Relating lexicographic smoothness and directed subdifferentiability

Kamil A. Khan

Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, USA

Abstract

Lexicographic derivatives developed by Nesterov and directed subdifferentials developed by Baier, Farkhi, and Roshchina are both essentially nonconvex generalized derivatives for nonsmooth nonconvex functions and satisfy strict calculus rules and mean-value theorems. This article aims to clarify the relationship between the two generalized derivatives. In particular, for scalar-valued functions that are locally Lipschitz continuous, lexicographic smoothness and directed subdifferentiability are shown to be equivalent, along with the necessary optimality conditions corresponding to each. For such functions, the visualization of the directed subdifferential—the Rubinov subdifferential—is shown to include the lexicographic subdifferential, and is also shown to be included in its convex hull. Implications of these results are discussed.

Keywords: nonsmooth analysis, generalized differentiation, directional derivative

2000 MSC: 90C31, 26B05, 49K10

1. Introduction

Several set-valued *generalized derivatives* have been developed for locally Lipschitz continuous functions $f : X \subset \mathbb{R}^n \to \mathbb{R}^m$ that are not differentiable everywhere, for use in methods for equation solving and optimization. These generalized derivatives include Clarke's generalized Jacobian [1] and the various generalized subdifferentials described in [2, 3, 4, 5, 6]. As summarized

Email address: `kkhan@anl.gov` (Kamil A. Khan)

by Mordukhovich [3], Clarke's generalized Jacobian is the smallest convex-valued generalized derivative that satisfies certain desirable properties and an inclusion-based sum rule; roughly, $\partial[f+g](x) \subset \partial f(x)+\partial g(x)$. Notably, two types of nonconvex generalized derivative have been proposed that satisfy calculus rules as equalities instead: the *lexicographic subdifferential* developed by Nesterov [7, 8] and the *directed* and *Rubinov subdifferentials* developed by Baier, Farkhi, and Roshchina [9, 10]. This article aims to clarify the relationship between these two approaches to constructing useful generalized derivatives and thereby to obtain new properties of each.

Intuition suggests that such a relationship ought to exist. Inspection of the directed subdifferential shows that it is motivated similarly to the *lexicographic-directional (LD-)derivative* described by Khan and Barton [11]; each is defined through the recursive application of directional derivatives, and each obeys similar sharp calculus rules. The LD-derivative is essentially a variant of the lexicographic derivative that satisfies calculus rules similarly to the classic directional derivative. Inspection of [12, Equation 21] and the construction of the directed subdifferential in [9, 10] suggests that the directed subdifferential is in some sense analogous to a collection of LD-derivatives. These observations motivate the developments in this article.

The main results obtained in this article are as follows. For scalar-valued locally Lipschitz continuous functions, lexicographic smoothness is shown to be equivalent to directional subdifferentiability, the necessary optimality conditions developed for lexicographic derivatives [8, Theorem 8] and directed subdifferentials [10, Equation 8] are shown to coincide, and the Rubinov subdifferential is shown to include the lexicographic subdifferential while being included in its convex hull. Various implications of these results are discussed.

This article is structured as follows. Section 2 presents relevant definitions, Section 3 develops the main results of the article, and Section 4 presents concluding remarks.

2. Mathematical background

Notational conventions used in this article are as follows. The space $\mathbb{R}^n$ is equipped with the Euclidean norm and inner product, which are denoted as $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$, respectively. The unit sphere in $\mathbb{R}^n$ is then $\mathcal{S}_{n-1} := \{d \in \mathbb{R}^n : \|d\| = 1\}$. The zero vector in $\mathbb{R}^n$ will be denoted as 0_n. The column space of a matrix $A \in \mathbb{R}^{m\times n}$ is $\mathcal{R}(A) := \{Ad : d \in \mathbb{R}^n\}$.

Given an open set $X \subset \mathbb{R}^n$ and a function $f : X \to \mathbb{R}^m$, the following limit, if it exists, is the *directional derivative* of f at $x \in X$ in the direction $d \in \mathbb{R}^n$:

$$f'(x; d) := \lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t}.$$

The function f is *directionally differentiable* at $x \in X$ if $f'(x; d)$ exists in $\mathbb{R}^m$ for each $d \in \mathbb{R}^n$. In this case, the mapping $f'(x; \cdot)$ is positively homogeneous on $\mathbb{R}^n$. If $f'(x; \cdot)$ is linear, then f is *(Gâteaux) differentiable* at x, with $Jf(x)$ denoting the corresponding (Gâteaux) derivative of f at x.

The *Fréchet subdifferential* and *Fréchet superdifferential* are described in detail by Kruger [6]; in particular, if $X \subset \mathbb{R}^n$ is open and $f : X \to \mathbb{R}$ is directionally differentiable at $x \in X$, then the Fréchet subdifferential of f at x is

$$\partial_{\mathrm{F}} f(x) = \{a \in \mathbb{R}^n : f'(x; d) \geq \langle a, d \rangle, \quad \forall d \in \mathbb{R}^n\},$$

and the Fréchet superdifferential of f at x is

$$\partial^+_{\mathrm{F}} f(x) = \{a \in \mathbb{R}^n : f'(x; d) \leq \langle a, d \rangle, \quad \forall d \in \mathbb{R}^n\}.$$

Observe that if both $\partial_{\mathrm{F}} f(x)$ and $\partial^+_{\mathrm{F}} f(x)$ are nonempty, then f is differentiable at x.

2.1. Lexicographic smoothness

Lexicographic derivatives were developed by Nesterov [8], and are defined as follows.

Definition 1 (from [8]). Given an open set $X \subset \mathbb{R}^n$ and a function $f : X \to \mathbb{R}^m$ that is Lipschitz continuous near $x \in X$, f is *lexicographically (L-)smooth* at x if, for each $p \in \mathbb{N}$ and each matrix $M := \begin{bmatrix} m_{(1)} & \cdots & m_{(p)} \end{bmatrix} \in \mathbb{R}^{n \times p}$, the following directional derivative mappings are well-defined

$$\begin{aligned}
f^{(0)}_{x,M} &: \mathbb{R}^n \to \mathbb{R}^m : d \mapsto f'(x; d), \\
f^{(1)}_{x,M} &: \mathbb{R}^n \to \mathbb{R}^m : d \mapsto [f^{(0)}_{x,M}]'(m_{(1)}; d), \\
&\vdots \\
f^{(p)}_{x,M} &: \mathbb{R}^n \to \mathbb{R}^m : d \mapsto [f^{(p-1)}_{x,M}]'(m_{(p)}; d).
\end{aligned}$$

In this case, if $\mathcal{R}(M) = \mathbb{R}^n$, then $f^{(p)}_{x,M}$ is linear, and the *lexicographic (L-)derivative* of f at x in the directions M is $J_{\mathrm{L}}f(x; M) := Jf^{(p)}_{x,M}(0_n) \in \mathbb{R}^{m\times n}$. The *lexicographic (L-)subdifferential* of f at x is then

$$\partial_{\mathrm{L}} f(x) := \{J_{\mathrm{L}}f(x; M) : p \in \mathbb{N},\ M \in \mathbb{R}^{n\times p},\ \mathcal{R}(M) = \mathbb{R}^n\} \subset \mathbb{R}^{m\times n}.$$

For notational consistency with the other generalized derivatives in this article, when $m = 1$, the elements of $\partial_{\mathrm{L}} f(x)$ will be transposed and will be considered to be elements of $\mathbb{R}^n$.

The class of L-smooth functions is closed under composition and includes all continuously differentiable functions, convex functions on open sets, parametric integrals with L-smooth integrands [8], solutions of parametric ordinary differential equation systems with L-smooth right-hand side functions [13], and functions that are piecewise differentiable [11] in the sense of Scholtes [14]. This article will show that quasidifferentiable functions are also L-smooth. Observe that L-smoothness is defined for vector-valued functions.

2.2. Directed subdifferentiability

Directed subdifferentials were introduced by Baier et al. for delta-convex functions [15] and were ultimately extended to the class of *directed subdifferentiable functions* [9, 10]. Although directed subdifferentials were developed in the framework of *directed sets* [12], this article proceeds instead in terms of the Fréchet subdifferential, the Fréchet superdifferential, and the *Rubinov subdifferential* that is described in [16] as the *visualization* of the directed subdifferential.

As an intermediate construct, for any $n \in \{2, 3, \ldots\}$ and $\ell \in \mathcal{S}_{n-1}$, let $\Pi_{n-1,\ell} \in \mathbb{R}^{(n-1)\times n}$ denote a fixed matrix whose columns constitute an orthonormal basis for the subspace $\{d \in \mathbb{R}^n : \langle \ell, d\rangle = 0\}$. Thus,

- $\Pi_{n-1,\ell}\, \ell = 0_{n-1}$, and
- $\Pi^{\mathrm{T}}_{n-1,\ell}\Pi_{n-1,\ell}\, d = d - \langle d, \ell\rangle\, \ell$ for each $d \in \mathbb{R}^n$.

Let $P_{n-1,\ell}$ denote the linear transformation $z \mapsto \Pi_{n-1,\ell}\, z$.

The following definition of *directed subdifferentiability* is the special case of [9, Definition 5.1] in which the functions considered are locally Lipschitz continuous. In this case, the various continuity and boundedness requirements of the original definition are shown to hold in [14, Chapter 3]. The Rubinov subdifferential is described subsequently.

Definition 2 (adapted from [9, 10]). Consider any open set $X \subset \mathbb{R}^n$ and any function $f : X \to \mathbb{R}$ that is Lipschitz continuous near $x \in X$. If $n = 1$, then f is *directed subdifferentiable* at x if f is directionally differentiable at x. If $n > 1$, then f is directed subdifferentiable at x if both of the following conditions are satisfied:

- f is directionally differentiable at x, and
- for each $\ell \in \mathcal{S}_{n-1}$, the function $f_\ell : \mathbb{R}^{n-1} \to \mathbb{R} : d \mapsto f'(x; \ell + \Pi^{\mathrm{T}}_{n-1,\ell}\, d)$ is directed subdifferentiable at 0_{n-1}.

Certain functions that are not locally Lipschitz continuous may also be directed subdifferentiable according to [9] but are not considered in this article. Observe that all directed subdifferentiable functions are scalar valued.

Definition 3 (adapted from [10] and Definitions 4.4 and 4.8 in [16]). Consider any open set $X \subset \mathbb{R}^n$ and any locally Lipschitz continuous function $f : X \to \mathbb{R}$ that is directed subdifferentiable at $x \in X$. Define sets $M_n(\vec{\partial} f(x)), \partial_{\mathrm{R}} f(x) \subset \mathbb{R}^n$ recursively over $n \in \mathbb{N}$ as follows. If $n = 1$, then

$$M_1(\vec{\partial} f(x)) := \emptyset \qquad \text{and} \qquad \partial_{\mathrm{R}} f(x) := \partial_{\mathrm{F}} f(x) \cup \partial^+_{\mathrm{F}} f(x).$$

If $n > 1$, then

$$\begin{aligned} M_n(\vec{\partial} f(x)) := \big\{\Pi^{\mathrm{T}}_{n-1,\ell}\, y + f'(x; \ell)\, \ell : \ell \in \mathcal{S}_{n-1},\ y \in \partial_{\mathrm{R}} f_\ell(0_{n-1})\big\} \\ \setminus (\partial_{\mathrm{F}} f(x) \cup \partial^+_{\mathrm{F}} f(x)), \end{aligned}$$

$$\text{and} \qquad \partial_{\mathrm{R}} f(x) := \partial_{\mathrm{F}} f(x) \cup \partial^+_{\mathrm{F}} f(x) \cup M_n(\vec{\partial} f(x)).$$

The set $\partial_{\mathrm{R}} f(x)$ is the *Rubinov subdifferential* of f at x.

Again, although not considered in this article, the Rubinov subdifferential is also defined in [10] for directed subdifferentiable functions that are not locally Lipschitz continuous.

3. Main results

This section presents the main results of this article: that L-smoothness and directed subdifferentiability are equivalent for locally Lipschitz continuous functions, that the corresponding optimality conditions are also equivalent, and that the L-subdifferential is related to the Rubinov subdifferential in a certain way.

3.1. Relating lexicographic smoothness and directed subdifferentiability

The following theorem shows that L-smoothness and directed subdifferentiability are equivalent for locally Lipschitz continuous functions. This theorem is proved at the end of the subsection.

Theorem 3.1. *Given an open set $X \subset \mathbb{R}^n$ and a function $f : X \to \mathbb{R}$, f is L-smooth at some $x \in X$ if and only if f is both Lipschitz continuous near x and directed subdifferentiable at x.*

The following result was claimed in [8], under the assumption that any quasidifferentiable function is also delta-convex. Although Baier et al. [16, Example 3.5] provide a counterexample for this assumption, Theorem 3.1 yields the desired conclusion nevertheless.

Corollary 3.2. *Given an open set $X \subset \mathbb{R}^n$, any quasidifferentiable function $f : X \to \mathbb{R}$ is L-smooth.*

Proof. Quasidifferentiable functions are both locally Lipschitz continuous and directed subdifferentiable and thus satisfy the hypotheses of Theorem 3.1.

The following corollary generalizes similar results in [10, Section 4] and follows immediately from Theorem 3.1 and [11, Proposition 2.2].

Corollary 3.3. *Given open sets $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}$ and locally Lipschitz continuous functions $g : X \to Y$ and $f : Y \to \mathbb{R}$, suppose that g is directed subdifferentiable at $x \in X$ and that f is directed subdifferentiable at $g(x)$. The composite function $h : X \to \mathbb{R} : z \mapsto f(g(z))$ is also directed subdifferentiable at x.*

Similarly, Theorem 3.1 implies that the various examples of L-smooth functions provided in [8, 13, 11] are all directed subdifferentiable.

The following corollary shows that if an objective function in an unconstrained optimization problem is locally Lipschitz continuous, then necessary optimality conditions developed by Baier et al. and by Nesterov coincide.

Corollary 3.4. *Given an open set $X \subset \mathbb{R}^n$, consider an L-smooth function $f : X \to \mathbb{R}$, and choose $\hat{x} \in X$. The point $\hat{x}$ satisfies the necessary optimality condition given by [10, Equation 8] if and only if it satisfies the necessary optimality condition given by [8, Theorem 8].*

Proof. By Theorem 3.1, f is directed subdifferentiable on X. The required result follows from inspection of the recursive definitions of the "$\succ$" relation in [8] and the "$\geq$" relation for directed sets in [10] and from inspection of the two necessary optimality conditions.

The following lemma is similar in spirit to [10, Proposition 2.4] and will be used in the subsequent proof of Theorem 3.1.

Lemma 3.5. *Given an open set $X \subset \mathbb{R}^n$ and a locally Lipschitz continuous function $f : X \to \mathbb{R}$, if f is directed subdifferentiable at $x \in X$, then the directional derivative mapping $f'(x;\cdot)$ is both directed subdifferentiable and Lipschitz continuous on $\mathbb{R}^n$.*

Proof. For notational convenience, define $\phi := f'(x;\cdot)$. The lemma will be proved by induction on $n \in \mathbb{N}$. As the base case of the inductive argument, if $n = 1$, then $\phi(d)$ can be computed directly for any $d \in \mathbb{R}$ to be

$$\phi(d) = \begin{cases} f'(x;1)\,d, & \text{if } d \geq 0, \\ -f'(x;-1)\,d, & \text{if } d < 0. \end{cases}$$

Thus, ϕ is piecewise linear in the sense of Scholtes [14] and is therefore Lipschitz continuous and directionally differentiable on $\mathbb{R}$, as required.

Next, as the inductive step, suppose that $n := m > 1$, and assume that the lemma has been demonstrated for the case in which $n := m - 1$. The function ϕ is Lipschitz continuous according to [14]; it remains to be shown that ϕ is also directed subdifferentiable at some arbitrary $y \in \mathbb{R}^m$. If $y = 0_m$, then this result follows immediately from [10, Proposition 2.4]. Thus, assume that $y \neq 0_m$, in which case there exist $\beta > 0$ and $\ell \in \mathcal{S}_{m-1}$ for which $y = \beta\ell$. Define a mapping f_ℓ as in Definition 2; since f is directed subdifferentiable at x, f_ℓ is directed subdifferentiable at 0_{m-1} and is Lipschitz continuous according to [14].

To establish the directional differentiability of ϕ at y, choose any $d \in \mathbb{R}^m$. Define $\alpha := \langle \ell, d\rangle \in \mathbb{R}$ and $v := d - \alpha\ell \in \mathbb{R}^m$, and observe that $\langle \ell, v\rangle = 0$. For any sufficiently small $t > 0$, $(\beta + t\alpha) > 0$, and so the positive homogeneity of ϕ yields

$$\begin{aligned}
&\phi(y + td) - \phi(y) \\
&\quad = (\beta + t\alpha)\left(\phi(\ell + (\tfrac{t}{\beta + t\alpha})v) - \phi(\ell)\right) + t\alpha\,\phi(\ell) \\
&\quad = (\beta + t\alpha)\left(f_\ell((\tfrac{t}{\beta + t\alpha})\Pi_{m-1,\ell}\,d) - f_\ell(0_{m-1})\right) + t\phi(\ell)\,\langle \ell, d\rangle.
\end{aligned}$$

Since f_ℓ is directionally differentiable at 0_{m-1}, the above equation implies that

$$\lim_{t\to 0^+} \frac{\phi(y+td)-\phi(y)}{t} = [f_\ell]'(0_{m-1}; \Pi_{m-1,\ell}\, d) + \phi(\ell)\,\langle \ell, d\rangle. \tag{1}$$

Since d was chosen arbitrarily, the directional differentiability of ϕ at y is thereby established.

It remains to be shown that for any $h \in \mathcal{S}_{m-1}$, the mapping $\phi_h : \mathbb{R}^{m-1} \to \mathbb{R} : z \mapsto \phi'(y; h+\Pi^{\mathrm{T}}_{m-1,h}\, z)$ is directed subdifferentiable at 0_{m-1}. Thus, choose any $h \in \mathcal{S}_{m-1}$, and define the affine transformations:

$$T_H : \mathbb{R}^{m-1} \to \mathbb{R}^m : z \mapsto h + \Pi^{\mathrm{T}}_{m-1,h}\, z \quad \text{and} \quad T_L : \mathbb{R}^m \to \mathbb{R} : z \mapsto \phi(\ell)\,\langle \ell, z\rangle.$$

Since f_ℓ is Lipschitz continuous and directed subdifferentiable at 0_{m-1}, the mapping $[f_\ell]'(0_{m-1}; \cdot)$ is Lipschitz continuous and directed subdifferentiable on $\mathbb{R}^{m-1}$ according to the inductive assumption. Propositions 2.6 and 4.3 in [10] then imply that the mapping

$$\psi : \mathbb{R}^{m-1} \to \mathbb{R} : z \mapsto \big[[f_\ell]'(0_{m-1}; \cdot)\big] \circ P_{m-1,\ell} \circ T_H(z) + T_L \circ T_H(z)$$

is also directed subdifferentiable at 0_{m-1}. Inspection of (1) and the definitions of f_ℓ and ϕ_h show that ψ is equivalent to ϕ_h, and so ϕ_h is directed subdifferentiable at 0_{m-1}. This completes the inductive step.

Proof of Theorem 3.1. First, suppose that f is L-smooth at $x \in X$, in which case f is both Lipschitz continuous near x and directionally differentiable at x. It will be shown by induction on the domain dimension $n \in \mathbb{N}$ that f is also directed subdifferentiable at x. The $n = 1$ case is trivial. Thus, suppose that $n > 1$. As the inductive assumption, suppose that given any open set $Y \subset \mathbb{R}^{n-1}$ and function $g : Y \to \mathbb{R}$, if g is L-smooth at some $y \in Y$, then g is also directed subdifferentiable at y. It remains to be shown that for arbitrary $\ell \in \mathcal{S}_{n-1}$, f_ℓ is directed subdifferentiable at 0_{n-1}. Define the affine transformation

$$T_F : \mathbb{R}^{n-1} \to \mathbb{R}^n : d \mapsto \ell + \Pi^{\mathrm{T}}_{n-1,y}\, d,$$

and observe that $f_\ell \equiv [f'(x;\cdot)] \circ T_F$. Since f is L-smooth at x, $f'(x;\cdot)$ is L-smooth on $\mathbb{R}^n$; Theorem 5 in [8] then implies that f_ℓ is L-smooth on $\mathbb{R}^{n-1}$. The inductive assumption, applied with $g := f_\ell$, shows that f_ℓ is directed subdifferentiable on 0_{n-1}. This completes the inductive argument and thereby demonstrates the "only if" claim of the theorem.

Next, for the "if" claim of the theorem, suppose that f is both Lipschitz continuous near x and directed subdifferentiable at x. Consider any $p \in \mathbb{N}$ and $M \in \mathbb{R}^{n \times p}$. It will be shown by induction on $k \in \{0, 1, \ldots, p\}$ that the directional derivative $f_{x,M}^{(k)}$ is well defined, Lipschitz continuous, and directed subdifferentiable on $\mathbb{R}^n$. Since $f_{x,M}^{(0)} \equiv f'(x; \cdot)$, the $k = 0$ case is demonstrated by Lemma 3.5. For the inductive step, assume that $f_{x,M}^{(j-1)}$ is well defined, Lipschitz continuous, and directed subdifferentiable on $\mathbb{R}^n$ for some $j \in \{1, \ldots, p\}$. Combined with this inductive assumption, Lemma 3.5 shows that $f_{x,M}^{(j)} \equiv [f_{x,M}^{(j-1)}]'(m_{(j)}; \cdot)$ is well defined, Lipschitz continuous, and directed subdifferentiable on $\mathbb{R}^n$, as required. This completes the inductive argument.

3.2. Relating the lexicographic and Rubinov subdifferentials

The following theorem relates the lexicographic and Rubinov subdifferentials for L-smooth functions. It is motivated by the similar properties of the LD-derivatives described in [11] and the directed subdifferential described in [9, 10]. This theorem is proved at the end of the subsection.

Theorem 3.6. *Given an open set $X \subset \mathbb{R}^n$ and a function $f : X \to \mathbb{R}$, if f is L-smooth at $x \in X$, then*

$$\partial_{\mathrm{L}} f(x) \subset \partial_{\mathrm{R}} f(x) \subset \operatorname{conv} \partial_{\mathrm{L}} f(x).$$

The inclusion $\partial_{\mathrm{L}} f(x) \subset \partial_{\mathrm{R}} f(x)$ may be strict; for example, with $f : \mathbb{R} \to \mathbb{R}$ defined as the absolute-value mapping $x \mapsto |x|$, one can readily verify that $\partial_{\mathrm{L}} f(0) = \{-1, +1\}$ and $\partial_{\mathrm{R}} f(0) = [-1, 1]$. Since the Rubinov subdifferential is the visualization of the directed subdifferential, Example 3.20 in [17] suggests that the Rubinov subdifferential of a delta-convex function can be nonconvex; if this is indeed the case, then the inclusion $\partial_{\mathrm{R}} f(x) \subset \operatorname{conv} \partial_{\mathrm{L}} f(x)$ would also be strict. Nevertheless, the following corollary is an immediate consequence of Theorem 3.6.

Corollary 3.7. *Given an open set $X \subset \mathbb{R}^n$ and a function $f : X \to \mathbb{R}$ that is L-smooth at $x \in X$, $\partial_{\mathrm{R}} f(x) = \operatorname{conv} \partial_{\mathrm{L}} f(x)$ if and only if $\partial_{\mathrm{R}} f(x)$ is convex.*

The following lemmas provide intermediate results that are used in the proof of Theorem 3.6.

Lemma 3.8. *Given an open set $X \subset \mathbb{R}^n$ and a function $f : X \to \mathbb{R}$ that is L-smooth at $x \in X$,*

$$\partial_{\mathrm{F}} f(x) \cup \partial^{+}_{\mathrm{F}} f(x) \subset \operatorname{conv} \partial_{\mathrm{L}} f(x).$$

Proof. Choose any $a \in \partial_{\mathrm{F}} f(x)$. To obtain a contradiction, suppose that $a \notin \operatorname{conv} \partial_{\mathrm{L}} f(x)$. Since $\operatorname{conv} \partial_{\mathrm{L}} f(x)$ is closed and convex, there exists $h \in \mathbb{R}^n$ for which $\langle a, h\rangle > \sup\{\langle g, h\rangle : g \in \partial_{\mathrm{L}} f(x)\}$. According to [8, Lemma 10], then, $\langle a, h\rangle > f'(x; h)$. This implies that $a \notin \partial_{\mathrm{F}} f(x)$, contradicting the choice of a. Thus, $\partial_{\mathrm{F}} f(x) \subset \operatorname{conv} \partial_{\mathrm{L}} f(x)$. A similar argument shows that $\partial^{+}_{\mathrm{F}} f(x) \subset \operatorname{conv} \partial_{\mathrm{L}} f(x)$ as well.

Lemma 3.9. *Given $d \in \mathbb{R}^n$ and a positively homogeneous function $f : \mathbb{R}^n \to \mathbb{R}^m$ that is L-smooth at 0_n, f is also L-smooth at d, $\partial_{\mathrm{L}} f(d) \subset \partial_{\mathrm{L}} f(0_n)$, and $f'(0_n; d) = f(d) = Ad$ for each $A \in \partial_{\mathrm{L}} f(d)$.*

Proof. Since $f \equiv f'(0_n; \cdot)$, the L-smoothness of f at d and the inclusion $\partial_{\mathrm{L}} f(d) \subset \partial_{\mathrm{L}} f(0_n)$ follow from the L-subdifferential's satisfaction of [8, Axiom 2]. Next, choose any $A \in \partial_{\mathrm{L}} f(d)$; by definition of the L-subdifferential, there exists $p \in \mathbb{N}$ and a matrix $M \in \mathbb{R}^{n \times p}$ for which $f^{(p)}_{d,M}(v) = Av$ for each $v \in \mathbb{R}^n$. Define a matrix $N := \begin{bmatrix} d & M \end{bmatrix} \in \mathbb{R}^{n \times (p+1)}$; the positive homogeneity of f implies that $f^{(0)}_{d,M} \equiv f^{(1)}_{0_n,N}$, and so $f^{(p)}_{d,M}(v) \equiv f^{(p+1)}_{0_n,N}(v)$ for each $v \in \mathbb{R}^n$. Thus, $f^{(p+1)}_{d,N}(d) = Ad$; the definition of N, the positive homogeneity of f, and Lemma 2.1 in [13] then yield

$$Ad = f^{(p+1)}_{0_n,N}(d) = f^{(0)}_{0_n,N}(d) = f'(0_n; d) = f(d),$$

as required.

Lemma 3.10. *Let $e_{(1)}$ denote the leftmost column of the identity matrix $I \in \mathbb{R}^{n \times n}$. Given $\ell \in \mathcal{S}_{n-1}$ and a positively homogeneous function $f : \mathbb{R}^n \to \mathbb{R}^m$ that is L-smooth at 0_n,*

$$\partial_{\mathrm{L}} f(\ell) \subset \{J_{\mathrm{L}} f(0_n; M) : M \in \mathbb{R}^{n \times n},\ M^{\mathrm{T}} M = I,\ M e_{(1)} = \ell\}.$$

Proof. Choose any $A \in \partial_{\mathrm{L}} f(\ell)$. For some $p \in \mathbb{N}$ and $B \in \mathbb{R}^{n \times p}$ with $\mathcal{R}(B) = \mathbb{R}^n$, $f^{(p)}_{\ell,B}(d) = Ad$ for each $d \in \mathbb{R}^n$. Define a matrix $N := \begin{bmatrix} \ell & B \end{bmatrix} \in \mathbb{R}^{n \times (p+1)}$; evidently $\mathcal{R}(N) = \mathbb{R}^n$. The positive homogeneity of f implies that $f^{(0)}_{\ell,B} \equiv f^{(1)}_{0_n,N}$, and so $f^{(p)}_{\ell,B} \equiv f^{(p+1)}_{0_n,N}$. Thus, $A = J_{\mathrm{L}} f(0_n; N)$. Since the leftmost column of N is $\ell \in \mathcal{S}_{n-1}$, Lemma 4 in [8] implies the existence of an orthonormal matrix $M \in \mathbb{R}^{n \times n}$ that satisfies $M e_{(1)} = \ell$ and $A = J_{\mathrm{L}} f(0_n; M)$.

With these intermediate results, Theorem 3.6 can be proved as follows.

PROOF OF THEOREM 3.6. According to Theorem 3.1, f is both Lipschitz continuous near x and directed subdifferentiable at x. Inspection of the definitions of the lexicographic and Rubinov subdifferentials shows that $\partial_{\mathrm{L}} f(x) = \partial_{\mathrm{L}}[f'(x;\cdot)](0_n)$ and $\partial_{\mathrm{R}} f(x) = \partial_{\mathrm{R}}[f'(x;\cdot)](0_n)$. Thus, it will be assumed without loss of generality that f is positively homogeneous, $X = \mathbb{R}^n$, and $x = 0_n$. Since f is positively homogeneous, $f \equiv f'(0_n;\cdot)$.

If $n = 1$, then, using the positive homogeneity of f, one may compute the Fréchet subdifferential and superdifferential of f at 0 directly to be

$$\partial_{\mathrm{F}} f(0) = \begin{cases} [-f(-1), f(1)] = \operatorname{conv} \partial_{\mathrm{L}} f(0), & \text{if } -f(-1) \leq f(1), \\ \emptyset & \text{otherwise,} \end{cases}$$
$$\partial_{\mathrm{F}}^{+} f(0) = \begin{cases} [f(1), -f(-1)] = \operatorname{conv} \partial_{\mathrm{L}} f(0), & \text{if } -f(-1) \geq f(1), \\ \emptyset & \text{otherwise.} \end{cases}$$

Thus, $\partial_{\mathrm{R}} f(0) = \partial_{\mathrm{F}} f(0) \cup \partial_{\mathrm{F}}^{+} f(0) = \operatorname{conv} \partial_{\mathrm{L}} f(0) \supset \partial_{\mathrm{L}} f(0)$, as required.

Next, the inclusion $\partial_{\mathrm{L}} f(0_n) \subset \partial_{\mathrm{R}} f(0_n)$ will be proved by induction on $n \in \mathbb{N}$; the case in which $n = 1$ was established in the previous paragraph. Suppose that $n := m > 1$, and assume that the inclusion has already been established for the case in which $n := m - 1$. Choose any $a \in \partial_{\mathrm{L}} f(0_m)$; it will be shown that $a \in \partial_{\mathrm{R}} f(0_m)$. If $a \in \partial_{\mathrm{F}} f(0_m) \cup \partial_{\mathrm{F}}^{+} f(0_m)$, then this result is trivial; hence, assume that $a \notin \partial_{\mathrm{F}} f(0_m) \cup \partial_{\mathrm{F}}^{+} f(0_m)$. It will be shown that $a \in M_m(\vec{\partial} f(0_m)) \subset \partial_{\mathrm{R}} f(0_m)$. Since $a \in \partial_{\mathrm{L}} f(0_m)$, Lemma 4 in [8] implies that there exists an orthonormal matrix $M \in \mathbb{R}^{m \times m}$ for which $f^{(m)}_{0_m,M}(v) = \langle a, v\rangle$ for each $v \in \mathbb{R}^m$. Define $\ell \in \mathcal{S}_{m-1}$ to be the leftmost column of M, and define $b := \Pi_{m-1,\ell}\, a \in \mathbb{R}^{m-1}$. Lemma 3.9 and the definition of $\Pi_{m-1,\ell}$ then yield

$$a = \Pi^{\mathrm{T}}_{m-1,\ell}\, b + \langle a, \ell\rangle\, \ell = \Pi^{\mathrm{T}}_{m-1,\ell}\, b + f'(0_m;\ell)\, \ell.$$

Thus, in order to show that $a \in M_m(\vec{\partial} f(0_m))$, it suffices to show that $b \in \partial_{\mathrm{R}} f_\ell(0_{m-1})$. The inductive assumption implies $\partial_{\mathrm{L}} f_\ell(0_{m-1}) \subset \partial_{\mathrm{R}} f_\ell(0_{m-1})$, so it suffices in turn to show that $b \in \partial_{\mathrm{L}} f_\ell(0_{m-1})$. Applying [8, Theorem 5] to the definition of f_ℓ yields $\partial_{\mathrm{L}} f_\ell(0_{m-1}) = \{\Pi_{m-1,\ell}\, v : v \in \partial_{\mathrm{L}} f(\ell)\}$. Moreover, Lemma 3.10 shows that $a \in \partial_{\mathrm{L}} f(\ell)$; thus, $b = \Pi_{m-1,\ell}\, a \in \partial_{\mathrm{L}} f_\ell(0_{m-1})$, as required. This completes the inductive step and thereby establishes the inclusion $\partial_{\mathrm{L}} f(0_n) \subset \partial_{\mathrm{R}} f(0_n)$.

Lastly, the inclusion $\partial_{\mathrm{R}} f(0_n) \subset \operatorname{conv} \partial_{\mathrm{L}} f(0_n)$ will be proved by induction on $n \in \mathbb{N}$; the case in which $n = 1$ has already been established. For the inductive step, suppose that $n := m > 1$, and assume that the inclusion has already been established for the case in which $n := m - 1$. Choose any $a \in \partial_{\mathrm{R}} f(0_m)$; it will be shown that $a \in \operatorname{conv} \partial_{\mathrm{L}} f(0_m)$. If $a \in \partial_{\mathrm{F}} f(0_m) \cup \partial_{\mathrm{F}}^{+} f(0_m)$, then Lemma 3.8 implies that $a \in \partial_{\mathrm{L}} f(0_m)$, as required. Otherwise, suppose that $a \notin \partial_{\mathrm{F}} f(0_m) \cup \partial_{\mathrm{F}}^{+} f(0_m)$, in which case $a \in M_m(\vec{\partial} f(0_m))$. Since f is positively homogeneous, $f'(0_m; \cdot) \equiv f$, which implies the existence of $\ell \in \mathcal{S}_{m-1}$ and $b \in \partial_{\mathrm{R}} f_\ell(0_{m-1})$ for which

$$a = \Pi_{m-1,\ell}^{\mathrm{T}} \, b + f'(0_m; \ell) \, \ell = \Pi_{m-1,\ell}^{\mathrm{T}} \, b + f(\ell) \, \ell.$$

By the inductive assumption, $\partial_{\mathrm{R}} f_\ell(0_{m-1}) \subset \operatorname{conv} \partial_{\mathrm{L}} f_\ell(0_{m-1})$, and so $b \in \operatorname{conv} \partial_{\mathrm{L}} f_\ell(0_{m-1})$. Applying the Carathéodory Theorem, one has $\lambda_i \in \mathbb{R}_+$ and $b_{(i)} \in \partial_{\mathrm{L}} f_\ell(0_{m-1})$ for each $i \in \{1, \ldots, m\}$, for which $\sum_{i=1}^{m} \lambda_i = 1$ and $\sum_{i=1}^{m} \lambda_i b_{(i)} = b$. For each $i \in \{1, \ldots, m\}$, define

$$a_{(i)} := \Pi_{m-1,\ell}^{\mathrm{T}} \, b_{(i)} + f(\ell) \, \ell; \tag{2}$$

it follows that $\sum_{i=1}^{m} \lambda_i a_{(i)} = a$. Now, to obtain a contradiction, suppose that $a \notin \operatorname{conv} \partial_{\mathrm{L}} f(0_m)$. Consequently, $a_{(j)} \notin \operatorname{conv} \partial_{\mathrm{L}} f(0_m)$ for some $j \in \{1, \ldots, m\}$. Since $\operatorname{conv} \partial_{\mathrm{L}} f(0_m)$ is convex and closed, there exists $h \in \mathbb{R}^m$ for which

$$\langle a_{(j)}, h \rangle > \sup\{\langle g, h \rangle : g \in \operatorname{conv} \partial_{\mathrm{L}} f(0_m)\} \geq \sup\{\langle g, h \rangle : g \in \partial_{\mathrm{L}} f(0_m)\}. \tag{3}$$

Applying [8, Theorem 5] to the definition of f_ℓ, one obtains $\partial_{\mathrm{L}} f_\ell(0_m) = \{\Pi_{m-1,\ell} \, v : v \in \partial_{\mathrm{L}} f(\ell)\}$. Thus, there exists $c \in \partial_{\mathrm{L}} f(\ell)$ for which $b_{(j)} = \Pi_{m-1,\ell} \, c$. Substituting this expression for $b_{(j)}$ into (2) with $i := j$ yields:

$$\langle a_{(j)}, h \rangle = \langle \Pi_{m-1,\ell}^{\mathrm{T}} \, \Pi_{m-1,\ell} \, c, h \rangle + f(\ell) \, \langle \ell, h \rangle = \langle c, h \rangle - \langle c, \ell \rangle \langle \ell, h \rangle + f(\ell) \, \langle \ell, h \rangle.$$

Lemma 3.9 shows that $\langle c, \ell \rangle = f(\ell)$ and $c \in \partial_{\mathrm{L}} f(0_m)$; these observations and the above equation imply that $\langle a_{(j)}, h \rangle \in \{\langle g, h \rangle : g \in \partial_{\mathrm{L}} f(0_m)\}$, which contradicts (3). So, $a \in \operatorname{conv} \partial_{\mathrm{L}} f(0_m)$, as required.

4. Conclusion

Theorem 3.1 demonstrates the equivalence between L-smoothness and directed subdifferentiability for scalar-valued functions that are locally Lipschitz continuous. Moreover, Theorem 3.6 shows that the Rubinov subdifferential is a particular superset of the L-subdifferential. These results suggest that the various benefits of each type of generalized derivative could be extended to the other, in the vein of Corollaries 3.2 and 3.3. For example, the class of L-smooth functions includes vector-valued functions and is known to include a broad variety of nonsmooth functions, such as convex functions on open sets [8] and the solutions of parametric ordinary differential equations with L-smooth right-hand side functions [13]. On the other hand, directed subdifferentiability extends to functions that are not locally Lipschitz continuous. Moreover, according to [18], when $f : \mathbb{R}^2 \to \mathbb{R}$ is the difference of two sublinear functions, the Rubinov subdifferential is identical to the Mordukhovich symmetric subdifferential [3], giving hope that L-subdifferentials and the various Mordukhovich subdifferentials could also be related.

Acknowledgment

This material was based on work supported by the U.S. Department of Energy, Office of Science, under contract DE-AC02-06CH11357.

References

[1] F. H. Clarke, Optimization and Nonsmooth Analysis, SIAM, Philadelphia, PA, 1990.

[2] R. T. Rockafellar, R. J.-B. Wets, Variational Analysis, A Series of Comprehensive Studies in Mathematics, Springer, Berlin, 1997.

[3] B. S. Mordukhovich, Variational Analysis and Generalized Differentiation I: Basic Theory, Springer, Berlin, 2006.

[4] T. H. Sweetser, III, A minimal set-valued strong derivative for vector-valued Lipschitz functions, J. Optimiz. Theory App. 23 (1977) 549–562.

[5] P. Michel, J.-P. Penot, A generalized derivative for calm and stable functions, Differential and Integral Equations 5 (2) (1992) 433–454.

[6] A. Y. Kruger, On Fréchet subdifferentials, J. Math. Sci. 116 (3) (2003) 3325–3358.

[7] Y. Nesterov, The technique of nonsmooth differentiation, Izvestija AN SSSR, Technicheskaja Kibernetika 1 (1987) 199–208, in Russian.

[8] Y. Nesterov, Lexicographic differentiation of nonsmooth functions, Math. Program. B 104 (2005) 669–700.

[9] R. Baier, E. Farkhi, V. Roshchina, Directed subdifferentiable functions and the directed subdifferential without delta-convex structure, J. Optim. Theory Appl. 160 (2014) 391–414.

[10] R. Baier, E. Farkhi, V. Roshchina, From quasidifferentiable to directed subdifferentiable functions: Exact calculus rules, preprint, http://www.arxiv.org/abs/1507.00174v2 (2015).

[11] K. A. Khan, P. I. Barton, A vector forward mode of automatic differentiation for generalized derivative evaluation, Optim. Method. Softw. 30 (6) (2015) 1185–1212.

[12] R. Baier, E. M. Farkhi, Differences of convex compact sets in the space of directed sets. Part I: The space of directed sets, Set-Valued Anal. 9 (2001) 217–245.

[13] K. A. Khan, P. I. Barton, Generalized derivatives for solutions of parametric ordinary differential equations with non-differentiable right-hand sides, J. Optimiz. Theory App. 163 (2) (2014) 355–386.

[14] S. Scholtes, Introduction to Piecewise Differentiable Equations, SpringerBriefs in Optimization, Springer, New York, 2012.

[15] R. Baier, E. Farkhi, The directed subdifferential of DC functions, in: A. Leizarowitz, B. S. Mordukhovich, I. Shafrir, A. J. Zaslavski (Eds.), Nonlinear Analysis and Optimization II: Optimization, Vol. 514, AMS and Bar-Ilan University, 2010, pp. 27–43.

[16] R. Baier, E. Farkhi, V. Roshchina, The directed and Rubinov subdifferentials of quasidifferentiable functions, Part I: Definition and examples, Nonlinear Anal. 75 (2012) 1074–1088.

[17] R. Baier, E. M. Farkhi, Differences of convex compact sets in the space of directed sets. Part II: Visualization of directed sets, Set-Valued Anal. 9 (2001) 247–272.

[18] R. Baier, E. Farkhi, V. Roshchina, On computing the Mordukhovich subdifferential using directed sets in two dimensions, in: R. S. Burachik,

J.-C. Yao (Eds.), Variational Analysis and Generalized Differentiation in Optimization and Control, Springer Optimization and its Applications, Springer, 2010, pp. 59–93.