TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 276, Number 1, March 1983

A GENERAL SUFFICIENCY THEOREM FOR NONSMOOTH NONLINEAR PROGRAMMING

BY

R. W. CHANEY[1]

ABSTRACT. Second-order conditions are given which are sufficient to guarantee that a given point be a local minimizer for a real-valued locally Lipschitzian function over a closed set in n-dimensional real Euclidean space. These conditions are expressed in terms of the generalized gradients of Clarke. The conditions provide a very general and unified framework into which many previous first- and second-order theorems fit.

1. Introduction. Let S be a closed subset of n-dimensional space R^n and let W be an open set in R^n. Suppose that the point x^* belongs to $S \cap W$ and let f be a real-valued locally Lipschitzian function on W. We consider here the problem

$$\text{P:} \qquad \text{Minimize } f(x), \text{ over } x \text{ in } S \cap W.$$

We shall present a general theorem which gives second-order conditions that are sufficient to guarantee that x^* be a local solution to problem P. This theorem both unifies and extends separate earlier results presented for this problem by the author [**2**]; specific remarks on this point are made below. The theorems in [**2**] have already been shown to yield both the classical theorem for problems with a finite number of functional constraints and Ioffe's theorem in finite dimensions; it can be shown that they yield other results, such as the sufficiency theorem of Fletcher and Watson [**8**] for certain nondifferentiable problems containing norms and a theorem proved by the author [**1**] for a certain class of nonsmooth problems.

The second-order sufficiency conditions given here are formulated in terms of the generalized gradients of Clarke. The generalized gradients of f at x are the members of the Clarke subdifferential $\partial f(x)$ of f at x. The Clarke subdifferential reduces to the subdifferential of convex analysis when f is convex and it contains only the gradient vector $\nabla f(x)$ when f is C^1 near x. We shall give below a brief summary of those basic properties of generalized gradients which we shall use. For more information about generalized gradients, one should examine such earlier papers of Clarke as [**4, 5, 6**], the work by Rockafellar [**23, 20, 21**], or the forthcoming book by

Received by the editors January 8, 1982. Presented to the Eleventh International Symposium on Mathematical Programming held in Bonn, West Germany, on August 23–27, 1982.

1980 *Mathematics Subject Classification*. Primary 49B34; Secondary 49B21.

Key words and phrases. Second-order sufficiency conditions, nonsmooth constrained optimization.

[1]The author wishes to thank the University of Florida for its generous support and hospitality during the period in which this research was begun.

©1983 American Mathematical Society
0002-9947/82/0000-1123/$03.75

Clarke [7]. In this paper, we shall—with one brief exception—consider only locally Lipschitzian functions. Of course, generalized gradients have been defined for more general classes of functions (see, e.g., [**5** and (especially) **21**]).

The classical method of proof for second-order sufficiency theorems uses second-order Taylor expansions. The method used here (and used first in [**2** and **3**]) is quite different. It proceeds according to the following outline: First, we deny the desired conclusion and from this initial step derive an infinite sequence of modified problems which do have minimizers approaching x^*. We then apply a first-order necessity theorem to each of these modified problems. We obtain infinite sequences of points and related "necessary conditions". From these sequences, a contradiction is derived "in the limit". There are various first-order theorems which are suitable for use in this proof. There are results proved by Clarke [**4**], Hiriart-Urruty [**11**], and Rockafellar [**20**, **22**]. Although it is the earliest of these results, the theorem of Clarke [**4**, Theorem 1] will serve us well in this article.

We obtain some particular results here for the case in which the set S is tangentially regular at x^*. This occurs when the contingent cone $K(S, x^*)$ and the Clarke tangent cone $T(S, x^*)$ coincide. The Clarke tangent cone $T(S, x^*)$ of S at x^* consists of all y in R^n such that, whenever one has sequences $\{t_k\}$ decreasing to 0 and $\{x_k\}$ converging to x^* with each x_k in S, then there exists a sequence $\{y_k\}$ convergent to y such that $x_k + t_k y_k$ belongs to S for all k. The contingent cone $K(S, x^*)$ of S at x^* consists of all y in R^n such that there exist sequences $\{t_k\}$ of positive numbers and $\{x_k\}$ convergent to x^* for which each x_k belongs to S and $\{(x_k - x^*)/t_k\}$ converges to y. (Hestenes [**9**, **10**] and others use the term "tangent cone" for what we call the "contingent cone".) The two cones $K(S, x^*)$ and $T(S, x^*)$ are both always closed but only $T(S, x^*)$ is necessarily convex (see [**18**, Theorem 1]). Moreover, we always have $T(S, x^*) \subseteq K(S, x^*)$ (see [**23**, p. 17]).

Before we proceed further, we need to set some additional notation. If x and y belong to R^n and if $\delta > 0$, then $|x|$ denotes the Euclidean norm of x, $x \cdot y$ the usual inner product of x and y, and $B(x, \delta)$ the set $\{z \in R^n: |z - x| \leqslant \delta\}$. If C is a closed convex set in R^n and if x belongs to C, we denote by $N(C, x)$ the normal cone to C at x [**19**, p. 15]. If C is merely a closed set, then the normal cone to C at x is given by $N(C, x) = N(T(C, x), 0)$.

2. The main theorem and its corollaries.

2.1 Remark. We list here several of the basic facts (taken from [**5**]) about generalized gradients which we shall use. We assume that f is locally Lipschitzian on an open set W in R^n and that x belongs to W.

(a) According to the theorem of Rademacher, f is differentiable a.e. on W. We denote by $\nabla f(x)$ the gradient of f at x (when it exists). Let E be the set of all points z in W for which f is differentiable at z. The (Clarke) subdifferential of f at x, denoted by $\partial f(x)$, is the convex hull of the set of all limits of convergent sequences $\{\nabla f(x_k)\}$ where $\{x_k\}_{k>1}$ is a sequence in E convergent to x.

Each member of the subdifferential is termed a generalized gradient of f at x. If K is a Lipschitzian constant for f on a neighborhood of x, then $|v| \leqslant K$ for every v in $\partial f(x)$.

(b) The generalized directional derivative of f at x in the direction d is defined by

$$f^0(x; d) = \limsup_{v \to 0;\, t \downarrow 0} \frac{f(x + v + td) - f(x + v)}{t}.$$

The function $f^0(x; \cdot)$ is convex and it is the support function of the convex (and compact) set $\partial f(x)$; i.e., for all x in W and d in R^n, we have

$$f^0(x; d) = \max\{v \cdot d : v \in \partial f(x)\}.$$

(c) The multifunction $x \to \partial f(x)$ is upper semicontinuous on W; thus, if $\{x_k\}$ and $\{v_k\}$ converge, respectively, to x in W and v in R^n and if v_k is in $\partial f(x_k)$ for each k, then v belongs to $\partial f(x)$.

2.2 Definition. Let $\{x_k\}$ be a sequence in R^n which converges to x and let d be a unit vector in R^n. Then $\{x_k\}$ converges to x in direction d in case the sequence $\{(x_k - x)/|x_k - x|\}$ converges to d.

2.3 Definition. Let x be in W and let d be a unit vector in R^n. We define $\partial_d f(x)$ to be the set of all v in R^n for each of which there exist sequences $\{x_k\}$ in W and $\{v_k\}$ in R^n such that:

(a) $\{x_k\}$ converges to x in direction d;

(b) $\{v_k\}$ converges to v;

(c) v_k belongs to $\partial f(x_k)$ for each k.

(Observe that, in view of 2.1(c), we have $\partial_d f(x) \subseteq \partial f(x)$. One may think of $\partial_d f(x)$ as the set of those generalized gradients of f at x which "arise" from the direction d.)

2.4 Remarks. In the main theorem, which follows, an auxiliary function g is introduced. One is free to choose g and one may choose it to be f itself or to be a certain Lagrangian associated somehow with P. Or, one may simply choose g so as to have a function whose subdifferential is simpler than that of f; a choice on these grounds is made in the recovery of Ioffe's sufficiency theorem [**12**, Theorem 2] from [**2**, Corollary 2.18] (see [**2**, Theorem 3.4]).

Next, we define $L(f, x^*)$ to be the set of all points td, where $d \in R^n$, $|d| = 1$, $t \geqslant 0$, and $v_0 \cdot d \leqslant 0$ for some v_0 in $\partial_d f(x^*)$. The set $L(f, x^*)$ is a closed cone.

2.5 Theorem. *Let S, W, x^* and f be as in the Introduction. Let g be a real-valued locally Lipschitzian function on W for which $g(x^*) = f(x^*)$ and $g(x) \leqslant f(x)$ for every x in $S \cap W$. Suppose that, to each unit vector d^* in the set $K(S, x^*) \cap L(f, x^*)$, there corresponds a closed convex cone $C(d^*)$ for which $d^* \in C(d^*)$. Finally we suppose also that*:

(a) *we have $w \cdot d \geqslant 0$ whenever d is any unit vector in $C(d^*)$ for some d^* in $K(S, x^*) \cap L(f, x^*)$ and w is any generalized gradient in $\partial_d g(x^*)$*;

(b) *there exists $m^* \geqslant 0$ such that*

$$\limsup w_k \cdot (x_k - x^*)/|x_k - x^*|^2 > m^*$$

whenever $\{x_k\}$ and $\{w_k\}$ are sequences and d^* and d are unit vectors for which

(i) *$\{x_k\}$ converges to x^* in direction d,*

(ii) *$d^* \in K(S, x^*) \cap L(f, x^*)$ and $d \in C(d^*)$,*

(iii) *$w_k \in \partial g(x_k)$ for each k,*

(iv) $\{w_k\}$ *converges to a point* w *in* $-N(C(d^*) + x^*, x^*)$.

Then there exists $\delta > 0$ *such that* $f(x) \geqslant f(x^*) + (m^*/2)|x - x^*|^2$ *whenever* x *belongs to* $B(x^*, \delta) \cap S$.

PROOF. Suppose that the desired conclusion is false and select a sequence $\{\delta_k\}$ of positive numbers decreasing to 0 with $\delta_1 < 1$. Given k, there exists z_k in $B(x^*, \delta_k) \cap S$ such that $f(z_k) - f(x^*) < (m^*/2)|z_k - x^*|^2$. We put

$$h(x) = g(x) - (m^*/2)|x - x^*|^2$$

and note that $z_k \neq z$ for each k. We have

$$\begin{aligned} h(z_k) &= g(z_k) - (m^*/2)|z_k - x^*|^2 \\ &\leqslant f(z_k) - (m^*/2)|z_k - x^*|^2 \leqslant f(x^*) = h(x^*). \end{aligned}$$

Setting $e_k = (z_k - x^*)/|z_k - x^*|$, we may assume that $\{e_k\}$ converges to a unit vector d^* in $K(S, x^*)$. By Lebourg's Mean Value Theorem [**13**, **14**], we have $f(z_k) - f(x^*) = v_k^* \cdot (z_k - x^*)$, where v_k^* belongs to $\partial f(\theta_k z_k + (1 - \theta_k)x^*)$ with $0 < \theta_k < 1$. By 2.1 and 2.3, we may assume that $\{v_k^*\}$ converges to v in $\partial_{d^*} f(x^*)$ and we clearly have $v \cdot d^* = \lim v_k^* \cdot e_k \leqslant 0$. Hence d^* belongs to $L(f, x^*)$.

For each k, we define a linear transformation T_k from R^n into itself by putting $T_k(x) = x + (x \cdot d^*)(e_k - d^*)$. With I as the identity transformation, we have $\|T_k - I\| \leqslant |e_k - d^*|$. We may assume $|e_k - d^*| < 0.5$ for all k and so it follows from the Perturbation Lemma [**17**, p. 45] that T_k is invertible for all k and that the sequence $\{T_k^{-1}\}$ is bounded. Set $A_k = T_k(C(d^*))$ for all k. Then A_k is a closed convex cone containing e_k. We know that z_k belongs to $B(x^*, \delta_k) \cap (A_k + \{x^*\})$ and so h attains its minimal value on $B(x^*, \delta_k) \cap (A_k + \{x^*\})$ at some point x_k which is different from x^*. By Clarke's theorem [**4**, Theorem 1] there exists v_k in $\partial h(x_k)$ such that $-v_k$ is normal to the convex set $B(x^*, \delta_k) \cap (A_k + \{x^*\})$ at x_k. For each k, let $t_k = |x_k - x^*| > 0$ and $d_k = (x_k - x^*)/t_k$. By [**19**, Corollary 23.8.1], there exist $c_k \geqslant 0$ and a vector u_k normal to $A_k + \{x^*\}$ at x_k such that $v_k + c_k d_k + u_k = 0$ for all k. According to 2.1(a), there exists w_k in $\partial g(x_k)$ such that $v_k = w_k - m^*(x_k - x^*)$ for all k and so

$$w_k - m^*(x_k - x^*) + c_k d_k + u_k = 0, \qquad k \geqslant 1. \tag{1}$$

Since x_k belongs to $A_k + \{x^*\}$, we know that d_k is in A_k. It follows that $x_k \pm t_k d_k$ belong to $A_k + \{x^*\}$ and so $u_k \cdot d_k = 0$. From (1), we obtain

$$w_k \cdot d_k + c_k = m^* t_k, \qquad k \geqslant 1. \tag{2}$$

We may assume that $\{d_k\}$ converges to a unit vector d in R^n. By 2.1 and 2.3, we may assume that $\{w_k\}$ converges to w in $\partial_d g(x^*)$ and so (2) implies that $\{c_k\}$ converges to a nonnegative number c. Hence, in view of (1), $\{u_k\}$ converges to a vector u.

We now wish to show that d belongs to $C(d^*)$. For each k, there exists d_k^* in $C(d^*)$ such that $d_k = T_k(d_k^*)$. Since $|d_k| = 1$ for all k and the T_k^{-1} are uniformly bounded, it follows $\{d_k^*\}$ is bounded and so we may assume that it converges to $d^{\sim}$

in $C(d^*)$. Since $|d_k^* - T_k(d_k^*)| \leqslant |e_k - d^*||d_k^*|$, we have

$$d^{\sim} - d = \lim(d_k^* - T_k(d_k^*)) = 0.$$

Hence d belongs to $C(d^*)$.

We infer from hypothesis (a) that $w \cdot d \geqslant 0$. From (2), we get $w \cdot d + c = 0$ and so, since $c \geqslant 0$, we must have $c = w \cdot d = 0$. From (1), we now obtain $w + u = 0$. To see that u belongs to $N(C(d^*) + x^*, x^*)$, we let e belong to $C(d^*)$. Then, given k, $T_k(e)$ belongs to A_k and so $(T_k(e) + x^* - x_k) \cdot u_k \leqslant 0$. Since $\{T_k(e)\}$ converges to e, we infer $e \cdot u \leqslant 0$ and so u does belong to $N(C(d^*) + x^*, x^*)$. Since conditions (i)–(iv) are all satisfied, we must have $\limsup w_k \cdot d_k/t_k > m^*$. But, from (2), we have $w_k \cdot d_k \leqslant m^* t_k$ for all k, and so we have reached a contradiction and the proof is complete.

2.6 Remarks. To apply Theorem 2.5, one needs to make fruitful choices for the cones $C(d^*)$. Various choices are possible and we shall now describe a few of them.

Notice first that we can choose $C(d^*) = R^n$ for each d^* in $K(S, x^*) \cap L(f, x^*)$. This choice is a natural one to make in the unconstrained case (i.e., when S is a neighborhood of x^*); in the unconstrained case, Theorem 2.5 reduces to Corollary 2.2 of [**2**]. Next, we consider the case in which we choose $C(d^*) = \{td^*: t \geqslant 0\}$ for each d^* in $K(S, x^*) \cap L(f, x^*)$. In the unconstrained case, this choice of $C(d^*)$ in Theorem 2.5 yields a slightly strengthened version of Corollary 2.18 of [**2**]. In the constrained case, it yields a useful and (probably) more versatile variant of Theorem 2.14 of [**2**]; this variant implies the classical sufficiency theorem for C^2 problems having a finite number of constraints (cf. [**2**, 2.17]).

The two choices for $C(d^*)$ just discussed correspond to extreme cases in our present framework, in the sense that they are, respectively, the largest and the smallest possible choices for $C(d^*)$.

Now Theorem 2.5 does not—as stated—imply Theorem 2.1 of [**2**] (because of the use of the auxiliary function M in Theorem 2.1 of [**2**]). However, the present approach can be modified to yield a generalization of that earlier theorem. We shall take up this matter briefly at the end of the paper.

If one checks [**2**, Theorem 3.4], one finds that Ioffe's sufficiency theorem [**12**, Theorem 2] is derivable from Theorem 2.5 with the choice $C(d^*) = \{td^*: t \geqslant 0\}$. As we have just mentioned, the same choice for $C(d^*)$ is made when one derives the classical result for C^2 problems. We offer a brief heuristic explanation as to why this choice of $C(d^*)$ is the proper one. Notice in Theorem 2.5 that, if $C(d^*)$ is enlarged, the restrictions (b)(i) and (b)(ii) become more demanding while the restriction (b)(iv) becomes less demanding. But, in a case in which $\partial g(x^*) = \{0\}$, the point w of (b)(iv) equals 0 and is always in $-N(C(d^*) + x^*, x^*)$ and so the restriction (b)(iv) is completely insensitive to the choice of $C(d^*)$. Thus, when $\partial g(x^*) = \{0\}$, one should choose $C(d^*)$ as small as possible in order to get the best result, since this choice makes hypothesis (b) as undemanding as possible.

Now, we turn to two other possible choices for $C(d^*)$ which lead to new results.

2.7 Corollary. *Let S, W, x^* and f be as in the Introduction. Suppose that g is a real-valued locally Lipschitzian function on W for which $g(x^*) = f(x^*)$ and $g(x) \leqslant f(x)$*

for all x in $S \cap W$. Suppose that S is tangentially regular at x^*. Suppose also that:

(a) *we have $w \cdot d \geqslant 0$ whenever d is any unit vector in $K(S, x^*)$ and w belongs to $\partial_d g(x^*)$*;

(b) *there exists $m^* \geqslant 0$ such that* $\limsup w_k \cdot (x_k - x^*)/|x_k - x^*|^2 > m^*$ *whenever $\{x_k\}$ and $\{w_k\}$ are sequences such that $\{x_k\}$ converges to x^* in direction d in $K(S, x^*)$, $w_k \in \partial g(x_k)$ for each k, and $\{w_k\}$ converges to a point in $-N(S, x^*)$.*

Then there exists $\delta > 0$ such that $f(x) \geqslant f(x^*) + (m^*/2)|x - x^*|^2$ whenever x belongs to $B(x^*, \delta) \cap S$.

PROOF. The cone $K(S, x^*) = T(S, x^*)$ is convex and so we may choose $C(d^*) = K(S, x^*)$ for each d^* in Theorem 2.5.

2.8 REMARKS. If the closed cone $K(S, x^*) \cap L(f, x^*)$ happens to be convex, it provides another natural choice for each $C(d^*)$. We turn our attention now to a special class of locally Lipschitzian functions for which this is true.

First, we must make several definitions. The real-valued locally Lipschitzian function F on W is subdifferentiably regular at x in W in case the directional derivative $F'(x; d)$ exists for all d in R^n and $F^0(x; d) = F'(x; d)$ for all d; actually, Rockafellar [**20**, p. 336] gives the definition for a more general case and shows that it is equivalent to the present one in the locally Lipschitzian case [**20**, p. 339]. The function F is said to be semismooth [**16**, Definition 1] at x in W in case the sequence $\{v_k \cdot d\}$ is always convergent whenever $\{x_k\}$ and $\{v_k\}$ are sequences such that $\{x_k\}$ converges to x in unit direction d and v_k belongs to $\partial F(x_k)$ for each k. Mifflin [**16**] has shown that, if F is semismooth at x, then, for each unit vector d, it is true that the directional derivative $F'(x; d)$ exists and equals $\lim v_k \cdot d$, where $\{v_k\}$ is *any* sequence chosen as in the definition just given. For more information about semismoothness, see [**16**]. For more information about subdifferentiable regularity (which Clarke terms "regularity"), see [**6**, **20** and **5**, Theorem 2.1]. Also Spingarn has shown [**24**, p. 82] that F is both semismooth and subdifferentiably regular at x if and only if $\partial F(x)$ is "submonotone" at x.

Now suppose that f is semismooth and subdifferentiably regular at x^*. Then, $L(f, x^*) = \{d \in R^n: f^0(x; d) \leqslant 0\}$, according to [**2**, Theorem 2.16]; hence, in view of 2.1(b), the cone $L(f, x^*)$ is convex. Therefore, if S is tangentially regular at x^*, it follows that the cone $K(S, x^*) \cap L(f, x^*)$ is convex. This leads to the following corollary.

2.9 COROLLARY. *Suppose that f is semismooth and subdifferentiably regular at x^*. Suppose that all of the hypotheses of Corollary* 2.7 *hold, provided that in* (a) *and* (b) *we replace $K(S, x^*)$ by $K(S, x^*) \cap L(f, x^*)$ and in* (b) *we replace $-N(S, x^*)$ by $-N(K(S, x^*) \cap L(f, x^*), 0)$.*

Then there exists $\delta > 0$ such that $f(x) \geqslant f(x^*) + (m^*/2)|x - x^*|^2$ whenever x belongs to $B(x^*, \delta) \cap S$.

(*Moreover, if g is also semismooth and subdifferentiably regular at x^*, then we can replace hypothesis* (a) *by the assumption that the set $\partial g(x^*) \cap -N(K(S, x^*) \cap L(f, x^*), 0)$ is nonempty.*)

PROOF. We choose $C(d^*) = K(S, x^*) \cap L(f, x^*)$ for each d^*. With these choices, we appeal to Theorem 2.5. It remains only to discuss the final parenthetical statement. Thus, suppose that v^* belongs to $\partial g(x^*) \cap -N(K(S, x^*) \cap L(f, x^*), 0)$. We must prove that (a) holds. Thus, if d belongs to $K(S, x^*) \cap L(f, x^*)$ and w belongs to $\partial_d g(x^*)$, we have, in view of 2.1(b),

$$w \cdot d = g'(x^*; d) = g^0(x^*; d) \geqslant v^* \cdot d \geqslant 0.$$

2.10 EXAMPLE. In Theorem 2.5, we can always choose $C(d^*) = \{td^*: t \geqslant 0\}$ for each d^*. We wish to indicate how other choices can be possible, even when S is not tangentially regular at x^*.

Let S be the set of all points in R^2 for which there exist polar coordinates (r, θ) with $-0.75\pi \leqslant \theta \leqslant 0.75\pi$ and $0 \leqslant r \leqslant \sqrt{2} + 2\cos\theta$. Thus, S is the set bounded by the outer loop of a certain "limacon". With $x^* = (0,0)$, we have $T(S, x^*) = \{(u, v) \in R^2: u \geqslant |v|\}$ while $K(S, x^*) = \{(u, v) \in R^2: u \geqslant -|v|\}$. Hence S is not tangentially regular at x^*. Given $d^* = (u^*, v^*)$ in $K(S, x^*)$, we may select $C(d^*) = \{(u, v) \in R^2: u \geqslant v\}$ provided $u^* \geqslant v^*$ and $C(d^*) = \{(u, v) \in R^2: u + v \geqslant 0\}$ otherwise. With these choices, each normal cone $N(C(d^*) + x^*, x^*)$ consists of one ray only.

2.11 EXAMPLE. We show that Theorem 2.5 is false if (b)(i) is replaced by "$\{x_k\}$ converges to x^* in direction d with $x_k \in S$ for each k".

Let $S = \{(x, y) \in R^2: -1 \leqslant x \leqslant 1 \text{ and } x^4 - x^2 \leqslant y \leqslant 1\}$. Define $f(x, y) = 3x^2 + (2y + 1)^2$ and take $g = f$. Let $x^* = (0,0)$. Observe that $K(S, x^*) = \{(c, d) \in R^2: d \geqslant 0\}$. We choose each $C(d^*)$ to be $K(S, x^*)$. Since f is C^1, we have $\partial f(x, y) = \{(6x, 8y + 4)\}$. If $d \geqslant 0$ then $(c, d) \cdot (0, 4) = 4d \geqslant 0$ and so (a) of Theorem 2.5 holds.

Now suppose that $\{(x_k, y_k)\}$ and $\{w_k\}$ are such that (i)–(iv) of (b) hold *and* suppose each (x_k, y_k) belongs to S. Then, with $z_k = (x_k, y_k)$ and noting that $w_k = (6x_k, 8y_k + 4)$, we have

$$w_k \cdot (z_k - x^*)/|z_k - x^*|^2 = 6 + \left(2y_k^2 + 4y_k\right)/\left(x_k^2 + y_k^2\right).$$

Since $y_k \geqslant x_k^4 - x_k^2$, we have

$$\limsup w_k \cdot (z_k - x^*)/\left(x_k^2 + y_k^2\right) \geqslant 4.$$

However, $x^* = (0,0)$ does not provide a local minimum for f over S. Indeed, take (x, y) with x small and positive and $y = x^4 - x^2$. Then

$$\begin{aligned} f(x, y) &= 3x^2 + (2x^4 - 2x^2 + 1)^2 \\ &= 1 - x^2\{1 - 8x^2 + 8x^4 - 4x^6\} < 1 = f(0,0), \end{aligned}$$

if x is close enough to 0.

3. Problems with a finite number of constraints.

3.1 REMARKS. We shall consider here problem P in the case in which $S = S_1 \cap S_2$, where S_2 is a given closed set and

$$S_1 = \bigcap_{i=1}^{m} \{x \in R^n: g_i(x) \leqslant 0\} \cap \bigcap_{i=m+1}^{q} \{x \in R^n: g_i(x) = 0\}; \tag{3}$$

it is assumed here that each of the functions g_i is locally Lipschitzian on W. Necessary conditions for x^* to be a local minimizer for this problem have been given successively in [**4**, **11** and **22**]. According to Rockafellar's result [**22**, Theorem 1], it is true that if problem P is "calm" at x^* (see [**4**, p. 172 or **22**, Proposition 1]) then there exist multipliers $a_1,\ldots,a_q$ such that

$$a_i \geqslant 0 \quad \text{for } i = 1,\ldots,m, \tag{4}$$

$$a_i g_i(x^*) = 0 \quad \text{for } i = 1,\ldots,m, \tag{5}$$

$$0 \in \partial\{f + a_1 g_1 + a_2 g_2 + \cdots + a_q g_q + \psi_2\}(x^*). \tag{6}$$

In (6), ψ_2 is the indicator function of the set S_2; we have $\psi_2(x) = 0$ if $x \in S_2$ and $\psi_2(x) = +\infty$ otherwise. Notice that in (6) we are considering the subdifferential of a function which is not locally Lipschitzian.

If multipliers $a_1,\ldots,a_q$ satisfying (4)–(6) exist, we define the Lagrangian L by $L = f + a_1 g_1 + \cdots + a_q g_q$ and observe that we may take the auxiliary function g in Theorem 2.5 to be L. We shall not write out in detail the special case of Theorem 2.5 which is produced by this choice for g. We shall, however, state a specific theorem which we can obtain when the functions of the problem are both semismooth and subdifferentiably regular at x^*.

3.2 THEOREM. *Suppose that the functions $f, g_1,\ldots,g_m$ are both semismooth and subdifferentiably regular at x^* and that the functions $g_{m+1},\ldots,g_q$ are C^1 near x^*. Let I be the set of all indices i for which $1 \leqslant i \leqslant m$ and $g_i(x^*) = 0$. Suppose that* 0 *does not belong to the convex hull of the union of the sets $\partial g_i(x^*)$ for i in I and the sets $\partial g_i(x^*) \cup -\partial g_i(x^*)$ for $i = m+1,\ldots,q$. Suppose that $T(S_1, x^*) \cap \operatorname{int} T(S_2, x^*)$ is nonvoid or that $\operatorname{int} T(S_1, x^*) \cap T(S_2, x^*)$ is nonvoid. Suppose also that S_2 is tangentially regular at x^*.*

Next, suppose that there exist multipliers $a_1,\ldots,a_q$ satisfying (4)–(6) *and put $L = f + a_1 g_1 + \cdots + a_q g_q$. Suppose finally that $m^* \geqslant 0$ exists such that it is true that*

$$\limsup w_k \cdot (x_k - x^*)/|x_k - x^*|^2 > m^*$$

whenever $\{x_k\}$ and $\{w_k\}$ are sequences such that:

(v) *$\{x_k\}$ converges to x^* in direction d in $T(S, x^*)$ for which $g_i'(x^*; d) = 0$ for all i in I such that $a_i > 0$*;

(vi) *w_k belongs to $\partial L(x_k)$ for each k*;

(vii) *$\{w_k\}$ converges to a point in $-N(L(f, x^*) \cap T(S, x^*), 0)$.*

Then there exists $\delta > 0$ such that $f(x) \geqslant f(x^*) + (m^*/2)|x - x^*|^2$ for every x in $B(x^*, \delta) \cap S$.

PROOF. We derive this theorem from Theorem 2.5. Notice first that $f(x^*) = L(x^*)$ and that $L \leqslant f$ on $S \cap W$. It follows from [**2**, Proposition 2.9] that L is both semismooth and subdifferentiably regular at x^*. As noted above in 2.8, the set $L(f, x^*)$ is a closed convex cone. It follows from [**20**, Corollary 2 to Theorem 5] that S_1 is tangentially regular at x^* and hence from [**20**, Corollary 4 to Theorem 2] that S is tangentially regular at x^*. Hence the closed cone $K(S, x^*) \cap L(f, x^*)$ is convex; for each d^* in $K(S, x^*) \cap L(f, x^*)$, we put $C(d^*) = K(S, x^*) \cap L(f, x^*)$.

To verify that (a) holds in Theorem 2.5, take d in $K(S, x^*) \cap L(f, x^*)$ and w in $\partial_d L(x^*)$. From [**20**, Corollary 2 to Theorem 2 and **20**, equation (2.10)], we infer that (6) implies the existence of v^* in $\partial L(x^*)$ such that $-v^*$ belongs to $N(S_2, x^*)$. Since L is both semismooth and subdifferentiably regular at x^*, we have

$$w \cdot d = L'(x^*; d) = L^0(x^*; d) \geqslant v^* \cdot d \geqslant 0. \tag{7}$$

It remains only to check that (b) of Theorem 2.5 holds in the present situation. So, suppose that $\{x_k\}$ and $\{w_k\}$ satisfy (i)–(iv) of Theorem 2.5 with g taken to be L. Then (vi) and (vii) of the present theorem obviously hold. As in (7), we have $w \cdot d = L'(x^*; d) \geqslant 0$ and so

$$f'(x^*; d) + a_1 g_1'(x^*; d) + \cdots + a_q g_q'(x^*; d) \geqslant 0. \tag{8}$$

Since d belongs to $L(f, x^*)$, we have $f'(x^*; d) \leqslant 0$ and, since d belongs to $K(S, x^*)$, we have $g_i'(x^*; d) = 0$ for $i > m$ and $g_i'(x^*; d) \leqslant 0$ for i in I. Hence, we infer from (8) that $g_i'(x^*; d) = 0$ for all i in I for which $a_i > 0$. It follows that $\{x_k\}$ and d satisfy (v). Therefore, from (v)–(vii) we infer

$$\limsup w_k \cdot (x_k - x^*)/|x_k - x^*|^2 > m^*$$

and so (b) must hold.

3.3 Remarks. The requirement in Theorem 3.2 that 0 not be in the convex hull of the union of certain subdifferentials is a constraint qualification. It is a generalization of the Mangasarian-Fromovitz constraint qualification (see [**15** or **23**, p. 15]). Of course, the classical sufficiency theorem for C^2 functions requires no such qualification and so cannot be a special case of Theorem 3.2. However, as we have remarked above, the proper way to recover the classical theorem is to choose $C(d^*) = \{td^*: t \geqslant 0\}$ for each d^*.

3.4 Remarks. It is possible to construct variants of Theorem 2.5 by use of certain additional auxiliary functions. We shall discuss here one such example. We continue to work with the problem of this section, where $S = S_1 \cap S_2$, with S_1 given by (3).

Let us suppose that positive numbers r and m^* are given. Let x^* belong to $S \cap W$ as before. We define a function M on W by letting $M(x)$ be the largest of the numbers $g_1(x), \ldots, g_m(x)$, and

$$f(x) - f(x^*) - (m^*/2)|x - x^*|^2 + r|g_{m+1}(x)| + \cdots + r|g_q(x)|.$$

Notice that $M(x^*) = 0$, since x^* belongs to S. The use of the function M is motivated by this observation: If x^* minimizes M on $B(x^*, \delta) \cap S_2$ then we have $f(x) \geqslant f(x^*) + (m^*/2)|x - x^*|^2$ for every x in $B(x^*, \delta) \cap S$. The theorem which follows provides a generalization of Theorems 2.1 and 2.14 of [**2**] in the same way that Theorem 2.5 provides a generalization of Corollaries 2.2 and 2.18 of [**2**].

3.5 Theorem. *Let W, f and x^* be as in the Introduction and let $S = S_1 \cap S_2$ as above. Let g be a real-valued locally Lipschitzian function on W for which $g(x^*) = M(x^*) = 0$ and $g(x) \leqslant M(x)$ for every x in $S \cap W$. Suppose that, to each unit vector d^* in $K(S, x^*) \cap L(f, x^*)$, there corresponds a closed convex cone $C(d^*)$ for which*

$d^* \in C(d^*)$. We suppose also that:

(a) *we have $w \cdot d \geqslant 0$ whenever d is any unit vector in $C(d^*)$ for some d^* in $K(S, x^*) \cap L(f, x^*)$ and w is any generalized gradient in $\partial_d g(x^*)$;*

(b) *we have $\limsup w_k \cdot (x_k - x^*)/|x_k - x^*|^2 > 0$ whenever $\{x_k\}$ and $\{w_k\}$ are sequences and d^* and d are unit vectors for which*

(i) *$\{x_k\}$ converges to x^* in direction d,*

(ii) *$d^* \in K(S, x^*) \cap L(f, x^*)$ and $d \in C(d^*)$,*

(iii) *$w_k \in \partial g(x_k)$ for each k,*

(iv) *$\{w_k\}$ converges to a point w in $-N(C(d^*) + x^*, x^*)$.*

Then there exists $\delta > 0$ such that $f(x) \geqslant f(x^*) + (m^*/2)|x - x^*|^2$ whenever x belongs to $B(x^*, \delta) \cap S$.

PROOF. Suppose that the desired conclusion is false and pick a sequence $\{\delta_k\}$ of positive numbers decreasing to 0 with $\delta_1 < 1$. Given k, there exists z_k in $B(x^*, \delta_k) \cap S$ such that $f(z_k) - f(x^*) < (m^*/2)|z_k - x^*|^2$. Put $h = g$ and note that $z_k \neq x^*$ for each k. We have $h(z_k) = g(z_k) \leqslant M(z_k) \leqslant 0 = M(x^*) = h(x^*)$ for each k.

The rest of the proof can be taken directly from the proof of Theorem 2.5 provided minor changes are made. In equations (1) and (2), for example, the summands involving m^* are replaced by 0; these modifications are typical of the necessary minor changes.

REFERENCES

1. R. W. Chaney, *Second-order optimality conditions for a class of nonlinear programming problems*, J. Math. Anal. Appl. **76** (1980), 516–534.

2. _____, *Second-order sufficiency conditions for nondifferentiable programming problems*, SIAM J. Control Optim. **7** (1982), 463–475.

3. _____, *On sufficient conditions in nonsmooth optimization*, Math. Oper. Res. **20** (1982), 20–33.

4. F. H. Clarke, *A new approach to Lagrange multipliers*, Math. Oper. Res. **1** (1976), 165–174.

5. _____, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.

6. _____, *Generalized gradients of Lipschitz functionals*, Adv. in Math. **40** (1981), 52–67.

7. _____, *Nonsmooth analysis and optimization* (forthcoming).

8. R. Fletcher and G. A. Watson, *First and second order conditions for a class of nondifferentiable optimization problems*, Math. Programming **18** (1980), 291–307.

9. M. R. Hestenes, *Calculus of variations and optimal control theory*, Wiley, New York and London, 1966.

10. _____, *Optimization theory: The finite dimensional case*, Interscience, New York, 1975.

11. J.-B. Hiriart-Urruty, *Refinements of necessary optimality conditions in nondifferentiable programming. I*, Appl. Math. Optim. **5** (1979), 63–82.

12. A. D. Ioffe, *Necessary and sufficient conditions for a local minimum 3: Second-order conditions and augmented duality*, SIAM J. Control Optim. **17** (1979), 266–288.

13. G. Lebourg, *Generic differentiability of Lipschitzian functions*, Trans. Amer. Math. Soc. **256** (1979), 125–144.

14. _____, *Valeur moyenne pour gradient généralisé*, C. R. Acad. Sci. Paris Ser. A **281** (1975), 795–797.

15. O. L. Mangasarian and S. Fromovitz, *The Fritz John necessary optimality conditions in the presence of equality and inequality constraints*, J. Math. Anal. Appl. **17** (1967), 37–47.

16. R. Mifflin, *Semismooth and semiconvex functions in constrained optimization*, SIAM J. Control Optim. **15** (1977), 959–972.

17. J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York, 1970.

18. R. T. Rockafellar, *Clarke's tangent cones and the boundaries of closed sets in R^n*, Nonlinear Anal. Theor. Meth. Appl. **3** (1979), 145–154.

19. _____, *Convex analysis*, Princeton Univ. Press, Princeton, N.J., 1970.

20. _____, *Directionally Lipschitzian functions and subdifferential calculus*, Proc. London Math. Soc. **39** (1979), 331–355.

21. _____, *Generalized directional derivatives and subgradients of nonconvex functions*, Canad. J. Math. **32** (1980), 257–280.

22. _____, *Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming*, Math. Programming Study **17** (1982), 28–66.

23. _____, *La théorie des sous-gradients et ses applications: Fonctions convexes et non convexes*, Collection Chaire Aisenstadt, Presses de l'Université de Montréal, Montreal, 1979.

24. J. E. Spingarn, *Submonotone subdifferentials of Lipschitz functions*, Trans. Amer. Math. Soc. **264** (1981), 77–89.

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, WESTERN WASHINGTON UNIVERSITY, BELLINGHAM, WASHINGTON 98225