NONSMOOTH CALCULUS IN FINITE DIMENSIONS*

D. E. WARD† AND J. M. BORWEIN‡

Abstract. The notion of subgradient, originally defined for convex functions, has in recent years been extended, via the “upper subderivative,” to cover functions that are not necessarily convex or even continuous. A number of calculus rules have been proven for these generalized subgradients. This paper develops the finite-dimensional generalized subdifferential calculus for (strictly) lower semicontinuous functions under considerably weaker hypotheses than those previously used. The most general finite-dimensional convex subdifferential calculus results are recovered as corollaries. Other corollaries given include new necessary conditions for optimality in a nonsmooth mathematical program. Various chain rule formulations are considered. Equality in the subdifferential calculus formulae is proven under hypotheses weaker than the usual “subdifferential regularity” assumptions.

Key words. Clarke tangent cone, upper subderivative, subgradient, contingent cone, subdifferential regularity

AMS(MOS) subject classifications. Primary 46G05; secondary 58C20, 90C30.

1. Introduction. Let E be a real locally convex Hausdorff topological vector space (l.c.t.v.s.) with dual space E^* . For $x \in E$, denote by $\mathcal{N}(x)$ the set of all neighborhoods of x .

DEFINITION 1.1. Let $C \subset E$ and $x_0 \in \text{cl } C$. The *Clarke tangent cone* to C at x_0 is the set

$$(1.1) \quad T_C(x_0) := \{y \in E \mid \forall Y \in \mathcal{N}(y), \exists X \in \mathcal{N}(x_0), \text{ and } \exists \lambda > 0 \text{ such that } \forall x' \in X \cap C \text{ and } \forall t \in (0, \lambda), \exists y' \in Y \text{ with } x' + ty' \in C\}.$$

The *normal cone* to C at x_0 is the set

$$(1.2) \quad N_C(x_0) := T_C(x_0)^0 := \{z \in E^* \mid \langle y, z \rangle \leq 0 \ \forall y \in T_C(x_0)\}.$$

$T_C(x_0)$ is always closed and contains the origin. More important, it is always a convex set. (See [17], [20, Chap. 1].)

DEFINITION 1.2. Let $f: E \rightarrow \bar{\mathbf{R}}$ be an extended-real-valued function on E . The *domain* of f is the set

$$(1.3) \quad \text{dom } f := \{x \in E \mid f(x) < +\infty\}.$$

The *epigraph* of f is the set

$$(1.4) \quad \text{epi } f := \{(x, r) \in E \times \mathbf{R} \mid f(x) \leq r\}.$$

Let $x_0 \in E$ be a point at which f is finite, and let $y \in E$. The upper subderivative of f at x_0 in the direction y is defined by

$$(1.5) \quad f^\uparrow(x_0; y) := \inf \{r \mid (y, r) \in T_{\text{epi } f}(x_0, f(x_0))\}.$$

Notice that $f^\uparrow(x_0; y)$ is defined precisely so that

$$(1.6) \quad \text{epi } f^\uparrow(x_0; y) = T_{\text{epi } f}(x_0, f(x_0)).$$

By (1.6), $f^\uparrow(x_0; \cdot)$ has several important properties. Since $T_{\text{epi } f}(x_0, f(x_0))$ is closed, $f^\uparrow(x_0; \cdot)$ is lower semicontinuous (l.s.c.), and since $T_{\text{epi } f}(x_0, f(x_0))$ is a convex cone, $f^\uparrow(x_0; \cdot)$ is convex and positively homogeneous. As a result, either

* Received by the editors April 3, 1984; accepted for publication (in revised form) February 6, 1986.

† Department of Mathematics and Statistics, Miami University, Oxford, Ohio 45056.

‡ Department of Mathematics, Statistics, and Computing Science, Dalhousie University, Halifax, Nova Scotia, Canada B3H 3J5.

(a) $f^\uparrow(x_0; 0) = 0$, in which case $f^\uparrow(x_0; \cdot)$ is *proper*, i.e., is somewhere finite and never $= -\infty$, or

(b) $f^\uparrow(x_0; 0) = -\infty$, in which case $f^\uparrow(x_0; \cdot)$ is equal to $-\infty$ throughout its domain. For further discussion of the idea of associating directional derivatives with approximating cones to epigraphs of functions, in particular the Clarke tangent cone, see [12], [7] and [20, Chap. 3].

DEFINITION 1.3. Let $f: E \rightarrow \bar{\mathbf{R}}$ and $x_0 \in E$ be such that $f(x_0)$ is finite. The *subgradient* of f at x_0 is the set

$$(1.7) \quad \partial f(x_0) := \{x^* \in E^* \mid \langle y, x^* \rangle \leq f^\uparrow(x_0; y) \forall y \in E\}.$$

The notation $\partial f(x_0)$ is the same as that for the subgradient of a convex function or the Clarke subgradient of a locally Lipschitzian function, and justifiably so. If f is locally Lipschitzian, $\partial f(x_0)$ coincides with the Clarke subgradient, which in turn coincides with the ordinary subgradient for convex functions if f is convex. (See [18], [19], [20].)

In [19], Rockafellar introduced the following direct characterization of $f^\uparrow(x_0; \cdot)$:

$$(1.8) \quad f^\uparrow(x_0; y) = \sup_{Y \in \mathcal{N}(y)} \inf_{\substack{N \in \mathcal{N}(x_0, f(x_0)) \\ \lambda > 0}} \sup_{\substack{(x, \alpha) \in \text{epi } f \cap N \\ t \in (0, \lambda)}} \inf_{y' \in Y} \frac{f(x + ty) - \alpha}{t}.$$

If f is l.s.c., this definition reduces to

$$(1.9) \quad \begin{aligned} f^\uparrow(x_0; y) &= \limsup_{\substack{x \rightarrow x_0 \\ t \downarrow 0}} \inf_{y' \rightarrow y} \frac{f(x + ty') - f(x)}{t} \\ &:= \sup_{y \in \mathcal{N}(y)} \inf_{\substack{X \in \mathcal{N}(x_0) \\ \lambda > 0}} \sup_{\substack{x \in X \\ t \in (0, \lambda)}} \inf_{y' \in Y} \frac{f(x + ty') - f(x)}{t}, \\ &\quad f(x) \equiv f(x_0) + \lambda \end{aligned}$$

If f is continuous, $f^\uparrow(x_0; y)$ may be written more simply as

$$(1.10) \quad \begin{aligned} f^\uparrow(x_0; y) &= \limsup_{\substack{x \rightarrow x_0 \\ t \downarrow 0}} \inf_{y' \rightarrow y} \frac{f(x + ty') - f(x)}{t} \\ &:= \sup_{Y \in \mathcal{N}(y)} \inf_{\substack{x \in \mathcal{N}(x_0) \\ \lambda > 0}} \sup_{\substack{x \in X \\ t \in (0, \lambda)}} \inf_{y' \in Y} \frac{f(x + ty') - f(x)}{t}. \end{aligned}$$

Rockafellar used characterization (1.8) in [18] to extend the subdifferential calculus to the subgradient of (1.7). (See also [6, § 2.9].)

In this paper, we show that if E is finite-dimensional and $f: E \rightarrow \bar{\mathbf{R}}$ is strictly l.s.c., the hypotheses in the subdifferential calculus theorems of [18] can be considerably weakened, and the hypotheses required in proving multiplier rules for nonsmooth mathematical programs can be correspondingly weakened. This is in analogy with the convex case, where in finite dimensions, interiority assumptions are replaced by assumptions about relative interiors. In fact, the convex subgradient calculus of § 23 of [16] can be entirely recaptured as a special case of the results presented here.

Our method of proof differs from that of [18] in that instead of working directly with $f^\uparrow(x_0; \cdot)$ as defined in (1.8), we use an “inversion theorem” (see [4]), the relationship in (1.6), and underlying properties of the Clarke tangent cone.

Here is an outline of the remainder of the paper. In § 2, we list the definitions and collect the preliminary results that we will use in proving subdifferential calculus

formulae. In § 3, we prove our two main subdifferential calculus formulae and give a number of corollaries. In § 4, we apply the results of § 3 to the study of necessary optimality conditions in nonsmooth mathematical programming. In particular, we prove a finite-dimensional version of Theorem 6 of [18] under weaker hypotheses and apply a “Dubovitskii–Milyutin” approach to prove a “Fritz John type” result as in [26]. We investigate in § 5 the possibility of extending a chain rule of Hiriart-Urruty [11, Chap. 8] to functions which are not necessarily locally Lipschitzian. In § 6 we prove inequalities involving the contingent [1] and Ursescu [9] directional derivatives. We then combine these results with those of § 3 to give conditions for equality in the subdifferential calculus theorems of § 3 that somewhat relax the usual “subdifferential regularity” conditions. We also give a generalization of one of our results of § 4 through the use of “upper convex approximates” to these directional derivatives.

An excellent background reference for this paper is Chapter 7 of Aubin and Ekeland’s recent book [3]. In particular, Corollary 3.4 and Propositions 3.10 and 3.14 are derived in [3] by an approach in many ways similar to that employed here.

Two entirely different approaches to this subject can be found in the significant papers [13] and [22]. In [13], Ioffe essentially proves Theorem 3.2 and its corollaries as special cases of corresponding formulae in the calculus of “approximate subdifferentials.” Rockafellar in [22] derives subgradient inclusions (3.3), (3.6), (3.30) and (4.2) by an approach that is “dual” to that taken here and in [18]. The methods used in [22] center around the concepts of the normal cone and proximal normals. Both [13] and [22] go further than this present work in some directions. On the other hand, the chain rule formulation in Theorem 3.17 and the material in § 6 are not discussed in either [13] or [22]. In fact, it is not clear how one might derive Theorem 3.17 by the methods of [13], and the methods of [22] do not readily yield conditions for equality in subdifferential calculus formulae. References [3], [6], [11], [13], [18], [22] and this work combine to form quite an extensive body of information about the generalized subdifferential calculus.

2. Preliminaries.

DEFINITION 2.1. Let E, E_1 be l.c.t.v.s. A function $G: E \rightarrow E_1$ is said to be *strictly differentiable* at $x_0 \in E$ if there exists a linear mapping $\nabla G(x_0): E \rightarrow E_1$ such that

$$(2.1) \quad \lim_{\substack{x \rightarrow x_0 \\ y' \rightarrow y \\ t \downarrow 0}} \frac{G(x + ty') - G(x)}{t} = \nabla G(x_0)y$$

for all $y \in E$.

DEFINITION 2.2. Let E be an l.c.t.v.s.

(a) The set $C \subset E$ is said to be closed near $x_0 \in C$ if there exists $X \in \mathcal{N}(x_0)$ such that $X \cap C$ is closed.

(b) The function $f: E \rightarrow \bar{\mathbf{R}}$ is *strictly l.s.c.* at $x_0 \in E$ if for some $\alpha > f(x_0)$, the function $\min\{f, \alpha\}$ is l.s.c.

It is observed in [21] that if f is strictly l.s.c. at x_0 , then the set $\text{epi } f$ is closed near $(x_0, f(x_0))$.

DEFINITION 2.3. $F: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ is *isotone* on $D \subset \mathbf{R}^n$ if $F(x) \leq F(y)$ whenever $x, y \in D$ and $x \leq y$ (with respect to the coordinate ordering). F is *strictly isotone* in the i th coordinate at $x := (x_1, \dots, x_n) \in \mathbf{R}^n$ if $F(x) < F(y)$ whenever $x \leq y$ and $x_i < y_i$.

The result that allows us to weaken the “constraint qualifications” of [18] is the following special case of Theorem 4.1 of [4], which itself follows from Ekeland’s variational principle [8] (see also [3, § 7.6]).

THEOREM 2.4. *Let $G: \mathbf{R}^p \rightarrow \mathbf{R}^q$ be strictly differentiable at $x_0 \in D \cap G^{-1}(0)$, where $D \subset \mathbf{R}^p$ is closed near x_0 . Assume that*

$$(2.2) \quad \nabla G(x_0) T_D(x_0) = \mathbf{R}^q.$$

Then

$$(2.3) \quad T_D(x_0) \cap \nabla G(x_0)^{-1}(0) \subset T_{D \cap G^{-1}(0)}(x_0).$$

We will use the following two properties of the Clarke tangent cone in our proofs in § 3. The first follows easily from (1.1).

PROPOSITION 2.5. *Let E, E_1 be l.c.t.v.s., and let $x_0 \in C \subset E$ and $y_0 \in D \subset E_1$. Then*

$$(2.4) \quad T_{C \times D}(x_0, y_0) = T_C(x_0) \times T_D(y_0).$$

PROPOSITION 2.6. *Let E, E_1 be l.c.t.v.s. and $A: E \rightarrow E_1$ be linear and continuous. Let $z_0 \in C \subset E$. Suppose that A is relatively open on C at z_0 ; i.e.,*

$$(2.5) \quad \text{For each } X \in \mathcal{N}(z_0), \text{ there exists } Z \in \mathcal{N}(Az_0) \text{ such that } Z \cap A(C) \subset A(X \cap C).$$

Then

$$(2.6) \quad A(T_C(z_0)) \subset T_{A(C)}(Az_0).$$

Proof. Let $y \in T_C(z_0)$ and $Y \in \mathcal{N}(A(y))$. Then $Y' := A^{-1}(Y)$ is in $\mathcal{N}(y)$. There exists $\lambda > 0$, $X \in \mathcal{N}(z_0)$ such that for all $t \in (0, \lambda)$ and for all $x' \in X \cap C$,

$$(x' + tY') \cap C \neq \emptyset.$$

By (2.5), there exists $Z \in \mathcal{N}(Az_0)$ such that

$$Z \cap A(C) \subset A(X \cap C).$$

Then for each $z \in Z \cap A(C)$, there exists $x' \in X \cap C$ such that $A(x') = z$. For such an x' , $(x' + tY') \cap C \neq \emptyset$ for all $t \in (0, \lambda)$. Hence $Ax' + tY \cap A(C) \neq \emptyset$ for all $t \in (0, \lambda)$, and so $A(y) \in T_{A(C)}(Az_0)$. Thus $A(T_C(z_0)) \subset T_{A(C)}(Az_0)$. $\square$

Remark 2.7. (a) Condition (2.5) holds in particular whenever A is open and one-to-one on C .

(b) Proposition 2.6 is a special case of Corollary 4.2 of [4] (see also [14]).

In § 3, we will establish calculus rules involving functions of two forms:

(a) $h := f_1 + f_2 \circ F$, where $f_1: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ is strictly l.s.c. at x_0 , $f_2: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ is strictly l.s.c. at $F(x_0)$, and $F: \mathbf{R}^n \rightarrow \mathbf{R}^m$ is strictly differentiable at $x_0 \in \text{dom } f_1 \cap F^{-1}(\text{dom } f_2)$.

(b) $h := F \circ f$, where $f = (f_1, \dots, f_n)$, each $f_i: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ is strictly l.s.c. at x_0 , and $F: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ is isotone.

In (b), we define $F(f(x))$ to be $+\infty$ whenever $f_i(x) = +\infty$ for some i . We also adopt the convention that if $f_i(x) = -\infty$ for some i ,

$$h(x) = \inf \{F(y) \mid f_i(x) < y_i, i = 1, \dots, n, y = (y_1, \dots, y_n)\}.$$

Under this convention, we have

$$(2.7) \quad \text{epi } h = \{(x, z) \mid \exists y_i \in \mathbf{R} \text{ with } f_i(x) \leq y_i, i = 1, \dots, n, F(y_1, \dots, y_n) \leq z\}$$

because F is isotone. We will use this fact in § 3.

The proofs of our calculus rules will consist of two stages:

Stage 1: Establish an inequality involving upper subderivatives with the help of (1.6), Proposition 2.6, Theorem 2.4 and Proposition 2.5.

Stage 2: Use that inequality and a convex subdifferential calculus formula to establish a corresponding inclusion for the subgradients of (1.7).

As part of stage 1, we will make use of the following lemma, which shows that upper subderivatives preserve isotonicity.

LEMMA 2.8. *Let $F: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be finite at $x_0 \in \mathbf{R}^n$ and isotone on a neighborhood of x_0 . Then $F^\uparrow(x_0; \cdot)$ is isotone on $\mathbf{R}^n$.*

Proof. Let $y_1, y_2 \in \mathbf{R}^n$ with $y_1 \leq y_2$ and $F^\uparrow(x_0; y_2) \leq d$. It suffices to show that $F^\uparrow(x_0; y_1) \leq d$. To this end, let $\varepsilon > 0$ and $U \in \mathcal{N}(0)$ be given. Suppose F is isotone on $X_1 \in \mathcal{N}(x_0)$. There exists $X_2 \in \mathcal{N}(x_0)$ and $\lambda_0 > 0$ such that

$$X_2 + (0, \lambda_0)(y_i + U) \subset X_1 \quad \text{for } i = 1, 2.$$

Since $(y_2, d) \in T_{\text{epi } F}(x_0, F(x_0))$, there exist $X \subset X_2$, $X \in \mathcal{N}(x_0)$, $\delta > 0$, and $\lambda \in (0, \lambda_0)$ such that for all $x \in X$, $r \in (F(x_0) - \delta, F(x_0) + \delta)$ with $F(x) \leq r$, and for all $t \in (0, \lambda)$, there exists $h \in U$ such that

$$\frac{F(x + t(y_2 + h)) - r}{t} \leq d + \varepsilon.$$

Now $x + t(y_1 + h) \leq x + t(y_2 + h)$ and both $x + t(y_1 + h)$ and $x + t(y_2 + h)$ are in X_1 . By isotonicity of F on X_1 ,

$$\frac{F(x + t(y_1 + h)) - r}{t} \leq \frac{F(x + t(y_2 + h)) - r}{t},$$

and hence $(y_1, d) \in T_{\text{epi } F}(x_0, F(x_0))$, as required. $\square$

In order to carry out Stage 2, we will require two convex subdifferential calculus formulae. The first is simply a combination of Theorems 23.8 and 23.9 of [16]. We provide a proof of the second result.

THEOREM 2.9 (cf. [16, Thms. 23.8, 23.9]). *Let $f_1: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$, $f_2: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be proper convex functions, and let $A: \mathbf{R}^m \rightarrow \mathbf{R}^n$ be linear. Let $x_0 \in \text{dom } f_1 \cap A^{-1}(\text{dom } f_2)$. Assume that*

$$(2.8) \quad A(\text{ri dom } f_1) \cap \text{ri dom } f_2 \neq \emptyset.$$

Then

$$(2.9) \quad \partial(f_1 + f_2 \circ A)(x_0) = \partial f_1(x_0) + A^T \partial f_2(Ax_0).$$

THEOREM 2.10. *Let $f_i: \mathbf{R}^{m_i} \rightarrow \bar{\mathbf{R}}$, $i = 1, \dots, n$, be proper convex functions and let $F: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be convex, proper and isotone. Call $f := (f_1, \dots, f_n)$. Assume $x_0 \in \bigcap_{i=1}^n \text{dom } f_i$ and $f(x_0) \in \text{dom } F$. Suppose that*

$$(2.10) \quad \text{int}(\text{dom } F) \cap \text{range } f \neq \emptyset.$$

Then

$$(2.11) \quad \partial(F \circ f)(x_0) = \{\partial(\lambda \cdot f)(x_0) \mid \lambda \in \partial F(f(x_0))\}.$$

If, in addition,

$$(2.12) \quad \bigcap_{i=1}^n \text{ri dom } f_i \neq \emptyset,$$

then

$$(2.13) \quad \partial(F \circ f)(x_0) = \{\lambda \cdot (\partial f_1(x_0), \dots, \partial f_n(x_0)) \mid \lambda \in \partial F(f(x_0))\}.$$

Proof. Let $\lambda \in \partial F(f(x_0))$ and $x^* \in \partial(\lambda \cdot f)(x_0)$. Then $\lambda \geq 0$ and

$$\begin{aligned} x^*(x - x_0) &\leq (\lambda \cdot f)(x) - (\lambda \cdot f)(x_0) \\ &= \lambda \cdot (f(x) - f(x_0)) \\ &\leq F(f(x)) - F(f(x_0)) \\ &= (F \circ f)(x) - (F \circ f)(x_0). \end{aligned}$$

Thus $x^* \in \partial(F \circ f)(x_0)$, so

$$\{\lambda \cdot (\partial f_1(x_0), \dots, \partial f_n(x_0)) \mid \lambda \in \partial F(f(x_0))\} \subset \{\partial(\lambda \cdot f)(x_0) \mid \lambda \in \partial F(f(x_0))\} \subset \partial(F \circ f)(x_0).$$

Conversely, suppose $x_0^* \in \partial(F \circ f)(x_0)$.

Call $h := F \circ f$. Since F is isotone, $h(x) = \min \{F(y) \mid f(x) \leq y\}$ for all $x \in \bigcap_{i=1}^n \text{dom } f_i$. Then $h^*(x_0) := \sup_x x_0^* \cdot x - h(x) = \sup_{x,y} \{x_0^* \cdot x - F(y) \mid y - f(x) \geq 0\}$. Assumption (2.10) guarantees the existence of a Slater point for the above concave program, so we may apply the Lagrange multiplier theorem of [16, § 28]: There exists $\lambda \geq 0$ such that

$$\begin{aligned} h^*(x_0^*) &= \sup_{x,y} x_0^* \cdot x - F(y) + \lambda \cdot (y - f(x)) \\ &= F^*(\lambda) + (\lambda \cdot f)^*(x_0^*). \end{aligned}$$

Then

$$\begin{aligned} x_0^* \cdot x &= h(x_0) + h^*(x_0^*) \\ &= F(f(x_0)) + (\lambda \cdot f)^*(x_0^*) + F^*(\lambda), \end{aligned}$$

and so

$$x_0^* \cdot x_0 + \lambda \cdot f(x_0) = \lambda \cdot f(x_0) + (\lambda \cdot f)^*(x_0^*) + F(f(x_0)) + F^*(\lambda).$$

We now have

$$x_0^* \cdot x_0 = \lambda \cdot f(x_0) + (\lambda \cdot f)^*(x_0^*) \quad \text{and} \quad \lambda \cdot f(x_0) = F(f(x_0)) + F^*(\lambda),$$

implying that

$$\lambda \in \partial F(f(x_0)) \quad \text{and} \quad x_0^* \in \partial(\lambda \cdot f)(x_0).$$

Thus $\partial(F \circ f)(x_0) \subset \{\partial(\lambda \cdot f)(x_0) \mid \lambda \in \partial F(f(x_0))\}$. If in addition (2.12) holds, Theorem 23.8 of [16] gives $x_0^* \in \partial(\lambda \cdot f)(x_0) = \lambda \cdot (\partial f_1(x_0), \dots, \partial f_n(x_0))$, so that (2.13) holds. $\square$

Remark 2.11. Since $\text{int}(\text{dom } F) \neq \emptyset$ (because F is proper and isotone), (2.10) is equivalent to the seemingly weaker condition

$$\text{ri}(\text{dom } F) \cap \text{ri}(\text{range } f) \neq \emptyset.$$

Remark 2.12. If $\lambda = 0$ in Theorem 2.10, λf_i should be interpreted as the indicator function of $\text{dom } f_i$ (Definition 3.3). This is in keeping with the convention that $(F \circ f)(x) = +\infty$ whenever some $f_i(x) = +\infty$. As a result, $\partial f_i(x_0)$ should be interpreted as $\partial^\infty f_i(x_0)$ in (2.13) and (3.27) (see Definition 3.13).

3. The main theorems and their corollaries. In Stage 1 of the proof of our first calculus formula, we will need a technical lemma verifying that condition (2.5) is satisfied for the appropriate A , C , and z_0 . In the proof of this lemma, for $x = (x_1, \dots, x_p) \in \mathbf{R}^p$ and $\varepsilon > 0$, we use the notation

$$B_\varepsilon(x) := \{y = (y_1, \dots, y_p) \in \mathbf{R}^p : |y_i - x_i| \leq \varepsilon, i = 1, \dots, p\}.$$

LEMMA 3.1. Let $f_1: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$, $f_2: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ be l.s.c., and let $F: \mathbf{R}^n \rightarrow \mathbf{R}^m$ be continuous at $x_0 \in \mathbf{R}^n$. Assume $f_1(x_0)$ and $f_2(F(x_0))$ are finite. Define $A: \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^m \times \mathbf{R} \rightarrow \mathbf{R}^n \times \mathbf{R}$ by $A(x, y, z, r) := (x, y + r)$ and $G: \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^m \times \mathbf{R} \rightarrow \mathbf{R}^m$ by $G(x, y, z, r) := F(x) - z$.

Then (2.5) is satisfied with A as above, $C := (\text{epi } f_1 \times \text{epi } f_2) \cap G^{-1}(0)$, and $z_0 := (x_0, f_1(x_0), F(x_0), f_2(F(x_0)))$.

Proof. Let $\varepsilon > 0$ be given, and let $X := B_\varepsilon(z_0)$. Since f_1, f_2 are l.s.c. and F is continuous, there exists $\delta \in (0, \varepsilon/3)$ such that for all $x \in B_\delta(x_0)$, we have

$$\begin{aligned} f_1(x) &\geq f_1(x_0) - \varepsilon/3, \\ F(x) &\in B_{\varepsilon/3}(F(x_0)), \\ (f_2 \circ F)(x) &\geq (f_2 \circ F)(x_0) - \varepsilon/3. \end{aligned}$$

Define $N := B_\delta(z_0)$. Since A is surjective, $Z := A(N) \in \mathcal{N}(A z_0)$. We will now verify that $Z \cap A(C) \subset A(X \cap C)$. To do so, let $(\bar{x}, \bar{r}) \in Z \cap A(C)$. Since $(\bar{x}, \bar{r}) \in Z$, there exists $(x, y, z, r) \in N$ with $x = \bar{x}$, $y + r = \bar{r}$. Since $(\bar{x}, \bar{r}) \in A(C)$, there exists $(x', y', z', r') \in C$ with $x' = \bar{x}$, $y' + r' = \bar{r}$, and $f_1(x') \leq y'$, $z' = F(x')$, $f_2(z') \leq r'$. Now $x' \in B_\delta(x_0)$, so $z' \in B_{\varepsilon/3}(F(x_0))$. Also $\bar{r} \in B_{\varepsilon/3}(f_1(x_0)) + B_{\varepsilon/3}(f_2(F(x_0)))$, so $\bar{r} \in B_{2\varepsilon/3}(f_1(x_0) + f_2(F(x_0)))$. Finally, since $y' + r' = \bar{r}$ and $y' \geq f_1(x_0) - \varepsilon/3$, we conclude that $y' \in B_\varepsilon(f_1(x_0))$ and $r' \in B_\varepsilon(f_2(F(x_0)))$. Thus $(x', y', z', r') \in X$, and so $(\bar{x}, \bar{r}) \in A(X \cap C)$. We conclude that $Z \cap A(C) \subset A(X \cap C)$. $\square$

We now proceed to the first of our two main theorems.

THEOREM 3.2. *Let $F: \mathbf{R}^n \rightarrow \mathbf{R}^m$ be strictly differentiable at x_0 , $f_1: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ finite and strictly l.s.c. at x_0 , and $f_2: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ finite and strictly l.s.c. at $F(x_0)$. Assume that*

$$(3.1) \quad \nabla F(x_0) \text{ dom } f_1^\dagger(x_0; \cdot) - \text{dom } f_2^\dagger(x_0; \cdot) = \mathbf{R}^m.$$

Then for all $y \in \mathbf{R}^n$,

$$(3.2) \quad (f_1 + f_2 \circ F)^\dagger(x_0; y) \leq f_1^\dagger(x_0; y) + f_2^\dagger(F(x_0); \nabla F(x_0)y).$$

Moreover,

$$(3.3) \quad \partial(f_1 + f_2 \circ F)(x_0) \subset \partial f_1(x_0) + (\nabla F(x_0))^T \partial f_2(F(x_0)).$$

Proof. Call $f := f_1 + f_2 \circ F$. Then

$$\begin{aligned} \text{epi } f &= \{(x_1, r) \in \mathbf{R}^n \times \mathbf{R} \mid f_1(x_1) \leq r_1, f_2(x_2) \leq r_2, r = r_1 + r_2, \\ &\quad F(x_1) - x_2 = 0, \text{ for some } x_2 \in \mathbf{R}^m, r_1, r_2 \in \mathbf{R}\}. \end{aligned}$$

Define A , G , and C , and z_0 as in Lemma 3.1, and define $D := \text{epi } f_1 \times \text{epi } f_2$. Then $A(D \cap G^{-1}(0)) = \text{epi } f$.

Now

$$\begin{aligned} \text{epi } f^\dagger(x_0; \cdot) &= T_{A(D \cap G^{-1}(0))}(x_0, f(x_0)) \quad \text{by (1.6)} \\ &\supset A(T_{D \cap G^{-1}(0)}(z_0)) \end{aligned}$$

by Lemma 3.1 and Proposition 2.6. Next observe that (3.1) and Proposition 2.5 ensure that $\nabla G(z_0)T_D(z_0) = \mathbf{R}^m$. We can therefore apply Theorem 2.4 to obtain

$$T_{D \cap G^{-1}(0)}(z_0) \supset T_D(z_0) \cap \nabla G(z_0)^{-1}(0).$$

Thus

$$\begin{aligned} A(T_{D \cap G^{-1}(0)}(z_0)) &\supset A(T_D(z_0) \cap \nabla G(z_0)^{-1}(0)) \\ &= A((\text{epi } f_1^\dagger(x_0; \cdot) \times \text{epi } f_2^\dagger(F(x_0); \cdot)) \\ &\quad \cap \nabla G(z_0)^{-1}(0)) \quad (\text{by Proposition 2.5}) \\ &= A(\{(h_1, r_1, h_2, r_2) \in \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^m \times \mathbf{R} \mid f_1^\dagger(x_0; h_1) \\ &\quad \leq r_1, f_2^\dagger(F(x_0); h_2) \leq r_2, \nabla F(x_0)h_1 = h_2\}) \\ &= \{(h, r_1 + r_2) \in \mathbf{R}^n \times \mathbf{R} \mid f_1^\dagger(x_0; h) \leq r_1, f_2^\dagger(F(x_0); \nabla F(x_0)h) \leq r_2\} \\ &= \text{epi } [f_1^\dagger(x_0; \cdot) + f_2^\dagger(F(x_0); \nabla F(x_0)(\cdot))]. \end{aligned}$$

Therefore $\text{epi } f^\uparrow(x_0; \cdot) \supset \text{epi } [f_1^\uparrow(x_0; \cdot) + f_2^\uparrow(F(x_0)(\cdot))]$, and so (3.2) holds. The rest of the proof is much as in [18]. Set $p_1 := f_1^\uparrow(x_0; \cdot)$ and $p_2 := f_2^\uparrow(F(x_0); \cdot)$. If either $p_1(0)$ or $p_2(0)$ is $-\infty$, equality holds in (3.3), since (3.2) shows that both sides of the inclusion are then empty. Assume, then, that $p_1(0) = p_2(0) = 0$.

Then

$$\begin{aligned} \partial f(x_0) &= \{z \in \mathbf{R}^n \mid (f_1 + f_2 \circ F)^\uparrow(x_0; y) \geq \langle y, z \rangle \forall y \in \mathbf{R}^n\} \\ &\subset \{z \mid p_1(y) + p_2(\nabla F(x_0)y) \geq \langle y, z \rangle \forall y \in \mathbf{R}^n\} \\ &= \partial(p_1 + p_2 \circ \nabla F(x_0))(0). \end{aligned}$$

Since p_1 and p_2 are convex and proper, and since (3.1) holds, we may apply Theorem 2.9 to obtain

$$\begin{aligned} \partial(p_1 + p_2 \circ \nabla F(x_0))(0) &= \partial p_1(0) + \nabla F(x_0)^T \partial p_2(0) \\ &= \partial f_1(x_0) + \nabla F(x_0)^T \partial f_2(F(x_0)) \end{aligned}$$

by definition, and so (3.3) holds. $\square$

We can now obtain improved versions, in the finite-dimensional case, of results in [17] and [18].

DEFINITION 3.3. Let $C \subset E$. The indicator function of C , denoted i_C , is defined by

$$i_C(x) = \begin{cases} 0 & \text{if } x \in C, \\ +\infty & \text{otherwise.} \end{cases}$$

Observe that for $x_0 \in C$, $i_C^\uparrow(x_0; \cdot) = i_{T_C(x_0)}(\cdot)$ and $\partial i_C(x_0) = N_C(x_0)$. The following is a strengthening of Theorem 5 of [17], which has also been proved by Aubin [2], [3].

COROLLARY 3.4. Let $F_2: \mathbf{R}^m \rightarrow \mathbf{R}^m$ be strictly differentiable at x_0 , and let $C_1 \subset \mathbf{R}^n$ be closed near x_0 and $C_2 \subset \mathbf{R}^m$ closed near $F(x_0)$. Suppose that

$$(3.4) \quad \nabla F(x_0) T_{C_1}(x_0) - T_{C_2}(F(x_0)) = \mathbf{R}^m.$$

Then

$$(3.5) \quad T_{C_1 \cap F^{-1}(C_2)}(x_0) \supset T_{C_1}(x_0) \cap \nabla F(x_0)^{-1} T_{C_2}(F(x_0))$$

and

$$(3.6) \quad N_{C_1 \cap F^{-1}(C_2)}(x_0) \subset N_{C_1}(x_0) + \nabla F(x_0)^T N_{C_2}(F(x_0)).$$

Proof. Let $f_1 := i_{C_1}$, $f_2 := i_{C_2}$ in Theorem 3.2. Then (3.1) becomes (3.4), and (3.5) and (3.6) follow from (3.2) and (3.3), respectively. $\square$

It is interesting to note that Theorem 2.4, which is used in proving Corollary 3.4, is itself a special case of Corollary 3.4.

Remark 3.5. If $F = A$ is linear, (3.4) can be weakened to

$$(3.7) \quad A(T_{C_1}(x_0)) - T_{C_2}(Ax_0) = \text{span}(AC_1 - C_2).$$

To see this, suppose $C \subset \mathbf{R}^p$, $x_0 \in C$. For a given affine set S with $\text{aff } C \subset S \subset \mathbf{R}^p$, denote by $T_C^S(x_0)$ the Clarke tangent cone of C at x_0 where C is considered as a subset of S rather than as a subset of $\mathbf{R}^p$. It is easy to see that $T_C^S(x_0) = T_C(x_0)$. Without loss of generality, we may assume $x_0 = 0$, since $T_{C-x_0}(0) = T_C(x_0)$ and $x_0 \in C_1 \cap A^{-1}(C_2)$ if and only if $0 \in (C_1 - x_0) \cap A^{-1}(C_2 - Ax_0)$. Call $V := \text{span } C_1$ and $W := \text{span}(AC_1 - C_2)$. If (3.7) holds, then $A(T_{C_1}(x_0)) - T_{C_2}^W(Ax_0) = W$ and A may be considered as $A: V \rightarrow W$. By Corollary 3.4, $T_{C_1 \cap A^{-1}(C_2)}^V(x_0) \supset T_{C_1}^V(x_0) \cap A^{-1} T_{C_2}^W(Ax_0)$, which is the same as (3.5).

COROLLARY 3.6. Let $f_1: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}, f_2: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be finite and strictly l.s.c. at x_0 . Assume that

$$(3.8) \quad \text{dom } f_1^\dagger(x_0; \cdot) - \text{dom } f_2^\dagger(x_0; \cdot) = \mathbf{R}^n.$$

Then for all $y \in \mathbf{R}^n$,

$$(3.9) \quad (f_1 + f_2)^\dagger(x_0; y) \leq f_1^\dagger(x_0; y) + f_2^\dagger(x_0; y).$$

In addition,

$$(3.10) \quad \partial(f_1 + f_2)(x_0) \subset \partial f_1(x_0) + \partial f_2(x_0).$$

Proof. Set $m = n$ and $F := I$ in Theorem 3.2. $\square$

Remark 3.7. (a) Corollary 3.6 is more general than the specialization of Theorem 2 of [18] to strictly l.s.c. functions with finite-dimensional domains, since Rockafellar's assumption

$$(3.11) \quad \text{dom } f_1^\dagger(x_0; \cdot) \cap \text{int dom } f_2^\dagger(x_0; \cdot) \neq \emptyset$$

implies, but is not implied by (3.8). Here is an example which satisfies (3.8) but not (3.11): Define $f_1: \mathbf{R}^2 \rightarrow \bar{\mathbf{R}}, f_2: \mathbf{R}^2 \rightarrow \bar{\mathbf{R}}$ by $f_1(x, y) := |x|^{1/2}, f_2(x, y) := |y|^{1/2}$. Then $\text{dom } f_1^\dagger(0; \cdot) = 0 \times \mathbf{R}$ and $\text{dom } f_2^\dagger(0; \cdot) = \mathbf{R} \times 0$, both of which have empty interior.

(b) Condition (3.8) can actually be weakened to

$$(3.12) \quad \text{dom } f_1^\dagger(x_0; \cdot) - \text{dom } f_2^\dagger(x_0; \cdot) = \text{span}(\text{dom } f_1 - \text{dom } f_2).$$

To see this, simply observe that we can again assume that $x_0 = 0$, and we can then replace $\mathbf{R}^n$ by $S := \text{span}(\text{dom } f_1 - \text{dom } f_2)$, considering f_1 and f_2 as $f_1: S \rightarrow \bar{\mathbf{R}}$ and $f_2: S \rightarrow \bar{\mathbf{R}}$.

It is not possible, however, to weaken the hypothesis still further to

$$(3.13) \quad \text{ri dom } f_1^\dagger(x_0; \cdot) \cap \text{ri dom } f_2^\dagger(x_0; \cdot) \neq \emptyset.$$

For example, let $C_1 := \{(x, y) \mid y = x^2\}, C_2 := \{(x, y) \mid y = -x^2\}$, and $f_1 := i_{C_1}, f_2 := i_{C_2}$. Let $x_0 = (0, 0)$. Then $f_1^\dagger(x_0; \cdot) = f_2^\dagger(x_0; \cdot) = i_{\mathbf{R} \times 0}(\cdot)$, while

$$(f_1 + f_2)^\dagger(x_0; \cdot) = i_{T_{C_1 \cap C_2}(x_0)}(\cdot) = i_{\{(0,0)\}}(\cdot).$$

Hence (3.9) does not hold for $y = (x, 0)$ with $x \neq 0$, although (3.13) is satisfied.

(c) If f_1 and f_2 are not strictly l.s.c., then (3.8) may hold without either (3.9) or (3.10) being satisfied. For example, let $C_1 := \mathbf{Q}$, the set of rational numbers, and let $C_2 = (\mathbf{R}/\mathbf{Q}) \cup \{0\}$. Define $f_1 := i_{C_1}, f_2 := i_{C_2}$, and let $x_0 = 0$. Then $T_{C_1}(0) = T_{C_2}(0) = \mathbf{R}$, so (3.8) holds, and $f_1^\dagger(x_0; \cdot) = f_2^\dagger(x_0; \cdot) = i_{\mathbf{R}}(\cdot)$. However, $C_1 \cap C_2 = \{0\}$, so we have $(f_1 + f_2)^\dagger(x_0; \cdot) = i_{\{0\}}(\cdot)$, and (3.9) does not hold for $y \neq 0$.

DEFINITION 3.8. Let $C \subset \mathbf{R}^p$. Define

$$\Delta^n C := \{(x_1, \dots, x_n) \mid x_i \in C, x_1 = x_2 = \dots = x_n\}.$$

DEFINITION 3.9. Let $C_i \subset \mathbf{R}^m, i = 1, \dots, n$ be convex sets. The sets $C_i, i = 1, \dots, n$, are said to be in *strong general position* [27] if

$$(3.14) \quad 0 \in \text{int} \left[\Delta^{n-1} C_1 - \prod_{j=2}^n C_j \right].$$

If the sets are cones, (3.14) is equivalent to

$$(3.15) \quad \Delta^{n-1} C_1 - \prod_{j=2}^n C_j = \mathbf{R}^{(n-1)m}.$$

Another equivalent way to write (3.14) is

$$(3.16) \quad 0 \in \text{int} \left[\Delta^n \mathbf{R}^m - \prod_{j=1}^n C_j \right].$$

(see [27] for a thorough discussion of this concept.) By (3.16), the order in which the sets in (3.14) are listed does not matter.

The following is an important special case of Corollary 3.4:

PROPOSITION 3.10 ([26], [3]). *Let $D_i \subset \mathbf{R}^m$, $i = 1, \dots, n$ be closed near $y_0 \in \bigcap_{i=1}^n D_i$. Assume $T_{D_i}(y_0)$, $i = 1, \dots, n$ are in strong general position. Then*

$$(3.17) \quad T_{D_1 \cap \dots \cap D_n}(y_0) \supset \bigcap_{i=1}^n T_{D_i}(y_0)$$

and

$$(3.18) \quad N_{D_1 \cap \dots \cap D_n}(y_0) \subset \sum_{i=1}^n N_{D_i}(y_0).$$

Proof. Call $C_1 := D_1 \times \dots \times D_n$, and let C_2 be the origin in $\mathbf{R}^{(n-1)m}$. Define $F: \mathbf{R}^{nm} \rightarrow \mathbf{R}^{(n-1)m}$ by $F(x_1, \dots, x_n) := (x_1 - x_2, \dots, x_1 - x_n)$. Apply Corollary 3.4 with $x_0 := (y_0, \dots, y_0)$. By Proposition 2.5, (3.4) reduces to $T_{D_i}(y_0)$, $i = 1, \dots, n$, being in strong general position. By (3.5),

$$T_{\Delta^n(D_1 \cap \dots \cap D_n)}(y_0, \dots, y_0) \supset (T_{D_1}(y_0) \times \dots \times T_{D_n}(y_0)) \cap \Delta^n \mathbf{R}^m, \quad \text{or}$$

$$\Delta^n T_{D_1 \cap \dots \cap D_n}(y_0) \supset \Delta^n \bigcap_{i=1}^n T_{D_i}(y_0).$$

Thus (3.17) holds. Finally, (3.18) follows from (3.6). $\square$

Remark 3.11. By (3.7), the strong general position assumption can be weakened to

$$(3.19) \quad \Delta^{n-1} T_{D_1}(y_0) - \prod_{i=2}^n T_{D_i}(y_0) = \prod_{j=2}^n \text{span}(D_1 - D_j).$$

We can derive from Proposition 3.10 a formula for the subgradient of $f(x) := \max_{1 \leq i \leq n} f_i(x)$, where $f_i: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ are strictly l.s.c. We start with a lemma which we will use to show that the condition “ $\text{dom } f_i^\uparrow(x_0; \cdot)$, $i = 1, \dots, n$, are in strong general position” is sufficient to guarantee this subgradient formula.

LEMMA 3.12. *Suppose $f_i: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$, $i = 1, \dots, n$ are such that*

$$\Delta^{n-1} \text{dom } f_1 - \prod_{i=2}^n \text{dom } f_i = \mathbf{R}^{m(n-1)}.$$

Then

$$\Delta^{n-1} \text{epi } f_1 - \prod_{i=2}^n \text{epi } f_i = (\mathbf{R}^m \times \mathbf{R})^{n-1}.$$

Proof. We must show $(\mathbf{R}^m \times \mathbf{R})^{n-1} \subset \Delta^{n-1} \text{epi } f_1 - \prod_{i=2}^n \text{epi } f_i$. Let $(y_1, s_1, \dots, y_{n-1}, s_{n-1}) \in (\mathbf{R}^m \times \mathbf{R})^{n-1}$. Then there exist $x_i \in \text{dom } f_i$, $i = 1, \dots, n$ such that $x_1 - x_i = y_{i-1}$, $i = 2, \dots, n$. Choose r_i , $i = 1, \dots, n$, such that $r_i \geq f_i(x_i)$. If $r_1 - r_2 < s_1$, replace r_1 by $s_1 + r_2$. If $r_1 - r_2 > s_1$, replace r_2 by $r_1 - s_1$. Now if $r_1 - r_3 > s_2$, replace r_3 by $r_1 - s_2$. If $r_1 - r_3 < s_2$, replace r_1 by $s_2 + r_3$ and r_2 by $s_2 + r_3 - s_1$. Proceed in this manner. After $k-1$ steps, we have $r_1 - r_i = s_{i-1}$, $i = 2, \dots, k$. If $r_1 - r_{k+1} > s_k$, replace r_{k+1} by $r_1 - s_k$. If $r_1 - r_{k+1} < s_k$, replace r_1 by $s_k + r_{k+1}$ and replace r_j , $j = 2, \dots, k$ by $s_k + r_{k+1} - s_{j-1}$.

After $n-1$ steps, we obtain $r_1 - r_i = s_{i-1}$, $i = 2, \dots, n$, and $(x_i, r_i) \in \text{epi } f_1$. Thus $(y_i, s_j) \in \Delta^{n-1} \text{epi } f_1 - \prod_{i=2}^n \text{epi } f_i$ and so $(\mathbf{R}^m \times \mathbf{R})^{n-1} \subset \Delta^{n-1} \text{epi } f_1 - \prod_{i=2}^n \text{epi } f_i$. $\square$

DEFINITION 3.13. Let $f: E \rightarrow \bar{\mathbf{R}}$ be finite at $x \in E$. The *asymptotic generalized gradient* of f at x is the set

$$\partial^\infty f(x) = \{z \in E^* \mid (z, 0) \in N_{\text{epi } f}(x, f(x))\}.$$

In our next result, we will use the fact that if $\partial f(x) \neq \emptyset$, then

$$(3.20) \quad N_{\text{epi } f}(x, f(x)) = \bigcup_{\lambda > 0} \lambda(\partial f(x), -1) \cup (\partial^\infty f(x), 0)$$

([18], [6]).

PROPOSITION 3.14. Let $f_i: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$, $i = 1, \dots, n$ be strictly l.s.c. and finite at x_0 . Define $f(x) := \max_{1 \leq i \leq n} f_i(x)$ and $I(x) := \{i \in \{1, \dots, n\} \mid f_i(x) = f(x)\}$. Suppose that f_i is continuous at x_0 for each $i \notin I(x_0)$, and suppose that $\text{dom } f_i^\uparrow(x_0; \cdot)$, $i = 1, \dots, n$, are in strong general position. Then for all $y \in \mathbf{R}^n$,

$$(3.21) \quad f^\uparrow(x_0; y) \leq \max_{I(x_0)} f_i^\uparrow(x_0; y).$$

If also $\partial f_i(x_0)$, $i \in I(x_0)$, are nonempty, then

$$(3.22) \quad \partial f(x_0) \subset \sum_{I(x_0)} (\lambda_i \partial f_i(x_0) \cup \partial^\infty f_i(x_0))$$

for some $\lambda_i \geq 0$ with $\sum_{I(x_0)} \lambda_i = 1$.

Proof. Call $D_i := \text{epi } f_i$, $i = 1, \dots, n$. Since f_j , $j \in I(x_0)$, are strictly l.s.c. and f_j , $j \notin I(x_0)$, are continuous at x_0 , $T_{D_j}(x_0, f(x_0)) = \mathbf{R}^{m+1}$ for all $j \notin I(x_0)$. Thus we only need to consider D_j with $j \in I(x_0)$. By Lemma 3.12, our assumption that $\text{dom } f_i^\uparrow(x_0; \cdot)$, $i = 1, \dots, n$, are in strong general position implies that $T_{D_i}(x_0, f(x_0))$, $i \in I(x_0)$, are also in strong general position. We may then apply Proposition 3.10. By (3.17),

$$\text{epi } f^\uparrow(x_0; \cdot) \supset \text{epi } \max_{i \in I(x_0)} f_i^\uparrow(x_0; \cdot),$$

which gives (3.21). If also $\partial f_i(x_0)$, $i \in I(x_0)$, are nonempty, we have by (3.18) and (3.20) that

$$\bigcup_{\lambda > 0} \lambda(\partial f(x_0), -1) \subset \sum_{i \in I(x_0)} \left(\bigcup_{\lambda_i > 0} \lambda_i(\partial f_i(x_0), -1) \cup (\partial^\infty f_i(x_0), 0) \right).$$

To obtain (3.22), set $\lambda = 1$ on the left-hand side of the above inclusion. $\square$

In the proof of our second main result, we will use another technical lemma to ensure that we may apply Proposition 2.6.

LEMMA 3.15. Let $f_i: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$, $i = 1, \dots, n$ be finite and strictly l.s.c. at x_0 , and let $F: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be l.s.c. Call $f := (f_1, \dots, f_n)$. Assume $F(f(x_0))$ is finite and F is isotone on $B_{t_0}(f(x_0)) + \mathbf{R}_+^n$ for some $t_0 > 0$. Define

$$A: \mathbf{R}^{nm+2n+1} \rightarrow \mathbf{R}^{m+1} \quad \text{by}$$

$$A(x_1, y_1, \dots, x_n, y_n, z_1, \dots, z_n, r) := (x_1, r).$$

(Here $x_i \in \mathbf{R}^m$, $y_i, z_i \in \mathbf{R}$.) Define

$$G: \mathbf{R}^{nm+2n+1} \rightarrow \mathbf{R}^{m(n-1)+n} \quad \text{by}$$

$$G(x_1, y_1, \dots, x_n, y_n, z_1, \dots, z_n, r) = (x_1 - x_2, \dots, x_1 - x_n, y_1 - z_1, \dots, y_n - z_n).$$

Assume for each $j \in \{1, \dots, n\}$ that either

- (1) f_j is continuous at x_0 , or
- (2) F is strictly isotone in the j th coordinate at $f(x_0)$.

Then (2.5) is satisfied with A as above,

$$C := (\text{epi } f_1 \times \dots \times \text{epi } f_n \times \text{epi } F) \cap G^{-1}(0), \text{ and}$$

$$z_0 := (x_0, f_1(x_0), \dots, x_0, f_n(x_0), f_1(x_0), \dots, f_n(x_0), F(f(x_0))).$$

Proof. For a given $\varepsilon > 0$, let $X := B_\varepsilon(z_0)$. Let $I \subset \{1, \dots, n\}$ be the set of coordinates in which F is strictly isotone at $f(x_0)$. For $t > 0$, define $U_t := \{(y_1, \dots, y_n) \in \mathbb{R}^n \mid y_i \geq f_i(x_0) - t, i = 1, \dots, n\}$. Suppose $i \in I$. We claim that there exists $\mu_i \in (0, t_0)$ such that

$$F(f_1(x_0) - \mu_i, \dots, f_i(x_0) + \varepsilon, \dots, f_n(x_0) - \mu_i) > F(f(x_0)).$$

If not, then for all $t \in (0, t_0)$,

$$F(f_1(x_0) - t, \dots, f_i(x_0) + \varepsilon, \dots, f_n(x_0) - t) \leq F(f(x_0)).$$

But since F is l.s.c., it follows that

$$F(f_1(x_0), \dots, f_i(x_0) + \varepsilon, \dots, f_n(x_0)) \leq F(f(x_0)).$$

This contradicts our assumption that $i \in I$. So for each $i \in I$, we can choose $\mu_i \in (0, t_0)$ with

$$\delta_i := F(f_1(x_0) - \mu_i, \dots, f_i(x_0) + \varepsilon, \dots, f_n(x_0) - \mu_i) - F(f(x_0)) > 0.$$

Let $\mu = \min_I \mu_i$, and let $\delta = \frac{1}{2} \min_I \delta_i$. Then if $r \in B_\delta(F(f(x_0))) \cap F(U_\mu)$, it follows from the isotonicity of F that $r = F(y)$ for some $y = (y_1, \dots, y_n)$ with $y_i \leq f_i(x_0) + \varepsilon$ for all $i \in I$.

Now by assumptions (1) and (2), there exists $\delta_0 \in (0, \min(\mu, \delta))$ such that $f_j(x) \in B_\mu(f_j(x_0))$ whenever $x \in B_{\delta_0}(x_0)$, $j \notin I$ and $f_j(x) \geq f_j(x_0) - \mu$ whenever $x \in B_{\delta_0}(x_0)$, $j \in I$. Let $N = B_{\delta_0}(z_0)$. Again $Z := A(N) \in \mathcal{N}(Az_0)$. We will now show that $Z \cap A(C) \subset A(X \cap C)$. Suppose $(\bar{x}, \bar{r}) \in Z \cap A(C)$. Since $(\bar{x}, \bar{r}) \in A(N)$, $\bar{x}$ must be in $B_{\delta_0}(x_0)$, and $\bar{r}$ in $B_{\delta_0}(F(f(x_0)))$. Since $(\bar{x}, \bar{r}) \in A(C)$, there exists $y \in \mathbb{R}^n$ with $f(\bar{x}) \leq y$ and $F(y) \leq \bar{r}$. Thus $\bar{y} = f(\bar{x})$ satisfies $F(\bar{y}) \leq \bar{r}$, $\bar{y}_i \in B_\mu(f_j(x_0))$ for all $j \notin I$, and $f_j(x_0) - \mu \leq \bar{y}_j \leq f_j(x_0) + \varepsilon$ for all $j \in I$. We conclude that $(\bar{x}, \bar{y}_1, \dots, \bar{x}, \bar{y}_n, \bar{y}_1, \dots, \bar{y}_n, \bar{r}) \in X \cap C$, and so $(\bar{x}, \bar{r}) \in A(X \cap C)$. $\square$

Remark 3.16. The hypotheses of Lemma 3.15 hold in the following important cases:

- (a) $F(y_1, \dots, y_n) := \sum_{i=1}^n y_i$ and each f_n strictly l.s.c. at x_0 .
- (b) $F(y_1, \dots, y_n) := \prod_{i=1}^n y_i$ and each f_j positive and strictly l.s.c. at x_0 .
- (c) $F(y_1, \dots, y_n) := \max_{1 \leq j \leq n} y_j$ and each f_j strictly l.s.c. at x_0 , with f_j continuous at x_0 for all $j \notin I(x_0)$, where $I(x_0) := \{i \in \{1, \dots, n\} \mid f_i(x_0) = \max_{1 \leq j \leq n} f_j(x_0)\}$.

THEOREM 3.17. Let $f_i: \mathbb{R}^m \rightarrow \bar{\mathbb{R}}$, $i = 1, \dots, n$ be finite and strictly l.s.c. at x_0 , and define $f := (f_1, \dots, f_n)$. Let $F: \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be finite at $f(x_0)$, isotone on the union of some neighborhood of $f(x_0)$ and $\text{Range } f + \mathbb{R}_+^n$, and l.s.c. Assume that for each $j \in \{1, \dots, n\}$, either

$$(3.23) \quad f_j \text{ is continuous at } x_0, \text{ or}$$

$$(3.24) \quad F \text{ is strictly isotone in the } j\text{th coordinate at } f(x_0).$$

Assume also that

$$(3.25) \quad (\Delta^n \mathbb{R}^n \times \text{dom } F^\uparrow(f(x_0); \cdot)) - S = \mathbb{R}^{nm+n}$$

where

$$S := \{(y_1, \dots, y_n, r_1, \dots, r_n) \mid (y_i, r_i) \in \text{epi } f_i^\uparrow(x_0; \cdot), i = 1, \dots, n\}.$$

Then for all $y \in \mathbf{R}^m$,

$$(3.26) \quad (F \circ f)^\uparrow(x_0; y) \leq F^\uparrow(f(x_0); f_1^\uparrow(x_0; y), \dots, f_n^\uparrow(x_0; y)).$$

If in addition each $f_i^\uparrow(x_0; \cdot)$ is proper, then

$$(3.27) \quad \partial(F \circ f)(x_0) \subset \{\lambda \cdot (\partial f_1(x_0), \dots, \partial f_n(x_0)) \mid \lambda \in \partial F(f(x_0))\}.$$

Proof. Call $h := F \circ f$. Since F is isotone on $\text{Range } f + \mathbf{R}_+^n$,

$$\text{epi } h = \{(x, r) \in \mathbf{R}^m \times \mathbf{R}, \exists (y_1, \dots, y_n) \in \mathbf{R}^n \text{ with } F(y_1, \dots, y_n) \leq r, f_i(x) \leq y_i, 1 \leq i \leq n\}.$$

Define the functions G and A as in Lemma 3.15, and define

$$D := \text{epi } f_1 \times \dots \times \text{epi } f_n \times \text{epi } F,$$

$$C := D \cap G^{-1}(0).$$

Then $\text{epi } h = A(C)$. Assumptions (3.23) and (3.24) allow us to apply Lemma 3.15 and Proposition 2.6. Thus

$$\text{epi } h^\uparrow(x_0; \cdot) = T_{A(C)}(x_0, h(x_0)) \quad (\text{by (1.6)})$$

$$\supset A(T_C(z_0)) \quad \text{where } z_0$$

$$:= (x_0, f_1(x_0), \dots, x_0, f_n(x_0), f_1(x_0), \dots, f_n(x_0), h(x_0))$$

(by Lemma 3.15 and Proposition 2.6)

$$\supset A(T_D(z_0) \cap \nabla G(z_0)^{-1}(0))$$

by Theorem 2.4, since (3.25) says exactly that

$$\nabla G(z_0) T_D(z_0) = \mathbf{R}^{(n-1)m+n},$$

$$= \{(x, r) \in \mathbf{R}^m \times \mathbf{R} \mid \exists y \in \mathbf{R}^n \text{ with } f_i^\uparrow(x_0; x) \leq y_i, F^\uparrow(f(x_0); y) \leq r\}$$

$$= \{(x, r) \mid F^\uparrow(f(x_0); f_1^\uparrow(x_0; x), \dots, f_n^\uparrow(x_0; x)) \leq r\}$$

since $F^\uparrow(f(x_0); \cdot)$ is itself isotone (Lemma 2.8). Thus $\text{epi } h^\uparrow(x_0; \cdot) \supset \text{epi } F^\uparrow(f(x_0); f_1^\uparrow(x_0; \cdot), \dots, f_n^\uparrow(x_0; \cdot))$, and so (3.26) holds.

Now if $F^\uparrow(f(x_0); 0) = -\infty$, (3.26) shows that both sides of (3.27) are empty. Assume, then, that $F^\uparrow(f(x_0); \cdot)$ is proper. Since each $f_i^\uparrow(x_0; \cdot)$ is convex and proper and (3.25) implies that (2.10) and (2.12) hold for $F^\uparrow(f(x_0); \cdot)$ and $(f_1^\uparrow(x_0; \cdot), \dots, f_n^\uparrow(x_0; \cdot))$, we may apply Theorem 2.10. We have

$$\begin{aligned} \partial(F \circ f)(x_0) &= \{z \in \mathbf{R}^m \mid (F \circ f)^\uparrow(x_0; y) \geq \langle y, z \rangle \forall y \in \mathbf{R}^m\} \\ &\subset \{z \in \mathbf{R}^m \mid F^\uparrow(f(x_0); f_1^\uparrow(x_0; y), \dots, f_n^\uparrow(x_0; y)) \geq \langle y, z \rangle \forall y \in \mathbf{R}^m\} \\ &= \partial((F^\uparrow(f(x_0); \cdot)) \circ (f_1^\uparrow(x_0; \cdot), \dots, f_n^\uparrow(x_0; \cdot)))(0) \\ &= \{\lambda \cdot (\partial f_1^\uparrow(x_0; \cdot)(0), \dots, \partial f_n^\uparrow(x_0; \cdot)(0)) \mid \lambda \in \partial F^\uparrow(f(x_0); \cdot)(0)\} \\ &= \{\lambda \cdot (\partial f_1(x_0), \dots, \partial f_n(x_0)) \mid \lambda \in \partial F(f(x_0))\}. \end{aligned}$$

□

Remark 3.18. (a) Assumption (3.25) reduces to a simple, familiar-looking form in important special cases. For example, if $n = 1$, (3.25) becomes

$$(3.28) \quad \text{Range } f_1^\uparrow(x_0; \cdot) - \text{dom } F^\uparrow(f_1(x_0); \cdot) = \mathbf{R}.$$

If F is as in Remark 3.16(a) or (b), (3.25) reduces to the assumption that $\text{dom } f_i^\uparrow(x_0; \cdot)$, $i = 1, \dots, n$ are in strong general position.

(b) The assumption that $f_i^\uparrow(x_0; 0) = 0$, $i = 1, \dots, n$ is not needed in important special cases. For example, if F is as in Remark 3.16(a) or (b) and $f_i^\uparrow(x_0; 0) = -\infty$ for some i , then both sides of (3.27) will equal $\emptyset$.

(c) Corollaries of Theorem 3.17 include Corollary 3.6, Proposition 3.10, and Proposition 3.14 (without our having to use Lemma 3.12).

(d) Any isotone function $F: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ is directionally Lipschitzian for all $y \leq 0$; i.e.,

$$\inf_{\substack{Y \in \mathcal{N}(y) \\ N \in \mathcal{N}(x_0, F(x_0)) \\ \lambda > 0}} \sup_{\substack{y' \in Y \\ (x, \alpha) \in N \cap \text{epi } F \\ t \in (0, \lambda)}} \frac{F(x + ty') - \alpha}{t} < +\infty$$

for all $y \leq 0$ (see [19, Prop. 4]). This property plays a crucial role in the subdifferential calculus results of [18]. However, there is no analogue of Theorem 3.17 for F merely directionally Lipschitzian, as we will see in § 5.

From Theorem 3.17 we can derive an extension of Corollary 3.6 to n functions, and a product rule for (locally) nonnegative functions.

COROLLARY 3.19. *Let $f_i: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$, $i = 1, \dots, n$ be strictly l.s.c. and finite at x_0 , and suppose that $\text{dom } f_i^\uparrow(x_0; \cdot)$, $i = 1, \dots, n$ are in strong general position. Then for all $y \in \mathbf{R}^m$,*

$$(3.29) \quad (f_1 + \dots + f_n)^\uparrow(x_0; y) \leq \sum_{i=1}^n f_i^\uparrow(x_0; y).$$

Moreover,

$$(3.30) \quad \partial(f_1 + \dots + f_n)(x_0) \subset \sum_{i=1}^n \partial f_i(x_0).$$

Proof. Let $F(x_1, \dots, x_n) = \sum_{i=1}^n x_i$ in Theorem 3.17. F is continuous and strictly isotone in each coordinate. Assumption (3.25) in this case reduces to $\text{dom } f_i^\uparrow(x_0; \cdot)$, $i = 1, \dots, n$ being in strong general position. As explained in Remark 3.18(b), the assumption that each $f_i^\uparrow(x_0; 0) = 0$ is not needed. Then (3.29) follows from (3.26) and (3.30) from (3.27). $\square$

COROLLARY 3.20. *Let $f_i: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$, $i = 1, \dots, n$ be nonnegative on $\mathbf{R}^m$ and strictly l.s.c. and positive at $x_0 \in \bigcap_{i=1}^n \text{dom } f_i$. Suppose that $\text{dom } f_i^\uparrow(x_0; \cdot)$, $i = 1, \dots, n$, are in strong general position. Then for all $y \in \mathbf{R}^n$,*

$$(3.31) \quad \left(\prod_{i=1}^n f_i \right)^\uparrow(x_0; y) \leq \sum_{i=1}^n \left(\prod_{j \neq i} f_j(x_0) \right) f_i^\uparrow(x_0; y).$$

Moreover,

$$(3.32) \quad \partial \left(\prod_{i=1}^n f_i \right)(x_0) \subset \sum_{i=1}^n \left(\prod_{j \neq i} f_j(x_0) \right) \partial f_i(x_0).$$

Proof. Let $F(x_1, \dots, x_n) = \prod_{i=1}^n x_i$ in Theorem 3.17. F is continuous and strictly isotone in each coordinate since f_i is nonnegative. Condition (3.25) again reduces to $\text{dom } f_i^\uparrow(x_0; \cdot)$, $i = 1, \dots, n$, being in strong general position. As explained in Remark 3.18(b), the assumption that each $f_i^\uparrow(x_0; \cdot) = 0$ is not needed. Then (3.31) follows from (3.26) and (3.32) from (3.27). $\square$

One special case of Corollary 3.20 is a quotient rule where the denominator is continuous and positive.

LEMMA 3.21. Suppose $g: E \rightarrow \bar{\mathbf{R}}$ is continuous and positive at $x_0 \in \text{dom } g$. Then

$$(3.33) \quad \left(\frac{1}{g}\right)^\uparrow(x_0; y) = \frac{(-g)^\uparrow(x_0; y)}{(g(x_0))^2} \quad \text{for all } y \in E,$$

and

$$(3.34) \quad \partial\left(\frac{1}{g}\right)(x_0) = \frac{1}{(g(x_0))^2} \partial(-g)(x_0).$$

Proof. First observe that (3.34) follows immediately from (3.33) because $1/(g(x_0))^2 > 0$. To prove (3.33), use (1.10) to write

$$\begin{aligned} \left(\frac{1}{g}\right)^\uparrow(x_0; y) &= \limsup_{\substack{x \rightarrow x_0 \\ t \downarrow 0}} \inf_{y' \rightarrow y} \frac{(1/g(x + ty')) - (1/g(x))}{t} \\ &= \limsup_{\substack{x \rightarrow x_0 \\ t \downarrow 0}} \inf_{y' \rightarrow y} \frac{(-g)(x + ty') - (-g)(x)}{t} \cdot \frac{1}{g(x + ty')g(x)} \\ &= \limsup_{\substack{x \rightarrow x_0 \\ t \downarrow 0}} \inf_{y' \rightarrow y} \frac{(-g)(x + ty') - (-g)(x)}{t} \lim_{\substack{x \rightarrow x_0 \\ t \downarrow 0 \\ y' \rightarrow y}} \frac{1}{g(x + ty')g(x)} \end{aligned}$$

(by Lemma 5.2, which we prove in § 5).

$$= \frac{1}{(g(x_0))^2} (-g)^\uparrow(x_0; y). \quad \square$$

PROPOSITION 3.22. Let $f: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ be nonnegative on $\text{dom } f$ and strictly l.s.c. and positive at x_0 , and let $g: \mathbf{R}^m \rightarrow \mathbf{R}$ be positive and continuous at $x_0 \in \text{dom } g \cap \text{dom } f$. Assume that

$$(3.35) \quad \text{dom } f^\uparrow(x_0; \cdot) - \text{dom } (-g)^\uparrow(x_0; \cdot) = \mathbf{R}^m.$$

Then

$$(3.36) \quad \left(\frac{f}{g}\right)^\uparrow(x_0; y) \leq \frac{f(x_0)(-g)^\uparrow(x_0; y) + g(x_0)f^\uparrow(x_0; y)}{(g(x_0))^2} \quad \text{for all } y \in \mathbf{R}^m$$

and

$$(3.37) \quad \partial\left(\frac{f}{g}\right)(x_0) \subset \frac{f(x_0)\partial(-g)(x_0) + g(x_0)\partial f(x_0)}{(g(x_0))^2}.$$

Proof. By Lemma 3.21, $\text{dom } (1/g)^\uparrow(x_0; \cdot) = \text{dom } (-g)^\uparrow(x_0; \cdot)$, so (3.35) ensures that we can apply Corollary 3.20 with $n = 2$, $f_1 := f$, and $f_2 := 1/g$. By (3.31) and (3.33),

$$\begin{aligned} \left(\frac{f}{g}\right)^\uparrow(x_0) &\leq f(x_0)\left(\frac{1}{g}\right)^\uparrow(x_0; y) + \frac{1}{g(x_0)}f^\uparrow(x_0; y) \\ &= \frac{f(x_0)}{(g(x_0))^2}(-g)^\uparrow(x_0; y) + \frac{1}{g(x_0)}f^\uparrow(x_0; y) \\ &= \frac{f(x_0)(-g)^\uparrow(x_0; y) + g(x_0)f^\uparrow(x_0; y)}{(g(x_0))^2}. \end{aligned}$$

Similarly, by (3.32) and (3.34),

$$\begin{aligned}\partial\left(\frac{f}{g}\right)(x_0) &\subset f(x_0)\partial\left(\frac{1}{g}\right)(x_0) + \frac{1}{g(x_0)}\partial f(x_0) \\ &= \frac{f(x_0)}{(g(x_0))^2}\partial(-g)(x_0) + \frac{1}{g(x_0)}\partial f(x_0) \\ &= \frac{f(x_0)\partial(-g)(x_0) + g(x_0)\partial f(x_0)}{(g(x_0))^2}.\end{aligned}\quad \square$$

Remark 3.23. If g is locally Lipschitzian near x_0 , then $(-g)^\uparrow(x_0; y) = -g^\uparrow(x_0; y)$, $\partial(-g)(x_0) = -\partial g(x_0)$, and (3.35) is automatically satisfied. If in addition f is locally Lipschitzian near x_0 , Proposition 3.22 almost reduces to the quotient rule of [11, Chap. 8] and [6, Chap. 2]. (The hypotheses of Proposition 3.22 include the assumption that $f(x_0) \neq 0$, which is not needed in [6] or [11].)

4. Application to constrained optimization problems.

PROPOSITION 4.1 ([18], [20]). *Suppose $f: E \rightarrow \bar{\mathbf{R}}$ has a local minimum at $x_0 \in E$. Then $0 \in \partial f(x_0)$.*

Combining Proposition 4.1 with subdifferential calculus results of the preceding section, we can give a strengthened version, in finite dimensions, of [18, Thm. 6].

PROPOSITION 4.2. *Suppose the problem*

$$\min \{f(x) \mid x \in C\}$$

has a local minimum at $x_0 \in C$, where $f: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ is strictly l.s.c. at x_0 and $C \subset \mathbf{R}^n$ is closed near x_0 . Assume

$$(4.1) \quad \text{dom } f^\uparrow(x_0; \cdot) - T_C(x_0) = \mathbf{R}^n.$$

Then

$$(4.2) \quad 0 \in \partial f(x_0) + N_C(x_0).$$

Proof. Let $f_1 := f$, $f_2 := i_C$ in Corollary 3.6. The function $f + i_C$ is minimized at x_0 , so by Proposition 4.1, $0 \in \partial(f + i_C)(x_0)$. Assumption (4.1) is (3.8) in this case, so $\partial(f + i_C)(x_0) \subset \partial f(x_0) + \partial i_C(x_0) = \partial f(x_0) + N_C(x_0)$, and $0 \in \partial f(x_0) + N_C(x_0)$. $\square$

We will next consider a particular constraint set, the set $C := \{x \mid g(x) \leq 0\}$ for a strictly l.s.c. function $g: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$.

PROPOSITION 4.3 (cf. [18, Thm. 5]). *Suppose $g: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ is strictly l.s.c. at $x_0 \in g^{-1}(0)$, and let $C := \{x \mid g(x) \leq 0\}$. Assume $0 \notin \partial g(x_0)$. Then*

$$(4.3) \quad T_C(x_0) \subset \{y \mid g^\uparrow(x_0; y) \leq 0\}.$$

If in addition $\partial g(x_0)$ is nonempty, then

$$(4.4) \quad N_C(x_0) \subset \left(\bigcup_{\lambda > 0} \lambda \partial g(x_0) \right) \cup \partial^\infty g(x_0).$$

Proof. In Proposition 3.10, let $n = 2$, $D_1 := \{(z, u) \in \mathbf{R}^m \times \mathbf{R} \mid u = 0\}$, and $D_2 := \text{epi } g$. To apply Proposition 3.10, we must verify that $T_{D_1}(x_0, 0) - T_{D_2}(x_0, 0) = \mathbf{R}^{m+1}$; i.e., that $\{(z, u) \mid u = 0\} - \text{epi } g^\uparrow(x_0; \cdot) = \mathbf{R}^{m+1}$. To do so, suppose that $(x', r') \in \mathbf{R}^m \times \mathbf{R}$. Since $0 \notin \partial g(x_0)$, there exists $y \in \mathbf{R}^m$ with $g^\uparrow(x_0; y) < 0$. Now $g^\uparrow(x_0; \cdot)$ is positively homogeneous, so for any $\bar{r} \in \mathbf{R}$, there exists $\bar{y} \in \mathbf{R}^m$ with $(\bar{y}, \bar{r}) \in \text{epi } g^\uparrow(x_0; \cdot)$. Let y' be such that $g^\uparrow(x_0; y') \leq -r'$. Then $(x', r') = (x' + y', 0) - (y', -r')$. Thus $\{(z, u) \mid u = 0\} - \text{epi } g^\uparrow(x_0; \cdot) = \mathbf{R}^{m+1}$. Now by (3.17), $T_{D_1 \cap D_2}(x_0, 0) \supset T_{D_1}(x_0, 0) \cap T_{D_2}(x_0, 0)$; i.e.,

$T_{C \times 0}(x_0, 0) \supset (\mathbf{R}^m \times 0) \cap \text{epi } g^\dagger(x_0; \cdot)$. Hence $T_C(x_0) \supset \{y \mid g^\dagger(x_0; y) \leq 0\}$. By (3.18), $N_{D_1 \cap D_2}(x_0, 0) \subset N_{D_1}(x_0, 0) + N_{D_2}(x_0, 0)$; i.e., $N_{C \times 0}(x_0, 0) \subset 0 \times \mathbf{R} + N_{D_2}(x_0, 0)$. If $\partial g(x_0) \neq \emptyset$, we have by (3.20) that

$$N_{D_2}(x_0) = \bigcup_{\lambda > 0} \lambda(\partial g(x_0), -1) \cup (\partial^\infty g(x_0), 0).$$

Hence

$$N_{C \times 0}(x_0, 0) \subset (0 \times \mathbf{R}) + \left(\bigcup_{\lambda > 0} \lambda(\partial g(x_0), -1) \cup (\partial^\infty g(x_0), 0) \right)$$

and (4.4) holds.

By combining Propositions 4.2 and 4.3 and applying Proposition 3.10, we can obtain the following Kuhn–Tucker Theorem:

THEOREM 4.4. *Let $f: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ and $g_i: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$, $i = 1, \dots, n$, be strictly l.s.c. at x_0 , $D \subset \mathbf{R}^m$ closed near x_0 , and $G: \mathbf{R}^m \rightarrow \mathbf{R}^p$ strictly differentiable at x_0 , a local minimizer of*

$$(P) \quad \min \{f(x) \mid g_i(x) \leq 0, i = 1, \dots, n, G(x) = 0, x \in D\}.$$

Call $I(x) := \{i \in \{1, \dots, n\} \mid g_i(x) = 0\}$. Assume that $0 \notin \partial g_i(x_0)$ and $\partial g_i(x_0) \neq \emptyset$ for each $i \in I(x_0)$, that g_i is continuous at x_0 for each $i \notin I(x_0)$, and that $\nabla G(x_0)\mathbf{R}^m = \mathbf{R}^p$. In addition, assume that $\text{dom } f^\dagger(x_0; \cdot)$, $\nabla G(x_0)^{-1}(0)$, $T_D(x_0)$, and $\{y \mid g^\dagger(x_0; y) \leq 0\}$, $i \in I(x_0)$, are in strong general position. Then

$$(4.5) \quad 0 \in \partial f(x_0) + \sum_{i \in I(x_0)} (\lambda_i \partial g_i(x_0) \cup \partial^\infty g_i(x_0)) + \lambda \nabla G(x_0) + N_D(x_0)$$

for some $\lambda_i \geq 0$ and $\lambda \in \mathbf{R}$.

Proof. Denote by C the constraint set in Program (P). Since $\nabla G(x_0)\mathbf{R}^m = \mathbf{R}^p$, $\nabla G(x_0)^{-1}(0) \subset T_{G^{-1}(0)}(x_0)$ by Theorem 2.4. (In fact, these two sets are equal). By our strong general position assumption, we have

$$\{y \mid g^\dagger(x_0; y) \leq 0, i \in I(x_0), y \in T_D(x_0), \nabla G(x_0)y = 0\} \subset T_C(x_0)$$

by Propositions 4.3 and 3.10. Thus the strong general position assumption guarantees that (4.1) holds. By Proposition 4.2,

$$0 \in \partial f(x_0) + N_C(x_0),$$

and (4.5) then follows immediately from (3.18) of Proposition 3.10 and Proposition 4.3. $\square$

Remark 4.5. In the case in which f and g_i , $i \in I(x_0)$ are strictly differentiable, our general position assumption and $\nabla G(x_0)\mathbf{R}^m = \mathbf{R}^p$ comprise the familiar Mangasarian–Fromovitz constraint qualification.

It is also possible to derive necessary optimality conditions for (P) via the following “Dubovitskii–Milyutin” result, due to Watkins [26]:

THEOREM 4.6. *Let $C_i \subset \mathbf{R}^m$, $i = 1, \dots, p$ be closed near x_0 with $\bigcap_{i=1}^p C_i = \{x_0\}$, and suppose that at least one of the sets $T_{C_i}(x_0)$, $i = 1, \dots, p$ is not a subspace. Then there exist $a_i \in N_{C_i}(x_0)$, $i = 1, \dots, p$, not all equal to zero, such that $\sum_{i=1}^p a_i = 0$.*

In [26], a “Fritz John type” theorem is derived from Theorem 4.6 for a constrained optimization problem involving locally Lipschitzian functions. Using Proposition 4.3, we can generalize this result beyond the locally Lipschitzian case.

THEOREM 4.7. *Posit the assumptions of Theorem 4.4, with the exception of the strong general position hypothesis. Assume, in addition, that f is directionally Lipschitzian at x_0 , that $\partial f(x_0)$ is nonempty, and that $0 \notin \partial f(x_0)$. Then*

$$(4.6) \quad 0 \in (\lambda_0 \partial f(x_0) \cup \partial^\infty f(x_0)) + \sum_{i \in I(x_0)} (\lambda_i \partial g_i(x_0) \cup \partial^\infty g_i(x_0)) + \lambda \nabla G(x_0) + N_D(x_0)$$

for some $\lambda_0 \geq 0$, $\lambda_i \geq 0$, and $\lambda \in \mathbf{R}$. Moreover, λ_0 , λ_i , $i \in I(x_0)$, λ , and the elements of $\partial^\infty f(x_0)$, $\partial^\infty g_i(x_0)$, and $N_D(x_0)$ in (4.6) are not all equal to zero.

Proof. Again call C the feasible set in (P). There exists $\varepsilon > 0$ such that $f(x) \geq f(x_0)$ for all $x \in B_\varepsilon(x_0)$. Let

$$C_0 = \{x \in B_\varepsilon(x_0) \mid f(x) \leq f(x_0) - \|x - x_0\|^2\},$$

$$C_i := \{x \in \mathbf{R}^n \mid g_i(x) \leq 0\}, i = 1, \dots, n,$$

$$C_{n+1} := \{x \in \mathbf{R}^n \mid G(x) = 0\} \quad \text{and} \quad C_{n+2} = D.$$

Since x_0 is a local minimizer for (P), $\bigcap_{i=0}^{n+2} C_i = \{x_0\}$. Consider first the case in which $T_{C_0}(x_0)$ is not a subspace. Then by Theorem 4.6, $0 \in \sum_{i=0}^{n+2} N_{C_i}(x_0)$. Call $\tilde{f}(x) := f(x) - f(x_0) + \|x - x_0\|^2$. By Corollary 3.6, $\partial \tilde{f}(x_0) \subset \partial f(x_0)$ and $\partial^\infty \tilde{f}(x_0) = \partial^\infty f(x_0)$ (see [22, Prop. 2.4]). Thus

$$\begin{aligned} N_{C_0}(x_0) &\subset \bigcup_{\lambda > 0} \lambda \partial \tilde{f}(x_0) \cup \partial^\infty \tilde{f}(x_0) \\ &\subset \bigcup_{\lambda > 0} \lambda \partial f(x_0) \cup \partial^\infty f(x_0). \end{aligned}$$

Again $\nabla G(x_0)^{-1}(0) = T_{C_{n+1}}(x_0)$. For $i \in \{1, \dots, n\} \setminus I(x_0)$, $N_{C_i}(x_0) = \{0\}$. Combining these facts and (4.4), we obtain (4.6).

Now suppose $T_{C_0}(x_0)$ is a subspace. Since f is directionally Lipschitzian at x_0 and $0 \notin \partial f(x_0)$, $T_{C_0}(x_0)$ has nonempty interior [18], which means that $T_{C_0}(x_0) = \mathbf{R}^n$. Hence $x_0 \in \text{int } C_0$ [17], contradicting the fact that x_0 is a local minimizer for (P). $\square$

The additional assumption that $\nabla G(x_0)^{-1}(0)$, $T_D(x_0)$, and $\{y \mid g_i^*(x_0; y) \leq 0\}$, $i \in I(x_0)$ are in strong general position is enough to guarantee that the element of $\lambda_0 \partial f(x_0) \cup \partial^\infty f(x_0)$ in (4.6) is nonzero (see [26, Remarks 1, 2]). This produces another Kuhn–Tucker type result.

5. Limitations of the generalized subdifferential calculus. In § 3, we obtain “chain rules” for $(F \circ f)^1(x_0; \cdot)$ and $\partial(F \circ f)(x_0)$ with F and f of these types:

- (a) $F: \mathbf{R}^m \rightarrow \mathbf{R}$ l.s.c., $f: \mathbf{R}^m \rightarrow \mathbf{R}^n$ strictly differentiable at x_0 .
 - (b) $F: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ l.s.c. and isotone, $f = (f_1, \dots, f_n)$ with each $f_i: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ l.s.c., and (3.23) or (3.24) satisfied.
- (In fact, (a) can be encompassed by (b)—e.g. [18, Thm. 3]).

Are other general chain rules possible? In particular, is there a chain rule for $F: \mathbf{R}^n \rightarrow \mathbf{R}$ directionally Lipschitzian (or strictly differentiable) and $f_i: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ l.s.c.? We answer this question by considering possible extensions of the following chain rule for Clarke subgradients. (See [5, Prop. 10] and [11, Chap. 8].)

THEOREM 5.1. *Let E be a normed space and $f: E \rightarrow \bar{\mathbf{R}}$ be locally Lipschitzian near x_0 , and let $F: \mathbf{R} \rightarrow \mathbf{R}$ be continuously differentiable near x_0 . Then*

$$(5.1) \quad (F \circ f)^0(x_0; y) = F'(f(x_0))f^0(x_0; y) \quad \text{for all } y \in E,$$

and

$$(5.2) \quad \partial(F \circ f)(x_0) = F'(f(x_0))\partial f(x_0).$$

In attempting to extend Theorem 5.1, it does not seem possible to apply the techniques used in the proofs of Theorems 3.2 and 3.17. We will proceed instead by examining limits of the form (1.10). We start by deriving a technical lemma involving these limits which we have already applied in the proof of Lemma 3.21.

LEMMA 5.2. Let E, E_1 be l.c.t.v.s., and let $g: E \times E_1 \rightarrow \bar{\mathbf{R}}$ and $h: E \times E_1 \rightarrow \bar{\mathbf{R}}$ be extended-real-valued functions. Suppose

$$\lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} h(x, y) = A \quad \text{and} \quad \limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} g(x, y) = B.$$

Then

- (i) $\limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} h(x, y)g(x, y) = AB$ if $0 < A < +\infty$.
- (ii) The same equation holds for $-\infty < A \leq 0$ if g is bounded in a neighborhood of (x_0, y_0) .

Proof of (i). For any given $\lambda \in (0, A)$, there exist $X_1 \in \mathcal{N}(x_0)$, $Y_1 \in \mathcal{N}(y_0)$ such that $A - \lambda \leq h(x, y) \leq A + \lambda$ for all $x \in X_1$, $y \in Y_1$. We first establish $\limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} h(x, y)g(x, y) \leq AB$.

Case 1. $0 \leq B < +\infty$. Let $Y \in \mathcal{N}(y)$ and $\varepsilon > 0$ be given. Choose $\lambda \in (0, A)$ small enough so that $(A + B)\lambda + \lambda^2 < \varepsilon$ and X_1, Y_1 as above. Then there exists $X_2 \in \mathcal{N}(x_0)$ such that for all $x \in X_2$, there exists $y \in Y \cap Y_1$ with $g(x, y) \leq B + \lambda$. Thus for all $x \in X_1 \cap X_2$, there exists $y \in Y \cap Y_1$ with $h(x, y)g(x, y) \leq (A + \lambda)(B + \lambda) = AB + (A + B)\lambda + \lambda^2 \leq AB + \varepsilon$. Since ε was arbitrary, $\limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} h(x, y)g(x, y) \leq AB$.

Case 2. $-\infty < B < 0$. Let $Y \in \mathcal{N}(y)$ and $\varepsilon > 0$ be given. Let $\lambda \in (0, \min(A, -B))$ be such that $(A - B)\lambda \leq \varepsilon$, and choose X_1, Y_1 as above. There exists $X_2 \in \mathcal{N}(x_0)$ such that for each $x \in X_2$, there exists $y \in Y \cap Y_1$ with $g(x, y) \leq B + \lambda$. So for each $x \in X_1 \cap X_2$, there exists $y \in Y \cap Y_1$ with $h(x, y)g(x, y) \leq (A - \lambda)(B + \lambda) = AB + (A - B)\lambda - \lambda^2 \leq AB + \varepsilon$. Since ε was arbitrary, $\limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} h(x, y)g(x, y) \leq AB$.

Case 3. $B = -\infty$. Let $m < 0$ and $Y \in \mathcal{N}(y)$ be given, and choose $\lambda \in (0, A)$ and X_1, Y_1 as above. There exists $X_2 \in \mathcal{N}(x_0)$ such that for all $x \in X_2$, there exists $y \in Y \cap Y_1$ with $g(x, y) \leq m/(A - \lambda)$. Then for all $x \in X_1 \cap X_2$, there exists $y \in Y \cap Y_1$ such that $h(x, y)g(x, y) \leq (A - \lambda)(m/(A - \lambda)) = m$. Since m was arbitrary, $\limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} h(x, y)g(x, y) = -\infty = AB$.

Case 4. $B = +\infty$. Choose $\lambda \in (0, A)$, and let $X \in \mathcal{N}(x_0)$ and $M > 0$ be given. Choose X_1, Y_1 as above. There exists $Y \in \mathcal{N}(y_0)$ such that for some $x \in X \cap X_1$, $g(x, y) \geq M$ for all $y \in Y$. So there exists $x \in X \cap X_1$ such that for all $y \in Y \cap Y_1$, $h(x, y)g(x, y) \geq M(A - \lambda)$. Since M and X were arbitrary, $\lim_{x \rightarrow x_0} \sup \inf_{y \rightarrow y_0} g(x, y)h(x, y) = +\infty = AB$. We have now established that $\limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} g(x, y)h(x, y) \leq AB$. It then follows that

$$\begin{aligned} B &= \limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} g(x, y) \\ &\leq \lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} \frac{1}{h(x, y)} \limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} g(x, y)h(x, y) \\ &= \frac{1}{A} \limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} g(x, y)h(x, y). \end{aligned}$$

Hence $\limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} g(x, y)h(x, y) \geq AB$, and (i) holds.

We leave the proof of (ii), which is similar, to the reader. We make the important observation that if $A \leq 0$, local boundedness of g is required in both the proof of $\limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} h(x, y)g(x, y) \leq AB$ and that of $\limsup_{x \rightarrow x_0} \inf_{y \rightarrow y_0} h(x, y)g(x, y) \geq AB$. $\square$

We can now prove what seems to be the best possible extension of Theorem 5.1.

THEOREM 5.3. *Let E be an l.c.t.v.s. and $f: E \rightarrow \bar{\mathbf{R}}$ continuous at $x_0 \in E$, and let $F: \mathbf{R} \rightarrow \mathbf{R}$ be continuously differentiable near $f(x_0)$. Assume $F'(f(x_0)) > 0$. Then for all $y \in E$,*

$$(5.3) \quad (F \circ f)^\uparrow(x_0; y) = F'(f(x_0))f^\uparrow(x_0; y).$$

Moreover,

$$(5.4) \quad \partial(F \circ f)(x_0) = F'(f(x_0))\partial f(x_0).$$

Proof.

$$\begin{aligned} (F \circ f)^\uparrow(x_0; y) &= \limsup_{\substack{x \rightarrow x_0 \\ t \downarrow 0}} \inf_{y' \rightarrow y} \frac{(F \circ f)(x + ty') - (F \circ f)(x)}{t} \\ &= \limsup_{\substack{x \rightarrow x_0 \\ t \downarrow 0}} \inf_{y' \rightarrow y} F'(z(t, x, y)) \frac{f(x + ty') - f(x)}{t} \end{aligned}$$

(for some $z(t, x, y)$ between $f(x)$ and $f(x + ty')$ by the classical mean value theorem)

$$= \lim_{\substack{x \rightarrow x_0 \\ t \downarrow 0 \\ y' \rightarrow y}} F'(z(t, x, y)) \limsup_{\substack{x \rightarrow x_0 \\ t \downarrow 0}} \inf_{y' \rightarrow y} \frac{f(x + ty') - f(x)}{t}$$

(by Lemma 5.2(i))

$$= F'(f(x_0))f^\uparrow(x_0; y)$$

since f is continuous at x_0 and F is continuously differentiable near $f(x_0)$.

Now if $f^\uparrow(x_0; 0) = -\infty$, $(F \circ f)^\uparrow(x_0; 0)$ will by (5.3) also equal $-\infty$, and both sides of (5.4) will be empty. Otherwise,

$$\begin{aligned} \partial(F \circ f)(x_0) &= \partial(F \circ f)^\uparrow(x_0; \cdot)(0) \\ &= \partial(F'(f(x_0))f^\uparrow(x_0; \cdot))(0) \\ &= F'(f(x_0))\partial f^\uparrow(x_0; \cdot)(0) \end{aligned}$$

since $\partial f^\uparrow(x_0; \cdot)(0)$ is nonempty

$$= F'(f(x_0))\partial f(x_0). \quad \square$$

Remark 5.4. (a) If $F'(f(x_0)) < 0$, neither inclusion of (5.4) will hold in general. For example, define $F: \mathbf{R} \rightarrow \mathbf{R}$ by $F(x) := -x$ and $f: \mathbf{R} \rightarrow \mathbf{R}$ by $f(x) := -|x|^{1/2}$. Then $\partial(F \circ f)(0) = \mathbf{R}$, but $\partial f(0) = \emptyset$. On the other hand, if $F(x) := -x$ but $f(x) := |x|^{1/2}$, $\partial(F \circ f)(x_0) = \emptyset$ while $F'(f(x_0))\partial f(x_0) = \mathbf{R}$.

(b) There are also counterexamples for $F'(f(x_0)) = 0$. Define $F: \mathbf{R} \rightarrow \mathbf{R}$ by $F(x) := x^2$ and $f: \mathbf{R} \rightarrow \mathbf{R}$ as in (a). Then $\partial(F \circ f)(0) = [-1, 1]$, but $F'(f(x_0)) = 0$ and $\partial f(0) = \emptyset$. On the other hand, if $F(x) := -x^{4/3}$ and $f(x) := |x|^{1/2}$, then $\partial(F \circ f)(0) = \emptyset$ while $F'(f(0)) = 0$ and $\partial f(0) = \mathbf{R}$.

We now have at least a partial answer to our original questions. Even if F is continuously differentiable and f is continuous, the requirement that F be locally isotone seems to be needed in order to prove a chain rule for upper subderivatives.

The difficulty seems to lie in the “nonradial” nature of ∂f for general f . If f is locally Lipschitzian, $f^0(x_0; -y) = (-f)^0(x_0; y)$ and $\partial(-f)(x_0) = -\partial f(x_0)$. These relationships would allow us to extend Theorem 5.3 to $F'(f(x_0)) < 0$ if they held in general, but they do not—e.g. consider $f: \mathbf{R} \rightarrow \mathbf{R}$ defined by $f(x) := |x|^{1/2}$ and $x_0 = 0$.

A more thorough discussion of the subject of this section is given in [24, § 3.5].

6. Contingent derivatives, Ursescu derivatives and subdifferential regularity. In this section, we consider the tangent cones

$$(6.1) \quad K_C(x_0) := \{y \in E \mid \forall Y \in \mathcal{N}(y), \forall \lambda > 0, \exists t \in (0, \lambda), \\ \exists y' \in Y \text{ such that } x_0 + ty' \in C\}$$

and

$$(6.2) \quad k_C(x_0) := \{y \in E \mid \forall Y \in \mathcal{N}(y), \exists \lambda > 0 \text{ such that } \forall t \in (0, \lambda), \exists y' \in Y \text{ with } x_0 + ty' \in C\}.$$

$K_C(x_0)$ is the well-known *contingent cone*. For more information on these cones, see [23], [7], [25], [24].

From (6.1) and (6.2), we see that

$$T_C(x_0) \subset k_C(x_0) \subset K_C(x_0)$$

for any $C \subset E$ and $x_0 \in \text{cl } C$. If C is a convex set or a differentiable manifold, all three cones coincide ([20, Chap. 2] and Theorem 2.4).

Let $f: E \rightarrow \bar{\mathbf{R}}$ be an extended real-valued function. Paralleling the development in § 1, define

$$(6.3) \quad f_+(x_0; y) := \inf \{r \mid (y, r) \in K_{\text{epi } f}(x_0, f(x_0))\}$$

and

$$(6.4) \quad f_{\square}(x_0; y) := \inf \{r \mid (y, r) \in k_{\text{epi } f}(x_0, f(x_0))\}.$$

The directional derivative $f_+(x_0; y)$ (respectively, $f_{\square}(x_0; y)$) has been called the *contingent derivative* (*Ursescu derivative*) of f at x_0 in the direction y [1], [9]. Since $T_{\text{epi } f}(x_0, f(x_0)) \subset k_{\text{epi } f}(x_0, f(x_0)) \subset K_{\text{epi } f}(x_0, f(x_0))$, $f_+(x_0; \cdot) \leq f_{\square}(x_0; \cdot) \leq f^{\uparrow}(x_0; \cdot)$. Definitions (6.3) and (6.4) are made exactly so that

$$(6.5) \quad \text{epi } f_+(x_0; \cdot) := K_{\text{epi } f}(x_0, f(x_0))$$

and

$$(6.6) \quad \text{epi } f_{\square}(x_0; \cdot) := k_{\text{epi } f}(x_0, f(x_0)).$$

There are also direct characterizations of $f_+(x_0; y)$ and $f_{\square}(x_0; y)$. It is well known that

$$(6.7) \quad f_+(x_0; y) = \liminf_{\substack{t \downarrow 0 \\ y' \rightarrow y}} \frac{f(x_0 + ty') - f(x_0)}{2}$$

(e.g., [20, Chap. 3]), and it is not difficult to show [7], [25], [9] that

$$(6.8) \quad f_{\square}(x_0; y) = \limsup_{t \downarrow 0} \inf_{y' \rightarrow y} \frac{f(x_0 + ty') - f(x_0)}{t}.$$

$K_C(x_0)$ and $k_C(x_0)$ are closed cones, so $f_+(x_0; \cdot)$ and $f_{\square}(x_0; \cdot)$ are l.s.c. and positively homogeneous. $K_C(x_0)$ and $k_C(x_0)$ are not in general convex, so it is not possible to directly associate convex subgradient sets with $f_+(x_0; \cdot)$ and $f_{\square}(x_0; \cdot)$ as we did with $f^{\uparrow}(x_0; \cdot)$ in Definition 1.3. It is possible, however, to prove calculus formulas for $f_+(x_0; \cdot)$ and $f_{\square}(x_0; \cdot)$, and in particular to derive the subdifferential calculus for convex functions by combining these formulae with those in § 3. We now proceed to do so, beginning with a counterpart to Theorem 3.17. Notice that it holds quite generally, with no need for the assumption of finite dimensionality or conditions (3.23), (3.24) and (3.25).

PROPOSITION 6.1. Let E be an l.c.t.v.s., let $f_i: E \rightarrow \bar{\mathbf{R}}, i = 1, \dots, n$ be finite at x_0 , and call $f := (f_1, \dots, f_n)$. Let $F: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be finite at $f(x_0)$ and isotone on the union of $\text{Range } f + \mathbf{R}_+^n$ and $B_\varepsilon(f(x_0))$ for some $\varepsilon > 0$. Then if $y \in E$ is such that $(f_i)_+(x_0; y), i = 1, \dots, n$ are finite,

$$(6.9) \quad (F \circ f)_+(x_0; y) \geq F_+(f(x_0); (f_1)_+(x_0; y), \dots, (f_n)_+(x_0; y)).$$

Alternatively, if $(f_1)_\square(x_0; y), (f_i)_+(x_0; y), i = 2, \dots, n$ are finite, then

$$(6.10) \quad (F \circ f)_\square(x_0; y) \geq F_+(f(x_0); (f_1)_\square(x_0; y), (f_2)_+(x_0; y), \dots, (f_n)_+(x_0; y)).$$

Proof. Let $(y, r) \in K_{\text{epi}(F \circ f)}(x_0, (F \circ f)(x_0))$. To prove (6.9), it suffices to show that

$$((f_1)_+(x_0; y), \dots, (f_n)_+(x_0; y), r) \in K_{\text{epi } F}(f(x_0), F(f(x_0))).$$

To this end, let $\lambda > 0$ and $\delta > 0$ be given. Call $z_i = (f_i)_+(x_0; y) - \delta$, and choose $\lambda_0 \in (0, \lambda)$ such that for all $i, f_i(x_0) + [0, \lambda_0]z_i \geq f_i(x_0) - \varepsilon$. Then there exist $Y_i \in \mathcal{N}(y), \lambda_i \in (0, \lambda_0)$ such that for all $t \in (0, \lambda_i)$ and all $y' \in Y_i$,

$$\frac{f_i(x_0 + ty') - f_i(x_0)}{t} \geq z_i, \quad i = 1, \dots, n.$$

In addition, there exist $\bar{y} \in \bigcap_{i=1}^n Y_i, \bar{t} \in (0, \min_i \lambda_i)$ such that

$$\frac{(F \circ f)(x_0 + \bar{t}\bar{y}) - (F \circ f)(x_0)}{\bar{t}} \leq r + \delta.$$

Now since $f_i(x_0) + \bar{t}z_i \leq f_i(x_0 + \bar{t}\bar{y})$ and F is isotone,

$$\frac{F(f(x_0) + \bar{t}(z_1, \dots, z_n)) - F(f(x_0))}{\bar{t}} \leq r + \delta.$$

Thus $((f_1)_+(x_0; y), \dots, (f_n)_+(x_0; y), r) \in K_{\text{epi } F}(f(x_0), F(f(x_0)))$, and so (6.9) holds.

Now let $(y, r) \in k_{\text{epi}(F \circ f)}(x_0, (F \circ f)(x_0))$. To prove (6.10), it suffices to show that $((f_1)_\square(x_0; y), (f_2)_+(x_0; y), \dots, (f_n)_+(x_0; y), r) \in K_{\text{epi } F}(f(x_0), F(f(x_0)))$. To this end, let $\delta > 0$ and $\lambda > 0$ be given. Call $z_1 = (f_1)_\square(x_0; y) - \delta, z_i = (f_i)_+(x_0; y) - \delta, i = 2, \dots, n$. Again choose $\lambda_0 \in (0, \lambda)$ such that $f_i(x_0) + [0, \lambda_0]z_i \geq f_i(x_0) - \varepsilon$. There exist $Y_i \in \mathcal{N}(y), \lambda_i \in (0, \lambda_0)$ such that for all $t \in (0, \lambda_i)$ and all $y' \in Y_i, ((f_i(x_0 + ty') - f_i(x_0))/t) \geq z_i, i = 2, \dots, n$, and there exists $Y_1 \in \mathcal{N}(y)$ such that for all $t > 0$, there exists $t' \in (0, t)$ with $((f_1(x_0 + t'y) - f_1(x_0))/t') \geq z_1$ for all $y' \in Y_1$. In addition, there exists $\bar{y} \in \bigcap_{i=1}^n Y_i$ and $\lambda_1 \in (0, \lambda_0)$ such that for all $t \in (0, \lambda_1), (((F \circ f)(x_0 + t\bar{y}) - (F \circ f)(x_0))/t) \leq r + \delta$. Thus for some $\bar{t} \in (0, \min_i \lambda_i), f_i(x_0) + \bar{t}z_i \leq f_i(x_0 + \bar{t}\bar{y})$. By the isotonicity of $F, ((F(f(x_0) + \bar{t}(z_1, \dots, z_n)) - F(f(x_0)))/\bar{t}) \leq r + \delta$. Hence $((f_1)_\square(x_0; y), (f_2)_+(x_0; y), \dots, (f_n)_+(x_0; y), r) \in K_{\text{epi } F}(f(x_0), F(f(x_0)))$, and so (6.10) holds. $\square$

Inequality (6.9) does not hold with “+” replaced by “ $\square$ ”. For example, let $n = 2$ and $F(x_1, x_2) := x_1 + x_2$; and consider $f_1: \mathbf{R} \rightarrow \mathbf{R}$ and $f_2: \mathbf{R} \rightarrow \mathbf{R}$ defined by

$$f_1(x) := \begin{cases} 0 & \text{if } x \leq 0, \\ x & \text{if } \frac{1}{2^{n+1}} < x \leq \frac{1}{2^n} \text{ and } n \text{ is odd,} \\ 2x & \text{if } \frac{1}{2^{n+1}} < x \leq \frac{1}{2^n} \text{ and } n \text{ is even,} \end{cases}$$

$$f_2(x) := \begin{cases} 0 & \text{if } x \leq 0, \\ x & \text{if } \frac{1}{2^{n+1}} < x \leq \frac{1}{2^n} \text{ and } n \text{ is even,} \\ 2x & \text{if } \frac{1}{2^{n+1}} < x \leq \frac{1}{2^n} \text{ and } n \text{ is odd.} \end{cases}$$

Then if $y \geq 0$,

$$(f_1)_{\square}(0; y) = (f_2)_{\square}(0; y) = 2y$$

while

$$(f_1 + f_2)(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ 3x & \text{if } x > 0, \end{cases}$$

so that $(f_1 + f_2)_{\square}(0; y) = 3y$. Thus for $y > 0$,

$$(f_1 + f_2)_{\square}(0; y) = 3y < 4y = (f_1)_{\square}(0; y) + (f_2)_{\square}(0; y).$$

PROPOSITION 6.2. *Let E, E^1 be l.c.t.v.s., let $F: E \rightarrow E^1$ be strictly differentiable at $x_0 \in E$, and let $f: E^1 \rightarrow \bar{\mathbf{R}}$ be finite at $F(x_0)$. Then for all $y \in E^1$,*

$$(6.11) \quad (f \circ F)_+(x_0; y) \geq f_+(F(x_0); \nabla F(x_0)y)$$

and

$$(6.12) \quad (f \circ F)_{\square}(x_0; y) \geq f_{\square}(F(x_0); \nabla F(x_0)y).$$

Proof. Let $(y, r) \in K_{\text{epi}(f \circ F)}(x_0, (f \circ F)(x_0))$. To prove (6.11), it suffices to show that $(\nabla F(x_0)y, r) \in K_{\text{epi } f}(F(x_0), f(F(x_0)))$. Let $Z \in \mathcal{N}(\nabla F(x_0)y)$, $\lambda > 0$, and $\delta > 0$ be given. There exist $Y \in \mathcal{N}(y)$ and $\lambda_1 \in (0, \lambda)$ such that $((F(x_0 + ty') - F(x_0))/t) \in Z$ for all $y' \in Y$ and $t \in (0, \lambda_1)$. In addition, there exist $\bar{y} \in Y$ and $\bar{t} \in (0, \lambda_1)$ such that $((f \circ F)(x_0 + \bar{t}\bar{y}) - (f \circ F)(x_0))/\bar{t} \leq r + \delta$. Call $\bar{z} := ((F(x_0 + \bar{t}\bar{y}) - F(x_0))/\bar{t})$. Then $\bar{z} \in Z$ and $((f(F(x_0) + \bar{t}\bar{z}) - f(F(x_0)))/\bar{t}) \leq r + \delta$. Hence $(\nabla F(x_0)y, r) \in K_{\text{epi } f}(F(x_0), f(F(x_0)))$, and (6.11) holds. The proof of (6.12) is very similar to the proof of (6.11), and we leave it to the reader. $\square$

We can now give conditions under which equality holds in Theorems 3.2 and 3.17.

DEFINITION 6.3. The function $f: E \rightarrow \bar{\mathbf{R}}$ is said to be *subdifferentially regular* at $x_0 \in \text{dom } f$ if $f_+(x_0; y) = f^{\dagger}(x_0; y)$ for all $y \in E$. It is said to be *subdifferentially weakly regular* at x_0 if $f_{\square}(x_0; y) = f^{\dagger}(x_0; y)$ for all $y \in E$.

PROPOSITION 6.4. *Suppose the hypotheses of Theorem 3.17 are satisfied, including the assumption that each $f_i^{\dagger}(x_0; \cdot)$ is proper. Assume in addition that F is subdifferentially regular at $f(x_0)$, $f_2, \dots, f_n$ are subdifferentially regular at x_0 , and f_1 is subdifferentially weakly regular at x_0 . Then equality holds in (3.26) and (3.27).*

Proof. By (6.10) and our regularity hypothesis, for all $y \in \mathbf{R}^n$,

$$\begin{aligned} (F \circ f)^{\dagger}(x_0; y) &\geq (F \circ f)_{\square}(x_0; y) \\ &\geq F_+(f(x_0); (f_1)_{\square}(x_0; y), (f_2)_+(x_0; y), \dots, (f_n)_+(x_0; y)) \\ &= F^{\dagger}(f(x_0); f_1^{\dagger}(x_0; y), \dots, f_n^{\dagger}(x_0; y)), \end{aligned}$$

so equality holds in (3.26). Equality in (3.27) follows from equality in (3.26). $\square$

PROPOSITION 6.5. *Suppose the hypotheses of Theorem 3.2 are satisfied, and that either*

$$(6.13) \quad f_1 \text{ is subdifferentially regular at } x_0 \text{ and } f_2 \text{ is subdifferentially weakly regular at } F(x_0), \text{ or}$$

$$(6.14) \quad f_1 \text{ is subdifferentially weakly regular at } x_0, \text{ and } f_2 \text{ is subdifferentially regular at } F(x_0).$$

Then equality holds in (3.3). If in addition $f_1^\uparrow(x_0; 0) = f_2^\uparrow(F(x_0); 0) = 0$, equality holds in (3.2) also.

Proof. If either $f_1^\uparrow(x_0; 0)$ or $f_2^\uparrow(F(x_0); 0)$ is $-\infty$, (3.2) implies that $(f_1 + f_2 \circ F)^\uparrow(x_0; 0) = -\infty$, and both sides of (3.3) are empty. Otherwise apply Proposition 6.1 with $n = 2$ and $F: \mathbf{R}^2 \rightarrow \mathbf{R}$ defined by $F(x_1, x_2) = x_1 + x_2$. If (6.13) holds, then for all $y \in \mathbf{R}^n$,

$$\begin{aligned} (f_1 + f_2 \circ F)^\uparrow(x_0; y) &\geq (f_1 + f_2 \circ F)_\square(x_0; y) \\ &\geq (f_1)_+(x_0; y) + (f_2 \circ F)_\square(x_0; y) \quad (\text{by (6.10)}) \\ &\geq (f_1)_+(x_0; y) + (f_2)_\square(F(x_0); \nabla F(x_0)y) \quad (\text{by (6.12)}) \\ &= (f_1)^\uparrow(x_0; y) + f_2^\uparrow(F(x_0); \nabla F(x_0)y). \end{aligned}$$

If (6.14) holds, then similarly, for all $y \in \mathbf{R}^n$,

$$\begin{aligned} (f_1 + f_1 \circ F)^\uparrow(x_0; y) &\geq (f_1 + f_2 \circ F)_\square(x_0; y) \\ &\geq (f_1)_\square(x_0; y) + (f_2 \circ F)_+(x_0; y) \quad (\text{by (6.10)}) \\ &\geq (f_1)_\square(x_0; y) + (f_2)_+(F(x_0); \nabla F(x_0)y) \quad (\text{by (6.11)}) \\ &= (f_1)^\uparrow(x_0; y) + f_2^\uparrow(F(x_0); \nabla F(x_0)y). \end{aligned}$$

In either case, equality holds in (3.2), and equality in (3.3) follows. $\square$

Remark 6.6. The conditions in Propositions 6.4 and 6.5 guaranteeing equality in the subdifferential calculus formulae of § 3 are less stringent than those given in [18]. For example, let $C_1 := \bigcup_{n=0}^\infty [2^{2n}, 2^{2n+1}]$ and $C := C_1 \cup (-C_1) \cup \{0\}$. Then $T_C(0) = k_C(0) = \{0\}$, but $K_C(0) = \mathbf{R}$. In Proposition 6.5, let $F := I$ and $x_0 := 0$. Define $f_1: \mathbf{R} \rightarrow \mathbf{R}$ by $f_1(x) := 0$, and let $f_2: \mathbf{R} \rightarrow \bar{\mathbf{R}}$ be i_C . Since (3.8) and (6.13) are satisfied, equality holds in (3.9) and (3.10), even though the condition for equality given in Theorem 2 of [18]—that both f_1 and f_2 be subdifferentially regular at x_0 —is not satisfied. Observe also that the regularity conditions in Propositions 6.4 and 6.5 are still valid in the infinite-dimensional setting of [18].

Since convex functions are subdifferentially regular, Propositions 6.4 and 6.5 enable us to rederive convex subdifferential calculus results, in particular Theorems 2.9 and 2.10. Thus Theorems 3.2 and 3.17, which are proved by means of Theorems 2.9 and 2.10, combine with Propositions 6.4 and 6.5 to yield Theorems 2.9 and 2.10 as special cases. We list below one further convex subdifferential calculus result.

PROPOSITION 6.7. *Let $F: \mathbf{R}^n \rightarrow \mathbf{R}^m$ be strictly differentiable at x_0 , let $f_1: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be convex, proper, and strictly l.s.c. at x_0 , and let $f_2: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ be convex, proper, and strictly l.s.c. at $F(x_0)$, where $x_0 \in \text{dom } f_1 \cap F^{-1}(\text{dom } f_2)$. Assume that*

$$(6.15) \quad \nabla F(x_0)x_0 - F(x_0) \in \text{int}(\nabla F(x_0) \text{dom } f_1 - \text{dom } f_2).$$

Then

$$(6.16) \quad \partial(f_1 + f_2 \circ F)(x_0) = \partial f_1(x_0) + \nabla F(x_0)^T \partial f_2(x_0).$$

Proof. By Theorem 3.2 and Proposition 6.5, equality will hold in (6.16) as long as $\nabla F(x_0) \text{dom } f_1^\uparrow(x_0; \cdot) - \text{dom } f_2^\uparrow(F(x_0); \cdot) = \mathbf{R}^m$, which in this case is equivalent to

$$\nabla F(x_0) \text{cone}(\text{dom } f_1 - x_0) - \text{cone}(\text{dom } f_2 - F(x_0)) = \mathbf{R}^m$$

(see [16], § 23). This reduces to

$$\text{cone}(\nabla F(x_0) \text{dom } f_1 - \text{dom } f_2 - \nabla F(x_0)x_0 + F(x_0)) = \mathbf{R}^m,$$

which is equivalent to (6.15). $\square$

COROLLARY 6.8 (cf. [17], Thm. 23.8]). Let $f_1: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ and $f_2: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ be convex, proper and l.s.c., and let $x_0 \in \text{dom } f_1 \cap \text{dom } f_2$. Assume that

$$(6.17) \quad \text{ri dom } f_1 \cap \text{ri dom } f_2 \neq \emptyset.$$

Then

$$(6.18) \quad \partial(f_1 + f_2)(x_0) = \partial f_1(x_0) + \partial f_2(x_0).$$

Proof. By Proposition 6.7 with $m = n$ and $F := I$, (6.18) will hold if $\text{cone}(\text{dom } f_1 - \text{dom } f_2) = \mathbf{R}^m$. By (3.12) of Remark 3.7 (b), this assumption can be weakened to $\text{cone}(\text{dom } f_1 - \text{dom } f_2) = \text{span}(\text{dom } f_1 - \text{dom } f_2)$, which is equivalent to (6.17). $\square$

REMARK 6.9. We saw in Remark 3.7 (b) that hypothesis (3.8) of Corollary 3.6 cannot be weakened to (3.13). The reason is that $\text{aff dom } f^\uparrow(x_0; \cdot)$ can be of smaller dimension than $\text{aff dom } f$, as in the example given there. If f is convex and proper, however, $\text{aff dom } f^\uparrow(x_0; \cdot) = \text{aff dom } f$, and assumption (6.17) is sufficient to give (6.18).

It is possible by the techniques of § 3 to prove analogues of (3.2) and (3.26) involving contingent and Ursescu directional derivatives. To do so, we need counterparts of Theorem 2.4 and Propositions 2.5 and 2.6.

THEOREM 6.10. Let $G: \mathbf{R}^p \rightarrow \mathbf{R}^q$ be strictly differentiable at $x_0 \in G^{-1}(0)$, and let $D \subset \mathbf{R}^p$ be closed near x_0 . Assume (2.2) holds. Then

$$(6.19) \quad K_D(x_0) \cap \nabla G(x_0)^{-1}(0) \subset K_{D \cap G^{-1}(0)}(x_0)$$

and

$$(6.20) \quad k_D(x_0) \cap \nabla G(x_0)^{-1}(0) \subset k_{D \cap G^{-1}(0)}(x_0).$$

Proof. The inclusion (6.19) is a special case of [4, Thm. 4.1(a)]. The proof of (6.20) is entirely analogous to that of [4, Thm. 4.1]. $\square$

PROPOSITION 6.11. Let E, E_1 be l.c.t.v.s., and let $x_0 \in C \subset E$ and $y_0 \in D \in E_1$. Then

$$(6.21) \quad k_{C \times D}(x_0, y_0) = k_C(x_0) \times k_D(y_0),$$

$$(6.22) \quad K_{C \times D}(x_0, y_0) \subset K_C(x_0) \times K_D(y_0),$$

$$(6.23) \quad K_{C \times D}(x_0, y_0) \supset K_C(x_0) \times k_D(y_0).$$

The proof of Proposition 6.11 is a straightforward consequence of (6.1) and (6.2). That equality does not in general hold in (6.22) is demonstrated by an example in § 2 of [4].

PROPOSITION 6.12. Let E, E_1 be l.c.t.v.s. and $A: E \rightarrow E_1$ be linear and continuous. Let $z_0 \in C \subset E$. Then

$$(6.24) \quad A(K_C(z_0)) \subset K_{A(C)}(Az_0)$$

and

$$(6.25) \quad A(k_C(z_0)) \subset k_{A(C)}(Az_0).$$

Inclusion (6.24) is well known (see, for example, [1], [3]). The proof of (6.25) parallels that of (6.24).

THEOREM 6.13. Let $f_i: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$, $i = 1, \dots$, be finite and strictly l.s.c. at x_0 , and call $f := (f_1, \dots, f_n)$. Let $F: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be finite and strictly l.s.c. at $f(x_0)$ and isotone on the union of $\text{Range } f + \mathbf{R}_+^n$ and $B_\varepsilon(f(x_0))$ for some $\varepsilon > 0$. Suppose (3.25) holds. Then for all $y \in \mathbf{R}^n$,

$$(6.26) \quad (F \circ f)_\square(x_0; y) \leq F_\square(f(x_0); (f_1)_\square(x_0; y), \dots, (f_n)_\square(x_0; y))$$

and

$$(6.27) \quad (F \circ f)_+(x_0; y) \leq F_{\square}(f(x_0); (f_1)_+(x_0; y), (f_2)_{\square}(x_0; y), \dots, (f_n)_{\square}(x_0; y)).$$

Equality holds in (6.26) if $(f_i)_{\square}(x_0; y) = (f_i)_+(x_0; y)$, $i = 2, \dots, n$ for all $y \in \mathbf{R}^m$, $F_{\square}(f(x_0); z) = F_+(f(x_0); z)$ for all $z \in \mathbf{R}^n$, and $(f_i)_+(x_0; y)$, $i = 1, \dots, n$ are never $-\infty$.

Proof of (6.26). Let $h := F \circ f$, and define A and G as in Lemma 3.15 and Theorem 3.17. Define

$$D := \text{epi } f_1 \times \dots \times \text{epi } f_n \times \text{epi } F \text{ and } C := D \cap G^{-1}(0).$$

As in Theorem 3.17, $\text{epi } h = A(C)$. Then

$$\begin{aligned} \text{epi } h_{\square}(x_0; \cdot) &= k_{A(C)}(x_0, h(x_0)) \\ &\supset A(k_C(z_0)) \end{aligned}$$

where $z_0 := (x_0, f_1(x_0), \dots, x_0, f_n(x_0), f_1(x_0), \dots, f_n(x_0), h(x_0))$, by (6.25). Now assumption (3.25) guarantees that $\nabla G(x_0) T_D(z_0) = \mathbf{R}^{(n-1)m+n}$, so we may apply Theorem 6.10. By (6.20) and (6.21),

$$\begin{aligned} A(k_C(z_0)) &\supset A(k_D(z_0) \cap \nabla G(z_0)^{-1}(0)) \\ &= \{(x, r) \in \mathbf{R}^m \times \mathbf{R} \mid \exists y \in \mathbf{R}^n \text{ with} \\ &\quad (f_i)_{\square}(x_0; x) \leq y_i, F_{\square}(f(x_0); y) \leq r\} \\ &= \text{epi } F_{\square}(f(x_0); (f_1)_{\square}(x_0; \cdot), \dots, (f_n)_{\square}(x_0; \cdot)) \end{aligned}$$

since $F_{\square}(f(x_0); \cdot)$ is itself isotone (as in Lemma 2.8). Hence $\text{epi } h_{\square}(x_0; \cdot) \supset \text{epi } F_{\square}(f(x_0); (f_1)_{\square}(x_0; \cdot), \dots, (f_n)_{\square}(x_0; \cdot))$, and (6.26) holds. The proof of (6.27) is similar and uses (6.24), (6.19), and (6.23). That equality holds in (6.26) and (6.27) under the given assumptions is a consequence of Proposition 6.1. $\square$

THEOREM 6.14. Let $F: \mathbf{R}^n \rightarrow \mathbf{R}^m$ be strictly differentiable at x_0 , let $f_1: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ be finite and strictly l.s.c. at x_0 , and let $f_2: \mathbf{R}^m \rightarrow \bar{\mathbf{R}}$ be finite and strictly l.s.c. at $F(x_0)$. Assume (3.1) holds. Then for all $y \in \mathbf{R}^n$,

$$(6.28) \quad (f_1 + f_2 \circ F)_{\square}(x_0; y) \leq (f_1)_{\square}(x_0; y) + (f_2)_{\square}(F(x_0); \nabla F(x_0)y)$$

and

$$(6.29) \quad (f_1 + f_2 \circ F)_+(x_0; y) \leq (f_1)_{\square}(x_0; y) + (f_2)_+(F(x_0); \nabla F(x_0)y),$$

$$(6.30) \quad (f_1 + f_2 \circ F)_+(x_0; y) \leq (f_1)_+(x_0; y) + (f_2)_{\square}(F(x_0); \nabla F(x_0)y).$$

If $(f_1)_+(x_0; \cdot)$ and $(f_2)_+(F(x_0); \nabla F(x_0)(\cdot))$ are never $-\infty$, equality holds in (6.28) and (6.29) if $(f_1)_{\square}(x_0; \cdot) = (f_1)_+(x_0; \cdot)$ and in (6.30) if $(f_2)_{\square}(F(x_0); \nabla F(x_0)(\cdot)) = (f_2)_+(F(x_0); \nabla F(x_0)(\cdot))$.

Proof. The proofs of (6.28), (6.29) and (6.30) are analogous to that of Theorem 3.2, using Theorem 6.10 and Propositions 6.11 and 6.12. That equality holds under the given assumptions is a consequence of Propositions 6.1 and 6.2. $\square$

We give here two consequences of Theorem 6.14.

COROLLARY 6.15. Let $C_1 \subset \mathbf{R}^n$ be closed near x_0 , let $F: \mathbf{R}^n \rightarrow \mathbf{R}^m$ be strictly differentiable at x_0 , and let $C_2 \subset \mathbf{R}^m$ be closed near $F(x_0)$. Assume that (3.4) holds. Then

$$(6.31) \quad k_{C_1 \cap F^{-1}(C_2)}(x_0) = k_{C_1}(x_0) \cap \nabla F(x_0)^{-1} k_{C_2}(F(x_0)),$$

$$\begin{aligned} (6.32) \quad K_{C_1}(x_0) \cap \nabla F(x_0)^{-1} K_{C_2}(F(x_0)) &\supset K_{C_1 \cap F^{-1}(C_2)}(x_0) \\ &\supset k_{C_1}(x_0) \cap \nabla F(x_0)^{-1} K_{C_2}(F(x_0)), \end{aligned}$$

and

$$(6.33) \quad K_{C_1 \cap F^{-1}(C_2)}(x_0) \supset K_{C_1}(x_0) \cap \nabla F(x_0)^{-1} k_{C_2}(F(x_0)).$$

Proof. In (6.31), $k_{C_1 \cap F^{-1}(C_2)}(x_0) \subset k_{C_1}(x_0) \cap \nabla F(x_0)^{-1} k_{C_2}(F(x_0))$ always holds. The opposite inclusion follows from (6.28) with $f_1 := i_{C_1}$ and $f_2 := i_{C_2}$. The first inclusion in (6.32) always holds, while the second is (6.29) with $f_1 := i_{C_1}$ and $f_2 := i_{C_2}$. Similarly, (6.33) follows from (6.30). $\square$

COROLLARY 6.16. *Let $f_1: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$, $f_2: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be finite and strictly l.s.c. at $x_0 \in \mathbf{R}^n$. Assume (3.8) holds. Then for all $y \in \mathbf{R}^n$,*

$$(6.34) \quad (f_1 + f_2)_{\square}(x_0; y) \leq (f_1)_{\square}(x_0; y) + (f_2)_{\square}(x_0; y)$$

and

$$(6.35) \quad (f_1 + f_2)_{+}(x_0; y) \leq (f_1)_{+}(x_0; y) + (f_2)_{\square}(x_0; y).$$

Equality holds in (6.34) and (6.35) if $(f_2)_{\square}(x_0; \cdot) = (f_2)_{+}(x_0; y)$ and if $(f_1)_{+}(x_0; \cdot)$ and $(f_2)_{+}(x_0; \cdot)$ are never equal to $-\infty$.

Proof. Take $m = n$ and $F := I$ in Theorem 6.14. $\square$

In the special case in which $f_{\square}(x_0; \cdot)$ is convex, we can define $\partial^k f(x_0) := \partial(f_{\square}(x_0; \cdot))(0)$ and can obtain results like the following:

PROPOSITION 6.17. *Let $f_1: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ and $f_2: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be finite and strictly l.s.c. at x_0 . Assume that $(f_1)_{\square}(x_0; \cdot)$ and $(f_2)_{\square}(x_0; \cdot)$ are convex, and that (3.8) holds. Then*

$$(6.36) \quad \partial^k(f_1 + f_2)(x_0) \subset \partial^k f_1(x_0) + \partial^k f_2(x_0).$$

Equality holds in (6.36) if $(f_1)_{\square}(x_0; \cdot) = (f_1)_{+}(x_0; \cdot)$ or $(f_2)_{\square}(x_0; \cdot) = (f_2)_{+}(x_0; \cdot)$ —in particular, if f_1 or f_2 is convex.

Proof. Since (3.8) holds, so does (6.34), and (6.36) follows from (6.34). Equality under the given assumptions follows from Corollary 6.16 and the fact that if either $(f_1)_{\square}(x_0; \cdot)$ or $(f_2)_{\square}(x_0; \cdot)$ is not proper, both sides of (6.36) are empty. $\square$

It is interesting to note that assumptions involving $f^{\uparrow}(x_0; \cdot)$ are required in these results on $f_{+}(x_0; \cdot)$ and $f_{\square}(x_0; \cdot)$. The same is true of Ioffe's results on approximate subdifferentials [13].

We now present an example which satisfies the hypothesis of Proposition 6.17 even though f_1 and f_2 are not convex.

Example 6.18. Define $f_1: \mathbf{R} \rightarrow \mathbf{R}$ by

$$f_1(x) := \begin{cases} 0 & \text{if } x \leq 0, \\ \frac{1}{2^{n+1}} & \text{if } \frac{1}{2^{n+1}} < x \leq \frac{1}{2^n}, \quad n = 0, \pm 1, \pm 2, \dots, \end{cases}$$

and define $f_2: \mathbf{R} \rightarrow \mathbf{R}$ by

$$f_2(x) := \begin{cases} 0 & \text{if } x \leq 0, \\ \frac{1}{4^{(n+1)}} & \text{if } \frac{1}{4^{(n+1)}} < x \leq \frac{1}{4^n}, \quad n = 0, \pm 1, \pm 2, \dots \end{cases}$$

It is not hard to see that

$$f_1^{\uparrow}(0; y) = f_2^{\uparrow}(0; y) = \begin{cases} 0 & \text{if } y \leq 0, \\ +\infty & \text{if } y > 0, \end{cases}$$

$$(f_1)_\square(0; y) = (f_2)_\square(0; y) = \begin{cases} 0 & \text{if } y \leq 0, \\ y & \text{if } y > 0, \end{cases}$$

$$(f_1 + f_2)_\square(0; y) = \begin{cases} 0 & \text{if } y \leq 0, \\ 2y & \text{if } y > 0, \end{cases}$$

$$(f_1)_+(0; y) = \begin{cases} 0 & \text{if } y \leq 0, \\ y/2 & \text{if } y > 0, \end{cases}$$

$$(f_2)_+(0; y) = \begin{cases} 0 & \text{if } y \leq 0, \\ y/4 & \text{if } y > 0, \end{cases}$$

$$(f_1 + f_2)_+(0; y) = \begin{cases} 0 & \text{if } y \leq 0, \\ 3y/4 & \text{if } y > 0. \end{cases}$$

Then $\text{dom } f_1^\uparrow(0; \cdot) = \text{dom } f_2^\uparrow(0; \cdot) = \{y \mid y \leq 0\}$, so (3.8) is satisfied and (6.36) holds. In fact $\partial_\square(f_1 + f_2)(0) = [0, 2]$ while $\partial_\square f_1(0) = \partial_\square f_2(0) = [0, 1]$, so equality holds in (6.36) even though $(f_i)_+(x_0; \cdot) \neq (f_i)_\square(x_0; \cdot)$, $i = 1, 2$.

If $f_\square(x_0; \cdot)$ is not convex, it is still possible to define a notion of subgradient for f by means of upper convex approximates.

DEFINITION 6.19 (cf. [15]). The function $h: E \rightarrow \bar{\mathbf{R}}$ is an *upper convex approximate* for $f: E \rightarrow \bar{\mathbf{R}}$ at $x_0 \in \text{dom } f$ if

(a) $h(y) \leq f_\square(x_0; y)$ for all $y \in E$, and

(b) $h(\cdot)$ is convex and positively homogeneous. The set $\partial_h^k(x_0) := \{x' \in E^* \mid h(y) \leq \langle y, x' \rangle \forall y \in E\}$ is called an *h -subgradient* of f at x_0 .

Specific examples of upper convex approximates include $f^\uparrow(x_0; \cdot)$, $f^0(x_0; \cdot)$, the asymptotic epiderivatives of [9], [10], and the upper convex approximations of [15]. As one might expect, not as much can be said about general upper convex approximates as can be determined about specific ones like $f^0(x_0; \cdot)$ or $f^\uparrow(x_0; \cdot)$; however, necessary conditions for minimality can be expressed in terms of upper convex approximates [24], [25]. We will not develop this concept in detail here, except to mention the following generalizations of Propositions 4.1 and 4.2:

PROPOSITION 6.20. Suppose $f: E \rightarrow \bar{\mathbf{R}}$ has a local minimum at $x_0 \in E$ and $h(x_0; \cdot): E \rightarrow \bar{\mathbf{R}}$ is an upper convex approximate for f at x_0 . Then

$$0 \leq f_\square(x_0; y) \leq h(x_0; y) \quad \text{for all } y \in E$$

and

$$0 \in \partial_h^k f(x_0).$$

PROPOSITION 6.21. Let $f: \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be strictly l.s.c. at x_0 , and let $C \subset \mathbf{R}^n$ be closed near x_0 . Suppose $x_0 \in X$ is a local minimizer of the problem $\min \{f(x) \mid x \in C\}$. Assume (4.1) holds. Let $h(x_0; \cdot): \mathbf{R}^n \rightarrow \bar{\mathbf{R}}$ be an upper convex approximate for f at x_0 such that $h(x_0; y) \leq f^\uparrow(x_0; y)$ for all $y \in \mathbf{R}^n$, and let D be a convex subcone of $k_C(x_0)$ satisfying $T_C(x_0) \subset D \subset k_C(x_0)$. Then

$$(6.37) \quad 0 \in \partial_h^k f(x_0) + D^0.$$

Proof. $f + i_C$ is locally minimized at x_0 , so for all $y \in \mathbf{R}^n$,

$$\begin{aligned} 0 &\leq (f + i_C)_\square(x_0; y) \\ &\leq f_\square(x_0; y) + (i_C)_\square(x_0; y) \end{aligned}$$

(by Corollary 6.16, since (4.1) holds)

$$\begin{aligned} &= f_\square(x_0; y) + i_{k_C}(x_0)^{(y)} \\ &\leq h(x_0; y) + i_D(y). \end{aligned}$$

Now $\text{dom } h(x_0; \cdot) \supset \text{dom } f^\dagger(x_0; \cdot)$ and $D \subset T_C(x_0)$ by hypothesis, so (4.1) implies that $\text{dom } h(x_0; \cdot) - D = \mathbf{R}^n$. Hence

$$0 \in \partial(h(x_0; \cdot) + i_D(\cdot))(0) = \partial_h f(x_0) + D^0$$

by Theorem 23.8 of [16] and so (6.37) holds. $\square$

Acknowledgments. We are grateful to G. G. Watkins for his help with the proof of Theorem 4.7, and to the referee for suggestions which improved the presentation of this work.

REFERENCES

- [1] J.-P. AUBIN, *Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions*, in *Advances in Mathematics Supplementary Studies 7A*, L. Nachbin, ed., Academic Press, New York, 1981, pp. 159–229.
- [2] ———, *Lipschitz behavior of solutions to convex minimization problems*, *Math. Oper. Res.*, 9 (1984), pp. 87–111.
- [3] J.-P. AUBIN AND I. EKELAND, *Applied Nonlinear Analysis*, John Wiley, New York, 1984.
- [4] J. M. BORWEIN, *Stability and regular points of inequality systems*, *J. Optim. Theory Appl.*, 48 (1986), pp. 9–52.
- [5] F. H. CLARKE, *A new approach to Lagrange multipliers*, *Math. Oper. Res.*, 1 (1976), pp. 165–174.
- [6] ———, *Optimization and Nonsmooth Analysis*, John Wiley, New York, 1983.
- [7] S. DOLECKI, *Tangency and differentiation: some applications of convergence theory*, *Ann. Mat. Pura Appl.*, CXXX (1982), pp. 223–255.
- [8] I. EKELAND, *On the variational principle*, *J. Math. Anal. Appl.*, 47 (1974), pp. 324–353.
- [9] H. FRANKOWSA, *Inclusions adjointes associées aux trajectoires minimales d'inclusions différentielles*, *C. R. Acad. Sci. Paris Sér. A-B*, 296 (1983), pp. 721–724.
- [10] ———, *Necessary conditions for the Bolza problem*, *Math. Oper. Res.*, 10 (1985), pp. 361–366.
- [11] J.-B. HIRIART-URRUTY, *Contributions à la programmation mathématique: déterministe et stochastique*, Thèse, Université de Clermont-Ferrand II, 1977.
- [12] ———, *Tangent cones, generalized gradients and mathematical programming in Banach spaces*, *Math. Oper. Res.*, 4 (1979), pp. 79–97.
- [13] A. D. IOFFE, *Approximate subdifferentials and applications I: the finite-dimensional theory*, *Trans. Amer. Math. Soc.*, 281 (1984), pp. 389–416.
- [14] A. G. KUSRAEV, *On a general method of subdifferentiation*, *Soviet Math. Dokl.*, 23 (1981), pp. 367–371.
- [15] B. N. PSHENICHNYI AND R. A. KHACHATRYAN, *Constraints of equality type in nonsmooth optimization problems*, *Soviet Math. Dokl.*, 26 (1982), pp. 659–662.
- [16] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
- [17] ———, *Clarke's tangent cone and the boundaries of closed sets in $\mathbf{R}^n$* , *Nonlinear Anal. Theory Methods Appl.*, 3 (1979), pp. 145–154.
- [18] ———, *Directionally Lipschitzian functions and subdifferential calculus*, *Proc. London Math. Soc.*, 39 (1979), pp. 331–355.
- [19] ———, *Generalized directional derivatives and subgradients of nonconvex functions*, *Canad. J. Math.*, 32 (1980), pp. 257–280.
- [20] ———, *The Theory of Subgradients and its Applications to Problems of Optimization of Convex and Nonconvex Functions*, Heldermann Verlag, Berlin, 1981.
- [21] ———, *Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming*, *Math. Programming Stud.*, 17 (1982), pp. 28–66.
- [22] ———, *Extensions of subgradient calculus with applications to optimization*, *Nonlinear Anal. Theory Methods Appl.*, 9 (1985), pp. 665–698.
- [23] C. URSESCU, *Tangent sets' calculus and necessary conditions for extremality*, *this Journal*, 20 (1982), pp. 563–574.
- [24] D. E. WARD, *Tangent cones, generalized subdifferential calculus, and optimization*, Ph.D. thesis, Dalhousie University, Halifax, Nova Scotia, Canada, 1984.
- [25] ———, *Isotone tangent cones and nonsmooth optimization*, preprint, 1985.
- [26] G. G. WATKINS, *Nonsmooth Milyutin–Dubovitskii theory and Clarke's tangent cone*, *Math. Oper. Res.*, 11 (1986), pp. 70–80.
- [27] C. ZALINESCU, *On convex sets in general position*, *Linear Algebra Appl.*, 64 (1985), pp. 191–198.