Slope Intervals, Generalized Gradients, Semigradients, Slant Derivatives,and Csets

Humberto Muñoz (hmunoz@math.subr.edu)* and R. B. Kearfott (rbk@louisiana.edu)†

Abstract. Many practical optimization problems are nonsmooth, and derivative-type methods cannot be applied. To overcome this difficulty, there are different concepts to replace the derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$: interval slopes, semigradients, generalized gradients, and slant derivatives are some examples. These approaches generalize the success of convex analysis, and are effective in optimization. However, with the exception of interval slopes, it is not clear how to automatically compute these; having a general analogue to the chain rule, interval slopes can be computed with automatic differentiation techniques. In this paper we study the relationships among these approaches for nonsmooth Lipschitz optimization problems in finite dimensional Euclidean spaces. Inclusion theorems concerning the equivalence of these concepts when there exist one sided derivatives in one dimension and in multidimensional cases are proved separately. Valid enclosures are produced. Under containment set (cset) theory, for instance, the cset of the gradient of a locally Lipschitz function f near x is included in its generalized gradient.

Keywords: generalized gradient, slope interval, semigradient, slant derivative, subdifferential, subgradient, cset, symmetric slope interval, nonsmooth optimization methods

1. Introduction

The purpose of this work is to delineate practical relationships among five different generalizations to the derivative or gradient of functions $f : D \subset \mathbb{R}^n \to \mathbb{R}$: slope interval, generalized gradient, subdifferential (set of subgradients), semigradient, slant derivative, and cset. These generalizations are used in the solution of nonsmooth constrained or unconstrained optimization problems.

Different techniques have been developed for nonsmooth optimization problems. For nonsmooth convex functions, the subdifferential in the sense of convex analysis is introduced in [26], and the stability of the optimal solution with subdifferential is studied in [8]. For constrained optimization problems with locally Lipschitz function, the generalized gradient is introduced in [3]. For unconstrained opti-

* Department of Mathematics, Southern University, P.O. Box 9440, Baton Rouge, LA 70815

† Department of Mathematics, University of Louisiana at Lafayette, Box 4-1010, Lafayette, LA 70504-1010

mization with locally Lipschitz continuous objective function, bundle methods are standard solution methods which provide interior approximations of the generalized gradient [12, 31]. In the French and Russian literature are the works of [1, 4] who also worked with Lipschitz functions. The first works on nonsmooth optimization using tools of interval arithmetic were presented in [20]. A combination of bundle methods and interval extensions (outer approximations) of the generalized gradient, based in Goldstein's approach and developed in [21] and [33].

The following goals motivate generalization of gradients: 1) to generalize the clear success of convex analysis; 2) to be effective in optimization; and 3) to be easy to compute, particularly in the composition of functions with chain rules. Computing a slope interval is easy with automatic slope computation, see Section 6.1 on [11] and [9, 24]. It is not possible to compute sharp slopes, in general, but we can compute enclosures for non-smooth functions easily enough, while computing the other generalizations of the gradient seems hard. The other two goals are achievable with all approaches.

In Section 2, basic notation of interval arithmetic as well as the notation used in the rest of the paper are given. In Section 3, the slope interval, directional slope, and symmetric interval slope are defined, and some results related to these concepts are shown. Clarke's generalized gradient and the main relationships between slope intervals and generalized gradients are established in Section 4. In Section 5, generalizations of the *Karush-Kuhn-Tucker* optimality conditions are presented. The semigradient and its relationships with slope intervals and symmetric interval slopes are given in Section 6. Cset theory and some results are given in Section 7. Section 8 deals with slant derivative and its relationship with slope interval. Some general conclusions appear in Section 9.

2. Notation

Real interval arithmetic, introduced in its modern form in [15], is based on arithmetic within the set of real closed intervals. A real bounded and closed *interval* is defined by

$$\boldsymbol{x} \equiv [\underline{x}, \bar{x}] := [\inf \boldsymbol{x}, \sup \boldsymbol{x}] \in \mathbb{IR},$$

where $\mathbb{IR}$ denotes the set of compact intervals. Ocasionally, in addition to using boldface to denote intervals, we use uppercase boldface to denote sets.

Let $\check{x}$ and $w(\boldsymbol{x})$ denote the *midpoint* and the *width* of $\boldsymbol{x}$, respectively, that is,

$$\check{x} := \frac{\underline{x} + \bar{x}}{2}, \qquad w(\boldsymbol{x}) := \bar{x} - \underline{x}.$$

If S is a subset of $\mathbb{R}^n$, $co(S)$ denotes the *convex hull* of S, and $\square S$ denotes the *interval hull* of S. The *interval union* or hull of two intervals is defined by

$$\boldsymbol{x} \underline{\cup} \mathbf{y} := [\min\{\underline{x}, \underline{y}\}, \max\{\bar{x}, \bar{y}\}],$$

where the bar under the union symbol means the convex hull. An n-dimensional *interval vector* (also called *box*) is defined by

$$\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)^T \in \mathbb{IR}^n,$$

where $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ are real intervals and $\mathbb{IR}^n$ is the set of real interval vectors. An *interval matrix* $\boldsymbol{A} = (\boldsymbol{a}_{ij})$ is a matrix all of whose entries $\boldsymbol{a}_{ij}$ are intervals. For interval vectors $\boldsymbol{x} \in \mathbb{IR}^n$, the midpoint $\check{x}$, the width $w(\boldsymbol{x})$, and the *hull* operation are defined componentwise. We use $w(\boldsymbol{x})$ in the context of $\|w(\boldsymbol{x})\| = \|w(\boldsymbol{x})\|_\infty$. The basic interval operations $(+, -, \cdot, /)$ and elementary interval functions can be defined operationally [11], [18], [23],[16].

Let $f : \mathbb{R} \to \mathbb{R}$ be a function, and x a given real value. We denote

$$f(x)^- := \lim_{y \uparrow x} f(y) \quad \text{and} \quad f(x)^+ := \lim_{y \downarrow x} f(y).$$

Let $\mathbb{R}^*$ be the set of extended real numbers consisting of the reals augmented with $-\infty$ and $+\infty$. $\pm\infty$ are always accepted as values of the lim-operators.

3. Slope Interval

Although there are different ways of defining slope intervals [13], [14], [7], [5], we will limit ourselves to the following.

DEFINITION 3.1 (Interval slope matrix, [11], p. 27). *Let $f : \mathbb{R}^n \to \mathbb{R}$ and let $\boldsymbol{x}$ be an interval vector. A set $\mathbf{S}$ of vectors in $\mathbb{R}^n$ is said to be a slope set for f over $\boldsymbol{x}$ and centered on the interval vector $\check{\boldsymbol{x}}$ (usually, $\check{\boldsymbol{x}}$ is a point or a very small box) if, for every $x \in \boldsymbol{x}$, and $\check{x} \in \check{\boldsymbol{x}}$,*

$$f(x) - f(\check{x}) = s^T \cdot (x - \check{x}) \quad \textit{for some } s \in \mathbf{S}.$$

Any smallest such slope set will be denoted by $\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{\boldsymbol{x}})$. The smallest interval vector that contains $\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{\boldsymbol{x}})$ is called the slope interval of f over $\boldsymbol{x}$, and it is denoted by $\square\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{\boldsymbol{x}})$.

EXAMPLE 3.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be the continuous function $f(x) = x_1 x_2$, and $\check{x} = (1,1)$. Then*

$$f(x) - f(\check{x}) = (1, x_1) \cdot (x - \check{x}) = (x_2, 1) \cdot (x - \check{x}).$$

Thus, $(1, x_1)$ is a slope of f at $\check{x}$, and so is $(x_2, 1)$. Hence the slope is not unique.

Note. It is not necessary that f be differentiable to obtain $\square\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{\boldsymbol{x}})$.

Interval Newton iteration with slope intervals has potential in global optimization and nonlinear systems solvers, especially when the derivatives of the objective function f have jump discontinuities, such as when f contains terms involving $\|\cdot\|$ or max, [9], [10], [25], [29].

LEMMA 3.1. *Let $f : \mathbb{R} \to \mathbb{R}$, and let $\boldsymbol{x}$ be an interval vector containing $\check{x}$. Then the limiting slope interval is given by*

$$\lim_{w(\boldsymbol{x}) \to 0} \square\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = \left[\liminf_{x \to \check{x}} \frac{f(x) - f(\check{x})}{x - \check{x}}, \limsup_{x \to \check{x}} \frac{f(x) - f(\check{x})}{x - \check{x}} \right].$$

Proof. By definition, $\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x})$ is the smallest set such that

$$\{a : f(x) - f(\check{x}) = a(x - \check{x}), x \in \boldsymbol{x}, x \neq \check{x}\} \subseteq \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}).$$

Thus, any $a \in \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x})$ satisfies

$$\inf_{\substack{x \in \boldsymbol{x} \\ x \neq \check{x}}} \frac{f(x) - f(\check{x})}{x - \check{x}} \leq a \leq \sup_{\substack{x \in \boldsymbol{x} \\ x \neq \check{x}}} \frac{f(x) - f(\check{x})}{x - \check{x}}$$

and

$$\square\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = \left[\inf_{\substack{x \in \boldsymbol{x} \\ x \neq \check{x}}} \frac{f(x) - f(\check{x})}{x - \check{x}}, \sup_{\substack{x \in \boldsymbol{x} \\ x \neq \check{x}}} \frac{f(x) - f(\check{x})}{x - \check{x}} \right].$$

Since $x \to \check{x}$ is equivalent to $w(\boldsymbol{x}) \to 0$, the result holds. □

EXAMPLE 3.2. *Let $f : \mathbb{R} \to \mathbb{R}$ be the continuous function defined by*

$$f(x) = \begin{cases} (x-1)^2, & x \geq 1, \\ 1 - x^2, & x < 1, \end{cases}$$

and $\boldsymbol{x} = [0, 2]$, $\check{x} = 1$. Then

$$\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = [-2, -1] \cup [0, 1], \quad \text{and} \quad \square\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = [-2, 1].$$

EXAMPLE 3.3. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function defined by*

$$f(x, y) = |x - 2| - y^2, \quad \boldsymbol{x} = ([1, 3], [0, 2])^T, \quad \text{and } \check{x} = (2, 1).$$

In this particular example, we can compute a partial slope enclosure with respect to one variable by substituting point values for the other variable, treating the other variable as constant. Thus, the first component of the slope vector is computed by

$$S_1 = \mathbf{S}^\sharp(|x-2|, [1,3], 2) = \{-1, 1\}.$$

Similarly, the second component of the slope vector is

$$S_2 = \mathbf{S}^\sharp(-y^2, [0,2], 1) = [-3, -1].$$

Thus,

$$\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = \mathbf{S} = (\{-1, 1\}, [-3, -1])^T,$$

and

$$\square\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = ([-1, 1], [-3, -1])^T.$$

DEFINITION 3.2 (Directional Slope). *Let $f : \mathbb{R}^n \to \mathbb{R}$, let x be a vector in $\mathbb{R}^n$, and v be any other vector in $\mathbb{R}^n$ with $\|v\| = 1$ (with the Euclidean norm in $\mathbb{R}^n$). The directional slope for f at x in the direction v with step size t is defined by*

$$\mathbf{S}_v(f, t, x) = \frac{f(x+tv) - f(x)}{t}.$$

THEOREM 3.1. *Let v be any unitary vector in $\mathbb{R}^n$, i.e. $\|v\| = 1$ and let $t > 0$, and let $\boldsymbol{x}$ be an interval vector containing x and $x + tv$, and let $\mathbf{S}^\sharp(f, \boldsymbol{x}, x)$ be any minimal slope set for f at x over $\boldsymbol{x}$. Then there exists some $s \in \mathbf{S}^\sharp(f, \boldsymbol{x}, x)$ such that*

$$\mathbf{S}_v(f, t, x) = s^T \cdot v. \tag{1}$$

(s will be denoted by $s_v(t)$, and it is not necessarily unique).

Proof. Let v be any unitary vector in $\mathbb{R}^n$ and let t be small enough to have $y = x + tv \in \boldsymbol{x}$. By the definition of $\mathbf{S}^\sharp(f, \boldsymbol{x}, x)$, there exists some $s \in \mathbf{S}^\sharp(f, \boldsymbol{x}, x)$ such that

$$t\mathbf{S}_v(f, t, x) = f(y) - f(x) = s^T \cdot (y - x) = s^T \cdot tv.$$

Dividing by t, we get (1). □

DEFINITION 3.3 (Directional Derivative). *Let $f : \mathbb{R}^n \to \mathbb{R}$. The usual (one–sided) directional derivative of f at x in the direction $v \in \mathbb{R}^n$ is*

$$f'(x; v) = \lim_{t \downarrow 0} \frac{f(x+tv) - f(x)}{t}$$

when this limit exists.

Note. Observe that $f'(x;v) = \lim_{t\downarrow 0} \mathbf{S}_v(f,t,x)$, when this limit exists.

EXAMPLE 3.4. *Let* $f:\mathbb{R}^2 \to \mathbb{R}$ *be the function defined by*

$$f(x_1,x_2) = x_1 x_2, \quad \boldsymbol{x} = ([1.5,2],[0.5,1.5])^T, \ \text{and} \ \check{x} = (1,1).$$

Then for any $v = (v_1,v_2)$ *we have*

$$\begin{aligned}\frac{f(\check{x}+tv)-f(\check{x})}{t} &= \frac{(1+tv_1)(1+tv_2)-1}{t}\\ &= v_1 + v_2 + tv_1v_2\\ &= (1+tv_1,1)^T\cdot v = (1,1+tv_2)^T\cdot v,\end{aligned}$$

so $s_v(t) = (1+tv_1,1)^T$ *or* $s_v(t) = (1,1+tv_2)^T$. *Thus,* $s_v(t)$ *is not necessarily unique. However, the directional slope*

$$\mathbf{S}_v(f,t,\check{x}) = s_v(t)\cdot v$$

is unique. Taking the limit when $t\downarrow 0$, *we get*

$$f'(\check{x};v) = \lim_{t\downarrow 0} \mathbf{S}_v(f,t,\check{x}) = v_1 + v_2 = (1,1)^T\cdot v.$$

We introduce a new concept of slope, Symmetric Interval Slope, which is an extension of the slope interval for discontinuous functions. The symmetric slope interval is calculated considering slopes with respect to both points $(\check{x}, \lim_{t\uparrow 0} f(\check{x}+te_i))$ and $(\check{x}, \lim_{t\downarrow 0} f(\check{x}+te_i))$, where $e_i, \quad i=1,\ldots,n$ is the i-th coordinate vector.

DEFINITION 3.4. *Let* $f:\mathbb{R}^n \to \mathbb{R}$. *The vector* $\mathbf{SS}$ *is said to be a* symmetric slope set *for* f *over* $\boldsymbol{x}$ *and centered on the interval vector* $\check{\boldsymbol{x}}$ *if, for each coordinate vector* e_i, $x\in\boldsymbol{x}$ *and* $\check{x}\in\check{\boldsymbol{x}}$,

$$f(x) - \lim_{t\uparrow 0} f(\check{x}+te_i) = S_{i1}^T\cdot(x-\check{x})$$

and

$$f(x) - \lim_{t\downarrow 0} f(\check{x}+te_i) = S_{i2}^T\cdot(x-\check{x}),$$

for some $S_{i1}, S_{i2} \in \mathbf{SS}, \quad i=1,\ldots,n$. *Any smallest such set of vectors satisfying this condition will be denoted by* $\mathbf{SS}^\sharp(f,\boldsymbol{x},\check{\boldsymbol{x}})$. *The smallest interval vector that contains* $\mathbf{SS}^\sharp(f,\boldsymbol{x},\check{\boldsymbol{x}})$, $\square\mathbf{SS}^\sharp(f,\boldsymbol{x},\check{\boldsymbol{x}})$, *is called* symmetric slope interval *of* f *over* $\boldsymbol{x}$.

In the one-dimensional case, the symmetric slope interval is calculated considering slopes with respect to both points $(\check{x}, f(\check{x})^-)$ and $(\check{x}, f(\check{x})^+)$. Note that the symmetric slope interval is the same as the slope interval for continuous functions at $\check{x}$.

LEMMA 3.2. *Let $f : \mathbb{R} \to \mathbb{R}$, and let $\boldsymbol{x}$ be an interval vector containing $\check{x}$. Then the limiting symmetric slope interval is given by*

$$\begin{aligned}\lim_{w(\boldsymbol{x})\to 0} [\!]\mathbf{SS}^\sharp(f, \boldsymbol{x}, \check{\boldsymbol{x}}) &= \Big[\min\Big\{\liminf_{x\to\check{x}} \tfrac{f(x)-f(\check{x})^-}{x-\check{x}}, \liminf_{x\to\check{x}} \tfrac{f(x)-f(\check{x})^+}{x-\check{x}}\Big\}, \\ &\qquad \max\Big\{\limsup_{x\to\check{x}} \tfrac{f(x)-f(\check{x})^-}{x-\check{x}}, \limsup_{x\to\check{x}} \tfrac{f(x)-f(\check{x})^+}{x-\check{x}}\Big\}\Big] \\ &= [s, S].\end{aligned}$$

Proof. The proof is analogous to the proof of Lemma 3.1. □

EXAMPLE 3.5. *Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by*

$$f(x) = \begin{cases} -x, & x \leq 0, \\ x - 1, & x > 0, \end{cases}$$

and let $\boldsymbol{x}$ be any interval centered at $\check{x} = 0$. Then

$$\lim_{w(\boldsymbol{x})\to 0} [\!]\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = [-\infty, -1], \qquad \lim_{w(\boldsymbol{x})\to 0} [\!]\mathbf{SS}^\sharp(f, \boldsymbol{x}, \check{x}) = [-\infty, 1].$$

Thus, this example illustrates that for discontinuous functions we have

$$\lim_{w(\boldsymbol{x})\to 0} [\!]\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) \subset \lim_{w(\boldsymbol{x})\to 0} [\!]\mathbf{SS}^\sharp(f, \boldsymbol{x}, \check{x}).$$

4. Generalized gradient

This section presents an alternate construction, the *generalized gradient*, of *locally Lipschitz* functions $f : \mathbb{R}^n \to \mathbb{R}$. The generalized gradient is originally defined for Banach spaces X (See [3] or [34]) and it has been studied to obtain generalized versions of the classical Hamiltonian and Euler-Lagrange equations of the calculus of variations so as to encompass problems in optimal control [27] and to show Lipschitz–type stability in nonsmooth convex problems [8]. In this work, we concentrate our attention on the case $X = \mathbb{R}^n$.

DEFINITION 4.1. $f : \mathbb{R}^n \to \mathbb{R}$ *is* Lipschitz of rank K near $x \in \mathbb{R}^n$ *(or f is* locally Lipschitz at $x \in \mathbb{R}^n$*) if $K \geq 0$ and there is an $\epsilon > 0$ such that*

$$|f(x'') - f(x')| \leq K\|x'' - x'\| \qquad \forall x'', x' \in x + \epsilon B, \tag{2}$$

where B is the unit ball in $\mathbb{R}^n$.

DEFINITION 4.2 (Generalized Directional Derivative [3], p. 25). *Let f be Lipschitz near a given point x, and let v be any other vector in*

$\mathbb{R}^n$. The generalized directional derivative *of f at x in the direction v, denoted $f^\circ(x;v)$, is defined as follows:*

$$f^\circ(x;v) = \limsup_{\substack{y\to x \\ t\downarrow 0}} \frac{f(y+tv)-f(y)}{t}, \tag{3}$$

where $y \in \mathbb{R}^n$ and t is a positive scalar.

It is clear that $f^\circ(x;0) = 0$. This definition does not presuppose the existence of any limit, since it involves an upper limit only.

DEFINITION 4.3. *A function $f : \boldsymbol{x} \to \mathbb{R}$ is* positively homogeneous *on $\boldsymbol{x}$ if $f(\lambda x) = \lambda f(x)$ for all $\lambda > 0$ and $x \in \boldsymbol{x}$.*

PROPOSITION 4.1 ([3], p. 25). *Let f be Lipschitz of rank K near x. Then*

(a) The function $v \to f^\circ(x;v)$ is finite, positively homogeneous, and subadditive on $\mathbb{R}^n$, and satisfies

$$\|f^\circ(x;v)\| \le K\|v\|.$$

(b) $f^\circ(x;v)$ is upper semicontinuous as a function of (x,v) and, as a function of v alone, is Lipschitz of rank K on $\mathbb{R}^n$.

(c) $f^\circ(x;-v) = (-f)^\circ(x;v)$.

Under the conditions of Proposition 4.1 and the Hahn-Banach Theorem there is at least one linear functional $\zeta : \mathbb{R}^n \to \mathbb{R}$ such that, for all v in $\mathbb{R}^n$, one has $f^\circ(x;v) \ge \zeta(v)$. ζ is bounded, so it belongs to the dual space $(\mathbb{R}^n)^*$ of continuous linear functionals on $\mathbb{R}^n$. In finite dimensional spaces, a space and its dual have the same dimension. Hence, they are isomorphic and homeomorphic, and the strong topology (generated at the original space) and the weak topology (generated at the dual space) are identified. So weak compactness is the same as compactness. Henceforth, we identify $(\mathbb{R}^n)^*$ with $\mathbb{R}^n$, and we adopt the convention of $\langle\zeta, v\rangle = \langle v, \zeta\rangle = \zeta(v)$.

DEFINITION 4.4 (Generalized Gradient[3], p. 27).
The generalized gradient *of f at x, denoted $\partial f(x)$, is the subset of $\mathbb{R}^n$ given by*

$$\partial f(x) = \{\zeta \in \mathbb{R}^n : f^\circ(x;v) \ge \langle\zeta, v\rangle, \quad \forall v \in \mathbb{R}^n\}. \tag{4}$$

The following proposition summarizes some properties of the generalized gradient in $\mathbb{R}^n$.

PROPOSITION 4.2 ([3], p. 27). *Let f be Lipschitz of rank K near x, and consider the norm in $\mathbb{R}^n$ defined by*

$$\|\zeta\| := \sup\{\langle \zeta, v\rangle : v \in \mathbb{R}^n, \|v\| \leq 1\}.$$

Then

(a) $\partial f(x)$ is a nonempty, convex, compact subset of $\mathbb{R}^n$ and $\|\zeta\| \leq K$ for every ζ in $\partial f(x)$.

(b) For every v in $\mathbb{R}^n$, one has

$$f^\circ(x; v) = \max\{\langle \zeta, v\rangle : \zeta \in \partial f(x)\}. \tag{5}$$

From (4) and (5) it is equivalent to know the set $\partial f(x)$ or the function $f^\circ(x; .)$; each is obtainable from the other. The next proposition relates the generalized gradient and the subdifferential of a convex function.

PROPOSITION 4.3 ([3], p. 36). *Let D be an open convex subset of $\mathbb{R}^n$. When f is convex on D and Lipschitz near x, then $\partial f(x)$ coincides with the subdifferential at x in the sense of convex analysis, and $f^\circ(x; v)$ coincides with the directional derivative $f'(x; v)$ for each v. The elements of the subdifferential $\partial f(x)$ are called subgradients of f at x.*

EXAMPLE 4.1 ([3], p. 28). *Let $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = |x|$, and $\check{x} = 0$. Since $||x| - |x'|| \leq |x - x'|$ for all $x, x' \in \mathbb{R}$, f is Lipschitz at any x with $K = 1$. Using slope sets on the interval $\boldsymbol{x}$ whose midpoint is $\check{x}$, we know that*

$$\lim_{w(\boldsymbol{x})\to 0} [\!]\mathbf{S}^\sharp(f, \boldsymbol{x}, 0) = [-1, 1].$$

Now, we compute $f^\circ(0; v)$ and $\partial f(0)$.

$$f^\circ(0; v) = \limsup_{\substack{y\to 0\\ t\downarrow 0}} \frac{|y + tv| - |y|}{t}.$$

For $v \geq 0$ we have the following three cases:

$$\left.\begin{array}{lll} y > 0, y + tv > 0 & \to & \frac{y+tv-y}{t} = v \\ y < 0, y + tv > 0 & \to & \frac{y+tv+y}{t} = \frac{2y}{t} + v \to [-\infty, v] \\ y < 0, y + tv < 0 & \to & \frac{-(y+tv)+y}{t} = -v \end{array}\right\} \to f^\circ(0; v) = v.$$

Similarly, for $v < 0$ we have the following three cases:

$$\left.\begin{array}{lll} y < 0, y + tv < 0 & \to & \frac{-(y+tv)+y}{t} = -v \\ y > 0, y + tv > 0 & \to & \frac{y+tv-y}{t} = v \\ y > 0, y + tv < 0 & \to & \frac{-(y+tv)-y}{t} = -\frac{2y}{t} - v \to [-\infty, -v] \end{array}\right\}.$$

This implies $f^\circ(0; v) = -v$. *Hence,* $f^\circ(0; v) = |v|$, *and*

$$\partial f(0) = \{\zeta \in \mathbb{R} : |v| \geq \zeta v\} = [-1, 1].$$

Thus, in this example we have

$$\partial f(0) = \lim_{w(\boldsymbol{x}) \to 0} \square \mathbf{S}^\sharp(f, \boldsymbol{x}, 0) = [-1, 1].$$

EXAMPLE 4.2. *Let* f *be the function defined by* $f(-x) = f(x)$ *for all* $x \in \mathbb{R}$, *let* $\boldsymbol{x}$ *be any interval containing* $\check{x} = 0$, *and*

$$f(x) = \begin{cases} x - \frac{1}{2}, & \frac{1}{2} \leq x \leq 1, \\ \frac{1}{2} - x, & \frac{1}{4} \leq x < \frac{1}{2}, \\ x, & \frac{1}{8} \leq x < \frac{1}{4}, \\ \frac{1}{2} - x - \sum_{i=2}^{k} \frac{1}{2^{(2i-1)}}, & \frac{1}{2^{(2k)}} \leq x \leq \frac{1}{2^{(2k-1)}}, k = 2, 3, \ldots \\ x + \sum_{i=2}^{k} \frac{1}{2^{(2i-1)}}, & \frac{1}{2^{(2k+1)}} \leq x \leq \frac{1}{2^{(2k)}}, k = 2, 3, \ldots \end{cases}$$

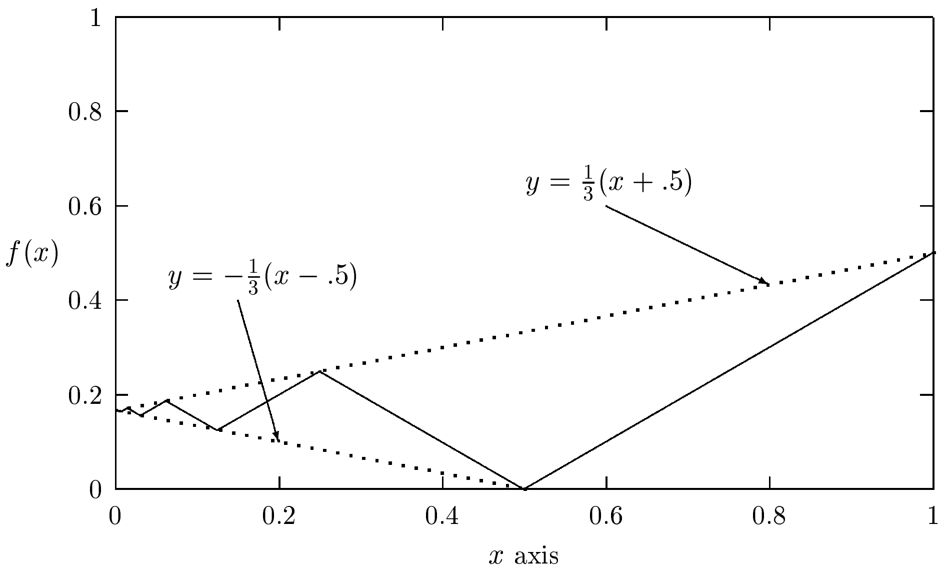

Figure 1. Graph of $f(x)$ in Example 4.2

The graph of f *as* $x \geq 0$ *is shown in Figure 1. Since* f *is linear in each interval with slopes either* 1 *or* -1, f *is Lipschitz near* $\check{x} = 0$. *Taking the sequence of points* $x_i = \frac{3}{2^i}$, *we have* $f'(x_i) = (-1)^{i+1}$ *and* $\partial f(0) = [-1, 1]$. *Also, the upper and lower bounds for the slopes occur at members of the sequence* $x = x_k = \frac{1}{2^k}$. *The points corresponding to the upper bound on the slope are on the straight line* $y = \frac{1}{3}(x + \frac{1}{2})$

and the points corresponding to the lower bound on the slope are in the straight line $y = -\frac{1}{3}(x - \frac{1}{2})$. Thus, in this example we have

$$\lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = \left[-\frac{1}{3}, \frac{1}{3}\right] \subset [-1, 1] = \partial f(\check{x}).$$

This example is not typical, but it shows that for f Lipschitz near $\check{x}$ not always is true that

$$\partial f(\check{x}) \subset \lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}).$$

Fundamental relationships between the concepts directional slope and generalized directional derivative imply inclusion relationships between slope sets and generalized gradients. Within the following class of functions, these inclusions become equalities.

DEFINITION 4.5. *f is said to be* regular *at $\check{x}$ provided*

(a) For all v, the usual one–sided directional derivative

$$f'(\check{x} : v) = \lim_{t\downarrow 0} \frac{f(\check{x} + tv) - f(\check{x})}{t} = \lim_{t\downarrow 0} \mathbf{S}_v(f, t, \check{x}),$$

exists.

(b) For all v, $f'(\check{x} : v) = f^\circ(\check{x}; v)$.

In [2] $f'(\check{x} : v)$ is denoted by $\delta^+ f(\check{x}; v)$. We recall Rademacher's Theorem, which states that a function which is locally Lipschitz on an open subset of $\mathbb{R}$ is differentiable almost everywhere (a.e.) (in the sense of Lebesgue measure) on that subset. The set of points where f is not differentiable is denoted by Ω_f. If f is locally Lipschitz near $\check{x}$, f is Lipschitz in a delta neighbourhood B of $\check{x}$. This implies that f is absolutely continuous in B and for any x in B we have that

$$f(x) - f(\check{x}) = \int_{\check{x}}^{x} f'(y)dy.$$

Examples 4.1 and 4.2 illustrate the following result that characterizes the relationship between generalized gradient and slope interval in one dimension.

THEOREM 4.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be Lipschitz near $\check{x}$ and let $\boldsymbol{x}$ be any interval centered at $\check{x}$. Then*

$$\lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) \subseteq \partial f(\check{x}). \tag{6}$$

If f is regular at $\check{x}$, then equality holds.

Proof. Let $\delta > 0$ be so small that f is Lipschitz in the open set $B_\delta = \{x : 0 < |x - \check{x}| < \delta,\}$. Then $\Omega_f \cap B_\delta$ has Lebesgue measure 0. For $x \in B_\delta$ and $x > \check{x}$ we have that

$$\frac{f(x) - f(\check{x})}{x - \check{x}} v = \frac{1}{x - \check{x}} \int_{\check{x}}^{x} f'(y) v dy \leq \frac{1}{x - \check{x}} \int_{\check{x}}^{x} \sup_{y \in [\check{x}, x] - \Omega_f} f'(y) v dy,$$

since from Rademacher's theorem f' exist a.e on $(\check{x}, x)$. Similar inequality holds for $x < \check{x}$. From (5), for any $\epsilon > 0$, there is a $\delta > 0$ such that $x \in B_\delta$ and

$$\frac{f(x) - f(\check{x})}{x - \check{x}} v \leq f^\circ(\check{x}; v) + \epsilon.$$

Since ϵ is arbitrary, we obtain

$$\limsup_{x \to \check{x}} \frac{f(x) - f(\check{x})}{x - \check{x}} v \leq f^\circ(\check{x}; v).$$

From Lemma 3.1, this implies that

$$f^\circ(\check{x}; v) \geq \begin{cases} v \limsup_{x \to \check{x}} \frac{f(x) - f(\check{x})}{x - \check{x}}, & v \geq 0, \\ v \liminf_{x \to \check{x}} \frac{f(x) - f(\check{x})}{x - \check{x}}, & v < 0, \end{cases} = \begin{cases} vM, & v \geq 0, \\ vm, & v < 0. \end{cases}$$

Thus,

$$f^\circ(\check{x}; v) \geq mv, \quad \text{and} \quad f^\circ(\check{x}; v) \geq Mv, \qquad \text{for all} \quad v \in \mathbb{R}.$$

Therefore $m, M \in \partial f(\check{x})$. Since $\partial f(\check{x})$ is a convex set, we have

$$\lim_{w(\boldsymbol{x}) \to 0} \square\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = [m, M] \subseteq \partial f(\check{x}),$$

so (6) holds. If f is regular at $\check{x}$, then we have

$$f^\circ(\check{x}; v) = \begin{cases} vM, & v \geq 0, \\ vm, & v < 0, \end{cases}$$

and since $m \leq M$, we obtain

$$\begin{aligned} \partial f(\check{x}) &= \{\zeta \in \mathbb{R} : f^\circ(\check{x}; v) \geq \zeta v \quad \text{for all} \quad v \in \mathbb{R}\} \\ &= \{\zeta \in \mathbb{R} : Mv \geq \zeta v, \quad v \geq 0\} \cap \{\zeta \in \mathbb{R} : mv \geq \zeta v, \quad v < 0\} \\ &= [-\infty, M] \cap [m, +\infty] \\ &= [m, M] \\ &= \lim_{w(\boldsymbol{x}) \to 0} \square\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}). \end{aligned}$$

Thus, (6) holds with equality. $\square$

The next theorem gives us the relationship between gradients and generalized gradients of Lipschitz functions.

THEOREM 4.2 ([3], p. 63). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz near x, and suppose S is any set of Lebesgue measure 0 in $\mathbb{R}^n$. Then*

$$\partial f(x) = co\left\{\lim \nabla f(x_i) : x_i \to x, \quad x_i \notin S \cup \Omega_f\right\}.$$

COROLLARY 4.1 ([3], p. 64).

$$f^\circ(\check{x}; v) = \limsup_{y \to \check{x}} \left\{\nabla f(y) \cdot v : y \notin S \cup \Omega_f\right\}.$$

The only way we presently know to obtain a relationship between generalized gradient and slope interval in multiple dimensions is to consider the relationship between the projections of these two concepts. Let $\boldsymbol{x} = \boldsymbol{x}_1 \times \ldots \times \boldsymbol{x}_n$, where $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are intervals in $\mathbb{R}$.

DEFINITION 4.6. *Let $f : \boldsymbol{x} \subset \mathbb{R}^n \to \mathbb{R}$ be Lipschitz near $\check{x}$. For $i = 1, \ldots, n$, the i-th projection $\lim_{w(\boldsymbol{x})\to 0} \pi_i \square \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x})$ is defined as the set*

$$\lim_{w(\boldsymbol{x})\to 0} co\{s_i \in \mathbb{R} : \exists s \in \mathbb{R}^n,\ f(x) - f(\check{x}) = s \cdot (x - \check{x}),\ for\ x \in \boldsymbol{x}, x \neq \check{x}\},$$

and the i-th projection of the generalized gradient of f at $\check{x}$, $\pi_i \partial f(\check{x})$ is defined by

$$\pi_i \partial f(\check{x}) = \{x_i \in \mathbb{R} : \exists x = (x_1, \ldots, x_i, \ldots, x_n) \in \partial f(\check{x})\}.$$

LEMMA 4.1. *Let $f : \boldsymbol{x} \subset \mathbb{R}^n \to \mathbb{R}$ be Lipschitz of rank K near $\check{x}$. Then for any $v \in \mathbb{R}^n$, $\limsup_{t\downarrow 0} \mathbf{S}_v(f, t, \check{x})$ is bounded and*

$$\limsup_{t \downarrow 0} \mathbf{S}_v(f, t, \check{x}) \leq f^\circ(\check{x}; v).$$

Proof. Let v any vector in $\mathbb{R}^n$. By Proposition 4.1, we have

$$-K\|v\| \leq \limsup_{t\downarrow 0} \mathbf{S}_v(f, t, \check{x}) = \limsup_{t \downarrow 0} \frac{f(\check{x} + tv) - f(\check{x})}{t}$$

$$\leq \limsup_{\substack{t\downarrow 0 \\ y \to \check{x}}} \frac{f(y + tv) - f(y)}{t} = f^\circ(\check{x}; v) \leq K\|v\|.$$

The second inequality holds since the set $\{y : y \to \check{x}\}$ contains $\{\check{x}\}$, and by a monotonicity property of lim sup. $\square$

LEMMA 4.2. *Let $f : \boldsymbol{x} \subset \mathbb{R}^n \to \mathbb{R}$ be Lipschitz of rank K near $\check{x}$. Then*

$$\lim_{w(\boldsymbol{x})\to 0} \pi_i \square \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) \subseteq \pi_i \partial f(\check{x}).$$

If f is regular at $\check{x}$, then equality holds.

Proof. Considering $t = w(\boldsymbol{x})$, and vectors $v \in \mathbb{R}^n$ with $\|v\| = 1$, we get

$$\begin{aligned}\lim_{w(\boldsymbol{x})\to 0} \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) &= \lim_{t\downarrow 0} \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) \\ &= \lim_{t\downarrow 0} \left\{ s_v(\zeta) : \mathbf{S}_v(f, \zeta, \check{x}) = s_v(\zeta)^T \cdot v; 0 < \zeta \le t, \forall v \in \mathbb{R}^n \right\}.\end{aligned}$$

Since $\limsup_{t\downarrow 0} \left\{ s_v(t)^T \cdot v \right\}$ is bounded for all $v \in \mathbb{R}^n$ by Lemma 4.1, then $\lim_{t\downarrow 0} \{s_v(\zeta)\}$ is bounded.

Suppose there exists an $s_i \in \lim_{t\downarrow 0} \pi_i \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x})$ and $s_i \notin \pi_i \partial f(\check{x})$. Then there exists an unitary vector $v \in \mathbb{R}^n$ and $s = \lim_{t\downarrow 0} \mathbf{s}_v(t)$ such that

$$s \in \lim_{t\downarrow 0} \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}), \quad s \notin \partial f(\check{x}), \quad \text{and} \quad \lim_{t\downarrow 0} s_{v_i}(t) = s_i.$$

This means that

$$f^\circ(\check{x}; v) < s^T \cdot v = \limsup_{t\downarrow 0} s_v(t)^T \cdot v \le \limsup_{t\downarrow 0} \mathbf{S}_v(f, t, \check{x}).$$

This contradicts Lemma 4.1. Thus,

$$\lim_{t\downarrow 0} \pi_i \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) \subseteq \pi_i \partial f(\check{x}).$$

Conversely, if f is regular at $\check{x}$, and $s_i \in \pi_i \partial f(\check{x})$, there exists an $s \in \partial f(\check{x})$ with $s = (s_1, \cdots, s_i, \cdots, s_n)$, and $\forall v \in \mathbb{R}^n$, we have

$$s^T \cdot v \le f^\circ(\check{x}; v) = \lim_{t\downarrow 0} \mathbf{S}_v(f, t, \check{x}).$$

By Theorem 3.1, for any $v \in \mathbb{R}^n$ there exist $s_v(t) \in \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x})$ such that $\mathbf{S}_v(f, t, \check{x}) = s_v(t)^T \cdot v$. Then we have

$$s^T \cdot v \le \lim_{t\downarrow 0} s_v(t)^T \cdot v.$$

If $v = v_i e_i$, for all $v_i \in \mathbb{R}$ we have reduced the problem to the one-dimensional case. Since

$$s_i v_i \le \lim_{t\downarrow 0} s_{v\,i}(t) v_i \le \limsup_{w(\boldsymbol{x})\downarrow 0} s_{v\,i}(t) v_i \le \begin{cases} \limsup_{w(\boldsymbol{x})\to 0} \pi_i \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) v_i, & v_i \ge 0 \\ \liminf_{w(\boldsymbol{x})\to 0} \pi_i \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) v_i, & v_i < 0, \end{cases}$$

then $\quad s_i \in \lim_{w(\boldsymbol{x})\to 0} \pi_i \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x})$. $\square$

For f regular at $x = (x_1, x_2)$, is shown in [3] that

$$\partial f(x_1, x_2) \subset \partial_1 f(x_1, x_2) \times \partial_2 f(x_1, x_2). \tag{7}$$

THEOREM 4.3. *Let $f : \boldsymbol{x} \subset \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz near $\check{x}$ and regular at $\check{x}$. Then*

$$\lim_{w(\boldsymbol{x})\to 0} \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) \subseteq \partial f(\check{x}) \subseteq \pi_1 \partial f(\check{x}) \times \cdots \times \pi_n \partial f(\check{x}) = \lim_{w(\boldsymbol{x})\to 0} \square \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}).$$

Proof. The second inclusion follows by extending (7) to dimension n, and applying Lemma 4.2. For the first inclusion suppose there exists s such that $s \in \lim_{w(\boldsymbol{x})\to 0} \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x})$, and $s \notin \partial f(\check{x})$. Then for some unitary vector $v \in \mathbb{R}^n$, and $t = w(\boldsymbol{x})$, $s = \lim_{t\downarrow 0} \mathbf{s}_v(t)$ and

$$f^\circ(\check{x}; v) < s^T \cdot v = \lim_{t\downarrow 0} s_v(t)^T \cdot v \le \limsup_{t\downarrow 0} \mathbf{S}_v(f, t, \check{x}).$$

This contradicts Lemma 4.1. □

The next example illustrates this theorem.

EXAMPLE 4.3. *Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$, defined by*

$$f(x, y) = \max\{\min\{x, -y\}, y - x\}.$$

Taking $\mathbb{R}^2 = C_1 \cup C_2 \cup C_3$, with

$$C_1 = \{(x, y) : y \le 2x \quad \textit{and} \quad y \le -x\},$$

$$C_2 = \{(x, y) : y \le x/2 \quad \textit{and} \quad y \ge -x\},$$

$$C_3 = \{(x, y) : y \ge 2x \quad \textit{or} \quad y \ge x/2\},$$

we have

$$f(x, y) = \begin{cases} x, & (x, y) \in C_1, \\ -y, & (x, y) \in C_2, \\ y - x, & (x, y) \in C_3. \end{cases}$$

(a) Let $\check{x} = (0, 0)$. f is Lipschitz near $\check{x}$ since it is linear in each region of $\mathbb{R}^2$. However, f is not regular at $\check{x}$. In fact,

$$\begin{aligned} f^\circ(\check{x}; v) &= \limsup\{(0, -1) \cdot v, (1, 0) \cdot v, (-1, 1) \cdot v : v \in \mathbb{R}^2\} \\ &= \begin{cases} -v_2, & v \in C_1, \\ v_1, & v \in C_2, \\ v_2 - v_1, & v \in C_3, \end{cases} \end{aligned}$$

and

$$\begin{aligned} f'(\check{x}; v) &= \lim_{t\downarrow 0} \frac{f(\check{x} + tv) - f(\check{x})}{t} \\ &= \begin{cases} v_1, & v \in C_1, \\ -v_2, & v \in C_2, \\ v_2 - v_1, & v \in C_3, \end{cases} \end{aligned}$$

exists for all $v \in \mathbb{R}^2$, but $f'(\check{x}; v) \ne f^\circ(\check{x}; v)$.

(b) The set $S = \partial C_1 \cup \partial C_2 \cup \partial C_3 = \Omega_f$, has Lebesgue measure 0. If $(x, y) \notin S$, then $\nabla f(x, y) \in \{(1,0), (0,-1), (-1,1)\}$.

By Theorem 4.2, $\partial f(\check{x}) = co\{(1,0), (0,-1), (-1,1)\}$, which is the triangle with these three vertices. Taking partial generalized gradients, we have

$$\begin{array}{ll} f(x,0) = \max\{0,-x\} & \to \partial_x f(\check{x}) = [-1,0], \\ f(0,y) = \max\{0,y\} & \to \partial_y f(\check{x}) = [0,1], \\ \pi_1 \partial f(\check{x}) = [-1,1], & \\ \pi_2 \partial f(\check{x}) = [-1,1]. & \end{array}$$

So, we have

$$\partial_x f(\check{x}) \times \partial_y f(\check{x}) \not\subset \partial f(\check{x}) \not\subset \partial_x f(\check{x}) \times \partial_y f(\check{x}).$$

(c) Let $\boldsymbol{x}$ be any box centered at $\check{x} = (0,0)$.

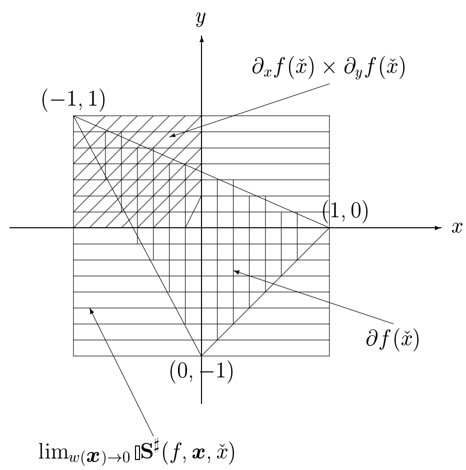

Figure 2. $\partial_x f(\check{x}) \times \partial_y f(\check{x}), \partial f(\check{x})$, and $\lim_{w(\boldsymbol{x})\to 0} \square\mathbf{S}^{\sharp}(f, \boldsymbol{x}, \check{x})$ for Example 4.3

Letting $x = (x_1, x_2)$, $f(\check{x}) = 0$, we consider the slope sets.

$$\begin{array}{lll} x \in C_1 : & f(x) = x_1 = (1,0)(x_1,x_2)^T & \to s = (1,0)^T, \\ x \in C_2 : & f(x) = -x_2 = (0,-1)(x_1,x_2)^T & \to s = (0,-1)^T, \\ x \in C_3 : & f(x) = x_2 - x_1 = (-1,1)(x_1,x_2)^T & \to s = (-1,1)^T. \end{array}$$

Thus,

$$\lim_{w(\boldsymbol{x})\to 0} \mathbf{S}^{\sharp}(f, \boldsymbol{x}, \check{x}) = \{(1,0)^T, (0,-1)^T, (-1,1)^T\},$$

and

$$\lim_{w(\boldsymbol{x})\to 0} \square\mathbf{S}^{\sharp}(f, \boldsymbol{x}, \check{x}) = [-1,1] \times [-1,1].$$

Clearly, the sharpest slope set is easily obtainable and

$$\partial f(\check{x}) \subset \pi_1 \partial f(\check{x}) \times \pi_2 \partial f(\check{x}) = \lim_{w(\boldsymbol{x})\to 0} [\![\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = [-1,1] \times [-1,1].$$

Here, the interval hull of $\partial f(\check{x})$ *is*

$$\lim_{w(\boldsymbol{x})\to 0} [\![\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}).$$

These relationships for Example 4.3 are illustrated in Figure 2.

5. Optimality conditions

This section contains generalizations of the *Karush-Kuhn-Tucker* optimality conditions for nonsmooth constrained optimization problems.

Let $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $(i = 1, 2, \ldots, k)$, and $h_j : \mathbb{R}^n \to \mathbb{R}$, $(j = 1, 2, \ldots, m)$ be continuous functions in a convex set $D \subseteq \mathbb{R}^n$. Let $\boldsymbol{x} \subseteq D$, and consider the constrained optimization problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & \\ g_i(x) \leq 0, & i = 1, 2, \ldots, k, \\ h_j(x) = 0, & j = 1, 2, \ldots, m, \\ x \in \boldsymbol{x} & \end{array} \tag{8}$$

If all functions are Lipschitz functions near $\check{x}$, the generalized gradients yield a generalization of the *Karush-Kuhn-Tucker* neccesary conditions of optimality [3],[34]. In fact, if $\check{x}$ is a solution of (8) there exist $\lambda_i \geq 0$, $i = 1, \ldots, m$, and $\mu_j \geq 0$, $j = 1, 2, \ldots, k$ satisfying

$$\left\{ \begin{array}{l} 0 \ \in \partial f(\check{x}) + \sum_{i=1}^m \lambda_i \partial g_i(\check{x}) + \sum_{j=1}^k \mu_j \partial h_j(\check{x}), \\ \quad \lambda_i g_i(\check{x}) = 0, \quad i = 1, \ldots, m, \end{array} \right. \tag{9}$$

In [22], [6], and [30] are introduced interval techniques used in the solution of (8) and based in inclusions of the generalized gradients. The next theorem shows that interval slopes also yield a practical generalization of the *Karush-Kuhn-Tucker conditions.*

THEOREM 5.1. *If* $\check{x}$ *is a solution of (8)there exist* $\lambda_i \geq 0$, $i = 1, \ldots, m$, *and* $\mu_j \geq 0$, $j = 1, 2, \ldots, k$ *satisfying*

$$\left\{ \begin{array}{rl} 0 & \in \lim_{w(\boldsymbol{x})\to 0} [\![\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) + \sum_{i=1}^m \lambda_i \lim_{w(\boldsymbol{x})\to 0} [\![\mathbf{S}^\sharp(g_i, \boldsymbol{x}, \check{x}) \\ & + \sum_{j=1}^k \mu_j \lim_{w(\boldsymbol{x})\to 0} [\![\mathbf{S}^\sharp(h_j, \boldsymbol{x}, \check{x}) \\ \lambda_i g_i(\check{x}) & = 0, \quad i = 1, \ldots, m, \end{array} \right. \tag{10}$$

Proof. If all functions are Lipschitz functions near $\check{x}$ and regular at $\check{x}$, Theorem 4.3 and (9) prove (10).

In [20] is shown that interval methods have no dificulties in handling nonsmooth problems, because neither the construction of inclusion functions nor the application of monotonicity tests depends on the smoothness of the objective function. □

6. Semigradient

The *semigradient* is a generalization of the concept of generalized gradient to arbitrary not necessarily continuous functions. This extension to arbitrary functions $f : \mathbb{R}^n \to \mathbb{R}^*$ was introduced in [17].

In [3], $\partial f(x)$ is defined for extended-valued functions $f : \mathbb{R}^n \to \mathbb{R}^*$, as long as f is finite at x. The generalized gradient is characterized geometrically in terms of normals to the *epigraph* of f at the point $(x, f(x))$, which is denoted and defined by

$$\text{epi}f = \left\{(x, r) \in \mathbb{R}^{n+1} : f(x) \leq r\right\}$$

$$T_{\text{epi}f}(x, f(x)) = \text{epi}f^{\circ}(x; \cdot),$$

$$N_{\text{epi}f}(x, f(x)) = \left\{\zeta : (\zeta, v) \leq 0 \text{ for all } v \text{ in } T_{\text{epi}f}(x, f(x))\right\},$$

$$\partial f(x) = \left\{\zeta : (\zeta, -1) \in N_{\text{epi}f}(x, f(x))\right\},$$

where $T_{\text{epi}f}(x, f(x))$ and $N_{\text{epi}f}(x, f(x))$ are the *tangent cone and normal cone* to the epi$f$ at $(x, f(x))$ respectively. In the non-Lipschitz case, the direct characterization of f° involves some complicated limits for which the following notion due to Rockafellar [26] is in order. The expression $(y, \alpha) \downarrow_f x$ shall mean that $(y, \alpha) \in \text{epi}f, y \to x, \alpha \to f(x)$. The function $f^{\circ}(x; .) : \mathbb{R}^n \to \mathbb{R}^*$ is defined as follows

$$f^{\circ}(x; v) = \lim_{\epsilon \downarrow 0} \limsup_{\substack{(y,\alpha)\downarrow_f x \\ t \downarrow 0}} \inf_{w \in v + \epsilon B} \frac{f(y + tw) - \alpha}{t}.$$

If f is lower semicontinuous at x, then $f^{\circ}(x; v)$ is given by the slightly simpler expression

$$f^{\circ}(x; v) = \lim_{\epsilon \downarrow 0} \limsup_{\substack{y \downarrow_f x \\ t \downarrow 0}} \inf_{w \in v + \epsilon B} \frac{f(y + tw) - f(y)}{t},$$

where $y \downarrow_f x$ signifies that y and $f(y)$ converge to x and $f(x)$ respectively. The following result shows that the extended f° plays the same role vis-à-vis ∂f as it did in the Lipschitz case.

COROLLARY 6.1 (Rockafellar [26]). *One has* $\partial f(x) = \emptyset$ *iff* $f^\circ(x;0) = -\infty$. *Otherwise, one has*

$$\partial f(x) = \{\zeta \in \mathbb{R}^n : f^\circ(x;v) \geq \langle \zeta, v \rangle, \quad \forall v \in \mathbb{R}^n\},$$

and

$$f^\circ(x;v) = \sup\{\langle \zeta, v \rangle : \zeta \in \partial f(x)\}.$$

If f is discontinuous and finite at x, $\partial f(x)$ only incorporates information of the epigraph of f near to the point $(x, f(x))$, and in some cases $\partial f(x)$ discards the behavior of f near x. The semigradient takes into account the behavior of f near x.

The following proposition, due to Moreau [17], provides an alternative characterization of $f^\circ(x;v)$.

PROPOSITION 6.1 ([17]). $f^\circ(x;v)$ *is given by*

$$f^\circ(x;v) = \max\left\{\limsup_{i\to\infty} \frac{f(x_i + t_i v) - f(x_i)}{t_i} : \forall\, x_i \to x, \quad t_i \downarrow 0\right\},$$

where the maximum is obtained over all sequences $\{x_i\}_{i=1}^\infty$ *in* $\mathbb{R}^n$ *converging to* x *and* $\{t_i\}_{i=1}^\infty$ *in* $(0,\infty)$ *converging to* 0.

DEFINITION 6.1 ([17]). *Assume an arbitrary function* $f : \mathbb{R}^n \to \mathbb{R}^*$. *Let* $x, v \in \mathbb{R}^n$. *Let* $x \in \mathbb{R}^n$. *The* semigradient *of* f *at* x, *denoted* $\mathbf{SG}f(x)$, *is a subset of* $\mathbb{R}^n$ *defined by*

$$\mathbf{SG}f(x) = \{\zeta \in \mathbb{R}^n : f^\circ(x;v) \geq \langle \zeta, v \rangle, \quad \forall v \in \mathbb{R}^n\},$$

where f° *is as in Definition 4.2.*

$\mathbf{SG}f(x)$ is as $\partial f(x)$ a closed convex set, possibly empty. The next theorem relates these two concepts.

THEOREM 6.1 ([17]). *Let* $x \in \mathbb{R}^n$. *For arbitrary functions* $f : \mathbb{R}^n \to \mathbb{R}^*$, *the following inclusion holds:*

$$\partial f(x) \subseteq \mathbf{SG}f(x).$$

In the particular case that $f : \mathbb{R}^n \to \mathbb{R}$ *is locally Lipschitz, this inclusion reduces to*

$$\partial f(x) = \mathbf{SG}f(x).$$

The relationships between symmetric slope intervals and semigradients in this section extend the results presented in the previous section to arbitrary functions.

THEOREM 6.2. *Let $f:\mathbb{R}\to\mathbb{R}$ be an arbitrary function and let $\boldsymbol{x}$ be any interval centered at $\check{x}$. Then*

$$\mathbf{SG}f(\check{x}) = \lim_{w(\boldsymbol{x})\to 0} \square\mathbf{SS}^\sharp(f,\boldsymbol{x},\check{x}). \tag{11}$$

Proof. From Proposition 6.1 we get

$$f^\circ(\check{x};v) = \max\left\{\limsup_{i\to\infty}\frac{f(x_i+t_iv)-f(x_i)}{t_i} : \forall\, x_i\to\check{x},\quad t_i\downarrow 0\right\}.$$

since $x_i\to\check{x}$ and $t_i\downarrow 0$, $y_i = x_i+t_iv\to\check{x}$. Then we have

$$\begin{aligned}
f^\circ(\check{x};v) &= \max\left\{\limsup_{i\to\infty}\frac{f(y_i)-f(x_i)}{y_i-x_i}v : \forall\, x_i\to\check{x},\quad t_i\downarrow 0,\quad y_i = x_i+t_iv\right\}\\
&\geq \max\left\{\limsup_{i\to\infty}\frac{f(y_i)-f(\check{x})^-}{y_i-\check{x}}v, \limsup_{i\to\infty}\frac{f(y_i)-f(\check{x})^+}{y_i-\check{x}}v : \forall\, y_i\to\check{x}\right\}\\
&\geq \max\left\{\limsup_{x\to\check{x}}\frac{f(x)-f(\check{x})^-}{x-\check{x}}v, \limsup_{x\to\check{x}}\frac{f(x)-f(\check{x})^+}{x-\check{x}}v\right\}\\
&= \begin{cases} v\max\left\{\limsup_{x\to\check{x}}\frac{f(x)-f(\check{x})^-}{x-\check{x}}, \limsup_{x\to\check{x}}\frac{f(x)-f(\check{x})^+}{x-\check{x}}\right\}, & v\geq 0,\\ v\min\left\{\liminf_{x\to\check{x}}\frac{f(x)-f(\check{x})^-}{x-\check{x}}, \liminf_{x\to\check{x}}\frac{f(x)-f(\check{x})^+}{x-\check{x}}\right\}, & v<0,\end{cases}\\
&= \begin{cases} vS, & v\geq 0,\\ vs, & v<0,\end{cases}
\end{aligned}$$

where s and S are given in Lemma 3.2. Thus, $f^\circ(\check{x};v)\geq sv$ and $f^\circ(\check{x};v)\geq Sv$ for all $v\in\mathbb{R}$, this implies that s and S belong to $\mathbf{SG}f(\check{x})$. Since $\mathbf{SG}f(\check{x})$ is a closed convex set, we get

$$\lim_{w(\boldsymbol{x})\to 0}\square\mathbf{SS}^\sharp(f,\boldsymbol{x},\check{x}) = [s,S]\subseteq\mathbf{SG}f(\check{x}).$$

In the other hand, let $\zeta\in\mathbf{SG}f(\check{x})$, and

$$\begin{aligned}
f^\circ(\check{x};v) &= \max\left\{\limsup_{i\to\infty}\frac{f(y_i)-f(x_i)}{y_i-x_i}v : \forall\, x_i\to\check{x},\quad t_i\downarrow 0,\quad y_i = x_i+t_iv\right\}\\
&= \begin{cases} v\limsup_{i\to\infty}\frac{f(y_i)-f(x_i)}{y_i-x_i}, & v\geq 0,\\ v\liminf_{i\to\infty}\frac{f(y_i)-f(x_i)}{y_i-x_i}, & v<0.\end{cases}
\end{aligned}$$

From definition of $\mathbf{SG}f(\check{x})$, we have $f^\circ(x;v)\geq\zeta v$, $\forall v\in\mathbb{R}$. Thus,

$$\liminf_{i\to\infty}\frac{f(y_i)-f(x_i)}{y_i-x_i}\leq\zeta\leq\limsup_{i\to\infty}\frac{f(y_i)-f(x_i)}{y_i-x_i},\quad \forall x_i,y_i\to\check{x}.$$

Lemma 3.2 and the previous inequalities imply that, $s\leq\zeta\leq S$, and we obtain

$$\mathbf{SG}f(\check{x})\subseteq[s,S] = \lim_{w(\boldsymbol{x})\to 0}\square\mathbf{SS}^\sharp(f,\boldsymbol{x},\check{x}).$$

This completes the proof of (11). $\square$

The following examples illustrate this result.

EXAMPLE 6.1. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be the function in the Example 3.5 defined by*

$$f(x) = \begin{cases} -x, & x \le 0, \\ x-1, & x > 0, \end{cases}$$

and $\boldsymbol{x}$ *is any interval centered at* $\check{x} = 0$*. Note that* f *is upper semi-*

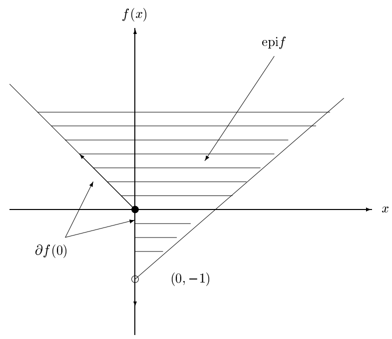

Figure 3. Graph of epi$f$ near $(0,0)$ and $\partial f(0)$ in Example 5.1

continuous at $\check{x} = 0$ *and has a local infimum (which is actually global) at* $\check{x}$ *see Figure 3. The Clarke generalized gradient of* f *only incorporates information of the epigraph of* f *near* $(0,0)$ *and thus discards the behavior of* f *for* $x > 0$*, which is crucial for* f *to have an infimum or not. Therefore, although this function has an infimum at* $\check{x}$*,* $0 \notin \partial f(0) = [-\infty, -1]$*. On the other hand,* $0 \in \mathbf{SG}f(0) = [-\infty, 1] = \lim_{w(\boldsymbol{x})\to 0} \square\mathbf{SS}^\sharp(f, \boldsymbol{x}, \check{x})$*. Thus, for this example we have*

$$\lim_{w(\boldsymbol{x})\to 0} \square\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = \partial f(0) \subset \mathbf{SG}f(0) = \lim_{w(\boldsymbol{x})\to 0} \square\mathbf{SS}^\sharp(f, \boldsymbol{x}, \check{x}).$$

EXAMPLE 6.2. *Consider the function* $f : \mathbb{R} \to \mathbb{R}$ *defined by*

$$f(x) = \begin{cases} x, & x < 0, \\ -x-1, & x \ge 0, \end{cases}$$

and $\boldsymbol{x}$ *is any interval centered at* $\check{x} = 0$*. Note that* f *is lower semi-continuous at* $\check{x} = 0$ *and has a local supremum (which is actually global) at* $\check{x}$ *see Figure 4. The Clarke generalized gradient of* f *only incorporates information of the epigraph of* f *near* $(0,-1)$ *and thus discards the behavior of* f *for* $x < 0$*, which is crucial for* f *to have a supremum or not. Therefore, although this function has a supremum at*

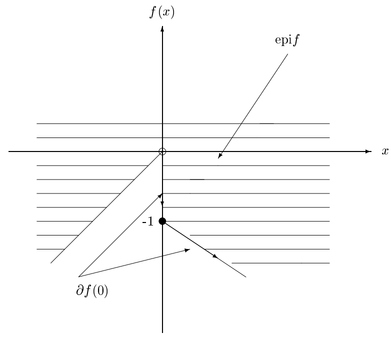

Figure 4. Graph of epi$f$ near $(0,-1)$ and $\partial f(0)$ in Example 5.2

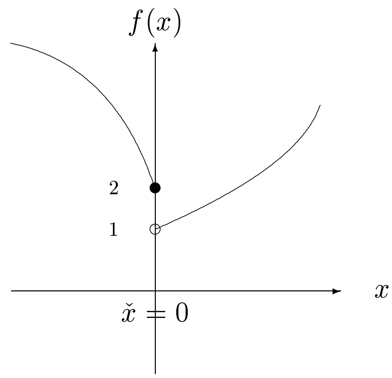

Figure 5. Graph of $f(x)$ in Example 5.3

$\check{x}$, $0 \notin \partial f(0) = [-\infty, -1]$. *On the other hand,* $0 \in \mathbf{SG}f(0) = [-\infty, 1] = \lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{SS}^\sharp(f, \boldsymbol{x}, \check{x})$. *Thus, for this example we have*

$$\lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = \partial f(0) \subset \mathbf{SG}f(0) = \lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{SS}^\sharp(f, \boldsymbol{x}, \check{x}).$$

EXAMPLE 6.3. *Consider the function* $f : \mathbb{R} \to \mathbb{R}$ *defined by*

$$f(x) = \begin{cases} x^2 + x + 1, & x > 0, \\ -(x+1)^2 + 3, & x \le 0, \end{cases}$$

and $\boldsymbol{x}$ *is any interval centered at* $\check{x} = 0$. *The function* f *is upper semicontinuous at* $\check{x}$ *and has a local infimum (which is actually global) at* $\check{x}$, *which is illustrated in Figure 5, and one verifies easily that*

$$\lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = [-\infty, -2] \subset \mathbf{SG}f(\check{x}) = [-\infty, 1].$$

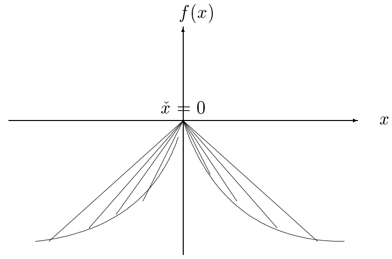

Figure 6. Graph of $f(x) = -\sqrt{|x|}$ with $\check{x} = 0$ in Example 5.4

Since f is neither convex nor Lipschitz at $\check{x}$, $\partial f(\check{x}) = \emptyset$. Observe that f is not regular at $\check{x}$, and the symmetric slope interval is

$$\lim_{w(\boldsymbol{x})\to 0} \square\mathbf{SS}^\sharp(f, \boldsymbol{x}, \check{x}) = [-\infty, 1].$$

EXAMPLE 6.4. *Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = -\sqrt{|x|}$, and let $\boldsymbol{x}$ be any interval centered at $\check{x} = 0$. Figure 6 shows that f has a local maximum (which is actually global) at $\check{x} = 0$, and one verifies easily that $0 \in \mathbb{R}^* = \mathbf{SG}f(\check{x}) = \lim_{w(\boldsymbol{x})\to 0} \square\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x})$. Since f is neither convex nor Lipschitz at $\check{x}$, $\partial f(\check{x}) = \emptyset$. Observe that f is regular at $\check{x}$ with $f^\circ(\check{x}; v) = \infty \quad \forall v \in \mathbb{R}$.*

In multiple dimensions, the next theorem extends Theorem 4.3 to arbitrary functions.

THEOREM 6.3. *Let $f : \boldsymbol{x} \subset \mathbb{R}^n \to \mathbb{R}$ be an arbitrary function that is regular at $\check{x}$. Then*

$$\lim_{w(\boldsymbol{x})\to 0} \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) \subseteq \mathbf{SG}f(\check{x}) \subseteq \lim_{w(\boldsymbol{x})\to 0} \square\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}). \tag{12}$$

Proof. Considering $w(\boldsymbol{x}) = t$, and vectors $v \in \mathbb{R}^n$ with $\|v\| = 1$, we have

$$\begin{aligned} \lim_{w(\boldsymbol{x})\to 0} \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) &= \lim_{t\downarrow 0} \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) \\ &= \lim_{t\downarrow 0} \left\{ s_v(\zeta) : \mathbf{S}_v(f, \zeta, \check{x}) = s_v(\zeta) \cdot v; 0 < \zeta \le t, \forall v \in \mathbb{R}^n \right\}. \end{aligned}$$

Since f is regular at $\check{x}$, we have

$$\lim_{t\downarrow 0} \mathbf{S}_v(f, t, \check{x}) = f^\circ(\check{x}; v).$$

Let $s \in \lim_{w(\boldsymbol{x})\to 0} \mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x})$. From Theorem 3.1, for some unitary vector $v \in \mathbb{R}^n$, we have

$$s \cdot v = \lim_{t\downarrow 0} \mathbf{S}_v(f, t, \check{x}) = f^\circ(\check{x}; v),$$

and it follows that $s \in \mathbf{SG}f(\check{x})$.

On the other hand, suppose that there exists $s \in \mathbb{R}^n$ such that $s \in \mathbf{SG}f(\check{x})$ and $s \notin \lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x})$. Then, for some unitary vector $v \in \mathbb{R}^n$, we get

$$s \cdot v > \lim_{t \downarrow 0} s_v(t) \cdot v = \lim_{t \downarrow 0} \mathbf{S}_v(f, t, \check{x}) = f^\circ(\check{x}; v),$$

which contradicts that $s \in \mathbf{SG}f(\check{x})$, and (12) is hold. $\square$

Example 4.3 illustrates this Theorem. In fact, f is Lipschitz near $\check{x}$, then

$$\mathbf{SG}f(\check{x}) = \partial f(\check{x}) = co\{(1,0),(0,-1),(-1,1)\},$$

which is the triangle with these three vertices, and

$$\mathbf{SG}f(\check{x}) \subset \lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x}) = [-1,1] \times [-1,1].$$

7. Containment-set of f

The containment-set or cset of a function was introduced in [35].

DEFINITION 7.1 ([35]). *Let $f : (\mathbb{R}^*)^n \to \mathbb{R}$ be an arbitrary function. The containment-set of f at a point $a = (a_1, \ldots, a_n) \in (\mathbb{R}^*)^n$ is the set of all possible limits of values of $f(x_i)$ where the vectors x_i converge to a as $i \to \infty$. The containment-set (or* cset*) is formally denoted* $\mathrm{cset}(f, a)$.

Arithmetic with the interval hull of csets is termed *cset-based interval arithmetic*, or *extended interval arithmetic* because it uses the extended real number system $\mathbb{R}^*$. The cset of $f' = \nabla f$ at $\check{x}$ is defined componentwise. The next two results relate the generalized gradient and semigradient to the cset of f'. Basic elements of this technique can be found in [21].

THEOREM 7.1. *For a locally Lipschitz function $f : (\mathbb{R}^*)^n \to \mathbb{R}$ near $\check{x}$ we have*

$$\mathrm{cset}(f', \check{x}) \subseteq \partial f(\check{x}) = co(\mathrm{cset}(f', \check{x})) \subseteq \Box\mathrm{cset}(f', \check{x}) = \Box\partial f(\check{x}),$$

where the third relation is an equality when $n = 1$.

Proof. It follows from Theorem 4.2 and Definition 6.1. All equalities hold when f is differentiable at $\check{x}$. When $n = 1$, $\Box\mathrm{cset}(f', \check{x})$, and $\partial f(\check{x})$

are equal to the interval constructed from the lower and upper bounds of the derivatives near to $\check{x}$, i.e.,

$$\partial f(\check{x}) = \Box\mathrm{cset}(f', \check{x}) = [\underline{f}'(\check{x}), \overline{f}'(\check{x})],$$

where

$$\overline{f}'(\check{x}) = \limsup_{x \to \check{x}} f'(x), \quad \text{and} \quad \underline{f}'(\check{x}) = \liminf_{x \to \check{x}} f'(x). \quad \Box$$

The next examples illustrate this theorem.

EXAMPLE 7.1. *Consider the function f in the Example 4.3, and $\check{x} = (0,0)$. We have*

$$\begin{aligned} \mathrm{cset}(\nabla f, \check{x}) &= \{(1,0),(0,-1),(-1,1)\} \\ &\subset \partial f(\check{x}) = co(\mathrm{cset}(\nabla f, \check{x})) = co\{(1,0),(0,-1),(-1,1)\} \\ &\subset \Box\mathrm{cset}(\nabla f, \check{x}) = [-1,1] \times [-1,1] = \Box\partial f(\check{x}) \\ &= \lim\nolimits_{w(\boldsymbol{x}) \to 0} \Box\mathbf{S}^{\sharp}(f, \boldsymbol{x}, \check{x}) \end{aligned}$$

EXAMPLE 7.2. *Consider the function f in the Example 4.2, $\check{x} = 0$, and $\boldsymbol{x}$ is any interval centered at $\check{x}$. We have*

$$\mathrm{cset}(f', \check{x}) = \{-1, 1\} \subset \partial f(\check{x}) = \Box\mathrm{cset}(f', \check{x}) = [-1, 1],$$

and

$$\lim_{w(\boldsymbol{x}) \to 0} \Box\mathbf{S}^{\sharp}(f, \boldsymbol{x}, \check{x}) = [-1/3, 1/3] \subset \Box\mathrm{cset}(f', \check{x}).$$

When f is an arbitrary function, the first inclusion of Theorem 7.1 can be generalized with the semigradient, and we have the following result.

THEOREM 7.2. *Let f an arbitrary function, and $\check{x}$ is in its domain, we have*

$$\Box\mathrm{cset}(f', \check{x}) \subseteq \mathbf{SG}f(\check{x}).$$

EXAMPLE 7.3. *Let $\boldsymbol{x}$ be any interval centered at $\check{x} = 0$. Consider the following Riemann function defined by*

$$f(x) = \begin{cases} 1, & \text{if } x \text{ is rational}, \\ 0, & \text{if } x \text{ is irrational}. \end{cases}$$

Since this function is not differentiable in any real value, then

$$\begin{aligned} \Box\mathrm{cset}(f', \check{x}) &= \emptyset \\ &\subset \mathbf{SG}f(\check{x}) = \lim\nolimits_{w(\boldsymbol{x}) \to 0} \Box\mathbf{SS}^{\sharp}(f, \boldsymbol{x}, \check{x}) = [-\infty, \infty]. \end{aligned}$$

EXAMPLE 7.4. *Consider the function in the Example 6.3, and $\boldsymbol{x}$ is any interval centered at $\check{x} = 0$. The derivative, $f'(x)$, is defined by*

$$f'(x) = \begin{cases} 2x + 1, & x > 0, \\ -2(x + 1), & x < 0, \end{cases}$$

then

$$\begin{aligned} \mathrm{cset}(f', \check{x}) &= \{-2, 1\} \\ &\subset \square\mathrm{cset}(f', \check{x}) = [-2, 1] \\ &\subset \mathbf{SG}f(\check{x}) = \lim_{w(\boldsymbol{x})\to 0} \square\mathbf{SS}^{\sharp}(f, \boldsymbol{x}, \check{x}) = [-\infty, 1]. \end{aligned}$$

EXAMPLE 7.5. *Consider the function in the Example 6.2, and $\boldsymbol{x}$ is any interval centered at $\check{x} = 0$. The derivative, $f'(x)$, is defined by*

$$f'(x) = \begin{cases} -1, & x > 0, \\ 1, & x < 0, \end{cases}$$

then

$$\begin{aligned} \mathrm{cset}(f', \check{x}) &= \{-1, 1\} \\ &\subset \square\mathrm{cset}(f', \check{x}) = [-1, 1] \\ &\subset \mathbf{SG}f(\check{x}) = \lim_{w(\boldsymbol{x})\to 0} \square\mathbf{SS}^{\sharp}(f, \boldsymbol{x}, \check{x}) = [-\infty, 1]. \end{aligned}$$

8. Slant differentiability

The definition of *slant differentiability* is introduced in [2]. This concept is an extension of Clarke's generalized derivative for locally Lipschitz functions in finite dimensional Euclidean spaces to infinite dimensional spaces. We present relationships between slant derivatives and slope sets for Lipschitzian functions in finite dimensional Euclidean spaces. Let X and Y be Banach spaces, and let D be an open domain in X. $L(X, Y)$ denotes the set of all bounded linear operators on X into Y.

DEFINITION 8.1 ([2]). *A function $f : D \subset X \to Y$ is said to be B-differentiable at the point x if it is one-sided directionally differentiable at x, and*

$$\lim_{h\to 0} \frac{f(x + h) - f(x) - f'(x; h)}{\|h\|} = 0.$$

In this case, we call $f'(x; .)$ the B-derivative of f at x.

DEFINITION 8.2 ([2]). *A function $f : D \subset X \to Y$ is said to be slantly differentiable at $x \in D$ if there exists a mapping $\hat{f}^{\circ} : D \to L(X, Y)$*

such that the family $\{\hat{f}^\circ(x+h) : h \in X, x+h \in D\}$ of bounded linear operators is uniformly bounded in the operator norm for h sufficiently small, and

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - \hat{f}^\circ(x+h)h}{\|h\|} = 0.$$

The function $\hat{f}^\circ$ is called a slanting function *for f at x.*

In finite dimensional Euclidean spaces, Shapiro [28] showed that a locally Lipschitz function f is B-differentiable at x if and only if it is directionally (Gateaux) differentiable at x.

DEFINITION 8.3 ([2]). *Suppose that $\hat{f}^\circ : D \to L(X,Y)$ is a slanting function for f at $x \in D$. We call the set*

$$\partial_S f(x) = \{\lim_{x_k \to x} \hat{f}^\circ(x_k)\}$$

the slant derivative *of f associated with $\hat{f}^\circ$ at $x \in D$. The limit is taken for any sequence $\{x_k\} \subset D$ such that $x_k \to x$ and the limit exists. (Note that $\hat{f}^\circ(x) \in \partial_S f(x)$, so $\partial_S f(x)$ is always nonempty.)*

When f is locally Lipschitz and regular at x, there exists a relationship between the generalized directional derivative $f^\circ(x;v)$ and the slant derivative $\hat{f}^\circ(x)$. Indeed, we have

$$f^\circ(x;v) = \lim_{t \downarrow 0} \hat{f}^\circ(x+tv)v.$$

EXAMPLE 8.1. *Let $X = Y = \mathbb{R}$ and $f(x) = \max(0,x)$. Let δ be a real number. Then the function*

$$\hat{f}^\circ(x) = \begin{cases} 1 & x > 0 \\ \delta & x = 0 \\ 0 & x < 0, \end{cases}$$

is a slanting function for f in X. The slant derivative of f for $x \in X$ is

$$\partial_S f(x) = \begin{cases} 1 & x > 0 \\ \{0, \delta, 1\} & x = 0 \\ 0 & x < 0. \end{cases}$$

In this example we get

$$\partial_S f(0) = \{0, \delta, 1\} \not\subseteq [0,1] = \partial f(0) = \mathbf{SG} f(0) = \lim_{w(\boldsymbol{x}) \to 0} [\![\mathbf{S}^\sharp(f, \boldsymbol{x}, 0),$$

for any $\boldsymbol{x} \subset X$, centered at $\check{x} = 0$, and $\delta \notin [0,1]$.

DEFINITION 8.4 ([2]). *We say that f is semismooth at x if there is a slanting function $\hat{f}^\circ$ for f in a neighborhood N_x of x, such that $\hat{f}^\circ$ and the associated derivative satisfy the following two conditions.*

(a) $\lim_{t\to 0^+} \hat{f}^\circ(x+th)h$ *exists for every* $h \in X$ *and*

$$\lim_{\|h\|\to 0} \frac{\lim_{t\to 0^+} \hat{f}^\circ(x+th)h - \hat{f}^\circ(x+h)h}{\|h\|} = 0.$$

(b)

$$\hat{f}^\circ(x+th)h - Vh = \circ(\|h\|) \qquad \textit{for all} \qquad V \in \partial_S f(x+h).$$

In finite dimensional spaces, Qi and Sun [19] showed that f is semismooth at x if and only if f is B-differentiable at x.

THEOREM 8.1 ([2]). *f is slantly differentiable at x if and only if f is Lipschitz continuous at x.*

Remark 1([2]) For a locally Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}^m$, if f is semismooth at x, then any single-valued selection of the Clarke Jacobian is a slanting function of f at x. Then, the slant derivative $\partial_S f(x)$ of f associated with $\hat{f}^\circ$ at x satisfies

$$\partial_S f(x) \subseteq \partial f(x).$$

COROLLARY 8.1 ([2] Mean Value Theorem). *Suppose that $f : D \subset X \to Y$ is slantly differentiable at x. Then for any $h \neq 0$ such that $x + h \in D$, there exists a slanting function $\hat{f}^\circ$ for f at x such that*

$$f(x+h) - f(x) = \hat{f}^\circ(x+h)h.$$

THEOREM 8.2. *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz continuous at $\check{x}$. Let $\boldsymbol{x}$ and $\check{\boldsymbol{x}}$ be interval boxes in D such that $\check{x} \in \check{\boldsymbol{x}} \subseteq \boldsymbol{x}$. Then there exists a slanting function $\hat{f}^\circ$ for f at $\check{x}$, such that the slant derivative $\partial_S f(x)$ of f associated with $\hat{f}^\circ$ at x satisfies*

$$\partial_S f(\check{x}) \subseteq \lim_{w(\boldsymbol{x})\to 0} []\mathbf{S}^\sharp(f, \boldsymbol{x}, \check{x})$$

Proof. Since f is Lipschitz continuous at $\check{x}$, from Theorem 6.5 there exists a slanting function $\hat{f}^\circ$ for f at $\check{x}$ and its associated slant derivative $\partial_S f(\check{x})$. Let $y \in \partial_S f(\check{x})$, then there is a sequence of points $\{x_k\} \subset D$ such that $x_k \to \check{x}$ and $\hat{f}^\circ(x_k) \to y$. There is a K such that $x_k \in \boldsymbol{x}$ for all $k \geq K$. From Corollary 6.6. we have

$$f(x_k) - f(\check{x}) = \hat{f}^\circ(x_k)(x_k - \check{x}).$$

Thus, $\hat{f}^{\circ}(x_k) \in \Box\mathbf{S}^{\sharp}(f, \boldsymbol{x}, \check{x})$, for all $k \geq K$. Since the slope interval is compact, we have

$$y \in \lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{S}^{\sharp}(f, \boldsymbol{x}, \check{x}). \quad \square$$

9. Conclusions

Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$ be an arbitrary function. The results presented above lead to the following inclusions. These inclusions are satisfied except on a set of measure zero. Since these concepts are generalizations of the gradient, all relations are equations when f is differentiable at $\check{x}$; in that case, all quantities are equal to the singleton set $\{f'(\check{x})\}$.

(a) From Theorems 4.1, 6.2 and 8.2 for $n = 1$, we can choose any particular value of $\lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{S}^{\sharp}(f, \boldsymbol{x}, \check{x})$ as slanting function of f at $\check{x}$, and we get

$$\partial_S f(\check{x}) \subseteq \lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{S}^{\sharp}(f, \boldsymbol{x}, \check{x}) \subseteq \partial f(\check{x}) \subseteq \lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{SS}^{\sharp}(f, \boldsymbol{x}, \check{x}) = \mathbf{SG}f(\check{x}).$$

Equality holds in the last two inclusions when f is locally Lipschitz and regular at $\check{x}$.

(b) From Theorems 4.3, 6.1, 6.3, and 8.2, for $n > 1$ and f regular at $\check{x}$, we get

$$\partial_S f(\check{x}) \subseteq \partial f(\check{x}) \subseteq \mathbf{SG}f(\check{x}) \subseteq \lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{S}^{\sharp}(f, \boldsymbol{x}, \check{x}) \subseteq \lim_{w(\boldsymbol{x})\to 0} \Box\mathbf{SS}^{\sharp}(f, \boldsymbol{x}, \check{x}).$$

The first inclusion is obtained if any-single value of $\partial f(\check{x})$ is chosen to define the slanting function associated to $\partial_S f(\check{x})$. Equality holds in the second and the fourth inclusions when f is locally Lipschitz near $\check{x}$.

(c) From Theorem 7.1, for f locally Lipschitz, we get

$$\mathrm{cset}(f', \check{x}) \subseteq \partial f(\check{x}) = co(\mathrm{cset}(f', \check{x})) \subseteq \Box\mathrm{cset}(f', \check{x}) = \Box\partial f(\check{x}),$$

where the third equality holds when $n = 1$.

(d) From Theorem 7.2, for $n \geq 1$, we get

$$\Box\mathrm{cset}(f', \check{x}) \subseteq \mathbf{SG}f(\check{x}).$$

Equality holds when f is differentiable at $\check{x}$.

(e) From Theorems 4.1 and 7.1, for $n = 1$ and f locally Lipschitz, we get

$$\lim_{w(\boldsymbol{x})\to 0} \square\mathbf{S}^{\sharp}(f, \boldsymbol{x}, \check{x}) \subseteq \square\mathrm{cset}(f', \check{x}).$$

Equality holds when f is regular at $\check{x}$.

Other important result in this work is Theorem 5.1, which shows that interval slopes yield a practical generalization of the *Karush-Kuhn-Tucker conditions* in constrained optimization.

Acknowledgements. The authors wish to thank both referees for their careful and laborious reports, which exhibited a clear understanding of the content and made suggestions to improve the exposition.

References

1. Aubin, J., P.: Lipschitz behavior of the solution to convex optimization problems, *Mat. Oper. Res.*, 9:87–111, 1984.
2. Chen, X., Nashed, Z., and Qi, L.: Smoothing Methods and Semismooth Methods for Nondifferentiable operator Equations, *SIAM, J. Numer. Anal.*, 38(4):1200–1216, 2000.
3. Clarke, Frank H.: *Optimization and Nonsmooth Analysis*. Classics in Applied Mathematics. SIAM, Philadelphia, 1990.
4. Dem'janov, V. F. and Rubinov, A. M.: On quasidifferentiable functionals. *Math. Dokl.*, 21:14–17, 1980.
5. Hansen, E. R.: In Global Optimization Using Interval Analysis: The Multidimensional Case. *Numer. Math.*, 34(3):247–270, 1980.
6. Hansen, E. R. and Sengupta, S.: Global Constrained Optimization Using Interval Analysis. In *Interval Mathematics 1980*, K. Nickel (ed.), Academic Press, New York, 1980.
7. Hansen, E. R.: *Global Optimization Using Interval Analysis*. Marcel Dekker, Inc., New York, 1992.
8. Janin R. and Gauvin J.: Lipschitz–Type Stability in Nonsmooth Convex Programs. *SIAM J. Control Optim.*, 38(1):124–137, 1999.
9. Kearfott, R. B.: Interval Extensions of Non-Smooth Functions for Global Optimization and Nonlinear Systems Solvers, *Computing*, 57:149–162, 1996.
10. Kearfott, R. B.: Treating Non-Smooth Functions as Smooth Functions in Global Optimization and Nonlinear Systems Solvers, *Scientific Computing and Validated Numerics*, G. Alefeld and A. Frommer and B. Lang, Mathematical Research, 90:160–172, 1996.
11. Kearfott, R. B.: *Rigorous Global Search: Continuous Problems*, Kluwer Academic Publishers, Netherlands, 1996.
12. Kiwiel, K. C.: *Methods of Descent in Nonsmooth Optimization*, Springer Verlag, Berlin, 1985.
13. Krawczyk, R.: Intervallsteigungen für rationale Funktionen und zugeordnete zentrishe Formen. *Freiburger Intervall-Berichte*, 83(2):1–30, 1983.
14. Krawczyk, R. and Neumaier, A.: Interval Slopes for Rational Functions and Associated Centered Forms. *SIAM J. Numer. Anal.*, 22(3):604–616, 1985.

15. Moore, R. E.: Interval Arithmetic and Automatic Error Analysis in Digital Computing: PhD dissertation. *UMI*, Stanford University, 1962.
16. Moore, R. E.: *Methods and Applications of Interval Analysis.* SIAM, Philadelphia, 1979.
17. Moreau Luc and Aeyels Dirk.: Optimization of Discontinuous Functions: A Generalized Theory of Differentiation, *SIAM Journal of Optimization*, 11(1):53–69, 2000.
18. Neumaier, A.: *Interval Methods for Systems of Equations.* Cambridge University Press, Cambridge, England, 1990.
19. Qi, L. and Sun, J.: A nonsmooth version of Newton's method. *Math. Programming*, 58:353–367, 1993.
20. Ratschek, H.: Some Recent Aspects of Interval Algorithms for Global Optimizations. *Reliability in Computing 'The Role of Interval Methods in Scientific Computing'*, Academic Press, 325–339, 1988.
21. Ratschek, H. and Görges, C.: Global Interval Methods for Local Nonsmooth Optimization. *Journal of Global Optimization*, 14:157–179, 1999.
22. Ratschek, H. and Rokne, J.: Interval Tools for Global Optimization. *Computers Math. Applic.*, 1:1–11, 1990.
23. Ratschek, H. and Rokne, J.: *Computer Methods for the Range of Functions.* Horwood,Chichester, England, 1984.
24. Ratz, D.: *Automatische Ergebnisverifikation bei globalen Optimierungsproblemen. Ph. D. Dissertation*,Universität Karlsruhe, 1992.
25. Ratz, D.: *Automatic Slope Computation and its Application in Nonsmooth Global Optimization.* Shaker Verlag, Aachen, 1998.
26. Rockafellar, R. T.: Directionally Lipschitzian functions and subdifferential calculus, *Proc. London Math. Soc.*, 39:331–355, 1979.
27. Rockafellar, R. T.: Equivalent Subgradient Versions of Hamiltonian and Euler–Lagrange equations in Variational Analysis. *SIAM J. Control and Optimization*, 34(4):1301–131, 1996.
28. Shapiro, A.: On concepts of directional differentiability, *J. Optim. Theory Appl.*, 66:477–487, 1990.
29. Shen Z., Neumaier, A. and Eiermann, M. C.: Solving Minimax Problems by Interval Methods, *BIT*, 30:742–751, 1990.
30. Sheng Z., Wolfe M. A.: An Interval Algorithm for Nondifferentiable Global Optimization, *Applied Mathematics and Computation*, 63:101–122, 1994.
31. Shor, N. Z.: *Minimization Methods for Nondifferentiable Functions. Springer-Verlag.* Berlin, 1985.
32. Simons S.: The least slope of a convex function and the maximal monotonicity of its subdifferential. *J. of Optimation Theory and Applications*, 71(1):127 – 126,1991.
33. Sonnenschein, C.: Zur Minimierung nichtdifferenzierbarer programmierbarer Funktionen. *Ph. D. Dissertation (in German)*, 1999.
34. Treiman, J. S.: Lagrange Multipliers for Nonconvex Generalized Gradients with Equality, Inequality, and Set Constraints. *SIAM J. Control Optim.*, 37(5):1313–1329, 1999.
35. Walster G. W., Hansen E. R., Pryce J. D.: Practical, Exception-free Interval Arithmetic on the Extended Reals. *Submitted to the SIAM Journal of Scientific Computing*, 2001.

Address for Offprints:
Dr. Humberto Muñoz
Department of Mathematics
Southern University and A& M College
P.O. Box 9440, Baton Rouge, LA 70813
Fax: + 225 7714762

Dr. R. Baker Kearfott
Department of Mathematics
University of Louisiana at Lafayette
Box 4-1010, Lafayette
LA 70504-1010