JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **158**, 519–538 (1991)

Chain Rules for Nonsmooth Functions

DOUG WARD

Miami University, Oxford, Ohio 45056-1641

Submitted by Frank H. Clarke

Received March 6, 1989

A systematic method is presented for the derivation of chain rules for compositions of functions $F \circ f$, where F is nondecreasing. This method is valid for directional derivatives and subgradients associated with any tangent cone having a short list of properties. Some major special cases are examined in detail; in particular, calculus rules are derived for Rockafellar's epi-derivatives and Clarke generalized gradients.

1. INTRODUCTION

The concept of subgradient, originally defined for convex functions, has in recent years been generalized to arbitrary functions. (The history of the work in this area is summarized in the introduction to [14].) One important idea in this development is that of defining generalized directional derivatives via tangent cones to epigraphs of functions [3, 5, 11–14]. For the Clarke tangent cone, calculus rules for the directional derivative and its associated subgradient set were established by Rockafellar in [13] (see also [2, Sect. 2.9]). The cornerstone of this calculus is a "sum formula" [13, Theorem 2], from which a number of other calculus rules may be derived (e.g., [13, Theorems 3, 4]).

In this paper, we describe a systematic method of establishing chain rules for compositions $F \circ f$, where

(i) $f = (f_1, ..., f_n)$ and each $f_i \colon E \to \mathbb{R} \cup \{+\infty\}$ for E a real, locally convex, Hausdorff topological vector space (l.c.s.);

(ii) $F \colon \mathbb{R}^n \to \bar{\mathbb{R}}$ is nondecreasing with respect to the usual ordering on $\mathbb{R}^n$ (see Definition 2.1).

Special cases of such chain rules include sum formulas as well as calculus rules not discussed in [13]—product and quotient rules for positive-valued functions, for example. This method is valid not only for the directional derivative and subgradient associated with the Clarke tangent cone, but for

0022-247X/91 $3.00

those associated with any tangent cone possessing certain desirable properties. Combining techniques and ideas of [12, 16], it reduces the proof of a chain rule to that of a corresponding chain rule for tangent cones to graphs of compositions of set-valued mappings.

Here is a brief summary of this paper: After reviewing some necessary definitions and notation, we outline our method in Section 2. In Section 3, we apply it to the Ursescu [17] tangent cone. The resulting chain rule has applications in optimization theory [21] and contains much information about the epi-derivatives defined by Rockafellar in [15]. We conclude Section 3 with a discussion of the calculus of epi-derivatives. The special case of the Clarke tangent cone is covered in Section 4. The chain rule for Clarke generalized gradients essentially subsumes the calculus of [13].

Throughout the paper, we will let E be a l.c.s. with dual space E'. We denote the class of *neighborhoods* of $x \in E$ by $\mathcal{N}(x)$. For a nonempty set $C \subset E$, we will use int C for the *interior* of C. The *recession cone* of C is the set

$$0^+C := \{y \in E \mid x + ty \in C \; \forall t \geq 0, x \in C\}.$$

For a function $f\colon E \to \bar{\mathbb{R}}$, we denote the *epigraph* of f by epi f. By the *effective domain* of f, we mean the set

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

We say f is *proper* if dom f is nonempty and f never takes on the value $-\infty$. If f is convex and $f(x)$ is finite, $\partial f(x)$ will denote the *subgradient* of f at x.

By a *tangent cone*, we mean a mapping $A\colon 2^E \times E \to 2^E$ such that $A(C, x)$ is a (possibly empty) cone and $0^+C \subset 0^+A(C, x)$ for each nonempty $C \subset E$ and $x \in C$. We will say that A has a certain property if $A(C, x)$ has that property for every nonempty $C \subset E$ and $x \in C$. When C is an epigraph, it will be convenient to use the notation

$$A(f, x) := A(\operatorname{epi} f, (x, f(x))).$$

For a function $f\colon E \to \bar{\mathbb{R}}$ that is finite at x and a tangent cone A, we define the *A directional derivative* of f at x with respect to y by

$$f^A(x; y) := \inf\{r \mid (y, r) \in A(f, x)\}.$$

Note that the inclusion

$$A(f, x) \subset \operatorname{epi} f^A(x; \cdot)$$

always holds; if A is a closed tangent cone, then in fact

$$A(f, x) = \operatorname{epi} f^A(x; \cdot).$$

We also associate with a tangent cone A the A *subgradient* of f at x, defined by

$$\partial^A f(x) := \{x' \in E' \mid \langle x', y \rangle \leqslant f^A(x; y)\ \forall y \in E\}.$$

We observe that if A is convex, then $f^A(x; \cdot)$ is sublinear, so that $\partial^A f(x) = \partial f^A(x; \cdot)(0)$ whenever $f^A(x; 0) = 0$. For convex A, one can prove calculus rules for A by establishing the appropriate inequalities for f^A and using them along with the corresponding results from the subdifferential calculus for convex functions. That is the strategy followed in [13, 18, 20, 22], and we will have use for it here, too. For the chain rule in this paper, the relevant chain rule for convex functions is given in [7–9, 22]. We will review it in Section 2.

It will be convenient in the sequel to use some terminology from the theory of set-valued mappings. Specifically, if E and E^1 are l.c.s. and $H: E \to 2^{E^1}$, we define the *domain* of H to be the set

$$D(H) := \{x \in E \mid H(x) \neq \varnothing\},$$

and the *graph* of H to be the set

$$\operatorname{Gr}(H) := \{(x, y) \mid y \in H(x)\}.$$

The *inverse* of H is the mapping $H^{-1}: E^1 \to 2^E$ defined by

$$H^{-1}(y) := \{x \mid y \in H(x)\},$$

and the *range* of H is the set $R(H) := D(H^{-1})$. If E^2 is another l.c.s. and $G: E^1 \to 2^{E^2}$, the *composition* $G \circ H$ is defined by

$$(G \circ H)(x) := \{z \in E^2 \mid \exists y \in H(x) \cap G^{-1}(z)\}.$$

For tangent cones A and B and $y \in H(x) \cap G^{-1}(z)$, we will employ the notation

$$A(H; x, y) := A(\operatorname{Gr}(H), (x, y))$$

and

$$B(G; y, z) \circ A(H; x, y) := \{(d, r) \mid \exists s \text{ with } (d, s) \in A(H; x, y),\ (s, r) \in B(G; y, z)\}.$$

2. A Systematic Chain Rule Proof Technique

We begin this section with a more detailed discussion of the type of composition of functions considered here.

Definition 2.1. Let $x := (x_1, ..., x_n)$ and $y := (y_1, ..., y_n)$ be elements of $\mathbb{R}^n$. We say $x \leqslant y$ if $x_i \leqslant y_i$ for each i. The function $F\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ is said to be *isotone* on $S \subset \mathbb{R}^n$ if $F(x) \leqslant F(y)$ whenever $x, y \in S$ and $x \leqslant y$.

Definition 2.2. Let E be a l.c.s., and call

$$\bar{\mathbb{R}}^m := \{(x_1, ..., x_m) \mid x_i \in \bar{\mathbb{R}}\ \forall i\}.$$

For a function $f\colon E \to \bar{\mathbb{R}}^m$, we define the mapping $Mf\colon E \to 2^{\mathbb{R}^m}$ by

$$Mf(x) := \{y \in \mathbb{R}^m \mid f(x) \leqslant y\}.$$

We will prove chain rules for compositions of the form $F \circ f$, where $f(x) := (f_1(x), ..., f_n(x))$, each $f_i\colon E \to \mathbb{R} \cup \{+\infty\}$, and $F\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ is isotone on $R(Mf) \cup Y$ for some $Y \in \mathcal{N}(f(x))$. In such a composition, we must specify how $(F \circ f)(x)$ is to be interpreted if some $f_i(x) = +\infty$. We will adopt the convention here that $(F \circ f)(x) = +\infty$ whenever $f_i(x) = +\infty$ for one or more i.

Remark 2.3. Another ambiguity arises in interpreting the expression $(0f)(x)$ for $f\colon E \to \mathbb{R} \cup \{+\infty\}$. In keeping with the above convention, we define

$$(0f)(x) = \begin{cases} 0 & \text{if} \quad x \in \operatorname{dom} f \\ +\infty & \text{otherwise.} \end{cases}$$

Thus if f is convex and $x_0 \in \operatorname{dom} f$, we have

$$\partial(0f)(x_0) = \{x' \in E' \mid \langle x', x - x_0 \rangle \leqslant 0\ \forall x \in \operatorname{dom} f\}.$$

In order for the relationship $\partial(\lambda f)(x_0) = \lambda \partial f(x_0)$ to be valid for all $\lambda \geqslant 0$ whenever $\partial f(x_0) \neq \varnothing$, we will then adopt the additional convention that

$$0\partial f(x_0) := \{x' \in E' \mid \langle x', x - x_0 \rangle \leqslant 0\ \forall x \in \operatorname{dom} f\}.$$

Note that $0\partial f(x_0) = 0^+\partial f(x_0)$ whenever $\partial f(x_0)$ is nonempty.

Compositions of the above form have several noteworthy properties:

(i) If F is isotone on $R(Mf)$, the set $\operatorname{epi}(F \circ f)$ is exactly the same as $\operatorname{Gr}(MF \circ Mf)$.

(ii) If F is isotone on some $Y \in \mathcal{N}(f(x))$, the function $F^A(f(x); \cdot)$ is isotone on all of $\mathbb{R}^n$ for many tangent cones A [18, 21, 22].

(iii) If F and f are convex and F is isotone on $R(Mf)$, then $F \circ f$ is convex.

Properties (i) and (ii) will play a key role in the proofs of chain rules for directional derivatives given in this paper. Property (iii) makes possible corresponding chain rules for subgradients.

In the statement and proof of our directional derivative chain rules, we will use the notation

$$S_i := \{(x, y_1, ..., y_n) \in E \times \mathbb{R}^n \mid f_i(x) \leqslant y_i\}$$

for $i = 1, ..., n$. Observe that $\operatorname{epi} f = \operatorname{Gr}(Mf) = \bigcap_{i=1}^n S_i$.

Let $x_0 \in E$ be such that $f(x_0)$ and $F(f(x_0))$ are finite, and define $z_0 := (x_0, f(x_0))$. Our chain rules will be valid for tangent cones A, B, and A_i, $i = 1, ..., n$ that satisfy these inclusions:

$$\{(y, r_1, ..., r_n) \mid f_i^{A_i}(x_0; y) \leqslant r_i\} \subset A_i(S_i, z_0), \qquad i = 1, ..., n \tag{1}$$

$$\bigcap_{i=1}^n A_i(S_i, z_0) \subset A(Mf, z_0) \tag{2}$$

$$B(MF; f(x_0), F(f(x_0))) \circ A(Mf; z_0) \subset A(MF \circ Mf; x_0, F(f(x_0))). \tag{3}$$

In the derivation of these chain rules, the major step will usually be the verification of (3). Inclusion (2) is known to hold for a number of tangent cones under mild hypotheses on f [12, 18–22], and (1) is usually true for arbitrary f if the A_i are closed.

Our proof technique is summarized in the following theorem:

THEOREM 2.4. *Let f, F, and x_0 be defined as above, with F isotone on $R(Mf) \cup Y$ for some $Y \in \mathcal{N}(f(x_0))$. Suppose that A, B, and A_i, $i = 1, ..., n$ are closed tangent cones such that $F^B(f(x_0)); \cdot)$ is isotone on $\mathbb{R}^n$, each $f_i^{A_i}(x_0; \cdot)$ is proper, and inclusions* (1), (2), (3) *hold. Then for all $y \in E$,*

$$(F \circ f)^A\,(x_0; y) \leqslant F^B(f(x_0); f_1^{A_1}(x_0; y), ..., f_n^{A_n}(x_0; y)). \tag{4}$$

Proof. Set $\Omega := \{(y, r) \mid F^B(f(x_0); f_1^{A_1}(x_0; y), ..., f_n^{A_n}(x_0; y) \leqslant r\}$. Then inequality (4) is equivalent to the inclusion $\Omega \subset A(F \circ f, x_0)$. To prove this inclusion, we first observe that since $F^B(f(x_0); \cdot)$ is isotone on $\mathbb{R}^n$ and each $f_i^{A_i}(x_0; \cdot)$ is proper,

$$\Omega = \{(y, r) \mid \exists d \in \mathbb{R}^n \text{ with } F^B(f(x_0); d) \leqslant r, f_i^{A_i}(x_0; y) \leqslant d_i\, \forall i\}.$$

By (2) and (1),

$$\begin{aligned} A(Mf; z_0) &= A\left(\bigcap_{i=1}^{n} S_i, z_0\right) \\ &\supset \bigcap_{i=1}^{n} A_i(S_i, z_0) \\ &\supset \bigcap_{i=1}^{n} \{(y, r_1, ..., r_n) \mid f_i^{A_i}(x_0; y) \leqslant r_i\}. \end{aligned}$$

Hence

$$\begin{aligned} \Omega &\subset \{(y, r) \mid \exists d \in \mathbb{R}^n \text{ with } F^B(f(x_0); d) \leqslant r,\ (y, d) \in A(Mf; z_0)\} \\ &= B(MF; f(x_0), F(f(x_0))) \circ A(Mf; z_0) \\ &\subset A(MF \circ Mf; x_0, F(f(x_0))) \qquad \text{by (3)} \\ &= A(F \circ f, x_0), \end{aligned}$$

where the last equation follows from the fact that F is isotone on $R(Mf)$, as mentioned in (i). Therefore, (4) holds. ∎

Remark 2.5. (a) Some possible choices for F in Theorem 2.4 are $F(x_1, ..., x_n) := \sum x_i$; $F(x_1, ..., x_n) := \max(x_i)$; and for $x \geqslant 0$, $F(x_1, ..., x_n) := \prod x_i$.

(b) Theorem 2.4 outlines a step-by-step method for proving chain rules: Given tangent cones A, A_i, and B, simply find conditions under which (1), (2), and (3) are satisfied. Many nonsmooth calculus formulae in the literature, including those of [12, 13, 18, 20–22], can be derived by this method. We will discuss two examples in detail in Sections 3 and 4.

If the tangent cones A_i and B in Theorem 2.4 are convex, then chain rules for A subgradients can also be derived. The proofs of such chain rules depend upon inequality (4) and the following chain rule for convex functions:

THEOREM 2.6 (See [7–9, 22]). *Let $f_i: E \to \mathbb{R} \cup \{+\infty\}$, $i = 1, ..., n$ be finite at x_0 and convex; and let $F: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be finite at $f(x_0)$, convex, and isotone on $R(Mf)$. If*

$$R(Mf) \cap \operatorname{int} \operatorname{dom} F \neq \varnothing, \tag{5}$$

then

$$\partial(F \circ f)(x_0) = \left\{\partial\left(\sum \lambda_i f_i\right)(x_0) \mid (\lambda_1, ..., \lambda_n) \in \partial F(f(x_0))\right\}.$$

If in addition each f_i, $i=2, \ldots, n$ is continuous on the interior of its domain and

$$\operatorname{dom} f_1 \cap \bigcap_{i=2}^{n} \operatorname{int} \operatorname{dom} f_i \neq \varnothing, \tag{6}$$

then

$$\partial(F \circ f)(x_0) = \left\{ \sum \lambda_i \partial f_i(x_0) \mid (\lambda_1, \ldots, \lambda_n) \in \partial F(f(x_0)) \right\},$$

where $\lambda_i \partial f_i(x_0)$ is interpreted as in Remark 2.3 if $\lambda_i = 0$.

Theorem 2.6 can be applied to the convex functions $F^B(f(x_0); \cdot)$ and $f_i^{A_i}(x_0; \cdot)$ to produce chain rules for A subgradients:

THEOREM 2.7. *Let $f_i \colon E \to \mathbb{R} \cup \{\infty\}$, $i=1, \ldots, n$ be finite at x_0 and $F \colon \mathbb{R}^n \to \bar{\mathbb{R}}$ finite at $f(x_0)$. Suppose A, B, and A_i, $i=1, \ldots, n$ are closed tangent cones such that A_i and B are convex, each $f_i^{A_i}(x_0; \cdot)$ is proper and continuous on the interior of its domain, $F^B(f(x_0); \cdot)$ is isotone on $\mathbb{R}^n$, and* (4) *holds. Assume that*

$$\operatorname{dom} f_1^{A_1}(x_0; \cdot) \cap \bigcap_{i=2}^{n} \operatorname{int} \operatorname{dom} f_i^{A_i}(x_0; \cdot) \neq \varnothing; \tag{7}$$

and that for some $v \in E$ and $s_i \geqslant f_i^{A_i}(x_0; v)$, $i=1, \ldots, n$,

$$(s_1, \ldots, s_n) \in \operatorname{int} \operatorname{dom} F^B(f(x_0); \cdot). \tag{8}$$

Then

$$\partial^A(F \circ f)(x_0) \subset \left\{ \sum \lambda_i \partial^{A_i} f_i(x_0) \mid (\lambda_1, \ldots, \lambda_n) \in \partial^B F(f(x_0)) \right\}. \tag{9}$$

Proof. First observe that if $F^B(f(x_0); 0) = -\infty$, inequality (4) implies that both sides of (9) are empty. We may assume, then, that $F^B(f(x_0); \cdot)$ is proper. Assumptions (8) and (7) guarantee that (5) and (6) are satisfied with $F^B(f(x_0); \cdot)$ and $f_i^{A_i}(x_0; \cdot)$ playing the roles of F and f_i, respectively. We can now use Theorem 2.6 (with $x_0 := 0$) along with (4) to deduce (9), just as in [18, Sect. 4] and [22, Theorem 3.17]. ∎

As we will see in Theorems 3.6 and 4.10, conditions (7) and (8) are often exactly what is needed to guarantee inclusions (2) and (3) as well as (9).

3. Example 1: The Ursescu Tangent Cone

Definition 3.1. Let $C \subset E$ and $x \in C$. The *Ursescu tangent cone* [17] to C at x is the set

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

The Ursescu tangent cone is a closed tangent cone that is quite useful in optimization theory [10, 12, 17–22]. In this section, we examine the special case of Theorem 2.4, where $A = B = A_i := k$. We then discuss one corollary of the resulting chain rule, a chain rule for the "epi-derivatives" defined by Rockafellar in [15].

Unfortunately, (2) is not always true for $A = A_i := k$. For example, let $n = 2$ and $E = \mathbb{R}$, and consider the sets

$$\Gamma_1 := \{0\} \cup \{1/(2n+1) \mid n \text{ is an integer}\};$$
$$\Gamma_2 := \{0\} \cup \{1/2n \mid n \text{ is a nonzero integer}\}.$$

Define f_i to be the *indicator function* of Γ_i (i.e., the function that is equal to 0 on Γ_i and $+\infty$ elsewhere). Then if $x_0 := 0$, we have $k(S_i, z_0) = \{(x, y, z) \mid y \geqslant 0, z \geqslant 0\}$, while $k(Mf, z_0) = \{(0, y, z) \mid y \geqslant 0, z \geqslant 0\}$, so that (2) does not hold.

It is the case, however, that under mild conditions on the f_i, (2) will be satisfied with $A = A_i := k$. The following lemma will help us derive such conditions. Its hypotheses involve the convex tangent cones

$$k^\infty(C, x) := \{ y \mid k(C, x) + y \subset k(C, x) \}$$

and

$$Ik^\infty(C, x) := \{ y \mid \forall z \in k(C, x), \exists V \in \mathcal{N}(0) \text{ and } \lambda > 0 \text{ such that } \forall (t, v) \in (0, \lambda) \times V \text{ with } x + t(z + v) \in C, \; x + t(z+v) + t(y + V) \subset C \}.$$

(For more information on these tangent cones, see [10, 12, 18–21].)

Lemma 3.2 [12, 19]. *Let $C_i \subset E$, $i = 1, \ldots, n$, and $x \in \bigcap_{i=1}^n C_i$. If*

$$k^\infty(C_1, x) \cap \bigcap_{i=2}^{n} Ik^\infty(C_i, x) \neq \varnothing, \tag{10}$$

then

$$\bigcap_{i=1}^{n} k(C_i, x) = k\left(\bigcap_{1=1}^{n} C_i, x\right) \tag{11}$$

and

$$\bigcap_{i=1}^{n} k^{\infty}(C_i, x) \subset k^{\infty}\left(\bigcap_{1=1}^{n} C_i, x\right). \tag{12}$$

We next turn out attention to inclusion (3). The following example shows that (3) is not generally true for $B = A := k$.

EXAMPLE 3.3. Let $n = 1$, $E = \mathbb{R}$, and $f(x) := x^3$, and define $F(x)$ to be the indicator function of the set $\{x \mid x \leqslant 0\}$. Then with $x_0 := 0$, (3) does not hold, since $(1, 0) \in k(MF; 0, 0) \circ k(Mf; 0, 0)$ but $(1, 0) \notin k(MF \circ Mf; 0, 0)$.

It turns out that (3) is satisfied for $B = A := k$ under conditions rather similar to those given in (10). We will show this in two steps, adopting a technique of [12]. In the first step, we make use of the *interior Ursescu cone*, defined by

$$Ik(C, x) := \{y \mid \exists Y \in \mathcal{N}(y), \exists \lambda > 0 \text{ such that } x + (0, \lambda) Y \subset C\}.$$

We observe that Ik is an open tangent cone with $Ik \subset k$ [17]. More importantly, (3) is true in general for $B := Ik$, $A := k$, as we now demonstrate.

LEMMA 3.4. *Let $H: E \to 2^{E^1}$, $G: E^1 \to 2^{E^2}$, and $y \in H(x) \cap G^{-1}(z)$. Then*

$$Ik(G; y, z) \circ k(H; x, y) \subset k(G \circ H; x, z).$$

Proof. Suppose $(d, s) \in Ik(G; y, z) \circ k(H; x, y)$, and let $D \in \mathcal{N}(d)$, $S \in \mathcal{N}(s)$ be given. There exists $w \in E^1$ such that $(d, w) \in k(H; x, y)$ and $(w, s) \in Ik(G; y, z)$. By the definition of Ik, there exist $W \in \mathcal{N}(w)$, $S' \in \mathcal{N}(s)$ with $S' \subset S$, and $\lambda_1 > 0$ such that

$$(y, z) + (0, \lambda_1)(W \times S') \subset \operatorname{Gr} G.$$

By the definition of k, there exists $\lambda \in (0, \lambda_1)$ such that for all $t \in (0, \lambda)$,

$$[(x, y) + t(D \times W)] \cap \operatorname{Gr} H \neq \varnothing.$$

Thus for all $t \in (0, \lambda)$, there exist $d' \in D$ and $w' \in W$ such that for all $s' \in S'$,

$$(x + td', y + tw') \in \operatorname{Gr} H$$

and

$$(y+tw', z+ts') \in \mathrm{Gr}\, G.$$

It follows that $(x, z)+t(d', s') \in \mathrm{Gr}(G \circ H)$, and so $(d, s) \in k(G \circ H; x, z)$ and the proof is complete. ▮

In the second step, we build upon Lemma 3.4 to derive (3) for $B = A := k$ under conditions involving k^∞ and Ik^∞. In the proof, the inclusion

$$k(C, x) + Ik^\infty(C, x) \subset Ik(C, x), \tag{13}$$

which holds for all C and $x \in C$ [12, Proposition 4.6; 10, Lemma 2.6] will play a key role.

PROPOSITION 3.5. *Let $H: E \to 2^{E^1}$, $G: E^1 \to 2^{E^2}$, and $y \in H(x) \cap G^{-1}(z)$. If*

$$\mathrm{Gr}[Ik^\infty(G; y, z) \circ k^\infty(H; x, y)] \neq \varnothing, \tag{14}$$

then

$$k(G; y, z) \circ k(H; x, y) \subset k(G \circ H; x, z).$$

Proof. Suppose $(d, s) \in k(G; y, z) \circ k(H; x, y)$. Then there exists $w \in E^1$ such that $(d, w) \in k(H; x, y)$ and $(w, s) \in k(G; y, z)$. By hypothesis, there exist u, v, r such $(u, v) \in k^\infty(H; x, y)$ and $(v, r) \in Ik^\infty(G; y, z)$. Now for all $t > 0$, we have

$$(d, w) + t(u, v) \in k(H; x, y)$$

by the definition of k^∞. In addition,

$$(w, s) + t(v, r) \in Ik(G; y, z)$$

by (13) and the fact that Ik^∞ is a cone. It follows from Lemma 3.4 that $(d, s) + t(u, r) \in k(G \circ H; x, z)$ for all $t > 0$. Since k is closed, we conclude that $(d, s) \in k(G \circ H; x, z)$, and the proof is complete. ▮

We can apply Lemma 3.2 and Proposition 3.5 in the proof of a chain rule for the directional derivative associated with k.

THEOREM 3.6. *Let $f_i: E \to \mathbb{R} \cup \{\infty\}$, $i = 1, ..., n$ be finite at x_0 and $F: \mathbb{R}^n \to \bar{\mathbb{R}}$ finite at $f(x_0)$ and isotone on $R(Mf) \cup Y$ for some $Y \in \mathcal{N}(f(x_0))$. Suppose that $f_i^k(x_0; \cdot)$ is proper for each i; that*

$$\mathrm{dom}\, f_1^{k^\infty}(x_0; \cdot) \cap \bigcap_{i=2}^{n} \mathrm{dom}\, f_i^{Ik^\infty}(x_0; \cdot) \neq \varnothing; \tag{15}$$

and that for some $v \in E$ *and* $s_i \geqslant f_i^{k^\infty}(x_0; v)$, $i = 1, \ldots, n$,

$$(s_1, \ldots, s_n) \in \operatorname{dom} F^{Ik^\infty}(f(x_0); \cdot). \tag{16}$$

Then for all $y \in E$,

$$(F \circ f)^k (x_0; y) \leqslant F^k(f(x_0); f_1^k(x_0; y), \ldots, f_n^k(x_0; y)). \tag{17}$$

Proof. Let $B = A = A_i := k$ in Theorem 2.4. Since F is isotone on some $Y \in \mathcal{N}(f(x_0))$, it is easy to show, just as in [22, Lemma 2.8], that $F^k(f(x_0); \cdot)$ is isotone on $\mathbb{R}^n$. It is also simple to demonstrate that (1) holds with $A_i := k$ or $A_i := k^\infty$, and that

$$\{(y, r_1, \ldots, r_n) \mid f_i^{Ik^\infty}(x_0; y) < r_i\} \subset Ik^\infty(S_i, z_0). \tag{18}$$

By (15), there exists $(y, r) \in E \times \mathbb{R}$ such that $f_1^{k^\infty}(x_0; y) \leqslant r$, $f_i^{Ik^\infty}(x_0; y) < r$, $i = 2, \ldots, n$. It then follows from (1) and (18) that (10) holds with $C_i := S_i$ and $x := z_0$. Thus inclusion (2) holds for $A = A_i := k$ by Lemma 3.2.

It remains to verify inclusion (3) for $B = A := k$. By hypothesis, there exist $v \in E$ and $s = (s_1, \ldots, s_n) \in \mathbb{R}^n$ such that (16) is satisfied and $f_i^{k^\infty}(x_0; v) \leqslant s_i$. Now inclusion (1) with $A_i := k^\infty$ and inclusion (12) imply that

$$(v, s_1, \ldots, s_n) \in \bigcap_{i=1}^{n} k^\infty(S_i, z_0) \subset k^\infty(Mf, z_0);$$

and (16) implies that there exists $d \in \mathbb{R}$ such that $(s_1, \ldots, s_n, d) \in Ik^\infty(MF; f(x_0), F(f(x_0)))$. It follows that (14) holds with $G := MF$, $H := Mf$, $x := x_0$, $y := f(x_0)$, and $z := F(f(x_0))$. Therefore, by Proposition 3.5, (3) holds. We have now verified all the conditions of Theorem 2.4, so inequality (17) is true. ∎

Remark 3.7. (a) The hypotheses of Theorem 3.6 are satisfied in a variety of situations. For example, $\operatorname{dom} F^{Ik^\infty}(f(x_0); \cdot) = \mathbb{R}^n$ for each of the functions F mentioned in Remark 2.5(a), so that (16) is satisfied automatically in these cases whenever (15) holds. The condition $\operatorname{dom} f_i^{Ik^\infty}(x_0; \cdot) \neq \varnothing$ holds true for a class of functions that includes the directionally Liptschitzian functions of Rockafellar [1, 13–14]. A function that is not directionally Lipschitzian but satisfies $\operatorname{dom} f_i^{Ik^\infty}(x_0; \cdot) \neq \varnothing$ is given in [21, Example 2.1].

(b) The methods of this section can also be used to prove a chain rule for the case $A = A_1 := K$ (defined below), $B = A_i := k$, $i = 2, \ldots, n$.

(c) The inequality in (17) may be strict, as shown by the example given after Proposition 6.1 of [22].

When $n=1$, Theorem 3.6 reduces to the following result:

COROLLARY 3.8. *Let $f: E \to \mathbb{R} \cup \{\infty\}$ be finite at x_0 and $F: \mathbb{R} \to \bar{\mathbb{R}}$ finite at $f(x_0)$ and isotone on $R(Mf) \cup Y$ for some $Y \in \mathcal{N}(f(x_0))$. Suppose that $f^k(x_0; \cdot)$ is proper and*

$$R(Mf^{k^\infty}(x_0; \cdot)) \cap \operatorname{dom} F^{lk^\infty}(f(x_0); \cdot) \neq \varnothing. \tag{19}$$

Then for all $y \in E$,

$$(F \circ f)^k\,(x_0; y) \leqslant F^k(f(x_0); f^k(x_0; y)). \tag{20}$$

Proof. Let $n=1$ in Theorem 3.6. Then (15) is simply the assumption that $\operatorname{dom} f^{k^\infty}(x_0; \cdot) \neq \varnothing$. Since $0 \in k^\infty(C, x)$ for all C and $x \in C$, $0 \in \operatorname{dom} f^{k^\infty}(x_0; \cdot)$ is true in general, and so $\operatorname{dom} f^{k^\infty}(x_0; \cdot) \neq \varnothing$. Condition (16) reduces to (19) in this case, and hence (20) follows from Theorem 3.6. ∎

Another interesting corollary of Theorem 3.6 is a chain rule for the epi-differentiable functions defined by Rockafellar in [15].

DEFINITION 3.9. (a) Let $C \in E$ and $x \in C$. The *contingent cone* to C at x is the set

$$K(C, x) := \{y \mid \forall Y \in \mathcal{N}(y), \forall \lambda > 0, [x + (0, \lambda) Y] \cap C \neq \varnothing\}.$$

(b) Let $f: E \to \bar{\mathbb{R}}$ be finite at x. Then f is said to be *epi-differentiable* at x if $f^K(x; \cdot)$ is proper and $f^k(x; \cdot) = f^K(x; \cdot)$. For such functions, $f^k(x; \cdot)$ is called the *epi-derivative* of f at x and is denoted $f'_x(\cdot)$.

This definition of epi-differentiability is different from, but equivalent to, that given in [15]. The equivalence of the two definitions is clear from [15, (2.13)] and the fact [3, 19] that the k and K directional derivatives of any function f can be written as

$$f^k(x; y) = \sup_{Y \in \mathcal{N}(y)} \inf_{\lambda > 0} \sup_{t \in (0, \lambda)} \inf_{v \in Y} (f(x + tv) - f(x))/t;$$

$$f^K(x; y) = \sup_{Y \in \mathcal{N}(y)} \sup_{\lambda > 0} \inf_{t \in (0, \lambda)} \inf_{v \in Y} (f(x + tv) - f(x))/t.$$

Rockafellar shows in [15, Sect. 2] that the types of functions usually studied in applications to optimization are epi-differentiable. For example, Frechet differentiable functions and convex functions are epi-differentiable. The following chain rule will give us more information about this class of functions.

COROLLARY 3.10. *Let* $f_i\colon E \to \mathbb{R} \cup \{\infty\}$, $i = 1, ..., n$ *be epi-differentiable at* x, *and let* $F\colon \mathbb{R}^n \to \bar{\mathbb{R}}$ *be epi-differentiable at* $f(x)$ *and isotone on* $R(Mf) \cup Y$ *for some* $Y \in \mathcal{N}(f(x))$. *Suppose that* (15) *and* (16) *hold with* $x_0 := x$, *and that* $(F \circ f)^K(x; y) = +\infty$ *whenever some* $f_i^K(x; y) = +\infty$. *Then* $F \circ f$ *is epi-differentiable at* x *and for all* $y \in E$,

$$(F \circ f)'_x(y) = F'_x(f(x); (f_1)'_x(y), (f_2)'_x(y), ..., (f_n)'_x(y)). \tag{21}$$

Proof. Let $y \in E$. If each $f_i^K(x; y)$ is finite, then by [22, Proposition 6.1],

$$(F \circ f)^K(x; y) \geqslant F^K(f(x); f_1^K(x; y), ..., f_n^K(x; y)). \tag{22}$$

If on the other hand some $f_i^K(x; y) = +\infty$, then both sides of (22) are equal to $+\infty$ by hypothesis, and the inequality is still valid.

Since $k \subset K$, it follows that $(F \circ f)^k(x; \cdot) \geqslant (F \circ f)^K(x; \cdot)$. Thus by (22) and Theorem 3.6, we have

$$\begin{aligned} F^k(f(x); f_1^k(x; \cdot), ..., f_n^k(x; \cdot)) &\geqslant (F \circ f)^k(x; \cdot) \\ &\geqslant (F \circ f)^K(x; \cdot) \\ &\geqslant F^K(f(x); f_1^K(x; \cdot), ..., f_n^K(x; \cdot)). \end{aligned}$$

Finally, the hypothesis that F is epi-differentiable at $f(x)$ and each f_i is epi-differentiable at x implies that

$$F^K(f(x); f_1^K(x; \cdot), ..., f_n^K(x; \cdot)) = F^k(f(x); f_1^k(x; \cdot), ..., f_n^k(x; \cdot)).$$

Therefore, $F \circ f$ is epi-differentiable at x and (21) holds. ∎

Corollary 3.10 shows that a finite sum of epi-differentiable functions is epi-differentiable if condition (15) is satisfied. Specifically, we have the following result:

COROLLARY 3.11. *Let* $f_i\colon E \to \mathbb{R} \cup \{\infty\}$, $i = 1, ..., n$ *be epi-differentiable at* x, *and suppose that* (15) *holds with* $x_0 := x$. *Then* $\sum f_i$ *is epi-differentiable at* x *and for all* $y \in E$,

$$\left(\sum f_i\right)'_x(y) = \sum (f_i)'_x(y). \tag{23}$$

Proof. Let $F\colon \mathbb{R}^n \to \mathbb{R}$ be defined by $F(x_1, ..., x_n) = \sum x_i$. Then F is epi-differentiable at $f(x)$, and since $\operatorname{dom} F'^{k^\infty}(f(x); \cdot) = \mathbb{R}^n$, (16) holds with $x_0 := x$. Since each $f_i^K(x; \cdot)$ is proper, it is easy to show that if some $f_i^K(x; y) = +\infty$, then $(\sum f_i)^K(x; y) = +\infty$ also. The assertion then follows from Corollary 3.10. ∎

Remark 3.12. (a) If condition (15) is not satisfied, Eq. (23) may not be true. For example, let $n = 2$ and $E = \mathbb{R}$, and let f_1 and f_2 be the indicator functions of the sets Γ_1 and Γ_2 defined at the beginning of this section. Then $f_i^K(0; y) = f_i^k(0; y) = 0$ for all $y \in \mathbb{R}$, but $(f_1 + f_2)^K(0; y) = (f_1 + f_2)^k(0; y) = +\infty$ for all nonzero y.

(b) It is also possible that a sum of epi-differentiable functions will not be epi-differentiable if (15) is not satisfied. For example, let f_i, $i = 1, 2$, be the indicator functions of the sets $\Gamma_i \cup \{\pm 2^{-n} \mid n \text{ is an integer}\}$. Then $f_i^K(0; y) = f_i^K(0; y) = (f_1 + f_2)^K(0; y) = 0$ for all $y \in \mathbb{R}$, but $(f_1 + f_2)^k(0; y) = +\infty$ for all nonzero y, so that $f_1 + f_2$ is not epi-differentiable at 0.

(c) If E is finite-dimensional, then (21) can be derived under different hypotheses that are neither weaker nor stronger than those of Corollary 3.10 (see [22, Theorem 6.13]).

(d) We may also deduce from Corollary 3.10 the fact that the pointwise maximum of a finite number of continuous epi-differentiable functions is epi-differentiable if condition (15) is satisfied.

4. Example 2: The Clarke Tangent Cone

Definition 4.1 [13, 14]. Let $C \subset E$ and $x \in C$. The *Clarke tangent cone* to C at x is the set

$$\mathrm{T}(C, x) := \{y \in E \mid \forall Y \in \mathcal{N}(y), \exists X \in \mathcal{N}(x) \text{ and } \lambda > 0 \text{ such that } \forall v \in X \cap C \text{ and } t \in (0, \lambda), (v + tY) \cap C \neq \varnothing\}.$$

The Clarke tangent cone, a closed convex tangent cone, plays a key role in nonsmooth analysis and optimization [2]. The calculus of its associated directional derivative (often denoted $f^{\uparrow}$) and subgradient have been discussed in [13, 1, 2, 8, 18, 22]. In this section, we derive a chain rule for this directional derivative and subgradient via Theorems 2.4 and 2.7.

As is the case for the Ursescu tangent cone, inclusion (2) is not always satisfied for $A = A_i := T$ (see the examples in [22, Remark 3.7(b), (c)]). Inclusion (2) is valid under conditions involving the *interior Clarke tangent cone*

$$\mathrm{IT}(C, x) := \{y \in E \mid \exists Y \in \mathcal{N}(y), X \in \mathcal{N}(x), \lambda > 0 \text{ such that } (X \cap C) + (0, \lambda) Y \subset C\}.$$

Such conditions can be developed via the following analogue of Lemma 3.2:

LEMMA 4.2 [13]. *Let* $C_i \subset E$, $i = 1, \ldots, n$, *and* $x \in \bigcap_{i=1}^{n} C_i$. *If*

$$\mathrm{T}(C_1, x) \cap \bigcap_{i=2}^{n} \mathrm{IT}(C_i, x) \neq \varnothing, \tag{24}$$

then

$$\bigcap_{i=1}^{n} \mathrm{T}(C_i, x) \subset \mathrm{T}\left(\bigcap_{1=1}^{n} C_i, x\right). \tag{25}$$

Remark 4.3. If $\mathrm{IT}(C, x) \neq \varnothing$, the set C is said to be *epi-Lipschitzian* at x [1–2, 13–14]. If C is epi-Lipschitzian at x, then $\mathrm{IT}(C, x) = \operatorname{int} \mathrm{T}(C, x)$. This follows directly from the general inclusion

$$\mathrm{IT}(C, x) + \mathrm{T}(C, x) \subset \mathrm{IT}(C, x), \tag{26}$$

which is derived in [14, 12]. If C_i, $i = 2, \ldots, n$ are epi-Lipschitzian at x, (24) is then equivalent to

$$\mathrm{T}(C_1, x) \cap \bigcap_{i=2}^{n} \operatorname{int} \mathrm{T}(C_i, x) \neq \varnothing.$$

We next consider inclusion (3). Example 3.3 shows that (3) is not generally true for $A = B := T$. As in Section 3, though, we can give a two-stage proof of this inclusion under certain conditions. One of these conditions was identified by Thibault in [16]:

DEFINITION 4.4 [16]. The relation $G: E \to 2^{E^1}$ is said to be *lower semi-continuous* at $(x, y) \in \operatorname{Gr} G$ relative to a set $\Omega \subset E$ containing x if for each $Y \in \mathcal{N}(y)$, there exists $X \in \mathcal{N}(x)$ such that $Y \cap G(x') \neq \varnothing$ for all $x' \in \Omega \cap X$.

The first stage of the proof of (3) for $A = B := T$ is an analogue of Lemma 3.4.

LEMMA 4.5 (Cf. [16, Proposition 4.2]). *Let* $H: E \to 2^{E^1}$, $G: E^1 \to 2^{E^2}$, *and* $y \in H(x) \cap G^{-1}(z)$. *If the mapping* $(u, v) \mapsto H(u) \cap G^{-1}(v)$ *is lower semi-continuous at* $((x, z), y)$ *relative to* $\operatorname{Gr}(G \circ H)$, *then*

$$\mathrm{IT}(G; y, z) \circ \mathrm{T}(H; x, y) \subset \mathrm{T}(G \circ H; x, z).$$

Proof. Suppose $(d, s) \in \mathrm{IT}(G; y, z) \circ \mathrm{T}(H; x, y)$, and let $D \in \mathcal{N}(d)$, $S \in \mathcal{N}(s)$ be given. There exists $w \in E^1$ such that $(d, w) \in \mathrm{T}(H; x, y)$ and $(w, s) \in \mathrm{IT}(G; y, z)$. By the definition of IT, there exist $Y' \in \mathcal{N}(y)$, $Z' \in \mathcal{N}(z)$, $W \in \mathcal{N}(w)$, $S' \in \mathcal{N}(s)$ and $\lambda_1 > 0$ such that $S' \subset S$ and

$$(Y' \times Z') + (0, \lambda_1)(W \times S') \subset \operatorname{Gr} G. \tag{27}$$

Similarly, the definition of T implies that there exist $X' \in \mathcal{N}(x)$, $Y \in \mathcal{N}(y)$ with $Y \subset Y'$, and $\lambda \in (0, \lambda_1)$ such that for all $(x', y') \in (X' \times Y) \cap \operatorname{Gr} H$ and $t \in (0, \lambda)$, there exists $(d', w') \in D \times W$ with

$$(x', y') + t(d', w') \in \operatorname{Gr} H. \tag{28}$$

By hypothesis, there exist $X \in \mathcal{N}(x)$, $Z \in \mathcal{N}(z)$ such that $X \subset X'$, $Z \subset Z'$ and for all $(x', z') \in (X \times Z) \cap \operatorname{Gr}(G \circ H)$,

$$Y \cap H(x') \cap G^{-1}(z') \neq \varnothing. \tag{29}$$

Now suppose $(x', z') \in (X \times Z) \cap \operatorname{Gr}(G \circ H)$ and $t \in (0, \lambda)$. By (29), there exists $y' \in Y$ such that $(x', y') \in \operatorname{Gr} H$, $(y', z') \in \operatorname{Gr} G$. It follows from (28) that there exists $(d', w') \in D \times W$ with $(x', y') + t(d', w') \in \operatorname{Gr} H$, and by (27), $(y', z') + t(w', s) \in \operatorname{Gr} G$. Hence, $(x', z') + t(d', s) \in \operatorname{Gr}(G \circ H)$, and so $(d, s) \in T(G \circ H; x, z)$. ▮

The second stage parallels Proposition 3.5. In fact, one can prove the following result by simply replacing k and k^{∞} with T, Ik and Ik^{∞} with IT, and (13) with (26) in the proof of Proposition 3.5.

PROPOSITION 4.6 (Cf. [8, Theorem 3.2.2; 16, Corollary 4.5]). *Let* $H: E \to 2^{E^1}$, $G: E^1 \to 2^{E^2}$, *and* $y \in H(x) \cap G^{-1}(z)$. *If the mapping* $(u, v) \mapsto H(u) \cap G^{-1}(v)$ *is lower semicontinuous at* $((x, z), y)$ *relative to* $\operatorname{Gr}(G \circ H)$ *and*

$$\operatorname{Gr}[\mathrm{IT}(G; y, z) \circ \mathrm{T}(H; x, y)] \neq \varnothing, \tag{30}$$

then

$$\mathrm{T}(G; y, z) \circ \mathrm{T}(H; x, y) \subset \mathrm{T}(G \circ H; x, z).$$

To establish (3) for $A = B := T$, it remains for us to find conditions under which the mapping $(x, z) \mapsto Mf(x) \cap (MF)^{-1}(z)$ is lower semicontinuous relative to $\operatorname{Gr}(MF \circ Mf)$. We begin with some definitions.

DEFINITION 4.7. The function $g: E \to \mathbb{R} \cup \{+\infty\}$ is said to be *lower semicontinuous* (l.s.c.) at x_0 if for every $\varepsilon > 0$, there exists $X \in \mathcal{N}(x_0)$ such that $f(x) \geqslant f(x_0) - \varepsilon$ for all $x \in X$.

DEFINITION 4.8. The function $F: \mathbb{R}^n \to \overline{\mathbb{R}}$ is said to be *strictly isotone* in the ith coordinate at $x := (x_1, ..., x_n)$ if $F(x) < F(y)$ whenever $x \leqslant y$ and $x_i < y_i$.

LEMMA 4.9 (Cf. [22, Lemma 3.15]). *Let* $f_i: E \to \mathbb{R} \cup \{+\infty\}$, $i = 1, ..., n$

be finite and l.s.c. at x_0, *and let* $F: \mathbb{R}^n \to \bar{\mathbb{R}}$ *be finite at* $f(x_0)$ *and isotone and l.s.c. on* $\prod_{i=1}^n (f_i(x_0) - \delta, +\infty)$ *for some* $\delta > 0$. *Suppose for each* $i \in \{1, ..., n\}$ *that either*

(i) f_i *is continuous at* x_0; *or*

(ii) F *is strictly isotone in the* ith *coordinate at* $f(x_0)$.

Then the mapping $(x, z) \mapsto Mf(x) \cap (MF)^{-1}(z)$ *is lower semicontinuous at* $((x_0, F(f(x_0))), f(x_0))$ *relative to* $\mathrm{Gr}(MF \circ Mf)$.

Proof. Let $\varepsilon > 0$ be given, and call $Y := \prod_{i=1}^n (f_i(x_0) - \varepsilon, f_i(x_0) + \varepsilon)$. Let $I \subset \{1, ..., n\}$ be the set of coordinates in which F is strictly isotone at $f(x_0)$. For each $i \in I$, the fact that F is l.s.c. implies, as in the proof of [22, Lemma 3.15], that there exists $\mu_i \in (0, \min(\delta, \varepsilon/2))$ such that

$$\alpha_i := F(f_1(x_0) - \mu_i, ..., f_i(x_0) + \varepsilon/2, ..., f_n(x_0) - \mu_i) - F(f(x_0)) > 0.$$

Define $\mu := \min_I \mu_i$, $\alpha := \min_I \alpha_i/2$ if I is nonempty and $\mu = \alpha := \varepsilon/2$ otherwise. Call $Z := (F(f(x_0)) - \alpha, F(f(x_0)) + \alpha)$. By hypothesis, there exists $X \in \mathcal{N}(x_0)$ such that for all $x \in X$,

$$\begin{aligned} &f_i(x_0) - \mu \leqslant f_i(x) && \text{for all} \quad i \in I; \\ &f_i(x_0) - \mu \leqslant f_i(x) \leqslant f_i(x_0) + \mu && \text{for all} \quad i \in \{1, ..., n\} \setminus I. \end{aligned}$$

Now suppose $(x, z) \in (X \times Z) \cap \mathrm{Gr}(MF \circ Mf)$. Then there exists $y := (y_1, ..., y_n)$ such that $f(x) \leqslant y$ and $F(y) \leqslant z$. If $y_i \geqslant f_i(x_0) + \varepsilon$ for any $i \in I$, then the isotonicity of F implies that $F(y) > F(f(x_0)) + \alpha$, contradicting the fact that $F(y) \leqslant z$. Thus $y \in Y$, and it follows that the mapping $(x, z) \mapsto Mf(x) \cap (MF)^{-1}(z)$ is lower semicontinuous at $((x_0, F(f(x_0))), f(x_0))$ relative to $\mathrm{Gr}(MF \circ Mf)$. ∎

The stage is now set for the proof of the following chain rule:

THEOREM 4.10. *Let* $f_i: E \to \mathbb{R} \cup \{\infty\}$, $i = 1, ..., n$ *and* $F: \mathbb{R}^n \to \bar{\mathbb{R}}$ *satisfy the hypotheses of Lemma* 4.9. *Suppose that* $f_i^{\mathrm{T}}(x_0; \cdot)$ *is proper for each* i; *that*

$$\operatorname{dom} f_1^{\mathrm{T}}(x_0; \cdot) \cap \bigcap_{i=2}^{n} \operatorname{dom} f_i^{\mathrm{IT}}(x_0; \cdot) \neq \varnothing; \tag{31}$$

and that for some $v \in E$ *and* $s_i \geqslant f_i^{\mathrm{T}}(x_0; v)$, $i = 1, ..., n$,

$$(s_1, ..., s_n) \in \operatorname{dom} F^{\mathrm{IT}}(f(x_0); \cdot). \tag{32}$$

Then for all $y \in E$,

$$(F \circ f)^{\mathrm{T}}(x_0; y) \leqslant F^{\mathrm{T}}(f(x_0); f_1^{\mathrm{T}}(x_0; y), ..., f_n^{\mathrm{T}}(x_0; y)). \tag{33}$$

Moreover,

$$\partial^{\mathrm{T}}(F\circ f)(x_0)\subset\left\{\sum\lambda_i\partial^{\mathrm{T}}f_i(x_0)\mid(\lambda_1,...,\lambda_n)\in\partial^{\mathrm{T}}F(f(x_0))\right\}. \tag{34}$$

Proof. Let $B=A=A_i:=T$ in Theorem 2.4. Since F is isotone on a neighborhood of $f(x_0)$, $F^{\mathrm{T}}(f(x_0);\cdot)$ is isotone on $\mathbb{R}^n$ by [22, Lemma 2.8]. It is easy to show that (1) holds with $A_i:=T$, and that

$$\{(y,r_1,...,r_n)\mid f_i^{\mathrm{IT}}(x_0;y)<r_i\}\subset\mathrm{IT}(S_i,z_0). \tag{35}$$

By (31), there exists $(y,r)\in E\times\mathbb{R}$ such that $f_1^{\mathrm{T}}(x_0;y)\leqslant r$, $f_i^{\mathrm{IT}}(x_0;y)<r$, $i=2,...,n$. It then follows from (1) and (35) that (24) holds with $C_i:=S_i$ and $x:=z_0$. Thus inclusion (2) holds for $A=A_i:=T$ by Lemma 4.2.

Next verify inclusion (3) for $B=A:=T$. By hypothesis, there exist $v\in E$ and $s=(s_1,...,s_n)\in\mathbb{R}^n$ such that (32) is satisfied and $f_i^{\mathrm{T}}(x_0;v)\leqslant s_i$. Now inclusion (1) with $A_i:=T$ and inclusion (25) imply that

$$(v,s_1,...,s_n)\in\bigcap_{i=1}^{n}\mathrm{T}(S_i,z_0)\subset\mathrm{T}(Mf,z_0);$$

and (32) implies that there exists $d\in\mathbb{R}$ such that $(s_1,...,s_n,d)\in\mathrm{IT}(MF;f(x_0),F(f(x_0)))$. It follows that (30) holds with $G:=MF$, $H:=Mf$, $x:=x_0$, $y:=f(x_0)$, and $z:=F(f(x_0))$. Therefore, by Proposition 4.6 and Lemma 4.9, (3) holds. We have now verified all the conditions of Theorem 2.4, so inequality (33) is true.

To prove (34), we first note that $\operatorname{dom}F^{\mathrm{IT}}(f(x_0);\cdot)\neq\varnothing$ and $\operatorname{dom}f_i^{\mathrm{IT}}(x_0;\cdot)\neq\varnothing$ for $i=2,...,n$ by (32) and (31). It is shown in [14, Theorem 3] that $f_i^{\mathrm{T}}(x_0;\cdot)$ is then continuous on the interior of its domain and that $\operatorname{dom}F^{\mathrm{IT}}(f(x_0;\cdot)=\operatorname{int}\operatorname{dom}F^{\mathrm{T}}(f(x_0);\cdot)$, $\operatorname{dom}f_i^{\mathrm{IT}}(x_0;\cdot)=\operatorname{int}\operatorname{dom}f_i^{\mathrm{T}}(x_0;\cdot)$ for $i=2,...,n$. Inclusion (34) thus follows from Theorem 2.7. ∎

Remark 4.11. (a) Conditions guaranteeing equality in (33) and (34) are given in [22, Proposition 6.4].

(b) Assumptions (31) and (32) can be weakened if E is finite dimensional [22, Sect. 3]. One could rederive the chain rule of [22, Theorem 3.27] via Theorem 2.4 and the appropriate analogues of Lemma 4.2 and Proposition 4.6.

(c) Kusraev and Kutateladze state a chain rule for $\partial^{\mathrm{T}}(F\circ f)$ in the more general situation in which f takes values in an ordered l.c.s. [8, Theorem 3.2.4]. In this chain rule, f is required to be required to be continuous at x_0. In the simpler setting of Theorem 4.10, we need only assume (thanks to Lemma 4.9) that f is l.s.c. at x_0.

Theorem 4.10 essentially subsumes the calculus of [13]. However, it does not encompass all that is known about the calulus of $\partial^{\mathrm{T}}(F \circ f)$. If F and f are locally Lipschitzian, chain rules requiring no isotonicity assumption on F are valid [2, 5]. The proofs of such chain rules rely upon the mean value theorem of Lebourg (see [2, Chap. 2]) and the upper semicontinuity of ∂^{T}, tools that are not available for general F and f [6].

5. Conclusion

Theorems 2.4 and 2.7 provide a systematic method for the derivation of directional derivative and subgradient calculus formulas for non-Lipschitzian functions. New formulas can be discovered and known ones rediscovered by this method.

Acknowledgment

The author is grateful for the referee's helpful and insightful comments.

References

1. J. M. Borwein and H. M. Strojwas, Directionally Lipschitzian mappings on Baire spaces, *Canad. J. Math.* **36** (1984), 95–130.
2. F. H. Clarke, "Optimization and Nonsmooth Analysis," Wiley, New York, 1983.
3. S. Dolecki, Tangency and differentiation: Some applications of convergence theory, *Ann. Mat. Pura, Appl.* **130** (1982), 223–255.
4. I. Ekeland and R. Temam, "Convex Analysis and Variational Problems," North-Holland, Amsterdam, 1976.
5. J.-B. Hiriart-Urruty, New concepts in nondifferentiable programming, *Bull. Soc. Math. France Mémoire* **60** (1979), 57–85.
6. J.-B. Hiriart-Urruty, Mean value theorems in nonsmooth analysis, *Numer. Funct. Anal. Optim.* **2** (1980), 1–30.
7. A. D. Ioffe and V. L. Levin, Subdifferentials of convex functions, *Trans. Moscow Math. Soc.* **26** (1972), 1–72.
8. A. G. Kusraev and S. S. Kutateladze, Local convex analysis, *J. Soviet Math.* **19** (1984), 2048–2087.
9. B. Lemaire, Application of a subdifferential of a convex composite functional to optimal control in variational inequalities, *in* "Nondifferentiable Optimization: Motivations and Applications" (V. F. Demyanov and D. Pallaschke, Eds.), Springer-Verlag, Berlin, 1985.
10. R. R. Merkovsky and D. E. Ward, Upper D.S.L. approximates and nonsmooth optimization, *Optimization* **21** (1990), 163–177.
11. J.-P. Penot, Calcul sous-differentiel et optimisation, *J. Funct. Anal.* **27** (1978), 248–276.
12. J.-P. Penot, Variations on the theme of nonsmooth analysis: another subdifferential, *in* "Nondifferentiable Optimization: Motivations and Applications" (V. F. Demyanov and D. Pallaschke, Eds.), Springer-Verlag, Berlin, 1985.

13. R. T. Rockafellar, Directionally Lipschitzian functions and subdifferential calculus, *Proc. London Math. Soc.* **39** (1979), 331–355.
14. R. T. Rockafellar, Generalized directional derivatives and subgradients of nonconvex functions, *Canad. J. Math.* **32** (1980), 157–180.
15. R. T. Rockafellar, First- and second-order epi-differentiability in nonlinear programming, *Trans. Amer. Math. Soc.* **307** (1988), 75–108.
16. L. Thibault, Tangent cones and quasi-interiorly tangent cones to multifunctions, *Trans. Amer. Math. Soc.* **277** (1983), 601–621.
17. C. Ursescu, Tangent sets' calculus and necessary conditions for extremality, *SIAM J. Control Optim.* **20** (1982), 563–574.
18. D. E. Ward, Convex subcones of the contingent cone in nonsmooth calculus and optimization, *Trans. Amer. Math. Soc.* **302** (1987), 661–682.
19. D. E. Ward, Isotone tangent cones and nonsmooth optimization, *Optimization* **18** (1987), 769–783.
20. D. E. Ward, Which subgradients have sum formulas?, *Nonlinear Anal.* **12** (1988), 1231–1243.
21. D. E. Ward, Directional derivative calculus and optimality conditions in nonsmooth mathematical programming, *J. Inform. Optim. Sci.* **10** (1989), 81–96.
22. D. E. Ward and J. M. Borwein, Nonsmooth calculus in finite dimensions, *SIAM. J. Control Optim.* **25** (1987), 1312–1340.