Minimizing and Stationary Sequences of Merit Functions for Complementarity Problems and Variational Inequalities*

Masao Fukushima† and Jong-Shi Pang‡

September 19, 1995

Abstract

We study the relation between minimizing sequences and stationary sequences of two merit functions for nonlinear complementarity problems and variational inequalities, without assuming the boundedness of the sequences. Conditions are obtained under which these two types of sequences are equivalent. The results are applicable to complementarity problems and variational inequalities that do not necessarily possess solutions.

Keywords Complementarity problems, variational inequalities, merit functions, minimizing sequences.

AMS Subject Classifications 90C30, 90C33

1 Introduction

Since the work of Auchmuty [1] and Fukushima [12], merit functions have become an important tool for solving nonlinear complementarity problems (NCPs) and variational inequalities (VIs). The recent paper by Fukushima [13] gives a comprehensive survey of such merit functions and summarizes their basic properties; extensive references are also given therein.

A principal use of a merit function for NCPs and VIs is for the design of iterative descent methods for solving these problems. In the convergence analysis of such a method, the emphasis is typically placed on the accumulation points of the generated sequence of iterates and on establishing some desirable solution properties of such points under appropriate assumptions; the case where no accumulation point exists is usually not treated. This omission thus raises the question of whether some useful conclusions can still be obtained in the latter case. In the related setting of convex programming, the issue leads to the study of well-behaved optimization problems for which

*This work was based on research supported partly by a Japan-US collaborative grant awarded to the authors by the U.S. National Science Foundation under grant INT-9417085 and the Japan Society for the Promotion of Science. In addition, the first author's research was also partly supported by the Scientific Research Grant in Aid from the Ministry of Education, Science and Culture, Japan under grant 06650443. The second author's research was also partly supported by the National Science Foundation under grant CCR-9213739.

†Graduate School of Information Science, Nara Institute of Science and Technology, Takayama, Ikoma, Nara 630-01, Japan; Email: fuku@is.aist-nara.ac.jp

‡Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, Maryland 21218-2692, U.S.A; Email: jsp@vicp1.mts.jhu.edu

pioneering results have been obtained by Auslender, Crouzeix, and others; see [3, 4, 2]. A recent addition is the paper by Chou, Ng and Pang [5] which is an accompanying work to this one.

Throughout this paper, we shall focus on two equivalent formulations of the NCP and VI as a nonconvex, unconstrained, differentiable minimization problem. Similar results can be obtained for other minimization formulations, such as those mentioned in [13]; these extensions will be omitted. The main issue to be addressed in the present study is the connection between a "minimizing sequence" and a "stationary sequence" of two merit functions proposed recently for the NCP and the VI and the conditions under which these sequential properties are equivalent. Throughout the study, the boundedness of the sequences is *not* assumed. Presumably the results obtained herein have algorithmic implications similar to those for convex programs discussed in the work of Auslender and his colleagues.

The organization of the rest of this paper is as follows. In the next section, we formally define the fundamental concepts to be studied and introduce the two merit functions that will be the focus of the subsequent work. Section 3 is concerned with the NCP and Section 4 with the VI.

A word of notation is in order. For a real-valued differentiable function $f : \Re^n \to \Re$, $\nabla f(x)$ denotes the gradient of f at $x \in \Re^n$ as a column n-vector; for a vector-valued differentiable function $F : \Re^n \to \Re^m$, $F'(x)$ denotes the $m \times n$ Jacobian matrix whose i-th row is $\nabla F_i(x)^T$, where for $i = 1, \ldots, m$, F_i is the i-th component function of F. We let $\nabla F(x)$ be the transpose of $F'(x)$.

2 Definitions and Background

Let $\theta : \Re^n \to \Re$ be a continuously differentiable function satisfying

$$\theta_{\inf} \equiv \inf\{\theta(x) : x \in \Re^n\} > -\infty. \tag{1}$$

A sequence of vectors $\{x^k\} \subset \Re^n$ is said to be *minimizing* if

$$\lim_{k\to\infty} \theta(x^k) = \theta_{\inf}.$$

A sequence of vectors $\{x^k\} \subset \Re^n$ is said to be *stationary* if

$$\lim_{k\to\infty} \nabla\theta(x^k) = 0.$$

It is known that, unlike a minimizer, a minimizing sequence is not necessarily stationary or vice versa, even if θ is convex; counterexamples can be found in [3, 23, 5]. As explained in these and other references, it is useful to know, from a computational point of view at least, when a stationary sequence is minimizing and vice versa. When θ is convex, the principal reason why these two properties are not equivalent is due to the possible unboundedness of the sequence in question. (For an extended real-valued convex function, the gap between the two concepts can exist even when the sequence is bounded.) In the case of a nonconvex function θ, the difference between a minimizing sequence and a stationary sequence becomes potentially more pronounced. It is the primary objective of this work to investigate conditions under which a minimizing sequence is stationary and vice versa when θ is the merit function of an NCP/VI.

For a given function $F : \Re^n \to \Re^n$, the NCP (F) is to find a vector x satisfying

$$x \geq 0, \quad F(x) \geq 0, \quad x \circ F(x) = 0,$$

where the notation $\circ$ denotes the Hadamard product of two vectors. We shall be concerned with the equivalent formulation of the NCP as an unconstrained differentiable minimization problem using the Burmeister-Fischer functional $\varphi : \Re^2 \to \Re$ defined by

$$\varphi(a,b) \equiv \sqrt{a^2+b^2} - (a+b), \quad (a,b) \in \Re^2.$$

This functional φ possesses many interesting properties which we summarize as follows; see Fischer [9]: φ is convex, subadditive, positively homogeneous, and globally Lipschitz continuous; φ is continuously differentiable everywhere in $\Re^2$ except at the origin. It is easy to see that for $(a,b) \neq (0,0)$ and $\alpha = a, b$,

$$\frac{\partial \varphi(a,b)}{\partial \alpha} = \frac{\alpha}{\sqrt{a^2+b^2}} - 1 \in [-2,0].$$

In the context of complementarity problems, the most important property of φ is the following one:

$$\varphi(a,b) = 0 \Leftrightarrow \min(a,b) = 0. \tag{2}$$

Based on the functional φ, Fischer has developed iterative descent algorithms for solving optimization problems, linear and nonlinear complementarity problems [9, 10, 11]. Other authors have proposed related algorithms using the same functional [14, 6, 26]. Tseng [24] has obtained inequalities between φ and the min functional. In particular, Lemma 3.1 in this last reference implies that for all scalars a and b,

$$|\varphi(a,b)| \leq (2+\sqrt{2})\sqrt{a^2+b^2}. \tag{3}$$

For the NCP (F) with F continuously differentiable, consider the functional $\Psi : \Re^n \to \Re_+$ defined by

$$\Psi(x) \equiv \tfrac{1}{2}\textstyle\sum_{i=1}^n \varphi(x_i, F_i(x))^2, \quad x \in \Re^n.$$

This functional Ψ is continuously differentiable and nonnegative everywhere on $\Re^n$. For each $x \in \Re^n$, there exist vectors $a(x), b(x) \in \Re^n$ satisfying $a_i(x), b_i(x) \in [-2,0]$ for all $i = 1, \ldots, n$ such that

$$\nabla \Psi(x) = [D_a(x) + \nabla F(x) D_b(x)]\Phi(x), \tag{4}$$

where $D_a(x)$ and $D_b(x)$ are the diagonal matrices with diagonal entries equal to $a_i(x)$ and $b_i(x)$, respectively, and $\Phi(x)$ is the vector with the i-th component equal to $\varphi(x_i, F_i(x))$. Furthermore, if $(x_i, F_i(x)) \neq (0,0)$, then

$$a_i(x) = \frac{\partial \varphi(x_i, F_i(x))}{\partial a} \quad \text{and} \quad b_i(x) = \frac{\partial \varphi(x_i, F_i(x))}{\partial b};$$

if $(x_i, F_i(x)) = (0,0)$, then

$$a_i(x) = \xi_i - 1 \quad \text{and} \quad b_i(x) = \rho_i - 1,$$

for some $(\xi_i, \rho_i) \in \Re^2$ satisfying $\xi_i^2 + \rho_i^2 \leq 1$. From these expressions for $(a_i(x), b_i(x))$, it can be verified that

$$a_i(x)^2 + b_i(x)^2 \geq 3 - 2\sqrt{2} > 0.$$

Based on these properties, we state the following result which has an important role to play in the next section. The proof of this result is elementary and thus omitted.

Proposition 1 *Let $x \in \Re^n$ be arbitrary. Define*

$$v \equiv D_a(x)\Phi(x) \quad \textit{and} \quad z \equiv D_b(x)\Phi(x).$$

The following four properties are valid for each i:

$$\begin{aligned} z_i > 0 \Leftrightarrow v_i > 0 &\Leftrightarrow \varphi(x_i, F_i(x)) < 0, \\ z_i = 0 \Leftrightarrow v_i = 0 &\Leftrightarrow \varphi(x_i, F_i(x)) = 0, \\ z_i < 0 \Leftrightarrow v_i < 0 &\Leftrightarrow \varphi(x_i, F_i(x)) > 0, \\ z_i^2 + v_i^2 &\geq (3 - 2\sqrt{2})\, \varphi(x_i, F_i(x))^2. \end{aligned}$$

By (2), the NCP (F) has a solution if and only if the unconstrained minimization problem

$$\text{minimize} \;\; \Psi(x) \;:\; x \in \Re^n$$

attains its global minimum and the optimal objective value is zero. In general, since $\Psi(x)$ is nonnegative for all x,

$$\Psi_{\inf} \equiv \inf\{\Psi(x) \;:\; x \in \Re^n\} > -\infty.$$

In the next section, we shall study minimizing sequences and stationary sequences of Ψ, without assuming that these sequences are bounded or $\Psi_{\inf} = 0$. In particular our results neither require nor imply that the NCP (F) has a solution.

In a similar vein, we shall study the variational inequality using a merit function first considered by Peng [21] and later extended by Yamashita, Taji and Fukushima [27]. Specifically, let X be a closed convex set in $\Re^n$ and $F : \Re^n \to \Re^n$ a continuously differentiable function. The VI (X, F) is to find a vector $x \in X$ such that

$$(y - x)^T F(x) \geq 0, \quad \text{for all } y \in X.$$

For a scalar $\alpha > 0$, let f_α denote the regularized gap function introduced in [12]:

$$f_\alpha(x) \equiv \max_{y \in X} \left(F(x)^T(x - y) - \frac{\alpha}{2}\|y - x\|^2 \right),$$

where $\|\cdot\|$ denotes the Euclidean norm of vectors.

This function f_α is finite and continuously differentiable everywhere on $\Re^n$ with gradient given by

$$\nabla f_\alpha(x) \equiv F(x) + \nabla F(x)(x - y_\alpha(x)) + \alpha(y_\alpha(x) - x), \quad x \in \Re^n, \tag{5}$$

where $y_\alpha(x)$ is the unique vector in X attaining the maximum in the definition of $f_\alpha(x)$. For two scalars $\beta > \alpha > 0$, define the functional

$$g_{\alpha\beta}(x) \equiv f_\alpha(x) - f_\beta(x), \quad x \in \Re^n,$$

which will be called the D-gap function, where "D" stands for the word "difference". Clearly $g_{\alpha\beta}(x) \geq 0$ for all x; more importantly, it was shown in [27] that $g_{\alpha\beta}(x) = 0$ if and only if x is a solution of the VI (X, F). Motivated by the latter fact, we wish to study the minimizing sequences and stationary sequences of $g_{\alpha\beta}$ along the same line as for the functional Ψ. Incidentally, it has been

noted by Peng and Yuan [22] that when $X \equiv \Re^n_+$ and $\beta = 1/\alpha$, where $\alpha \in (0,1)$, the D-gap function $g_{\alpha\beta}$ reduces to the implicit Lagrangrian function for the NCP (F) introduced by Mangasarian and Solodov [18] and further studied in [16, 25, 7]. Moreover, it is easy to see that for all $x \in \Re^n$,

$$\frac{\beta-\alpha}{2}\|x - y_\alpha(x)\|^2 \;\geq\; g_{\alpha\beta}(x) \;\geq\; \frac{\beta-\alpha}{2}\|x - y_\beta(x)\|^2. \tag{6}$$

In the following lemma, we state some important properties of the solution function $y_\alpha(x)$ that will be used later.

Proposition 2 *Let X be a closed convex set in $\Re^n$. For arbitrary vectors $x, x' \in \Re^n$ and scalars $\beta > \alpha > 0$,*

$$\|y_\alpha(x) - y_\alpha(x')\| \;\leq\; \alpha^{-1}\,\|F(x) - F(x')\| + \|x - x'\|,$$

$$\|y_\beta(x) - y_\alpha(x)\| \;\leq\; \frac{\beta-\alpha}{\alpha}\,\|x - y_\beta(x)\|.$$

Proof. We note that the vector $y_\alpha(x)$ is the Euclidean projection of the vector $x - \alpha^{-1}F(x)$ onto the set X. Thus by the nonexpansiveness of the projector, the first inequality follows. To prove the second inequality, we deduce by the variational principle applied to $y_\alpha(x)$ and $y_\beta(x)$,

$$(y_\beta(x) - y_\alpha(x))^T\,[\,F(x) + \alpha(y_\alpha(x) - x)\,] \;\geq\; 0,$$

$$(y_\alpha(x) - y_\beta(x))^T\,[\,F(x) + \beta(y_\beta(x) - x)\,] \;\geq\; 0.$$

Adding these two inequalities and rearranging terms, we obtain

$$\alpha\,(y_\beta(x) - y_\alpha(x))^T(y_\alpha(x) - x) - \beta\,(y_\beta(x) - y_\alpha(x))^T(y_\beta(x) - x) \;\geq\; 0, \tag{7}$$

which can be rewritten as

$$\alpha\,(y_\beta(x) - y_\alpha(x))^T(y_\beta(x) - y_\alpha(x)) \;\leq\; (\beta-\alpha)\,(y_\beta(x) - y_\alpha(x))^T(x - y_\beta(x)).$$

Applying the Cauchy-Schwartz inequality to the right-hand inner product and canceling one factor $\|y_\beta(x) - y_\alpha(x)\|$, we easily obtain the second claimed inequality. □

We say that a function $G : \Re^n \to \Re^m$ is *uniformly continuous near a sequence* $\{x^k\} \subset \Re^n$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all k and all y,

$$\|y - x^k\| \;\leq\; \delta \;\Rightarrow\; \|G(y) - G(x^k)\| \;\leq\; \varepsilon.$$

Clearly if G is Lipschitz continuous on $\Re^n$, then G is uniformly continuous near any (bounded or unbounded) sequence. In particular, the latter conclusion holds if G is affine.

Following are two technical results which will be needed later; both results make use of the uniform continuity property near a sequence. The first result pertains to a general vector-valued function; the second result concerns specifically the Burmeister-Fisher functional φ and the associated vector-valued function Φ.

Proposition 3 *Let the function $F : \Re^n \to \Re^n$ be a Fréchet differentiable mapping. If the Fréchet derivative $F' : \Re^n \to \Re^{n\times n}$ is uniformly continuous near a sequence $\{x^k\}$ and $\{F'(x^k)\}$ is bounded, then F is uniformly continuous near $\{x^k\}$.*

Proof By the mean-value theorem [19], we may write

$$\begin{aligned} F(y) - F(x^k) &= \int_0^1 F'(x^k + t(y - x^k))(y - x^k)dt \\ &= \int_0^1 [F'(x^k + t(y - x^k)) - F'(x^k)]\,(y - x^k)dt + F'(x^k)(y - x^k). \end{aligned}$$

Since $\{F'(x^k)\}$ is bounded and F' is uniformly continuous near $\{x^k\}$, the uniform continuity of F near $\{x^k\}$ follows. □

Lemma 1 *Suppose that the function $F : \Re^n \to \Re^n$ is uniformly continuous near a sequence $\{x^k\} \subset \Re^n$ with $\{\Phi(x^k)\}$ bounded. For every $\varepsilon > 0$, there exists a $\delta > 0$ such that for all k*

$$\|y - x^k\| \le \delta \;\Rightarrow\; \|[D_a(y) - D_a(x^k)]\Phi(y)\| \le \varepsilon.$$

Proof By the Lipschitz continuity of the functional φ and the uniform continuity of F near $\{x^k\}$, it follows that Φ is uniformly continuous near $\{x^k\}$. Thus, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for all k and i,

$$\|y - x^k\| \le \delta \;\Rightarrow\; |\varphi(y_i, F_i(y)) - \varphi(x_i^k, F_i(x^k))| \le \frac{\varepsilon}{8}.$$

For simplicity in the proof, we assume that the sequence $\{\Phi(x^k)\}$ converges to a limit Φ^*. Consider an arbitrary component $\Phi_i^* = \lim_{k\to\infty} \varphi(x_i^k, F_i(x^k))$. On the one hand, if $\Phi_i^* = 0$, then for k sufficiently large,

$$|\varphi(x_i^k, F_i(x^k))| \le \frac{\varepsilon}{8};$$

moreover we have

$$\begin{aligned} &|[a_i(y) - a_i(x^k)]\varphi(y_i, F_i(y))| \\ &\le |[a_i(y) - a_i(x^k)]\varphi(x_i^k, F_i(x^k))| + |[a_i(y) - a_i(x^k)]\,[\varphi(y_i, F_i(y)) - \varphi(x_i^k, F_i(x^k))]| \le \varepsilon, \end{aligned}$$

where the last inequality follows because $|a_i(x)| \le 2$ for all $x \in \Re^n$. On the other hand, if $\Phi_i^* > 0$, then by adjusting δ if necessary, it follows from the uniform continuity of Φ near $\{x^k\}$ that for some constant $c > 0$, all k sufficiently large and all y satisfying $\|y - x^k\| \le \delta$, we have $|\varphi(y_i, F_i(y))| \ge c$ which implies by (3)

$$\sqrt{y_i^2 + F_i^2(y)} \ge \frac{c}{2 + \sqrt{2}}.$$

Hence we deduce

$$a_i(y) - a_i(x^k) = \frac{y_i}{\sqrt{y_i^2 + F_i^2(y)}} - \frac{x_i^k}{\sqrt{(x_i^k)^2 + F_i^2(x^k)}}.$$

Since the two denominators in the right-hand side are bounded away from zero and F_i is uniformly continuous near $\{x^k\}$, it follows that $|a_i(y) - a_i(x^k)|$ can be made arbitrarily small uniformly for all k, whenever y is sufficiently close to x^k; hence the same will be true for $|[a_i(y) - a_i(x^k)]\varphi(y_i, F_i(y))|$ because $\{\varphi(y_i, F_i(y))\}$ is bounded. Since i is arbitrary, the desired conclusion of the lemma holds. □

We conclude this section by citing a result proved in [5] which gives a sufficient condition for a minimizing sequence of a smooth function to be stationary. See the reference for a proof. Incidentally, the function in this result is *not* assumed to be convex.

Theorem 1 *Let $\theta : \Re^n \to \Re$ be a Fréchet differentiable function satisfying (1). If $\{x^k\}$ is a minimizing sequence of θ and $\nabla\theta$ is uniformly continuous near $\{x^k\}$, then*

$$\lim_{k\to\infty} \nabla\theta(x^k) = 0.$$

3 The NCP

In this section, we study the NCP function Ψ defined in the last section. Let

$$\Psi_{\inf} \equiv \inf\{\Psi(x) : x \in \Re^n\} \geq 0.$$

Throughout the discussion, we assume that the function F defining the NCP is continuously differentiable on $\Re^n$; this assumption implies the same property for Ψ as noted before. Our goal is to investigate situations when a minimizing sequence of Ψ is stationary and vice versa. We first give some examples to illustrate that there are situations in which this is not true.

Borrowing an example from [5], we consider a 2-dimensional NCP with the following function:

$$F(x_1, x_2) \equiv \begin{pmatrix} e^{x_1^2 - x_2} \\ 0 \end{pmatrix}, \quad (x_1, x_2) \in \Re^2.$$

The solution set of this NCP (F) is $\{(x_1, x_2) : x_1 = 0, x_2 \geq 0\}$. Consider the sequence $\{x^k\}$ with

$$x^k = (x_1^k, x_2^k) \equiv (k, k^2 + 0.5\log k), \quad \text{for all } k.$$

Trivially, $F_1(x_1^k, x_2^k) = 1/\sqrt{k}$; thus $F(x^k) \to 0$. By an easy calculation, it can be verified that

$$\lim_{k\to\infty} \Psi(x^k) = 0 \quad \text{and} \quad \lim_{k\to\infty} \frac{\partial\Psi(x^k)}{\partial x_1} = 2.$$

Thus the sequence $\{x^k\}$ is minimizing but not stationary for Ψ.

Conversely, consider the 2-dimensional function

$$F(x_1, x_2) \equiv \begin{pmatrix} 1 \\ \frac{1}{(x_1+x_2-1)^2+1} - 1 \end{pmatrix}, \quad (x_1, x_2) \in \Re^2.$$

The NCP (F) has the unique solution $(x_1, x_2) = (0, 1)$. Let $\{x^k\}$ be the sequence such that

$$x^k = (x_1^k, x_2^k) \equiv (c, k), \quad \text{for all } k,$$

where c is an arbitrary positive constant. Then, by direct calculation, it is easy to see that

$$\lim_{k\to\infty} \nabla\Psi(x^k) = 0.$$

Hence the sequence $\{x^k\}$ is stationary. Moreover, we have

$$\frac{\partial\Psi(x)}{\partial x_2} < 0 \tag{8}$$

for all x such that $x_1 = c$ and x_2 is sufficiently large. To see this, first note that

$$\frac{\partial \Psi(x)}{\partial x_2} = \frac{\partial F_1(x)}{\partial x_2} b_1(x)\varphi(x_1, F_1(x)) + \left[a_2(x) + \frac{\partial F_2(x)}{\partial x_2} b_2(x)\right] \varphi(x_2, F_2(x)). \tag{9}$$

Since $\frac{\partial F_1(x)}{\partial x_2} = 0$, the first term vanishes. Next, note that

$$\begin{aligned} a_2(x) &= \frac{x_2}{\sqrt{x_2^2 + F_2(x)^2}} - 1 = \frac{-F_2(x)^2}{\sqrt{x_2^2 + F_2(x)^2}\left(x_2 + \sqrt{x_2^2 + F_2(x)^2}\right)}, \\ \frac{\partial F_2(x)}{\partial x_2} &= -\frac{2(x_1 + x_2 - 1)}{((x_1 + x_2 - 1)^2 + 1)^2}, \\ b_2(x) &= \frac{F_2(x)}{\sqrt{x_2^2 + F_2(x)^2}} - 1. \end{aligned}$$

Now let us fix $x_1 = c$ and consider the limit $x_2 = k \to \infty$. Then it is easy to see that

$$a_2(x) < 0, \quad \frac{\partial F_2(x)}{\partial x_2} < 0, \quad b_2(x) < 0$$

and

$$a_2(x) = O(k^{-2}), \quad \frac{\partial F_2(x)}{\partial x_2} = O(k^{-3}), \quad b_2(x) \to -1,$$

from which we can deduce that the term in the square bracket in (9) eventually becomes negative. Moreover we have

$$\varphi(x_2, F_2(x)) = \frac{-2x_2 F_2(x)}{\sqrt{x_2^2 + F_2(x)^2} + x_2 + F_2(x)} \to -\lim_{\substack{x_1 = c \\ x_2 \to \infty}} F_2(x) = 1.$$

Hence it follows from (9) that (8) is eventually satisfied. Therefore $\{\Psi(x^k)\}$ is a decreasing sequence. On the other hand, we have

$$\lim_{k \to \infty} \Psi(x^k) = \frac{1}{2}\left((\sqrt{1 + c^2} - (1 + c))^2 + 1\right) > 0.$$

Thus the sequence $\{x^k\}$ is not minimizing, because the NCP (F) has a solution, i.e., $\Psi_{\inf} = 0$.

Motivated by results of [3] concerning well-behaved convex functions, we introduce an important property for sequences associated with the NCP (F). When specialized to constant sequences (i.e. single vectors), this property is closely related to the concept of an s-regular vector defined by Moré [15] which in turn is motivated by the notion of s-regularity introduced by Pang and Gabriel [20]; other related regularity concepts for a vector have been used by Ferris and Ralph [8] and De Luca, Facchinei and Kanzow [6].

Definition 1 *Let $\{x^k\} \subset \Re^n$ be a sequence with $\{\Psi(x^k)\}$ bounded. We say that $\{x^k\}$ is asymptotically regular (with respect to φ and F) if for every scalar $\lambda > \Psi_{\inf}$ and every convergent sequence $\{z^k\} \subset \Re^n$ with limit z^* satisfying*

(a) $z^* \circ \Phi^* \neq 0$ *for some accumulation point* Φ^* *of* $\{\Phi(x^k)\}$, *and*

(b) for all k *and all* i,

$$z_i^k > 0 \Leftrightarrow \varphi(x_i^k, F_i(x^k)) > 0,$$
$$z_i^k = 0 \Leftrightarrow \varphi(x_i^k, F_i(x^k)) = 0,$$
$$z_i^k < 0 \Leftrightarrow \varphi(x_i^k, F_i(x^k)) < 0,$$

there exist a convergent sequence $\{y^k\}$ *with limit* y^*, *a scalar* $\gamma > 0$ *and a function* $\delta : \Re_+ \to \Re_+$ *with the property,*

$$[\lim_{k\to\infty} t_k = t_* > 0] \Rightarrow \liminf_{k\to\infty} \delta(t_k) > 0, \tag{10}$$

such that

(A1) $y^k \circ z^k \geq 0$ *for all* k *sufficiently large,*

(A2) $y^* \circ \Phi^* \neq 0$ *for some accumulation point* Φ^* *of* $\{\Phi(x^k)\}$, *and*

(A3) $(y^k)^T \nabla F(x^k) z^k \geq \delta((\Psi(x^k) - \lambda)_+)\,((y^k)^T z^k)^\gamma$ *for all* k *sufficiently large.*

If $\{x^k\}$ is a sequence such that

$$\lim_{k\to\infty} \Psi(x^k) = 0,$$

then $\{x^k\}$ is trivially asymptotically regular because there exists no sequence $\{z^k\}$ satisfying the requirements in the above definition. Separately, if $\{x^k\}$ is such that there exists a constant $\tau > 0$ satisfying

$$(z^k)^T \nabla F(x^k) z^k \geq \tau\,(z^k)^T z^k, \quad \text{for all } k,$$

for all convergent sequences $\{z^k\}$ with properties (a) and (b) in Definition 1, then $\{x^k\}$ is asymptotically regular; it suffices to choose $y^k \equiv z^k$ for all k, $\gamma = 2$, and

$$\delta(t) = \begin{cases} t & \text{if } t \in [0, \tau] \\ \tau & \text{otherwise.} \end{cases}$$

The latter situation includes the case where $\{x^k\}$ is a constant sequence with $x^k \equiv x$ for all k and $\nabla F(x)$ is positive definite.

Theorem 2 *Let* $F : \Re^n \to \Re^n$ *be a continuously differentiable function and* $\{x^k\} \subset \Re^n$ *be an arbitrary sequence with* $\{\Psi(x^k)\}$ *and* $\{\nabla F(x^k)\}$ *bounded. The following two statements are valid.*

(i) If $\{x^k\}$ *is asymptotically regular and*

$$\lim_{k\to\infty} \nabla\Psi(x^k) = 0, \tag{11}$$

then

$$\lim_{k\to\infty} \Psi(x^k) = \Psi_{\inf}. \tag{12}$$

(ii) Conversely, if ∇F *is uniformly continuous near* $\{x^k\}$ *and (12) holds, then (11) holds.*

Proof To prove the first statement, suppose that $\{x^k\}$ is asymptotically regular, stationary but not minimizing for Ψ. Without loss of generality, we may assume that the sequence $\{\Phi(x^k)\}$ has a limit Φ^* and

$$\Psi^* \equiv \tfrac{1}{2}\|\Phi^*\|^2 > \Psi_{\inf}. \tag{13}$$

Choose a scalar λ satisfying $\Psi^* > \lambda > \Psi_{\inf}$; let

$$v^k \equiv D_a(x^k)\Phi(x^k) \quad \text{and} \quad z^k \equiv D_b(x^k)\Phi(x^k).$$

Then, by (4), we have

$$\nabla\Psi(x^k) = v^k + \nabla F(x^k)z^k.$$

Since $\{D_a(x^k)\}$ and $\{D_b(x^k)\}$ are bounded, we may assume without loss of generality that $\{v^k\}$ and $\{z^k\}$ converge to some vectors v^* and z^* respectively. By Proposition 1, $\{z^k\}$ satisfies condition (b) in Definition 1. We argue that condition (a) in this definition is also satisfied. Assume not; then $z^* \circ \Phi^* = 0$. We first show $z^* = 0$. In fact, if $z_i^* \neq 0$ for some i, then $z^* \circ \Phi^* = 0$ would imply $\Phi_i = 0$. But this together with the definition of z^k and the boundedness of $D_b(x^k)$ in turn implies $z_i^k \to 0$, contradicting $z_i^* \neq 0$. Now, since

$$v^k = \nabla\Psi(x^k) - \nabla F(x^k)z^k$$

and $\{\nabla F(x^k)\}$ is bounded, it follows, by passing to the limit $k \to \infty$, that $v^* = 0$. Therefore, for each $i = 1, \ldots, n$, we have

$$0 = (v_i^*)^2 + (z_i^*)^2 \geq (3 - 2\sqrt{2})(\Phi_i^*)^2.$$

Thus $\Phi^* = 0$ which contradicts the assumed inequality (13). Consequently, the two conditions (a) and (b) in Definition (1) are satisfied.

By asymptotic regularity, it follows that there exist a convergent sequence $\{y^k\}$ with limit y^*, a scalar $\gamma > 0$, and a function δ satisfying conditions (A1), (A2) and (A3). Moreover, since $\lim_{k\to\infty} \Phi(x^k) = \Phi^*$ and $\Psi^* = \frac{1}{2}\|\Phi^*\|^2 > \lambda$ by assumption, we have $(\Psi(x^k) - \lambda)_+ = \Psi(x^k) - \lambda > 0$ for all k sufficiently large. Therefore we have

$$\begin{aligned} 0 &= \lim_{k\to\infty}(y^k)^T\nabla\Psi(x^k) \\ &= \lim_{k\to\infty}(y^k)^T[\,v^k + \nabla F(x^k)z^k\,] \\ &\geq \liminf_{k\to\infty}[(y^k)^T v^k + \delta(\Psi(x^k) - \lambda)((y^k)^T z^k)^\gamma]. \end{aligned}$$

Since $y^k \circ z^k \geq 0$ for all k sufficiently large, it follows that $y^* \circ z^* \geq 0$. Moreover, by Proposition 1, it follows that $y^k \circ v^k \geq 0$ for all k sufficiently large; thus $y^* \circ v^* \geq 0$. Since $\liminf_{k\to\infty} \delta(\Psi(x^k) - \lambda) > 0$, we deduce

$$y^* \circ v^* = y^* \circ z^* = 0,$$

which implies for all $i = 1, \ldots, n$,

$$0 = (y_i^*)^2[(z_i^*)^2 + (v_i^*)^2] \geq (3 - 2\sqrt{2})(y_i^*)^2(\Phi_i^*)^2.$$

Thus $y^* \circ \Phi^* = 0$ which contradicts condition (A2). This contradiction establishes statement (i).

To prove statement (ii), it suffices to verify that $\nabla\Psi$ is uniformly continuous near $\{x^k\}$, by Theorem 1. For this purpose, we evaluate

$$\nabla\Psi(y) - \nabla\Psi(x^k) \;=\; [D_a(y)\Phi(y) - D_a(x^k)\Phi(x^k)] + [\nabla F(y)D_b(y)\Phi(y) - \nabla F(x^k)D_a(x^k)\Phi(x^k)].$$

The term within the first bracket of the right-hand side is equal to

$$D_a(x^k)[\Phi(y) - \Phi(x^k)] + [D_a(y) - D_a(x^k)]\Phi(y).$$

By Proposition 3, it follows that F is uniformly continuous near $\{x^k\}$; this in turn implies that Φ is also uniformly continuous near $\{x^k\}$, by the Lipschitz continuity of φ; Since $\{D_a(x^k)\}$ is bounded, it follows that the first summand in the above expression can be made arbitrarily small in norm uniformly in k whenever y is sufficiently close to x^k; the same conclusion holds for the second summand by Lemma 1. By a similar argument, it can be shown that the term

$$\nabla F(y)D_b(y)\Phi(y) - \nabla F(x^k)D_a(x^k)\Phi(x^k)$$

can also be made arbitrarily small in norm uniformly in k whenever y is sufficiently close to x^k. Consequently $\nabla\Psi$ is uniformly continuous near $\{x^k\}$. □

The boundedness assumption of $\{\nabla F(x^k)\}$ is satisfied if either F is affine or each $\nabla F(x^k)$ is positive semidefinite plus as defined in [17]. Ideally, it would be nice if Theorem 2 would remain valid without this assumption.

4 The VI

We next turn our attention to the VI (X, F). The following result for the D-gap function $g_{\alpha\beta}$ is the analog of Theorem 2.

Theorem 3 *Let X be a closed convex subset of $\Re^n$ and $F : \Re^n \to \Re^n$ be a continuously differentiable function. Let $\beta > \alpha > 0$ be given scalars and $\{x^k\} \subset \Re^n$ be an arbitrary sequence with $\{\nabla F(x^k)\}$ bounded. The following two statements are valid.*

(i) Suppose that for all $\lambda > g_{\inf} \equiv \inf_{x\in\Re^n} g_{\alpha\beta}(x)$, there exist a scalar $\gamma > 0$ and a function $\delta : \Re_+ \to \Re_+$ with property (10) such that for all k,

$$\begin{aligned}(y_\beta(x^k) - y_\alpha(x^k))^T \nabla F(x^k)(y_\beta(x^k) - y_\alpha(x^k)) \\ \geq\; \delta((g_{\alpha\beta}(x^k) - \lambda)_+)\, \|y_\beta(x^k) - y_\alpha(x^k)\|^\gamma.\end{aligned} \tag{14}$$

If $\{g_{\alpha\beta}(x^k)\}$ is bounded and

$$\lim_{k\to\infty} \nabla g_{\alpha\beta}(x^k) \;=\; 0, \tag{15}$$

then

$$\lim_{k\to\infty} g_{\alpha\beta}(x^k) \;=\; g_{\inf}. \tag{16}$$

(ii) Conversely, if ∇F is uniformly continuous near $\{x^k\}$ and (16) holds, then (15) holds.

Proof Similar to the proof of Theorem 3, we assume for convenience that $\{g_{\alpha\beta}(x^k)\}$ has a limit g^*; we further assume by way of contradiction that g^* is greater than $g_{\inf}$. Choose a scalar λ satisfying $g^* > \lambda > g_{\inf}$. Let the scalar γ and function δ be as stated. By (5), we deduce

$$\nabla g_{\alpha\beta}(x^k) \;=\; \nabla F(x^k)(y_\beta(x^k) - y_\alpha(x^k)) + \alpha(y_\alpha(x^k) - x^k) - \beta(y_\beta(x^k) - x^k). \tag{17}$$

We claim that $y_\beta(x^k) - y_\alpha(x^k)$ is bounded. By Proposition 2, it suffices to show that $\{x^k - y_\beta(x^k)\}$ is bounded. But the latter sequence is clearly bounded by (6) and the boundedness of $\{g_{\alpha\beta}(x^k)\}$. Consequently, (15) yields

$$\lim_{k\to\infty} (y_\beta(x^k) - y_\alpha(x^k))^T \nabla g_{\alpha\beta}(x^k) \;=\; 0.$$

Premultiplying (17) by $(y_\beta(x^k) - y_\alpha(x^k))^T$, using (14) and (7), passing to the limit $k \to \infty$, and using the assumption $g^* > \lambda$ and the property (10), we deduce

$$\lim_{k\to\infty} (y_\beta(x^k) - y_\alpha(x^k)) \;=\; 0. \tag{18}$$

Since $\{x^k - y_\beta(x^k)\}$ is bounded, we may assume without loss of generality that this sequence converges. From (18), it follows that

$$\lim_{k\to\infty} (x^k - y_\beta(x^k)) \;=\; \lim_{k\to\infty} (x^k - y_\alpha(x^k)). \tag{19}$$

This yields by (6)

$$\lim_{k\to\infty} g_{\alpha\beta}(x^k) \;=\; \frac{\beta - \alpha}{2} \lim_{k\to\infty} \|x^k - y_\alpha(x^k)\|^2. \tag{20}$$

By (15), (17), (18), (19), the boundedness of $\{\nabla F(x^k)\}$ and the fact that $\beta > \alpha$, we obtain

$$\lim_{k\to\infty} (x^k - y_\beta(x^k)) \;=\; 0,$$

which together with (20) implies that $\{g_{\alpha\beta}(x^k)\}$ converges to zero. This is a contradiction. Hence (16) holds.

To prove statement (ii), we show that $\nabla g_{\alpha\beta}$ is uniformly continuous near $\{x^k\}$. By Propositions 3 and 2, it follows that both y_α and y_β are uniformly continuous near $\{x^k\}$. Hence by the expression (17), it suffices to show that $\nabla F(x)(y_\beta(x) - y_\alpha(x))$ is uniformly continuous near $\{x^k\}$. We may write

$$\begin{aligned}
&\nabla F(x')(y_\beta(x') - y_\alpha(x')) - \nabla F(x^k)(y_\beta(x^k) - y_\alpha(x^k)) \\
&\quad = (\nabla F(x') - \nabla F(x^k))(y_\beta(x^k) - y_\alpha(x^k)) + \nabla F(x')[(y_\beta(x') - y_\beta(x^k)) - (y_\alpha(x') - y_\alpha(x^k))].
\end{aligned}$$

As above the sequence $\{y_\beta(x^k) - y_\alpha(x^k)\}$ is bounded; moreover, since ∇F is uniformly continuous near $\{x^k\}$, it follows that $\{\nabla F(x')\}$ is bounded for x' sufficiently near x^k. From the above expression, the desired uniform continuity of $\nabla F(x)(y_\beta(x) - y_\alpha(x))$ follows. □

References

[1] G. Auchmuty, *Variational principles for variational inequalities*, Numerical Functional Analysis and Optimization 10 (1989), pp. 863-874.

[2] A. Auslender, *Convergence of stationary sequences for variational inequalities with maximal monotone operators*, Applied Mathematics and Optimization 28 (1993), pp. 161-172.

[3] A. Auslender and J.-P. Crouzeix, *Well behaved asymptotical convex functions*, Analyse non-linéare, Gauthier-Villars, Paris, 1989, pp. 101-122.

[4] A. Auslender, R. Cominetti and J.-P. Crouzeix, *Convex functions with unbounded level sets and applications to duality theory*, SIAM Journal on Optimization 3 (1993), pp. 669-687.

[5] C.C. Chou, K.F. Ng and J.S. Pang, *Minimizing and stationary sequences for optimization problems*, manuscript, Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, Maryland, August 1995.

[6] T. De Luca, F. Facchinei and C. Kanzow, *A semismooth equation approach to the solution of nonlinear complementarity problems* , Technical Report, Dipartimento di Informatica e Sistemistica, Università di Roma "La Sapienza", Rome, Italy, January, 1995.

[7] F. Facchinei and C. Kanzow, *When are the (un)constrained stationary points of the implicit Lagrangian global solutions?* , Technical Report, Dipartimento di Informatica e Sistemistica, Università di Roma "La Sapienza", Rome, Italy, August, 1995.

[8] M.C. Ferris and D. Ralph, *Projected gradient methods for nonlinear complementarity problems via normal maps*, in D.-Z. Du, L. Qi and R.S. Womersley, eds., Recent Advances in Nonsmooth Optimization, World Scientific Publishers, Singapore (1995), to appear.

[9] A. Fischer, *A special Newton-type optimization method*, Optimization 24 (1992), pp. 269-284.

[10] A. Fischer, *A Newton-type method for positive semidefinite linear complementarity problems*, Journal of Optimization Theory and Applications, to appear.

[11] A. Fischer, *An NCP-function and its use for the solution of complementarity problems*, in D.-Z. Du, L. Qi and R.S. Womersley, eds., Recent Advances in Nonsmooth Optimization, World Scientific Publishers, Singapore (1995), to appear.

[12] M. Fukushima, *Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems*, Mathematical Programming, 53 (1992), pp. 99-110.

[13] M. Fukushima, *Merit functions for variational inequality and complementarity problems*, Technical Report 95022, Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan, June, 1995.

[14] H. Jiang and L. Qi, *A new nonsmooth equations approach to nonlinear complementarities*, AMR94/31, School of Mathematics, The University of New South Wales, Sydney, Australia, October, 1994.

[15] J.J. Moré, *Global methods for nonlinear complementarity problems*, Preprint MCS-P429-0494, Argonne National Laboratory, Argonne, Illinois, April, 1994.

[16] Z.-Q. Luo, O.L. Mangasarian, J. Ren and M.V. Solodov, *New error bounds for the linear complementarity problem*, Mathematics of Operations Research, 19 (1994), pp. 880-892.

[17] Z.-Q. Luo and P. Tseng, *A decomposition property for a class of square matrices*, Applied Mathematics Letters, 4 (1991), pp. 67-69.

[18] O.L. Mangasarian and M.V. Solodov, *Nonlinear complementarity as unconstrained and constrained minimization*, Mathematical Programming, 62 (1993) pp. 277-297.

[19] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, NY, 1970.

[20] J.-S. Pang and S.A. Gabriel, *NE/SQP: A robust algorithm for the nonlinear complementarity problem*, Mathematical Programming 60 (1993), pp. 295-337.

[21] J.M. Peng, *Equivalence of variational inequality problems to unconstrained optimization*, Technical Report, State Key Laboratory of Scientific and Engineering Computing, Academia Sinica, Beijing, China, April, 1995.

[22] J.M. Peng and Y.X. Yuan, *Unconstrained methods for generalized complementarity problems*, Technical Report, State Key Laboratory of Scientific and Engineering Computing, Academia Sinica, Beijing, China, December, 1994.

[23] M.J. Todd, *On convergence properties of algorithms for unconstrained minimization*, IMA Journal on Numerical Analysis 9 (1989) pp. 435-441.

[24] P. Tseng, Growth behaviour of a class of merit functions for the nonlinear complementarity problem, Technical Report, Department of Mathematics, University of Washington, Seattle, May, 1994.

[25] N. Yamashita and M. Fukushima, *On stationary points of the implicit Lagrangian for nonlinear complementarity problems*, Journal of Optimization Theory and Applications 84 (1995) 653-663.

[26] N. Yamashita and M. Fukushima, *Modified Newton methods for solving semismooth reformulations of monotone complementarity problems*, Technical Report 95021, Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan, May, 1995.

[27] N. Yamashita, K. Taji and M. Fukushima, *Unconstrained optimization reformulations of variational inequality problems*, Technical Report 95024, Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan, August, 1995.