REGULARITY PROPERTIES OF A SEMISMOOTH REFORMULATION OF VARIATIONAL INEQUALITIES[1]

Francisco Facchinei[1], Andreas Fischer[2] and Christian Kanzow[3]

[1] Università di Roma "La Sapienza"
Dipartimento di Informatica e Sistemistica
Via Buonarroti 12, I-00185 Roma, Italy
e-mail: soler@dis.uniroma1.it

[2] Technical University of Dresden
Institute of Numerical Mathematics
D-01062 Dresden, Germany
e-mail: fischer@math.tu-dresden.de

[3] University of Hamburg
Institute of Applied Mathematics
Bundesstrasse 55, D-20146 Hamburg, Germany
e-mail: kanzow@math.uni-hamburg.de

December 1995 (revised December 1996)

Abstract: Variational inequalities over sets defined by systems of equalities and inequalities are considered. A new reformulation of the KKT-conditions of the variational inequality as a system of equations is proposed. A related unconstrained minimization reformulation is also investigated. As a by-product of the analysis, a new characterization of strong regularity of KKT-points is given.

Key Words: Variational inequality problem, KKT-conditions, strong regularity.

1 Introduction

We consider the variational inequality problem, $VIP(X, F)$ for short, which is to find a vector $x^* \in X$ such that

$$F(x^*)^T(x - x^*) \geq 0 \qquad \text{for all } x \in X, \tag{1}$$

where $X \subseteq \mathbb{R}^n$ is a closed set and $F : \mathbb{R}^n \to \mathbb{R}^n$ is any given function which we assume to be at least continuously differentiable.

Variational inequality problems include as special cases complementarity problems and necessary optimality conditions for constrained optimization problems and have many applications in the engineering and economic sciences, see, e.g., [13, 22].

This has motivated an increasing interest in variational inequalities, from both the theoretical and the algorithmic point of view. One of the main paths followed in the theoretical study of variational inequalities has been their reduction to nonsmooth systems of (generalized) equations, which opens the way to the application of some sophisticated machinery developed in the latter field. In particular, the reformulation as a generalized equation led to a dramatic improvement in our understanding of the properties of variational inequalities, see, for example, [19, 27, 28]. Another important kind of reformulation, which attracted much attention and also gave important insights into variational inequality problems, is the one based on the normal map, see [7, 20, 31, 32] and references therein.

Naturally associated to the equation reformulation is the optimization approach, where the square of the norm of the system of equations is minimized in order to find numerically a solution of the variational inequality. Also in this case the use of the normal map equation reformulation led to the development of robust and effective algorithms, see [2, 6, 26, 33, 34].

In this paper, we propose a new equation reformulation of the KKT-conditions of $VIP(X, F)$ (see below) and study some relevant properties of this reformulation and of the associated optimization problem. In order to explain this approach, a few words on the KKT-system associated to $VIP(X, F)$ are in order.

We assume that the feasible set X can be represented as follows:

$$X := \{x \in \mathbb{R}^n \,|\, h(x) = 0, \, g(x) \geq 0\}, \tag{2}$$

where $h : \mathbb{R}^n \to \mathbb{R}^p$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable functions. Then the following equations and inequalities are called the *KKT-conditions* of problem $VIP(X, F)$:

$$\begin{aligned}
F(x) + \nabla h(x)y - \nabla g(x)z &= 0, \\
h(x) &= 0, \\
g(x) \geq 0, \, z \geq 0, \, z^T g(x) &= 0.
\end{aligned} \tag{3}$$

A triple $(x, y, z) \in \mathbb{R}^{n+p+m}$ satisfying (3) is called a *KKT-point*.

There is a strong relationship between the KKT-conditions (3) and the variational inequality problem (1). If $x^* \in \mathbb{R}^n$ is a (local) solution of $VIP(X, F)$ and if a constraint

qualification holds (e.g., the Mangasarian-Fromovitz condition), then multiplier vectors $y^* \in \mathbb{R}^p$ and $z^* \in \mathbb{R}^m$ exist such that the vector $w^* := (x^*, y^*, z^*) \in \mathbb{R}^{n+p+m}$ is a KKT-point of VIP(X, F). Conversely, if all component functions h_i $(i = 1, \ldots, p)$ are affine and all component functions g_i $(i = 1, \ldots, m)$ are concave (so that X is a convex set) then the x-part of every KKT-point $w^* = (x^*, y^*, z^*)$ is a solution of VIP(X, F), see [17]. Note that F is not required to be monotone for this relationship to hold. Moreover, if X is a polyhedral set, then each solution of VIP(X, F) corresponds to a KKT-point and vice versa.

In order to reformulate the KKT-conditions of VIP(X, F) as a system of equations, we use the function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\varphi(a, b) := \sqrt{a^2 + b^2} - a - b.$$

This function has first been used by Fischer [14] in order to reformulate the KKT-conditions of an inequality constrained optimization problem. Since then it has become quite popular in the fields of linear and nonlinear complementarity, constrained optimization and variational inequality problems, see, e.g., the survey paper [15]. The main property of this function is the following characterization of its zeros:

$$\varphi(a, b) = 0 \iff a \geq 0, b \geq 0, ab = 0.$$

Therefore, the KKT-conditions (3) can equivalently be written as the nonlinear system of equations

$$\Phi(w) := \Phi(x, y, z) = 0, \tag{4}$$

where $\Phi : \mathbb{R}^{n+p+m} \to \mathbb{R}^{n+p+m}$ is defined by

$$\Phi(w) = \Phi(x, y, z) := \begin{pmatrix} L(x, y, z) \\ h(x) \\ \phi(g(x), z) \end{pmatrix}$$

with

$$
\begin{aligned}
L(x, y, z) &:= F(x) + \nabla h(x)y - \nabla g(x)z, \\
\phi(g(x), z) &:= (\varphi(g_1(x), z_1), \ldots, \varphi(g_m(x), z_m))^T \in \mathbb{R}^m.
\end{aligned}
$$

Associated to the nonlinear system of equations (4), we can consider the problem of minimizing its natural merit function, i.e.,

$$\Psi(w) \to \min, \tag{5}$$

where

$$\Psi(w) := \frac{1}{2}\Phi(w)^T\Phi(w) = \frac{1}{2}\|\Phi(w)\|^2.$$

Note that the function φ is not differentiable in the origin, so that the system (4) is a nonsmooth reformulation of the KKT-conditions (3). However, it can be seen that Φ is a

locally Lipschitz-continuous operator, while, remarkably, if h and g are C^2-functions, Ψ is continuously differentiable, see Proposition 4.1 below.

Our main motivation for studying the reformulations (4) and (5) of the KKT-conditions (3) comes from the field of algorithms for complementarity and optimization problems. There, it was recently shown that reformulations based on the function φ possess many valuable properties, and allow the development of efficient solution methods [5, 8]. Moreover, we think that the approach proposed in this paper also gives additional insight in the variational inequality problem that could fruitfully be exploited in the analysis of, e.g., stability and sensitivity properties. In particular we point out that we shall establish a new, appealing characterization of strongly regular KKT-points of VIP(X, F). We refer the reader to [5, 15] for a more detailed discussion on approaches based on the function φ and to [9, 10, 11] for applications of the results of this paper to the development of algorithms for the solution of problem VIP(X, F).

In this paper we shall concentrate on two specific topics which, in the past few years, emerged as crucial points in reformulations of complementarity and variational inequality problems as systems of equations or as optimization problems:

(a) conditions which ensure the nonsingularity of all the generalized Jacobians of Φ (or at least ensure the existence of a nonsingular element in the generalized Jacobian), and

(b) conditions which guarantee that a stationary point of Ψ is a solution of the variational inequality VIP(X, F).

Points (a) and (b) are very relevant to the development of efficient algorithms [5, 9, 10] while point (a) could also pave the way to an application of the equation reformulation to the study of stability and sensitivity of solutions of variational inequalities.

The paper is organized as follows. In the next section we collect some basic definitions and results on coherent orientation. In Section 3 we study conditions for the nonsingularity of the generalized Jacobians of Φ. In Section 4 we are concerned with the important question of establishing conditions under which a stationary point of our merit function Ψ is a KKT-point of VIP(X, F). We present several conditions depending on the structure of the feasible set X.

Notation. A function $G : \mathbb{R}^n \to \mathbb{R}^t$ is called a C^k-function if it is k times continuously differentiable, and LC^k-function if it is a C^k-function whose kth derivative is locally Lipschitz-continuous. If $G : \mathbb{R}^n \to \mathbb{R}^t$ is a C^1-function, we denote by $\nabla G(x)$ the gradient of G (i.e., the transpose of the Jacobian) at x. If G is locally Lipschitzian at x, $\partial G(x)$ indicates Clarke's [3] generalized Jacobian, a set of $t \times n$ matrices. When $t = 1$, $\partial G(x)$ is the generalized gradient and is usually regarded as a set of column vectors. For a matrix $A \in \mathbb{R}^{n \times t}$, $A = (a_{ij})$, and subsets $I \subseteq \{1, \ldots, n\}$ and $J \subseteq \{1, \ldots, t\}$, we denote by A_{IJ} the submatrix of A consisting of the elements $a_{ij}, i \in I, j \in J$. Similarly, given any vector $v \in \mathbb{R}^n, v = (v_1, \ldots, v_n)^T$, and any subset $I \subseteq \{1, \ldots, n\}$, we denote by v_I the vector in $\mathbb{R}^{|I|}$ having the components $v_i, i \in I$. Let $A \in \mathbb{R}^{n \times n}, I \subseteq \{1, \ldots, n\}$ and $\bar{I} := \{1, \ldots, n\} \backslash I$

be the complementary subset of I. If we write

$$A = \begin{pmatrix} A_{II} & A_{I\bar{I}} \\ A_{\bar{I}I} & A_{\bar{I}\bar{I}} \end{pmatrix}$$

and assume that the principal submatrix A_{II} is nonsingular, then the Schur-complement of A_{II} in A is defined by $A/A_{II} := A_{\bar{I}\bar{I}} - A_{\bar{I}I} A_{II}^{-1} A_{I\bar{I}}$. A matrix $A \in \mathbb{R}^{n \times n}$ is a P_0-matrix (P-matrix) if the determinant of each of its principal submatrices is nonnegative (positive). We recall that every positive semidefinite (positive definite) matrix is a P_0-matrix (P-matrix), but the converse is not true in general. Finally, throughout the paper, $\| \cdot \|$ denotes the Euclidean vector norm.

2 Faces and Coherently Oriented Matrices

Let B be a square t-dimensional matrix. The *orientation* of B is defined to be the sign of the determinant of B, and can therefore assume three possible values: $-1, 0$ and $+1$. Let L be a t-dimensional subspace of $\mathbb{R}^n$ and Q be a matrix whose columns form a basis for L. If A is an $n \times n$ matrix, then the matrix $Q^T A Q$ is called the *section* of A in the subspace L. The orientation of A on the subspace L is defined to be the orientation of $Q^T A Q$. It is easy to see that, as long as the columns of Q form a basis of L, the sign of $Q^T A Q$ does not depend on the particular choice of the basis for L, so that the orientation of A on a subspace is well defined.

For arbitrarily chosen finite index sets I, J and vectors $a_i \in \mathbb{R}^n, b_j \in \mathbb{R}^n$ $(i \in I, j \in J)$, let C be the polyhedral cone defined by

$$C := \{v \in \mathbb{R}^n | a_i^T v = 0 \ (i \in I), b_j^T v \leq 0 \ (j \in J)\}.$$

We recall that a *face* of C is any set which can be obtained as $H \cap C$, where H is a supporting hyperplane of C (a hyperplane H is a *supporting hyperplane* of C if C is entirely contained in one of the two closed halfspaces defined by H). We also recall that the intersection of faces is a face.

The following proposition collects some facts that will be useful in this paper.

Proposition 2.1 *Let $\tilde{J}$ be a subset of J (the empty set is a possible subset). We define the following linear space:*

$$L(\tilde{J}) := \{v \in \mathbb{R}^n | a_i^T v = 0 \ (i \in I), b_j^T v = 0 \ (j \in \tilde{J})\}.$$

Then the following facts hold.

(a) For every $\tilde{J} \subseteq J$, $L(\tilde{J}) \cap C$ is a face of C.

(b) If F is a face of C then there exists an index set $\tilde{J} \subseteq J$ such that $F = L(\tilde{J}) \cap C$.

(c) Suppose that all the vectors a_i with $i \in I$ and b_j with $j \in J$ are linearly independent, and let $F = L(\tilde{J}) \cap C$ be a given face of C. Then the linear space generated by F is $L(\tilde{J})$.

Proof. (a) $L(\tilde{J})$ is an intersection of supporting hyperplanes and hence $L(\tilde{J}) \cap C$, being the intersection of faces, is a face.

(b) This derives immediately from, e.g., Theorem 7.27 in [1]. For the proof of point (c) it is important to note that, according to Theorem 7.27 in [1], the set $\tilde{J}$ can be taken equal to $J(F)$, where $J(F) := \{j \in J \,|\, b_j^T v = 0 \,\forall v \in F\}$.

(c) Let $\tilde{J}$ be given, and let $F = L(\tilde{J}) \cap C$. By the observation we made in the proof of part (b), we know that $F = C \cap L(J(F))$. On the other hand, by Lemma 7.26 in [1], we also know that the linear space generated by F is $L(J(F))$. So to complete the proof it will suffice to show that $\tilde{J} = J(F)$. By definition we have

$$
\begin{aligned}
F &= C \cap L(\tilde{J}) \\
&= \{v \in \mathbb{R}^n \,|\, a_i^T v = 0 \ (i \in I), b_j^T v = 0 \ (j \in \tilde{J}), b_j^T v \le 0 \ (j \in J \setminus \tilde{J})\}.
\end{aligned}
\tag{6}
$$

On the other hand, we also have

$$
\begin{aligned}
F &= C \cap L(J(F)) \\
&= \{v \in \mathbb{R}^n \,|\, a_i^T v = 0 \ (i \in I), b_j^T v = 0 \ (j \in J(F)), b_j^T v \le 0 \ (j \in J \setminus J(F))\}.
\end{aligned}
\tag{7}
$$

Suppose now, by contradiction, that $\tilde{J} \ne J(F)$. By the definitions of F and $J(F)$, $\tilde{J} \subseteq J(F)$; this means that an index k exists which is in $J(F)$ but not in $\tilde{J}$. Since the vectors a_i with $i \in I$ and b_j with $j \in J$ are linearly independent, we can find a point $\bar{v}$ such that

$$
a_i^T \bar{v} = 0 \ (i \in I), \quad b_j^T \bar{v} = 0 \ (j \in \tilde{J}), \quad b_j^T \bar{v} = -1 \ (j \in J \setminus \tilde{J}).
$$

By (6) $\bar{v}$ belongs to F, but this contradicts (7) since we should have $b_k^T \bar{v} = 0$, and thus the proof is complete. $\square$

We shall now introduce the concept of a *coherently oriented* matrix over a cone and some related results, see [2, 21, 29, 30] and references therein.

Definition 2.2 *A matrix $A \in \mathbb{R}^{n \times n}$ is said to be* positively coherently oriented *(positively semicoherently oriented, negatively coherently oriented, negatively semicoherently oriented) on a polyhedral cone C if the orientation of the section of A in all linear spaces spanned by the faces of C is $+1$ ($+1$ or $0, -1, -1$ or 0).*

A is said to be coherently oriented *(semicoherently oriented) on C if A is either positively coherently oriented (semicoherently oriented) or negatively coherently oriented (semicoherently oriented) on C.*

It may be interesting to stress that the property of, let us say, positive coherent orientation, is not hereditary, i.e., if A is positively coherently oriented on C and $C' \subseteq C$, then this does not imply that A is positively coherently oriented on C'. This is intuitively obvious since the property of being positively coherently oriented is connected to the matrix A *and* to the structure of the faces of C, and a cone C' contained in C can have a completely different structure. However, the following result, of which we omit the elementary proof, holds.

Proposition 2.3 *Suppose that A is positive definite (positive semidefinite) on a subspace L. If C is a cone contained in L, then A is positively coherently oriented (positively semicoherently oriented) on C.*

We note that the property of positive coherent orientation differs and is more general than that of positive definiteness. However, it may be interesting to note that if the linear transformation A can be represented by a symmetric matrix, then the positive coherent orientation of A on a polyhedral convex cone containing no lines reduces to positive definiteness, see [31].

We finally restate a result due to Liu [21, Lemma 3.6] which will be used in our analysis of Section 3 to determine the orientation of a matrix over a subspace.

Lemma 2.4 *Consider the following square matrix*

$$M = \begin{pmatrix} N & B \\ -B^T & 0 \end{pmatrix},$$

where $N \in \mathbb{R}^{n \times n}$ is arbitrary and $B \in \mathbb{R}^{n \times m}$, with $m \leq n$, has full column rank. Then, the orientation of M is equal to the orientation of the section of N in the subspace $Ker(B^T)$.

3 Nonsingularity Conditions for $\partial \Phi$

In this section, we present several conditions guaranteeing that all elements, or at least one element, in the generalized Jacobian $\partial\Phi(w)$ are nonsingular. The results of this section are of interest since, on the one hand, they relate the nonsingularity of $\partial\Phi$ to Robinson's [28] strong regularity condition and, on the other hand, they allow us to define superlinear convergent Newton methods for the solution of the nonlinear system $\Phi(w) = 0$, see [9, 11] for more details.

Our first aim is to better understand the structure of the generalized Jacobian of Φ. The next result gives some insight into the structure of the generalized Jacobian $\partial\Phi(w)$.

Proposition 3.1 *Suppose that F is a C^1-function and g and h are C^2-functions. Let $w = (x, y, z) \in \mathbb{R}^{n+p+m}$. Then, each element $H \in \partial\Phi(w)^T$ can be represented as follows:*

$$H = \begin{pmatrix} \nabla_x L(w) & \nabla h(x) & \nabla g(x) D_a(w) \\ \nabla h(x)^T & 0 & 0 \\ -\nabla g(x)^T & 0 & D_b(w) \end{pmatrix},$$

where $D_a(w) := diag(a_1(w), \ldots, a_m(w)), D_b(w) := diag(b_1(w), \ldots, b_m(w)) \in \mathbb{R}^{m \times m}$ are diagonal matrices whose ith diagonal elements are given by

$$a_i(w) = \frac{g_i(x)}{\sqrt{g_i(x)^2 + z_i^2}} - 1, \quad b_i(w) = \frac{z_i}{\sqrt{g_i(x)^2 + z_i^2}} - 1$$

if $(g_i(x), z_i) \neq 0$, and by

$$a_i(w) = \xi_i - 1, b_i(w) = \rho_i - 1 \text{ for any } (\xi_i, \rho_i) \text{ with } \|(\xi_i, \rho_i)\| \leq 1$$

if $(g_i(x), z_i) = 0$.

Proof. The first $n+p$ components of the vector function Φ are continuously differentiable, so the expression for the first $n+p$ columns of H readily follows. Then, consider the last m columns. Known rules on the evaluation of the generalized Jacobian (see [3, Proposition 2.6.2 (e)]) yield

$$\partial\phi(g(x), z)^T \subseteq \partial\varphi(g_1(x), z_1) \times \cdots \times \partial\varphi(g_m(x), z_m).$$

If i is such that $(g_i(x), z_i) \neq (0, 0)$, then φ is continuously differentiable at $(g_i(x), z_i)$ and, again, the expression of the $(n + p + i)$th column of H readily follows. If instead $(g_i(x), z_i) = (0, 0)$ then, using the theorem on the generalized gradient of a composite function (see [3, Theorem 2.3.9 (iii)]) and recalling that the generalized gradient of the Euclidean norm evaluated at the origin is the closed unit ball, we get

$$\partial\varphi(g_i(x), z_i)^T \subseteq \{((\xi_i - 1)\nabla g_i(x)^T, 0, \ldots, 0, (\rho_i - 1)e_i^T) \mid \|(\xi_i, \rho_i)\| \leq 1\}. \qquad (8)$$

Hence, the proposition follows from the first part of the proof, (8) and the definition of $a_i(w)$ and $b_i(w)$. $\qquad \square$

The results of this section are strongly related to Robinson's definition of a strongly regular KKT-point [28]. The original definition of strong regularity is somewhat involved; since we will not use it explicitly, we do not report it here. Instead, we shall take advantage of some equivalent definitions due to Robinson [28] and Liu [21]. In order to state these equivalent formulations, we need some more notation, in particular index set definitions. These index sets depend on the point at hand, $w = (x, y, z) \in \mathbb{R}^{n+p+m}$. Some of them also depend on the particular element of the generalized Jacobian $\partial\Phi(w)$ at hand. This dependency will always be clear from the context and is therefore not explicitly indicated. Now let us define the following index sets :

$I := \{1, \ldots, m\},$

$I_0 := \{i \in I \mid g_i(x) = 0, z_i \geq 0\}, \quad I_> := \{i \in I \mid g_i(x) > 0, z_i = 0\}, \quad I_R := I \setminus (I_0 \cup I_>).$

Moreover, we also need the following further index sets:

$I_{00} := \{i \in I_0 \mid z_i = 0\}, \quad I_+ := \{i \in I_0 \mid z_i > 0\},$

$I_{01} := \{i \in I_{00} \mid \rho_i = 1\}, \quad I_{02} := \{i \in I_{00} \mid \rho_i < 1, \xi_i < 1\}, \quad I_{03} := \{i \in I_{00} \mid \xi_i = 1\},$

$I_{R2} := I_R \cup I_{02}$

(see Proposition 3.1 for the definitions of ξ_i and ρ_i; these numbers depend on the element of the generalized Jacobian of $\Phi(w)$). The following relationships between these index sets can easily be seen to hold:

$$I = I_0 \cup I_> \cup I_R, \qquad I_0 = I_{00} \cup I_+, \qquad I_{00} = I_{01} \cup I_{02} \cup I_{03}.$$

Taking into account the definition of these index sets and Proposition 3.1, every element $H \in \partial\Phi(w)^T$ has the following structure (the dependence on w is suppressed for simplicity):

$$H = \begin{pmatrix} \nabla_x L & \nabla h & -\nabla g_+ & -\nabla g_{01} & \nabla g_{R2}(D_a)_{R2} & 0 & 0 \\ \nabla h^T & 0 & 0 & 0 & 0 & 0 & 0 \\ -\nabla g_+^T & 0 & 0 & 0 & 0 & 0 & 0 \\ -\nabla g_{01}^T & 0 & 0 & 0 & 0 & 0 & 0 \\ -\nabla g_{R2}^T & 0 & 0 & 0 & (D_b)_{R2} & 0 & 0 \\ -\nabla g_{03}^T & 0 & 0 & 0 & 0 & -I & 0 \\ -\nabla g_>^T & 0 & 0 & 0 & 0 & 0 & -I \end{pmatrix}, \tag{9}$$

where $(D_a)_{R2}$ and $(D_b)_{R2}$ are negative definite diagonal matrices. Note that we abbreviated g_{I_+} etc. by g_+ etc. in (9). This notation will be used frequently throughout the paper.

It will also be useful to define some matrices which turn out to be closely related to the strong regularity condition:

$$M(J) := \begin{pmatrix} \nabla_x L & \nabla h & \nabla g_+ & \nabla g_J \\ -\nabla h^T & 0 & 0 & 0 \\ -\nabla g_+^T & 0 & 0 & 0 \\ -\nabla g_J^T & 0 & 0 & 0 \end{pmatrix},$$

where J is any subset of $I_{00} \cup I_R$. We are now in the position to summarize some known equivalent conditions for a KKT-point of VIP(X, F) to be strongly regular. The reader is referred to Robinson [28] and Liu [21] for the proofs. We point out that, as a simple consequence of Theorem 3.7, we will be able to give a new characterization of strong regularity.

Theorem 3.2 *Let $w^* = (x^*, y^*, z^*) \in \mathbb{R}^{n+p+m}$ be a KKT-point of VIP(X, F). Then the following assertions are equivalent:*

(a) *w^* is a strongly regular KKT-point.*

(b) *For every $J \subseteq I_{00}$, the matrices $M(J)$ have the same nonzero orientation.*

(c) *The gradients $\nabla h_i(x^*)$ $(i \in K)$ and $\nabla g_i(x^*)$ $(i \in I_0)$ are linearly independent, and the matrix $\nabla_x L(w^*)$ is coherently oriented on the cone*

$$C(w^*) := \{v \in \mathbb{R}^n \,|\, \nabla h(x^*)^T v = 0, \nabla g_{I_+}(x^*)^T v = 0, \nabla g_{I_{00}}(x^*)^T v \leq 0\}.$$

(d) *$M(\emptyset)$ is nonsingular and either I_{00} is empty or $M(I_{00})/M(\emptyset)$ is a P-matrix.*

By point (b) of this theorem the following definition seems very natural.

Definition 3.3 *The* Extended Strong Regularity (ESR) *condition is said to hold at a point w if the sign of the determinant of $M(J)$*

(i) is a nonzero constant for all index sets J such that $J \subseteq I_{00}$;

(ii) cannot have opposite signs if J is such that $J \subseteq I_{00} \cup I_R$ (but can possibly be 0).

More precisely, Definition 3.3 (i) means that

$$\text{sign}(\det M(J_1)) \cdot \text{sign}(\det M(J_2)) > 0$$

for every $J_1, J_2 \subseteq I_{00}$, whereas part (ii) is equivalent to

$$\text{sign}(\det M(J_1)) \cdot \text{sign}(\det M(J_2)) \geq 0$$

for every $J_1, J_2 \subseteq I_{00} \cup I_R$.

This definition and its name are motivated by the fact that if $w = w^*$ is a KKT-point, then $I_R = \emptyset$ and the ESR condition amounts to the requirement that, for all $J \subseteq I_{00}$, the sign of $M(J)$ is equal to the same nonzero constant $(-1$ or $+1)$. By Theorem 3.2 (b) this is equivalent to the strong regularity of w^*. Thus condition ESR can be viewed as an *algebraic extension* of the concept of strong regularity to nonstationary points.

One of the main results we can prove is the following.

Theorem 3.4 *Let* $w = (x, y, z) \in \mathbb{R}^{n+p+m}$ *be given, and suppose that the ESR condition holds at* w. *Then all elements in the generalized Jacobian* $\partial \Phi(w)$ *are nonsingular.*

Proof. Consider an arbitrary, but fixed element in $\partial \Phi(w)^T$. This element has the structure indicated in (9), and is obviously nonsingular if and only if the following matrix is nonsingular:

$$
\begin{pmatrix}
\nabla_x L & \nabla h & \nabla g_+ & \nabla g_{01} & \nabla g_{R2} & 0 & 0 \\
-\nabla h^T & 0 & 0 & 0 & 0 & 0 & 0 \\
-\nabla g_+^T & 0 & 0 & 0 & 0 & 0 & 0 \\
-\nabla g_{01}^T & 0 & 0 & 0 & 0 & 0 & 0 \\
-\nabla g_{R2}^T & 0 & 0 & 0 & (D_b)_{R2}(D_a)_{R2}^{-1} & 0 & 0 \\
-\nabla g_{03}^T & 0 & 0 & 0 & 0 & I & 0 \\
-\nabla g_>^T & 0 & 0 & 0 & 0 & 0 & I
\end{pmatrix}.
$$

In turn this matrix is nonsingular if and only if the following matrix is nonsingular:

$$
\begin{pmatrix}
\nabla_x L & \nabla h & \nabla g_+ & \nabla g_{01} & \nabla g_{R2} \\
-\nabla h^T & 0 & 0 & 0 & 0 \\
-\nabla g_+^T & 0 & 0 & 0 & 0 \\
-\nabla g_{01}^T & 0 & 0 & 0 & 0 \\
-\nabla g_{R2}^T & 0 & 0 & 0 & (D_b)_{R2}(D_a)_{R2}^{-1}
\end{pmatrix}. \tag{10}
$$

The matrix (10) can be written as the sum of the matrix $M(J)$, with $J = I_{01} \cup I_{R2}$, and of the diagonal matrix (the dimensions of the blocks are the same as in (10))

$$
D := \begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & (D_b)_{R2}(D_a)_{R2}^{-1}
\end{pmatrix}.
$$

10

We now recall that, given a square matrix $\bar{A}$ of dimension r and a diagonal matrix $\bar{D}$ (with the same dimension), one has (see [4, p. 60])

$$\det(\bar{D} + \bar{A}) = \sum_{\alpha} \det \bar{D}_{\alpha\alpha} \det \bar{A}_{\bar{\alpha}\bar{\alpha}},$$

where the summation ranges over all subsets α of $\{1, \ldots, r\}$ (with complements $\bar{\alpha} = \{1, \ldots, r\} \setminus \alpha$), and where it is assumed that the determinant of an "empty" matrix is 1. Exploiting this formula, the determinant of (10) can be written as

$$\det M(J) + \sum_{\alpha \subseteq I_{R2}} \det D_{\alpha\alpha} \det M(J)_{\bar{\alpha}\bar{\alpha}}, \tag{11}$$

where the first term corresponds to $\alpha = \emptyset$. Moreover, we have taken into account that if α contains an element which does not belong to I_{R2}, then the determinant of $D_{\alpha\alpha}$ is 0. Since the nonzero diagonal elements of the matrix D are all positive, it follows that the determinants of $D_{\alpha\alpha}$ in (11) are all positive. Then to show that the determinant of (10) is nonzero and hence to conclude the proof, it will now be sufficient to show that the determinants of $M(J)$ and of all $M(J)_{\bar{\alpha}\bar{\alpha}}$ in (11) never have opposite signs, and that at least one of them is nonzero. These matrices can be written as $M(J \setminus \alpha)$, with $\alpha \subseteq I_{R2}$ (empty set included). If $\alpha = I_{R2}$, then $J \setminus I_{R2} = I_{01} \subseteq I_{00}$ while, in general, for every $\alpha \subseteq I_{R2}$ we have $J \setminus \alpha \subseteq I_{00} \cup I_R$. The theorem then follows directly from the definition of the ESR condition. $\qquad \square$

Remark 3.5 This result has interesting algorithmic consequences. In particular it implies that the semismooth Newton method [24, 25] applied to the system of equations (4) is locally quadratically convergent to a KKT-point w^* of VIP(X, F) if w^* is strongly regular (see [9, 11] for details). Even in the particular case in which the variational inequality VIP(X, F) arises from a nonlinear program with inequality constraints, strong regularity compares favourably with the assumptions required by existing superlinear convergent methods with linear equations subproblems.

We see that the ESR condition is a sufficient condition for the nonsingularity of all the elements of the generalized Jacobian of Φ. If the point at hand is a KKT-point, then it turns out that this is also a necessary condition. To prove this result, we first need the following technical proposition.

Proposition 3.6 Let $w = (x, y, z) \in \mathbb{R}^{n+p+m}$ be an arbitrary vector and assume that all elements in the generalized Jacobian $\partial\Phi(w)$ are nonsingular. Then, given any subset $J \subseteq I_{00}$, there exists an element in $\partial\Phi(w)$ such that $I_{01} = J, I_{02} = \emptyset$ and $I_{03} = I_{00} \setminus J$.

Proof. In the first part of the proof we show that the nonsingularity of all elements in $\partial\Phi(w)$ implies the linear independence of the gradients $\nabla g_i(x)$ $(i \in I_{00})$. To this end, let

$\{x^k\}$ be any sequence converging to x, and define $w^k = (x^k, y^k, z^k)$, where $y^k = y$ for all k and

$$z_i^k = \begin{cases} z_i & \text{if } i \in I \setminus I_{00}, \\ 1/k & \text{if } i \in I_{00} \text{ and } g_i(x^k) = 0, \\ \sqrt{|g_i(x^k)|} & \text{if } i \in I_{00} \text{ and } g_i(x^k) \neq 0. \end{cases}$$

Then it is easy to see that the sequence $\{w^k\} = \{(x^k, y^k, z^k)\} \subset \mathbb{R}^{n+p+m}$ converges to w such that Φ is differentiable at each w^k and that $\{\nabla\Phi(w^k)\}$ converges to an element $H \in \partial\Phi(w)^T$ such that $I_{01} = I_{00}$ and $I_{02} = I_{03} = \emptyset$. But then, since H is nonsingular by assumption, the linear independence of the gradients $\nabla g_i(x)$ $(i \in I_{00})$ follows immediately from the general structure of the (transposed) generalized Jacobian given in (9).

For the second part of the proof, let $J \subseteq I_{00}$ be any given subset. We now build another sequence $\{w^k\} = \{(x^k, y^k, z^k)\} \subset \mathbb{R}^{n+p+m}$ converging to w such that Φ is differentiable at each w^k and $\{\nabla\Phi(w^k)\}$ converges to an element $H \in \partial\Phi(w)^T$ such that H enjoys the properties stated in the proposition. By the first part of the proof, we know that the gradients $\nabla g_i(x)$ $(i \in I_{00})$ are linearly independent. Hence it follows from the Implicit Function Theorem that there is a sequence $\{x^k\}$ converging to x such that $g_i(x^k) > 0$ for all $i \in I_{00}$ and all k. Then take $w^k = (x^k, y^k, z^k)$ with x^k being the kth iterate of the above sequence, $y^k = y$ for all k and

$$z_i^k = \begin{cases} z_i & \text{if } i \in I \setminus I_{00}, \\ \sqrt{g_i(x^k)} & \text{if } i \in J, \\ g_i(x^k)^2 & \text{if } i \in I_{00} \setminus J. \end{cases}$$

Note that $\{w^k\}$ tends to w and that Φ is differentiable at w^k because there is no index i such that the case $g_i(x^k) = z_i^k = 0$ occurs. Furthermore, using Proposition 3.1 and the continuity of the functions involved, it easily follows that $\{\nabla\Phi(w^k)\}$ converges to an element H having the desired properties. $\qquad\square$

Theorem 3.7 *Let $w^* = (x^*, y^*, z^*) \in \mathbb{R}^{n+p+m}$ be a KKT-point of VIP(X, F). Then the ESR condition is satisfied at w^* if and only if all elements in the generalized Jacobian $\partial\Phi(w^*)$ are nonsingular.*

Proof. The necessity part of the theorem is just Theorem 3.4. We therefore turn to the sufficiency part.

Assume that all elements in the generalized Jacobian $\partial\Phi(w^*)$ are nonsingular. We first prove that the nonzero orientation of all these matrices is the same. In fact, suppose, by contradiction, that there are $H_1, H_2 \in \partial\Phi(w^*)^T$ such that $\det(H_1) < 0$ and $\det(H_2) > 0$. Since the determinant is a continuous function, a convex combination $\bar{H}$ of H_1 and H_2 exists such that $\det(\bar{H})=0$. Because $\partial\Phi(w^*)$ is convex by definition, the singular matrix $\bar{H}$ belongs to $\partial\Phi(w^*)^T$, thus contradicting the nonsingularity of all elements in the generalized Jacobian $\partial\Phi(w^*)$.

Because w^* is a KKT-point, the subset I_R must be empty. Therefore, to be sure that the ESR condition holds, we only need that the matrices $M(J)$ have the same nonzero

orientation for all $J \subseteq I_{00}$. According to Proposition 3.6, for each $J \subseteq I_{00}$, we can find an $H(J) \in \partial\Phi(w^*)^T$ such that $J = I_{01}$ and $I_{03} = I_{00} \setminus I_{01}$. With regard to the structure of the elements of $\partial\Phi(w^*)^T$ indicated in (9), and recalling that $h : \mathbb{R}^n \to \mathbb{R}^p$ and $g : \mathbb{R}^n \to \mathbb{R}^m$, we obtain

$$0 \neq \mathrm{sign}(\det H(J)) = (-1)^{p+m}\mathrm{sign}\left(\det \begin{pmatrix} \nabla_x L & \nabla h & \nabla g_+ & \nabla g_J & 0 & 0 \\ -\nabla h^T & 0 & 0 & 0 & 0 & 0 \\ -\nabla g_+^T & 0 & 0 & 0 & 0 & 0 \\ -\nabla g_J^T & 0 & 0 & 0 & 0 & 0 \\ -\nabla g_{03}^T & 0 & 0 & 0 & I & 0 \\ -\nabla g_>^T & 0 & 0 & 0 & 0 & I \end{pmatrix} \right)$$

which, in turn, implies

$$0 \neq \mathrm{sign}(\det H(J)) = (-1)^{p+m}\mathrm{sign}\left(\det \begin{pmatrix} \nabla_x L & \nabla h & \nabla g_+ & \nabla g_J \\ -\nabla h^T & 0 & 0 & 0 \\ -\nabla g_+^T & 0 & 0 & 0 \\ -\nabla g_J^T & 0 & 0 & 0 \end{pmatrix} \right)$$

$$= (-1)^{p+m}\mathrm{sign}(\det M(J)).$$

Since, in view of the first part of this proof, all matrices $H \in \partial\Phi(w^*)^T$ have the same nonzero orientation, it follows that the ESR condition is satisfied. $\square$

As we already observed, if $w = w^*$ is a KKT-point, then the ESR condition is equivalent to the strong regularity condition. We therefore get as an immediate consequence of Theorem 3.7 the following new characterization of a strongly regular KKT-point.

Corollary 3.8 *A KKT-point w^* of VIP(X, F) is strongly regular if and only if all the elements in the generalized Jacobian $\partial\Phi(w^*)$ are nonsingular.*

We now illustrate with some examples the results obtained so far.

Example 3.9 Consider VIP(X, F) with $n = 1$, $p = 0$ and $m = 1$, $F(x) = x - 1$ and $g_1(x) = x$. For this problem we have

$$\Phi(x, z) = \begin{pmatrix} x - 1 - z \\ \sqrt{x^2 + z^2} - x - z \end{pmatrix}.$$

Consider the point $(x, z) = (0, 0)$; in this point $I_{00} = \{1\}$ while $I_+ = I_R = \emptyset$. We have

$$\det M(\emptyset) = \det(1) = 1, \qquad \det M(\{1\}) = \det \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} = 1,$$

so that the ESR condition is satisfied. From the expression of Φ it can be easily seen that the elements of $\partial\Phi(0, 0)$ are given by

$$\begin{pmatrix} 1 & -1 \\ \alpha - 1 & \beta - 1 \end{pmatrix} \tag{12}$$

with $(\alpha, \beta)^T$ belonging to the generalized gradient of $\sqrt{x^2 + z^2}$ in $(0,0)$. Therefore, $\|(\alpha, \beta)\| \le 1$ so that all the matrices in (12) are nonsingular, as expected.

Example 3.10 Consider VIP(X, F) with $n = 1$, $p = 0$ and $m = 1$, $F(x) = -x$ and $g_1(x) = x$. For this problem we have

$$\Phi(x, z) = \left(\begin{array}{c} -x - z \\ \sqrt{x^2 + z^2} - x - z \end{array} \right).$$

Consider the KKT-point $(x, z) = (0, 0)$; in this point $I_{00} = \{1\}$ while $I_+ = I_R = \emptyset$. We have

$$\det(M(\emptyset)) = \det(-1) = -1, \qquad \det(M(\{1\})) = \det \left(\begin{array}{cc} -1 & 1 \\ -1 & 0 \end{array} \right) = 1,$$

so that the ESR condition is not satisfied. From the expression of Φ it can be easily seen that the elements of $\partial\Phi(0, 0)$ are given by

$$\left(\begin{array}{cc} -1 & -1 \\ \alpha - 1 & \beta - 1 \end{array} \right) \tag{13}$$

with $(\alpha, \beta)^T$ belonging to the generalized gradient of $\sqrt{x^2 + z^2}$ in $(0, 0)$. If we take $\alpha = \beta$ such that $\|(\alpha, \beta)\| \le 1$, e.g., $\alpha = \beta = 1/\sqrt{2}$, we see that the corresponding element in (13) is singular, whereas it is nonsingular whenever $\alpha \ne \beta$.

Motivated by Theorem 3.2, our aim is now to obtain conditions which are necessary and sufficient for ESR to hold by extending those of points (c) and (d) in Theorem 3.2.

We shall first consider conditions extending (c). To this end, we introduce the following notation:

$$C(w; I_1) := \{v \in \mathbb{R}^n \,|\, \nabla h(x)^T v = 0, \nabla g_{I_+}(x)^T v = 0, \nabla g_{I_1}(x)^T v \le 0\},$$

where I_1 is a subset of $I \setminus I_+$. Note that, with this notation, the polyhedral cone in point (c) of Theorem 3.2 corresponds to $C(w; I_{00})$.

Theorem 3.11 *Let $w = (x, y, z) \in \mathbb{R}^{n+p+m}$ be given. Then the ESR condition is satisfied at w if and only if the following three conditions hold:*

(a) The vectors $\nabla h_i(x)$ $(i \in K)$ and $\nabla g_i(x)$ $(i \in I_0)$ are linearly independent;

(b) the matrix $\nabla_x L(w)$ is positively (negatively) coherently oriented over $C(w; I_{00})$;

(c) the matrix $\nabla_x L(w)$ is positively (negatively) semicoherently oriented over $C(w; I_{00} \cup \tilde{I})$ for every $\tilde{I} \subseteq I_R$ such that the vectors $\nabla h_i(x)$ $(i \in K)$ and $\nabla g_i(x)$ $(i \in I_0 \cup \tilde{I})$ are linearly independent.

Proof. By Proposition 2.1 (a), (b), all and only the faces of $C(w; I_{00})$ can be obtained as $L(J) \cap C(w; I_{00})$, for all $J \subseteq I_{00}$, where

$$L(J) := \{v \in \mathbb{R}^n | \nabla h(x)^T v = 0, \nabla g_{I_+}(x)^T v = 0, \nabla g_J(x)^T v = 0\}.$$

Then, by condition (a) and Proposition 2.1 (c), all and only the linear subspaces generated by the faces of $C(w; I_{00})$ can be obtained as $L(J)$, for all $J \subseteq I_{00}$. Taking into account that $L(J) = Ker\left((\nabla h(x), \nabla g_{I_+}(x), \nabla g_J(x))^T\right)$ for $J \subseteq I_{00}$, we have, by Lemma 2.4 (take $B = (\nabla h(x), \nabla g_{I_+}(x), \nabla g_J(x)))$ and (a), that the sign of the determinant of $M(J)$ is equal to that of the section of $\nabla_x L(w)$ in the subspace $L(J)$, so that part (i) of the definition of ESR follows by (b).

Assume now that $J \subseteq I_{00} \cup I_R$. If the vectors $\nabla h_i(x)$ $(i \in K)$ and $\nabla g_i(x)$ $(i \in I_+ \cup J)$ are not linearly independent, then $M(J)$ is obviously singular. If, otherwise, the set of vectors is linearly independent then we can consider the cone $C(w; J)$, and reason as before so that, by (c), we have that (ii) of ESR holds.

The opposite implication can easily be obtained by reversing the above arguments. $\square$

Remark 3.12 In the statement of Theorem 3.11 we have stressed the fact that the orientations on the faces of the cones considered in points (b) and (c) have to be the same, when not 0, however, this is superfluous, since this fact holds automatically.

In condition (c) of Theorem 3.11 we could have equivalently considered only those subsets $\tilde{I}$ for which there is no other subset $\tilde{I}'$ such that $\tilde{I} \subseteq \tilde{I}' \subseteq I_R$ and the vectors $\nabla h_i(x) \, (i \in K)$ and $\nabla g_i(x) \, (i \in I_0 \cup \tilde{I}')$ are linearly independent. In particular, if the vectors $\nabla h_i(x) \, (i \in K)$ and $\nabla g_i(x) \, (i \in I_0 \cup I_R)$ are linearly independent, then we could just consider the cone $C(w; I_{00} \cup I_R)$.

As a corollary, we get the following useful result.

Corollary 3.13 *Let $w = (x, y, z) \in \mathbb{R}^{n+p+m}$ be any given vector. Assume that*

(a) *the gradients $\nabla h_i(x) \, (i \in K)$ and $\nabla g_i(x) \, (i \in I_0)$ are linearly independent;*

(b) *$v^T \nabla_x L(w) v > 0$ for all $v \in \mathbb{R}^n, v \neq 0$, such that $\nabla h_i(x)^T v = 0 \, (i \in K)$ and $\nabla g_i(x)^T v = 0 \, (i \in I_+)$.*

Then all elements $H \in \partial \Phi(w)$ are nonsingular.

Proof. Condition (a) of Corollary 3.13 is obviously the same as (a) of Theorem 3.11. On the other hand, condition (b) of the corollary requires that $\nabla_x L(w)$ is positive definite on a subspace that contains all the cones considered in assumptions (b) and (c) of Theorem 3.11. By Proposition 2.3 this implies that (b) and (c) of Theorem 3.11 are fulfilled. $\square$

Note that, if $w = w^*$ is a KKT-point of VIP(X, F), then the above corollary reduces to a recent result by Jiang [18], see also Fischer [14].

We now pass to consider necessary and sufficient conditions for ESR to hold which are inspired by (d) in Theorem 3.2.

Theorem 3.14 *The ESR condition is satisfied in a point $w = (x, y, z) \in \mathbb{R}^{n+p+m}$ if and only if the following three conditions hold:*

  (a) $M(\emptyset)$ is nonsingular;

  (b) either I_{00} is empty or $M(I_{00})/M(\emptyset)$ is a P-matrix;

  (c) either $I_{00} \cup I_R$ is empty or $M(I_{00} \cup I_R)/M(\emptyset)$ is a P_0-matrix.

Proof. We only consider the case that $M(\emptyset)$ is nonsingular since, otherwise, neither the ESR condition nor (a) of Theorem 3.14 is satisfied. The principal submatrices of $M(I_{00})/M(\emptyset)$ are exactly the matrices given by $M(J)/M(\emptyset)$, for all $J \subseteq I_{00}$. Using the determinantal formula for the Schur complement, we get

$$\det M(J) = \det\left(M(J)/M(\emptyset)\right)\det M(\emptyset).$$

Hence the sign of $\det M(J)$ is equal to the sign of $\det M(\emptyset)$ for all $J \subseteq I_{00}$ if and only if all the principal minors of $M(I_{00})/M(\emptyset)$ are positive. This fact establishes the equivalence of condition (b) in Theorem 3.14 and of (i) in the ESR condition. The equivalence of (c) with (ii) can be shown in a similar way. We omit the details here. $\square$

In some cases (see, e.g., Proposition 4.2) it may also be of interest to determine conditions ensuring that at least one element in the generalized Jacobian $\partial\Phi(w)$ is nonsingular. We devote the last part of this section to this issue. We naturally expect that a weaker condition than ESR will be needed for establishing this result. Actually, the following condition will be useful.

Definition 3.15 *The* Extended Weak Regularity (EWR) *condition is said to hold at a point $w = (x, y, z) \in \mathbb{R}^{n+p+m}$ if a possibly empty index set $\tilde{J}$, with $\tilde{J} \subseteq I_{00}$, exists such that*

  (i) the sign of the determinant of $M(\tilde{J})$ is $+1$ (-1);

  (ii) each sign of the determinants of $M(J)$ is either 0 or $+1$ (-1) for all $J \subseteq I_{00} \cup I_R$.

Note that the EWR and the ESR conditions coincide if $I_{00} = \emptyset$. In particular, if $w = w^*$ is a KKT-point, the two conditions coincide if strict complementarity holds. In general, however, the EWR condition is weaker than the ESR condition. We illustrate this by a simple example.

Example 3.16 Consider VIP(X, F) with $n = 2$, $p = 0$, $m = 1$, $F(x) := (x_1, x_2^2)^T$ and $g_1(x) = x_2$. Consider the KKT-point $(x_1, x_2, z) = (0, 0, 0)$. We have $I_{00} = \{1\}$, while $I_+ = I_R = \emptyset$. We have

$$M(\emptyset) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad M(\{1\}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Therefore, $\det(M(\emptyset)) = 0$ and $\det(M(\{1\})) = 1$, so that the ESR condition is not satisfied, whereas the EWR condition obviously holds by taking $\tilde{J} = \{1\}$.

The following result is the direct counterpart of Theorem 3.11 and gives a complete characterization of the EWR condition in terms of the (semi)coherent orientation of a certain matrix.

Theorem 3.17 *Let $w = (x, y, z) \in \mathbb{R}^{n+p+m}$ be given. Then the EWR condition is satisfied at w if and only if there is a (possibly empty) index set $\tilde{J} \subseteq I_{00}$ such that the following three conditions hold:*

(a) *the gradients $\nabla h_i(x)$ $(i \in K)$ and $\nabla g_i(x)$ $(i \in I_+ \cup \tilde{J})$ are linearly independent;*

(b) *the orientation of the matrix $\nabla_x L(w)$ on the subspace*

$$\{v \in \mathbb{R}^n | \nabla h(x)^T v = 0, \nabla g_{I_+}(x)^T v = 0, \nabla g_{\tilde{J}}(x)^T v = 0\}$$

is positive (negative).

(c) *the matrix $\nabla_x L(w)$ is positively (negatively) coherently oriented over $C(w; I_{00} \cup \tilde{I})$ for every $\tilde{I} \subseteq I_R$ such that the vectors $\nabla h_i(x)$ $(i \in K)$ and $\nabla g_i(x)$ $(i \in I_0 \cup \tilde{I})$ are linearly independent.*

Proof. The proof is almost identical to the one of Theorem 3.11, so we omit the details. $\square$

Moreover, we can prove the following result.

Theorem 3.18 *Let $w = (x, y, z) \in \mathbb{R}^{n+p+m}$ be given, and suppose that the EWR condition holds at w. Then the generalized Jacobian $\partial \Phi(w)$ contains at least one nonsingular element.*

Proof. Let $\tilde{J} \subseteq I_{00}$ denote any index set which satisfies the conditions (i) and (ii) in Definition 3.15. We first show that there is a matrix $\tilde{H} \in \partial \Phi(w)^T$ with $I_{01} = \tilde{J}$. To this end, let us define a sequence $\{w^k\} = \{(x^k, y^k, z^k)\} \subset \mathbb{R}^{n+p+m}$ such that $(x^k, y^k) := (x, y)$ for all k and such that the sequence $\{z^k\}$ satisfies the following conditions: $z^k \to z$ and, for all k, $z_i^k > 0$ if $i \in \tilde{J}$, $z_i^k < 0$ if $i \in I_{00} \setminus \tilde{J}$ and $z_i^k = z_i$ if $i \in I \setminus I_{00}$. Note that, since $\tilde{J} \subseteq I_{00}$, it is always possible to find a sequence $\{z^k\}$ satisfying these conditions. Then, it can easily be verified that $\{w^k\}$ converges to w, that Φ is differentiable at each w^k and that we can assume, renumbering if necessary, that the sequence $\{\nabla \Phi(w^k)\}$ converges to a certain element $\tilde{H} \in \partial \Phi(w)^T$. Taking into account Proposition 3.1 it further follows that, for each w^k, $b_i(w^k) = 0$ if and only if $i \in \tilde{J}$. With regard to the definition of I_{01} we obtain that $I_{01} = \tilde{J}$ is valid for $\tilde{H}$.

The proof of the theorem is now almost a repetition of that of Theorem 3.4 in which we take into account that $I_{01} = \tilde{J}$. We report here the complete proof for the sake of clarity.

We shall show that the particular matrix $\tilde{H}$ constructed above is nonsingular. This matrix has the structure indicated in (9), and obviously it is nonsingular if and only if

the following matrix is nonsingular:

$$
\begin{pmatrix}
\nabla_x L & \nabla h & \nabla g_+ & \nabla g_{\tilde{J}} & \nabla g_{R2} & 0 & 0 \\
-\nabla h^T & 0 & 0 & 0 & 0 & 0 & 0 \\
-\nabla g_+^T & 0 & 0 & 0 & 0 & 0 & 0 \\
-\nabla g_{\tilde{J}}^T & 0 & 0 & 0 & 0 & 0 & 0 \\
-\nabla g_{R2}^T & 0 & 0 & 0 & (D_b)_{R2}(D_a)_{R2}^{-1} & 0 & 0 \\
-\nabla g_{03}^T & 0 & 0 & 0 & 0 & I & 0 \\
-\nabla g_{>}^T & 0 & 0 & 0 & 0 & 0 & I
\end{pmatrix},
$$

where, we recall, $I_{01} = \tilde{J}$ holds. Reasoning as in the proof of Theorem 3.4, we can see that this matrix is nonsingular if and only if

$$
\det M(J) + \sum_{\alpha \subseteq I_{R2}} \det D_{\alpha\alpha} \det M(J)_{\bar{\alpha}\bar{\alpha}} \neq 0 \tag{14}
$$

with $J = \tilde{J} \cup I_{R2}$. Again, as in the proof of Theorem 3.4, to show that (14) holds it will now be sufficient to show that the determinants of $M(J)$ and of all $M(J)_{\bar{\alpha}\bar{\alpha}}$ in (14) never assume opposite signs, and that at least one of them is nonzero. These matrices can be written as $M(J \setminus \alpha)$, with $\alpha \subseteq I_{R2}$ (empty set included). If $\alpha = I_{R2}$, then $J \setminus I_{R2} = I_{01} = \tilde{J}$ while, in general, for every $\alpha \subseteq I_{R2}$ we have $J \setminus \alpha \subseteq I_{00} \cup I_R$. The thesis then follows directly from the EWR condition (Definition 3.15). $\qquad \square$

Example 3.19 Consider the problem introduced in Example 3.16. We have

$$
\Phi(x_1, x_2, z) = \begin{pmatrix} x_1 \\ x_2^2 - z \\ \sqrt{x_2^2 + z^2} - x_2 - z \end{pmatrix},
$$

so that, in $(x_1, x_2, z) = (0, 0, 0)$, we have

$$
\partial\Phi(0, 0, 0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & \alpha - 1 & \beta - 1 \end{pmatrix},
$$

where $(\alpha, \beta)^T$ belongs to the generalized gradient of $\sqrt{x_2^2 + z^2}$ in $(0, 0)$. It is then easy to see that $\partial\Phi(0, 0, 0)$ contains a nonsingular element; for example, take $(\alpha, \beta) = (0, 1)$. On the other hand, it also contains a singular element; to this end, take $(\alpha, \beta) = (1, 0)$.

Remark 3.20 We proved in Theorem 3.7 that the ESR condition is satisfied in a KKT-point w if and only if all the elements in $\partial\Phi(w)$ are nonsingular. On the other hand, we proved in Theorem 3.18 that, if the EWR condition is satisfied in a (KKT-) point w, at least one nonsingular element exists in $\partial\Phi(w)$. It is then natural to ask whether the existence of a nonsigular element in $\partial\Phi(w)$ implies the EWR condition. The answer is in the negative, and this is shown by Example 3.10. In fact, the problem in this example does not satisfy the EWR condition in the KKT-point $(x, z) = (0, 0)$ considered there. However, it is sufficient to consider $(\alpha, \beta) = (0, 1)$ in (13) to obtain a nonsingular generalized Jacobian of $\partial\Phi(0, 0)$.

We next state a simple consequence of Theorem 3.18.

Corollary 3.21 *Let $w = (x, y, z) \in \mathbb{R}^{n+p+m}$ be an arbitrary vector. Assume that*

(a) *the gradients $\nabla h_i(x)$ $(i \in K)$ and $\nabla g_i(x)$ $(i \in I_+)$ are linearly independent;*

(b) *$v^T \nabla_x L(w) v > 0$ for all $v \in \mathbb{R}^n, v \neq 0$, such that $\nabla h_i(x)^T v = 0$ $(i \in K)$ and $\nabla g_i(x)^T v = 0$ $(i \in I_+)$.*

Then there is at least one nonsingular element in $\partial \Phi(w)$.

Proof. We first show that assumptions (a) and (b) imply that the EWR condition is satisfied at w. For $\tilde{J} := \emptyset$ assumption (a) of Theorem 3.17 is equivalent to assumption (a) of Corollary 3.21. Furthermore, all the cones which appear in assumptions (b) and (c) of Theorem 3.17 are contained in the linear space

$$\{v \in \mathbb{R}^n | \nabla h(x)^T v = 0, \nabla g_{I_+}(x)^T v = 0\}.$$

Hence, by Proposition 2.3, assumptions (b) and (c) of Theorem 3.17 follow from assumption (b) of this Corollary. Therefore the EWR condition is satisfied at w because of Theorem 3.17. Thus the assertion follows immediately from Theorem 3.18. □

Some explanations to Corollaries 3.13 and 3.21 are in order. If the arbitrary vector w in these results is a KKT-point of $\mathrm{VIP}(X, F)$ and if we borrow some terminology from the optimization literature, then Corollary 3.13 states that all elements in the generalized Jacobian $\partial \Phi(w)$ are nonsingular if

(a) the gradients of all active constraints are linearly independent and

(b) the strong second order condition is satisfied.

On the other hand, Corollary 3.21 guarantees the existence of at least one nonsingular element in $\partial \Phi(w)$ if condition (a) is replaced by the weaker assumption that only the gradients belonging to the strongly active constraints are linearly independent. We stress, however, that the vector w in our results is an arbitrary vector, not necessarily a KKT-point, and that we have obtained these corollaries as simple consequences of more general results.

In the following proposition, we state that if the strong second order condition is weakened to the second order condition, but if the gradients of all binding constraints are linearly independent, then the assertion of Corollary 3.21 remains true. The proof of this result is omitted for the sake of brevity.

Proposition 3.22 *Let $w = (x, y, z) \in \mathbb{R}^{n+p+m}$ be any given vector. Assume that*

(a) *the gradients $\nabla h_i(x)$ $(i \in K)$ and $\nabla g_i(x)$ $(i \in I_0)$ are linearly independent;*

(b) *$v^T \nabla_x L(w) v > 0$ for all $v \in \mathbb{R}^n, v \neq 0$, such that $\nabla h_i(x)^T v = 0$ $(i \in K)$ and $\nabla g_i(x)^T v = 0$ $(i \in I_0)$.*

Then there is at least one nonsingular element $H \in \partial \Phi(w)$.

4 Stationary Conditions for Ψ

In this section, after noting that Ψ is continuously differentiable, we study conditions which guarantee that $\nabla\Psi(w^*) = 0$ implies $\Psi(w^*) = 0$, i.e., that a stationary point of the merit function Ψ is a KKT-point of VIP(X, F). After considering the general case we will deal with variational inequalities on polyhedral sets X for which we are able to obtain stronger results. These conditions play a key role in the algorithms which seek KKT-points of VIP(X, F) through the minimization of Ψ. In fact, by using unconstrained minimization algorithms we can only hope to find stationary points of Ψ. Hence, the weaker the conditions which guarantee that a stationary point of Ψ is a global solution, the more effective the unconstrained minimization of Ψ will be.

As a first step, we state the smoothness of Ψ and give an expression of its gradient.

Proposition 4.1 *If F is a C^1-function and g, h are C^2-functions, then Ψ is continuously differentiable and $\nabla\Psi(w) = H\Phi(w)$ for every $H \in \partial\Phi(w)^T$.*

Proof. By the chain rule for the composition of a strictly differentiable function and a Lipschitz continuous function (see [3, Theorem 2.6.6]) we have $\partial\Psi(w) = \partial\Phi(w)^T\Phi(w)$. It is easy to check that $\partial\Phi(w)^T\Phi(w)$ is single valued everywhere because the zero components of $\Phi(w)$ cancel the "multivalued columns" of $\partial\Phi(w)^T$. Therefore we have, using the Corollary to Theorem 2.2.4 in [3], that Ψ is continuously differentiable and that $\nabla\Psi(w) = H\Phi(w)$ for every $H \in \partial\Phi(w)^T$. $\qquad\square$

4.1 General X

We first state a simple but useful consequence of Proposition 4.1.

Proposition 4.2 *Let $w^* = (x^*, y^*, z^*) \in \mathbb{R}^{n+p+m}$ be a stationary point of Ψ. Then w^* is a KKT-point of VIP(X, F) if there is at least one nonsingular element in the generalized Jacobian $\partial\Phi(w^*)$.*

Proof. Since w^* is a stationary point of Ψ, we have, by Proposition 4.1, that

$$0 = \nabla\Psi(w^*) = H\Phi(w^*) \tag{15}$$

for any $H \in \partial\Phi(w^*)^T$. But, by the assumptions made, we have that the matrix H can be chosen to be nonsingular. By (15), this implies $\Phi(w^*) = 0$ so that w^* is a KKT-point of VIP(X, F). $\qquad\square$

By the results of the previous section, this proposition implies, in particular, that if w^* is a stationary point of Ψ such that the ESR or the EWR condition holds at w^*, then w^* is a KKT-point of the variational inequality problem. It is, however, possible to give further conditions which still guarantee that every stationary point of the function Ψ is a KKT-point of VIP(X, F).

Theorem 4.3 *Let $w^* = (x^*, y^*, z^*) \in \mathbb{R}^{n+p+m}$ be a stationary point of Ψ. Assume that*

(a) $\nabla_x L(w^*)$ is positive semidefinite on $\mathbb{R}^n$;

(b) $\nabla_x L(w^*)$ is positive definite on $Ker(\nabla h(w^*)^T) \cap Ker(\nabla g(w^*)^T)$;

and either of the following two conditions holds:

(c1) $\nabla h(x^*)$ has full column rank;

(c2) h is an affine function and X is nonempty.

Then $\Psi(w^*) = 0$, i.e., w^* is a KKT-point of $VIP(X, F)$.

Proof. Suppose that $\nabla \Psi(w^*) = 0$. Using Propositions 4.1 and 3.1, this can be written as

$$\nabla_x L(w^*) L^* + \nabla h(x^*) h^* + \nabla g(x^*) D_a(w^*) \phi^* = 0, \tag{16}$$

$$\nabla h(x^*)^T L^* = 0, \tag{17}$$

$$-\nabla g(x^*)^T L^* + D_b(w^*) \phi^* = 0, \tag{18}$$

where (L^*, h^*, ϕ^*) denotes $(L(w^*), h(x^*), \phi(g(x^*), z^*))$. Moreover, if $\phi_i^* = 0$ then the value of the ith diagonal elements of the matrices $D_a(w^*)$ and $D_b(w^*)$ are immaterial to the value of the left-hand side of (16) and (18). Therefore we can assume, without loss of generality, that these diagonal elements are all equal to -1. With this convention we have that the diagonal matrices $D_a(w^*)$ and $D_b(w^*)$ are nonsingular with negative diagonal elements (note that a diagonal element of these two matrices can be 0 only if the corresponding element of ϕ^* is equal to 0, cf. Proposition 3.1).

We can now solve (18) with respect to ϕ^* and substitute in (16), thus obtaining

$$\left(\nabla_x L(w^*) + \nabla g(x^*) D \nabla g(x^*)^T \right) L^* + \nabla h(x^*) h^* = 0, \tag{19}$$

where we have indicated by D the positive diagonal matrix $D_a(w^*) D_b(w^*)^{-1}$. Premultiplying (19) by $(L^*)^T$ and taking (17) into account, we obtain

$$(L^*)^T \left(\nabla_x L(w^*) + \nabla g(x^*) D \nabla g(x^*)^T \right) L^* = 0. \tag{20}$$

Since D is positive definite and the matrix $\nabla g(x^*) D \nabla g(x^*)^T$ is positive semidefinite, it is easy to see, using assumptions (a) and (b) and the fact that $L^* \in Ker(\nabla h(x^*)^T)$ by (17), that (20) implies $L^* = 0$. Hence $\phi^* = 0$ follows from (18), and (19) becomes

$$\nabla h(x^*) h^* = 0. \tag{21}$$

We now want to show that $h^* = 0$. Assume first that (c1) holds. Then, since $\nabla h(x^*)$ has full column rank, we immediately get $h^* = h(x^*) = 0$ from (21). Assume now, instead, that (c2) holds. Since h is an affine function, the norm function $\frac{1}{2} \| h \|^2$ is convex with

gradient equal to the left-hand side of (21). Hence (21) implies that x^* is a minimum point of the norm function $\frac{1}{2}\| h \|^2$. Since the set X is nonempty, this in turn implies $h^* = h(x^*) = 0$.

Therefore, in both cases, $(L^*, h^*, \phi^*) = 0$, hence $\Phi(w^*) = 0$, so that w^* is a KKT-point of VIP(X, F). $\square$

4.2 Polyhedral X

A particularly interesting case is when X is *polyhedral*, i.e.,

$$h(x) = Ax - b, \qquad g(x) = Cx - d,$$

with $A \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^p$, $C \in \mathbb{R}^{m \times n}$ and $d \in \mathbb{R}^m$. A first important consequence of the polyhedrality of X is that the KKT-conditions (3) are both necessary and sufficient, so that any solution of the variational inequality problem corresponds to a certain solution of the nonsmooth system of equations (4) and vice versa. Furthermore, in this case some nice geometrical interpretations can be given to the conditions employed in Theorem 4.3. In particular we note that, in this case, $Ker(A) \cap Ker(C)$ is the lineality space of the set X. So, if this lineality space reduces to $\{0\}$, then conditions (a) and (b) of Theorem 4.3 collapse to the requirement that $\nabla_x L(w)$ is just positive semidefinite. This discussion can be summarized in the following theorem, where we have taken into account that in this case the Jacobian of the Lagrangian L is just $\nabla F(x)^T$.

Theorem 4.4 *Assume that X is polyhedral. Let $w^* = (x^*, y^*, z^*) \in \mathbb{R}^{n+p+m}$ be a stationary point of Ψ. Then w^* is a solution of VIP(X, F) if*

(a1) $\nabla F(x^*)$ is positive semidefinite on $\mathbb{R}^n$;

(b1) $\nabla F(x^*)$ is positive definite on the lineality space of X.

In particular, if either

(1) the set X is bounded or

(2) the set X is contained in the nonnegative orthant $\mathbb{R}^n_+$

is satisfied, then assumption (a1) is sufficient for the theorem to hold.

Proof. (a1) and (b2) are just (a) and (b) rewritten taking into account that, since the constraints are affine, then $\nabla_x L(w) = \nabla F(x)$ and recalling that the lineality space of X is $Ker(A) \cap Ker(C)$. The second part of the Theorem follows by observing that the conditions (1) or (2) obviously imply that the set X contains no lines and hence that the lineality space of X is $\{0\}$. $\square$

It may be interesting to note the following facts. Suppose that $X = X_1 \cap X_2$, where X_1 is convex and X_2 is polyhedral. If the set X_2 satisfies one of the conditions (1) or (2)

or, more in general, if its lineality space is $\{0\}$, then it is obvious that $Ker(\nabla h(x^*)^T) \cap Ker(\nabla g(x^*)^T) = \{0\}$ (g and h are the constraints defining X) so that condition (b) in Theorem 4.3 becomes superfluous. In turn, the previous observation can be applied in the following way. Suppose that we have a VIP(X, F) with X (described by general constraints) contained, for example, in the nonnegative orthant. In order to guarantee that a stationary point is a global minimum of Ψ we need conditions (a) and (b) of Theorem 4.3. However we can consider the variational inequality problem VIP($F, X \cap \{x \,|\, x \geq 0\}$). This latter problem is obviously equivalent to the original one, nevertheless, as observed before, requirement (b) of Theorem 4.3 is now superfluous.

In [10] we considered the case, often encountered in the applications, in which X is a rectangle. Then, weaker conditions can be obtained. In particular we showed in [10] that if $X = \mathbb{R}^n_+$, i.e., if the variational inequality problem reduces to a complementarity problem, then F being a P_0-function is a sufficient condition for guaranteeing that every stationary point of Ψ is a global solution.

References

[1] A. BACHEM AND W. KERN: *Linear Programming Duality. An Introduction to Programming Duality.* Springer-Verlag, Heidelberg, 1992.

[2] M. CAO AND M.C. FERRIS: *A pivotal method for affine variational inequalities.* Mathematics of Operations Research 21, 1996, pp. 44–64.

[3] F.H. CLARKE: *Optimization and Nonsmooth Analysis.* John Wiley and Sons, New York, 1983 (reprinted by SIAM, Philadelphia, 1990).

[4] R.W. COTTLE, J.-S. PANG AND R.E. STONE: *The Linear Complementarity Problem.* Academic Press, Boston, 1992.

[5] T. DE LUCA, F. FACCHINEI AND C. KANZOW: *A semismooth equation approach to the solution of nonlinear complementarity problems.* Mathematical Programming 75, 1996, pp. 407–439.

[6] S. P. DIRKSE AND M. C. FERRIS: *The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems.* Optimization Methods and Software 5, 1995, pp. 123–156.

[7] B. C. EAVES: *On the basic theorem of complementarity.* Mathematical Programming 1, 1971, pp. 68–75.

[8] F. FACCHINEI AND C. KANZOW: *A nonsmooth inexact Newton method for the solution of large-scale nonlinear complementarity problems.* Mathematical Programming 76, 1997, pp. 493–512.

[9] F. Facchinei, A. Fischer and C. Kanzow: *Inexact Newton methods for semismooth equations with applications to variational inequality problems.* In: *Nonlinear Optimization and Applications,* G. Di Pillo and F. Giannessi, editors, Plenum Press, New York, 1996, pp. 125–139.

[10] F. Facchinei, A. Fischer and C. Kanzow: *A semismooth Newton method for variational inequalities: The case of box constraints.* In: *Complementarity and Variational Problems. State of the Art,* M. C. Ferris and J.-S. Pang, editors, SIAM, Philadelphia, 1997, pp. 76–90.

[11] F. Facchinei, A. Fischer and C. Kanzow: *A semismooth Newton method for variational inequalities: Theoretical results and preliminary numerical results.* Technical Report 102, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany, 1995.

[12] F. Facchinei and J. Soares: *A new merit function for nonlinear complementarity problems and a related algorithm.* SIAM Journal on Optimization 7, 1997, pp. 225–247.

[13] M.C. Ferris and J.-S. Pang: *Engineering and economic applications of complementarity problems.* SIAM Review, to appear.

[14] A. Fischer: *A special Newton-type optimization method.* Optimization 24, 1992, pp. 269–284.

[15] A. Fischer: *An NCP-function and its use for the solution of complementarity problems.* In: *Recent Advances in Nonsmooth Optimization,* D.Z. Du, L. Qi and R.S. Womersley, editors, World Scientific Publishers, Singapore, 1995, pp. 88–105.

[16] M. Fukushima: *Merit functions for variational inequality and complementarity problems.* In: *Nonlinear Optimization and Applications,* G. Di Pillo and F. Giannessi, editors, Plenum Press, New York, 1996, pp. 155–170.

[17] P.T. Harker and J.-S. Pang: *Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications.* Mathematical Programming 48, 1990, pp. 161–220.

[18] H. Jiang: *Local properties of solutions of nonsmooth variational inequalities.* Optimization 33, 1995, pp. 119–132.

[19] A.J. King and R.T. Rockafellar: *Sensitivity analysis for nonsmooth generalized equations.* Mathematical Programming 55, 1992, pp. 193–212.

[20] M. Kojima: *Strongly stable stationary solutions in nonlinear programs.* In: *Analysis and Computation of Fixed Points,* S.M. Robinson, editor, Academic Press, New York, New York, 1980, pp. 93–138.

[21] J. Liu: *Strong stability in variational inequalities.* SIAM Journal on Control and Optimization 33, 1995, pp. 725–749.

[22] A. Nagurney: *Network Economics: A Variational Inequality Approach.* Kluwer Academic Publishers, Boston, 1993.

[23] J.-S. Pang: *Complementarity Problems.* In: *Handbook of Global Optimization,* R. Horst and P.M. Pardalos, editors, Kluwer Academic Publisher, Dordrecht, 1995, pp. 271–338.

[24] J.-S. Pang and L. Qi: *Nonsmooth equations: motivation and algorithms.* SIAM Journal on Optimization 3, 1993, pp. 443–465.

[25] L. Qi and J. Sun: *A nonsmooth version of Newton's method.* Mathematical Programming 58, 1993, pp. 353–367.

[26] D. Ralph: *Global convergence of damped Newton's method for nonsmooth equations via the path search.* Mathematics of Operations Research, 1994, pp. 352–389.

[27] S.M. Robinson: *Generalized equations and their solutions. Part I: Basic theory.* Mathematical Programming Study 10, 1979, pp. 128–141.

[28] S.M. Robinson: *Strongly regular generalized equations.* Mathematics of Operations Research 5, 1980, pp. 43–62.

[29] S.M. Robinson: *Normal maps induced by linear transformations.* Mathematics of Operations Research 17, 1992, pp. 691–714.

[30] S.M. Robinson: *Homeomorphism conditions for normal maps of polyhedra.* In: *Optimization and Nonlinear Analysis,* A. Ioffe, M. Marcus and S. Reich, editors, Longman, Harlow, England, 1992, pp. 240–248.

[31] S.M. Robinson: *Nonsingularity and symmetry for linear normal maps.* Mathematical Programming 62, 1993, pp. 415–425.

[32] S.M. Robinson: *Sensitivity analysis of variational inequalities by normal-map techniques.* In: *Nonlinear Variational Inequalities and Network Equilibrium Problems,* F. Giannessi and A. Maugeri, editors, Plenum Press, New York, 1995, pp. 257–269.

[33] H. Sellami and S.M. Robinson: *Homotopies based on nonsmooth equations for solving nonlinear variational inequalities.* In: *Nonlinear Optimization and Applications,* G. Di Pillo and F. Giannessi, editors, Plenum Press, New York, 1996, pp. 329–343.

[34] B. Xiao and P.T. Harker: *A nonsmooth Newton method for variational inequalities, parts I and II.* Mathematical Programming 65, 1994, pp. 151–216.