SIAM J. CONTROL AND OPTIMIZATION
Vol. 31, No. 5, pp. 1340–1359, September 1993

014

WEAK SHARP MINIMA IN MATHEMATICAL PROGRAMMING*

J. V. BURKE† AND M. C. FERRIS‡

Abstract. The notion of a *sharp*, or *strongly unique*, minimum is extended to include the possibility of a nonunique solution set. These minima will be called *weak sharp minima*. Conditions necessary for the solution set of a minimization problem to be a set of weak sharp minima are developed in both the unconstrained and constrained cases. These conditions are also shown to be sufficient under the appropriate convexity hypotheses. The existence of weak sharp minima is characterized in the cases of linear and quadratic convex programming and for the linear complementarity problem. In particular, a result of Mangasarian and Meyer is reproduced that shows that the solution set of a linear program is always a set of weak sharp minima whenever it is nonempty. Consequences for the convergence theory of algorithms are also examined, especially conditions yielding finite termination.

Key words. finite termination, strongly unique minima, sharp minima

AMS subject classifications. 90C20, 90C30, 65K05

1. Introduction.

Let $f\colon X \mapsto \overline{\mathbf{R}} := \mathbf{R}\bigcup\{-\infty,\infty\}$; we say that f has a sharp minimum at $\bar{x} \in \mathbf{R}^n$ if $f(x) \geq f(\bar{x}) + \alpha\,\|x-\bar{x}\|$, for all x near $\bar{x}$ and some $\alpha > 0$. The notion of a sharp minimum, or equivalently, a strongly unique local minimum, has far reaching consequences for the convergence analysis of many iterative procedures [1], [8], [11], [12], [17], [18]. In this article, we extend the notion of a sharp minimum to include the possibility of a nonunique solution set. We say that $\bar{S} \subset \mathbf{R}^n$ is a set of *weak sharp minima* for the function f relative to the set $S \subset \mathbf{R}^n$ where $\bar{S} \subset S$ if there is an $\alpha > 0$ such that

$$f(x) \geq f(y) + \alpha \operatorname{dist}(x \mid \bar{S}), \tag{1}$$

for all $x \in S$ and $y \in \bar{S}$ where

$$\operatorname{dist}(x \mid \bar{S}) := \inf_{z \in \bar{S}} \|x - z\|.$$

The constant α and the set $\bar{S}$ are called the modulus and domain of sharpness for f over S, respectively. Clearly, $\bar{S}$ is a set of global minima for f over S. The notion of weak sharp minima is easily localized. We will say that $\bar{x} \in \mathbf{R}^n$ is a local weak sharp minimum for f on $S \subset \mathbf{R}^n$ if there exists a set $\bar{S} \subset S$ and a parameter $\delta > 0$ with $\bar{x} \in \bar{S}$ such that the set $\bar{S} \cap \{x : \|x - \bar{x}\| \leq \delta\}$ is a set of weak sharp minima for the function

$$f_\delta(x) := \begin{cases} f(x), & \text{if } \|x-\bar{x}\| \leq \delta, \\ +\infty, & \text{otherwise,} \end{cases}$$

relative to the set S. Since the restriction to the local setting is straightforward, we will concentrate on the global definition.

The study of weak sharp minima is motivated primarily by applications in convex and convex composite programming, where such minima commonly occur. For example, such minima frequently occur in linear programming, linear complementarity,

* Received by the editors September 30, 1991; accepted for publication (in revised form) April 5, 1992. This material is based on research supported by National Science Foundation grants CCR-9157632 and DMS-9102059 and Air Force Office of Scientific Research grant AFOSR-89-0410.

† Department of Mathematics, GN–50, University of Washington, Seattle, Washington 98195.

‡ Computer Sciences Department, University of Wisconsin, Madison, Wisconsin 53706.

and least distance or projection problems. The goals of this study are to quantify this property, investigate its geometric structure, characterize its occurrence in simple convex programming problems, and, finally, to analyze its impact on the convergence of algorithms. Furthermore, although our primary interest is with convex programming, we also investigate the significance of weak sharp minima for nonconvex problems. However, in the latter case, rather strong regularity conditions are required to yield significant extentions of the convex case. Nonetheless, we do obtain some very interesting and significant results for differentiable problems with convex constraints. These results extend and refine earlier work of Al-Khayyal and Kyparisis [1] on the finite termination of algorithms at sharp minima. In a later study, we also show how these results can be applied to convex composite optimization problems to establish the quadratic rate of convergence of a variety of algorithms. This study builds on the work initiated in [9].

Our study begins in §2 with the derivation of first-order necessary conditions for the solution set of a problem to be a set of weak sharp minima. The unconstrained ($S = \mathbf{R}^n$) and constrained cases are treated separately. When the problem data is convex, it is shown that these conditions are also sufficient. In the third section these results are applied to three important classes of convex programs: quadratic programming, linear programming, and the linear complementarity problem. In the final section we examine certain tools for studying the convergence of algorithms in the presence of weak sharp minima. In particular, it is shown how we can attain finite convergence to weak sharp minima.

The notation that we employ is for the most part standard; however, a partial list is provided for the readers' convenience. The *inner product* on $\mathbf{R}^n$ is defined as the bilinear form

$$\langle y, x \rangle := \sum_{i=1}^{n} y_i x_i.$$

We denote a *norm* on $\mathbf{R}^n$ by $\|\cdot\|$. Each norm defines a norm dual to it and is given by

$$\|x\|_o := \sup_{\|y\| \le 1} \langle y, x \rangle .$$

The associated closed unit balls for these norms are denoted by $\mathbf{B}$ and $\mathbf{B}^\circ$, respectively. The 2-norm plays a special role in our development and is denoted by

$$\|x\|_2 := \sqrt{\langle x, x \rangle}.$$

If it is understood from the context that we are speaking of the 2-norm, then we will drop the subscript "2" from this notation.

Given two subsets A and B of $\mathbf{R}^n$ and $\beta \in \mathbf{R}$, we define

$$A \pm \beta B := \{a \pm \beta b : a \in A,\ b \in B\}.$$

On the other hand,

$$A \setminus B := \{a \in A : a \notin B\}.$$

If $A \subset \mathbf{R}^n$ then the *polar* of A is defined to be the set

$$A^\circ := \{x^* \in \mathbf{R}^n : \langle x^*, x \rangle \le 1\ \forall x \in A\}.$$

This notation is consistent with the definition of the dual unit ball B°. The *indicator* and *support* functions for A are given by

$$\psi(x \mid A) := \begin{cases} 0 & \text{if } x \in A, \\ +\infty & \text{otherwise} \end{cases}$$

and

$$\psi^*(x \mid A) := \sup\{\langle x^*, x\rangle : x^* \in A\},$$

respectively. Moreover, we write $\operatorname{int} A$ for the interior of A, $\operatorname{cl} A$ for the closure of A, and $\operatorname{span} A$ for the linear span of the elements of A. The *relative interior* of A, denoted $\operatorname{ri} A$, is the interior of A relative to the *affine hull* of A, which is given by

$$\operatorname{aff} A := \left\{ \sum_{k=1}^{s} \lambda_k x_k \;\middle|\; \begin{array}{c} s \in \{1, 2, \ldots\},\ x_k \in A \text{ and } \lambda_k \in \mathrm{R} \\ \text{for } k = 1, 2, \ldots, s, \text{ with } \sum_{k=1}^{s} \lambda_k = 1 \end{array} \right\}.$$

The subspace *perpendicular* to A is defined to be

$$A^\perp := \{y \in \mathrm{R}^n \,:\, \langle y, x\rangle = 0 \text{ for all } x \in A\}.$$

If A is closed, then we define the *projection* of a point $x \in \mathrm{R}^n$ onto the set A as the set of all points in A that are closest to A in a given norm. In this paper, we will only speak of the projection with respect to the 2-norm; it is denoted by

$$P(x \mid A) := \{\bar{y} \in A \,:\, \|x - \bar{y}\|_2 = \inf_{y \in A} \|x - y\|_2\}.$$

The projection is an example of a multivalued mapping on R^n. The set A is said to be *convex* if the line segment connecting any two points in A is also contained in A. The *convex hull* of the set A, denoted $\operatorname{co}(A)$, is the smallest convex set that contains A; that is, $\operatorname{co}(A)$ is the intersection of all convex sets that contain A. It is interesting to note that the projection operator can be used to characterize the closed convex subsets of R^n. That is, the set A is closed and convex if and only if the projection operator for A, $P(\cdot \mid A)$, is single valued on all of R^n [2], [16].

Given $x \in A$, we define the *normal cone* to A at x, denoted $N(x \mid A)$, to be the closure of the convex hull of all limits of the form

$$\lim_k t_k^{-1}(x_k - p_k),$$

where the sequences $\{t_k\} \subset \mathrm{R}$, $\{p_k\} \subset A$, and $\{x_k\} \subset \mathrm{R}^n$ satisfy $t_k \downarrow 0$, $p_k \in P(x_k \mid A)$, and $p_k \to x$. If A is convex, we can show that this definition implies that

$$N(x \mid A) = \{x^* \in \mathrm{R}^n \,:\, \langle x^*, y - x\rangle \le 0 \ \forall y \in A\}.$$

The *tangent cone* to A at x is defined dually by the relation

$$T(x \mid A) := N(x \mid A)^\circ.$$

If A is convex, we have the relation

$$T(x \mid A) = \operatorname{cl}\left[\cup_{\lambda \ge 0} \lambda(A - x)\right].$$

The *contingent cone* to A at x plays a role similar to that of the tangent cone but is, in general, larger. The contingent cone to A at x is given by

$$K(x \mid A) := \{d \in \mathrm{R}^n \, : \, \exists t_k \downarrow 0, \; d^k \to d, \text{with } x + t_k d^k \in A\} \, .$$

The set A is said to be *regular* at $x \in A$ if $T(x \mid A) = K(x \mid A)$. In particular, every convex set is regular.

Let $f\colon X \mapsto \overline{\mathrm{R}} := \mathrm{R} \cup \{+\infty\}$. The *domain* and *epigraph* of f are given by

$$\operatorname{dom} f := \{x \in \mathrm{R}^n \, : \, f(x) < +\infty\}$$

and

$$\operatorname{epi} f := \{(x, \lambda) \in \mathrm{R}^n \times \mathrm{R} \, : \, f(x) \le \lambda\},$$

respectively. Observe that f is lower semicontinuous if and only if $\operatorname{epi} f$ is closed. For $x \in \operatorname{dom} f$, we define the *subdifferential* of f at x to be the set

$$\partial f(x) := \{x^* \, : \, (x^*, -1) \in N((x, f(x)) \mid \operatorname{epi} f)\},$$

and the *singular subdifferential* of f at x to be the set

$$\partial^\infty f(x) := \{x^* \, : \, (x^*, 0) \in N((x, f(x)) \mid \operatorname{epi} f)\}.$$

The mappings ∂f and $\partial^\infty f$ are further examples of multivalued mappings on R^n. We observe that the set $\partial f(x) \cup \partial^\infty f(x)$ is always nonempty even though ∂f may be empty at certain points. Moreover, the function f is locally Lipschitzian on R^n if and only if ∂f is nonempty and compact valued on all of R^n. The domain of ∂f is the set

$$\operatorname{dom} \partial f := \{x^* \in \mathrm{R}^n \, : \, \partial f(x) \neq \emptyset\}.$$

If f is convex, then this subdifferential coincides with the usual subdifferential from convex analysis. The *generalized directional derivative* of f is the support function of $\partial f(x)$,

$$f^\circ(x; d) := \psi^*(d \mid \partial f(x)),$$

and the *contingent directional derivative* of f at x in the direction d is given by

$$f^-(x; d) := \liminf_{\substack{u \to d \\ t \downarrow 0}} \frac{f(x + tu) - f(x)}{t} \, .$$

The relation $f^-(x; d) \le f^\circ(x; d)$ always holds. The function f is said to be *regular* at x if $f^\circ(x; d) = f^-(x; d)$ in which case the usual directional derivative,

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

exists and equals this common value. See [7] for further details of subdifferential calculus.

2. Subdifferential geometry. We begin with a study of some geometric consequences of of weak sharp minima. Specifically, we are interested in first-order necessary conditions. The general unconstrained ($S = \mathbb{R}^n$) and constrained cases are treated separately. In both cases, it is shown that the necessary conditions are also sufficient under appropriate convexity hypotheses. The following preliminary result is required.

LEMMA 2.1. *Suppose $f\colon \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ is closed, proper, and convex, the sets $\bar{S} := \arg\min\{f(x) : x \in \mathbb{R}^n\}$ and C are nonempty, closed, and convex subsets of $\mathbb{R}^n$ with $C \subseteq \bar{S}$, and $\alpha > 0$. The following are equivalent:*

1. *$\alpha \mathbf{B}^\circ \bigcap N(x \mid \bar{S}) \subseteq \partial f(x)$, for all $x \in C$,*
2. *$\alpha \mathbf{B}^\circ \bigcap \bigcup_{x \in C} N(x \mid \bar{S}) \subseteq \bigcup_{x \in C} \partial f(x)$.*

Proof. [1 $\Longrightarrow$ 2]. Trivial.

[2 $\Longrightarrow$ 1]. Let $z \in C$ and $z^* \in \alpha \mathbf{B}^\circ \bigcap N(z \mid \bar{S})$. Then by hypothesis, $z^* \in \partial f(u)$ for some $u \in C$. Since $C \subseteq \bar{S}$ implies that $\partial f(u) \subseteq N(u \mid \bar{S})$, hence $z^* \in N(u \mid \bar{S})$, it follows from $z^* \in N(z \mid \bar{S})$ that

$$\langle z^*, u \rangle = \langle z^*, z \rangle . \tag{2}$$

However, $z^* \in \partial f(u)$ is by definition $f(y) - f(u) \geq \langle z^*, y - u \rangle$, for all y. Since u, $z \in \bar{S}$, $f(u) = f(z)$ so that (2) gives $f(y) - f(z) \geq \langle z^*, y - z \rangle$, for all y, or equivalently, $z^* \in \partial f(z)$. □

Necessary conditions for weak sharp minima in the unconstrained case now follow.

THEOREM 2.2. *Let $f\colon \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ be lower semicontinuous and $\alpha > 0$. Consider the following statements:*

1. *The set $\bar{S}$ is a set of weak sharp minima for the function f on $\mathbb{R}^n$ with modulus α.*
2. *For all $d \in \mathbb{R}^n$,*
$$f^-(x; d) \geq \alpha \operatorname{dist}(d \mid K(x \mid \bar{S})).$$
3. *For all $d \in \mathbb{R}^n$*
$$f^\circ(x; d) \geq \alpha \operatorname{dist}(d \mid T(x \mid \bar{S})).$$
4. *The inclusion*
$$\alpha \mathbf{B}^\circ \bigcap N(x \mid \bar{S}) \subseteq \partial f(x)$$
holds.
5. *The inclusion*
$$\alpha \mathbf{B}^\circ \bigcap \left[\bigcup_{x \in \bar{S}} N(x \mid \bar{S}) \right] \subseteq \bigcup_{x \in \bar{S}} \partial f(x)$$
holds.
6. *For all $y \in \mathbb{R}^n$,*
$$f'(p; y - p) \geq \alpha \operatorname{dist}(y \mid \bar{S}),$$
where $p \in P(y \mid \bar{S})$.

Statement 1 implies statement 2 for all $x \in \bar{S}$. Statement 2 implies statement 3 at points $x \in \bar{S}$ at which $\bar{S}$ is regular. Statements 3 and 4 are equivalent. If f is closed proper and convex and the set $\bar{S}$ is nonempty closed and convex, then statements 1–6 are equivalent with 2, 3, and 4 holding at every point of $\bar{S}$.

Proof.

[1 $\Longrightarrow$ 2]. Let $x \in \bar{S}$. The hypothesis guarantees that for all t and d'

$$f(x+td') - f(x) \geq \alpha \text{dist}\,(x+td' \mid \bar{S}),$$

which implies that

$$\frac{f(x+td') - f(x)}{t} \geq \alpha \frac{\text{dist}\,(x+td' \mid \bar{S}) - \text{dist}\,(x \mid \bar{S})}{t}.$$

By taking lim infs of both sides as $d' \to d$ and $t \downarrow 0$ and applying [4, Thm. 4], we obtain the result.

[(2 plus regularity) $\Longrightarrow$ 3]. Simply observe that regularity at $x \in \bar{S}$ implies the equivalence $T(x \mid \bar{S}) = K(x \mid \bar{S})$ and by definition $f^\circ(x;\cdot) \geq f^-(x;\cdot)$.

[3 $\Longleftrightarrow$ 4]. We recall from [5, Thm. 3.1] that if $K \subset \mathrm{R}^n$ is a nonempty closed convex cone, then

$$\text{dist}\,(x \mid K) = \psi^*(x \mid K^\circ \cap \mathrm{B}^\circ).$$

The result now follows from the fact that $f^\circ(x;\cdot) = \psi^*(\cdot \mid \partial f(x))$.

Observe that if f is closed proper and convex, and $\bar{S}$ is nonempty closed and convex, then f is regular on its domain and $\bar{S}$ is regular at each of its elements. Hence either one of the statements 1 or 2 implies both 3 and 4 for all $x \in \bar{S}$.

[(4 holds for all $x \in \bar{S}$) $\Longrightarrow$ 5]. Trivial.

[(5 plus convexity) $\Longrightarrow$ 4]. Convexity and Lemma 2.1 combine to establish that 5 implies 4.

[(5 plus convexity) $\Longrightarrow$ 1]. Given $y \in \mathrm{R}^n$, Theorem 1 in [4] implies the existence of a $x^* \in \alpha\mathrm{B}^\circ \bigcap N(P(y \mid \bar{S}) \mid \bar{S})$ such that $\alpha\text{dist}\,(y \mid \bar{S}) = \langle x^*, y\rangle - \psi^*(x^* \mid \bar{S})$. Thus, by hypothesis, there exists a $x \in \bar{S}$ with $x^* \in \partial f(x)$. Hence

$$\begin{aligned} f(y) &\geq f(x) + \langle x^*, y - x\rangle \\ &\geq f(x) + \langle x^*, y\rangle - \langle x^*, x\rangle \\ &\geq f(x) + \langle x^*, y\rangle - \psi^*(x^* \mid \bar{S}) \\ &= f(x) + \alpha\text{dist}\,(y \mid \bar{S}). \end{aligned}$$

Since $y \in \mathrm{R}^n$ is arbitrary, the result is obtained.

[(1 plus convexity) $\Longrightarrow$ 6]. Let y be given and define $p := P(y \mid \bar{S})$ so that $f(y) \geq f(p) + \alpha\text{dist}\,(y \mid \bar{S}) = f(p) + \alpha\,\|y-p\|$. Let $z = \lambda y + (1-\lambda)p$ for $\lambda \in [0,1]$. Then $p = P(z \mid \bar{S})$ and

$$f(z) \geq f(p) + \alpha\,\|z-p\| = f(p) + \alpha\lambda\,\|y-p\|$$

implying that

$$\frac{f(p+\lambda(y-p)) - f(p)}{\lambda} \geq \alpha\,\|y-p\|\,.$$

The result now follows in the limit.

[(6 plus convexity) $\Longrightarrow$ 1]. Since f is convex it follows that for all x and y

$$f'(x; y-x) = \inf_{t>0} \left[\frac{f(x+t(y-x)) - f(x)}{t} \right]$$

so that for any y we may take $x = P(y \mid \bar{S}) = p$, $t = 1$ and

$$f(p+y-p) - f(p) \geq f'(p; y-p) \geq \alpha \mathrm{dist}\,(y \mid \bar{S}). \qquad \square$$

COROLLARY 2.3. *Suppose f is closed proper and convex and has a set of weak sharp minima $\bar{S}$ that is nonempty, closed, convex, and compact. Then*

$$0 \in \mathrm{int} \bigcup_{\bar{x} \in \bar{S}} \partial f(\bar{x}).$$

Proof. The corollary follows if we can show that

$$\bigcup_{x \in \bar{S}} N(x \mid \bar{S}) = \mathbf{R}^n.$$

Clearly, $\bigcup_{x \in \bar{S}} N(x \mid \bar{S}) \subset \mathbf{R}^n$, so let $y \in \mathbf{R}^n$. By continuity of $\langle y, \cdot \rangle$ and compactness of $\bar{S}$

$$z^* \in \arg\max{}_{z \in \bar{S}} \langle y, z \rangle$$

so that $\langle y, z - z^* \rangle \leq 0$, for all $z \in \bar{S}$. Hence $y \in N(z^* \mid \bar{S})$. $\square$

In the constrained case, we must introduce a constraint qualification to guarantee the validity of the type of first-order optimality conditions that are required for our analysis. For the problem

$$\begin{array}{ll} \text{minimize} & f(x), \\ x \in S & \end{array} \tag{3}$$

these optimality conditions take the form

$$0 \in \partial f(x) + N(x \mid S). \tag{4}$$

Condition (4) is not always guaranteed to be valid even in the fully convex case, so a constraint qualification is required.

Example 2.4. Consider (3), where $f: \mathbf{R} \mapsto \overline{\mathbf{R}}$ is given by

$$f(x) := \begin{cases} -\sqrt{1+x^2}, & \text{for } x \in [-1, 1] \\ +\infty, & \text{otherwise,} \end{cases}$$

and $S := \{x : x \leq -1\}$. This is a convex program with a closed proper convex objective function having unique global solution $\bar{x} = -1$. However, (4) does not hold since $\partial f(\bar{x}) = \emptyset$.

For this reason we introduce the following constraint qualification due to Rockafellar [20].

DEFINITION 2.5. We say that the *basic constraint qualification* (BCQ) for (3) is satisfied at $x \in S$ if for every $u \in \partial^{\infty} f(x)$ and $v \in N(x \mid S)$ such that $u + v = 0$ it must be the case that $u = v = 0$. The BCQ is said to be satisfied on a set $\bar{S} \subset S$ if it is satisfied at every point of $\bar{S}$.

From Rockafellar [20, Cor. 5.2.1], we know that the optimality condition (4) is satisfied at every local solution to (3) at which the BCQ holds. In particular, if f is locally Lipschitzian on $\mathbf{R}^n$, then $\partial^\infty f(x) = \{0\}$ on all of $\mathbf{R}^n$; hence the BCQ is vacuously satisfied on all of S, so (4) holds at every local minima for (3).

THEOREM 2.6. *Suppose* $f: \mathbf{R}^n \mapsto \overline{\mathbf{R}}$ *is lower semicontinuous and* $\bar{S} \subset S$ *are nonempty closed subsets of* $\mathbf{R}^n$.

(a) *The inclusion*

$$\alpha \mathbf{B} \subset \partial f(\bar{x}) + \left[T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})\right]^\circ. \tag{5}$$

holds at $\bar{x} \in \bar{S}$ *if and only if*

$$f^\circ(\bar{x}; z) \geq \alpha \|z\| \ \ \forall\ z \in T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S}). \tag{6}$$

(b) *If* $\bar{S}$ *is a set of weak sharp minima for* f *over* S *with modulus* $\alpha > 0$ *such that the BCQ holds at every point of* $\bar{S}$, *then for each* $\bar{x} \in \bar{S}$ *at which* f, S, *and* $\bar{S}$ *are regular, we have the inclusion* (5).

(c) *If we further assume that* f *is closed proper and convex and the sets* $\bar{S}$ *and* S *are nonempty closed and convex, then* $\bar{S}$ *is a set of weak sharp minima for* f *over* S *with modulus* $\alpha > 0$ *if and only if the inclusion* (5) *holds for all* $\bar{x} \in \bar{S}$.

Proof. (a) We show that (5) and (6) are equivalent. Clearly, both statements are false if $\partial f(\bar{x})$ is empty, so we assume it to be nonempty. First note that (6) is equivalent to

$$\sup\{\langle x^*, z\rangle \mid x^* \in \partial f(\bar{x})\} \geq \alpha \|z\| \,\forall z \in T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S}). \tag{7}$$

We show this is equivalent to

$$\sup\left\{\langle x^*, z\rangle \,\middle|\, x^* \in \partial f(\bar{x}) + \left[T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})\right]^\circ\right\} \geq \alpha \|z\| \ \ \forall z \in \mathbf{R}^n. \tag{8}$$

This is accomplished in two parts. First, it is shown that the supremum in (8) is infinite if $z \notin T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})$ and then it is shown that the suprema in (7) and (8) are equal if $z \in T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})$. Suppose $z \notin T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})$. Then there exists $z^* \in \left[T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})\right]^\circ$ such that $\langle z^*, z\rangle > 0$. Let $x^* \in \partial f(\bar{x})$, which is nonempty by assumption, and consider $x^* + \lambda z^*$ as $\lambda \to \infty$. Since $\langle x^* + \lambda z^*, z\rangle \uparrow +\infty$ as $\lambda \uparrow +\infty$, we see that the supremum in (8) is infinite. Suppose that $z \in T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})$. Then

$$\begin{aligned} &\sup\{\langle x^*, z\rangle \mid x^* \in \partial f(\bar{x})\} \\ &\qquad \leq \sup\left\{\langle x^*, z\rangle \,\middle|\, x^* \in \partial f(\bar{x}) + \left[T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})\right]^\circ\right\} \end{aligned}$$

since $0 \in \left[T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})\right]^\circ$. However

$$\begin{aligned} &\sup\left\{\langle x^*, z\rangle \,\middle|\, x^* \in \partial f(\bar{x}) + \left[T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})\right]^\circ\right\} \\ &\qquad = \sup\left\{\langle y^* + z^*, z\rangle \,\middle|\, y^* \in \partial f(\bar{x}), z^* \in \left[T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})\right]^\circ\right\} \\ &\qquad \leq \sup\{\langle y^*, z\rangle \mid y^* \in \partial f(\bar{x})\} \end{aligned}$$

by the definition of a polar cone.

Note that (8) is equivalent to

$$\psi^*(z \mid \alpha\mathbf{B}) \leq \psi^*(z \mid \partial f(\bar{x}) + \left[T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})\right]^\circ),$$

which is equivalent to

$$\alpha\mathbf{B} \subseteq \partial f(\bar{x}) + \left[T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})\right]^\circ,$$

which establishes the result.

(b) The definitions imply that $\bar{S}$ is a set of weak sharp minima for f over S with modulus $\alpha > 0$ if and only if $\bar{S}$ is a set of weak sharp minima for the function $h(x) := f(x) + \psi(x \mid S)$ over $\mathbf{R}^n$ with modulus $\alpha > 0$. We will show that this implies (6) for every $\bar{x} \in \bar{S}$ at which f, $\bar{S}$ and S are regular.

Let $\bar{x} \in \bar{S}$ be a point at which f, $\bar{S}$, and S are regular. Since $\bar{S}$ is a set of weak sharp minima for h over $\mathbf{R}^n$ with modulus $\alpha > 0$, Theorem 2.2 implies that

$$h'(\bar{x}; d) \geq \alpha \operatorname{dist}(d \mid T(\bar{x} \mid \bar{S})) \quad \text{for all } d\,.$$

Now, by the BCQ, [20, Cor. 8.1.2], and the regularity of S, we know that

$$h'(\bar{x}; d) \leq f'(\bar{x}; d) + \psi(\cdot \mid S)'(\bar{x}; d) = f'(\bar{x}; d) + \psi(d \mid T(\bar{x} \mid S)). \tag{9}$$

Therefore,

$$f'(\bar{x}; d) \geq \alpha \operatorname{dist}(d \mid T(\bar{x} \mid \bar{S})) \quad \text{for all } d \in T(\bar{x} \mid S).$$

This last inequality implies (6) since

$$\operatorname{dist}(d \mid T(\bar{x} \mid \bar{S})) = \|d\|$$

for every $d \in N(\bar{x} \mid \bar{S})$.

(c) Since convexity implies regularity, half of this result has already been established in part (b). It remains to show that (5) holding for all $\bar{x} \in \bar{S}$ implies that $\bar{S}$ is a set of weak sharp minima for f over S with modulus α.

Let $\bar{x} \in \bar{S}$. It was shown in part (a) that the statement (5) is equivalent to the statement (6). Thus we need only show that if (6) holds for all $\bar{x} \in \bar{S}$, then $\bar{S}$ is a set of weak sharp minima for f over S with modulus α. To this end, let $x \in \mathbf{R}^n$ be given and set $\bar{x} = P(x \mid \bar{S})$. By (9) we only need consider cases where $x - \bar{x} \in T(\bar{x} \mid S)$. From the definition of projection it follows that $x - \bar{x} \in N(\bar{x} \mid \bar{S})$. Therefore, $f'(\bar{x}; x - \bar{x}) \geq \alpha \|x - \bar{x}\|$, for all x and hence $h'(\bar{x}; x - \bar{x}) \geq \alpha \operatorname{dist}(x \mid \bar{S})$, for all x. By Theorem 2.2, $\bar{S}$ is a set of weak sharp minima for f over S with modulus α. □

COROLLARY 2.7. *Suppose $f: \mathbf{R}^n \mapsto \mathbf{R}$ is differentiable and $\bar{S} \subset S$ are nonempty closed subsets of $\mathbf{R}^n$.*

(a) *The inclusion*

$$\alpha\mathbf{B} \subset \nabla f(\bar{x}) + \left[T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S})\right]^\circ \tag{10}$$

holds at $\bar{x} \in \bar{S}$ if and only if

$$\langle \nabla f(\bar{x}), z \rangle \geq \alpha \|z\| \;\; \forall \; z \in T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S}).$$

(b) *If $\bar{S}$ is a set of weak sharp minima for f over S with modulus $\alpha > 0$, then for each $\bar{x} \in \bar{S}$ at which S and $\bar{S}$ are regular, we have the inclusion* (10).

(c) *If we further assume that f is closed proper and convex and the sets $\bar{S}$ and S are nonempty closed and convex, then $\bar{S}$ is a set of weak sharp minima for f over S with modulus $\alpha > 0$ if and only if*

$$-\nabla f(\bar{x}) \in \operatorname{int} \bigcap_{x \in \bar{S}} \left[T(x \mid S) \bigcap N(x \mid \bar{S})\right]^{\circ}.$$

Remark. The corollary given above is a strengthening of [1, Prop. 2.2]. In particular, the equivalence in part (a) is proven without assumptions on convexity of S. In fact, under the convexity assumptions in part (c), the condition given in [1] is equivalent to strong uniqueness. By relaxing strong uniqueness to the assumption of a weak sharp minimum, all the results of [1, Prop. 2.2] still follow, with the exception of uniqueness.

3. Some special cases.

We now examine three important classes of convex programming problems and characterize when these problems posses weak sharp minima. The problem classes considered are linear and quadratic programming and the linear complementarity problem.

3.1. Quadratic programming.

We will use the results on weak sharp minima from §2 to obtain a necessary and sufficient condition for weak sharp minima to occur in convex quadratic programs.

The quadratic programming problem is

$$\min_{x \in S} \ \tfrac{1}{2}\langle x, Qx\rangle + \langle c, x\rangle, \tag{11}$$

where S is polyhedral and $Q \in \mathbb{R}^{n\times n}$ is symmetric and positive semidefinite. The key to our characterization of when problem (11) has weak sharp minima is the relation (6) in Theorem 2.6. To apply this result, we must first obtain a tractable description of the tangent cone to the solution set of (11). This is accomplished by using the description of the solution set of a convex program given in [3], [14].

THEOREM 3.1. *Let $\bar{S}$ be the set of solutions to the problem* $\min\{f(x) \ : \ x \in S\}$ *where both $f\colon \mathbb{R}^n \mapsto \mathbb{R}$ and $S \subset \mathbb{R}^n$ are taken to be convex and choose $\bar{x} \in \bar{S}$. Then*

$$\bar{S} = \{x \in S \mid \nabla f(x) = \nabla f(\bar{x}), \langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0\}.$$

It is clear that for convex quadratic programs this gives the solution set as

$$\bar{S} = S \bigcap \{x \mid \langle \nabla f(\bar{x}), x - \bar{x}\rangle = 0\} \bigcap \{x \mid \nabla^2 f(\bar{x})(x - \bar{x}) = 0\}$$

and since S is polyhedral

$$T(x \mid \bar{S}) = T(x \mid S) \bigcap (\nabla f(\bar{x}))^{\perp} \bigcap \ker(\nabla^2 f(\bar{x})). \tag{12}$$

Note that $\nabla f(\bar{x})$ is constant on the solution set of a convex program and $\nabla^2 f(\bar{x})$ is constant for the problem (11). In the rest of this paper, we will use the notation $\nabla f(\bar{x})$, $\nabla^2 f(\bar{x})$ for these constants and span(d), ker(A) to represent the subspace generated by d and the nullspace of the matrix A, respectively.

THEOREM 3.2. *Let $\bar{S}$ be the set of solutions to* (11) *and assume that $\bar{S}$ is nonempty. Then $\bar{S}$ is a set of weak sharp minima for f over S if and only if*

$$\left(\ker(\nabla^2 f(\bar{x}))\right)^{\perp} \subseteq \operatorname{span}(\nabla f(\bar{x})) + N(x \mid S), \quad \forall x \in \bar{S},$$

or, equivalently,

$$(\nabla f(\bar{x}))^{\perp} \bigcap T(x \mid S) \subseteq \ker(\nabla^2 f(\bar{x})), \quad \forall x \in \bar{S},$$

where $\bar{x}$ is any element of $\bar{S}$.

Proof.

($\Longleftarrow$) We show that (6) holds. Let $x \in \bar{S}$ and $d \in T(x \mid S)$. Note that (12) and the hypothesis gives

$$\begin{aligned} K := T(x \mid \bar{S})^{\circ} &= N(x \mid S) + \operatorname{span}(\nabla f(\bar{x})) + \left(\ker(\nabla^2 f(\bar{x}))\right)^{\perp} \\ &= N(x \mid S) + \operatorname{span}(\nabla f(\bar{x})). \end{aligned}$$

Therefore

$$\begin{aligned} \alpha \operatorname{dist}(d \mid T(x \mid \bar{S})) &= \alpha \psi^*(d \mid \mathbf{B} \textstyle\bigcap T(x \mid \bar{S})^{\circ}) \\ &= \alpha \sup \left\{ \langle z, d \rangle \,\middle|\, z \in \mathbf{B} \textstyle\bigcap K \right\}. \end{aligned}$$

It follows from [21, p. 65] that $K = \operatorname{span}(\nabla f(\bar{x})) + (K \bigcap (\nabla f(\bar{x}))^{\perp})$, hence, $z \in \mathbf{B} \bigcap K$ implies $z = \lambda \nabla f(\bar{x}) + y$ with $|\lambda| \leq \eta$, where

$$\eta := \begin{cases} 1/\left\|\nabla f(\bar{x})\right\|, & \text{if } \|\nabla f(\bar{x})\| \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and $y \in K \bigcap (\nabla f(\bar{x}))^{\perp}$. Therefore

$$\begin{aligned} &\alpha \operatorname{dist}(d \mid T(x \mid \bar{S})) \\ &\quad = \alpha \sup \left\{ \langle \lambda \nabla f(\bar{x}) + y, d \rangle \,\middle|\, |\lambda| \leq \eta, y \in N(x \mid S) \textstyle\bigcap (\nabla f(\bar{x}))^{\perp} \right\} \\ &\quad \leq \alpha \eta \langle \nabla f(\bar{x}), d \rangle \\ &\quad \leq \langle \nabla f(\bar{x}), d \rangle = \langle \nabla f(x), d \rangle = f'(x; d) \end{aligned}$$

as required. The last two inequalities follow since d and y are polar to each other and by choosing $\alpha \leq \|\nabla f(\bar{x})\|$ when $\nabla f(\bar{x}) \neq 0$.

($\Longrightarrow$) Suppose that for some $x \in \bar{S}$, $T(x \mid S) \bigcap (\nabla f(\bar{x}))^{\perp} \not\subseteq \ker(\nabla^2 f(\bar{x}))$. Then there exists $d \in T(x \mid S) \bigcap (\nabla f(\bar{x}))^{\perp}$ with $d \notin \ker(\nabla^2 f(\bar{x}))$. Thus from (12), $d \notin T(x \mid \bar{S})$ and so

$$\alpha \operatorname{dist}(d \mid T(x \mid \bar{S})) > 0 = \langle \nabla f(\bar{x}), d \rangle = f'(x; d),$$

which, using (6), implies that (11) does not have a weak sharp minimum. □

It is possible to illustrate the theorem by means of adapting a simple example given in [18, p. 206].

Example 3.3. The problem is

$$\begin{array}{ll} \text{minimize}_{x \in \mathbf{R}^3} & \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \\ \text{subject to} & x_i \in [a_i, b_i], i = 1, 2, 3 \end{array}$$

for given a, $b \in \mathrm{R}^3$ with $a \leq b$. We let $S = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$. Note that

$$\bar{S} = \{(P(0 \mid [a_1, b_1]), P(0 \mid [a_2, b_2]), x_3) \mid x_3 \in [a_3, b_3]\}$$

and for each $\bar{x} \in \bar{S}$, $\nabla f(\bar{x}) = (\bar{x}_1, \bar{x}_2, 0)$. Also,

$$\nabla^2 f(\bar{x}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ so that } \ker \nabla^2 f(\bar{x}) = \{0\} \times \{0\} \times \mathrm{R}.$$

Furthermore, for $\bar{x} \in S$ we have

$$T(\bar{x} \mid S) = I_1(\bar{x}_1) \times I_2(\bar{x}_2) \times I_3(\bar{x}_3),$$

where

$$I_j(x_j) = \begin{cases} [0, +\infty) & \text{if } x_j = a_j, \\ \mathrm{R} & \text{if } a_j < x_j < b_j, \\ (-\infty, 0] & \text{if } x_j = b_j. \end{cases}$$

It follows that the second equivalence of Theorem 3.2 is satisfied exactly when $0 > b_i$ or $0 < a_i$ for $i = 1, 2$; that is, the box does not straddle the x_1 or x_2 axis. This is precisely when the problem has a weak sharp minimum.

A generalization of this result that does not require the set S to be polyhedral is easily obtained. Observe that the argument given above only employs the polyhedrality of S to establish that (12) holds. However, (12) also holds under the assumption

$$\mathrm{ri}\,(S - \bar{x}) \bigcap (\nabla f(\bar{x}))^\perp \bigcap \ker(\nabla^2 f(\bar{x})) \neq \emptyset$$

(see [21, Cors. 23.8.1 and 16.4.2]), so the following result is immediate.

THEOREM 3.4. *Let $\bar{S}$ be the solution set for* (11) *where it is no longer assumed that S is polyhedral. Suppose $\bar{x} \in \bar{S}$ is such that*

$$\mathrm{ri}\,(S - \bar{x}) \bigcap (\nabla f(\bar{x}))^\perp \bigcap \ker(\nabla^2 f(\bar{x})) \neq \emptyset.$$

Then $\bar{S}$ is a set of weak sharp minima for f over S if and only if

$$\left(\ker(\nabla^2 f(\bar{x}))\right)^\perp \subseteq \mathrm{span}\,(\nabla f(\bar{x})) + N(x \mid S) \quad \forall x \in \bar{S}.$$

3.2. Linear programming. It was shown in [15] that the solution set of a linear program is a set of weak sharp minima. We show below how it can be obtained as a corollary to Theorem 3.2.

The linear programming problem is

$$\text{(13)} \qquad \underset{x \in S}{\text{minimize}} \quad \langle c, x \rangle$$

where S is polyhedral.

THEOREM 3.5. *If* (13) *has a solution, then the set of solutions is a set of weak sharp minima for this problem.*

Proof. Let $\bar{x}$ be a solution of (13). We note that for linear programming $f(x) = \langle c, x\rangle$ so that

$$\left(\ker(\nabla^2 f(\bar{x}))\right)^{\perp} = \{0\}.$$

It follows that

$$\left(\ker(\nabla^2 f(\bar{x}))\right)^{\perp} \subseteq \operatorname{span}(\nabla f(\bar{x})) + N(x \mid S) \quad \forall x \in \bar{S},$$

so by Theorem 3.2, (13) has a weak sharp minimum. □

As was done in Theorem 3.4, we can generalize this result to the case where S is not assumed to be polyhedral.

Remark. It is tempting to consider parametric results for weak sharp minima. In fact, the following example shows that this is not too fruitful. Consider the linear programs $P(i)$, for $i = 1, \ldots, \infty$, given by

$$\begin{array}{ll} \text{minimize} & x_1/i + x_2 \\ \text{subject to} & x \geq 0 \end{array}$$

Then, as shown above, each of these problems has a weak sharp minimum. However, it is easy to show that there is no constant $\alpha > 0$ that will work for all of them.

As a simple application of this result, we have the following corollary.

COROLLARY 3.6. *Suppose* $f: \mathbb{R}^n \mapsto \overline{\mathbb{R}}$ *is a proper polyhedral convex function and the problem*

$$\min_{x \in \mathbb{R}^n} f(x) \tag{14}$$

has a nonempty solution set, $\bar{S}$. *Then* $\bar{S}$ *is a set of weak sharp minima for (14).*

Proof. It follows from the defintion of a polyhedral convex function that

$$f(x) = h(x) + \psi(x \mid C),$$

where

$$h(x) := \max\{\langle x, b_1\rangle - \beta_1, \ldots, \langle x, b_k\rangle - \beta_k\}$$

and

$$C := \{x : \langle x, b_{k+1}\rangle \leq \beta_{k+1}, \ldots, \langle x, b_m\rangle \leq \beta_m\}.$$

It is clear that (14) is equivalent to

$$\begin{array}{ll} \text{minimize}_x & h(x) \\ \text{subject to} & x \in C, \end{array} \tag{15}$$

which in turn is equivalent to the linear program

$$\begin{array}{llcll} \text{minimize}_{(x,\psi)} & & & \psi & \\ \text{subject to} & \psi & \geq & \langle x, b_i\rangle - \beta_i & i = 1, \ldots, k \\ & x & \in & C & \end{array} \tag{16}$$

and that the solution set of (16) is $\bar{S} \times \{h(\bar{x})\}$ for any $\bar{x} \in \bar{S}$. Theorem 3.5 implies the existence of $\alpha > 0$ such that

$$\begin{aligned} \psi - h(\bar{x}) &\geq \alpha \operatorname{dist}((x, \psi) \mid \bar{S} \times \{h(\bar{x})\}) \\ &\geq \alpha \operatorname{dist}(x \mid \bar{S}) \end{aligned}$$

for all (x, ψ) feasible for (16). It then follows that

$$h(x) - h(\bar{x}) \geq \alpha \text{dist}\,(x \mid \bar{S})$$

for all $x \in C$ since $(x, h(x))$ is feasible for (16). Thus (15) has a weak sharp minimum as required. □

3.3. Sharpness for linear complementarity problems. We will use the analysis given previously to show that nondegenerate monotone linear complementarity problems have weak sharp minima. This was proved in [13].

The linear complementarity problem is to find an $x \geq 0$ with $Mx+q \geq 0$ satisfying $\langle x, Mx + q\rangle = 0$. To study this we consider the related optimization problem

$$\text{(17)} \qquad \begin{array}{ll} \text{minimize} & \langle x, Mx + q\rangle \\ \text{subject to} & Mx + q \geq 0, x \geq 0\ . \end{array}$$

Given any feasible point x for (17), we define the sets

$$I(x) = \{i \mid M_i x + q_i = 0\} \quad \text{and} \quad J(x) = \{j \mid x_j = 0\}\,.$$

It is clear that any solution of (17) satisfies

$$I(x) \bigcup J(x) = \{1, \ldots, n\}\,.$$

We make a convexity(monotone) assumption that M is positive semidefinite and a nondegeneracy assumption that there is a solution of (17), $\hat{x}$, which satisfies

$$I(\hat{x}) \bigcap J(\hat{x}) = \emptyset.$$

Under these assumptions, it can be shown that any other solution of (17) satisfies $I(\hat{x}) \subseteq I(x)$ and $J(\hat{x}) \subseteq J(x)$, (see for instance [13, Lemma 2.2]).

THEOREM 3.7. *The solution set of a nondegenerate monotone linear complementarity problem* (17) *is a set of weak sharp minima for the problem* (17).

Proof. Let x be any solution of (17) and let $\hat{x}$ be the nondegenerate solution. By Theorem 3.2 we must show

$$(\nabla f(\hat{x}))^{\perp} \bigcap T(x \mid S) \subseteq \ker(\nabla^2 f(\hat{x})),$$

which, for this problem, means

$$\left\langle (M + M^T)\hat{x} + q, d \right\rangle = 0, \qquad \left. \begin{array}{r} M_{I(x)} d \geq 0 \\ d_{J(x)} \geq 0 \end{array} \right\rangle \implies (M + M^T) d = 0.$$

We note that

$$\begin{aligned} 0 &= \left\langle (M + M^T)\hat{x} + q, d \right\rangle \\ &= \langle M\hat{x} + q, d\rangle + \langle \hat{x}, Md\rangle \\ &= \sum_{i \in J(\hat{x})} (M\hat{x} + q)_i d_i + \sum_{j \in I(\hat{x})} \hat{x}_j (Md)_j. \end{aligned}$$

Since $I(\hat{x}) \subseteq I(x)$ and $J(\hat{x}) \subseteq J(x)$ and $M_{I(x)} d \geq 0$ and $d_{J(x)} \geq 0$ we see that

$$\sum_{i \in J(\hat{x})} (M\hat{x} + q)_i d_i = 0 \quad \text{and} \quad \sum_{j \in I(\hat{x})} \hat{x}_j (Md)_j = 0.$$

It now follows that $d_{J(\hat{x})} = 0$ and $(Md)_{I(\hat{x})} = 0$ so that $\langle d, Md \rangle = 0$. This is equivalent to $(M + M^T)d = 0$ as required. □

Note that in this result, we asume that the related optimization problem (17) has a weak sharp minimum, as opposed to an assumption of the form

$$-M\hat{x} - q \in \operatorname{int} N(\hat{x} \mid \mathbf{R}^n_+) \tag{18}$$

as made in [1]. Using Theorem 3.7 it is easy to construct examples that are sharp in the sense given above, but do not satisfy (18).

4. Finite termination of algorithms.

In this section we study the convergence properties of algorithms for solving problems of the form

$$\begin{array}{ll} \text{minimize} & f(x) \\ x \in S & \end{array} \tag{19}$$

where it is assumed that $f: \mathbf{R}^n \mapsto \mathbf{R}$ is differentiable and S is a nonempty closed convex subset of $\mathbf{R}^n$. Under the assumption that the solution set for (19), $\bar{S}$, is a set of weak sharp minima, we will examine certain tools for identifying an element of $\bar{S}$ in a finite number of iterations. Our approach is based on the techniques developed in [6]. Consequently, we need to introduce some elementary facts concerning the face structure of convex sets.

Recall that a nonempty convex subset $\widehat{C}$ of a closed convex set C in $\mathbf{R}^n$ is said to be a face of C if every convex subset of C whose relative interior meets $\widehat{C}$ is contained in $\widehat{C}$ (e.g., see [21, §18]). In fact, the relative interiors of the faces of C form a partition of C [21, Thm. 18.2]. Thus every point $x \in C$ can be associated with a unique face of C denoted by $F(x|C)$ such that $x \in \operatorname{ri}(F(x|C))$. A face $\widehat{C}$ of C is said to be exposed if there is a vector $x^* \in \mathbf{R}^n$ such that $\widehat{C} = E(x^* \mid C)$ where

$$E(x^* \mid C) := \arg\max\{\langle x^*, y \rangle : y \in C\}.$$

The vector x^* is said to expose the face $E(x^* \mid C)$. It is well known and elementary to show that every face $\widehat{C}$ of a polyhedron is exposed and that the exposing vectors are precisely the elements of $\operatorname{ri}(N(x|C))$ for any $x \in \operatorname{ri} \widehat{C}$.

With these notions in mind, we have the following key result.

THEOREM 4.1. *If $\bar{S}$ is a set of weak sharp minima for problem* (19) *that is regular, then the set*

$$K := \bigcap_{x \in \bar{S}} [T(x \mid S) \cap N(x \mid \bar{S})]^\circ$$

has nonempty interior and for each $z \in \operatorname{int} K$ we have the inclusion $E(z \mid S) \subset \bar{S}$. If it is further assumed that the function f is convex, then $\bar{S}$ is an exposed face of S with exposing vector $-\nabla f(\bar{x})$ for any $\bar{x} \in \bar{S}$.

Proof. The fact that the set K has nonempty interior follows immediately from Corollary 2.7, in particular, $-\nabla f(\bar{x}) \in \operatorname{int} K$ for any $\bar{x} \in \bar{S}$. Let $z \in \operatorname{int} K$ and choose $\delta > 0$ so that $z + \delta \mathbf{B} \subset K$. Then for each $\bar{x} \in \bar{S}$

$$\langle z + \delta \mathbf{B}, d \rangle \leq 0 \text{ for all } d \in T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S}),$$

or, equivalently,

$$\langle z, d \rangle \leq -\delta \|d\| \text{ for all } d \in T(\bar{x} \mid S) \bigcap N(\bar{x} \mid \bar{S}).$$

Hence, given $x \in S$ and $p \in P(x \mid \bar{S})$ we have

$$\langle z, x-p\rangle \leq -\delta \|x-p\|$$

since $(x-p) \in T(p \mid S) \bigcap N(p \mid \bar{S})$. Consequently, $E(z \mid S) \subset \bar{S}$.

It only remains to show that if f is convex, then $E(-\nabla f(\bar{x}) \mid S) = \bar{S}$ for any $\bar{x} \in \bar{S}$. First observe that

$$\nabla f(x) = \nabla f(y) \text{ for every } x,\ y \in \bar{S} \tag{20}$$

by Theorem 3.1. Moreover, it has been established that $E(-\nabla f(\bar{x}) \mid S) \subset \bar{S}$. Hence the result will follow if we can show that $\langle \nabla f(x), x\rangle = \langle \nabla f(x), y\rangle$ for any choice of $x,\ y \in \bar{S}$. But this follows immediately from Theorem 3.1. □

Remark. In Theorem 4.1, the nonemptiness of the set $\operatorname{int} K$ followed from the differentiability hypothesis on f. In the absence of such a differentiability hypothesis, the result would, in general, be false. Indeed, we need only consider the case $f(x) := \operatorname{dist}(x \mid \bar{S})$, where $\bar{S}$ is any nonempty closed set and S is any set that properly contains $\bar{S}$ in its interior.

In [10], the notion of minimum principle sufficiency was introduced. Assuming $\bar{x} \in \bar{S}$, we define

$$\hat{S} := \arg\min \{\nabla f(\bar{x})x \mid x \in S\}.$$

Minimum principle sufficiency (MPS) is the equality of the sets $\hat{S}$ and $\bar{S}$. Note that in the notation above $\hat{S} \equiv E(-\nabla f(\bar{x}) \mid S)$. Solodov [22] pointed out the following interesting corollary to Theorem 4.1.

THEOREM 4.2. *Suppose f is convex and differentiable and $S \subset \mathrm{R}^n$ is a closed convex set. If $\bar{S}$ is a set of weak sharp minima for* (19) *then MPS is satisfied. If S is polyhedral, then MPS is equivalent to $\bar{S}$ being a set of weak sharp minima for* (19).

Proof. The first statement of the theorem follows from Theorem 4.1. For the second part, note that for all $x \in S$

$$\begin{aligned} f(x) - f(P(x \mid \bar{S})) &\geq \nabla f(P(x \mid \bar{S}))(x - P(x \mid \bar{S})) && \text{by convexity of } f,\\ &= \nabla f(\bar{x})(x - P(x \mid \bar{S})) && \text{by Theorem 3.1,}\\ &= \nabla f(\bar{x})(x - P(x \mid \hat{S})) && \text{by MPS,}\\ &\geq \alpha \left\| x - P(x \mid \hat{S}) \right\| && \text{by Theorem 3.5,}\\ &= \alpha \left\| x - P(x \mid \bar{S}) \right\| && \text{by MPS,} \end{aligned}$$

where $\alpha > 0$ as required. □

The following simple example shows that the assumption of polyhedrality cannot be removed in the above.

Example 4.3. The problem

$$\begin{array}{ll} \text{minimize} & x_1 \\ \text{subject to} & (x_1 - 1)^2 + x_2^2 \leq 1 \end{array}$$

has a unique solution $(0, 0)$. It is easy to see that MPS is satisfied. However, for the problem to have a weak sharp minimum would require the existence of $\alpha > 0$ such that

$$x_1 \geq \alpha\sqrt{x_1^2 + x_2^2}$$

for all feasible points x. If we consider points x on the boundary of the circle, then it follows that

$$x_1^2 \geq 2\alpha^2 x_1,$$

which is not true for x_1 sufficiently small.

Another simple application of Theorem 4.1 results in the following strong upper semicontinuity result for linear programs that was first proven in [19, Lemma 3.5].

COROLLARY 4.4. *Let S be a polyhedral convex set in R^n. Let $c \in \mathrm{R}^n$ and $\bar{S} := \arg\max_{x \in S} \langle \bar{c}, x \rangle$. Then there is a neighborhood U of $\bar{c}$ such that if $c \in U$ then*

$$\arg\max_{x \in S} \langle c, x \rangle = \arg\max_{x \in \bar{S}} \langle c, x \rangle .$$

Proof. If $\bar{S} = \emptyset$, the result follows from the fact that a polyhedral set has a finite number of faces and the graph of the subdifferential of a closed proper convex function is closed. Otherwise, it follows from Theorem 3.5 that $\bar{S}$ is a set of weak sharp minima for

$$\max_{x \in S} \langle \bar{c}, x \rangle .$$

By Theorem 4.1, it follows that $\bar{S} = E(\bar{c} \mid S)$ and that for all c in a neighborhood of $\bar{c}$ that $E(c \mid S) \subset E(\bar{c} \mid S)$. The required equality $E(c \mid S) = E(c \mid \bar{S})$ now follows easily. □

As another immediate consequence of Theorem 4.1, we obtain the following generalization of a result found in [1].

COROLLARY 4.5. *Suppose $\bar{S}$ is a set of weak sharp minima for the problem* (19) *and let $\{x^k\} \subset \mathrm{R}^n$. If either*

(a) *f is convex and $\{x^k\}$ is any sequence for which* $\operatorname{dist}(x^k \mid \bar{S}) \to 0$ *and ∇f is uniformly continuous on an open set containing $\{x^k\}$, or*

(b) *the sequence $\{x^k\}$ converges to some $\hat{x} \in \bar{S}$, ∇f is continuous and $\bar{S}$ is regular,*

then there is a positive integer k_0 such that any solution of

$$\begin{array}{ll} \text{minimize} & \langle \nabla f(x^k), x \rangle \\ x \in S & \end{array} \tag{21}$$

solves (19).

Proof. Let us first assume that (a) holds. By Theorem 2.6,

$$-\nabla f(\bar{x}) + \alpha \mathrm{B} \in \bigcap_{x \in \bar{S}} \left[T(x \mid S) \cap N(x \mid \bar{S}) \right]^\circ \tag{22}$$

for every $\bar{x} \in \bar{S}$, where $\alpha > 0$ is the modulus of weak sharp minimization for the set $\bar{S}$. Also, by Theorem 3.1, $\nabla f(x) = \nabla f(y)$ for all $x,\ y \in \bar{S}$. Consequently, the hypotheses imply the existence of an integer k_0 such that $\left\| \nabla f(x^k) - \nabla f(\bar{x}) \right\| < \alpha$ for all $k \geq k_0$. Therefore, by Theorem 4.1, $E(\nabla f(x^k) \mid S) = \bar{S}$.

If (b) holds, then (22) is still valid for every point $\bar{x} \in \bar{S}$. The result follows just as it did under assumption (a) since $\left\| \nabla f(x^k) - \nabla f(\hat{x}) \right\| \to 0$. □

The proof of this result only requires the assumption (22) to hold. Part b) of the above corollary can then be proven under the hypothesis that (22) holds only at $\hat{x}$. This is a weakening of the hypotheses that $-\nabla f(\hat{x}) \in \operatorname{int} N(\hat{x} \mid S)$ in [1, Thm. 2.1].

Assuming that we can solve (21), Corollary 4.5 can be employed to construct hybrid iterative algorithms for solving problem (19) that will terminate finitely at weak sharp minima. All that needs to be done is to solve the problem (21) occasionally and if an optima is found, then stop. However, some algorithms do not require such a "fix" to locate weak sharp minima finitely. We show that when the objective function f is convex, we can characterize those algorithms that can identify weak sharp minima finitely. We begin with a result that relates the optimality condition given in Theorem 2.6 to the structure of convex subsets of the constraint region S.

LEMMA 4.6. *Let F be any nonempty closed convex subset of the closed convex set $S \subset \mathbb{R}^n$. Then*

$$F + \bigcap_{x \in F} [T(x \mid S) \cap N(x \mid F)]^\circ \subset \bigcup_{x \in F} [x + N(x \mid S)] =: K. \tag{23}$$

Proof. Let $\bar{x} \in F$. We need only show that

$$\bar{K} := \bar{x} + \bigcap_{x \in F} [T(x \mid S) \cap N(x \mid F)]^\circ \subset K.$$

Let $y \in \bar{K}$ and let $\bar{y}$ be the projection of $P(y \mid S)$ onto F. Since $y \in \bar{K}$, there is a $z \in [T(\bar{y} \mid S) \cap N(\bar{y} \mid F)]^\circ$ such that $y = \bar{x} + z$. Hence

$$\begin{aligned} 0 &= \langle y - y, P(y \mid S) - \bar{y} \rangle \\ &= \langle P(y \mid S) + (y - P(y \mid S)) - \bar{x} - z, P(y \mid S) - \bar{y} \rangle \\ &= \langle (P(y \mid S) - \bar{y}) + (y - P(y \mid S)) + (\bar{y} - \bar{x}) - z, P(y \mid S) - \bar{y} \rangle \\ &= \|P(y \mid S) - \bar{y}\|_2^2 + \langle y - P(y \mid S), P(y \mid S) - \bar{y} \rangle \\ &\quad + \langle \bar{y} - \bar{x}, P(y \mid S) - \bar{y} \rangle + \langle -z, P(y \mid S) - \bar{y} \rangle . \end{aligned}$$

Observe that each of the terms in the final sum is nonnegative. The second term is nonnegative since $(y - P(y \mid S)) \in N(P(y \mid S) \mid S)$ and $-(P(y \mid S) - \bar{y}) \in T(P(y \mid S) \mid S)$. The third term is nonnegative since $\bar{x} - \bar{y} \in T(\bar{y} \mid F)$ while $(P(y \mid S) - \bar{y}) \in N(\bar{y} \mid F)$. Finally, the fourth term is nonnegative since $(P(y \mid S) - \bar{y}) \in [T(\bar{y} \mid S) \cap N(\bar{y} \mid F)]$. Hence each term is zero so that $\bar{y} = P(y \mid S)$; that is, $y \in \bar{y} + N(\bar{y} \mid S) \subset K$. □

Remarks. 1. It should be noted that one can easily generate examples in which the inclusion (23) is strict.

2. In the fully convex and differentiable case, it was shown in Theorem 4.1 that the set of weak sharp minima $\bar{S}$ is an exposed face of the constraint region S. Consequently, the set F in the above lemma may be taken to be the set $\bar{S}$. In this case we may write

$$K = \bigcup_{x \in \bar{S}} [F(x \mid S) + N(x \mid S)].$$

Lemma 4.6 is now employed to show that the characterization given in [6] of those algorithms that identify the optimal face of S in a finite number of steps also characterizes those algorithms that identify weak sharp minima finitely.

THEOREM 4.7. *Suppose f is convex and let $\bar{S} \subset S$ be a set of weak sharp minima for* (19). *If $\{x^k\} \subset S$ is such that* $\operatorname{dist}(x^k \mid \bar{S}) \to 0$ *and ∇f is uniformly continuous on an open set containing $\{x^k\}$, then $x^k \in \bar{S}$ for all k sufficiently large if and only if*

$$P(-\nabla f(x^k) \mid T(x^k \mid S)) \to 0. \tag{24}$$

Proof. If $x^k \in \bar{S}$ for all k sufficiently large, then $-\nabla f(x^k) \in N(x^k \mid S)$ for all k sufficiently large so that (24) holds trivially. On the other hand, suppose (24) is satisfied. The Moreau decomposition of $-\nabla f(x^k)$ yields

$$-\nabla f(x^k) = P(-\nabla f(x^k) \mid T(x^k \mid S)) + P(-\nabla f(x^k) \mid N(x^k \mid S)).$$

From Theorem 3.1, we have that ∇f is constant on $\bar{S}$. Thus for any $\bar{x} \in \bar{S}$, the hypotheses imply that

$$\left\| \nabla f(\bar{x}) + P(-\nabla f(x^k) \mid N(x^k \mid S)) \right\| \to 0,$$

so

$$\operatorname{dist}(x^k + P(-\nabla f(x^k) \mid N(x^k \mid S)) \mid \bar{S} - \nabla f(\bar{x})) \to 0.$$

However, by Theorem 2.6,

$$\bar{S} - \nabla f(\bar{x}) \subset \operatorname{int} \left[\bar{S} + \bigcap_{x \in \bar{S}} [T(x \mid S) \cap N(x \mid \bar{S})]^\circ \right].$$

Thus Lemma 4.6 implies that

$$\begin{aligned} x^k + P(-\nabla f(x^k) \mid N(x^k \mid S)) &\in \operatorname{int} \left[\bar{S} + \bigcap_{x \in \bar{S}} [T(x \mid S) \cap N(x \mid \bar{S})]^\circ \right] \\ &\subset \bigcup_{x \in \bar{S}} [x + N(x \mid S)] \end{aligned}$$

for all k sufficiently large. Therefore,

$$\begin{aligned} x^k &= P\left(x^k + P(-\nabla f(x^k) \mid N(x^k \mid S)) \middle| S\right) \\ &\in P\left(\bigcup_{x \in \bar{S}} [x + N(x \mid S)] \middle| S \right) \\ &\subset \bigcup_{x \in \bar{S}} \{x\} \\ &= \bar{S} \end{aligned}$$

for all k sufficiently large. □

In [6], it was shown that the condition (24) is simple to check in certain cases. In particular, it was established that the standard sequential quadratic programming method and the gradient projection method both satisfy (24) and so will automatically generate sequences that terminate finitely at weak sharp minima. We should also note that Polyak[18, Exer. 2, p. 209] indicates that the gradient projection method terminates finitely at weak sharp minima.

REFERENCES

[1] F. A. Al-Khayyal and J. Kyparisis, *Finite convergence of algorithms for nonlinear programs and variational inequalities*, J. Optim. Theory Appl., 70 (1991), pp. 319–332.

[2] L. N. H. Bunt, *Bijdrage tot de Theorie der Convexe Puntverzamelingen*, Thesis, University of Groningen, Groningen, the Netherlands, 1934.

[3] J. V. Burke and M. C. Ferris, *Characterization of solution sets of convex programs*, Oper. Res. Lett., 10 (1991), pp. 57–60.

[4] J. V. Burke, M. C. Ferris, and M. Qian, *On the Clarke subdifferential of the distance function to a closed set*, J. Math. Anal. Appl., 166 (1992), pp. 199–213.

[5] J. V. Burke and S. P. Han, *A Gauss–Newton approach to solving generalized inequalities*, Math. Oper. Res., 11 (1986), pp. 632–643.

[6] J. V. Burke and J. J. Moré, *On the identification of active constraints*, SIAM J. Numer. Anal., 25 (1988), pp. 1197–1211.

[7] F. H. Clarke, *Optimization and Nonsmooth Analysis.* John Wiley, New York, 1983.

[8] L. Cromme, *Strong uniqueness*, Numer. Math., 29 (1978), pp. 179–193.

[9] M. C. Ferris, *Weak sharp minima and penalty functions in mathematical programming*, Tech. Report 779, Computer Sciences Department, University of Wisconsin, Madison, Wisconsin, June 1988.

[10] M. C. Ferris and O. L. Mangasarian, *Minimum principle sufficiency*, Math. Programming, 57 (1992), pp. 1–14.

[11] R. Hettich, *A review of numerical methods for semi–infinite optimization*, in Semi–Infinite Programming and Applciations, A. V. Fiacco and K. O. Kortanek, eds., Springer–Verlag, Berlin, 1983.

[12] K. Madsen, *Minimization of Non-Linear Approximation Functions*, Dr. Techn. Thesis, Istititute for Numerical Analysis, The Technical University of Denmark, Lyngby, Denmark, 1985.

[13] O. L. Mangasarian, *Error bounds for nondegenerate monotone linear complementarity problems*, Math. Programming, 48 (1990), pp. 437–446.

[14] ———, *A simple characterization of solution sets of convex programs*, Oper. Res. Lett., 7 (1988), pp. 21–26.

[15] O. L. Mangasarian and R. R. Meyer, *Nonlinear perturbation of linear programs*, SIAM J. Control Optim., 17 (1979), pp. 745–752.

[16] T. S. Motzkin, *Sur quelque propriétés caractéristiques des ensembles convexes*, Rend. Accad. Naz. Lincei, 21 (1935), pp. 562–567.

[17] B. T. Polyak, *Sharp Minima*, Institute of Control Sciences Lecture Notes, Moscow, USSR, 1979; Presented at the IIASA Workshop on Generalized Lagrangians and Their Applications, IIASA, Laxenburg, Austria, 1979.

[18] ———, *Introduction to Optimization*, Optimization Software, Inc., Publications Division, New York, 1987.

[19] S. M. Robinson, *Local structure of feasible sets in nonlinear programming, Part* II: *nondegeneracy*, Math. Programming Stud., 22 (1984), pp. 217–230.

[20] R. T. Rockafellar, *Extensions of subgradient calculus with applications to optimization*, Nonlinear Anal., 9 (1985), pp. 665–698.

[21] ———, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[22] M. V. Solodov, Private communication, 1992.