Characterizations of Strong Regularity for Variational Inequalities over Polyhedral Convex Sets*

A. L. Dontchev
Mathematical Reviews, Ann Arbor, MI 48107
and
R. T. Rockafellar
Dept. of Math., Univ. of Washington, Seattle, WA 98195

Revised version: August 17, 1995
(Accepted by the SIAM Journal on Optimization)

Abstract

Linear and nonlinear variational inequality problems over a polyhedral convex set are analyzed parametrically. Robinson's notion of strong regularity, as a criterion for the solution set to be a singleton depending Lipschitz continuously on the parameters, is characterized in terms of a new "critical face" condition and in other ways. The consequences for complementarity problems are worked out as a special case. Application is also made to standard nonlinear programming problems with parameters that include the canonical perturbations. In that framework a new characterization of strong regularity is obtained for the variational inequality associated with the Karush-Kuhn-Tucker conditions.

*This work was supported by National Science Foundation grants DMS 9404431 for the first author and DMS 9500957 for the second.

1. Introduction

For a map $f : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^n$ and a nonempty, polyhedral, convex set $C \subset \mathbb{R}^n$, we study the variational inequality problem in which a point x is sought such that

$$x \in C \ \text{ and } \ \langle z + f(w,x),\, x' - x\rangle \geq 0 \ \text{ for all } \ x' \in C.$$

(Here $\langle \cdot, \cdot \rangle$ refers to the scalar product of two vectors.) This problem is viewed as depending on $w \in \mathbb{R}^d$ and $z \in \mathbb{R}^n$ as parameter vectors (with z representing the "canonical perturbations"); we put them together as $p = (z, w)$. In terms of the normal cone $N_C(x)$ to C at x in convex analysis, which is given by

$$N_C(x) = \begin{cases} \{v \in \mathbb{R}^n \,|\, \langle v,\, x' - x\rangle \leq 0 \ \text{ for all } \ x' \in C\} & \text{if } x \in C, \\ \emptyset & \text{if } x \notin C, \end{cases}$$

the targeted variational inequality can be expressed conveniently as

$$z + f(w,x) + N_C(x) \ \ni \ 0. \tag{1}$$

For each $p = (z, w) \in \mathbb{R}^n \times \mathbb{R}^d$ let $S(p)$ be the (possibly empty) set of solutions x of (1). We concern ourselves with the local behavior of the map S around a fixed element $p_0 = (z_0, w_0)$ and a point $x_0 \in S(p_0)$. Specifically we are interested in the circumstances under which S is locally single-valued and Lipschitz continuous around (p_0, x_0), in the sense that there exist neighborhoods U of x_0 and V of p_0 such that the map $p \mapsto S(p) \cap U$ is single-valued and Lipschitz continuous relative to $p \in V$. In addressing this we assume here that:

(A) *f is differentiable with respect to x with Jacobian matrix $\nabla_x f(w, x)$ depending continuously on (w, x) in a neighborhood of (w_0, x_0);*

(B) *f is Lipschitz continuous in w uniformly in x around (w_0, x_0); that is, there exist neighborhoods U of x_0 and V of w_0 and a number $l > 0$ such that $\|f(w_1, x) - f(w_2, x)\| \leq l\|w_1 - w_2\|$ for all $x \in U$ and $w_1, w_2 \in V$.*

It has long been known, thanks to Robinson, that the analysis of S is closely tied to the linear variational inequality

$$q + Ax + N_C(x) \ \ni \ 0 \tag{2}$$

with canonical parameter vector q around q_0 in the case of

$$A = \nabla_x f(w_0, x_0), \qquad q_0 = z_0 + f(w_0, x_0) - \nabla_x f(w_0, x_0)x_0, \tag{3}$$

which serves to "linearize" (1) at (p_0, x_0). Let $L(q)$ denote the set of solutions x to (2). Then $x_0 \in L(q_0)$ by (3). In a landmark paper [23], Robinson proved that the solution map S for (1) is locally single-valued and Lipschitz continuous around (p_0, x_0) when the solution map L for (2) is locally single-valued and Lipschitz continuous around (q_0, x_0). Robinson called this property of L under (3) the *strong regularity* of the variational inequality (1) at (p_0, x_0). (His framework in [23] was somewhat broader than the one adopted here: C did not have to be polyhedral, and w could range over a parameter space other than $\mathbb{R}^d$; for subsequent extensions in such a mode see Robinson [25] and Dontchev and Hager [6].)

The strong regularity of (1) at (p_0, x_0) is identical by definition to the strong regularity of (2) at (q_0, x_0) under the choice of elements in (3). Our goal here is to characterize this strong regularity by a certain *critical face condition* on A and the closed faces of the critical cone K_0 consisting of the vectors in the tangent cone $T_C(x_0)$ to C at x_0 that are orthogonal to the normal vector

$$v_0 = -Ax_0 - q_0 \in N_C(x_0). \tag{4}$$

We further provide characterizations through a localized Lipschitz condition on L at (q_0, x_0) which we call the Aubin property, and also through the lower semicontinuity of L around (q_0, x_0).

The result that the lower semicontinuity of L around (q_0, x_0) thereby entails the local single-valuedness and Lipschitz continuity of L can be compared with the well known fact that a monotone map has to be single-valued and continuous wherever it is lower semicontinuous; we would have L monotone if the matrix A in (2) were monotone (i.e., positive semidefinite, not necessarily symmetric), but such monotonicity is not assumed. We do not know whether the lower semicontinuity of S around (p_0, x_0) likewise ensures the local single-valuedness and Lipschitz continuity of S around (p_0, x_0), but we verify that the Aubin property of S at (p_0, x_0) does yield it.

Throughout we denote by $\mathbb{B}_r(x)$ the closed ball centered at x with radius r and by $\mathbb{B}$ the closed unit ball. For a (potentially set-valued) map Γ from $\mathbb{R}^m$ to $\mathbb{R}^n$ we denote by $\operatorname{gph}\Gamma$ the graph of Γ, i.e., the set $\{(u, x) \,|\, u \in \mathbb{R}^m,\ x \in \Gamma(u)\}$. Recall that Γ is called lower semicontinuous *at* the pair

$(u_0, x_0) \in \operatorname{gph}\Gamma$ if for every sequence $u_i \to u_0$ there exists a sequence $x_i \to x_0$ such that $x_i \in \Gamma(u_i)$ for all $i = 1, 2, \ldots$ sufficiently high. If Γ is lower semicontinuous at every point $(u, x) \in \operatorname{gph}\Gamma$ with u belonging to an open set U, it is said to be lower semicontinuous on U. We say that Γ is lower semicontinuous *around* $(u_0, x_0) \in \operatorname{gph}\Gamma$ if there exists a neighborhood W of (u_0, x_0) such that Γ is lower semicontinuous at every $(u, x) \in \operatorname{gph}\Gamma \cap W$. We also employ the following concept of Aubin [1].

Definition 1. *A set-valued map Γ from $I\!R^m$ to the subsets of $I\!R^n$ has the* Aubin property *at $(u_0, x_0) \in \operatorname{gph}\Gamma$ with a constant M if there exist neighborhoods U of u_0 and V of x_0 such that*

$$\Gamma(u_1) \cap V \ \subset \ \Gamma(u_2) + M\|u_1 - u_2\| I\!B \ \textit{ for all } \ u_1, u_2 \in U.$$

Aubin himself referred to this as "pseudo-Lipschitz continuity." In actuality it is a fundamental property more important in general than Lipschitz continuity as usually interpreted with the Hausdorff metric: it readily characterizes the latter when Γ is locally bounded (see Rockafellar [27]), but it makes better sense in most cases when Γ is not locally bounded or in particular has unbounded images $\Gamma(u)$—all of which jars with the connotation of "pseudo" as "false." We prefer therefore to call this concept the Aubin property, giving credit where credit is due. This property of Γ is equivalent to Γ^{-1} having a "linear rate of openness" (hence providing a link to open map theorems) as well as to Γ^{-1} being metrically regular (a basic condition employed in the stability analysis of optimization problems); see Borwein and Zhuang [3] and Penot [21].

When C is the nonnegative orthant $I\!R^n_+$, the variational inequality (1) corresponds to the *complementarity* problem while (2) gives the *linear complementarity* problem, which seeks an $x \in I\!R^n$ such that

$$Ax + q \geq 0, \quad x \geq 0, \quad \langle x,\, Ax + q\rangle = 0. \tag{5}$$

The initial motivation for our efforts came from this case and the results that had been obtained for its solution map L_0, assigning to each q the set of all x that satisfy (5), if any.

Samelson, Thrall and Wesler [29] showed that L_0 is single-valued everywhere on $I\!R^n$ if and only if A is a P-matrix; that is, every principal minor of A has positive sign. Alternative descriptions of P-matrices have been provided in [4] and [19]. Mangasarian and Shiau [16] proved that when L_0 is single-valued everywhere it is automatically Lipschitz continuous everywhere as well. Gowda [8] proved that if $L_0(q) = \{0\}$ for some positive q, then the Lipschitz continuity of L_0 everywhere in the sense of the Hausdorff metric guarantees that L_0 is single-valued everywhere; see also Pang [20]. Murthy, Parthasarathy and Sabatini [18] dropped the requirement that $L_0(q) = \{0\}$ for some positive q, obtaining that L_0 is Lipschitz continuous everywhere if and only if it is single-valued everywhere. Gowda and Sznajder [9] noted that this result can be extended to the map L by using some recently discovered properties of normal maps. Such results got us interested in investigating also the question of local single-valuedness versus local Lipschitz continuity and how these properties could better be understood in relation to each other, and this led to the developments presented here. A product of this study, back on the global level, turns out to be that the lower semicontinuity of L everywhere on $I\!R^n$ is already enough to guarantee the single-valuedness of L everywhere.

We begin in Section 2 by considering the linear variational inequality (2) and establishing the equivalences that have been mentioned for its solution map L. In obtaining our critical face condition, a key step is the application to L of Mordukhovich's coderivative criterion in [17] for the Aubin property to hold. In Section 3 we return to the nonlinear variational inequality (1), putting the preceding results together and furnishing along the way an independent proof of Robinson's theorem, not based on a fixed point argument but utilizing instead our identification of strong regularity with the Aubin property, and noting further that strong regularity is not just sufficient but necessary for the local single-valuedness and Lipschitz continuity of S when the canonical perturbations z are present along with the general parameter element w.

As applications in Sections 4 and 5 we characterize strong regularity in the complementarity problem and for the variational inequality representing the first-order optimality conditions in a nonlinear programming problem. In the latter case we demonstrate that the combination of linear independence of the active constraint gradients with the strong second-order sufficient condition for local optimality is necessary as well as sufficient (in the presence of the canonical perturbations) for the Karush-Kuhn-Tucker map to be not just locally single-valued and Lipschitz continuous but such that its primal components are locally optimal solutions.

2. Strong Regularity and the Linear Problem.

As noted, the strong regularity of the nonlinear variational inequality (1) at (p_0, x_0) is the same as the strong regularity of the linear variational inequality (2) at (q_0, x_0) under the choice of elements in (3). Our plan of characterizing the strong regularity property in various ways can therefore be executed entirely in the linear context. Eventually the case of interest will be the one given by (3), but for now (q_0, x_0) can be any pair belonging to the graph of the solution map L for (2).

A key part in our investigation will be played by a lemma which reduces the linear variational inequality (2) over a polyhedral convex set C to a variational inequality over a polyhedral convex *cone* K. To formulate it, we introduce for each $x \in C$ and normal vector $v \in N_C(x)$ the cone

$$K(x, v) = \{x' \in T_C(x) \mid x' \perp v\}, \tag{6}$$

where $T_C(x)$ denotes the tangent cone to C at x. Here $T_C(x)$ is polyhedral convex because C is polyhedral convex, and this ensures that $K(x, v)$ is polyhedral convex too. Our interest will center especially on the *critical cone* associated with (2) for (q_0, x_0), which is

$$K_0 = K(x_0, v_0) \quad \text{for} \quad v_0 = -Ax_0 - q_0. \tag{7}$$

Reduction Lemma. *For any $(x, v) \in G = \operatorname{gph} N_C$ there is a neighborhood U of $(0, 0)$ in $\mathbb{R}^n \times \mathbb{R}^n$ such that for $(x', v') \in U$ one has*

$$v + v' \in N_C(x + x') \iff v' \in N_{K(x,v)}(x').$$

In particular, the tangent cone $T_G(x, v)$ to G at (x, v) is $\operatorname{gph} N_{K(x,v)}$.

Proof. This is a particular case of Lemma 3.5 in Robinson [24]; see also Theorem 5.6 in Rockafellar [28]. □

In the following theorem we show that the Aubin property of the map L is equivalent to the the strong regularity of the variational inequality (2). The key steps in the proof are the Reduction Lemma and a combination of some recently obtained characterizations of normal maps.

Theorem 1. *The following are equivalent:*

(i) *L is lower semicontinuous around (q_0, x_0);*

(ii) *L has the Aubin property at (q_0, x_0);*

(iii) *L is locally single-valued and Lipschitz continuous around* (q_0, x_0)*;*
(iv) *the linear variational inequality (2) is strongly regular at* (q_0, x_0).

Proof. Obviously, (iv) $\Leftrightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i). It will suffice therefore to show that (i) $\Rightarrow$ (iii). For the critical cone K_0 in (7) consider the variational inequality

$$q' + Ax' + N_{K_0}(x') \ni 0 \tag{8}$$

and denote its solution map by L': for each q', $L'(q')$ is the set of all x' satisfying (8). The Reduction Lemma tells us that as long as (x', v') is near enough to $(0, 0)$, we have $v_0 + v' \in N_C(x_0 + x')$ if and only if $v' \in N_{K_0}(x')$. Thus, $(q_0 + q') + A(x_0 + x') + N_C(x_0 + x') \ni 0$ if and only if $q' + Ax' + N_{K_0}(x') \ni 0$. In the shifted notation $x = x_0 + x'$ and $q = q_0 + q'$, therefore, we have $x \in L(q)$ if and only if $x' \in L'(q')$. In particular $0 \in L'(0)$, and the lower semicontinuity of L around (q_0, x_0) in (i) reduces to that of L' around $(0, 0)$. But L' is positively homogeneous by virtue of K_0 being a cone, so the lower semicontinuity of L' around $(0, 0)$ implies the lower semicontinuity and nonempty-valuedness of L' on all of $\mathbb{R}^n$. Our task comes down to proving that this implies L' is locally single-valued and Lipschitz continuous around $(0, 0)$ (and hence by positive homogeneity has these properties globally).

Let Π_{K_0} be the projection map onto K_0. We have

$$u - \Pi_{K_0}(u) \in N_{K_0}(\Pi_{K_0}(u)) \text{ for all } u.$$

Let h be the normal map associated with (8), namely

$$h(u) = [u - \Pi_{K_0}(u)] + A\Pi_{K_0}(u).$$

As a step toward applying known results of the theory of normal maps, we prove next that h is an open map: it maps open sets into open sets.

For this purpose fix any open set $O \subset \mathbb{R}^n$ and any point $h(u)$ with $u \in O$; it will be expedient to take $q' = -h(u)$. Consider any sequence $q'_i \to q'$ as $i \to \infty$. By demonstrating the existence of a sequence $u_i \to u$ with $q'_i = -h(u_i)$, we will confirm that eventually $-q'_i \in h(O)$ and therefore that $h(O)$ is open. From the definitions of L' and h along with the choice of q' we have for $x' = \Pi_{K_0}(u)$ that $-q' = [u - x'] + Ax'$ with $u - x' \in N_{K_0}(x')$, hence $x' \in L'(q')$ and $x' - q' - Ax' = u$. The nonempty-valuedness and lower semicontinuity of L' on $\mathbb{R}^n$ implies the existence of points $x'_i \in L'(q'_i)$ (for i sufficiently large) with $x'_i \to x'$. Since $x'_i \in L'(q'_i)$ we have $-q'_i - Ax'_i \in N_{K_0}(x'_i)$

and consequently for the points $u_i = x'_i - q'_i - Ax'_i$ that $\Pi_{K_0}(u_i) = x'_i$. But $x'_i - q'_i - Ax'_i \to x' - q - Ax'$, so we have $u_i \to u$ as demanded.

The rest of the proof is based on combining two known facts. The first is that a piecewise affine map (here h fits this category because Π_{K_0} is piecewise linear) is open if and only if it is coherently oriented; see Eaves and Rothblum [7], Lemma 6.12, and Scholtes [30], Proposition 2.3.7. The second is that the normal map corresponding to a linear variational inequality over a polyhedral convex set is coherently oriented if and only if it is one-to-one; see Robinson [26], Theorem 4.3, and also Ralph [22]. From these facts we deduce that h^{-1} is single-valued and Lipschitz continuous everywhere. The equivalence

$$x' \in L'(q') \iff x' = \Pi_{K_0}(h^{-1}(q'))$$

implies then that L' is single-valued and Lipschitz continuous everywhere. Thus we have arrived at (iii), the goal we had set out for. □

Remark 1. Corresponding to Theorem 1 on the global level is the fact that the following are equivalent:

(i) L is lower semicontinuous on $\mathbb{R}^n$;

(ii) $L(q)$ is a singleton set for every $q \in \mathbb{R}^n$.

This is easily derivable from known literature with a little help from the argument we have used in proving Theorem 1. Assuming (i), denote by h_C the normal map associated with (2). Tracing the argument in the proof of Theorem 1 but with h replaced by h_C and L' replaced by L, and relying on the references cited there, we obtain that h_C is open everywhere and consequently that L is single-valued everywhere. Conversely, under (ii) the map h_C is a homeomorphism, hence it is Lipschitz continuous everywhere. Then L is Lipschitz continuous and in particular lower semicontinuous everywhere. □

We proceed now toward our critical face condition. Recall that the closed faces F of any polyhedral convex cone K are the polyhedral convex cones of the form

$$F = \{x \in K \mid x \perp v\} \text{ for some } v \in K^*, \tag{9}$$

where K^* denotes the polar of K. The largest of these faces is K itself, while the smallest is $K \cap (-K)$, this being the maximal subspace of $\mathbb{R}^n$ included within K. Recall too that

$$v' \in N_K(x') \quad\iff\quad x' \in K, \quad v' \in K^*, \quad x' \perp v'. \tag{10}$$

Definition 2. *The* critical face condition *will be said to hold at* (q_0, x_0) *if for all choices of closed faces* F_1 *and* F_2 *of the critical cone* K_0 *with* $F_1 \supset F_2$,

$$u \in F_1 - F_2, \;\; A^\top u \in (F_1 - F_2)^* \quad \Longrightarrow \quad u = 0$$

(where $A^\top$ *denotes the transpose of* A*).*

Remark 2. When the critical cone K_0 happens to be a subspace, it has a unique closed face (namely itself). The critical face condition reduces then to a nonsingularity condition for A relative to this subspace:

$$u \in K_0, \;\; A^\top u \perp K_0 \quad \Longrightarrow \quad u = 0.$$

Theorem 2. *The solution map* L *for the linear variational inequality (2) has the Aubin property at* (q_0, x_0)*, and therefore the other equivalent properties of Theorem 1 as well, if and only if the critical face condition holds at* (q_0, x_0)*.*

Proof. According to the powerful criterion developed by Mordukhovich [17], we know that a necessary and sufficient condition for the Aubin property to hold for L at (q_0, x_0) is

$$A^\top u + D^* N_C(x_0|v_0)(u) \ni 0 \quad \Longrightarrow \quad u = 0,$$

where the map $D^* N_C(x_0|v_0)$ is the coderivative of the map N_C at the point (x_0, v_0) of $G = \operatorname{gph} N_C$. By definition, the graph of this coderivative map consists of all the pairs $(-u, r)$ such that $(r, u) \in \widetilde{N}_G(x_0, v_0)$, where $\widetilde{N}_G(x_0, v_0)$ is the generalized cone of normals to the (nonconvex) set G that is described below. In these terms the Mordukhovich criterion takes the form:

$$(A^\top u, u) \in \widetilde{N}_G(x_0, v_0) \quad \Longrightarrow \quad u = 0. \tag{11}$$

Everything hinges therefore on determining $\widetilde{N}_G(x_0, v_0)$.

In general, $\widetilde{N}_G(x_0, v_0)$ is defined as the "lim sup" of polar cones $T_G(x, v)^*$ as $(x, v) \to (x_0, v_0)$ in G, but because G is the union of finitely many polyhedral sets in $\mathbb{R}^{2n}$ (due to C being polyhedral), only finitely many cones can be manifested as $T_G(x, v)$ at points $(x, v) \in G$ near (x_0, v_0). Thus, we have for any sufficiently small neighborhood U of (x_0, v_0) that

$$\widetilde{N}_G(x_0, v_0) = \bigcup_{(x,v) \in U \cap G} T_G(x, v)^*. \tag{12}$$

Next we utilize the Reduction Lemma: we have $T_G(x,v) = \operatorname{gph} N_{K(x,v)}$, and therefore by (10) as applied to $K = K(x,v)$ that

$$T_G(x,v) = \{(x',v') \,|\, x' \in K(x,v),\ v' \in K(x,v)^*,\ x' \perp v'\}.$$

It follows that

$$\begin{aligned} T_G(x,v)^* &= \{(r,u) \,|\, \langle (r,u),(x',v')\rangle \le 0 \text{ for all } (x',v') \in T_G(x,v)\} \\ &= \{(r,u) \,|\, \langle r,x'\rangle + \langle u,v'\rangle \le 0 \text{ for all} \\ &\qquad\qquad x' \in K(x,v),\ v' \in K(x,v)^* \text{ with } x' \perp v'\}. \end{aligned}$$

It is evident from this (first in considering $v' = 0$, then in considering $x' = 0$) that actually

$$T_G(x,v)^* = K(x,v)^* \times K(x,v). \tag{13}$$

Hence $\widetilde{N}_G(x_0,v_0)$ is the union of all product sets $K^* \times K$ associated with cones K such that $K = K(x,v)$ for some $(x,v) \in G$ near enough to (x_0,v_0).

We claim now that the cones K arising in this manner are precisely the cones of the form $F_1 - F_2$ where F_1 and F_2 are closed faces of $K_0 = K(x_0,v_0)$ satisfying $F_1 \supset F_2$. This will be enough to prove the theorem by way of (11), (12) and (13).

For any vector $v \in I\!R^n$, let $[v] = \{\tau v \,|\, \tau \in I\!R\}$. Of course, this is a subspace of dimension 1 if $v \neq 0$, but just $\{0\}$ if $v = 0$. Accordingly, $[v]^\perp$ is a hyperplane through the origin if $v \neq 0$, but $[v]^\perp = I\!R^n$ if $v = 0$.

Because C is polyhedral, we know that for $x \in C$ sufficiently near to x_0 we are sure to have

$$\begin{aligned} T_C(x) = T_C(x_0) + [x - x_0] \supset T_C(x_0), \\ N_C(x) = N_C(x_0) \cap [x - x_0]^\perp \subset N_C(x_0). \end{aligned}$$

Furthermore, the vectors $x - x_0$ for $x \in C$ sufficiently near to x_0 are the vectors $x' \in T_C(x_0)$ having sufficiently small norm. On the other hand the cones of form $T_C(x_0) \cap [v]^\perp$ for $v \in N_C(x_0)$ are the closed faces of $T_C(x_0)$, while the "lim sup" of $T_C(x_0) \cap [v]^\perp$ as $v \to v_0$ with $v \in N_C(x_0)$ is included within $T_C(x_0) \cap [v_0]^\perp$. Since $T_C(x_0)$ has only finitely many closed faces, we must have $T_C(x_0) \cap [v]^\perp \subset T_C(x_0) \cap [v_0]^\perp$ for $v \in N_C(x_0)$ sufficiently close to v_0. Since the critical cone $K_0 = T_C(x_0) \cap [v_0]^\perp$ is itself a closed face of $T_C(x_0)$, any closed face of $T_C(x_0)$ within K_0 is also a closed face of K_0.

In light of this, the cones $K = K(v,x)$ at points $(x,v) \in G$ arbitrarily near to (x_0, v_0) are the cones having the form

$$K = (T_C(x_0) + [x']) \cap [v]^\perp \text{ for some } x' \in T_C(x_0) \cap [v_0]^\perp \text{ and } v \in N_C(x_0) \cap [x']^\perp$$

with v sufficiently close to v_0 and x' sufficiently close to 0 (with $x' = x - x_0$). We can write $K = (T_C(x_0) \cap [v]^\perp) + [x']$ equally well, because $x' \perp v$.

If K has this form, let $F_1 = T_C(x_0) \cap [v]^\perp$, this being a closed face of the polyhedral cone K_0 for reasons already given. We have $x' \in F_1$ and therefore actually $K = F_1 - F_2$, where F_2 is the closed face of F_1 having x' in its relative interior. Then F_2 is also a closed face of K_0, and the desired representation of K is achieved.

Conversely, if $K = F_1 - F_2$ for closed faces F_1 and F_2 of K_0 with $F_1 \supset F_2$, there must be a vector $v \in N_C(x_0)$ with $T_C(x_0) \cap [v]^\perp = F_1$. Then F_2 is a closed face of F_1. Let $x' \in \operatorname{ri} F_2$; in particular $x' \in T_C(x_0)$, so by taking the norm of x' sufficiently small we can arrange that the point $x = x_0 + x'$ lies in C. We have $x' \perp v$ and

$$\begin{aligned} F_1 - F_2 &= (T_C(x_0) \cap [v]^\perp) + [x'] = (T_C(x_0) + [x']) \cap [v]^\perp \\ &= (T_C(x_0) + [x - x_0]) \cap [v]^\perp = T_C(x) \cap [v]^\perp, \end{aligned}$$

which is the form required. □

Corollary 1. *A sufficient condition for the Aubin property to hold for L at (q_0, x_0), and therefore all the other equivalent properties in Theorem 1 as well, is that $\langle u, Au\rangle > 0$ for all vectors $u \neq 0$ in the subspace $K_0 - K_0$ spanned by the critical cone K_0.*

Proof. The inequality $\langle u, Au\rangle \leq 0$ is equivalent to $\langle u, A^\top u\rangle \leq 0$, which must hold in particular when u belongs to a cone $F_1 - F_2 \subset K_0 - K_0$ and $A^\top u \in (F_1 - F_2)^*$. In the circumstances described, this is impossible unless $u = 0$. □

Remark 3. In consequence of Theorem 2, the critical face condition is both necessary and sufficient for the coherent orientation of the normal map associated with the linear variational inequality (2).

3. The Nonlinear Problem

Now we extend our results to the nonlinear variational inequality (1), taking the linear variational inequality (2) to be its linearization as indicated by (3). We start by recording a background fact about our underlying assumption (A) in Section 1.

Strict Differentiability Lemma. *Under (A) there exist for any $\varepsilon > 0$ neighborhoods U of x_0 and V of w_0 such that, for all $x_1, x_2 \in U$ and $w \in V$,*

$$\|f(w,x_1) - f(w,x_2) - \nabla_x f(w_0,x_0)(x_1 - x_2)\| \;\le\; \varepsilon\|x_1 - x_2\|.$$

Proof. This is classical, but we supply the proof for completeness. For an arbitrary $e \in I\!R^n$ with $\|e\| = 1$ and any x_1, $x_2 \in I\!R^n$ and $w \in I\!R^d$ we can apply the mean value theorem to $\varphi(t) = \langle e, f(w, tx_1 + (1-t)x_2)\rangle$ to get a value $\tau \in (0,1)$ such that $\varphi(1) - \varphi(0) = \varphi'(\tau)$, i.e.,

$$\langle e, f(w,x_1)\rangle - \langle e, f(w,x_2)\rangle = \langle e, \nabla_x f(w, \tau x_1 + (1-\tau)x_2)(x_1 - x_2)\rangle.$$

Choose neighborhoods U of x_0 and V of w_0 such that U is convex and $\|\nabla_x f(w,x) - \nabla_x f(w_0,x_0)\| \le \varepsilon$ when $x \in U$ and $w \in V$, as is possible by virtue of the continuity of $\nabla_x f(w,x)$ in w and x that is assumed in (A). For all $x_1, x_2 \in U$ and $w \in V$ we have

$$\begin{aligned}
&\langle e,\, [f(w,x_1)\rangle - f(w,x_2) - \nabla_x f(w_0,x_0)(x_1 - x_2)]\,\rangle \\
&\quad = \langle e,\, [\nabla_x f(w, \tau x_1 + (1-\tau)x_2) - \nabla_x f(w_0,x_0)](x_1 - x_2)\,\rangle \\
&\quad \le \|\nabla_x f(w, \tau x_1 + (1-\tau)x_2) - \nabla_x f(w_0,x_0)\|\|x_1 - x_2\| \\
&\qquad\qquad \le \varepsilon\|x_1 - x_2\|.
\end{aligned}$$

This being true for all e with $\|e\| = 1$, we get the required estimate. □

Proposition 1. *The following are equivalent for the maps L and S:*
(i) *L has the Aubin property at (q_0, x_0);*
(ii) *S has the Aubin property at (p_0, x_0).*

Proof. This can be obtained at once from the observation that the Mordukhovich coderivative criterion for the Aubin property to hold, as invoked in the proof of Theorem 2, has the same form for (1) at (p_0, x_0) that it has for (2) at (q_0, x_0), because the coderivative map associated with S is the same as for L. But we proceed anyway with an independent proof which shows how this equivalence extends beyond such a framework; cf. Remark 4 below.

Let L have the Aubin property at (q_0, x_0) with a constant M; that is, for some $a > 0$ and $b > 0$ and for every $q', q'' \in \mathbb{B}_b(q_0)$ we have

$$L(q') \cap \mathbb{B}_a(x_0) \subset L(q'') + M\|q' - q''\|\mathbb{B}. \tag{14}$$

Let $\varepsilon > 0$ be such that $M\varepsilon < 1$. Choose $\alpha > 0$ and $\beta_1 > 0$ with

$$\alpha < \min\{a, b/\varepsilon\}$$

such that the inequality in the Strict Differentiability Lemma holds whenever $x', x'' \in \mathbb{B}_\alpha(x_0)$ and $w \in \mathbb{B}_{\beta_1}(w_0)$. Let $\beta > 0$ be such that

$$\beta \leq \min\left\{\beta_1, \frac{\alpha(1-\varepsilon M)}{4M(1+l)}, \frac{b-\varepsilon\alpha}{1+l}\right\}. \tag{15}$$

It will be demonstrated that S has the Aubin property at (p_0, x_0) with constant $M' = [M(l+1)/[1-\varepsilon M]$.

Fix $p', p'' \in \mathbb{B}_\beta(p_0)$, with $p' = (z', w')$ and $p'' = (z'', w'')$, and consider any $x' \in S(p') \cap \mathbb{B}_{\alpha/2}(x_0)$. Then

$$\begin{aligned} 0 &\in z' + f(w', x') + N_C(x') \\ &= [z' + f(w', x') - \nabla_x f(w_0, x_0)x'] + Ax' + N_C(x'), \end{aligned}$$

so that $x' \in L(q') \cap \mathbb{B}_{\alpha/2}(x_0)$ for $q' = z' + f(w', x') - \nabla_x f(w_0, x_0)x'$, where in terms of the linearization map

$$g(x) = f(w_0, x_0) + \nabla_x f(w_0, x_0)(x - x_0) = q_0 - z_0 + \nabla_x f(w_0, x_0)x \tag{16}$$

we can write $q' - q_0 = z' - z_0 + f(w', x') - g(x')$. Using (15) we have

$$\begin{aligned} \|q' - q_0\| &= \|z' - z_0 + f(w', x') - g(x')\| \\ &\leq \|z' - z_0\| + +\|f(w', x') - f(w', x_0) - \nabla_x f(w_0, x_0)(x' - x_0)\| \\ &\qquad +\|f(w', x_0) - f(w_0, x_0)\| \\ &\leq \|p' - p_0\| + \varepsilon\|x' - x_0\| + l\|w' - w_0\| \leq \beta(1+l) + \frac{\varepsilon\alpha}{2}, \end{aligned}$$

$$\text{so that } \|q' - q_0\| \leq b, \text{ that is, } q' \in \mathbb{B}_b(q_0). \tag{17}$$

Analogously, for the vector $q'' = z'' + f(w'', x') - \nabla_x f(w_0, x_0)x' = q_0 + z'' - z_0 + f(w'', x') - g(x')$ we have $q'' \in \mathbb{B}_b(q_0)$. Let $x_1 = x'$. On the basis of (14) there exists then an x_2 such that

$$z'' + f(w'', x_1) + \nabla_x f(w_0, x_0)(x_2 - x_1) + N_C(x_2) \ni 0$$

and

$$\begin{aligned}\|x_2 - x_1\| &\leq M\|q' - q''\| \leq M(\|z' - z''\| + \|f(w', x_1) - f(w'', x_1)\|) \\ &\leq M(\|z' - z''\| + l\|w' - w''\|) \leq M(l+1)\|p' - p''\|.\end{aligned}$$

Suppose that there exist points $x_2, x_3, \ldots, x_{n-1}$ with

$$z'' + f(w'', x_{i-1}) + \nabla_x f(w_0, x_0)(x_i - x_{i-1}) + N_C(x_i) \ni 0$$

and

$$\|x_i - x_{i-1}\| \leq M(l+1)\|p' - p''\|(M\varepsilon)^{i-2} \text{ for } i = 2, \ldots, n-1.$$

Then for every i we have

$$\begin{aligned}\|x_i - x_0\| &\leq \|x_1 - x_0\| + \sum_{j=2}^{i} \|x_j - x_{j-1}\| \\ &\leq \frac{\alpha}{2} + M(l+1)\|p' - p''\| \sum_{j=2}^{i} (M\varepsilon)^{j-2} \\ &\leq \frac{\alpha}{2} + \frac{M(l+1)}{1 - \varepsilon M}\|p' - p''\| \leq \frac{\alpha}{2} + \frac{2M\beta(l+1)}{1 - \varepsilon M} \leq \alpha,\end{aligned}$$

because of (15). Setting $q_i = z'' + f(w'', x_i) - \nabla_x f(w_0, x_0)x_i = q_0 + z'' - z_0 + f(w'', x_i) - g(x_i)$ for $i = 2, 3, \ldots, n-1$ we get

$$\begin{aligned}\|q_i - q_0\| &= \|z'' - z_0 + f(w'', x_i) - g(x_i)\| \\ &\leq \|z'' - z_0\| + \|f(w'', x_i) - f(w'', x_0) - \nabla_x f(w_0, x_0)(x_i - x_0)\| \\ &\qquad + \|f(w'', x_0) - f(w_0, x_0)\| \\ &\leq \|p' - p_0\| + \varepsilon\|x_i - x_0\| + l\|w'' - w_0\| \\ &\leq \beta(1 + l) + \varepsilon\alpha \leq b,\end{aligned}$$

so that $q_i \in I\!B_b(q_0)$. Since $x_{n-1} \in L(q_{n-2}) \cap I\!B_\alpha(x_0)$, we know from the Aubin property (14) that there exists x_n with

$$z'' + f(w'', x_{n-1}) + \nabla_x f(w_0, x_0)(x_n - x_{n-1}) + N_C(x_n) \ni 0 \qquad (18)$$

and

$$\begin{aligned}\|x_n - x_{n-1}\| &\leq M\|q_{n-1} - q_{n-2}\| \\ &\leq M\|f(w'', x_{n-1}) - f(w'', x_{n-2}) - \nabla_x f(w_0, x_0)(x_{n-1} - x_{n-2})\| \\ &\leq M\varepsilon\|x_{n-1} - x_{n-2}\| \leq M(l+1)\|p' - p''\|)(M\varepsilon)^{n-2}.\end{aligned}$$

The induction step is thereby joined. We obtain an infinite sequence of points $x_1, x_2, \ldots, x_n, \ldots$ in $\mathbb{B}_\alpha(x_0)$ that is a Cauchy sequence and therefore converges to some $x'' \in \mathbb{B}_\alpha(x_0)$. Since $f(w'', \cdot)$ is continuous in $\mathbb{B}_\alpha(x_0)$ and the normal cone map N_C has closed graph, it follows from (18) that $x'' \in S(p'')$. Moreover, since

$$\begin{aligned} \|x_n - x'\| &\leq \sum_{i=2}^{n} \|x_i - x_{i-1}\| \\ &\leq M(l+1)\|p' - p''\| \sum_{i=2}^{n} (M\varepsilon)^{i-2} \leq \frac{M(l+1)}{1-\varepsilon M}\|p' - p''\|, \end{aligned}$$

we obtain in passing to the limit that

$$\|x'' - x'\| \leq \frac{M(l+1)}{1-\varepsilon M}\|p' - p''\| = M'\|p' - p''\|.$$

The implication (i) $\Rightarrow$ (ii) is thereby established.

To prove the implication (ii) $\Rightarrow$ (i), suppose S has the Aubin property at (p_0, x_0) with constant M. Choose ε, α and β relative to M as above. It will be demonstrated that L has the Aubin property at (q_0, x_0) with constant $M' = M/(1 - \varepsilon M)$. Consider $q', q'' \in \mathbb{B}_\beta(q_0)$ and $x' \in L(q') \cap \mathbb{B}_{\alpha/2}(x_0)$:

$$q' + \nabla_x f(w_0, x_0)x' + N_C(x') \ni 0.$$

Then $x' \in S(p') \cap \mathbb{B}_{\alpha/2}(x_0)$ for the parameter element $p' = (z', w_0)$ with $z' = q' + \nabla_x f(w_0, x_0)x' - f(w_0, x') = z_0 + [q' - q_0] - [f(w_0, x') - g(x')]$. Now also let $p'' = (z'', w_0)$ for the vector $z'' = q'' + \nabla_x f(w_0, x_0)x' - f(w_0, x') = z_0 + [q'' - q_0] - [f(w_0, x') - g(x')]$. As in the chain of estimates leading to (17) we get $p', p'' \in \mathbb{B}_b(p_0)$. Then there exists x_2 such that

$$z'' + f(w_0, x_2) + f(w_0, x_0) + \nabla_x f(w_0, x_0)(x' - x_0) - f(w_0, x') + N_C(x_2) \ni 0,$$

$$\|x_2 - x'\| \leq M\|p' - p''\| = M\|z' - z''\|.$$

By emulating the argument in the first part of the proof, we obtain by induction a sequence $x' = x_1, x_2, \ldots, x_n, \ldots$ convergent to x'' and such that

$$z'' + f(w_0, x_n) + f(w_0, x_0) + \nabla_x f(w_0, x_0)(x_{n-1} - x_0) - f(w_0, x_{n-1}) + N_C(x_n) \ni 0,$$

$$\|x_n - x'\| \leq M\|q' - q''\| \sum_{i=2}^{n} (M\varepsilon)^{i-2}.$$

Passing to the limit we obtain that $x'' \in L(q'')$ and

$$\|x' - x''\| \leq \frac{M}{1 - \varepsilon M}\|q' - q''\| = M'\|q' - q''\|.$$

This finishes the proof. □

Remark 4. The result in Proposition 1 carries over to a much wider setting, as may be gleaned from the proof we have given. Let F be a set-valued map with closed graph from a complete metric space X into the subsets of a linear normed space Z, and let f be a function from $W \times X$ to Z, where W is a metric space. Suppose that $g : X \mapsto Z$ is a continuous function which strongly approximates f around (w_0, x_0) in the sense of Robinson [25], and that f satisfies condition (B) (with the metric of W replacing the norm). Consider the maps

$$\Sigma(p) = \{x \in X \,|\, 0 \in z + f(w, x) + F(x)\}$$

where $p = (z, w)$, and

$$\Lambda(z) = \{x \in X \,|\, 0 \in z + g(x) + F(x)\},$$

and let $x_0 \in \Sigma(p_0)$. Then Σ has the Aubin property at (p_0, x_0) if and only if Λ has it at (z_0, x_0). Prototypes of such a theorem are contained in [5] and [6], where a fixed point argument is utilized. □

Next we give a new proof of the original result of Robinson in [23] (for the present context). In contrast to Robinson's argument, we do not appeal to a fixed point theorem but rely on Proposition 1 instead. Furthermore, whereas Robinson focused on the implication from the property of L to the corresponding one for S, we point out that—in the presence of the canonical perturbation vector z alongside of w in the element $p = (z, w)$—the implication goes both ways and becomes an equivalence.

Proposition 2. *The following properties are equivalent:*
(i) *L is locally single-valued and Lipschitz continuous around (q_0, x_0);*
(ii) *S is locally single-valued and Lipschitz continuous around (p_0, x_0).*

Proof. Let (i) hold. In particular L has the Aubin property, hence S has it too by Proposition 2. To get (ii) it suffices therefore to verify that S is locally single-valued.

Suppose to the contrary that in every neighborhood V of p_0 and X of x_0 there exist $\bar{p} = (\bar{z}, \bar{w})$ and $x_1, x_2 \in S(\bar{p}) \cap X$ such that $x_1 \neq x_2$: in particular,

$$\bar{z} + f(\bar{w}, x_i) + N_C(x_i) \ni 0 \text{ for } i = 1, 2.$$

Let M be the Lipschitz constant of L around (q_0, x_0) and choose $\varepsilon > 0$ small enough that $M\varepsilon < 1$. Using the Strict Differentiability Lemma, choose neighborhoods U of x_0 and V of w_0 with $U \subset X$ such that

$$\|f(w, x') - f(w, x'') - \nabla_x f(w_0, x_0)(x' - x'')\| \le \varepsilon \|x' - x''\|$$

for every $x', x'' \in U$ and $w \in V$. Note that, for $i = 1, 2$, and for some sufficiently small neighborhood $U' \subset U$ of x_0, we have $L(q_i) \cap U' = \{x_i\}$ for $q_i = q' + \nabla_x f(w_0, x_0)x' - f(w_0, x') = q_0 + [\bar{z} - z_0] + [f(\bar{w}, x_i) - g(x_i)]$, where g is given as before by (16). Then

$$\begin{aligned}
\|x_1 - x_2\| &\le M\|q_1 - q_2\| \\
&= M\|[f(\bar{w}, x_1) - g(x_1)] - [f(\bar{w}, x_2) - g(x_2)]\| \\
&= M\|f(\bar{w}, x_1) - f(\bar{w}, x_2) - \nabla_x f(x_0, w_0)(x_1 - x_2)\| \\
&\le M\varepsilon\|x_1 - x_2\| < \|x_1 - x_2\|,
\end{aligned}$$

which is a contradiction. Hence S is locally single-valued around (p_0, x_0).

The converse implication (ii) $\Rightarrow$ (i) is established in the same way. □

Combining Propositions 1 and 2 with Theorems 1 and 2 we obtain the following result, in which the implication from (i) to (iv), already known from Robinson's theorem in [23], has ended up in a circle of equivalences.

Theorem 3. *The following properties are equivalent:*

(i) *The nonlinear variational inequality (1) is strongly regular at (p_0, x_0);*

(ii) *The critical face condition of Definition 2 holds at (q_0, x_0);*

(iii) *The solution map S has the Aubin property at (p_0, x_0);*

(iv) *The solution map S is locally single-valued and Lipschitz continuous around (p_0, x_0).*

4. Application to the Complementarity Problem.

Next we apply our results to the nonlinear complementarity problem with canonical perturbations, namely

$$x \ge 0, \quad f(w, x) + z \ge 0, \quad \langle x, f(w, x) + z \rangle = 0, \tag{19}$$

which is the special case of (1) with $C = I\!R^n_+$. Our general assumptions and notation for (1) continue in the analysis of this case, in particular (3) and (4). We associate with the vector $v_0 \in N_C(x_0)$ the index sets J_1, J_2, J_3 in $\{1, 2, \ldots, n\}$ given by

$$\begin{aligned} J_1 &= \{i \,|\, x_0^i > 0,\ v_0^i = 0\}, \\ J_2 &= \{i \,|\, x_0^i = 0,\ v_0^i = 0\}, \\ J_3 &= \{i \,|\, x_0^i = 0,\ v_0^i < 0\}. \end{aligned}$$

Proposition 3. *In the case of the nonlinear complementarity problem, the critical cone K_0 consists of the vectors x' satisfying*

$$\begin{cases} x'_i \text{ free} & \text{for } i \in J_1, \\ x'_i \geq 0 & \text{for } i \in J_2, \\ x'_i = 0 & \text{for } i \in J_3, \end{cases}$$

and the cones $F_1 - F_2$, where F_1 and F_2 are closed faces of K_0 with $F_1 \supset F_2$, are the cones K of the following form. There is a partition of $\{1, 2, \ldots, n\}$ into index sets J'_1, J'_2, J'_3 with $J_1 \subset J'_1 \subset J_1 \cup J_2$ and $J_3 \subset J'_3 \subset J_3 \cup J_2$, such that K consists of the vectors x' satisfying

$$\begin{cases} x'_i \text{ free} & \text{for } i \in J'_1, \\ x'_i \geq 0 & \text{for } i \in J'_2, \\ x'_i = 0 & \text{for } i \in J'_3. \end{cases} \tag{20}$$

The vectors $u \in K$ with $A^\top u \in K^*$ are then the ones such that

$$\begin{cases} u_i \text{ free},\ (A^\top u)_i = 0 & \text{for } i \in J'_1, \\ u_i \geq 0,\ (A^\top u)_i \leq 0 & \text{for } i \in J'_2, \\ u_i = 0,\ (A^\top u)_i \text{ free} & \text{for } i \in J'_3. \end{cases}$$

Proof. It is easy to see that K_0 has the form described, so we focus on analyzing its closed faces. Each such face F has the form $K_0 \cap [v']^\perp$ for some vector $v' \in K_0^*$. The vectors v in question are those with

$$\begin{cases} v'_i = 0 & \text{for } i \in J_1, \\ v'_i \leq 0 & \text{for } i \in J_2, \\ v'_i \text{ free} & \text{for } i \in J_3. \end{cases}$$

The closed faces F of K_0 correspond one-to-one therefore with the subsets of J_2: the face F corresponding to an index set J_2^F consists of the vectors x' such that

$$\begin{cases} x'_i \text{ free} & \text{for } i \in J_1, \\ x'_i \geq 0 & \text{for } i \in J_2 \setminus J_2^F, \\ x'_i = 0 & \text{for } i \in J_3 \cup J_2^F. \end{cases}$$

If F_1 and F_2 have $J_2^{F_1} \subset J_2^{F_2}$, so that $F_1 \supset F_2$, then $F_1 - F_2$ is given by (20) with $J'_1 = J_1 \cup [J_2 \setminus J_2^{F_2}]$, $J'_2 = J_2^{F_2} \setminus J_2^{F_1}$, $J'_3 = J_3 \cup J_2^{F_1}$. □

Theorem 4. *The general complementarity problem (19) is strongly regular at* (p_0, x_0) *if and only if the following condition holds for the entries* a_{ij} *of* A*: if* u_i *for* $i \in J_1 \cup J_2$ *are numbers satisfying*

$$\sum_{i \in J_1 \cup J_2} u_i a_{ij} \begin{cases} = 0 & \text{for } j \in J_1 \text{ and for } j \in J_2 \text{ with } u_j < 0, \\ \leq 0 & \text{for } j \in J_2 \text{ with } u_j > 0, \end{cases}$$

then $u_i = 0$ *for all* $i \in J_1 \cup J_2$.

Proof. This condition specializes the critical face condition to this setting, as seen from Proposition 3. It remains only to apply Theorem 3. □

Remark 5. The sufficiency of the condition in Theorem 4 for strong regularity can also be proved directly. Consider the linear complementarity problem (5) and its solution map L_0, which has $x_0 \in L_0(q_0)$. Assume temporarily that $J_1 = \emptyset$ and $J_3 = \emptyset$. Let a be a subset of $\{1, 2, \ldots, n\}$ and let A_{aa} be the corresponding submatrix of A. Let $u^a \in \ker A_{aa}^\top$ and

$$u_j = \begin{cases} u_j^a & \text{if } j \in a, \\ 0 & \text{otherwise} . \end{cases}$$

Then $(A^\top u)_j = 0$ for all j for which $u_j \neq 0$. Hence, from the condition displayed in Theorem 4, we have $u = 0$. Thus A_{aa} is nonsingular, and we see every principal submatrix of A is nonsingular. Furthermore, let $j \in \{1, 2, \cdots, n\}$ and let $u \in I\!R^n$ be such that $u_j = 1$ and $u_i = 0$ for all $i \neq j$. Then $(A^\top u)_j = a_{jj}$. If $a_{jj} \leq 0$, the condition in Theorem 4 implies that $u = 0$, a contradiction. Hence $a_{ii} > 0$ for all i.

There is a linear one-to-one correspondence between the graph of the solution map L_0 and the graph of the solution map of any principal pivotal transform of the linear complementarity problem (5); see e.g. [4]. Then the solution map of any principal pivotal transform of (5) has the Aubin property at $(0,0)$. By the above argument all diagonal entries of any principal pivotal transform of A are positive. This means that A is a P-matrix, see [19], p. 205. This in turn is equivalent to the condition that L_0 is single-valued everywhere.

Consider now the general case. First observe that the submatrix A_{11} corresponding to the set J_1 is nonsingular. (Take $u^1 \in \ker A_{11}$ and $u_j = u^1_j$ if $j \in J_1$, but $u_j = 0$ otherwise; then $(A^\top u)_j = 0$ for all $j \in J_1$, hence from the condition in Theorem 4, $u^1 = 0$.) Utilizing the Reduction Lemma we come to a complementarity problem of the form

$$\begin{aligned} 0 &= A_{11}x^1 + A_{12}x^2 + q^1 \\ z^2 &= A_{12}x^1 + A_{22}x^2 + q^2 \\ x^2 &\geq 0,\ z^2 \geq 0,\ \langle x^2, z^2\rangle = 0, \end{aligned}$$

where the superscripts of x correspond to the sets of indices J_1 and J_2. By solving the first equation and substituting to the second one we obtain a problem whose solution map has the Aubin property at $(0,0)$; that is, $J_1 = \emptyset$, $J_3 = \emptyset$. This case has already been treated. □

As a corollary of Theorem 4 we obtain the following characterization of P-matrices.

Corollary 2. *For an $n \times n$ matrix A, the following are equivalent:*
(i) *A is a P-matrix;*
(ii) *For $u \in I\!R^n$,*

$$(A^\top u)_j \begin{cases} = 0 & \text{for } j \text{ with } u_j < 0 \\ \leq 0 & \text{for } j \text{ with } u_j > 0 \end{cases} \quad \Longrightarrow \quad u = 0.$$

5. Application to Nonlinear Programming.

For a further illustration of our general results we consider the nonlinear programming problem

$$\text{minimize } g_0(w,x) + \langle v, x\rangle \text{ in } x \text{ subject to}$$

$$g_i(w,x) - u_i \begin{cases} = 0 & \text{for } i \in [1,r], \\ \leq 0 & \text{for } i \in [r+1,m], \end{cases} \tag{21}$$

for $\mathcal{C}^2$ functions $g_i : I\!R^d \times I\!R^n \to I\!R$, $i = 0, 1, \ldots, m$, where the vectors $w \in I\!R^d$, $v \in I\!R^n$ and $u = (u_1, \ldots, u_m) \in I\!R^m$ are parameters. In terms of

$$L(w, x, y) = g_0(w, x) + y_1 g_1(w, x) + \cdots + y_m g_m(w, x)$$

the first-order optimality conditions for this problem, namely the Karush-Kuhn-Tucker conditions, take the form

$$\begin{cases} v + \nabla_x L(w, x, y) = 0, \\ -u + \nabla_y L(w, x, y) \in N_Y(y) \quad \text{for } \; Y = I\!R^r \times I\!R^{m-r}_+. \end{cases} \tag{22}$$

These can be written together as the variational inequality

$$(v, u) + f(w, x, y) + N_C(x, y) \ni (0, 0) \tag{23}$$

under the choice of elements

$$f(w, x, y) = (\nabla_x L(w, x, y), -\nabla_y L(w, x, y)), \qquad C = I\!R^n \times Y. \tag{24}$$

(Obviously $\nabla_y L(w, x, y)$ is simply the vector in $I\!R^m$ having as its components the values $g_i(w, x)$ for $i = 1, \ldots, m$.) Here (x, y) replaces the point x of the general theory, while (v, u) corresponds to the canonical perturbation vector z. The set C is a polyhedral convex cone, and the map f satisfies our blanket assumptions (A) and (B). (Weaker conditions on the g_i's would suffice, but we leave that aside.)

Consider any pair (x_0, y_0) satisfying the KKT conditions (22)—or equivalently the variational inequality (23)—for given u_0, v_0, and w_0. We wish to work out what our results say about strong regularity in this variational inequality at $(u_0, v_0, w_0, x_0, y_0)$ and therefore about the local single-valuedness and Lipschitz continuity of the map S_{KKT} that assigns to each (u, v, w) the set of KKT pairs (x, y) in problem (21).

Associate with the given elements u_0, v_0, w_0, x_0, y_0, the index sets I_1, I_2, I_3 in $\{1, 2, \ldots, m\}$ defined by

$$\begin{aligned} I_1 &= \{i \in [r+1, m] \,|\, g_i(w_0, x_0) - u_{0i} = 0,\ y_{0i} > 0\} \cup \{1, \ldots, r\}, \\ I_2 &= \{i \in [r+1, m] \,|\, g_i(w_0, x_0) - u_{0i} = 0,\ y_{0i} = 0\}, \\ I_3 &= \{i \in [r+1, m] \,|\, g_i(w_0, x_0) - u_{0i} < 0,\ y_{0i} = 0\}. \end{aligned}$$

The tangent cone $T_C(x_0, y_0)$ consists of all the vectors $(x', y') \in I\!R^n \times I\!R^m$ such that

$$\begin{cases} x' \text{ free} & \\ y_i' \text{ free} & \text{for } i \in I_1, \\ y_i' \geq 0 & \text{for } i \in I_2 \cup I_3. \end{cases}$$

By definition the critical cone K_0 consists of all $(x', y') \in T_C(x_0, y_0)$ orthogonal to the vector $(v_0, u_0) + f(w_0, x_0, y_0)$, but the KKT conditions imply that all the components of this vector are 0 except for the ones at the end corresponding to inactive inequality constraints. From this it is apparent that the critical cone is given by

$$(x', y') \in K_0 \quad \Longleftrightarrow \quad \begin{cases} x' \text{ free} & \\ y_i' \text{ free} & \text{for } i \in I_1, \\ y_i' \geq 0 & \text{for } i \in I_2, \\ y_i' = 0 & \text{for } i \in I_3. \end{cases} \tag{25}$$

On the other hand, the matrix A in the critical face condition specializes to

$$A = \begin{bmatrix} H(w_0, x_0, y_0) & G(w_0, x_0)^\top \\ -G(w_0, x_0) & 0 \end{bmatrix} \tag{26}$$

for the second-derivative matrix $H(w, x, y) = \nabla^2_{xx} L(w, x, y)$ and the matrix $G(w, x) = \nabla^2_{yx} L(w, x, y)$ having as its rows the constraint gradient vectors $\nabla_x g_i(w, x)$ for $i = 1, \ldots, m$.

Theorem 5. *The variational inequality (23)–(24) associated with the KKT conditions (22) is strongly regular for $(u_0, v_0, w_0, x_0, y_0)$ if and only if the following two requirements, specializing the critical face condition to this setting, are fulfilled:*

(a) *The vectors $\nabla_x g_i(w_0, x_0)$ for $i \in I_1 \cup I_2$ are linearly independent;*

(b) *For each partition of $\{1, 2, \ldots, m\}$ into index sets I_1', I_2', I_3' with $I_1 \subset I_1' \subset I_1 \cup I_2$ and $I_3 \subset I_3' \subset I_3 \cup I_2$, the cone $K(I_1', I_2') \subset I\!R^n$ consisting of all the vectors x' satisfying*

$$\langle \nabla_x g_i(w_0, x_0), x' \rangle \begin{cases} = 0 & \text{for } i \in I_1', \\ \leq 0 & \text{for } i \in I_2', \end{cases}$$

should be such that

$$x' \in K(I_1', I_2'),\ \ \nabla^2_{xx} L(w_0, x_0, y_0) x' \in K(I_1', I_2')^* \quad \Longrightarrow \quad x' = 0.$$

Proof. From the analysis in Sections 3 and 4 and the observations just made, it is evident that the cones of form $F_1 - F_2$ in which F_1 and F_2 are closed faces of K_0 with $F_1 \supset F_2$ correspond one-to-one with the partitions (I_1', I_2', I_3') by

$$(x', y') \in F_1 - F_2 \iff \begin{cases} x' \text{ free} & \\ y_i' \text{ free} & \text{for } i \in I_1', \\ y_i' \geq 0 & \text{for } i \in I_2', \\ y_i' = 0 & \text{for } i \in I_3', \end{cases} \tag{27}$$

in which case

$$(x'', y'') \in (F_1 - F_2)^* \iff \begin{cases} x'' = 0 & \\ y_i'' = 0 & \text{for } i \in I_1', \\ y_i'' \leq 0 & \text{for } i \in I_2', \\ y_i'' \text{ free} & \text{for } i \in I_3'. \end{cases}$$

In view of the structure determined for A in (26) the critical face condition emerges as the requirement that whenever x' and y' satisfy (27) and have

$$H(w_0, x_0, y_0)x' - \sum_{i=1}^{m} y_i' \nabla_x g_i(w_0, x_0) = 0 \text{ with } x' \in K(I_1', I_2'),$$

then $x' = 0$ and $y' = 0$. The vectors of the form $\sum_{i=1}^{m} y_i' \nabla_x g_i(w_0, x_0)$ with y' satisfying (27) are of course the ones in the polar cone $K(I_1', I_2')^*$ (by Farkas' Lemma), so we see that the critical cone condition comes down to (b) along with the requirement that no $\sum_{i=1}^{m} y_i' \nabla_x g_i(w_0, x_0)$ with y' satisfying (27) can vanish unless $y' = 0$. Since the partition $I_1' = I_1 \cup I_2$, $I_2' = 0$, $I_3' = I_3$ can be taken as a special case, the latter means neither more nor less than the linear independence in (a). □

By combining Theorem 5 with second-order conditions we obtain a characterization of the case where the KKT map also gives local optimality.

Theorem 6. *The following are equivalent:*

(i) *The map S_{KKT} is locally single-valued and Lipschitz continuous around $(u_0, v_0, w_0, x_0, y_0)$, moreover with the property that for all $(u, v, w, x, y) \in$ gph S_{KKT} in some neighborhood of $(u_0, v_0, w_0, x_0, y_0)$, x is a locally optimal solution to the nonlinear programming problem (21) for (u, v, w);*

(ii) *The constraint gradients* $\nabla_x g_i(w_0, x_0)$ *for* $i \in I_1 \cup I_2$ *are linearly independent and the strong second-order sufficient condition for local optimality holds for* $(u_0, v_0, w_0, x_0, y_0)$*: one has*

$$\langle x', \nabla^2_{xx} L(w_0, x_0, y_0)x'\rangle > 0 \text{ for all } x' \neq 0 \text{ in the subspace}$$

$$M = \{x' \,|\, x' \perp \nabla_x g_i(w_0, x_0) \text{ for all } i \in I_1\}.$$

Proof. From Theorem 5 we already know that the local single-valuedness and Lipschitz continuity in (i) require the linear independence in (ii). On the other hand, the positive definiteness in (ii) suffices by Theorem 5 for S_{KKT} to be locally single-valued and Lipschitz continuous around $(u_0, v_0, w_0, x_0, y_0)$, because $\langle x', \nabla^2_{xx} L(w_0, x_0, y_0)x'\rangle \leq 0$ when x' and $\nabla^2_{xx} L(w_0, x_0, y_0)x'$ belong to cones that are polar to each other. Henceforth we therefore work in the picture of S_{KKT} being locally single-valued and Lipschitz continuous around $(u_0, v_0, w_0, x_0, y_0)$, with both (a) and (b) of Theorem 5 holding. In particular then we have

$$x' \in M, \quad \nabla^2_{xx} L(w_0, x_0, y_0)x' \in M^\perp \quad \Longrightarrow \quad x' = 0 \tag{28}$$

because $K(I_1', I_2') = M$ and $K(I_1', I_2') = M^\perp$ when $I_1' = I_1$, $I_2' = \emptyset$, $I_3' = I_2 \cup I_3$. The focus is on verifying that the local optimality in (i) corresponds in these circumstances to the positive definiteness in (ii).

For simplicity we denote by $S_0(u, v, w)$ the uniquely determined pair (x, y) in the local single-valuedness of $S_{\text{KKT}}(u, v, w)$. We limit attention to parameter elements (u, v, w) near enough to (u_0, v_0, w_0) for this to make sense.

Let $P(u, v, w)$ be the nonlinear programming problem associated with (u, v, w) in (21). This problem has $(x, y) = S_0(u, v, w)$ as a KKT pair, with $(x, y) \to (x_0, y_0)$ as $(u, v, w) \to (u_0, v_0, w_0)$. For (u, v, w) close to (u_0, v_0, w_0), the index sets $I_1(u, v, w)$, $I_2(u, v, w)$, and $I_2(u, v, w)$ that correspond to (x, y) as I_1, I_2 and I_3 do to (x_0, y_0) must satisfy

$$I_1 \subset I_1(u, v, w) \subset I_1 \cup I_2, \qquad I_3 \subset I_3(u, v, w) \subset I_3 \cup I_2. \tag{29}$$

In particular $I_1(u, v, w) \cup I_2(u, v, w) \subset I_1 \cup I_2$, so the gradients $\nabla_x g_i(w, x)$ for $i \in I_1(u, v, w) \cup I_2(u, v, w)$ must be linearly independent. We know then that for x to be locally optimal in $P(u, v, w)$ it is necessary that

$$\langle x', \nabla^2_{xx} L(w, x, y)x'\rangle \geq 0 \text{ for all } x' \text{ satisfying}$$

$$\langle \nabla_x g_i(w,x), x' \rangle \begin{cases} = 0 & \text{for } i \in I_1(u,v,w), \\ \leq 0 & \text{for } i \in I_2(u,v,w), \end{cases} \tag{30}$$

whereas a sufficient condition for x to be locally optimal in $P(u, v, w)$ is the same thing with "≥ 0" strengthened to "> 0" when $x' \neq 0$. (See [10], Theorem 10.1.)

The sufficiency just described leads immediately through (29) to the conclusion that the positive definiteness in (ii) entails the local optimality in (i). Conversely, if (i) holds the second-order necessary condition must be satisfied by $(x, y) = S_0(u, v, w)$ for all (u, v, w) near enough to (u_0, v_0, w_0). The gradient linear independence property ensures that we can find a sequence of points $x^k \to x_0$ with

$$g_i(w_0, x^k) - u_{0i} \begin{cases} = 0 & \text{for } i \in I_1, \\ < 0 & \text{for } i \in I_2 \cup I_3. \end{cases}$$

For $v^k = -\nabla_x L(w_0, x^k, y_0)$ we have the KKT conditions in $P(u_0, v^k, w_0)$ satisfied by (x^k, y_0), hence $(x^k, y_0) = S_0(u_0, v^k, w_0)$ (for k sufficiently large). Then the necessary condition in (30) holds for these elements, with $I_1(u_0, v^k, w_0) = I_1$ and $I_2(u_0, v^k, w_0) = \emptyset$ so that the condition is just

$$\langle x', \nabla^2_{xx} L(w_0, x^k, y_0) x' \rangle \geq 0 \text{ for all } x' \in M.$$

In the limit as $k \to \infty$ we obtain (from the continuity of the second derivatives of L) that

$$\langle x', \nabla^2_{xx} L(w_0, x_0, y_0) x' \rangle \geq 0 \text{ for all } x' \in M.$$

This positive semidefiniteness relative to the subspace M must actually be positive definiteness, for otherwise there would have to exist by the symmetry of $\nabla^2_{xx} L(w_0, x_0, y_0)$ a vector $x' \neq 0$ in M with $\nabla^2_{xx} L(w_0, x_0, y_0) x' \in M^\perp$. But that is impossible by (28). □

Corollary 3. *In the convex programming case (where $g_i(w, x)$ is affine in x for $i = 1, \ldots, r$ while $g_i(w, x)$ is convex in x for $i = 0$ and $i = r+1, \ldots, m$), condition (ii) of Theorem 6 is both necessary and sufficient for the map S_{KKT} to be locally single-valued and Lipschitz continuous.*

Proof. In this case the local optimality in Theorem 6(i) is automatic. □

Several algebraic characterizations of strong regularity in optimization have previously been available in the literature. In his original paper [23], Robinson characterized the strong regularity of linear KKT systems. Kojima [14] introduced the concept of strong stability of the KKT system of a nonlinear program, which roughly means that the map S_{KKT} is locally single-valued and continuous with respect to a sufficiently rich class of perturbations. He gave a characterization of this property and noted (in his Corollary 6.6.) that a KKT point (x_0, y_0), of the kind in which x_0 is a local minimizer and the gradients of the active constraints are linearly independent, is strongly stable if and only if the strong second-order sufficient condition holds. Jongen et al. [11] proved through a far-reaching inertia-type theorem that strong stability in the sense of Kojima and strong regularity in the sense of Robinson are equivalent properties. Another characterization of the strong regularity of a KKT point was obtained by Kummer [15], who based his argument on a general implicit-function-type theorem for a nonsmooth equation equivalent to the KKT system. A related approach was developed in Jongen, Klatte and Tammer [12] and Klatte and Tammer [13], the latter containing a survey of characterizations of strong regularity. Our critical face condition in Theorem 5 differs from the conditions of Robinson, Kojima and Kummer, and it is not clear to us how one could derive the equivalence between these various conditions directly.

The implication (ii) $\Rightarrow$ (i) in Theorem 6 was noted by Robinson in [23], Theorem 4.1. Note that the condition (i) in Theorem 6 implies, via Proposition 2, that x_0 is a local minimizer and (x_0, y_0) is a strongly regular KKT point. Furthermore, from the strong regularity one can obtain directly that the gradients of the active constraints are linearly independent. Then, by combining the results of Kojima [14] and Jongen et al. [11], one obtains (ii). Recently, Bonnans and Sulem [2] gave a different proof of this implication.

References

[1] J.-P. Aubin, Lipschitz behavior of solutions to convex minimization problems, Math. Oper. Res. **9** (1984), 87-111.

[2] J. F. Bonnans, A. Sulem, Pseudopower expansion of solutions of generalized equations and constrained optimization problems, Mathematical Programming, 1995.

[3] J. M. Borwein, D. M. Zhuang, Verifiable necessary and sufficient conditions for openness and regularity of set-valued maps, J. Math. Anal. Appl. **134** (1988), 441-459.

[4] R. W. Cottle, Jong-Shi Pang, R. E. Stone, *The linear complementarity problem*, Academic Press, Boston 1992.

[5] A. L. Dontchev, W. W. Hager, An inverse function theorem for set-valued maps, Proc. Amer. Math. Soc. **121** (1994), 481-489.

[6] A. L. Dontchev, W. W. Hager, Implicit functions, Lipschitz maps and stability in optimization, Math. of Oper. Res. **19** (1994), 753-768.

[7] B. C. Eaves, U. G. Rothblum, Relationship of properties of piecewise affine maps over ordered fields, Linear Algebra and Appl. **132** (1990), 1-63.

[8] M. S. Gowda, On the continuity of the solution map in linear complementarity problems, SIAM J. Optimization **2** (1992), 619-634.

[9] M. S. Gowda, R. Sznajder, On the Lipschitz properties of polyhedral multifunctions, preprint, Dept. of Math. and Stat., University of Maryland Baltimore County, November 1994.

[10] M. Hestenes, *Calculus of Variations and Optimal Control Theory*, Wiley, 1966.

[11] H. Th. Jongen, T. Möbert, J. Rückmann, K. Tammer, Implicit functions and sensitivity of stationary points, Linear Algebra and Appl. **95**(1987), 97–109.

[12] H. Th. Jongen, D. Klatte, K. Tammer, Implicit functions and sensitivity of stationary points, Mathematical Programming **19** (1990), 123–138.

[13] D. Klatte, K. Tammer, Strong stability of stationary solutions and Karush-Kuhn-Tucker points in nonlinear optimization, Annals of Oper. Research **27** (1990), 285–308.

[14] M. Kojima, Strongly stable stationary solutions in nonlinear programming, in S. M. Robinson, ed., *Analysis and computation of fixed points*, Academic Press, New York, 1988, 93–138.

[15] B. Kummer, Lipschitzian inverse functions directional derivatives and application in $\mathcal{C}^{1,1}$ optimization, J. Opt. Theory Appl. **70** (1991), 561–582.

[16] O. L. Mangasarian, T.-H. Shiau, Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems, SIAM J. Control Optim. **25** (1987), 582-595.

[17] B. Mordukhovich, Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis, Trans. Amer. Math. Soc. **343** (1994), 609-657.

[18] G. S. R. Murthy, T. Parthasarathy, M. Sabatini, On Lipschitzian **Q**-matrices, Technical Report 20, Statistical Quality Control and Operations Res. Unit, Indian Statistical Institute, July 1994.

[19] K. G. Murty, *Linear complementarity, linear and nonlinear programming*, Heldermann Verlag, Berlin, 1988.

[20] Jong-Shi Pang, A degree-theoretic approach to parametric nonsmooth equations with multivalued perturbed solution sets, Mathematical Programming **62** (1993), 359-383.

[21] J.-P. Penot, Metric regularity, openness and Lipschitz multifunctions, Nonlinear Anal. **13** (1989), 629-643.

[22] D. Ralph, A new proof of Robinson's homeomorphism theorem for PL-normal maps, Linear Algebra and Appl. **178** (1993), 249-260.

[23] S. M. Robinson, Strongly regular generalized equations, Math. Oper. Res. **5** (1980), 43-62.

[24] S. M. Robinson, Local structure of feasible sets in nonlinear programming, Part II: Nondegeneracy, Mathematical Programming Study **22** (1984), 217–230.

[25] S. M. Robinson, An implicit-function theorem for a class of nonsmooth functions, Math. Oper. Res. **16** (1991), 292-309.

[26] S. M. Robinson, Normal maps induced by linear transformations, Math. of Oper. Res. **17** (1992), 691-714.

[27] R. T. Rockafellar, Lipschitz properties of multifunctions, Nonlinear Analysis **9** (1985), 867-885.

[28] R. T. Rockafellar, Proto-differentiability of set-valued mappings and its applications in optimization, Ann. Inst. H. Poincaré, Analyse Non Linéaire **6** (1989), suppl. 449-482.

[29] H. Samelson, R. M. Thrall, O. Wesler, A partition theorem for Euclidean n-space, Proc. Amer. Math. Soc. **9** (1958), 805-807.

[30] S. Scholtes, Introduction to piecewise differentiable equations, preprint 53/1994, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe, May 1994.