SOLVING BOX CONSTRAINED VARIATIONAL INEQUALITIES BY USING THE NATURAL RESIDUAL WITH D-GAP FUNCTION GLOBALIZATION

Christian Kanzow[1,2] and Masao Fukushima[3,4]

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

[3] Department of Applied Mathematics and Physics
Graduate School of Engineering
Kyoto University
Kyoto 606-01
Japan
e-mail: fuku@kuamp.kyoto-u.ac.jp

November 21, 1997

Abstract. We present a new method for the solution of the box constrained variational inequality problem, BVIP for short. Basically, this method is a nonsmooth Newton method applied to a reformulation of BVIP as a system of nonsmooth equations involving the natural residual. The method is globalized by using the D-gap function. We show that the proposed algorithm is globally and fast locally convergent. Moreover, if the problem is described by an affine function, the algorithm has a finite termination property. Numerical results for some large-scale variational inequality problems are reported.

Key words: Variational inequality problem, mixed complementarity problem, natural residual, D-gap function, Newton's method, global convergence, quadratic convergence, finite termination.

[2]Current address (October 1, 1997 — September 30, 1998): Computer Sciences Department, University of Wisconsin — Madison, 1210 West Dayton Street, 53706 Madison, WI; e-mail: kanzow@cs.wisc.edu. The research of this author is being supported by the DFG (Deutsche Forschungsgemeinschaft).

[4]The work of this author is supported in part by the Scientific Research Grant-in-Aid from the Ministry of Education, Science and Culture, Japan.

1 Introduction

Let l and u be two n-dimensional vectors with components $l_i \in \mathbb{R} \cup \{-\infty\}$ and $u_i \in \mathbb{R} \cup \{\infty\}$ satisfying $l_i < u_i$, and denote by $\mathbb{B}$ the nonempty and possibly infinite box $[l, u] := \{x \in \mathbb{R}^n \,|\, l_i \leq x_i \leq u_i,\ i = 1, \ldots, n\}$. Then the *box constrained variational inequality problem*, BVIP for short, is to find a vector $x^* \in \mathbb{B}$ such that

$$F(x^*)^T(x - x^*) \geq 0 \quad \forall x \in \mathbb{B}, \tag{1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a given function, which we assume to be continuously differentiable throughout this note. This problem is also called the *mixed complementarity problem*, see [3, 4], where the reader may also find a number of interesting applications of this problem.

Most of the existing methods for the solution of BVIP are based on a suitable reformulation of this problem as an optimization problem, as a system of nonlinear equations, or as a fixed-point problem. We refer the interested reader to the recent paper [17] and the references therein. In particular, it is well-known that x^* solves BVIP if and only if it solves the nonlinear equation

$$r(x) = 0, \tag{2}$$

where $r : \mathbb{R}^n \to \mathbb{R}^n$ is the so-called *natural residual* of BVIP defined by

$$r(x) := x - \Pi_{\mathbb{B}}(x - F(x)),$$

where $\Pi_{\mathbb{B}}(z)$ denotes the Euclidean projection of a vector z on the set $\mathbb{B}$. Note that, thanks to the special structure of the box $\mathbb{B}$, the projection onto this set is trivial to compute. Indeed, this is one of the reasons why we restrict ourselves to BVIP rather than a more general variational inequality problem.

The major idea underlying the algorithm to be presented in this paper is to solve BVIP by applying a Newton-type method to the nonlinear system (2). Since this system is not differentiable, we have to take a nonsmooth Newton method. However, it is not a trivial matter to globalize the nonsmooth Newton method.

The globalization we suggest is based on another reformulation of problem (1). We first recall that the *regularized gap function* [7] for problem (1) is given by

$$g_\gamma(x) := \max_{y \in \mathbb{B}} \left\{ F(x)^T(x - y) - \frac{\gamma}{2}\|x - y\|^2 \right\},$$

where $\gamma > 0$ is any given parameter. Since the expression in the curly brackets is a strictly concave quadratic function in y, it has a unique maximizer, which we denote $y_\gamma(x)$. In fact, it is easy to see that this maximizer is given by

$$y_\gamma(x) = \Pi_{\mathbb{B}}\left(x - \frac{1}{\gamma}F(x)\right),$$

see [7]. Hence we can rewrite the regularized gap function as

$$g_\gamma(x) = F(x)^T(x - y_\gamma(x)) - \frac{\gamma}{2}\|x - y_\gamma(x)\|^2.$$

Our globalization strategy is now based on the *D-gap function*

$$g_{\alpha\beta}(x) := g_\alpha(x) - g_\beta(x),$$

where $0 < \alpha < \beta$, i.e., the D-gap function is just the difference of two regularized gap functions with different parameter values α and β. The D-gap function was first introduced by Peng [14] and later studied in, e.g., [19, 18, 8, 12, 15]. The D-gap function has a number of nice properties; here, we only recall that $g_{\alpha\beta}$ is continuously differentiable and nonnegative on the whole space $\mathbb{R}^n$, and that $g_{\alpha\beta}(x^*) = 0$ if and only if x^* solves problem (1). Hence it provides an unconstrained optimization reformulation

$$\min\ g_{\alpha\beta}(x), \quad x \in \mathbb{R}^n,$$

of problem (1).

Nevertheless, the direct minimization of $g_{\alpha\beta}$ is difficult since it is once, but not twice continuously differentiable. So we will use this function only to monitor convergence of the iterates generated by a nonsmooth Newton method applied to the nonlinear system (2).

The paper is organized as follows: In Section 2, we state some preliminary results concerning the relationship between the growth behaviour of the natural residual r and that of the D-gap function $g_{\alpha\beta}$. The algorithm along with brief convergence analysis is presented in Section 3. Preliminary numerical results on some large-scale problems are reported in Section 4. We conclude the paper with some remarks in Section 5.

We adopt the following notations. For an $n \times n$ matrix $A = (a_{ij})$ and index sets $\gamma, \delta \subseteq \{1, \dots, n\}$, $A_{\gamma\delta}$ denotes the submatrix of A that consists of elements a_{ij}, $i \in \gamma$, $j \in \delta$. Similarly, for an n-vector d, d_γ denotes the subvector of d with elements d_i, $i \in \gamma$. The ith component function of the vector-valued function $F : \mathbb{R}^n \to \mathbb{R}^n$ is denoted by $F_i : \mathbb{R}^n \to \mathbb{R}$ and the Jacobian of F at $x \in \mathbb{R}^n$ is denoted as $F'(x) \equiv \nabla F(x)^T$, where $\nabla F(x)$ is the $n \times n$ matrix with columns $\nabla F_i(x)$, $i = 1, \dots, n$.

2 Preliminaries

As announced in the introduction, the aim of this section is to provide a relationship between the growth behaviour of the natural residual r and that of the D-gap function $g_{\alpha\beta}$. We first restate a result from Yamashita et al. [19].

Lemma 2.1 *Let* $0 < \alpha < \beta$. *Then*

$$\frac{\beta - \alpha}{2}\|x - y_\beta(x)\|^2 \le g_{\alpha\beta}(x) \le \frac{\beta - \alpha}{2}\|x - y_\alpha(x)\|^2$$

for all $x \in \mathbb{R}^n$.

The following result follows immediately from the definition of y_γ and Lemma 2.2 in Calamai and Moré [1] (see also [9, 15]).

Lemma 2.2 *Let* $0 < \alpha < \beta$. *Then*

$$\alpha \| x - y_\alpha(x) \| \le \beta \| x - y_\beta(x) \|$$

for all $x \in \mathbb{R}^n$.

We are now able to state and prove the main result of this section. We note that this result has essentially been established by Peng and Fukushima [15] recently. However, the present proof is considerably more concise.

Proposition 2.3 *Let* $0 < \alpha < 1 < \beta$. *Then there exist constants* $c_1 > 0$ *and* $c_2 > 0$ *such that*

$$c_1 \| r(x) \|^2 \le g_{\alpha\beta}(x) \le c_2 \| r(x) \|^2$$

for all $x \in \mathbb{R}^n$.

Proof. In view of Lemma 2.2, we have

$$\alpha^2 \| x - y_\alpha(x) \|^2 \le \| x - y_1(x) \|^2 \le \beta^2 \| x - y_\beta(x) \|^2$$

for all $x \in \mathbb{R}^n$. Hence we obtain from Lemma 2.1 and $0 < \alpha < 1 < \beta$

$$\begin{aligned} \frac{\beta - \alpha}{2\beta^2} \| x - y_1(x) \|^2 &\le \frac{\beta - \alpha}{2} \| x - y_\beta(x) \|^2 \\ &\le g_{\alpha\beta}(x) \\ &\le \frac{\beta - \alpha}{2} \| x - y_\alpha(x) \|^2 \\ &\le \frac{\beta - \alpha}{2\alpha^2} \| x - y_1(x) \|^2. \end{aligned}$$

Since the natural residual can be written as

$$r(x) = x - y_1(x),$$

we obtain

$$c_1 \| r(x) \|^2 \le g_{\alpha\beta}(x) \le c_2 \| r(x) \|^2$$

with

$$c_1 := \frac{\beta - \alpha}{2\beta^2} \quad \text{and} \quad c_2 := \frac{\beta - \alpha}{2\alpha^2}.$$

This completes the proof. □

Proposition 2.3 is essential in establishing a fast local convergence property of the algorithm to be presented in the next section.

3 Algorithm and Convergence

In this section, we present the algorithm and investigate its global and local convergence properties.

Thanks to the special structure of the set $\mathbb{B} = [l, u]$, the ith component of the natural residual $r(x)$ can be written as

$$\begin{aligned} r_i(x) &= x_i - \text{mid}\,\{l_i, u_i, x_i - F_i(x)\} \\ &= \text{mid}\,\{x_i - l_i, x_i - u_i, F_i(x)\}, \end{aligned}$$

where $\text{mid}\{a, b, c\}$ denotes the median of three numbers a, b, c. Since F is continuously differentiable, r is piecewise smooth. Therefore the B-subdifferential [16] and the Clarke subdifferential, which are defined by

$$\partial_B r(x) = \{H \in \mathbb{R}^{n\times n} \,|\, H = \lim_{k\to\infty} F'(x^k),\ x = \lim_{k\to\infty} x^k,\ F'(x^k) \text{ exists for all } k\}$$

and

$$\partial r(x) = \text{conv}\ \partial_B r(x),$$

respectively, are well-defined for each $x \in \mathbb{R}^n$. Recall that the function r is semismooth [13, 16] if, for any $x \in \mathbb{R}^n$,

$$\lim_{\substack{H\in\partial r(x+td')\\ d'\to d,\, t\downarrow 0}} Hd'$$

exists for all $d \in \mathbb{R}^n$. In fact, since the 'mid' operator is a semismooth operator, we have the following results, which follow immediately from the fact that the composition of (strongly) semismooth functions is again (strongly) semismooth [13, 5].

Lemma 3.1 *The natural residual r is semismooth. If, in addition, F' is locally Lipschitzian, then the natural residual r is strongly semismooth.*

Basically, the proposed algorithm is a nonsmooth Newton method [16] applied to the system of semismooth equations (2). Specifically, the Newton direction is determined as a solution of the linear equation

$$H_k d = -r(x^k), \tag{3}$$

where H_k is an arbitrary element of $\partial_B r(x^k)$.

To obtain an explicit representation of $\partial_B r(x)$, we introduce the following three index sets:

$$\begin{aligned} \alpha(x) &:= \{i \,|\, x_i - F_i(x) \in (l_i, u_i)\}, \\ \beta(x) &:= \{i \,|\, x_i - F_i(x) \in \{l_i, u_i\}\}, \\ \gamma(x) &:= \{i \,|\, x_i - F_i(x) \notin [l_i, u_i]\}. \end{aligned}$$

Let us abbreviate the index sets $\alpha(x), \beta(x)$ and $\gamma(x)$ by α, β and γ, respectively. (Although the letters α and β are also used to define the D-gap function, we hope that this does not cause any confusion because both notations are standard in the literature.) Let H be an arbitrary element in $\partial_B r(x)$ and let $H_{i\cdot}$ denote the ith row of H. Then it follows immediately from the definition of r that

$$\begin{aligned} H_{i\cdot} = \nabla F_i(x)^T \quad & \text{if} \quad i \in \alpha, \\ H_{i\cdot} = \nabla F_i(x)^T \text{ or } H_{i\cdot} = e_i^T \quad & \text{if} \quad i \in \beta, \\ H_{i\cdot} = e_i^T \quad & \text{if} \quad i \in \gamma. \end{aligned}$$

This implies that, for some index set δ with $\emptyset \subseteq \delta \subseteq \beta$, H is expressed as

$$H = \begin{pmatrix} [F'(x)]_{\alpha\cup\delta,\alpha\cup\delta} & [F'(x)]_{\alpha\cup\delta,\bar{\delta}\cup\gamma} \\ 0_{\bar{\delta}\cup\gamma,\alpha\cup\delta} & I_{\bar{\delta}\cup\gamma,\bar{\delta}\cup\gamma} \end{pmatrix}, \tag{4}$$

where $\bar{\delta} := \beta \setminus \delta$ denotes the complement of δ in β. By exploiting this special structure of H, the Newton equation (3) can be written in the following reduced form:

$$\begin{aligned} [F'(x^k)]_{\alpha\cup\delta,\alpha\cup\delta}\, d_{\alpha\cup\delta} &= -r_{\alpha\cup\delta}(x^k) + [F'(x^k)]_{\alpha\cup\delta,\bar{\delta}\cup\gamma}\, r_{\bar{\delta}\cup\gamma}(x^k), \\ d_{\bar{\delta}\cup\gamma} &= -r_{\bar{\delta}\cup\gamma}(x^k), \end{aligned}$$

with $\alpha = \alpha(x^k), \beta = \beta(x^k)$ and $\gamma = \gamma(x^k)$.

The proposed algorithm uses the D-gap function to control the iterates generated by the Newton method. More precisely, we first try to accept the full step computed by the Newton equation (3). If this is not possible, we still try to accept the Newton direction. If this direction satisfies a sufficient descent condition, we accept it as a search direction used in the Armijo line search; otherwise, we switch to a steepest descent step.

The following is a precise description of the algorithm.

Algorithm 3.2 *(Globalized Nonsmooth Newton Method for BVIP)*

(S.0) Choose $x^0 \in \mathbb{R}^n, 0 < \alpha < 1 < \beta, \rho > 0, \lambda \in (0,1), \sigma \in (0,1/2), p > 1, \eta \in (0,1), \varepsilon \geq 0,$ *and set* $k := 0$.

(S.1) If $\|\nabla g_{\alpha\beta}(x^k)\| \leq \varepsilon$, *STOP.*

(S.2) Select an arbitrary element $H_k \in \partial_B r(x^k)$.

(a) Find a solution $d^k \in \mathbb{R}^n$ *of the Newton equation (3). If it is not solvable, then set* $d^k := -\nabla g_{\alpha\beta}(x^k)$ *and go to (S.3), else go to (b).*

(b) If

$$g_{\alpha\beta}(x^k + d^k) \leq \eta g_{\alpha\beta}(x^k),$$

then set $t_k := 1$ *and go to (S.4), else go to (c).*

(c) If d^k *does not satisfy the descent condition*

$$\nabla g_{\alpha\beta}(x^k)^T d^k \leq -\rho \|d^k\|^p,$$

then set $d^k := -\nabla g_{\alpha\beta}(x^k)$. *Go to (S.3).*

(S.3) Let t_k be the largest element in the set $\{1, \lambda, \lambda^2, \ldots\}$ such that

$$g_{\alpha\beta}(x^k + t_k d^k) \leq g_{\alpha\beta}(x^k) + \sigma t_k \nabla g_{\alpha\beta}(x^k)^T d^k.$$

(S.4) Set $x^{k+1} := x^k + t_k d^k, k \leftarrow k+1$, and go to (S.1).

Throughout this section, we assume that the termination parameter ε is equal to 0 and that Algorithm 3.2 generates an infinite sequence $\{x^k\}$.

An immediate consequence of the statement of Algorithm 3.2 is the fact that this method is well-defined for arbitrary BVIP as long as the function F is continuously differentiable.

The following convergence analysis borrows a number of ideas from some related papers. Our analysis will therefore be relatively short since we usually refer to these papers if the proofs are similar to existing ones.

For example, the global convergence result in the next theorem can be shown in exactly the same way as in [2] (see also Section 5.1 in [11]).

Theorem 3.3 *Every accumulation point of a sequence generated by Algorithm 3.2 is a stationary point of the D-gap function $g_{\alpha\beta}$.*

As shown in [12], a stationary point x^* of the D-gap function is a solution of BVIP if the Jacobian matrix $F'(x^*)$ is a P-matrix. This assumption holds, in particular, if F itself is a uniform P-function. Since the level sets of the D-gap function are compact for uniform P-functions (see [12]), it follows from Theorem 3.3 that, under this assumption, Algorithm 3.2 will generate a sequence that converges to the unique solution of BVIP.

To establish a fast local convergence property of the algorithm, we first need a preliminary result. We recall that a solution x^* of BVIP is said to be *b-regular* if the submatrices

$$[F'(x^*)]_{\alpha\cup\delta,\alpha\cup\delta}$$

with $\alpha = \alpha(x^*)$ and $\beta = \beta(x^*)$, are nonsingular for all index sets δ such that $\emptyset \subseteq \delta \subseteq \beta$. With this concept, we can prove the following nonsingularity result.

Proposition 3.4 *If x^* is a b-regular solution of BVIP, then all elements $H \in \partial_B r(x^*)$ are nonsingular.*

Proof. Observe that an arbitrary member H in $\partial_B r(x^*)$ can be represented as (4) with $\alpha = \alpha(x^*)$, $\beta = \beta(x^*)$ and $\emptyset \subseteq \delta \subseteq \beta$. Hence H is nonsingular if and only if the submatrix

$$[F'(x^*)]_{\alpha\cup\delta,\alpha\cup\delta}$$

is nonsingular; the latter readily follows from the very definition of a b-regular solution. □

Using the previous results, we can now prove the following local convergence results in the same way as in [2] (again, see also [11]). We stress, however, that the proof is based on the fact that the squared natural residual has the same growth behaviour as the D-gap function, see Proposition 2.3. In the following, we suppose that the parameters α and β involved in the definition of the D-gap function $g_{\alpha\beta}$ satisfy $0 < \alpha < 1 < \beta$.

Theorem 3.5 *Let $\{x^k\}$ be a sequence generated by Algorithm 3.2. Assume that x^* is an accumulation point of $\{x^k\}$ such that x^* is a b-regular solution of BVIP. Then*

(a) the entire sequence $\{x^k\}$ converges to x^*;

(b) the search direction d^k is eventually computed from the Newton equation (3);

(c) the test in Step (S.2) (b) is eventually accepted so that $t_k = 1$ for all k sufficiently large;

(d) the rate of convergence is Q-superlinear;

(e) if, in addition, F' is locally Lipschitzian, then the rate of convergence is Q-quadratic.

Finally we mention a finite termination property of the algorithm. The proof of this result can be carried out along the same lines as in [6] for the nonlinear complementarity problem. In particular, the proof heavily exploits the fact that the natural residual is piecewise smooth.

Theorem 3.6 *Let x^* be a b-regular solution of BVIP and F be an affine mapping. Then there is a neighbourhood of x^* such that, whenever an iterate x^k belongs to this neighbourhood, the next iterate x^{k+1} generated by Algorithm 3.2 is equal to the solution x^*.*

4 Preliminary Numerical Results

We implemented Algorithm 3.2 in a straightforward way using MATLAB 5.0 and tested it on a SUN SPARC 20 station on all large-scale problems from the MCPLIB test problem collection by Dirkse and Ferris [4]. In our numerical experiments, we replaced the standard Armijo rule in Step (S.3) of Algorithm 3.2 by a nonmonotone Armijo-rule [10]. We terminate the algorithm if

$$\|r(x^k)\| \le \varepsilon$$

holds for an iterate x^k.

The implementation uses the following parameters:

$$\alpha = 0.9,\ \beta = 1.1,\ \rho = 10^{-8},\ \lambda = 0.5,\ \sigma = 10^{-4},\ p = 2.1,\ \eta = 0.9,\ \varepsilon = 10^{-6}.$$

We report the results in Table 1. The columns of this table have the following meanings:

problem :	name of the test problem in MCPLIB
n :	number of variables
m :	number of (finite) bounds on the variables x_i
k :	number of iterations
F-eval. :	number of function evaluations
$g_{\alpha\beta}(x^f)$:	value of $g_{\alpha\beta}(x)$ at the final iterate $x = x^f$
$\|\nabla g_{\alpha\beta}(x^f)\|$:	value of $\|\nabla g_{\alpha\beta}(x)\|$ at the final iterate $x = x^f$
N :	number of Newton steps
G :	number of gradient (steepest descent) steps.

Table 1: Results for the large-scale problems from MCPLIB

problem	n	m	k	F-eval.	$g_{\alpha\beta}(x^f)$	$\|\nabla g_{\alpha\beta}(x^f)\|$	N	G
bert_oc	5000	2000	4	6	3.8e-28	6.1e-14	4	0
bratu	5625	11250	13	29	1.0e-20	1.9e-10	13	0
obstacle	2500	5000	10	11	3.3e-31	2.3e-15	10	0
opt_cont31	1024	1024	5	12	3.6e-30	4.3e-15	5	0
opt_cont127	4096	4096	8	32	1.4e-29	9.0e-15	8	0
opt_cont255	8192	8192	11	51	2.9e-29	1.2e-14	11	0
opt_cont511	16384	16384	12	52	6.0e-29	1.6e-14	12	0

We feel that the algorithm performs quite effectively on these large-scale problems. In particular, the number of iterations is pretty small for all test problems. Since we solve at each iteration just a system of linear equations, which, in fact, is of reduced dimension thanks to the special structure of the mapping r, the method seems to be very efficient.

On the other hand, we admit that the algorithm does not seem to be as robust as some other existing ones. In fact, when tested on smaller problems from MCPLIB, the algorithm sometimes converged to a mere local minimizer of the D-gap function. If, however, the conditions for global convergence are satisfied, the algorithm is expected to work quite well. This may be supported by our numerical experience since most of the large-scale problems in MCPLIB seem to satisfy those conditions.

5 Final Remarks

In this paper, we proposed a new algorithm for the solution of box constrained variational inequality problems (BVIP). The algorithm is a nonsmooth Newton-type method that is based on a reformulation of BVIP as a system of nonlinear equations and is endowed with a global convergence property by using the D-gap function.

The D-gap function is by no means the only choice for globalizing the algorithm. In fact, whenever we have a differentiable unconstrained minimization reformulation of BVIP, say

$$\min \ \Psi(x), \qquad x \in \mathbb{R}^n,$$

with a function $\Psi : \mathbb{R}^n \to \mathbb{R}$ such that Ψ has the same growth behaviour as the natural residual squared, i.e., such that a relationship like the one established in Proposition 2.3 holds, all convergence results remain valid even if we replace the D-gap function $g_{\alpha\beta}$ everywhere by the function Ψ. Proper choice of Ψ may actually improve the performance of Algorithm 3.2. This is an interesting topic for future research.

Finally, we mention that it seems possible to extend the results obtained in this paper to the general variational inequality problem where the feasible set is described by a general closed convex set. In fact, all results in Section 2 hold for the general problem, as well as most of the results in Section 3. The main problem with the general variational inequality problem is the calculation of the B-subdifferential of the projection operator on a convex set. One possibility would be to use the approximation scheme described in the paper [18] by Sun, Fukushima and Qi. We leave this also as a future research topic.

References

[1] P.H. Calamai and J.J. Moré, "Projected gradient methods for linearly constrained problems", *Mathematical Programming* **39** 93–116 (1987).

[2] T. De Luca, F. Facchinei and C. Kanzow, "A theoretical and numerical comparison of some semismooth algorithms for complementarity problems", Preprint, Computer Sciences Department, University of Wisconsin, Madison, WI, forthcoming.

[3] S.P. Dirkse and M.C. Ferris, "The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems", *Optimization Methods and Software* **5** 123–156 (1995).

[4] S.P. Dirkse and M.C. Ferris, "MCPLIB: A collection of nonlinear mixed complementarity problems", *Optimization Methods and Software* **5** 319–345 (1995).

[5] A. Fischer, "Solution of monotone complementarity problems with locally Lipschitzian functions", *Mathematical Programming* **76** 513–532 (1997).

[6] A. Fischer and C. Kanzow, "On finite termination of an iterative method for linear complementarity problems", *Mathematical Programming* **74** 279–292 (1996).

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

[8] M. Fukushima and J.-S. Pang, "Minimizing and stationary sequences of merit functions for complementarity problems and variational inequalities", in: M.C. Ferris and J.-S. Pang (eds.), *Complementarity and Variational Problems: State of the Art*, SIAM, Philadelphia, PA, 1997, pp. 91–104.

[9] E.M. Gafni and D.P. Bertsekas, "Two-metric projection methods for constrained optimization", *SIAM Journal on Control and Optimization* **22** 936–964, (1984).

[10] L. Grippo, F. Lampariello and S. Lucidi, "A nonmonotone linesearch technique for Newton's method", *SIAM Journal on Numerical Analysis* **23** 707–716 (1986).

[11] C. Kanzow, *Semismooth Newton-type Methods for the Solution of Nonlinear Complementarity Problems*, Habilitation Thesis, Institute of Applied Mathematics, University of Hamburg, Hamburg, Germany, April 1997.

[12] C. Kanzow and M. Fukushima, "Theoretical and numerical investigation of the D-gap function for box constrained variational inequalities", *Mathematical Programming*, to appear.

[13] R. Mifflin, "Semismooth and semiconvex functions in constrained optimization", *SIAM Journal on Control and Optimization* **15** 959–972 (1977).

[14] J.-M. Peng, "Equivalence of variational inequality problems to unconstrained optimization", *Mathematical Programming* **78** 347–356 (1997).

[15] J.-M. Peng and M. Fukushima, "A hybrid Newton method for solving the variational inequality problem via the D-gap function", Technical Report, Department of Applied Mathematics and Physics, Kyoto University, Kyoto, Japan, May 1997.

[16] L. Qi, "Convergence analysis of some algorithms for solving nonsmooth equations", *Mathematics of Operations Research* **18** 227–244 (1993).

[17] L. Qi, "Regular pseudo-smooth NCP and BVIP functions and globally and quadratically convergent generalized Newton methods for complementarity and variational inequality problems", Applied Mathematics Report 97/14, School of Mathematics, The University of New South Wales, Sydney, Australia, July 1997 (revised September 1997).

[18] D. Sun, M. Fukushima and L. Qi, "A computable generalized Hessian of the D-gap function and Newton-type methods for variational inequality problems", in: M.C. Ferris and J.-S. Pang (eds.): *Complementarity and Variational Problems: State of the Art*, SIAM, Philadelphia, PA, 1997, pp. 452–473.

[19] N. Yamashita, K. Taji and M. Fukushima, "Unconstrained optimization reformulations of variational inequality problems", *Journal of Optimization Theory and Applications* **92** 439–456 (1997).