

SUBMONOTONE SUBDIFFERENTIALS OF LIPSCHITZ FUNCTIONS¹

BY

JONATHAN E. SPINGARN

ABSTRACT. The class of “lower- C^1 ” functions, that is functions which arise by taking the maximum of a compactly indexed family of C^1 functions, is characterized in terms of properties of the generalized subdifferential. A locally Lipschitz function is shown to be lower- C^1 if and only if its subdifferential is “strictly submonotone”. Other properties of functions with “submonotone” subdifferentials are investigated.

0. Introduction. One of the nice features of convex optimization is the link with monotone mappings. A *monotone mapping* (Minty [5]) is a multivalued function $T: R^n \rightrightarrows R^n$ (one which maps points in R^n into subsets of R^n) having the property that $\langle x_1 - x_2, y_1 - y_2 \rangle \geq 0$ whenever $y_i \in T(x_i)$, $i = 1, 2$, where $\langle \cdot, \cdot \rangle$ denotes the usual inner product on R^n .

If $f: R^n \rightarrow R$ is convex, then ∂f , the subdifferential of f , is the monotone mapping defined by

$$\partial f(x) = \{ \zeta \in R^n : f(z) \geq f(x) + \langle \zeta, z - x \rangle \forall z \in R^n \}.$$

Similarly, saddle functions give rise in a natural way to monotone mappings. Due to this, the problems of minimizing convex functions or of finding saddle points of convex-concave functions can be rephrased as “variational problems” (i.e., finding a point x such that $0 \in T(x)$ for a monotone mapping T), often resulting in considerable simplification. For example, Rockafellar [8], [9] has exploited the link between monotone mappings and saddle functions to unify and simplify the existing theory of multiplier methods in convex programming.

If $f: R^n \rightarrow R$ is a locally Lipschitz function, then following Clarke [1] we take the *generalized subdifferential* of f to be the multivalued mapping $\partial f: R^n \rightrightarrows R^n$ defined by taking $\partial f(x)$ to be the convex hull of the set of all limit points of sequences of the form $(\nabla f(x_n))$, where $x_n \rightarrow x$ and f is differentiable at x_n for all n . The multifunction ∂f is locally bounded, upper semicontinuous, nonempty-, compact-, and convex-valued [1], and satisfies a mean-value property [3].

It is the aim of this paper to show that a concept closely related to monotonicity, e.g. “submonotonicity”, plays a natural role in the analysis of nondifferentiable,

Received by the editors November 14, 1979.

1980 *Mathematics Subject Classification*. Primary 26B25, 47H05; Secondary 26B05, 49A29.

Key words and phrases. Submonotone mapping, generalized gradient, lower- C^1 function, nondifferentiable optimization.

¹Presented at the Tenth International Symposium on Mathematical Programming in Montreal, August 27–31, 1979 sponsored by the Mathematical Programming Society and IFORS. This research was supported by the Air Force Office of Scientific Research under grant number F49620-79-C-0120.

nonconvex problems. We will do this by demonstrating how properties of locally Lipschitz functions can be related to monotone-type properties of their generalized subdifferentials.

Our most important result appears in §III, where a complete characterization is obtained, in terms of properties of the generalized subdifferential, for the class of "lower- C^1 " functions. A function $f: R^n \rightarrow R$ is *lower- C^1* provided that for each $\bar{x} \in R^n$, there are a compact set S , a neighborhood U of $\bar{x}$, and a function $g: U \times S \rightarrow R$ such that g and $\nabla_x g$ are continuous jointly in x and s , and such that $f(x) = \max_{s \in S} g(x, s)$ for all $x \in U$. It will be shown that the lower- C^1 functions are precisely those locally Lipschitz functions whose generalized subdifferentials are "strictly submonotone".

In §II, some implications of the submonotonicity property are developed, and several equivalent characterizations are given. This concept is then contrasted with properties that have been discussed by other authors. Among these are regularity in the sense of Clarke [2], quasi-differentiability in the sense of Pshenichnyi [6], lower semidifferentiability in the sense of Rockafellar [10], and semismoothness in the sense of Mifflin [4].

The closed unit ball in R^n is denoted by $B = \{x \in R^n: |x| \leq 1\}$. If $K \subset R^n$ is a compact convex set, then Ψ_K^* is the *support function* of K , defined by $\Psi_K^*(u) = \sup\{\langle u, x \rangle: x \in K\}$. For any $u \in R^n$, we let $K_u = \{x \in K: \langle u, x \rangle = \Psi_K^*(u)\}$. Thus K_u is the set of all points in K for which u is an outer normal.

A set-valued mapping $T: R^n \rightrightarrows R^n$ is *closed* provided the set $\{(x, y): y \in T(x)\}$ is closed. T is *locally bounded* if for every $\bar{x} \in R^n$ there are $\varepsilon > 0$ and $R > 0$ such that $y \in T(x)$, $|x - \bar{x}| < \varepsilon$ implies $|y| < R$.

We will say the sequence (x_n) *converges to x in the direction $u \in R^n$* , written $x_n \rightarrow_u x$, provided $x_n \rightarrow x$ and either $u = 0$ or $u \neq 0$, $(x_n - x)/|x_n - x| \rightarrow u/|u|$, and $x_n \neq x$ for all sufficiently large n .

If $f: R^n \rightarrow R$, the *directional derivative* of f at x (when it exists) is the function

$$f'(x; u) = \lim_{t \downarrow 0} \frac{f(x + tu) - f(x)}{t}.$$

I. Submonotonicity. In this section, $T: R^n \rightrightarrows R^n$ denotes a convex-valued closed and locally bounded multifunction. T will be called *submonotone* at $x \in R^n$ provided

$$\liminf_{\substack{x' \rightarrow x, x' \neq x \\ y \in T(x), y' \in T(x')}} \frac{\langle y' - y, x' - x \rangle}{|x' - x|} > 0.$$

(T is trivially submonotone at x if $T(x) = \emptyset$.) T is *directionally upper semicontinuous* (d.u.s.c.) at x provided that for all $u \in R^n$, whenever $x_k \rightarrow_u x$ and $y_k \in T(x_k)$ for all k , then for every $\varepsilon > 0$ there exists k_0 such that $T(x_k) \subset T(x)_u + \varepsilon B \forall k > k_0$. For $u = 0$, this is automatically satisfied since T is assumed to be closed and locally bounded. Equivalently, T is d.u.s.c. at x if and only if for all $u \neq 0$, whenever $x_k \rightarrow_u x$ and $T(x_k) \ni y_k \rightarrow y$, then $y \in T(x)_u$. If T is submonotone [respectively, d.u.s.c.] at all $x \in R^n$, then T is *submonotone* [resp., d.u.s.c.].

(1.1) THEOREM. Let $T: R^n \rightrightarrows R^n$ be convex-valued, closed, and locally bounded (as is the case if $T = \partial f$ with f locally Lipschitz). Then T is d.u.s.c. at x if and only if T is submonotone at x .

PROOF. If T is not submonotone at x , there is $\varepsilon > 0$ and there are sequences $x_n \rightarrow x$, $x_n \neq x$, $y_n \in T(x_n)$, $y'_n \in T(x)$, such that $\langle x_n - x, y_n - y'_n \rangle / |x_n - x| \leq -\varepsilon < 0$, $\forall n$. We may clearly assume $x_n \rightarrow_u x$ for some $u \neq 0$, and since T is closed and locally bounded, that $y_n \rightarrow y \in T(x)$ and $y'_n \rightarrow y' \in T(x)$. Then $\Psi_{T(x)}^*(u) > \langle u, y' \rangle - \varepsilon \geq \langle u, y \rangle$, so T is not d.u.s.c.

Suppose that T is submonotone at x . Let $x_n \rightarrow_u x$, $u \neq 0$, $y_n \in T(x_n)$, $y_n \rightarrow y$. Since T is closed and locally bounded, $y \in T(x)$ and we will be done if we can show $y \in T(x)_u$. If $z \in T(x)$,

$$(y - z) \cdot u = \lim \frac{\langle y_n - z, x_n - x \rangle}{|x_n - x|} > 0$$

since T is submonotone at x . Since this holds for all $z \in T(x)$, $y \cdot u \geq \Psi_{T(x)}^*(u)$, showing that T is d.u.s.c. at x . $\square$

Of course if $f: R^n \rightarrow R$ is convex, ∂f is monotone, and hence submonotone. The fact that ∂f is directionally upper semicontinuous in this case is proved by Rockafellar [7, Theorem 24.6].

The multifunction $T: R^n \rightrightarrows R^n$ will be called *strictly submonotone* at x provided

$$\liminf_{\substack{x_1 \neq x_2 \\ x_i \rightarrow x, i=1,2 \\ y_i \in T(x_i), i=1,2}} \frac{\langle x_1 - x_2, y_1 - y_2 \rangle}{|x_1 - x_2|} > 0.$$

Strict submonotonicity clearly implies submonotonicity.

Next, we state a characterization of strict submonotonicity similar to the one provided in Theorem 1.1 for submonotonicity. The proof is similar, so it has been omitted.

(1.2) THEOREM. Let $T: R^n \rightrightarrows R^n$ be convex-valued, closed, and locally bounded. Then T is strictly submonotone at x if and only if whenever $x_n \rightarrow x$, $x'_n \rightarrow x$, $x_n \neq x'_n$, $y_n \in T(x_n)$, $y'_n \in T(x'_n)$, $y_n \rightarrow y$, $y'_n \rightarrow y'$, $x_n - x'_n \rightarrow_o 0$, one also has $v \cdot y' \leq v \cdot y$.

II. Lipschitzian functions. Next, we turn our attention to a particular class of multifunctions, namely those that are the generalized subdifferential for a locally Lipschitz function $f: R^n \rightarrow R$. Thus, if $T = \partial f$, we ask what the submonotonicity of ∂f implies about f .

If f is locally Lipschitz, the Clarke derivative of f is the function

$$f^\circ(x; u) = \limsup_{\substack{t \downarrow 0 \\ h \rightarrow 0}} \frac{f(x + h + tu) - f(x + h)}{t}.$$

$f^\circ(x; \cdot)$ is a continuous sublinear function which is the support function of $\partial f(x)$. For every $u, v \in R^n$, $f^\circ(x; \cdot)$, being a finite convex function, possesses a finite directional derivative at u in the direction v which we denote by $f^\circ(x; u; v)$. Alternatively, we could define $f^\circ(x; u; \cdot)$ to be the support function of $\partial f(x)_u$.

Clearly $f^\circ(x; 0; \cdot) = f^\circ(x; \cdot)$. Let us also define

$$f^\rightarrow(x; u; v) = \begin{cases} \limsup_{\substack{h \rightarrow_u 0 \\ t/|h| \downarrow 0}} \frac{f(x + h + tv) - f(x + h)}{t} & \text{if } u \neq 0, \\ f^\circ(x; v) & \text{if } u = 0. \end{cases}$$

Clearly $f^\rightarrow(x; u; v) \leq f^\circ(x; v)$. Also, $f^\rightarrow(x; u; \cdot)$ is sublinear, so $f^\rightarrow(x; u; \cdot)$ is the support function of some subset of $\partial f(x)$. As we shall see, that subset is $\partial f(x)_u$ exactly when ∂f is submonotone or, equivalently, d.u.s.c. To see that $f^\rightarrow(x; u; \cdot)$ is sublinear, note that it is positively homogeneous and

$$\begin{aligned} f^\rightarrow(x; u; v_1 + v_2) &= \limsup \frac{f(x + h + t(v_1 + v_2)) - f(x + h)}{t} \\ &\leq \limsup \frac{f(x + (h + tv_1) + tv_2) - f(x + (h + tv_1))}{t} \\ &\quad + \limsup \frac{f(x + h + tv_1) - f(x + h)}{t} \\ &= f^\rightarrow(x; u; v_2) + f^\rightarrow(x; u; v_1). \end{aligned}$$

(2.1) THEOREM. Let $f: R^n \rightarrow R$ be locally Lipschitz. ∂f is d.u.s.c. at x if and only if $f^\circ(x; u; v) = f^\rightarrow(x; u; v)$ for all $u, v \in R^n$.

PROOF. ($\Leftarrow$) Let $u \neq 0$, $x_k \rightarrow_u x$, $\partial f(x_k) \ni y_k \rightarrow y$. To show ∂f is d.u.s.c., it must be demonstrated that $y \in \partial f(x)_u$. Fix an arbitrary $v \in R^n$. Then

$$v \cdot y_k \leq f^\circ(x_k; v) = \limsup_{\substack{h \rightarrow 0 \\ t \downarrow 0}} \frac{f(x_k + h + tv) - f(x_k + h)}{t}$$

so $h_k, t_k > 0$ can be found with

$$v \cdot y_k - \frac{1}{k} \leq \frac{f(x_k + h_k + t_k v) - f(x_k + h_k)}{t_k}$$

and $\max\{t_k, |h_k|\} \leq |x_k - x|/k$. Hence,

$$\begin{aligned} v \cdot y &= \lim_k v \cdot y_k \leq \limsup_k \frac{f(x_k + h_k + t_k v) - f(x_k + h_k)}{t_k} \\ &\leq f^\rightarrow(x; u; v), \end{aligned}$$

where the last inequality follows from the fact that

$$x_k - x + h_k \rightarrow_u 0 \quad \text{and} \quad t_k/|x_k - x + h_k| \downarrow 0.$$

But $f^\rightarrow(x; u; v) = f^\circ(x; u; v)$ by assumption, so $v \cdot y \leq f^\circ(x; u; v) = \Psi_{\partial f(x)_u}^*(v)$ for all v , which implies that $y \in \partial f(x)_u$.

($\Rightarrow$) Fix $u \neq 0$, $v \in R^n$. First we show that $f^\circ(x; u; v) \geq f^\rightarrow(x; u; v)$. Pick sequences $h_n \rightarrow_u 0$, $t_n/|h_n| \downarrow 0$ such that

$$f^\rightarrow(x; u; v) = \lim_{n \rightarrow \infty} \frac{f(x + h_n + t_n v) - f(x + h_n)}{t_n}.$$

By the mean-value property (Lebourg [3]), there is, for each n ,

$$y_n \in \partial f(x + h_n + c_n t_n v)$$

with $0 < c_n < 1$ such that

$$v \cdot y_n = \frac{f(x + h_n + t_n v) - f(x + h_n)}{t_n}.$$

Without loss of generality, we can assume that $y_n \rightarrow y$ for some $y \in \partial f(x)$. Since ∂f is assumed to be d.u.s.c. at x , we have $y \in \partial f(x)_u$. Hence $f^\rightarrow(x; u; v) = \lim v \cdot y_n = v \cdot y \leq \Psi_{\partial f(x)_u}^*(v) = f^\circ(x; u; v)$, as desired.

To prove the opposite inequality, fix $u \neq 0$, $v \in R^n$, $w \in \partial f(x)_u$, and we will show $w \cdot v \leq f^\rightarrow(x; u; v)$. From this, the desired inequality follows by taking the supremum in w .

By d.u.s.c., we may find $\delta_n > 0$ ($n = 1, 2, \dots$) such that $0 < \delta < \delta_n$ implies

$$\partial f(x + \delta(u + v/n)) \subset \partial f(x)_{u+v/n} + B/n^2.$$

Clearly we may assume $\delta_n \rightarrow 0$. Let $x_n = x + \delta_n(u + v/n)$ and choose $y_n \in \partial f(x_n)$. Then $x_n \rightarrow_u x$ and $y_n \in \partial f(x)_{u+v/n} + B/n^2$. Since $y_n \in \partial f(x_n)$, we may find $t_n > 0$ and $h_n \in R^n$ such that

$$v \cdot y_n - \frac{1}{n} \leq \frac{f(x_n + h_n + t_n v) - f(x_n + h_n)}{t_n},$$

$$\max\{|h_n|, t_n\} < |x_n - x|/n.$$

Next, we will show that $\liminf_n y_n \cdot v \geq w \cdot v$. Since

$$x_n + h_n \rightarrow_u x \quad \text{and} \quad t_n/|x_n - x + h_n| \downarrow 0,$$

this will imply

$$w \cdot v \leq \liminf_n \frac{f(x_n + h_n + t_n v) - f(x_n + h_n)}{t_n} \leq f^\rightarrow(x; u; v)$$

which is the desired result.

For each n , choose $y'_n \in \partial f(x)_{u+v/n}$ such that $|y_n - y'_n| \leq n^{-2}$. Then

$$\begin{aligned} y_n \cdot (u + v/n) &= y'_n \cdot (u + v/n) + (y_n - y'_n) \cdot (u + v/n) \\ &\geq w \cdot (u + v/n) - n^{-2}|u + v/n| \end{aligned}$$

(because $w \in \partial f(x)$, $y'_n \in \partial f(x)_{u+v/n}$)

$$\geq y'_n \cdot u + \frac{1}{n} w \cdot v - \frac{1}{n^2} \left| u + \frac{1}{n} v \right|$$

(because $w \in \partial f(x)_u$, $y'_n \in \partial f(x)$)

$$\geq y_n \cdot u + \frac{1}{n} w \cdot v - \frac{1}{n^2} \left(|u| + \left| u + \frac{1}{n} v \right| \right)$$

(because $|y_n - y'_n| \leq n^{-2}$). So

$$y_n \cdot v \geq w \cdot v - \frac{1}{n} \left(|u| + \left| u + \frac{1}{n} v \right| \right)$$

and hence $\liminf y_n \cdot v \geq w \cdot v$, as desired. $\square$

Combining our results so far, we obtain the following:

(2.2) COROLLARY. *If $f: R^n \rightarrow R$ is locally Lipschitz, then the following are equivalent*

- (i) ∂f is submonotone at x ,
- (ii) ∂f is d.u.s.c. at x ,
- (iii) $f^\circ(x; \cdot; \cdot) = f^\circ(x; \cdot; \cdot)$.

Now that we have acquired a better understanding of the submonotonicity property of ∂f and what it implies about f , a logical question to ask next is: Just how strong is this property? In other words, if we take a look at "regularity" or "subdifferentiability" properties that have been studied for nondifferentiable functions by other authors, then which of these implies or is implied by the submonotonicity of ∂f ?

A locally Lipschitz function $f: R^n \rightarrow R$ is said to be *semismooth* at $x \in R^n$ (Mifflin [4]) provided that $x_k \rightarrow_u x$ and $y_k \in \partial f(x_k)$ imply that $\langle u, y_k \rangle \rightarrow f'(x; u)$.

(2.3) PROPOSITION. *If ∂f is submonotone at x then f is semismooth at x .*

PROOF. If $x_k \rightarrow_u x$ and $y_k \in \partial f(x_k)$ then every subsequence of (y_k) has a subsequence converging to some point in $\partial f(x)_u$ by directional upper semicontinuity. Hence $\langle u, y_k \rangle \rightarrow \Psi_{\partial f(x)}^*(u)$. By Proposition 2.5, $\Psi_{\partial f(x)}^*(u) = f'(x; u)$. $\square$

The function $f(x) = -|x|$ is semismooth, but ∂f is not submonotone at $x = 0$, so the converse of 2.3 is false.

Following Pshenichnyi [6], let us say that f is *quasi-differentiable* at x if there is a closed convex set K such that $f'(x; \cdot) = \Psi_K^*(\cdot)$. The function $f(x) = -|x|$ is not quasi-differentiable, so it is natural to ask whether every locally Lipschitz function which is both semismooth and quasi-differentiable has a submonotone subgradient mapping. The answer is negative. Consider the function $f: R^2 \rightarrow R$ defined as follows:

$$f(a, b) = \begin{cases} 0 & \text{if } a \leq 0, \\ a^2/4 & \text{if } a > 0, |b| \geq a^2/2, \\ |b| - b^2/a^2 & \text{if } a > 0, |b| < a^2/2. \end{cases}$$

Then f is differentiable at all points where either $b \neq 0$ or $a < 0$. At all points $x = (a, 0)$ with $a > 0$, f is quasi-differentiable since $f'(x; \cdot) = \Psi_K^*(\cdot)$ with $K = [(0, -1), (0, 1)]$. f is also locally Lipschitz, and it is not hard to check that f is everywhere semismooth. However, ∂f is not d.u.s.c. since $\partial f(0) = K$ but $(0, 0) \in \partial f(0, b)$ for all $b \neq 0$.

A locally Lipschitz function $f: R^n \rightarrow R$ will be called *regular* at x (Clarke [2]) provided that $f'(x; \cdot) = \Psi_{\partial f(x)}^*(\cdot)$. Clearly this is a stronger property than quasi-differentiability. The function f of the previous paragraph is not regular at 0, so it is natural to ask whether semismoothness plus regularity implies the submonotonicity of ∂f . This time the answer is affirmative:

(2.4) PROPOSITION. *∂f is submonotone at x if and only if f is semismooth and regular at x .*

PROOF. Suppose f is semismooth and regular at x . If $x_n \rightarrow_u x$ ($u \neq 0$), $y_n \in \partial f(x_n)$, and $y_n \rightarrow y$ then $y \in \partial f(x)$ and

$$\begin{aligned}\langle y, u \rangle &= \lim \langle y_n, u \rangle \\ &= f'(x; u) \quad (\text{by semismoothness}) \\ &= \Psi_{\partial f(x)}^*(u) \quad (\text{by regularity}),\end{aligned}$$

so $y \in \partial f(x)_u$. Hence ∂f is d.u.s.c., hence submonotone at x . The other direction follows by Propositions 2.3 and 2.5. $\square$

Rockafellar [10] has defined $z \in R^n$ to be a *lower semigradient* for f at x if

$$\liminf_{\substack{v \rightarrow u \\ t \downarrow 0}} \frac{f(x + tv) - f(x)}{t} > \langle u, z \rangle \quad \forall u \in R^n.$$

If such a z exists, f is *lower semidifferentiable*.

(2.5) PROPOSITION. Let $f: R^n \rightarrow R$ be locally Lipschitz, ∂f submonotone at x . Then

$$\lim_{\substack{t \downarrow 0 \\ v \rightarrow u}} \frac{f(x + tv) - f(x)}{t} = \Psi_{\partial f(x)}^*(u) \quad \forall u \in R^n.$$

In particular, f is lower semidifferentiable at x and $\partial f(x)$ is the set of lower semigradients. Also, f is regular at x .

PROOF. If $u = 0$, equality follows easily from the fact that f is locally Lipschitz, so suppose $u \neq 0$. Let $t_n \downarrow 0$, $v_n \rightarrow u$. For each n , there is $c_n \in (0, 1)$ and $y_n \in \partial f(x + c_n t_n v_n)$ such that $(f(x + t_n v_n) - f(x))/t_n = y_n \cdot v_n$. Since $x + c_n t_n v_n \rightarrow_u x$, we must have $y_n \cdot u \rightarrow \Psi_{\partial f(x)}^*(u)$. Thus

$$\lim_{n \rightarrow \infty} \frac{f(x + t_n v_n) - f(x)}{t_n} = \lim y_n \cdot v_n = \lim y_n \cdot u = \Psi_{\partial f(x)}^*(u).$$

Hence f is lower differentiable and $\partial f(x)$ is the set of lower semigradients. It is then obvious that f is regular at x . $\square$

The converse of 2.5 is false: $f(x) = x^2 \sin x^{-1}$ is locally Lipschitz and differentiable but ∂f is not submonotone at $x = 0$.

It is also possible for a function to be regular and lower semidifferentiable but for ∂f not to be submonotone. Consider, for example, any function $f: R \rightarrow R$ satisfying the following properties:

- (i) $f(x) = x - x^2$ for $x = 1/2, 1/3, 1/4, \dots$
- (ii) f' exists and is decreasing on $(1/(n+1), 1/n)$, $f'_+(1/(n+1)) = 1$ and $f'_-(1/n) = 0$, $n = 2, 3, 4, \dots$
- (iii) $f(x) = \frac{1}{4}$ for $x \geq \frac{1}{2}$ and $f(0) = 0$,
- (iv) $f(-x) = f(x)$ for all x .

Since $|x| - x^2 \leq f(x) \leq |x|$ for all x , $f'(0; u) = |u|$ for all u . Also, $\partial f(0) = [-1, 1]$ so f is regular at 0. But ∂f is clearly not submonotone at 0. Note that the behavior of f is nice at all points $x \neq 0$.

Since the property of strict submonotonicity is central to this paper, it is useful to mention an example of a function $f: R^2 \rightarrow R^2$ such that ∂f is submonotone

everywhere, but is not strictly submonotone. The function is

$$f(x, y) = \begin{cases} |y| & \text{if } x \leq 0, \\ |y| - x^2 & \text{if } x > 0, |y| > x^2, \\ (x^4 - y^2)/2x^2 & \text{if } x > 0, |y| \leq x^2. \end{cases}$$

It is easily checked that f is locally Lipschitz, that ∂f is everywhere submonotone, and $\partial f(0, 0) = [(0, -1), (0, 1)]$. If we let $x_n = (1/n, 1/n^2)$, $x'_n = (1/n, -1/n^2)$, $y_n = (2/n, -1)$, $y'_n = (2/n, 1)$, $n = 1, 2, \dots$, and $u = (1, 0)$, then $x_n \rightarrow_u 0$, $x'_n \rightarrow_u 0$, $y_n \in \partial f(x_n)$, and $y'_n \in \partial f(x'_n)$ for all n . However,

$$\frac{\langle x_n - x'_n, y_n - y'_n \rangle}{|x_n - x'_n|} = -2 \quad \text{for all } n$$

so ∂f is not strictly submonotone.

III. Lower- C^1 functions. In this section, we characterize the class of lower- C^1 functions in terms of their generalized subdifferentials. In Theorem 3.9, it is demonstrated that a locally Lipschitz f is lower- C^1 if and only if ∂f is strictly submonotone. The term "lower- C^1 function" was suggested to us by Professor R. T. Rockafellar.

(3.1) LEMMA. *Let $f: R^n \rightarrow R$ be locally Lipschitz, $x, y \in R^n$. For every $\varepsilon > 0$, there are neighborhoods U of x and V of y such that if $x' \in U$ and $y' \in V$, then $|\Psi_{\partial f(x')}(y) - \Psi_{\partial f(x')}(y')| < \varepsilon$.*

PROOF. Let κ be a Lipschitz constant for f on a neighborhood U of x . Then $\partial f(x') \subset \kappa B$ for all $x' \in U$, and it follows that κ is a (global) Lipschitz constant for $\Psi_{\partial f(x')}(\cdot)$. Take V to be the open ball of radius ε/κ centered at y . $\square$

(3.2) LEMMA. *Let $f: R^n \rightarrow R$ be locally Lipschitz. Then*

$$\liminf_{\substack{x' \rightarrow x \\ t \downarrow 0}} \frac{f(x' + ty) - f(x')}{t} - \Psi_{\partial f(x')}(y) > 0 \quad \forall y \in R^n \quad (3.3)$$

if and only if, for any compact $K \subset R^n$ and any $\varepsilon > 0$, there is a neighborhood U of x and $\lambda > 0$ such that

$$\frac{f(x' + ty') - f(x')}{t} - \Psi_{\partial f(x')}(y') > -\varepsilon \quad (3.4)$$

whenever $x' \in U$, $y' \in K$, $0 < t < \lambda$.

PROOF. Assume (3.3) holds, and fix $K \subset R^n$ and $\varepsilon > 0$. Since f is locally Lipschitz, (3.3) implies

$$\liminf_{\substack{x' \rightarrow x \\ y' \rightarrow y \\ t \downarrow 0}} \frac{f(x' + ty') - f(x')}{t} - \Psi_{\partial f(x')}(y) > 0 \quad \forall y \in R^n.$$

This and Lemma 3.1 imply that for each $y \in K$ we may find neighborhoods U_y of x , V_y of y , and $\lambda_y > 0$ such that $\Psi_{\partial f(x)}^*(y) - \Psi_{\partial f(x)}^*(y') \geq -\varepsilon/2$ and

$$\frac{f(x' + ty') - f(x')}{t} - \Psi_{\partial f(x)}^*(y) \geq -\varepsilon/2$$

whenever $x' \in U_y$, $y' \in V_y$, and $0 < t < \lambda_y$. Pick a finite subcover $V_{y_1}, \dots, V_{y_m}$ for K , and let $U = U_{y_1} \cap \dots \cap U_{y_m}$ and $\lambda = \min\{\lambda_{y_1}, \dots, \lambda_{y_m}\}$. For any $x' \in U$, $y' \in K$, and $t \in (0, \lambda)$, let i be such that $y' \in V_{y_i}$, and we get

$$\begin{aligned} \frac{f(x' + ty') - f(x')}{t} - \Psi_{\partial f(x)}^*(y') &= \left(\frac{f(x' + ty') - f(x')}{t} - \Psi_{\partial f(x)}^*(y_i) \right) + (\Psi_{\partial f(x)}^*(y_i) - \Psi_{\partial f(x)}^*(y')) \\ &\geq -\varepsilon/2 - \varepsilon/2 = -\varepsilon, \end{aligned}$$

as desired. The opposite direction of the lemma is obvious. $\square$

(3.5) PROPOSITION. *If $f: R^n \rightarrow R$ is locally Lipschitz, then ∂f is strictly submonotone at x if and only if (3.3) holds.*

PROOF. ($\Rightarrow$) If $y = 0$, the assertion is trivial. Without any loss of generality, we may assume that $|y| = 1$. Fix $\varepsilon > 0$. Since ∂f is strictly submonotone at x , there is $r > 0$ such that

$$\frac{\langle x_1 - x_2, y_1 - y_2 \rangle}{|x_1 - x_2|} \geq -\varepsilon$$

whenever $|x_i - x| < 2r$, $y_i \in \partial f(x_i)$ for $i = 1, 2$, and $x_1 \neq x_2$. Let x' and t be chosen so that $|x' - x| < r$ and $0 < t < r$. We will complete the proof by showing that

$$\frac{f(x' + ty) - f(x')}{t} - \Psi_{\partial f(x)}^*(y) \geq -\varepsilon.$$

Choose any $y_1 \in \partial f(x')$. By the mean-value theorem of Lebourg [3], we may find $s \in (0, t)$ and $y_2 \in \partial f(x' + sy)$ such that $f(x' + ty) - f(x') = t\langle y, y_2 \rangle$. Letting $x_1 = x'$ and $x_2 = x' + sy$, we have

$$\begin{aligned} \frac{f(x' + ty) - f(x')}{t} - \Psi_{\partial f(x)}^*(y) &= \langle y, y_2 - y_1 \rangle \\ &= \frac{\langle x_2 - x_1, y_2 - y_1 \rangle}{|x_2 - x_1|} \geq -\varepsilon. \end{aligned}$$

($\Leftarrow$) Next, suppose (3.3) holds, and let $\varepsilon > 0$ be given. By Lemma 3.2, there is a neighborhood U of x and $\lambda > 0$ such that

$$\frac{f(x' + tu) - f(x')}{t} - \Psi_{\partial f(x)}^*(u) \geq -\varepsilon/2$$

whenever $x' \in U$, $|u| \leq 1$, and $0 < t < \lambda$. We may also assume that U is small enough so that $|z - z'| < \lambda$ for all $z, z' \in U$. Fix $x_i \in U$, $y_i \in \partial f(x_i)$ for $i = 1, 2$,

with $x_1 \neq x_2$. Let $t = |x_2 - x_1|$ and $u = (x_2 - x_1)/t$. Then

$$\begin{aligned} \frac{\langle x_1 - x_2, y_1 - y_2 \rangle}{|x_1 - x_2|} &= -\langle u, y_1 \rangle - \langle -u, y_2 \rangle \\ &> -\Psi_{\partial f(x_1)}^*(u) - \Psi_{\partial f(x_2)}^*(-u) \\ &= \frac{f(x_1 + tu) - f(x_1)}{t} - \Psi_{\partial f(x_1)}^*(u) \\ &\quad + \frac{f(x_2 - tu) - f(x_2)}{t} - \Psi_{\partial f(x_2)}^*(-u) \\ &> -\varepsilon/2 - \varepsilon/2 = -\varepsilon, \end{aligned}$$

which shows that ∂f is strictly submonotone at x . $\square$

(3.6) LEMMA. Let $f: R^n \rightarrow R$ be locally Lipschitz, let C and K be compact sets in R^n , and suppose that ∂f is strictly submonotone on C . Then

$$\liminf_{\substack{x \in C \\ y \in K \\ t \downarrow 0}} \frac{f(x + ty) - f(x)}{t} - \Psi_{\partial f(x)}^*(y) > 0.$$

PROOF. Let $\varepsilon > 0$ be given. By Proposition 3.5 and Lemma 3.2, for each $x \in C$, there is $\lambda_x > 0$ such that $(f(x' + ty) - f(x'))/t - \Psi_{\partial f(x')}^*(y) > -\varepsilon$ whenever $|x' - x| < \lambda_x$, $y \in K$, and $0 < t < \lambda_x$. Let $x_1, \dots, x_r \in C$ be such that for every $x \in C$ we have $|x - x_i| < \lambda_{x_i}$ for some i . Let $\lambda = \min(\lambda_{x_1}, \dots, \lambda_{x_r})$. Then for any $x \in C$, $y \in K$, we have

$$\frac{f(x + ty) - f(x)}{t} - \Psi_{\partial f(x)}^*(y) > -\varepsilon$$

whenever $0 < t < \lambda$. $\square$

(3.7) LEMMA. Let $\phi(t)$ be real-valued, defined for $t > 0$ sufficiently small, such that $\lim_{t \rightarrow 0} \phi(t) = 0$. Then there is a continuously differentiable function $\alpha(t)$ defined on $[0, a]$ for some $a > 0$ such that $\alpha(0) = \alpha'(0) = 0$, $\alpha(t) > t\phi(t)$, $\forall t \in (0, a]$.

PROOF. Let $a > 0$ be such that ϕ is bounded above on $(0, 2a]$, and let $a_k = a/2^k$, $k = 0, 1, \dots$. If β is the infimum of all affine functions $l: R \rightarrow R$ which satisfy $l(a_k) > \phi(t)$ for all $t \in (0, 2a_k]$ and all $k = 0, 1, 2, \dots$ then the following properties are easily checked:

β is continuous, concave, nondecreasing on $[0, a]$,

$\beta(0) = 0$,

$\beta > \phi$ on $(0, a]$,

β is affine on $[a_{k+1}, a_k]$, $k = 0, 1, 2, \dots$.

Also, β'_+ , the right derivative of β , has these properties:

β'_+ is finite, nonnegative, and nonincreasing on $(0, a)$,

β'_+ is constant on $[a_{k+1}, a_k]$, $k = 0, 1, 2, \dots$,

β'_+ is integrable on $[0, a]$.

This last assertion is proven as follows. Whenever $0 < u < v < a$,

$$\beta(v) - \beta(u) = \int_u^v \beta'_+(s) \, ds$$

(cf. Rockafellar [7, 24.2.1]). Since $\beta'_+ \geq 0$ and β is continuous,

$$\int_0^a \beta'_+(s) \, ds = \lim_{\substack{u \rightarrow 0 \\ v \rightarrow a}} \int_u^v \beta'_+(s) \, ds = \beta(a) - \beta(0) < \infty,$$

so β is integrable. Note that since $\beta(0) = 0$, $\beta(t) = \int_0^t \beta'_+(s) \, ds$ for all $t \in [0, a]$.

For each $k = 1, 2, \dots$, pick c_k such that $\frac{1}{2}(a_k + a_{k+1}) < c_k < a_k$, $(a_k - c_k)(\beta'_+(a_{k+1}) - \beta'_+(a_k)) < a_{k+1}$. Define $\mu: (0, a) \rightarrow \mathbb{R}$ to be the function that agrees with $1 + \beta'_+$ on the intervals $[a_{k+1}, c_k]$ ($k = 1, 2, \dots$) and on $[a_1, a_0]$, and is affine on the intervals $[c_k, a_k]$ ($k = 1, 2, \dots$). Then μ is continuous, nonnegative, and nonincreasing on $(0, a)$ and

$$\int_{a_{k+1}}^t \mu(s) - \beta'_+(s) \, ds \geq 0 \quad \text{for all } k = 0, 1, 2, \dots, t \in [a_{k+1}, a_k].$$

Since $0 \leq \mu \leq \beta'_+ + 1$ and β'_+ is integrable, it follows that μ is integrable. Then for all $t \in [0, a]$, $\int_0^t \mu(s) \, ds \geq \int_0^t \beta'_+(s) \, ds = \beta(t)$.

Define $\alpha(t) = t \int_0^t \mu(s) \, ds$ for all $t \in [0, a]$. Clearly,

α is continuously differentiable on $(0, a]$.

$\alpha(0) = 0$.

$\alpha(t) \geq t\phi(t)$ for $t \in (0, a]$.

It remains only to show that α is continuously differentiable at 0. We have

$$\alpha'(0) = \lim_{t \rightarrow 0^+} \frac{\alpha(t)}{t} = \lim_{t \rightarrow 0} \int_0^t \mu(s) \, ds = 0.$$

Also, for $t > 0$,

$$\begin{aligned} \alpha'(t) &= \int_0^t \mu(s) \, ds + t\mu(t) = \int_0^t (\mu(s) + \mu(t)) \, ds \\ &< 2 \int_0^t \mu(s) \, ds \quad (\text{since } \mu \text{ is nonincreasing}) \end{aligned}$$

so $\lim_{t \rightarrow 0} \alpha'(t) = 0$. $\square$

(3.8) PROPOSITION. Let $f: \mathbb{R}^n \rightarrow \mathbb{R}$ be locally Lipschitz. If ∂f is strictly submonotone then for every compact $C \subset \mathbb{R}^n$, there is a continuously differentiable $\alpha: [0, a] \rightarrow \mathbb{R}$ such that $\alpha(0) = \alpha'(0) = 0$ and

$$f(x + ty) \geq f(x) + t\Psi_{\partial f(x)}^*(y) - \alpha(t)$$

whenever $x \in C$, $|y| = 1$, and $0 < t \leq a$.

PROOF. For $t > 0$, define

$$\phi(t) = - \inf_{\substack{t' \leq t \\ x \in C \\ |y|=1}} \min \left(\frac{f(x + t'y) - f(x)}{t'} - \Psi_{\partial f(x)}^*(y), 0 \right).$$

Then $\phi > 0$ and by Lemma 3.6, $\lim_{t \rightarrow 0} \phi(t) = 0$. By Lemma 3.7, there is a real-valued function $\alpha(t)$ which is continuously differentiable on $[0, a]$ for some $a > 0$ such that $\alpha(0) = \alpha'(0) = 0$ and $\alpha(t) \geq t\phi(t)$ for all $t \in (0, a]$. It follows that $f(x + ty) \geq f(x) + t\Psi_{\partial f(x)}^*(y) - \alpha(t)$ whenever $x \in C$, $|y| = 1$ and $0 < t \leq a$. $\square$

(3.9) THEOREM. *Let $f: R^n \rightarrow R$ be locally Lipschitz. f is lower- C^1 if and only if ∂f is strictly submonotone.*

PROOF. ($\Rightarrow$) Suppose ∂f is strictly submonotone, and fix $\bar{x} \in R^n$. By Proposition 3.8, there are $a > 0$, and a C^1 function $\alpha: [0, a] \rightarrow R$ such that $\alpha(0) = \alpha'(0) = 0$ and $f(x + y) \geq f(x) + \langle \zeta, y \rangle - \alpha(|y|)$ whenever $|x - \bar{x}| \leq 1$, $|y| \leq a$, and $\zeta \in \partial f(x)$. Let $b = \min\{1, a/2\}$. Then

$$f(x) \geq f(x') + \langle x - x', \zeta \rangle - \alpha(|x - x'|)$$

whenever $|x - \bar{x}| \leq b$, $|x' - \bar{x}| \leq b$, and $\zeta \in \partial f(x')$. Let $U = \{x: |x - \bar{x}| < b\}$ and $S = \{(x', \zeta): |x' - \bar{x}| \leq b, \zeta \in \partial f(x')\}$. If we define

$$g(x, x', \zeta) = f(x') + \langle x - x', \zeta \rangle - \alpha(|x - x'|),$$

then g has the properties required for the definition of "lower- C^1 function".

($\Rightarrow$) Fix $\bar{x} \in R^n$, let U , S , and g be as indicated, and let $K \subset U$ be a compact convex neighborhood of $\bar{x}$. By compactness, $\nabla_x g(x, s)$ is uniformly continuous on $K \times S$. So, defining for $t > 0$

$$\eta(t) = \sup_{\substack{z, z' \in K \\ s \in S \\ |z - z'| \leq t}} |\nabla_x g(z, s) - \nabla_x g(z', s)|$$

we have $\lim_{t \downarrow 0} \eta(t) = 0$. By Lemma 3.7 there is, for some $a > 0$, a C^1 function $\alpha: [0, a] \rightarrow R$ such that $\alpha(0) = \alpha'(0) = 0$ and $\alpha(t) \geq t\eta(t)$ for all $t \in (0, a]$.

Fix $x, x' \in K$ such that $x \neq x'$. For each $s \in S$, by the mean-value theorem, there is $x'' \in K$ on the line segment (x, x') such that $g(x', s) - g(x, s) = (x' - x) \cdot \nabla_x g(x'', s)$. Then

$$\begin{aligned} & [g(x', s) - g(x, s) - (x' - x) \cdot \nabla_x g(x, s)] / |x' - x| \\ &= (\nabla_x g(x'', s) - \nabla_x g(x, s))(x' - x) / |x' - x| \\ &\geq -\eta(|x'' - x|) \geq -\eta(|x' - x|) \geq -\frac{\alpha(|x' - x|)}{|x' - x|}. \end{aligned}$$

Hence, for all $s \in S$,

$$g(x', s) \geq g(x, s) + (x' - x) \cdot \nabla_x g(x, s) - \alpha(|x' - x|).$$

Let $\zeta \in \partial f(x)$ be arbitrary. By Clarke [1, Theorem 2.1], we may find $s_1, \dots, s_k \in S$ and numbers $\lambda_1, \dots, \lambda_k$ such that

$$\zeta = \sum \lambda_i \nabla_x g(x, s_i), \quad \lambda_i \geq 0, \quad \sum \lambda_i = 1, \quad g(x, s_i) = f(x).$$

Then

$$\begin{aligned} f(x') &\geq \sum \lambda_i g(x', s_i) \\ &\geq \sum \lambda_i (g(x, s_i) + (x' - x) \cdot \nabla_x g(x, s_i) - \alpha(|x' - x|)) \\ &= f(x) + (x' - x) \cdot \zeta - \alpha(|x' - x|). \end{aligned}$$

Since this holds for all $\zeta \in f(x)$, we have shown that for all $x, x' \in K$ with $x \neq x'$, we have

$$f(x') \geq f(x) + \Psi_{\partial f(x)}^*(x' - x) - \alpha(|x' - x|).$$

It then follows easily that 3.3 holds at $x = \bar{x}$ and hence by Proposition 3.5 that ∂f is strictly submonotone at $\bar{x}$. $\square$

ACKNOWLEDGEMENT. We wish to thank Professor R. T. Rockafellar for sharing many valuable insights with us.

REFERENCES

1. F. H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.
2. ———, *Generalized gradients of Lipschitz functionals*, Adv. in Math. (to appear).
3. G. Lebourg, *Valeur moyenne pour gradient généralisé*, C. R. Acad. Sci. Paris Ser. A **281** (1975), 795–797.
4. R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optimization **15** (1977), 6.
5. G. J. Minty, *Monotone (nonlinear) operators in Hilbert space*, Duke Math. J. **29** (1962), 341–346.
6. B. N. Pshenichnyi, *Necessary conditions for an extremum*, Marcel Dekker, New York, 1971.
7. R. T. Rockafellar, *Convex analysis*, Princeton Univ. Press, Princeton, N. J., 1972.
8. ———, *The multiplier method of Hestenes and Powell applied to convex programming*, J. Optimization Theory Appl. **12** (1973), 6.
9. ———, *Augmented Lagrangians and applications of the proximal point algorithm in convex programming*, Math. Oper. Res. **1** (1976).
10. ———, *The theory of subgradients and its applications to problems of optimization*, Lecture Notes, Univ. of Montreal, Feb.-March, 1978.

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA, GEORGIA 30332