^k \to y$, $t_k \downarrow 0$. ■

COROLLARY 3.1.1. For p as in theorem 3.1, the assertions of proposition 2.1 are also valid if instead of (3.3) one takes

$$M(\bar{u}) = \bigcup_{\bar{x} \in X(\bar{u})} \{y \mid (0,y) \in \partial P(\bar{x}, \bar{u})\},$$

$$M^\infty(\bar{u}) = \bigcup_{\bar{x} \in X(\bar{u})} \{y \mid (0,y) \in \partial^\infty P(\bar{x}, \bar{u})\}. \tag{3.7}$$

In the special case where $X(\bar{u})$ consists of a unique $\bar{x}$, this yields

$$\partial p(\bar{u}) \subset \{y \mid (0,y) \in \partial P(\bar{x}, \bar{u})\}, \qquad \partial^\infty p(\bar{u}) \subset \{y \mid (0,y) \in \partial^\infty P(\bar{x}, \bar{u})\}. \tag{3.8}$$

Proof. The first is true because $\hat{\partial}P(\bar{x}, \bar{u}) \subset \partial P(\bar{x}, \bar{u})$ and $\hat{\partial}^\infty P(\bar{x}, \bar{u}) \subset \partial^\infty P(\bar{x}, \bar{u})$. The second is true because $\partial P(\bar{x}, \bar{u})$ and $\partial^\infty P(\bar{x}, \bar{u})$ are closed convex sets with $\partial P(\bar{x}, \bar{u}) + \partial^\infty P(\bar{x}, \bar{u}) = \partial P(\bar{x}, \bar{u})$. ■

COROLLARY 3.1.2. Let $D \subset R^n \times R^m$ be a closed set, and let E be its projection on R^m:

$$E = \{u \mid X(u) \neq \varnothing\}, \qquad \text{where} \quad X(u) = \{x \mid (x, u) \in D\}. \tag{3.9}$$

Let $\bar{u} \in E$ and suppose

$$\begin{gathered}\exists \eta > 0 \text{ such that the set} \\ \{x \mid \exists u \text{ with } |u - \bar{u}| \leqq \eta, (x, u) \in D\} \text{ is bounded.}\end{gathered} \tag{3.10}$$

Then E is closed relative to the neighborhood $\{u \mid |u - \bar{u}| \leqq \eta\}$, and for all $u \in E$ in this neighborhood the set $X(u)$ is compact. Whenever $x^k \in X(u^k)$ with $u^k \to \bar{u}$, then the sequence $\{x^k\}$ is bounded, and all of its cluster points belong to $X(\bar{u})$. Moreover for

$$M(\bar{u}) := \bigcup_{\bar{x} \in X(\bar{u})} \{y \mid (0, y) \in \hat{N}_D(\bar{x}, \bar{u})\} \tag{3.11}$$

all the assertions of corollary 2.1.2 are valid, and the same is true for

$$M(\bar{u}) := \bigcup_{\bar{x} \in X(\bar{u})} \{y \mid (0, y) \in N_D(\bar{x}, \bar{u})\}. \tag{3.12}$$

Proof. Take $P = \delta_D$ in theorem 3.1 and corollary 3.1.1. ■

The next theorem generalizes in several ways the main perturbation result of [11, theorem 2]. The latter corresponds to the case where f is locally Lipschitzian as well as F, and C has the special form in remark 3.2 below. The framework in [11] allows for a broader class of perturbations than the ones presently under consideration, however. Such perturbations can also be treated at the new level of generality (see theorem 8.3), but only after we have developed the machinery much further.

THEOREM 3.2. For closed sets $C \subset R^m$ and $D \subset R^n$, a locally Lipschitzian mapping $F: R^n \to R^m$ and lower semicontinuous function $f: R^n \to \bar{R}$, consider

$$\begin{aligned} p(u) &:= \inf_x \{f(x) \mid F(x) + u \in C, x \in D\}, \\ X(u) &:= \operatorname{argmin}_x \{f(x) \mid F(x) + u \in C, x \in D\}. \end{aligned} \tag{3.13}$$

Let $\bar{u}$ be a point where p is finite and the following holds:

$$\begin{gathered}\exists \eta > 0 \quad \text{and} \quad \bar{\alpha} > p(\bar{u}) \text{ such that the set} \\ \{x \in D \mid f(x) \leqq \bar{\alpha} \text{ and } \exists u \text{ with } |u - \bar{u}| \leqq \eta, F(x) + u \in C\} \text{ is bounded.}\end{gathered} \tag{3.14}$$

Then p is strictly lower semicontinuous at $\bar{u}$, and $X(\bar{u})$ is nonempty and compact. Moreover for

$$\begin{aligned} M(\bar{u}) &:= \bigcup_{\bar{x} \in X(\bar{u})} \{y \in N_C(F(\bar{x}) + \bar{u}) \mid 0 \in \partial(f + \delta_D)(\bar{x}) + y\partial F(\bar{x})\} \\ M^\infty(\bar{u}) &:= \bigcup_{\bar{x} \in X(\bar{u})} \{y \in N_C(F(\bar{x}) + \bar{u}) \mid 0 \in \partial^\infty(f + \delta_D)(\bar{x}) + y\partial F(\bar{x})\}, \end{aligned} \tag{3.15}$$

one has $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$ and $\hat{\partial}^\infty p(\bar{u}) \subset M^\infty(\bar{u})$, so all the assertions in proposition 2.1 are valid.

Remark 3.2. The condition $F(x) + u \in C$ reduces to the standard constraint system in mathematical programming when

$$C = \{w = (w_1, \ldots, w_m) \mid w_i \leqq 0 \text{ for } i = 1, \ldots, s, w_i = 0 \text{ for } i = s+1, \ldots, m\}. \tag{3.16}$$

Indeed, writing $F(x) = (f_1(x), \ldots, f_m(x))$ (with $f_i : R^n \to R$ locally Lipschitzian) and $u = (u_1, \ldots, u_m)$, one has $F(x) + u \in C$ if and only if

$$f_i(x) + u_i \begin{cases} \leqq 0 & \text{for } i = 1, \ldots, s, \\ = 0 & \text{for } i = s+1, \ldots, m. \end{cases} \tag{3.17}$$

Then if $\bar{x}$ and $\bar{u}$ are such that these constraints are satisfied, one has for $y = (y_1, \ldots, y_m)$ that

$$y \in N_C(F(\bar{x}) + \bar{u}) \Leftrightarrow \begin{cases} y_i \geqq 0 & \text{for } i \leqq s \text{ with } f_i(\bar{x}) + \bar{u}_i = 0, \\ y_i = 0 & \text{for } i \leqq s \text{ with } f_i(\bar{x}) + \bar{u}_i < 0. \end{cases} \tag{3.18}$$

Of course if $\bar{x} \in \operatorname{int} D$ (e.g. if $D = R^n$), the term $N_D(\bar{x})$ is superfluous in (3.15).

Proof of theorem 3.2. Let

$$P(x, u) = \begin{cases} f(x) & \text{if } F(x) + u \in C, x \in D, \\ +\infty \text{ otherwise.} \end{cases} \tag{3.19}$$

Then P is lower semicontinuous, and $p(u)$ and $X(u)$ are as in theorem 3.1. Also, condition (3.2) is satisfied by (3.14). Hence the semicontinuity properties in theorem 3.1 hold for $p(u)$ and $X(u)$.

We must show that the sets in (3.15) include $\hat{\partial} p(\bar{u})$ and $\hat{\partial}^\infty p(\bar{u})$ respectively. This will prove the theorem.

Suppose that y^k is a proximal subgradient of p at u^k, $u^k \to \bar{u}$, $p(u^k) \to p(\bar{u})$. Then for certain $r_k > 0$ the function (3.5) has a local minimum at u^k. We know from theorem 3.1 that when k is large enough there exists $x^k \in X(u^k)$, and that by passing to subsequences if necessary, we can arrange to have $x^k \to \bar{x} \in X(\bar{u})$, so that

$$f(x^k) = p(u^k) \to p(\bar{u}) = f(\bar{x}) \text{ (finite).} \tag{3.20}$$

The function (3.6) then has a local minimum at (x^k, u^k). Setting $w = F(x) + u$, $\bar{w} = F(\bar{x}) + \bar{u}$ and $w^k = F(x^k) + u^k \to \bar{w}$, we can express this as follows: the function

$$f_k(w, x) = f(x) - y^k[w - F(x)] + r_k|w - F(x) - w^k + F(x^k)|^2 + \delta_C(w) + \delta_D(x) \tag{3.21}$$

has a local minimum at (w^k, x^k).

To prove $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$, we suppose $y^k \to y$ and write $f_k = g + h_k$, where

$$g(w, x) = [-y \cdot w + \delta_C(w)] + [f(x) + y \cdot F(x) + \delta_D(x)], \tag{3.22}$$

$$h_k(w, x) = -(y^k - y) \cdot [w - F(x)] + r_k|w - F(x) - w^k + F(x^k)|^2, \tag{3.23}$$

and note that

$$\operatorname{Lip}_{h_k}(w^k, x^k) \leqq |y^k - y|(1 + \operatorname{Lip}_F(x^k)) \to 0.$$

(The "quadratic" term in the definition of h_k has Lipschitz modulus 0 at (w^k, x^k).) Then $g + h_k$ has a local minimum at (w^k, x^k), and the hypothesis of proposition 2.3 is satisfied, so we obtain $(0, 0) \in \partial g(\bar{w}, \bar{x})$. (Clearly g is lower semicontinuous, while h is finite everywhere and locally Lipschitzian. Although g could have the value $-\infty$ somewhere if f did, we do know that $f(\bar{x})$ is finite and f is lower semicontinuous, and therefore that $f(x) > -\infty$ when $|x - \bar{x}| \leqq \delta$, say. Redefining $f(x) = +\infty$ when $|x - \bar{x}| > \delta$ would ensure that g is nowhere $-\infty$ and would lead to the same conclusion, since $\partial g(\bar{w}, \bar{x})$ is unaffected by any modification of g outside of a neighborhood of $(\bar{w}, \bar{x})$.)

We next apply proposition (2.5) to reduce the condition $(0, 0) \in \partial g(\bar{w}, \bar{x})$ to get $0 \in \partial g_1(\bar{w})$ and $0 \in \partial g_2(\bar{w})$, where $g_1(w)$ and $g_2(x)$ are the two terms on the right in (3.22). Since the function $w \to -y \cdot w$ is differentiable, we have

$$\partial g_1(\bar{w}) = -y + \partial\delta_C(\bar{w}) = -y + N_C(\bar{w}) \tag{3.24}$$

by corollary 2.4.2 and (1.8). Thus $0 \in \partial g_1(\bar{w})$ if and only if $y \in N_C(\bar{w}) = N_C(F(\bar{x}) + \bar{u})$, while $0 \in \partial g_2(\bar{x})$ means that $0 \in \partial(f + y \cdot F + \delta_D)(\bar{x})$. But

$$\partial(f + y \cdot F + \delta_D)(\bar{x}) \subset \partial(f + \delta_D)(\bar{x}) + y\partial F(x) \tag{3.25}$$

by proposition 2.4, corollary 2.4.1, and proposition 2.6. We have demonstrated therefore that any $y \in \hat{\partial} p(\bar{u})$ belongs to one of the sets in the union defining $M(\bar{u})$ in (3.15). Thus $\partial p(\bar{u}) \subset M(\bar{u})$.

To prove that $\partial^\infty p(\bar{u}) \subset M_\infty(\bar{u})$, we take the same reasoning down a slightly different path. Instead of $y_k \to y \in \hat{\partial} p(\bar{u})$ we suppose $t_k y^k \to y \in \hat{\partial}^\infty p(\bar{u})$, where $t_k \downarrow 0$. The fact that the function f_k in (3.21) has a local minimum at (w^k, x^k) is interpreted as saying that the function

$$\begin{aligned}\tilde{f}_k(w, x, \alpha) = {} & t_k\alpha + \delta_{\operatorname{epi} f}(x, \alpha) - t_k y^k \cdot [w - F(x)] \\ & + t_k r_k |w - F(x) - w^k + F(x^k)|^2 + \delta_C(w) + \delta_D(x)\end{aligned}$$

has a local minimum at $(\bar{w}, \bar{x}, f(\bar{x}))$. We write $\tilde{f}_k = \tilde{g} + \tilde{h}_k$, where

$$\tilde{g}(w, x, \alpha) = [-y \cdot w + \delta_C(w)] + [y \cdot F(x) + \delta_D(x) + \delta_{\operatorname{epi} f}(x, \alpha)], \tag{3.26}$$

$$\tilde{h}_k(w, x, \alpha) = t_k\alpha - (t_k y^k - y) \cdot [w - F(x)] + t_k r_k |w - F(x) - w^k + F(x^k)|^2. \tag{3.26$'$}$$

Then $\tilde{g} + \tilde{h}_k$ has a local minimum at $(w^k, x^k, f(x^k))$ with $(w^k, x^k, f(x^k)) \to (\bar{w}, \bar{x}, f(\bar{x}))$ and

$$\operatorname{Lip}_{\tilde{h}_k}(w^k, x^k, f(x^k)) \leqq t_k + |t_k y^k - y|(1 + \operatorname{Lip}_F(\bar{x})) \to 0,$$

and we obtain from proposition 2.3 that $(0, 0, 0) \in \partial\tilde{g}(\bar{w}, \bar{x}, f(\bar{x}))$. The latter reduces by proposition 2.5 to $0 \in \partial\tilde{g}_1(\bar{w})$ and $(0, 0) \in \partial\tilde{g}_2(\bar{x}, f(\bar{x}))$, where $\tilde{g}_1(w)$ and $\tilde{g}_2(x, \alpha)$ are the two expressions in (3.26). The first condition is again equivalent to $y \in N_C(\bar{w}) = N_C(F(\bar{x}) + \bar{u})$ by (3.24). On the other hand we can write

$$\tilde{g}_2(x, \alpha) = y \cdot F(x) + \delta_{\operatorname{epi}\varphi}(x, \alpha) \qquad \text{for} \quad \varphi = f + \delta_D$$

and deduce from corollary 2.4.1 and proposition 2.6 that

$$\partial\tilde{g}_2(\bar{x}, f(\bar{x})) \subset (y\partial F(\bar{x}), 0) + N_{\operatorname{epi}\varphi}(\bar{x}, f(\bar{x})).$$

The condition $(0,0) \in \partial\bar{g}_2(\bar{x}, f(\bar{x}))$ therefore implies the existence of some $z \in y\partial F(\bar{x})$ such that $(-z, 0) \in N_{\text{epi}\,\varphi}(\bar{x}, f(\bar{x}))$, i.e. $-z \in \partial^{\infty}\varphi(\bar{x})$ (cf. (1.5)). Thus it implies $0 \in \partial^{\infty}(f + \delta_D)(\bar{x}) + y\partial F(\bar{x})$. This establishes that $y \in M^{\infty}(\bar{u})$ and finishes the proof of the theorem. ■

Remark 3.3. The proof of theorem 3.2 reveals that the same conclusions would hold if one took in place of (3.15):

$$
\begin{aligned}
M(\bar{u}) &:= \bigcup_{\bar{x} \in X(\bar{u})} \{y \in N_C(F(\bar{x}) + \bar{u}) \,|\, 0 \in \partial(f + yF + \delta_D)(\bar{x})\},\\
M^{\infty}(\bar{u}) &:= \bigcup_{\bar{x} \in X(\bar{u})} \{y \in N_C(F(\bar{x}) + \bar{u}) \,|\, 0 \in \partial^{\infty}(f + yF + \delta_D)(\bar{x})\}.
\end{aligned}
\tag{3.27}
$$

This set $\overline{M(u)}$ is sometimes smaller than the one in (3.15) (never larger), so the result is slightly sharper when stated in this way. (The set $M^{\infty}(\bar{u})$ is the same in (3.15) and (3.27) by virtue of proposition 2.4.)

COROLLARY 3.2.1. If f is locally Lipschitzian in theorem 3.2, the same conclusions hold with (3.15) replaced by

$$
\begin{aligned}
M(\bar{u}) &:= \bigcup_{\bar{x} \in X(\bar{u})} \{y \in N_C(F(\bar{x}) + \bar{u}) \,|\, 0 \in \partial f(\bar{x}) + y\partial F(\bar{x}) + N_D(\bar{x})\},\\
M^{\infty}(\bar{u}) &:= \bigcup_{\bar{x} \in X(\bar{u})} \{y \in N_C(F(\bar{x}) + \bar{u}) \,|\, 0 \in y\partial F(\bar{x}) + N_D(\bar{x})\}.
\end{aligned}
\tag{3.28}
$$

Proof. The sets (3.28) include the ones in (3.15) in this case by corollary 2.4.1. ■

Remark 3.4. Under any circumstances guaranteeing for all $\bar{x} \in \bar{X}(\bar{u})$ that

$$
\partial(f + \delta_D)(\bar{x}) \subset \partial f(\bar{x}) + N_D(\bar{x}) \quad \text{and} \quad \partial^{\infty}(f + \delta_D)(\bar{x}) \subset \partial^{\infty} f(\bar{x}) + N_D(\bar{x}), \tag{3.29}
$$

the conclusions of theorem 3.2 also hold with (3.15) replaced by

$$
\begin{aligned}
M(\bar{u}) &:= \bigcup_{\bar{x} \in X(\bar{u})} \{y \in N_C(F(\bar{x}) + \bar{u}) \,|\, 0 \in \partial f(\bar{x}) + y\partial F(\bar{x}) + N_D(\bar{x})\},\\
M^{\infty}(\bar{u}) &:= \bigcup_{\bar{x} \in X(\bar{u})} \{y \in N_C(F(\bar{x}) + \bar{u}) \,|\, 0 \in \partial^{\infty} f(\bar{x}) + y\partial F(\bar{x}) + N_D(\bar{x})\},
\end{aligned}
\tag{3.30}
$$

since the latter sets are then larger, if anything. We will prove in corollary 8.1.2 that (3.29) does hold if

$$
\nexists \text{ nonzero } z \in \partial^{\infty} f(\bar{x}) \quad \text{with} \quad -z \in N_D(\bar{x}). \tag{3.31}
$$

4. ALTERNATIVE RESULT IN LIPSCHITZIAN CASE

A variant of theorem 3.2 can be proved in the case where the function f is locally Lipschitzian. It employs a device that F. H. Clarke has made much of, and it leads, as we shall see in the next section, to an alternative multiplier rule with a stability property that can be helpful. Here we set

$$
d_D(x) = \text{dist}(x, D) = \min_{x' \in D} |x - x'|. \tag{4.1}
$$

THEOREM 4.1. Suppose in theorem 3.2 that f too is locally Lipschitzian. Let λ_f, λ_F, and $\eta > 0$, be numbers such that

$$\lambda_f > \operatorname{Lip}_f(\bar{x}) \quad \text{and} \quad \lambda_F - \eta > \operatorname{Lip}_F(\bar{x}) \quad \text{for all} \quad \bar{x} \in Z(\bar{u}), \tag{4.2}$$

and let Δ_C and Δ_D be any lower semicontinuous functions such that $d_c \leqq \Delta_C \leqq \delta_C$ and $d_D \leqq \Delta_D \leqq \delta_D$. Then the assertions of proposition 2.1 are all valid also for

$$\begin{aligned} M(\bar{u}) :&= \bigcup_{\bar{x} \in X(\bar{u})} \{y \in \eta^{-1}[\lambda_f + |y|\lambda_F]\partial\Delta_C(F(\bar{x}) + \bar{u}) \,| \\ &\qquad 0 \in \partial f(\bar{x}) + y\partial F(\bar{x}) + [\lambda_f + |y|\lambda_F]\partial\Delta_D(\bar{x})\} \\ M^\infty(\bar{u}) :&= \bigcup_{\bar{x} \in X(\bar{u})} \{y \in \eta^{-1}|y|\lambda_F\partial\Delta_C(F(\bar{x}) + \bar{u}) \,|\, 0 \in y\partial F(\bar{x}) + |y|\lambda_F\partial\Delta_D(\bar{x})\}. \end{aligned} \tag{4.3}$$

Note that the conclusions of theorem 4.1 reduce to those of theorem 3.2 when $\Delta_C = \delta_C$ and $\Delta_D = \delta_D$; the sets (4.3) are then the same as the ones in (3.15), since $\partial\Delta_C(F(\bar{x}) + \bar{u})$ and $\partial\Delta_D(\bar{x})$ become the cones $N_C(F(\bar{x}) + \bar{u})$ and $N_D(\bar{x})$, which are closed with respect to multiplication by positive scalars. In other cases, however, an extra feature is obtained. For instance when Δ_C and Δ_D are locally Lipschitzian (as when $\Delta_C = d_C$ and $\Delta_D = d_D$), one has multifunctions $\partial\Delta_C$ and $\partial\Delta_D$ that are of closed graph, a property not universally possessed by the multifunctions N_C and N_D when C and D are not convex. (See [8] for the closed graph property of N_C, N_D, in the convex case, and [10, p. 150] for a counter example in the nonconvex case.)

The following fact will be needed in the proof of theorem 4.1.

LEMMA (Clarke [3].) Suppose $\bar{v}$ gives a local minimum of $g(v)$ relative to $v \in E$, where $E \subset R^n$ is closed and g is Lipschitz continuous on a neighborhood of $\bar{v}$. Let $\lambda > \operatorname{Lip}_g(\bar{v})$. Then $\bar{v}$ gives a local unconstrained minimum of $g + \lambda d_E$.

Proof. For $\varepsilon > 0$ sufficiently small, λ is a Lipschitz constant for g relative to the ball $\bar{v} + 2\varepsilon B$ (B = closed unit ball), and at the same time $\bar{v}$ minimizes g relative to $E \cap (\bar{v} + 2\varepsilon B)$. Consider any v in $\bar{v} + \varepsilon B$ and let v' be a point of E nearest to v. Then $|\bar{v} - v'| = d_E(v) \leqq |v - \bar{v}| \leqq \varepsilon$, so v lies in the ball $\bar{v} + 2\varepsilon B$. It follows that $g(\bar{v}) \leqq g(v') \leqq g(v) + \lambda|v - v'|$, so $(g + \lambda d_E)(\bar{v}) \leqq (g + \lambda d_E)(v)$. ■

Proof of theorem 4.1. We repeat the argument of theorem 3.2 word for word until the point of observing that the function f_k in (3.21) has a local minimum at (w^k, x^k). Here we make a change of variables, substituting $\bar{w} + \varepsilon z$ for w and $\bar{w} + \varepsilon z^k$ for w^k, where $z^k \to 0$. The function

$$\varphi_k(z, x) = f(x) - y^k[\varepsilon z - F(x)] + r_k|\varepsilon z - F(x) - \varepsilon z^k + F(x^k)|^2 \tag{4.4}$$

then has a local minimum at (z^k, x^k) relative to the set $A_\varepsilon \times D$, where

$$A_\varepsilon = \{z \,|\, \bar{w} + \varepsilon z \in C\} = \varepsilon^{-1}[C - \bar{w}], \tag{4.5}$$

$$d_{A_\varepsilon}(z) = \varepsilon^{-1}d_C(\bar{w} + \varepsilon z) \leqq \varepsilon^{-1}\Delta_C(\bar{w} + \varepsilon z) =: \Delta_{A_\varepsilon}(z). \tag{4.6}$$

Obviously φ_k is locally Lipschitzian with

$$\operatorname{Lip}_{\varphi_k}(z^k, x^k) \leqq \operatorname{Lip}_f(x^k) + |y^k|[\varepsilon + \operatorname{Lip}_F(x^k)]. \tag{4.7}$$

(The last term in the formula for φ_k has Lipschitz modulus 0 at (z^k, x^k)).

To prove that the set $M(\bar{u})$ in (4.3) includes $\partial\hat{p}(u)$, we need only demonstrate now that if $y^k \to y$ we have the conditions which define $M(\bar{u})$ satisfied by y in terms of the point $\bar{x} = \lim x^k$. From $|y^k| \to |y|$ we get

$$\mathrm{Lip}_{\varphi_k}(z^k, x^k) < \lambda_f + |y|\lambda_f \quad \text{for each } k$$

in (4.7) by virtue of (4.2) and (4.4). The lemma above then asserts that the locally Lipschitzian function

$$\varphi_k + (\lambda_f + |y|\lambda_F)d_{A_\varepsilon \times D}$$

on $R^m \times R^n$ has a local minimum at (z^k, x^k). But

$$\begin{aligned} d_{A_\varepsilon \times D}(z, x) &= [d_{A_\varepsilon}(z)^2 + d_D(x)^2]^{1/2} \\ &\leqq d_{A_\varepsilon}(z) + d_D(x) \leqq \Delta_{A_\varepsilon}(z) + \Delta_D(x). \end{aligned}$$

with Δ_{A_ε} defined as in (4.6). Therefore the function

$$\psi_k(z, x) = \varphi_k(z, x) + (\lambda_f + |y|\lambda_F)[\Delta_{A_\varepsilon}(z) + \Delta_D(x)]$$

likewise has a local minimum at (z^k, x^k), so that

$$(0, 0) \in \partial\psi_k(z^k, x^k).$$

Write $\psi_k = g + h_k$, where (referring to (4.5))

$$g(z, x) = (\lambda_f + |y|\lambda_F)\Delta_{A_\varepsilon}(z) - y \cdot \varepsilon z + f(x) + y \cdot F(x) + [(\lambda_f + |y|\lambda_F)\Delta_D](x), \tag{4.8}$$

$$h_k(z, x) = (y - y^k) \cdot [\varepsilon z - F(x)] + r_k|\varepsilon z - F(x) - \varepsilon z^k + F(x^k)|^2. \tag{4.9}$$

Calculating with proposition 2.4, we derive

$$(0, 0) \in \partial g(z^k, x^k) + \partial h_k(z^k, x^k).$$

In other words,

$$\exists (s^k, v^k) \in \partial g(z^k, x^k) \qquad \text{with} \quad -(s^k, v^k) \in \partial h_k(z^k, x^k).$$

But then

$$|(s^k, v^k)| \leqq \mathrm{Lip}_{h_k}(z^k, x^k) \leqq |y - y^k|(\varepsilon + \mathrm{Lip}_F(x^k)) \to 0,$$

so $(s^k, v^k) \to (0, 0)$. Since $(z^k, x^k) \to (0, \bar{x})$ and g is Lipschitz continuous around $(0, \bar{x})$, it follows from $(s^k, v^k) \in \partial g(z^k, x^k)$ that $(0, 0) \in \partial g(0, \bar{x})$ (cf. proposition 2.2). We next apply proposition 2.5 (and corollary 2.4.1) to translate the latter condition into

$$0 \in (\lambda_f + |y|\lambda_F)\partial\Delta_{A_\varepsilon}(0) - \varepsilon y, \tag{4.10}$$

$$0 \in \partial[f + y \cdot F + (\lambda_f + |y|\lambda_F)\Delta_D](\bar{x}). \tag{4.11}$$

Observe now that the definition (4.6) of Δ_{A_ε} gives

$$\partial\Delta_{A_\varepsilon}(0) = \varepsilon^{-1}[\varepsilon\partial\Delta_C(\bar{w})] = \partial\Delta_C(\bar{w}). \tag{4.12}$$

Condition (4.10) thus yields

$$y \in \varepsilon^{-1}(\lambda_f + |y|\lambda_F)\partial\Delta_C(\bar{w}). \tag{4.13}$$

At the same time, (4.11) implies via propositions 2.4 and 2.6 that

$$0 \in \partial f(\bar{x}) + y\partial F(\bar{x}) + [\lambda_f + |y|\lambda_F]\partial\Delta_D(\bar{x}). \tag{4.14}$$

In summary, if $y^k \to y$ then (4.13) and (4.14) hold, so $y \in M(\bar{u})$. This establishes that $\hat{\partial}p(\bar{u}) \subset M(\bar{u})$.

The argument demonstrating $\hat{\partial}^\infty p(\bar{u}) \subset M^\infty(\bar{u})$ differs only slightly from the one just given. We look at the case where $t_k y^k \to y$ with $t_k \downarrow 0$. The function $t_k\varphi_k$ has a local minimum at (z^k, x^k) relative to $A_\varepsilon \times D$, and

$$\mathrm{Lip}_{t_k\varphi_k}(z^k, x^k) \leqq t_k \,\mathrm{Lip}_f(x^k) + |t_k y^k|(\varepsilon + \mathrm{Lip}_F(x^k)),$$

which implies

$$\mathrm{Lip}_{t_k\varphi_k}(z^k, x^k) < |y|\lambda_F \quad \text{for large } k.$$

Then by the lemma we have a local minimum of

$$t_k\varphi_k + |y|\lambda_F d_{A_\varepsilon \times D}$$

at (z^k, x^k), hence also a local minimum of

$$\psi_k(z, x) = \varphi_k(z, x) + |y|\lambda_F[\Delta_{A_\varepsilon}(z) + \Delta_D(x)]$$

at (z^k, x^k), where $\psi_k = g + h_k$ with

$$g(z, x) = [|y|\lambda_F\Delta_{A_\varepsilon}(z) - \varepsilon y \cdot z] + [y \cdot F + |y|\lambda_F\Delta_D](\bar{x}),$$

$$h_k(z, x) = (y - t_k y^k) \cdot [\varepsilon z - F(z)] + t_k r_k |\varepsilon z - F(x) - \varepsilon z^k + F(x^k)|^2.$$

As before we get

$$(0, 0) \in \partial\psi_k(z^k. x^k) \subset \partial g(z^k, x^k) + \partial h_k(z^k, x^k)$$

and deduce that $(0, 0) \in \partial g(0, \bar{x})$, from which the desired conditions

$$y \in \varepsilon^{-1}\lambda_F|y|\partial\Delta_C(F(\bar{x}) + \bar{u}) \qquad \text{and} \quad 0 \in y\partial F(\bar{x}) + |y|\lambda_F\partial\Delta_D(\bar{x})$$

follow and show that $y \in M^\infty(\bar{u})$. ■

Remark. The proof of theorem 4.1 shows that the conclusions would still hold if the conditions defining $M(\bar{u})$ and $M^\infty(\bar{u})$ were strengthened to

$$0 \in \partial[f + y \cdot F + (\lambda_f + |y|\lambda_F)\Delta_D](\bar{x}) \quad \text{and} \quad 0 \in \partial[y \cdot F + |y|\lambda_F\Delta_D](\bar{x})$$

respectively.

5. LAGRANGE MULTIPLIER RULES

The vectors y appearing in theorems 3.1, 3.2, and 4.1 can be regarded as generalized Lagrange multiplier vectors associated with necessary conditions for optimality. Such necessary conditions in fact are consequences of the theorems mentioned, as we now demonstrate. We first present a rule for a very general optimization problem depending on a parameter vector u.

THEOREM 5.1. Let $\bar{x}$ be a locally optimal solution to the problem

$$\text{minimize } g(x, \bar{u}) \text{ over } x \in R^n, \tag{5.1}$$

where $g: R^n \times R^m \to \bar{R}$ is lower semicontinuous and $g(\bar{x}, \bar{u})$ is finite. Either suppose that the problem is calm at $\bar{x}$ with respect to perturbations of $\bar{u}$, in the sense that

$$\nexists (x^k, u^k) \to (\bar{x}, \bar{u}) \quad \text{with} \quad u_k \neq \bar{u} \quad \text{and} \quad \frac{g(x^k, u^k) - g(\bar{x}, \bar{u})}{|u^k - \bar{u}|} \to -\infty, \tag{5.2}$$

or suppose that $\bar{x}$ satisfies the constraint qualification

$$\nexists \text{ nonzero } y \text{ with } (0, y) \in \partial^\infty g(\bar{x}, \bar{u}). \tag{5.3}$$

Then

$$\exists\, y \text{ with } (0, y) \in \partial g(\bar{x}, \bar{u}). \tag{5.4}$$

Proof. Taking any $\mu > 0$ small enough that $\bar{x}$ gives the minimum in (5.1) relative to the ball $\{x \mid |x - \bar{x}| \leqq \mu\}$, we define

$$P(x, u) = \begin{cases} g(x, u) + |x - \bar{x}|^2 & \text{if} \quad |x - \bar{x}| \leqq \mu \\ +\infty \text{ otherwise} \end{cases} \tag{5.5}$$

and observe that the hypothesis of theorem 3.1 is satisfied, moreover with $X(\bar{u}) = \{\bar{x}\}$. Then corollary 3.1.1 furnishes the inclusions (3.8), where $p(u) = \inf_x P(x, u)$. By corollary 2.4.2:

$$\partial P(\bar{x}, \bar{u}) = \partial g(\bar{x}, \bar{u}), \qquad \partial^\infty P(\bar{x}, \bar{u}) = \partial^\infty g(\bar{x}, \bar{u}). \tag{5.6}$$

If the constraint qualification (5.3) holds, we obtain $\partial^\infty p(\bar{u}) = \{0\}$ by (3.8). Then $\partial p(\bar{u}) \neq \emptyset$, so (5.4) must hold by (3.8).

We must show this conclusion is also valid if in place of (5.3) we assume (5.2). The latter is equivalent to the assumption that for small enough μ in definition (5.5), one will have a lower bound

$$p(u) \geqq p(\bar{u}) - \rho|u - \bar{u}| \qquad \text{when} \quad |u - \bar{u}| \leqq \varepsilon \tag{5.7}$$

for certain $\rho > 0$, $\varepsilon > 0$. Then the function

$$r(u) = p(u) + q(u), \qquad \text{where} \quad q(u) = \rho|u - \bar{u}|,$$

has a local minimum at $\bar{u}$. Since p is strictly lower semicontinuous at $\bar{u}$ and q is Lipschitzian, it follows from proposition 2.4 that

$$0 \in \partial r(\bar{u}) \subset \partial p(\bar{u}) + \partial q(\bar{u}).$$

In particular $\partial p(\bar{u}) = \emptyset$, and from (5.6) and (3.8) we can again conclude (5.4). ■

We turn now to the case of problems with explicit constraints.

THEOREM 5.2. Let $\bar{x}$ be a locally optimal solution to the problem

$$\text{minimize } f(x) \text{ subject to } F(x) + \bar{u} \in C, x \in D,$$

where $f: R^n \to \bar{R}$ is lower semicontinuous with $f(\bar{x})$ finite, $F: R^n \to R^m$ is locally Lipschitzian, and $C \subset R^m$ and $D \subset R^n$ are closed. Suppose either that the problem is calm at $\bar{u}$, in the sense that

$$\nexists (x^k, u^k) \to (\bar{x}, \bar{u}) \text{ with } F(x^k) + u^k \in C, x^k \in D, \text{ such that}$$
$$u^k \neq \bar{u} \text{ and } [f(x^k) - f(\bar{x})]/|u^k - \bar{u}| \to -\infty, \tag{5.8}$$

or that $\bar{x}$ satisfies the constraint qualification

$$\nexists \text{ nonzero } y \in N_C(F(\bar{x}) + \bar{u}) \text{ with } 0 \in \partial^\infty(f + \delta_D)(x) + y\partial F(\bar{x}). \tag{5.9}$$

Then

$$\exists y \in N_C(F(\bar{x}) + u) \text{ with } 0 \in \partial(f + \delta_D)(\bar{x}) + y\partial F(\bar{x}). \tag{5.10}$$

Remark 5.1. The subgradient conditions in (5.9) and (5.10) can be replaced respectively by

$$0 \in \partial^\infty(f + yF + \delta_D)(\bar{x}), \qquad 0 \in \partial(f + yF + \delta_D)(\bar{x}), \tag{5.11}$$

to obtain a slightly sharper result. This is true because the proof given below works for any sets $M(\bar{u})$ and $M^\infty(\bar{u})$ lending validity to theorem 3.2. See remark 3.3.

Remark 5.2. If f happens to be locally Lipschitzian, the subgradient conditions in (5.9) and (5.10) can for the same reason be replaced respectively by

$$0 \in y\partial F(\bar{x}) + N_D(\bar{x}), \qquad 0 \in \partial f(\bar{x}) + y\partial F(\bar{x}) + N_D(\bar{x}). \tag{5.12}$$

See corollary 3.2.1.

Remark 5.3. More generally the subgradient conditions in (5.9) and (5.10) can be replaced by

$$0 \in \partial^\infty f(\bar{x}) + y\partial F(\bar{x}) + N_D(\bar{x}), \qquad 0 \in \partial f(\bar{x}) + y\partial F(\bar{x}) + N_D(\bar{x}), \tag{5.13}$$

whenever f is such that

$$\partial^\infty(f + \delta_D)(\bar{x}) \subset \partial^\infty f(\bar{x}) + N_D(\bar{x}), \qquad \partial(f + \delta_D)(\bar{x}) \subset \partial f(\bar{x}) + N_D(\bar{x}).$$

Results in Section 8 will show that this is correct when $\bar{x}$ satisfies the further constraint qualification that

$$\nexists \text{ nonzero } z \in \partial^\infty f(\bar{x}) \text{ with } -z \in N_D(\bar{x}) \tag{5.14}$$

(see corollary 8.1.2).

Proof of theorem 5.2. The argument parallels the proof of the preceding theorem, but this time we instead apply theorem 3.2 to

$$p(u) = \inf_x\{\bar{f}(x) \mid F(x) + u \in C, x \in \bar{D}\},$$

where $\bar{f}(x) = f(x) + |x - \bar{x}|^2$, $\bar{D} = \{x \in D \,|\, |x - \bar{x}| \leqq \mu\}$. For small enough μ the set $X(\bar{u})$ consists of just $\bar{x}$, so

$$\partial p(\bar{u}) \subset \{y \in N_C(F(\bar{x}) + \bar{u}) \,|\, 0 \in \partial(\bar{f} + \delta_{\bar{D}})(x) + y\partial F(\bar{x})\},$$

$$\partial^\infty p(\bar{u}) \subset \{y \in N_C(F(\bar{x}) + \bar{u}) \,|\, 0 \in \partial^\infty(\bar{f} + \delta_{\bar{D}})(\bar{x}) + y\partial F(\bar{x})\},$$

where by corollary 2.4.2: $\partial(\bar{f} + \delta_D)(\bar{x}) = \partial(f + \delta_D)(\bar{x})$ and $\partial^\infty(\bar{f} + \delta_D)(\bar{x}) = \partial^\infty(f + \delta_D)(\bar{x})$.

Thus the desired conclusion (5.10) will follow if $\partial p(\bar{u}) \neq \emptyset$. Furthermore, the constraint qualification (5.9) implies $\partial^\infty p(\bar{u}) = \{0\}$, which does ensure $\partial p(\bar{u}) \neq \emptyset$ (cf. the basic facts in Section 1). The calmness condition (5.8), on the other hand, is equivalent to a lower bound of type (5.7) holding when μ is sufficiently small, and this too ensures $\partial p(\bar{u}) \neq 0$ as verified in the preceding proof. ■

Remark 5.4. The case of theorem 5.2 that can be compared most easily with Lagrange multiplier results already in the literature is the one where f is itself locally Lipschitzian (cf. remark 5.2) and the condition $F(x) + \bar{u} \in C$ represents a standard constraint system as described in remark 3.2. This case was first treated by Clarke [5] in terms of generalized subgradients but using a different technique that relied on Ekeland's variational principle [6] and did not provide an interpretation of the multipliers in terms of the differential effect of certain perturbations, such as we have here by virtue of theorem 3.2. In Clarke's multiplier rule the set $y_1\partial f_1(\bar{x}) + \ldots + y_m\partial f_m(\bar{x})$ appears in place of the smaller set $y\partial F(\bar{x})$, where $y = (y_1, \ldots, y_m)$ and $F = (f_1, \ldots, f_m)$, but on the other hand a multiple of $\partial d_D(\bar{x})$ appears in place of the larger set $N_D(\bar{x})$. Hiriart-Urruty [7] showed how to consolidate the expression $\partial f(\bar{x}) + y_1\partial f_1(\bar{x}) + \ldots + y_m\partial f_m(\bar{x})$ to $\partial(f + yF)(\bar{x})$. Rockafellar [11] showed further that the rule could be written in terms of $\partial(f + yF + \delta_D)$ (cf. the extension indicated in remark 5.1) and be validated under a weaker calmness assumption than Clarke's, namely (5.8).

COROLLARY 5.2.1. Let $\bar{x}$ be a locally optimal solution to the problem

$$\text{minimize } f(x) \text{ subject } x \in C,$$

where $f: R^n \to \bar{R}$ is lower semicontinuous with $f(\bar{x})$ finite, and $C \subset R^n$ is closed. Suppose either that the problem is calm at $\bar{x}$ in the sense that

$$\nexists x^k \to \bar{x} \text{ with } x^k \notin C, [f(x^k) - f(x)]/d_C(x^k) \to -\infty,$$

or that $\bar{x}$ satisfies the constraint qualification

$$\nexists \text{ nonzero } y \in N_C(\bar{x}) \text{ with } -y \in \partial^\infty f(\bar{x}).$$

Then

$$\exists y \in N_C(\bar{x}) \text{ with } -y \in \partial f(\bar{x}).$$

Proof. Specialize theorem 5.2 to $D = R^n = R^m$, $F =$ identity, $\bar{u} = 0$. ■

COROLLARY 5.2.2. Let $\bar{x}$ be a locally optimal solution to the problem

$$\text{minimize } f(x) \text{ subject to } F(x) = 0,$$

where $F: R^n \to R^m$ is locally Lipschitzian. Suppose either that the problem is calm at $\bar{x}$ in the sense that

$$\nexists x^h \to \bar{x} \text{ with } F(\bar{x}^k) \neq 0 \text{ and } [f(x^k) - f(\bar{x})]/|F(x^k)| \to -\infty,$$

or that $\bar{x}$ satisfies the constraint qualification

$$\nexists \text{ nonzero } y \in R^m \text{ with } 0 \in \partial^\infty f(\bar{x}) + y\partial F(\bar{x}).$$

Then

$$\exists y \in R^m \text{ with } 0 \in \partial f(\bar{x}) + y\partial F(\bar{x}).$$

Proof. Specialize theorem 5.2 to $D = R^n$, $C = \{0\}$, $\bar{u} = 0$. ■

COROLLARY 5.2.3. Let $\bar{x}$ be a locally optimal solution to the problem

$$\text{minimize } f(x) + g(F(x)) \text{ over all } x \in R^n,$$

where $f: R^n \to \bar{R}$ and $g: R^m \to \bar{R}$ are lower semicontinuous with $f(\bar{x})$ and $g(F(\bar{x}))$ finite, and $F: R^n \to R^m$ is locally Lipschitzian. Suppose either that the problem is calm at $\bar{x}$ in the sense that

$$\nexists (x^k, u^k) \to (\bar{x}, 0) \text{ in } R^n \times R^m \text{ with } u^k \neq 0 \quad \text{and}$$
$$[f(x^k) + g(F(x^k) + u^k) - f(\bar{x}) - g(F(\bar{x}))]/|u^k| \to -\infty,$$

or that $\bar{x}$ satisfies the constraint qualification

$$\nexists \text{ nonzero } y \in \partial^\infty g(F(\bar{x})) \text{ with } 0 \in \partial^\infty f(\bar{x}) + y\partial F(\bar{x}).$$

Then

$$\exists y \in \partial g(F(\bar{x})) \text{ with } 0 \in \partial f(\bar{x}) + y\partial F(\bar{x}).$$

Proof. Reformulate the problem as

$$\text{minimize } \bar{f}(x, w) \text{ subject to } \bar{F}(x, w) = 0,$$

where $\bar{f}(x, w) = f(x) + g(w)$, $\bar{F}(x, w) = F(x) - w$. This has a local minimum at $(\bar{x}, \bar{w})$, where $\bar{w} = F(\bar{x})$. Apply the preceding corollary and invoke proposition 2.5 to calculate $\partial\bar{f}(\bar{x}, \bar{w})$ and $\partial^\infty\bar{f}(\bar{x}, \bar{w})$. ■

The next theorem furnishes an alternative to theorem 5.2 that allows for a different treatment of the constraint $x \in D$ along the lines followed by Clarke [12], as mentioned in remark 5.4.

THEOREM 5.3. Let $\bar{x}$ be a locally optimal solution to the problem

$$\text{minimize } f(x) \text{ subject to } F(x) + \bar{u} \in C, x \in D,$$

where $f: R^n$ and $F: R^n \to R^m$ are locally Lipschitzian, and $C \subset R^m$ and $D \subset R^n$ are closed. Let λ_f, λ_F and $\varepsilon > 0$ be any numbers such that

$$\lambda_f > \mathrm{Lip}_f(\bar{x}) \text{ and } \lambda_F - \varepsilon > \mathrm{Lip}_F(\bar{x}), \tag{5.15}$$

and let Δ_C and Δ_D be any lower semicontinuous functions such that $d_C \leqq \Delta_C \leqq \delta_C$ and $d_D \leqq \Delta_D \leqq \delta_D$. Then

$$\exists(\eta, y) \neq (0,0) \text{ with } \eta \geqq 0, y \in \varepsilon^{-1}[\eta\lambda_f + |y|\lambda_F]\partial\Delta_C(F(\bar{x}) + \bar{u}),$$
$$\text{and } 0 \in \eta\partial f(\bar{x}) + y\partial F(\bar{x}) + [\eta\lambda_f + |y|\lambda_F]\partial\Delta_D(\bar{x}). \tag{5.16}$$

If the calmness condition (5.8) is fulfilled, then one can take $\eta > 0$ (hence $\eta = 1$).

Proof. The pattern of reasoning is identical to the proof of theorem 5.2, except that theorem 4.1 is invoked rather than theorem 3.2. The conclusion is stated slightly differently, however: (5.16) is equivalent to the assertion that if

$$\nexists \text{ nonzero } y \in \varepsilon^{-1}|y|\lambda_F\partial\Delta_C(F(\bar{x}) + \bar{u})$$
$$\text{with } 0 \in y\partial F(\bar{x}) + |y|\lambda_F\partial\Delta_D(\bar{x}),$$

then

$$\exists y \in \varepsilon^{-1}[\lambda_f + |y|\lambda_F]\partial\Delta_C(F(\bar{x}) + \bar{u})$$
$$\text{with } 0 \in \partial f(\bar{x}) + y\partial f(\bar{x}) + [\lambda_f + |y|\lambda_F]\partial\Delta_D(\bar{x}).$$

The case of the possible constraint qualification (5.9) in theorem 5.2 is thus incorporated in this version in another form. ■

Theorem 5.3 turns into the multiplier rule of Clarke [3, theorem 6.1.1] if we choose $\Delta_C = \delta_C$, $\Delta_D = d_D$, and specialize C to the case of remark 3.2, so that the condition $F(x) + \bar{u} \in C$ represents a mixed system of equality and inequality constraints of the usual sort. Clarke's result, however, is also valid for x belonging to a Banach space, not just R^n.

6. CONSEQUENCES OF THE PERTURBATION THEOREMS

We turn now to the application of the preceding results to the development of further rules for subdifferentiation. We begin with formulas that can be derived from the perturbation theorems alone.

THEOREM 6.1. Consider

$$p(u) = \inf_x\{f(x) \mid G(x) = u\},$$
$$X(u) = \operatorname{argmin}_x\{f(x) \mid G(x) = u\}, \tag{6.1}$$

where $f: R^n \to \bar{R}$ is lower semicontinuous and $G: R^n \to R^m$ is locally Lipschitzian. Let $\bar{u}$ be a point where p is finite and the following holds:

$$\exists \varepsilon > 0 \text{ and } \bar{\alpha} > p(\bar{u}) \text{ such that the set}$$
$$\{x \mid f(x) \leqq \bar{\alpha}, |G(x) - \bar{u}| \leqq \varepsilon\} \text{ is bounded.} \tag{6.2}$$

Then p is strictly lower semicontinuous at $\bar{u}$ and $X(\bar{u})$ is nonempty and compact. Moreover for

$$M(\bar{u}) := \{y \mid \exists \bar{x} \in X(\bar{u}) \quad \text{with} \quad \partial f(\bar{x}) \cap y\partial G(\bar{x}) \neq \emptyset\},$$
$$M^{\infty}(\bar{u}) := \{y \mid \exists \bar{x} \in X(\bar{u}) \quad \text{with} \quad \partial^{\infty} f(\bar{x}) \cap y\partial G(\bar{x}) \neq \emptyset\}, \tag{6.3}$$

one has $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$ and $\hat{\partial}^{\infty} p(\bar{u}) \subset M^{\infty}(\bar{u})$, so all the assertions in proposition 2.1 are valid.

Proof. Simply apply theorem 3.2 to $F = -G$, $C = \{0\}$, $D = R^n$ (i.e. $\delta_D \equiv 0$), and the given f. ∎

COROLLARY 6.1.1. Let $E = G(D)$, where $D \subset R^n$ is closed and $G: R^n \to R^m$ is locally Lipschitzian. Let $\bar{u} \in E$ be a point where the following holds:

$$\exists \varepsilon > 0 \text{ such that the set } \{x \in D \mid |G(x) - \bar{u}| \leqq \varepsilon\} \text{ is bounded.} \tag{6.4}$$

Then E is closed relative to some neighborhood of $\bar{u}$, and for the cone

$$M(\bar{u}) := \{y \mid \exists \bar{x} \in G^{-1}(\bar{u}) \cap D \quad \text{with} \quad N_D(\bar{x}) \cap y\partial G(\bar{x}) \neq \emptyset\} \tag{6.5}$$

all the assertions in corollary 2.1.1 are valid.

Proof. Apply theorem 6.1 to $f = \delta_D$, which gives $p = \delta_E$. ∎

Our next result concerns the operation of infimal convolution of extended-real-valued functions.

THEOREM 6.2. Consider

$$p(u) = \inf_{x_1,\ldots,x_r} \{f_1(x_1) + \ldots + f_r(x_r) \mid x_1 + \ldots + x_r = u\},$$
$$X(u) = \operatorname*{argmin}_{x_1,\ldots,x_r} \{f_1(x_1) + \ldots + f_r(x_r) \mid x_1 + \ldots + x_r = u\}, \tag{6.6}$$

where $f_i: R^m \to \bar{R}$ is lower semicontinuous. (The convention $\infty - \infty = \infty$ is used to handle the extended arithmetic in these formulas, when required.) Let $\bar{u}$ be a point where $p(\bar{u})$ is finite, and suppose

$$\exists \varepsilon > 0 \quad \text{and} \quad \bar{\alpha} > p(\bar{u}) \text{ such that the set}$$
$$\{(x_1, \ldots, x_r) \mid f_1(x_1) + \ldots + f_r(x_r) \leqq \bar{\alpha}, |x_1 + \ldots + x_r - u| \leqq \varepsilon\} \text{ is bounded.} \tag{6.7}$$

Then p is strictly lower continuous at $\bar{u}$ and $X(\bar{u})$ is nonempty and compact. Moreover for

$$M(\bar{u}) := \bigcup_{(\bar{x}_1,\ldots,\bar{x}_r) \in X(\bar{u})} [\partial f_1(\bar{x}_1) \cap \ldots \cap \partial f_r(\bar{x}_r)]$$
$$M^{\infty}(\bar{u}) := \bigcup_{(\bar{x}_1,\ldots,\bar{x}_r) \in X(\bar{u})} [\partial^{\infty} f_1(\bar{x}_1) \cap \ldots \cap \partial^{\infty} f_r(\bar{x}_r)] \tag{6.8}$$

one has $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$ and $\hat{\partial}^{\infty} p(\bar{u}) \subset M^{\infty}(\bar{u})$, so all the assertions of proposition 2.1 are valid.

Proof. For $x = (x_1, \ldots, x_r) \in (R^m)^r$ define $f(x) = f_1(x_1) + \ldots + f_r(x_r)$ and $G(x) = x_1 + \ldots + x_r$. The situation is thereby reduced to that of theorem 6.1. Moreover

$$\partial f(x) = (\partial f_1(x_1), \ldots, \partial f_r(x_r)), \qquad \partial^\infty f(x) = (\partial^\infty f_1(x_1), \ldots, \partial^\infty f_r(x_r)),$$

by proposition 2.5, and $y\partial G(\bar{x})$ consists of just the vector $(y, \ldots, y) \in (R^m)^r$. The sets (6.3) are therefore expressed by (6.8) in this case, and the conclusions of theorem 6.1 give us what we want. ■

COROLLARY 6.2.1. Let $E = C_1 + \ldots + C_r$, where $C_i \subset R^m$ is closed. Let $\bar{u} \in E$ be such that

$$\exists \varepsilon > 0 \text{ such that the set}$$
$$\{(x_1, \ldots, x_r) \mid x_i \in C_i, |x_1 + \ldots + x_r - \bar{u}| \leqq \varepsilon\} \text{ is bounded.} \tag{6.9}$$

Then C is closed relative to some neighborhood of $\bar{u}$, and for

$$M(\bar{u}) := \bigcup_{\substack{\bar{x}_1 + \ldots + \bar{x}_r = \bar{u} \\ \bar{x}_i \in C_i}} [N_{C_1}(\bar{x}_1) \cap \ldots \cap N_{C_r}(\bar{x}_r)] \tag{6.10}$$

one has $\hat{N}_E(\bar{u}) \subset M(\bar{u})$, so that all the assertions of corollary 2.1.1 are valid.

Proof. Take $f_i = \delta_{C_i}$ in the theorem. ■

COROLLARY 6.2.2. Consider

$$p(u) = \inf_{x \in C} g(u - x), \qquad X(u) = \operatorname*{argmin}_{x \in C} g(u - x), \tag{6.11}$$

where $C \subset R^m$ is closed and $g: R^m \to \bar{R}$ is lower semicontinuous. Let $\bar{u}$ be a point where p is finite and

$$\exists \alpha > p(u) \quad \text{and} \quad \varepsilon > 0 \text{ such that the set}$$
$$\{(x, u) \mid x \in C, |u - \bar{u}| \leqq \varepsilon, g(u - x) \leqq \bar{\alpha}\} \text{ is bounded.} \tag{6.12}$$

Then g is strictly lower semicontinuous at $\bar{u}$ and $X(\bar{u})$, is nonempty and compact. Moreover for

$$\begin{aligned} M(\bar{u}) &:= \bigcup_{\bar{x} \in X(\bar{u})} N_C(\bar{x}) \cap \partial g(\bar{u} - \bar{x}), \\ M^\infty(\bar{u}) &:= \bigcup_{\bar{u} \in X(\bar{u})} N_C(\bar{x}) \cap \partial^\infty g(\bar{u} - \bar{x}), \end{aligned} \tag{6.13}$$

one has $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$ and $\hat{\partial}^\infty p(\bar{u}) \subset M^\infty(\bar{u})$, so that all the assertions of proposition 2.1 are valid.

Proof. In theorem 6.2 take $r = 2, f_1 = \delta_C, f_2 = g$. In place of x_1 and x_2 write x and $u - x$. ■

Remark 6.2. Choices of C that are of particular interest in corollary 6.2.2 for general g are $C = \varepsilon B$ (closed ball of radius ε around 0) or $C = R^N_+$, which yield

$$p(u) = \inf_{|v - u| \leqq \varepsilon} g(v) \quad \text{or} \quad p(u) = \inf_{v \leqq u} g(v).$$

Noteworthy choices of g for general C are $g(v) = |v|$ and $g(v) = \frac{1}{2}|v|^2$. These yield $p = d_C$ and $p = \frac{1}{2}d_C^2$, respectively.

THEOREM 6.3. Let $p(u) = \min\{f_1(u), \ldots, f_r(u)\}$, where $f_i: R^n \to \bar{R}$ is lower semicontinuous. Let $\bar{u}$ be a point where $p(\bar{u})$ is finite, and let $I(\bar{u}) = \{i \mid f_i(\bar{u}) = f(\bar{u})\}$. Then p is lower semicontinuous, and for

$$M(\bar{u}): = \bigcup_{i \in I(\bar{u})} \partial f_i(\bar{u}), \qquad M^\infty(\bar{u}): = \bigcup_{i \in I(\bar{u})} \partial^\infty f_i(\bar{u}), \tag{6.14}$$

one has $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$ and $\hat{\partial}^\infty p(\bar{u}) \subset M^\infty(\bar{u})$, so that all the conclusions of proposition 2.1 are valid.

Proof. Define $P(x, u) = f_i(u)$ if $x = i \in \{i, \ldots, m\} \subset R$, $P(x, u) = \infty$ otherwise. The hypothesis of theorem 3.1 is then satisfied, and the conclusion of that theorem translates into the result given here. (It actually yields a slightly stronger conclusion in which $\partial f_i(\bar{u})$ and $\partial^\infty f_i(\bar{u})$ are replaced in (6.14) by the smaller sets $\hat{\partial} f_i(\bar{u})$ and $\hat{\partial}^\infty f_i(\bar{u})$.) ■

COROLLARY 6.3.1. Let $C = C_1 \cup \ldots \cup C_r$, where $C_i \subset R^n$ is closed. Let $\bar{u} \in C$ and let $I(\bar{u}) = \{i \mid \bar{u} \in C_i\}$. Then C is closed, and for

$$M(\bar{u}): = \bigcup_{i \in I(\bar{u})} N_{C_i}(\bar{u}) \tag{6.15}$$

one has $\hat{N}_C(\bar{u}) \subset M(\bar{u})$, so the conclusions of corollary 2.1.1 hold.

Proof. Let $f_i = \delta_{C_i}$ in theorem 6.3. ■

7. CHAIN RULES

The subdifferentiation formulas that we tackle next depend not only on the perturbation theorems, but at a crucial stage also on the Lipschitzian version of the Lagrange multiplier rule, namely theorem 5.3. This is because of a limiting process in the proof which requires a semicontinuity property that is available for elements satisfying the conditions in theorems 5.3, but not necessarily in the case of the conditions in theorem 5.2, at least not without additional assumptions.

THEOREM 7.1. Let $p(u) = g(G(u))$, where $G: R^m \to R^d$ is locally Lipschitzian and $g: R^d \to \bar{R}$ is lower semicontinuous. Let $\bar{u}$ be such that $p(\bar{u})$ is finite. Assume

$$\nexists \text{ nonzero } y \in \partial^\infty g(G(\bar{u})) \quad \text{with} \quad 0 \in y \partial G(\bar{u}). \tag{7.1}$$

Then for the sets

$$M(\bar{u}): = \partial g(G(\bar{u}))\partial G(\bar{u}), \qquad M^\infty(\bar{u}): = \partial^\infty g(G(\bar{u}))\partial G(\bar{u}), \tag{7.2}$$

one has $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$ and $\hat{\partial}^\infty p(\bar{u}) \subset M^\infty(\bar{u})$, so that all the assertions in proposition 2.1 are valid.

Remark 7.1. Condition (7.1) can be stated dually as follows, in terms of the convex cone K which is polar to $\partial^\infty g(G(\bar{u}))$: for no $A \in \partial G(\bar{u})$ can K be separated from the (linear) range space of A. The chain rule previously proved by the author in [13, theorem 3] requires that $\partial G(\bar{u})$ consists of a single $\bar{A}$ whose range space meets int K, a condition that is obviously more restrictive. On the other hand, this earlier chain rule is stated in a "directionally Lipschitzian" form that holds true when R^n is replaced by an infinite-dimensional space.

Proof of theorem 7.1. Write

$$p(u) = \inf\{\alpha \mid (v, \alpha) \in \operatorname{epi} g, v - G(u) = 0\}$$
$$= \inf_{v,\alpha}\{\varphi(u, v, \alpha) \mid F(u, v, \alpha) \in C, (u, v, \alpha) \in D\}, \tag{7.3}$$

where

$$\varphi(u, v, \alpha) = \alpha, \quad C = \{0\}, \quad D = R^m \times \operatorname{epi} g, \quad F(u, v, \alpha) = G(u) - v. \tag{7.4}$$

The functions φ and F are locally Lipschitzian, and the minimizing set $X(u)$ in (7.3) is trivially just the singleton $\{(u, G(u), g(G(u)))\}$. Obviously p is lower semicontinuous.

Suppose z^k is a proximal subgradient of p at u^k, where $u^k \to \bar{u}$ and $p(u^k) \to p(\bar{u})$, i.e. $g(G(u^k)) \to g(G(\bar{u}))$. We show first that if $z^k \to z$, then $z \in M(\bar{u})$ as defined in (7.2).

For some $r_k > 0$ we have a local minimum at u^k for the function

$$u \to p(u) - z^k \cdot u + r_k|u - u^k|^2,$$

and this amounts to a local minimum at $(u^k, G(u^k), g(G(u^k)))$ for the function

$$F_k(u, v, \alpha) = \alpha - z^k \cdot u + r_k|u - u^k|^2 \tag{7.5}$$

over $(u, v, \alpha) \in D$ subject to $F(u, v, \alpha) \in C$. We wish to invoke the multiplier rule in theorem 5.3 at this local minimum, and we can do so because f_k is locally Lipschitzian:

$$\operatorname{Lip}_{f_k}(u^k, G(u^k), g(G(u^k))) \leqq |z^k| + 1 \to |z| + 1.$$

It is essential to note in this that the same values λ_f and λ_F can be made to work for all k, if chosen large enough. There is no need to consider ε, because we take $\Delta_C = \delta_C = \delta_0$, so that $\partial\Delta_C(0)$ is the normal cone to $\{0\}$ at 0, i.e. the whole space R^d. On the other hand, we take $\Delta_D = d_D$.

The implication of theorem 5.3 then is that there exists $(\eta_k, y^k) \neq (0, 0)$ with $\eta_k \geqq 0$, such that

$$(0,0,0) \in \eta_k \partial f_k(u^k, G(u^k), g(G(u^k))) + y^k \partial F(u^k, G(u^k), g(G(u^k)))$$
$$+ [\eta_k\lambda_f + |y^k|\lambda_f]\partial d_D(u^k, G(u^k), g(G(u^k))). \tag{7.6}$$

We can normalize to have $|(\eta_k, y^k)| = 1$. Thus by passing to subsequences if necessary, we can arrange that (η_k, y^k) converges to some pair $(\eta, y) \neq (0, 0)$ with $\eta \geqq 0$.

In order to investigate the limiting condition satisfied by the pair (η, y), we apply the rules of subdifferential calculus to the sets in (7.6). It is apparent from (7.5) that

$$\partial f_k(u^k, G(u^k), g(G(u^k))) = \{(-z^k, 0, \eta_k)\}. \tag{7.7}$$

Furthermore from (7.4),

$$\partial F(u^k, G(u^k), g(G(u^k))) = (\partial G(u^k), -I, 0),$$
$$d_D(u, v, \alpha) = d_E(v, \alpha) \quad \text{for} \quad E = \operatorname{epi} g.$$

Therefore (7.6) is equivalent to

$$\eta_k z^k \in y^k \partial G(u^k), \quad (y^k, -\eta_k) \in [\eta_k \lambda_f + |y^k| \lambda_F] \partial d_E(w^k, g(w^k)), \tag{7.8}$$

where the multifunctions ∂G and ∂d_E are of "closed graph" because G and d_E are locally Lipschitzian (cf. propositions 2.2 and 2.6); this is the crucial property that has been obtained by invoking the more complicated multiplier rule in theorem 5.3 rather than the one in theorem 5.2.

Limits may therefore be taken in (7.8). We get

$$\eta z \in y \partial G(u), \quad (y, -\eta) \in [\eta \lambda_f + |y| \lambda_F] \partial d_E(\bar{w}, g(\bar{w})), \tag{7.9}$$

where $\bar{w} = G(\bar{u})$. Inasmuch as

$$\mathrm{cl} \bigcup_{\lambda \geqq 0} \lambda \partial d_E(\bar{w}, G(\bar{w})) = N_E(\bar{w}, G(\bar{w}))$$

(Clarke [9]), the second condition in (7.9) yields

$$(y, -\eta) \in N_{\mathrm{epi} g}(G(\bar{u}), g(G(\bar{u}))). \tag{7.10}$$

We know $(\eta, y) \neq (0, 0)$, $\eta \geqq 0$. If $\eta = 0$, then $y \neq 0$ and the first condition in (7.9) says $0 \in y \partial G(\bar{u})$, while (7.10) says $y \in \partial^\infty g(G(\bar{u}))$. This is impossible by assumption (7.1). Hence $\eta > 0$, and replacing y by $\eta^{-1} y$ if necessary we can reduce to the case where $\eta = 1$. Then the first condition in (7.9) says $z \in y \partial G(\bar{u})$, while (7.10) says $y \in \partial g(G(\bar{u}))$. In other words, $z \in \partial g(G(\bar{u})) \partial G(\bar{u}) = M(\bar{u})$.

A similar argument covers the case where instead of $z^k \to z$, we assume $t_k z^k \to z$ with $t_k \downarrow 0$ and aim at proving $z \in M^\infty(\bar{u})$. Then $t_k f_k$ has a local minimum at $(u^k, G(u^k), g(G(u^k)))$ relative to $(u, v, \alpha) \in D$, $F(u, v, \alpha) \in C$, and

$$\mathrm{Lip}_{t_k f_k}(u^k, G(u^k), g(G(u^k))) \leqq |t_k z^k| + t_k \to |z|.$$

Once again we may apply the multiplier rule in theorem 5.3 for fixed λ_f and λ_F, and this yields the same condition as in (7.6) but with $\eta_k t_k$ in place of the first η_k. It can be assumed that $(\eta_k, y^k) \to (\eta, y) \neq (0, 0)$, $\eta \geqq 0$. We reduce (7.6) as before using subgradient calculus, obtaining this time in place of (7.8)

$$\eta_k t_k z^k \in y^k \partial G(u^k), \quad (y^k, -\eta_k t_k) \in [\eta \lambda + |y^k| \lambda_F] \partial d_E(w^k, g(w^k))$$

for $w^k = G(u^k)$. In the limit this yields

$$\eta z \in y \partial G(\bar{u}), \quad (y, 0) \in [\eta \lambda + |y| \lambda_F] \partial d_E(\bar{w}, g(\bar{w})),$$

where the last condition implies

$$(y, 0) \in N_E(\bar{w}, \quad g(\bar{w})), \quad \text{i.e.} \quad y \in \partial^\infty g(G(\bar{u})).$$

If $\eta = 0$, then $y \neq 0$ and (7.1) would be contradicted. Therefore $\eta > 0$, and we can take $\eta = 1$. Thus we get the existence of some $y \in \partial^\infty g(G(\bar{u}))$ such that $z \in y \partial G(\bar{u})$, and we conclude $z \in M^\infty(\bar{u})$. ■

COROLLARY 7.1.1. Let $\bar{u} \in E = \{u \mid G(u) \in D\} = G^{-1}(D)$, where $G: R^m \to R^d$ is locally Lipschitzian and $D \subset R^d$ is closed. Assume

$$\nexists \text{ nonzero } y \in N_D(G(\bar{u})) \quad \text{with} \quad 0 \in y \partial G(\bar{u}). \tag{7.11}$$

Then for the set

$$M(\bar{u}) = N_D(G(\bar{u}))\partial G(\bar{u}) \tag{7.12}$$

one has $\hat{N}_E(\bar{u}) \subset M(\bar{u})$, so that all the assertions of corollary 2.1.1 are valid.

Proof. Take $g = \delta_D$ in the theorem. ■

Remark 7.2. Especially interesting in corollary 7.1.1 is the case where $D = \{0\}$. Then with the notation $G(u) = (g_1(u), \ldots, g_d(u)) \in R^d$ we have

$$E = \{u \,|\, g_i(u) = 0 \ \text{ for } i = 1, \ldots, m\}$$

(where each g^i is locally Lipschitzian from R^m to R). Furthermore $N_D(G(\bar{u})) = N_0(0) = R^d$ for any $\bar{u} \in E$. so the result says the following. If there does not exist a vector $y = (y_1, \ldots, y_d) \neq (0, \ldots, 0)$ with

$$0 \in \partial[y_1 g_1 + \ldots + y_m g_m](\bar{u})$$

(the latter condition being equivalent to $0 \in \partial G(\bar{u})$ by proposition 2.6), then for the set

$$M(\bar{u}) := \bigcup_{y \in R^m} \partial[y_1 g_1 + \ldots + y_m g_m](\bar{u})$$

one has $N_E(\bar{u}) \subset \text{cl co } M(\bar{u})$ as well as the other estimates in corollary 2.1.1.

More generally in place of $D = \{0\}$ one can consider

$$D = \{w = (w_1, \ldots, w_s, 0, \ldots, 0) \,|\, w_i \leqq 0 \quad \text{for} \quad i = 1, \ldots, s\},$$

so that

$$E = \{u \,|\, g_i(u) \leqq 0 \quad \text{for } i = 1, \ldots, s, \quad \text{and} \quad g_i(u) = 0 \quad \text{for} \quad i = s+1, \ldots, d\}.$$

Then for any $\bar{u} \in E$ one has

$$N_D(G(\bar{u})) = \{y \,|\, y_i \geqq 0 \quad \text{for} \quad i = 1, \ldots, s \quad \text{with} \quad g_i(\bar{u}) = 0,$$
$$y_i = 0 \quad \text{for} \quad i = 1, \ldots, s \quad \text{with}\, g_i(\bar{u}) < 0\}.$$

COROLLARY 7.1.2. Let $p(u) = g(G(u))$, where $G: R^m \to R^d$ is locally Lipschitzian and $g: R^d \to \bar{R}$ is lower semicontinuous. If G is strictly differentiable at $\bar{u}$ and

$$\nexists \text{ nonzero } y \in \partial^\infty g(G(\bar{u})) \quad \text{with} \quad y \nabla G(\bar{u}) = 0, \tag{7.13}$$

then

$$\partial p(\bar{u}) \subset \partial g(G(\bar{u}))\nabla G(\bar{u}), \quad \partial^\infty p(\bar{u}) \subset \partial^\infty g(G(\bar{u}))\nabla G(\bar{u}). \tag{7.14}$$

Proof. The set $\partial G(\bar{u})$ reduces in this case to the single matrix $\bar{A} = \nabla G(\bar{u})$. The sets $\partial g(G(\bar{u}))\bar{A}$ and $\partial^\infty g(G(\bar{u}))\bar{A}$ are then convex, in fact closed because of (7.13) (cf. [8, theorem 9.1]): for $\bar{w} = G(\bar{u})$, $\partial^\infty g(\bar{w})$ is the recession cone of the closed convex set $\partial g(\bar{w}) + \partial^\infty g(\bar{w})$ is always equal to $\partial g(\bar{w})$ by these definitions. The conclusions of proposition 2.1, which the theorem guarantees, then yield

$$\partial p(\bar u) \subset \operatorname{cl}\operatorname{co}[M(\bar u) + M^\infty(\bar u)] = \operatorname{cl}\operatorname{co}[\partial g(\bar w)\bar A + \partial^\infty g(\bar w)\bar A]$$
$$= \operatorname{cl}\operatorname{co}[[\partial g(\bar w) + \partial^\infty g(\bar w)]\bar A] = \operatorname{cl}\operatorname{co}[\partial g(\bar w)\bar A] = \partial g(\bar w)\bar A.$$

If $\partial p(\bar u) \neq \emptyset$, it follows from this inclusion that the recession cone of $\partial p(\bar u)$, which is $\partial^\infty p(\bar u)$, is included in the recession cone of $\partial g(\bar w)\bar A$, which (by (7.13), cf. [8, theorem 9.1]) is $\partial^\infty g(\bar w)\bar A$. If $\partial p(\bar u) = \emptyset$, one gets the same result via proposition 2.1:

$$\partial^\infty p(\bar u) \subset \operatorname{cl}\operatorname{co} M^\infty(\bar u) = \operatorname{cl}\operatorname{co}[\partial^\infty g(\bar w)\bar A] = \partial^\infty g(\bar w)\bar A.$$

Either way, the inclusions (7.14) are both correct. ■

COROLLARY 7.1.3. Let $G: R^m \to R^d$ be locally Lipschitzian, and let $D \subset R^d$ be closed. Suppose $\bar u \in E = \{u \mid G(u) \in D\} = G^{-1}(D)$ is a point where G is strictly differentiable and

$$\nexists \text{ nonzero } y \in N_D(G(\bar u)) \quad \text{with} \quad y\nabla G(\bar u) = 0. \tag{7.15}$$

Then

$$N_E(\bar u) \subset N_D(G(\bar u))\nabla G(\bar u). \tag{7.16}$$

Proof. This is the case of the preceding corollary where g is the indicator δ_D. ■

COROLLARY 7.1.4. (Clarke [3, p. 72].) Let $p(u) = g(G(u))$, where $G: R^m \to R^d$ is locally Lipschitzian and $g: R^d \to \bar R$ is locally Lipschitzian around $G(\bar u)$ for a certain $\bar u \in R^m$. Then p is locally Lipschitzian around $\bar u$ with

$$\partial p(\bar u) \subset \operatorname{co}[\partial g(G(\bar u))\partial G(\bar u)]. \tag{7.17}$$

If in fact g is strictly differentiable at $G(\bar u)$, then

$$\partial p(\bar u) \subset \nabla g(G(\bar u))\partial G(\bar u). \tag{7.18}$$

Proof. The assumption of Lipschitz continuity on g means that $\partial^\infty g(G(\bar u)) = \{0\}$; also, the convex set $\partial g(G(\bar u))$ is nonempty and compact. Condition (7.1) is satisfied vacuously in this case, and one has $M^\infty(\bar u) = \{0\}$ in (7.2). The last part of proposition 2.1 then asserts that p is locally Lipschitzian around $\bar u$ (actually this could also be verified directly) and (7.17) holds. If g is strictly differentiable at $G(\bar u)$, the set $\partial g(G(\bar u))$ consists of a single vector $\nabla g(G(\bar u))$, and since $\partial G(\bar u)$ is a convex set of matrices the convex hull operation in (7.17) can be dropped. ■

The result in corollary 7.1.4 is just a special case of a more general chain rule for locally Lipschitzian mappings which can be derived by the same method.

COROLLARY 7.1.5. (Clarke [3, p. 75].) Let $F(u) = H(G(u))$, where $G: R^m \to R^d$ and $H: R^d \to R^q$ are locally Lipschitzian. Then F is locally Lipschitzian and for all $\bar u$ one has

$$\partial F(\bar u)v \subset \operatorname{co}[\partial H(G(u))\partial G(u)]v \qquad \text{for} \quad v \in R^m. \tag{7.19}$$

Proof. For arbitrary $z \in R^q$ apply corollary 7.1.4 to $g(w) = z \cdot H(w)$. Draw the conclusion that $z \cdot F$ is locally Lipschitzian with

$$\begin{aligned}\partial(z \cdot F)(\bar{u}) &\subset \mathrm{co}\,[\partial(z \cdot H)(G(\bar{u})\partial G(\bar{u})]\\ &= \mathrm{co}\,[(z\partial H(g(\bar{u}))\partial G(\bar{u})] = z(\mathrm{co}\,[\partial H(g(\bar{u}))\partial G(\bar{u})]).\end{aligned} \tag{7.20}$$

Therefore F is locally Lipschitzian and $\partial(z \cdot F)(\bar{u}) = z\partial F(\bar{u})$ (proposition 2.6). The latter, together with (7.20) for all z, implies (7.19). ■

COROLLARY 7.1.6. Let $p(u) = \varphi(f(u))$, where $f: R^m \to R$ is locally Lipschitzian and $\varphi: R \to \bar{R}$ is lower semicontinuous. Let $\bar{u}$ be a point such that φ is finite at $f(\bar{u})$ and $0 \notin \partial f(\bar{u})$. Then for

$$M(\bar{u}) := \partial\varphi(f(\bar{u}))\partial f(\bar{u}), \qquad M^{\infty}(\bar{u}) := \partial^{\infty}\varphi(f(\bar{u}))\partial f(\bar{u}), \tag{7.21}$$

one has $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$ and $\hat{\partial}^{\infty} p(\bar{u}) \subset M^{\infty}(\bar{u})$, so that all the assertions of proposition 2.1 are valid.

If φ is nondecreasing on R, or if f is strictly differentiable at $\bar{u}$, then

$$\partial p(\bar{u}) \subset \partial\varphi(f(\bar{u}))\partial f(\bar{u}), \qquad \partial^{\infty} p(\bar{u}) \subset \partial^{\infty}\varphi(f(\bar{u}))\partial f(\bar{u}) \tag{7.22}$$

(where in the second of these cases $\partial f(\bar{u})$ reduces to $\nabla f(\bar{u})$).

Proof. Take $G = f, g = \varphi$ (the case of $R^d = R$). Specialize theorem 7.1 and corollary 7.1.2. Observe that when φ is nondecreasing, the sets $\partial\varphi(f(\bar{u}))$ and $\partial^{\infty}\varphi(f(\bar{u}))$ are subintervals of $[0, \infty)$, so $M(\bar{u})$ and $M^{\infty}(\bar{u})$ in (7.21) are convex. Then the convex hull operation is superfluous. ■

8. OTHER SUBGRADIENT FORMULAS AND A PARAMETRIC MULTIPLIER RULE

The chain rules in theorem 7.1 and its corollaries lead to other rules of subdifferentiation through the technique of representing a given kind of function as the composition of some other function with a Lipschitzian transformation. We demonstrate this first with a generalization of the rule for subgradients of sums of functions (cf. proposition 2.4 and more generally [13, theorem 2]).

THEOREM 8.1. Let $f = f_1 + \dots + f_r$, where $f_i: R^n \to \bar{R}$ is lower semicontinuous, and let $\bar{x}$ be a point where all the functions f_i are finite. Suppose

$$\nexists y_i \in \partial^{\infty} f_i(\bar{x}) \quad \text{with} \quad y_1 + \dots + y_r = 0, \text{ except } \quad y_1 = \dots = y_r = 0. \tag{8.1}$$

Then

$$\begin{aligned}\partial f(\bar{x}) &\subset \partial f_1(\bar{x}) + \dots + \partial f_r(\bar{x}),\\ \partial^{\infty} f(\bar{x}) &\subset \partial^{\infty} f_1(\bar{x}) + \dots + \partial^{\infty} f_r(\bar{x}).\end{aligned} \tag{8.2}$$

Proof. Define $g(u_1, \dots, u_r) = g_1(u_1) + \dots + g_r(u_r)$ on $(R^n)^r$, and define $G: R^n \to (R^n)^r$ by $G(u) = (u, \dots, u)$. Then $f(u) = g(G(u))$, and we are in the realm of corollary 7.1.2 with assumptions (7.13) fulfilled at $u = \bar{x}$. (A vector y of the kind forbidden in (7.13) would correspond to an r-tuple $(y_1, \dots, y_r) \in (R^n)^r$ of the kind forbidden in (8.1).) The inclusions in (7.14) reduce for $p = f$ to the ones in (8.2). ■

Remark 8.1. Proposition 2.4 is, of course, the case of theorem 8.1 where $f_1 = g, f_2 = h$, $\partial^\infty f_2(\bar{x}) = \{0\}$. In general for $r = 2$, condition (8.1) means that the nonempty convex cones $D_i = \{w | f_i^\uparrow(\bar{x}; w) < \infty\}$, where $f_i^\uparrow(\bar{x}; w)$ is the directional subdirective defined in [14], cannot be separated: $\partial^\infty f_i(\bar{x})$ is the cone polar to D_i (cf. argument used in proving proposition 2.4). In particular D_1 and D_2 cannot be separated if $D_1 \cap \operatorname{int} D_2 \neq \emptyset$. In this case the formulas for $\partial(f_1 + f_2)(\bar{x})$ and $\partial^\infty(f_1 + f_2)(\bar{x})$ are covered by theorem 2 of [13], a result which is valid in infinite-dimensional spaces and also provides conditions under which equality holds in (8.2). The technique of proof used here, while it allows a weakening of the separation hypothesis in a finite-dimensional setting, does not seem to provide corresponding results about the possibility of equality.

COROLLARY 8.1.1. Let $\bar{x} \in C = C_1 \cap \ldots \cap C_r$, where $C_i \subset R^n$ is closed. Suppose

$$\nexists y_i \in N_{C_i}(\bar{x}) \quad \text{with} \quad y_1 + \ldots + y_r = 0, \quad \text{except } y_1 = \ldots = y_r = 0. \tag{8.3}$$

Then

$$N_C(\bar{x}) \subset N_{C_1}(\bar{x}) + \ldots + N_{C_r}(\bar{x}). \tag{8.4}$$

Proof. Take f_i in theorem 8.1 to be the indicator function δ_{C_i}, cf. (1.8). ■

COROLLARY 8.1.2. Let $f: R^n \to \bar{R}$ be lower semicontinuous and let $D \subset R^n$ be closed. Let $\bar{x} \in D$ be a point where f is finite and such that

$$\nexists \text{ nonzero } z \in \partial^\infty f(\bar{x}) \quad \text{with} \quad -z \in N_D(\bar{x}).$$

Then

$$\partial(f + \delta_D)(\bar{x}) \subset \partial f(\bar{x}) + N_D(\bar{x}), \qquad \partial^\infty(f + \delta_D)(\bar{x}) \subset \partial^\infty f(\bar{x}) + N_D(\bar{x}).$$

Proof. This specializes the theorem again in terms of (1.8). ■

COROLLARY 8.1.3. Let $h(x) = f(x) + g(F(x))$, where $f: R^n \to \bar{R}$ and $g: R^m \to \bar{R}$ are lower semicontinuous, and $F: R^n \to R^m$ is continuously differentiable. Let $\bar{x}$ be a point where h is finite, and suppose

$$\nexists \text{ nonzero } y \in \partial^\infty g(F(\bar{x})) \quad \text{with} \quad 0 \in \partial^\infty f(\bar{x}) + y\nabla F(\bar{x}).$$

Then

$$\partial h(\bar{x}) \subset \partial f(\bar{x}) + \nabla g(F(\bar{x}))\nabla F(\bar{x}), \qquad \partial^\infty h(\bar{x}) \subset \partial^\infty f(\bar{x}) + \partial g(F(\bar{x}))\nabla F(\bar{x}).$$

Proof. This combines the two-function case of theorem 8.1 with the chain rule in corollary 7.1.2. ■

COROLLARY 8.1.4. Let $E = D \cap F^{-1}(C)$, where $D \subset R^n$ and $C \cap R^m$ are closed sets, and $F: R^n \to R^m$ is continuously differentiable. Let $\bar{x}$ be a point of E such that

$$\nexists \text{ nonzero } y \in N_C(F(\bar{x})) \quad \text{with} \quad 0 \in N_D(\bar{x}) + y\nabla F(\bar{x}). \tag{8.5}$$

Then

$$N_E(\bar{x}) \subset N_D(\bar{x}) + N_C(F(\bar{x}))\nabla F(\bar{x}). \tag{8.6}$$

Proof. Take $f = \delta_D$ and $g = \delta_C$ in corollary 8.13. ■

Remark 8.2. In terms of the polars of the cones $N_D(\bar{x})$ and $N_C(F(\bar{x}))$, which are the tangent cones $T_D(\bar{x})$ and $T_C(F(\bar{x}))$ (see Clarke [3, p. 51]), the assertion of corollary 8.1.4 yields the following: if the convex cones $T_C(F(\bar{x}))$ and $\nabla F(\bar{x})T_D(\bar{x})$ cannot be separated (i.e. if $T_C(F(\bar{x})) - \nabla F(\bar{x})T_D(\bar{x})$ is all of R^n), then

$$T_E(\bar{x}) \supset T_D(\bar{x}) \cap \nabla F(\bar{x})^{-1}T_C(F(\bar{x}))$$

(the inverse being taken in the sense of an inverse multifunction, not necessarily single-valued). This result has been proved by Aubin [1, Section 4].

Proof. This specializes the theorem again in terms of (1.8). ■

THEOREM 8.2. Let $p(u) = g(\bar{x}, u)$, where $g: R^n \times R^m \to \bar{R}$ is lower semicontinuous and $\bar{x} \in R^n$ is fixed. Let $\bar{u} \in R^m$ be such that $g(\bar{x}, \bar{u})$ is finite and

$$\nexists \text{ nonzero } z \quad \text{with} \quad (z, 0) \in \partial^\infty g(\bar{x}, \bar{u}). \tag{8.7}$$

Then

$$\partial p(\bar{u}) \subset \{y \mid \exists z \quad \text{with} \quad (z, y) \in \partial g(\bar{x}, \bar{u})\},$$

$$\partial^\infty p(\bar{u}) \subset \{y \mid \exists z \quad \text{with} \quad (z, y) \in \partial^\infty g(\bar{x}, \bar{u})\}. \tag{8.8}$$

Proof. Define $G: R^m \to R^n \times R^m$ by $G(u) = (\bar{x}, u)$. Then $p(u) = g(G(u))$, and the hypothesis of corollary 7.1.2 is satisfied: one has $(z, y)\nabla G(\bar{u}) = y$, so this image is 0 for a nonzero element (z, y) of $\partial^\infty g(\bar{x}, \bar{u})$ if and only if the element is of the form $(z, 0)$ with $z \neq 0$. Relations (7.14) turn into (8.8). ■

Remark 8.3. Theorem 8.2 strengthens our previous result on this matter in [11, proposition 4], which gave these conclusions only under the additional assumption that $\partial^\infty g(\bar{x}, \bar{u})$ is pointed.

COROLLARY 8.2.1. Let $E = \{u \mid (\bar{x}, u) \in D\}$, where $D \subset R^n \times R^m$ is closed and $\bar{x} \in R^n$ is fixed. Let $\bar{u} \in R^m$ be such that $(\bar{x}, \bar{u}) \in D$ and

$$\nexists \text{ nonzero } z \quad \text{with} \quad (z, 0) \in N_D(\bar{x}, \bar{u}). \tag{8.9}$$

Then

$$N_E(\bar{u}) \subset \{y \mid \exists z \quad \text{with} \quad (z, y) \in N_D(\bar{x}, \bar{u})\}. \tag{8.10}$$

Proof. Take $g = \delta_D$ in theorem 8.2. Then $p = \delta_E$. ■

COROLLARY 8.2.2. Let $E = \{u \in R^m \mid f(u) \leqq 0\}$, where $f: R^m \to \bar{R}$ is lower semicontinuous. Let $\bar{u}$ be a point where $f(\bar{u}) = 0$ but $0 \notin \partial f(\bar{u})$. Then

$$N_E(\bar{u}) \subset \left[\bigcup_{\lambda > 0} \lambda \partial f(\bar{u})\right] \cup \partial^\infty f(\bar{u}).$$

Proof. We shall invoke the preceding corollary. Let $D = \text{epi} f \subset R^m \times R$, so that $E = \{u \mid (u, 0) \in D\}$. Recall that

$$N_D(\bar{u}, 0) = N_{\text{epi} f}(\bar{u}, f(\bar{u}))$$
$$= \Big[\bigcup_{\lambda > 0} \lambda(\partial f(\bar{u}), -1)\Big] \cup (\partial^\infty f(\bar{u}), 0).$$

There does not exist $z \neq 0$ with $(0, z) \notin N_D(\bar{u}, 0)$, for then there would exist $\lambda > 0$ with $(0, -\lambda) \in N_D(\bar{u}, 0)$, and we would have $0 \in \partial f(\bar{u})$, contrary to hypothesis. Therefore

$$N_E(\bar{u}) \subset \{y \mid \exists \lambda \in R \quad \text{with} \quad (y, -\lambda) \in N_D(\bar{u}, 0)\}.$$

This inclusion reduces to the one claimed. ∎

The normal cone estimate in corollary 8.2.2 generalizes the result obtained by the author in [13, theorem 5]. That result, valid in an infinite-dimensional setting, requires that f be directionally Lipschitzian at $\bar{u}$, a condition equivalent in the finite-dimensional case to $\partial f(u)$ being nonempty but not including any whole line. However, that version also provides a criterion for the inclusion to hold as an equation.

Our final result is an extension of theorem 3.2 to a more general class of perturbations.

THEOREM 8.3. Consider

$$\begin{aligned} p(u) &= \inf_x \{f(x, u) \mid F(x, u) \in C, (x, u) \in D\}, \\ X(u) &= \operatorname{argmin}_x \{f(x, u) \mid F(x, u) \in C, (x, u) \in D\}, \end{aligned} \tag{8.11}$$

where $C \subset R^m$ and $D \subset R^n \times R^d$ are closed, $F: R^n \times R^d \to R^m$ is locally Lipschitzian, and $f: R^n \times R^d \to \bar{R}$ is lower semicontinuous. Suppose $\bar{u}$ is a point where $p(\bar{u})$ is finite, and that

$$\exists\, \varepsilon > 0 \quad \text{and} \quad \bar{\alpha} > p(\bar{u}) \quad \text{such that the set}$$
$$\{(x, u) \mid f(x, u) \leqq \bar{\alpha}, F(x, u) \in C, (x, u) \in D, |u - \bar{u}| \leqq \varepsilon\} \text{ is bounded.} \tag{8.12}$$

Then p is strictly lower semicontinuous at $\bar{u}$, and $X(\bar{u})$ is nonempty and compact. If in addition each such $\bar{x} \in X(\bar{u})$ satisfies the constraint qualification

$$\nexists (z_1, v_1) \in \partial^\infty f(\bar{x}, \bar{u}), \quad (z_2, v_2) \in N_D(\bar{x}, \bar{u}), \quad y \in N_C(F(\bar{x}, \bar{u})) \tag{8.13}$$

such that $(z_1, v_1, z_2, v_2, y) \neq (0, 0, 0, 0, 0)$ but $-(z_1 + z_2, v_1 + v_2) \in y\partial F(\bar{x}, \bar{u})$, then for the sets

$$\begin{aligned} M(\bar{u}) = \{v \mid \exists y \in N_C(F(\bar{x}, \bar{u})) \quad &\text{with} \\ (0, v) \in \partial f(\bar{x}, \bar{u}) + y\partial F(\bar{x}, \bar{u}) &+ N_D(\bar{x}, \bar{u})\}, \\ M^\infty(\bar{u}) = \{v \mid \exists y \in N_C(F(\bar{x}, \bar{u})) \quad &\text{with} \\ (0, v) \in \partial^\infty f(\bar{x}, \bar{u}) + y\partial F(\bar{x}, \bar{u}) &+ N_D(\bar{x}, \bar{u})\} \end{aligned} \tag{8.14}$$

one has $\hat{\partial} p(\bar{u}) \subset M(\bar{u})$ and $\hat{\partial}^\infty p(\bar{u}) \subset M^\infty(\bar{u})$, so all the assertions of proposition 2.1 are valid.

Proof. Define

$$P(x,u)=\begin{cases}f(x,u) & \text{if } F(x,u)\in C,(x,u)\in D,\\ +\infty & \text{otherwise.}\end{cases} \tag{8.15}$$

Then P and p fit the pattern of theorem 3.1, all the assumptions in that result being fulfilled. What we need to demonstrate is that under (8.3), every pair $(0,v)\in\hat{\partial}P(\bar{x},\bar{u})$ satisfies the condition

$$\exists y\in N_C(F(\bar{x},\bar{u}))\quad\text{with}\quad (0,v)\in\partial f(\bar{x},\bar{u})+y\partial F(\bar{x},\bar{u})+N_D(\bar{x},\bar{u}), \tag{8.16}$$

while every $(0,v)\in\hat{\partial}^\infty P(\bar{x},\bar{u})$ satisfies the corresponding condition where $\partial^\infty f(\bar{x},\bar{u})$ appears in place of $\partial f(\bar{x},\bar{u})$. To this end we write $P(x,u)=g(G(x,u))$, where

$$G(x,u)=(x,u,x,u,F(x,u))\in R^n\times R^d\times R^n\times R^d\times R^m, \tag{8.17}$$

$$g(x_1,u_1,x_2,u_2,w)=f(x_1,u_1)+\delta_D(x_2,u_2)+\delta_C(w).$$

The chain rule in theorem 7.1 will be applied. We note that

$$\partial g(G(\bar{x},\bar{u}))=(\partial f(\bar{x},\bar{u}),N_D(\bar{x},\bar{u}),N_C(F(\bar{x},\bar{u}))),$$

$$\partial^\infty g(G(\bar{x},\bar{u}))=(\bar{\partial}^\infty f(\bar{x},\bar{u}),N_D(\bar{x},\bar{u}),N_C)F(\bar{x},\bar{u}))). \tag{8.18}$$

On the other hand,

$$(z_1,v_1,z_2,v_2,y)\partial G(\bar{x},\bar{u})=(z_1+z_2,v_1+v_2)+y\partial F(\bar{x},\bar{u}).$$

Assumption (7.1) in theorem 7.1 thus becomes (8.13), which is assumed here. It follows that

$$\hat{\partial}P(\bar{x},\bar{u})\subset\partial g(G(\bar{x},\bar{u}))\partial G(\bar{x},\bar{u}),$$

$$\hat{\partial}^\infty P(\bar{x},\bar{u})\subset\partial^\infty g(G(\bar{x},\bar{u}))\partial G(\bar{x},\bar{u}),$$

and by virtue of (8.17) and (8.18) this is all we had to show. ■

Remark 8.4. Theorem 8.3 reduces to the version of theorem 3.2 in remark 3.4 in the special case where $R^d=R^m$, $F(x,u)=F_0(\bar{x})+u$, $f(x,u)=f_0(x)$, $D=D_0\times R^m$.

THEOREM 8.4. (Parametric multiplier rule.) Let $\bar{x}$ be a locally optimal solution to the problem

$$\text{minimize } f(x,\bar{u}) \text{ over all } x \text{ satisfying } F(x,\bar{u})\in C,\ (x,\bar{u})\in D,$$

where $C\subset R^n$ and $D\subset R^n\times R^d$ are closed, $F:R^n\times R^d\to R^m$ is locally Lipschitzian, and $f:R^n\times R^d\to\bar{R}$ is lower semicontinuous. Suppose $f(\bar{x},\bar{u})$ is finite and either that the problem satisfies the calmness condition

$$\nexists(x^k,u^k)\to(\bar{x},\bar{u})\quad\text{with}\quad F(x^k,u^k)\in C,\ (x^k,u^k)\in D,$$

$$\text{such that } u^k\neq\bar{u}\quad\text{and}\quad [f(x^k,u^k)-f(\bar{x},\bar{u})]/|u^k-\bar{u}|\to-\infty, \tag{8.19}$$

or that $\bar{x}$ satisfies the constraint qualification (8.13). Then

$$\exists y\in N_C(F(\bar{x},\bar{u}))\quad\text{and}\quad v\in R^d\quad\text{with}$$

$$(0,v)\in\partial f(\bar{x},\bar{u})+y\partial F(\bar{x},u)+N_D(\bar{x},\bar{u}). \tag{8.20}$$

Proof. The argument is the same as the proof of theorem 5.2, but with theorem 8.3 used in place of theorem 3.2. ■

Remark 8.5. In the special case mentioned in remark 8.3, the multiplier rule in theorem 8.4 reduces to the one of theorem 5.3 as expanded in remark 5.3. The new result generalizes the parametric multiplier rule given by the author in [11, theorem 2].

REFERENCES

1. AUBIN J.-P., Lipschitz behavior of solutions to convex minimization problems, *Math. Operat. Res.* **9**, 87–111 (1984).
2. CLARKE F. H., Necessary conditions for nonsmooth problems in optimal control and the calculus of variations, Ph.D. thesis, University of Washington, Seattle (1973).
3. CLARKE F. H., *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York (1983).
4. CLARKE F. H., Generalized gradients and applications, *Trans. Am. math. Soc.* **205**, 247–262 (1975).
5. CLARKE F. H., A new approach to Lagrange multipliers, *Math. Operat. Res.* **1**, 165–174 (1976).
6. EKELAND I., On the variational principle, *J. math. Analysis Applic.* **47**, 324–353 (1974).
7. HIRIART-URRUTY J.-B., Refinements of necessary optimality conditions in nondifferentiable programming I, *Appl. Math. Optim.* **5**, 63–82 (1979).
8. ROCKAFELLAR R. T., *Convex Analysis*, Princeton University Press, Princeton, NJ (1980).
9. ROCKAFELLAR R. T., *The Theory of Subgradients and its Applications to Problems of Optimization: Convex and Nonconvex Functions*, Heldermann, Berlin (1981).
10. ROCKAFELLAR R. T., Clarke's tangent cones and the boundaries of closed sets in R^n, *Nonlinear Analysis* **3**, 145–154 (1979).
11. ROCKAFELLAR R. T., Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming, *Math. Prog. Study* **17**, 28–66 (1982).
12. ROCKAFELLAR R. T., Proximal subgradients, marginal values, and augmented Lagrangians in conconvex optimization, *Math. Operat. Res.* **6**, 427–437 (1981).
13. ROCKAFELLAR R. T., Directionally Lipschitzian functions and subdifferential calculus, *Proc. Lond. Math. Soc.* **39**, 331–355 (1979).
14. ROCKAFELLAR R. T., Generalized directional derivatives and subgradients of nonconvex functions, *Can. J. Math.* **32**, 257–280 (1980).
15. TREIMAN J., A new characterization of Clarke's tangent cone and its applications to subgradient analysis and optimization, Ph.D. thesis, University of Washington (1983).
16. WETS R. J.-B., On inf-compact mathematical programs, Fifth Conference on Optimization Techniques, Part 1, *Lecture Notes in Computer Sciences* **3**, 426–436, Springer, Berlin (1973).