Journal of Global Optimization 6: 401–425, 1995.

On numerical solution of hemivariational inequalities by nonsmooth optimization methods

M. MIETTINEN and M.M. MKEL
Laboratory of Scientific Computing
Department of Mathematics, University of Jyvskyl
P.O. Box 35, FIN-40351 Jyvskyl, Finland

and

J. HASLINGER
Faculty of Mathematics and Physics, Charles University, Prague
Ke Karlovu 5, 12000 Praha 2, Czech Republic
and
Laboratory of Scientific Computing
Department of Mathematics, University of Jyvskyl
P.O. Box 35, FIN-40351 Jyvskyl, Finland

Abstract. In this paper we consider numerical solution of hemivariational inequalities (HVI) by using nonsmooth, nonconvex optimization methods. First we introduce a finite element approximation of (HVI) and show that it can be transformed to a problem of finding a substationary point of the corresponding potential function. Then we introduce a proximal budle method for nonsmooth nonconvex and constrained optimization. Numerical results of a nonmonotone contact problem obtained by the developed methods are also presented.

Key words: hemivariational inequalities, finite element method, substationary point, nonsmooth optimization, bundle methods, nonmonotone contact problem

1. Introduction

Hemivariational inequalities, generalizations of variational inequalities, were presented by Panagiotopoulos [20]-[22]. Their origin is in nonsmooth mechanics of solid, especially in nonmonotone contact problems. We refer the reader to [18],[22] and references therein for the mathematical theory and the applications of them.

In this paper we present a fully discrete approximation model of (HVI) based on the finite element technique and show that it can be numerically realized by using nonsmooth, nonconvex optimization method. This model was introduced first for scalar–valued (HVI) in [10],[11],[13],[14] and then it was extended for vector–valued (HVI) in [12]. It is applicable for the unconstrained (HVI) problems and the constrained ones with a nonempty, closed, convex constraint set. Furthermore, it can also used for the so-called variational–hemivariational inequalitites to which class the constrained (HVI) problems belong as special cases (see [13]). In the case of variational inequalities it can be shown that our approximation is equivalent to the classical one presented in [4],[5]. The common property of the (HVI) problems, to which we have applied our approximation, is that the nonsmooth, nonmonotone behaviour of the problem is concentrated on the lower order terms, i.e. on the terms which do not contain the highest order derivatives. This is because the main tool, which is used for showing the convergence of the nonsmooth, nonmonotone terms in our approximation as the discretization parameter (which is in the finite element technique the size of the mesh of the partitions) tends to zero, is that the generalized directional derivative is upper semicontinuous (see [3]). And this approach cannot use for the terms which contain the highest order derivatives which converge only weakly in the considered function spaces. This is the main drawback and restriction which we have in the approximation theory of the (HVI) problems (and also in the general theory of (HVI)) compared to the corresponding one of variational inequalities in which one can exploit effectively the monotone nature, and consequently the convex nature of the problems.

The outline of this paper is as follows. In the second section we formulate the considered vector–valued (HVI) which can also have constraints. For simplicity we have restricted ourselves to the case having a polynomial growth condition for the nonsmooth, nonmonotone term. For more general cases we refer to [18],[19]. Then we present a fully discrete FEM–approximation for it. It can be shown that the solutions of the discrete problems converge strongly on subsequences to the solutions of the continuous one (see the proof in [12]). In the third section we study the substationary points of the corresponding nonconvex potential functions of the continuous and discrete (HVI) problems. By a substationary point we mean that 0 belongs to the sum of the generalized gradient of the potential function and the normal cone of the constraint set. We show that the substationary points of the potential functions are also the solutions of the (HVI) problems and this holds for the both problems, the continuous and discrete ones. This is gives us the theoretical basis to numerically solve the discrete (HVI) problem by transforming it to a problem of finding of a substationary point of the nonconvex locally Lipschitz continuous function. Finally we consider the question if the substationary points are preserved as the discretization parameter tends to zero: we can only show that the global minima are preserved. The subsection 4 is devoted to the question how to generate substationary points of the locally Lipschitz continuous function. We introduce a proximal bundle method for nonsmooth nonconvex and constrained optimization. Our method is a generalization of the proximal bundle method by [9] to the nonconvex constrained case. It is based on the method derived in [15] and it has also close relationship with the bundle trust method of [23]. In the last section we study in detail an example of (HVI), namely a linear elastic contact problem with a nonmonotone frictionless foundation or with a rigid frictionless foundation and a nonmonotone layer above it. They can be formulated matematically as a unconstrained (HVI) problem or a constrained (HVI) problem, respectively. We apply our approximation model to them and solve the discrete problems numerically by transforming them to nonsmooth minimization problems and using then the proximal bundle method introduced in the fourth section. For the other numerical methods which can be applied to the presented discrete (HVI) problems we refer to [22].

2. Hemivariational inequalities and their finite element approximation

2.1. Formulation of the continuous problem

Let V be a real Hilbert space and $\Omega \subset {}^N$ be a bounded domain with Lipschitz boundary Γ. We shall denote by $\|\cdot\|_V$ the norm of V, V' the dual space of V and $\langle\cdot,\cdot\rangle_V$ the corresponding duality pairing. It will be supposed that

$$V \text{ is compactly imbedded in } L^2(\Omega_0; {}^M),\ \Omega_0 \subset \Omega \tag{1}$$

or

$$V \text{ is compactly imbedded in } L^2(\Gamma_0; {}^M),\ \Gamma_0 \subset \Gamma. \tag{2}$$

We shall also have a nonempty, closed, convex subset K of V. Let j be a locally Lipschitz continuous function from ^{M} to satisfying firstly *the generalized sign condition* which is expressed by means of the generalized directional derivative (see [3])

$$j^\circ(\xi; -\xi) \le C_1 + C_2|\xi| \quad \forall \xi \in {}^M, \tag{3}$$

and secondly *the growth condition* expressed by means of the generalized gradient of j (see [3])

$$\eta \in \partial j(\xi) \Longrightarrow |\eta| \le C_3(1 + |\xi|), \tag{4}$$

where C_1, C_2 and C_3 are positive constants independent of ξ and η. Let $a : V \times V \to$ be a bilinear form satisfying the continuity and the coerciveness conditions (α, m positive constants):

$$|a(v,w)| \le m\|v\|_V\|w\|_V \qquad \forall v, w \in V; \tag{5}$$

$$a(v,v) \ge \alpha\|v\|_V^2 \qquad \forall v \in V, \tag{6}$$

and g be an element of V'. By a hemivariational inequality we mean the following problem (if (1) holds):

$$\begin{cases} \text{find } u \in K \text{ and } \mathcal{X}(u) \in L^2(\Omega_0; {}^M) \text{ such that} \\ a(u, v-u) + \int_{\Omega_0} \mathcal{X} \cdot (v-u)\,dx \ge \langle g, v-u\rangle_V \quad \forall v \in K \\ \text{and } \mathcal{X}(x) \in \partial j(u(x)) \quad \text{a.e. in } \Omega_0 \end{cases} \tag{P1}$$

or (if (2) holds)

$$\begin{cases} \text{find } u \in K \text{ and } \mathcal{X}(u) \in L^2(\Gamma_0; {}^M) \text{ such that} \\ a(u, v-u) + \int_{\Gamma_0} \mathcal{X} \cdot (v-u)\, dx \geq \langle g, v-u \rangle_V \quad \forall v \in K \\ \text{and } \mathcal{X}(x) \in \partial j(u(x)) \quad \text{a.e. in } \Gamma_0. \end{cases} \tag{P2}$$

In the sequel we shall use the symbol (P) if we mean the both problems (P1) and (P2) (we shall use this convention also in other notations).

THEOREM 1. *There exists at least one solution of the problem (P).*

For the proof of the above theorem we refer to [18],[19].

2.2. Formulation of the discrete problem

The approximation of the problem (P) is constructed by using the finite element technique. Let $h \in (0,1)$ be a discretization parameter which is related to the mesh size of the partitions used for the constructions of FEM-spaces. First we introduce finite-dimensional approximations V_h and Y_h of the spaces V and $Y_1 = L^2(\Omega_0; {}^M)$ (or $Y_2 = L^2(\Gamma_0; {}^M)$). In order to construct V_h we can use the standard FEM–approach: Let $\{V_h\}_{h \in (0,1)}$, $V_h \subset C(\overline{\Omega}; {}^M)$, be a family of finite-dimensional subspaces of V satisfying the condition

$$\forall v \in V \ \{v_h\},\ v_h \in V_h : v_h \to v \text{ in } V \text{ as } h \to 0_+. \tag{7}$$

If $\Omega \subset {}^2$ is a polygon, V_h can be, for example, a space of piecewise linear functions over some regular triangulation $\mathcal{T}_h$ of $\overline{\Omega}$ (see [2]). Furthermore, we need to approximate the convex set K: Let $\{K_h\}_{h \in (0,1)}$ be a family of nonempty, closed, convex subsets of V_h satisfying

$$\forall v \in K \ \exists \{v_h\},\ v_h \in K_h : v_h \to v \text{ in } V \text{ as } h \to 0_+; \tag{8}$$

$$\{v_h\},\ v_h \in K_h : v_h \rightharpoonup v \text{ in } V \text{ as } h \to 0_+ \Longrightarrow v \in K \tag{9}$$

(see [4]–[6]).

For constructing the FEM–space Y_h we have to be more careful. As in the approximation of the variational inequalities of the second kind (see [4],[5]) we first fix a quadrature formula

$$\int_{\Omega_0} f(x)\, dx \approx \sum_{i=1}^{m_h} c_h^i f(x_h^i) \quad \left(\text{or } \int_{\Gamma_0} f(x)\, dx \approx \sum_{i=1}^{m_h} c_h^i f(x_h^i)\right), \tag{10}$$

where c_h^i are weights and x_h^i are nodal points of the quadrature formula, which we use to approximate the integral $\int_{\Omega_0} \mathcal{X} \cdot v\, dx$ (or $\int_{\Gamma_0} \mathcal{X} \cdot v\, dx$) (in [4],[5] it is used for approximating the convex function $J(u) = \int_{\Omega_0} j(u(x))\, dx$ (or $J(u) = \int_{\Gamma_0} j(u(x))\, dx$)). Then we define another partition $\mathcal{T}_h'$ of $\overline{\Omega}_h$ (or Γ_0), $\Omega_0 \subset \Omega_h$ satisfying

(i) $\overline{\Omega}_h = \cup_{i=1}^{m_h} K_h^i$;
(ii) $\max_{i=1,\dots,m_h}\{$ diameter of $K_h^i\} \leq h$;
(iii) int $K_h^i \cap$ int $K_h^j = \emptyset \; \forall i \neq j$;
(iv) K_h^i is closed and has a nonempty interior for each $i = 1, \dots, m_h$;
(v) for each $i = 1, \dots, m_h$ there is exactly one point $x_h^i \in$ int $K_h^i \cap \overline{\Omega}$;
(vi) $m_N($ int $K_h^i \cap \Omega) = c_h^i$, $i = 1, \dots, m_h$ (m_N is the Lebesgue measure in ^{N}).

The space Y_h is defined such that it contains all restrictions to Ω_0 of piecewise constant functions over $\mathcal{T}_h'$, i.e.,

$$Y_h = \{f = (f_1, \dots, f_M) \in L^\infty(\Omega_0; {}^M) : \exists \tilde{f} = (\tilde{f}_1, \dots, \tilde{f}_M) : \Omega_h \to {}^M,$$
$$\tilde{f}_j|_{\text{int } K_h^i} \text{ is constant } i = 1, \dots, m_h, j = 1, \dots, M, f = \tilde{f}|_{\Omega_0}\}.$$

We define in a similar way X_h a space of functions components of which are piecewise continuous functions over the partition $\mathcal{T}_h'$:

$$X_h = \{f = (f_1, \dots, f_M) \in L^\infty(\Omega_0; {}^M) : \exists \tilde{f} = (\tilde{f}_1, \dots, \tilde{f}_M) : \Omega_h \to {}^M,$$
$$\tilde{f}_j|_{\text{int } K_h^i} \text{ is continuous } i = 1, \dots, m_h, j = 1, \dots, M, f = \tilde{f}|_{\Omega_0}\}.$$

To define the approximation problem we have to define also a linear mapping $P_h : X_h \to Y_h$, the so called mass lumping operator:

$$(P_h f)(x) = \sum_{i=1}^{m_h} f(x_h^i)(\mathcal{X}_{\text{int } K_h^i})(x), \quad x \in \Omega_0,$$

where $\mathcal{X}_{\text{int } K_h^i}$ is the characteristic function of int K_h^i. The following consistency conditions between the spaces V_h and Y_h are assumed:

$$v_h \rightharpoonup v \text{ in } V \text{ as } h \to 0_+ \implies \text{there exists} \tag{11}$$
$$\text{a subsequence of } \{v_h\} \text{ such that } P_{h'} v_{h'} \to v \text{ in } Y_1 \text{ as } h' \to 0_+;$$
$$\|P_h\|_{\mathcal{L}(V_h, Y_h)} \leq C_4, \tag{12}$$

where C_4 is a positive constant independent of h. For the problem (P2) we can define the partition $\mathcal{T}_h'$ over Γ_0, the spaces Y_h,X_h and the linear operator P_h in a similar way.

It remains only to approximate the bilinear form a and the linear form $\langle g, \cdot \rangle_V$. This can be done by using the standard approach, i.e. using appropriate numerical integration formulae (see [2]): Let $a_h : V_h \times V_h \to$ be an approximation of a satisfying the following properties:

$$\exists \tilde{m} > 0 : |a_h(u_h, v_h)| \leq \tilde{m}\|u_h\|_V \|v_h\|_V \;\; \forall u_h, v_h \in V_h, \forall h \in (0,1); \tag{13}$$
$$\exists \tilde{\alpha} > 0 : a_h(v_h, v_h) \geq \tilde{\alpha}\|v_h\|_V^2 \;\; \forall v_h \in V_h, \;\; \forall h \in (0,1); \tag{14}$$
$$u_h \rightharpoonup u, v_h \to v \text{ in } V \text{ as } h \to 0_+, \;\; u_h, v_h \in V_h \implies \tag{15}$$
$$a_h(u_h, v_h) \to a(u,v) \text{ and } a_h(v_h, u_h) \to a(v,u) \text{ as } h \to 0+$$

and let $g_h \in V_h'$ be an approximation of g such that

$$\exists \tilde{\beta} > 0 : |\langle g_h, v_h \rangle_{V_h}| \leq \tilde{\beta} \|v_h\|_V \ \forall v_h \in V_h, \ \forall h \in (0,1); \tag{16}$$

$$v_h \rightharpoonup v \text{ in } V \text{ as } h \to 0_+, \ v_h \in V_h \Longrightarrow \tag{17}$$

$$\langle g_h, v_h \rangle_{V_h} \to \langle g, v \rangle_V \text{ as } h \to 0_+,$$

where V_h' is the dual space of V_h and $\langle \cdot, \cdot \rangle_{V_h}$ the corresponding duality pairing.

Using the defined notations we are now able to define fully discrete FEM–approximations of the problems (P1) and (P2) as follows:

$$\left\{ \begin{array}{l} \text{find } u_h \in K_h \text{ and } \mathcal{X}_h(u_h) \in Y_h \text{ such that} \\ a_h(u_h, v_h - u_h) + \int_{\Omega_0} \mathcal{X}_h \cdot (P_h v_h - P_h u_h)\, dx \\ \geq \langle g_h, v_h - u_h \rangle_{V_h} \quad \forall v_h \in K_h \\ \text{and } \mathcal{X}_h(x) \in \partial j((P_h u_h)(x)) \quad \text{a.e. in } \Omega_0 \end{array} \right. \tag{P1$_h$}$$

and

$$\left\{ \begin{array}{l} \text{find } u_h \in K_h \text{ and } \mathcal{X}_h(u_h) \in Y_h \text{ such that} \\ a_h(u_h, v_h - u_h) + \int_{\Gamma_0} \mathcal{X}_h \cdot (P_h v_h - P_h u_h)\, dx \\ \geq \langle g_h, v_h - u_h \rangle_{V_h} \quad \forall v_h \in K_h \\ \text{and } \mathcal{X}_h(x) \in \partial j((P_h u_h)(x)) \quad \text{a.e. in } \Gamma_0. \end{array} \right. \tag{P2$_h$}$$

It is possible to show that firstly the discrete problems are solvable and secondly that the discrete problems are closed on subsequences to the continuous ones, i.e.

THEOREM 2. *There exists at least one solution $(u_h, \mathcal{X}_h(u_h))$ of (P)$_h$ for all $h \in (0,1)$ and we can find a subsequence of $\{(u_h, \mathcal{X}_h(u_h))\}$ such that $u_{h'}$ converges strongly to u in V and $\mathcal{X}_{h'}(u_{h'})$ converges weakly to $\mathcal{X}$ in Y. Moreover, $(u, \mathcal{X})$ is a solution of (P).*

For the proof of this theorem we refer to [12].

3. Substationary points of the corresponding nonconvex energy functions

Throughout this section we shall assume that the bilinear form a and its approximation a_h, $h \in (0,1)$, are symmetric. Let us define a function L from V to as follows:

$$L(v) = \frac{1}{2} a(v,v) - \langle g, v \rangle_V + J(v), \tag{18}$$

where

$$J(v) = \int_{\Omega_0} j(v(x))\, dx, \quad (\text{or } J(v) = \int_{\Gamma_0} j(v(x))\, dx), \tag{19}$$

and its approximation $L_h : V_h \to , h \in (0,1)$:

$$L_h(v_h) = \frac{1}{2}a_h(v_h, v_h) - \langle g_h, v_h\rangle_{V_h} + J_h(v_h), \tag{20}$$

where $J_h(v_h) = J(P_h v_h)$ for all $v_h \in V_h$. Since the function j is locally Lipschitz continuous and it satisfies (4), it is easy to see that J and J_h are also locally Lipschitz continuous, and consequently L and L_h.

The main aim in this section is to show that all substationary points of L on K and L_h on K_h are solutions of (P) and $(\mathrm{P})_h$, respectively. By a substationary point we shall mean the following:

DEFINITION 1. *Let f be a locally Lipschitz continuous function from a Banach space X to . A point $x \in X$ is called a substationary point of f iff $0 \in \partial f(x) + N_K(x)$, where $\partial f(x)$ is the generalized gradient of f at x and $N_K(x)$ the normal cone to K at x.*

In the sequel we shall consider only the case (P1), because (P2) can be treated in a similar way.

PROPOSITION 1. *It holds that every substationary point of L is a solution of (P1).*

PROOF: Let u be a substationary point of L on K, i.e.

$$0 \in \partial L(u) + N_K(u) = Au + \partial J(u) - g + N_K(u), \tag{21}$$

where $A : V \to V'$ is defined by $\langle Av, w\rangle_V = a(v, w)$ for all $v, w \in V$. The equality holds in (21) due to Corollary 1 of Proposition 2.3.3 in [3]. Therefore, there exist $\mathcal{X} \in \partial J(u)$ and $w \in N_K(u)$ such that

$$0 = Au + \mathcal{X} - g + w. \tag{22}$$

¿From [1] we know that $\mathcal{X}$ satisfies the relation

$$\mathcal{X}(x) \in \partial j(u(x)) \quad \text{a.e. in } \Omega_0. \tag{23}$$

Using the definition of the normal cone, i.e. $\langle w, v\rangle_V \leq 0$ for all $v \in T_K(u)$, $T_K(u)$ the tangent cone of K at u, and the fact that K is a convex set we have that

$$\langle w, v - u\rangle_V \leq 0 \quad \forall v \in K. \tag{24}$$

Substituting (23),(24) to (22) we obtain

$$\begin{cases} 0 = \langle Au + \mathcal{X} - g, v - u\rangle_V + \langle w, v - u\rangle_V \\ \leq \langle Au + \mathcal{X} - g, v - u\rangle_V \quad \forall v \in K \\ \text{and } \mathcal{X}(x) \in \partial j(u(x)) \quad \text{a.e. in } \Omega_0, \end{cases}$$

i.e. u solves (P1). Thus the proof is complete. □

Next we shall show that the corresponding result holds also in the discrete case. First we shall express the problem $(\mathrm{P1})_h$ in a matrix form (as the discretization parameter is fixed, we shall skip it in many situations). Let $\{\varphi^j\}_{j=1}^n$ be the basis of V_h. We shall make identifications $V_h \equiv {}^n$ and $Y_h \equiv [{}^M]^m$, where m is the number of the nodal points of the quadrature formula (10). Moreover, we shall identify $v = \sum_{j=1}^n v_j \varphi^j \in V_h$ with the nodal vector $\mathrm{v} = (v_j)_{j=1}^n \in {}^n$. Let us define an $m \times n$–matrix $\mathcal{P} = (\mathbf{P}_{ij})$, $\mathbf{P_{ij}} \in {}^{\mathbf{M}}$, as follows

$$\mathbf{P}_{ij} = (P_h \varphi^j)(x^i), \quad i = 1, ..., m, j = 1, ..., n, \tag{25}$$

where $x^i, i = 1, ..., m$ are the nodal point of (10). Let us use the following notations: $\mathbf{K} = \{\mathrm{v} \in {}^n : v \in K_h\}$, $\mathbf{A} = (a_h(\varphi^i, \varphi^j))_{i,j=1}^n$ an $n \times n$–matrix and $\mathbf{g} = (\langle g_h, \varphi^j \rangle_{V_h})_{j=1}^n \in {}^n$. Then the problem $(\mathrm{P1})_h$ is equivalent to

$$\begin{cases} \text{find } \mathbf{u} = (u_1, ..., u_n) \in \mathbf{K} \text{ and } \mathrm{s} = (\mathrm{s}_1, ..., \mathrm{s}_m) \in [{}^M]^m \text{ such that} \\ (\mathrm{v} - \mathbf{u})^T \mathbf{A} \mathbf{u} + (\mathcal{P}\mathrm{v} - \mathcal{P}\mathbf{u})^T \mathrm{s} \geq (\mathrm{v} - \mathbf{u})^T \mathbf{g} \quad \forall \mathrm{v} \in \mathbf{K} \\ \text{and } \mathrm{s}_i \in c_i \partial j((\mathcal{P}\mathrm{u})_i) \quad i = 1, ..., m. \end{cases}$$

PROPOSITION 2. *It holds that every substationary point of L_h is a solution of $(P1)_h$.*

PROOF: The function $\mathbf{J} : {}^n \to {}$, $\mathbf{J}(\mathrm{v}) = J_h(v)$ for all $v \in V_h$, can be written as a composite function $\hat{\mathbf{J}} \circ \mathcal{P}$, where $\mathcal{P}$ is the $m \times n$–matrix defined by (25) and $\hat{\mathbf{J}} : [{}^M]^m \to {}$ is a function defined by

$$\hat{\mathbf{J}}(\mathrm{v}) = \sum_{i=1}^m c_i j(\mathrm{v}_i). \tag{26}$$

The generalized directional derivative of $\hat{\mathbf{J}}$ can be estimated as follows:

$$\begin{aligned} \hat{\mathbf{J}}^\circ(\mathrm{v}; \mathrm{w}) &= \limsup_{\mathbf{z} \to 0, t \to 0_+} \frac{\hat{\mathbf{J}}(\mathrm{v} + \mathrm{z} + t\mathrm{w}) - \hat{\mathbf{J}}(\mathrm{v} + \mathrm{z})}{t} \\ &= \limsup_{\mathbf{z} \to 0, t \to 0_+} \frac{\sum_{i=1}^m c_i j(\mathrm{v}_i + \mathrm{z}_i + t\mathrm{w}_i) - \sum_{i=1}^m c_i j(\mathrm{v}_i + \mathrm{z}_i)}{t} \\ &\leq \sum_{i=1}^m c_i \limsup_{\mathbf{z}_i \to 0, t_i \to 0_+} \frac{j(\mathrm{v}_i + \mathrm{z}_i + t_i \mathrm{w}_i) - j(\mathrm{v}_i + \mathrm{z}_i)}{t_i} \\ &= \sum_{i=1}^m c_i j^\circ(\mathrm{v}_i; \mathrm{w}_i) \end{aligned} \tag{27}$$

This implies that the following holds: if $\mathrm{s} \in \partial \hat{\mathbf{J}}(\mathrm{v})$ its components necessarily satisfy the relation

$$\mathrm{s}_i \in c_i \partial j(\mathrm{v}_i) \quad i = 1, ..., m. \tag{28}$$

Furthermore, applying Theorem 2.3.10 of [3] we get that every element $\mathrm{s}' \in \partial \mathbf{J}(\mathrm{w})$ can be decomposed into the form

$$\mathrm{s}' = \mathcal{P}^T \mathrm{s}, \tag{29}$$

where $\mathrm{s} \in \partial \hat{\mathbf{J}}(\mathcal{P}\mathrm{w})$.

Let $\mathbf{u}$ be a substationary point of L_h on $\mathbf{K}$. Using the same arguments as in the proof of Proposition 1 we get firstly that

$$0 \in \mathbf{A}\mathbf{u} + \partial \mathbf{J}(\mathbf{u}) - \mathbf{g} + N_{\mathbf{K}}(\mathbf{u}), \tag{30}$$

and secondly that there exists $\mathrm{s}' \in \partial \mathbf{J}(\mathbf{u})$ such that

$$(\mathrm{v} - \mathbf{u})^T \mathbf{A}\mathbf{u} + (\mathrm{v} - \mathbf{u})^T \mathrm{s}' \geq (\mathrm{v} - \mathbf{u})^T \mathbf{g} \quad \forall \mathrm{v} \in \mathbf{K}. \tag{31}$$

Substituting (28) and (29) into (31) we get that every substationary point of L_h is a solution of the following problem

$$\begin{cases} \text{find } \mathbf{u} = (u_1, ..., u_n) \in \mathbf{K} \text{ and } \mathrm{s} = (\mathrm{s}_1, ..., \mathrm{s}_m) \in [^M]^m \text{ such that} \\ (\mathrm{v} - \mathbf{u})^T \mathbf{A}\mathbf{u} + (\mathcal{P}\mathrm{v} - \mathcal{P}\mathbf{u})^T \mathrm{s} \geq (\mathrm{v} - \mathbf{u})^T \mathbf{g} \quad \forall \mathrm{v} \in \mathbf{K} \\ \text{and } \mathrm{s}_i \in c_i \partial j((\mathcal{P}\mathbf{u})_i) \quad i = 1, ..., m, \end{cases}$$

which is nothing else that $(\mathrm{P1})_h$. Thus the proof of this proposition is complete. □

Combining Theorem 2 and Proposition 2 we see that one possible way to numerically solve the problem (P) is to transform the discrete problem $(\mathrm{P})_h$ to the problem of finding a substationary point of the corresponding potential function L_h, because it holds that the substationary points of L_h tend to the solutions of (P) on subsequences. And now because local minima of a locally Lipschitz continuous functions are its substationary points, we can use nonsmooth optimization methods (see the next section) for finding some of the solutions of (P). On the other hand if we want to numerically solve the problem of finding a substationary point of the function L, which is more restricted than the problem (P), we cannot use exactly the same approach. This is because of the fact that the limits (not even limits of subsequences) of the substationary points of L_h are not necessarily substationary points of the function L. The only thing what we can show is that the global minima are preserved on subsequences as h tends to 0. That is why one has to use global nonsmooth optimization methods for solving the substationary point problem of L.

For the completeness let us prove the above mentioned result of the global minima.

PROPOSITION 3. *The global mimina of the function L_h, $h \in (0,1)$, converge strongly in V on subsequences to the global minima of L.*

PROOF: Using the well–known result of the calculus of variations that a weakly lower semicontinuos and coercive function defined on a nonempty, closed, convex set of a reflexive Banach space has at least one minimum point, we obtain that the functions L and L_h, $h \in (0,1)$ have a minimum point. The coerciveness holds due to the sign condition (3), the coerciveness conditions of a,a_h (6),(14) and the weak lower semicontinuity is satisfied because of the compact imbeddings (1),(2), the weak lower semicontinuity of $\frac{1}{2}a(\cdot,\cdot)$,$\frac{1}{2}a_h(\cdot,\cdot)$.

Let $\{u_h\}$, $u_h \in V_h$ be a sequence such that u_h is a minimum point of L_h on K_h. Then

$$\begin{aligned} &\frac{1}{2}a_h(u_h,u_h) - \langle g_h, u_h\rangle_{V_h} + J(P_h u_h) \\ &\le \frac{1}{2}a_h(v_h,v_h) - \langle g_h, v_h\rangle_{V_h} + J(P_h v_h) \quad \forall v_h \in K_h. \end{aligned} \tag{32}$$

¿From Theorem 2 and Proposition 2 it follows that we have a subsequence $\{u_{h'}\}$ which converges strongly to some element $u \in K$ in V. Passing again to a subsequence if necessary we also have that $\{P_{h'}u_{h'}\}$ converges strongly to u in Y due to (11). Now it is easy to show that u is a minimum point of L on K. Indeed: Let $v \in K$ be given. Because of (7) there exists a sequence $\{v_h\}$, $v_h \in K_h$ such that $v_h \to v$ in V. Letting $h \to 0_+$ in (32) we get that

$$\begin{aligned} &\frac{1}{2}a(u,u) - \langle g, u\rangle_V + J(u) \\ &\le \frac{1}{2}a(v,v) - \langle g, v\rangle_V + J(v) \quad \forall v \in K, \end{aligned} \tag{33}$$

i.e. u is a global minumum point of L on K. □

4. Proximal bundle method for nonsmooth nonconvex optimization

In this section we concentrate on the question how to generate substationary points of the locally Lipschitz continuous function f from ^{n} to on the subset $K \subset {}^n$.

PROPOSITION 4. *It holds that every local minimizer is a substationary point of f on K.*

For the proof we refer to [3].

Due to the Proposition 4 we consider the following nonsmooth, nonconvex and constrained optimization problem

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to} & x \in K. \end{cases} \tag{GP}$$

In what follows our feasible set K has a more specific structure, i.e.

$$K = \{x \in {}^n \mid g(x) = \max_{i=1,\dots,m} g_i(x) \leq 0\},$$

where each g_i from ^{n} to is a locally Lipschitz continuous function. We suppose that at each $x \in {}^n$ we can evaluate the function values $f(x)$, $g(x)$ and arbitrary subgradients $\xi^f \in \partial f(x)$, $\xi^g \in \partial g(x)$.

The nonsmooth optimization methods for solving (GP) can be divided into two main classes: (Kiev) subgradient methods and bundle methods. The principle behind subgradient methods is to generalize smooth gradient or quasi-Newton methods by replacing the gradient by an arbitrary subgradient. This simple idea leads, however, to difficulties with *a priori* choice of the step size in line search operation and the lack of an implementable stopping criterion.

In this paper we construct a bundle method for the problem (GP). It is a generalization of the method introduced in [9] to the nonconvex constrained case, and is based on the method derived in [15]. For further study of bundle methods we refer to [16] and [17].

4.1. Direction finding

The idea of our method is to form a simpler approximation for the problem (GP). Suppose that the starting point x_1 is feasible and at the k-th iteration of the algorithm we have the current iteration point $x_k \in {}^n$, some auxiliary points $y_j \in {}^n$ previous iterations and corresponding subgradients $\xi_j^f \in \partial f(y_j)$ for $j \in J_f^k$ and $\xi_j^g \in \partial g(y_j)$ for $j \in J_g^k$, where the index sets J_f^k, $J_g^k \subset \{1, \dots, k\}$ are assumed to be nonempty. We define the linearizations at $x \in {}^n$ by

$$\bar{f}_j(x) = f(y_j) + (\xi_j^f)^T(x - y_j) \quad \text{for all} \quad j \in J_f^k \quad \text{and} \tag{34}$$

$$\bar{g}_j(x) = g(y_j) + (\xi_j^g)^T(x - y_j) \quad \text{for all} \quad j \in J_g^k. \tag{35}$$

Note that we do not need to store the auxiliary points y_j, since by denoting $f_j^k = \bar{f}_j(x_k)$ and $g_j^k = \bar{g}_j(x_k)$ the we obtain the following recursive updating formula

$$f_j^{k+1} = f_j^k + (\xi_j^f)^T(x_{k+1} - x_k) \quad \text{for all} \quad j \in J_f^k \quad \text{and} \tag{36}$$

$$g_j^{k+1} = g_j^k + (\xi_j^g)^T(x_{k+1} - x_k) \quad \text{for all} \quad j \in J_g^k. \tag{37}$$

Furthermore, for all $x \in {}^n$ we define the polyhedral approximations by

$$\hat{f}^k(x) = \max\{\bar{f}_j(x) \mid j \in J_f^k\} \tag{38}$$

$$\hat{g}^k(x) = \max\{\bar{g}_j(x) \mid j \in J_g^k\} \quad \text{and} \tag{39}$$

$$\hat{H}^k(x) = \max\{\hat{f}^k(x) - f(x_k),\ \hat{g}^k(x)\}, \tag{40}$$

which all are convex functions.

As in the classical cutting plane method [7] we replace the original objective and constraint functions by their polyhedral approximations. By using the function $\hat{H}^k$ we get over the constraints. In order to avoid the poor convergence rate of the cutting plane method, we add to the objective function a regularizing quadratic penalty term $\frac{1}{2}||d||^2$. To improve further the method we exploit the proximal bundle idea due to [9] and [23], and we multiply this penalty term by a weight $u_k > 0$ to obtain the following unconstrained cutting plane approximation of (GP)

$$\begin{cases} \text{minimize} & \hat{H}^k(x_k + d) + \frac{u_k}{2}||d||^2 \\ \text{subject to} & d \in {}^n. \end{cases} \tag{CP}$$

Due to the nonconvexity we define the distance measure by

$$s_j^k = ||x_j - y_j|| + \sum_{i=j}^{k-1} ||x_{i+1} - x_i|| \quad \text{for} \quad j = 1, \dots, k-1, \tag{41}$$

$$s_k^k = ||x_k - y_k|| \tag{42}$$

and the subgradient locality measures by

$$\beta_{f,j}^k = \max\{|f(x_k) - f_j^k|,\ \gamma_f(s_j^k)^2\} \quad \text{for all} \quad j \in J_f^k, \tag{43}$$

$$\beta_{g,j}^k = \max\{|g_j^k|,\ \gamma_g(s_j^k)^2\} \quad \text{for all} \quad j \in J_g^k, \tag{44}$$

where $\gamma_f \geq 0$ and $\gamma_g \geq 0$ are the distance measure parameters ($\gamma_f = 0$ if f is convex and $\gamma_g = 0$ if g is convex).

Note that the problem (CP) still is a nonsmooth (piecewise linear) optimization problem. However, due to special minmax-structure, it can be rewritten as

$$\begin{cases} \text{minimize} & v + \frac{u_k}{2}||d||^2 \\ \text{subject to} & -\beta_{f,j}^k + (\xi_j^f)^T d \leq v \quad \text{for all} \quad j \in J_f^k \\ \text{and} & -\beta_{g,j}^k + (\xi_j^g)^T d \leq v \quad \text{for all} \quad j \in J_g^k, \end{cases} \tag{BP}$$

which is a quadratic (smooth) problem. In convex case ($\gamma_f = \gamma_g = 0$) the problems (CP) and (BP) are equivalent. For computational reasons it is more effective to solve the dual problem of (BP), i.e. we find multipliers λ_j^k for $j \in J_f^k$ and μ_j^k for $j \in J_g^k$ that solve the problem

$$\begin{cases} \text{minimize} & \frac{1}{2u_k}||\sum_{j\in J_f^k} \lambda_j \xi_j^f + \sum_{j \in J_g^k} \mu_j \xi_j^g||^2 \\ & + \sum_{j \in J_f^k} \lambda_j \beta_{f,j}^k + \sum_{j\in J_g^k} \mu_j \beta_{g,j}^k \\ \text{subject to} & \sum_{j\in J_f^k} \lambda_j + \sum_{j \in J_g^k} \mu_j = 1 \\ \text{and} & \lambda_j, \mu_j, \geq 0. \end{cases} \tag{DP}$$

THEOREM 3. *Problems (BP) and (DP) are equivalent, and they have unique solutions* (d_k, v_k) *and* (λ_j^k, μ_j^k)*, respectively, such that*

$$d_k = -\frac{1}{u_k}\Big[\sum_{j\in J_f^k} \lambda_j^k \xi_j^f + \sum_{j\in J_g^k} \mu_j^k \xi_j^f\Big], \tag{45}$$

$$v_k = -u_k\|d_k\|^2 - \sum_{j\in J_f^k} \lambda_j^k \beta_j^k - \sum_{j\in J_g^k} \mu_j^k \beta_j^k. \tag{46}$$

For the proof of the above theorem we refer to [15].

4.2. Subgradient aggregation

We note that the larger the index set J_f^k and J_g^k are the more accurate the polyhedral approximation $\hat{H}^k$ is. Thus the simplest strategy is to choose $J_f^k = J_g^k = \{1, \ldots, k\}$. However, in practice this choise presents serious problems with storage and computation time after a large number of iterations.

Next we shall present the subgradient aggregation strategy (cf. [8]) for keeping the dimension of the problem (DP) bounded. Let λ_j^k for $j \in J_f^k$ and μ_j^k for $j \in J_g^k$ be the Lagrange multipliers of the problem (BP) at iteration k and denote $\lambda_f^k = \sum_{j\in J_f^k} \lambda_j^k$ and $\mu_g^k = \sum_{j\in J_g^k} \mu_j^k$. We define the scaled multipliers for all $j \in J_f^k$ and $j \in J_g^k$ by

$$\tilde{\lambda}_j^k = \begin{cases} \lambda_j^k/\lambda_f^k, & \text{if } \lambda_f^k > 0 \\ 1/|J_f^k|, & \text{if } \lambda_f^k = 0 \end{cases} \qquad \text{and} \qquad \tilde{\mu}_j^k = \begin{cases} \mu_j^k/\mu_g^k, & \text{if } \mu_g^k > 0 \\ 1/|J_g^k|, & \text{if } \mu_g^k = 0 \end{cases}$$

and the aggregate subgradients by

$$(p_f^k, \tilde{f}_p^k) = \sum_{j\in J_f^k} \tilde{\lambda}_j^k(\xi_j^f, f_j^k) \qquad \text{and} \qquad (p_g^k, \tilde{g}_p^k) = \sum_{j\in J_g^k} \tilde{\mu}_j^k(\xi_j^g, g_j^k).$$

Our aim is to add into the problem (BP) the aggregate constraints

$$-\tilde{\beta}_{f,p}^k + (p_f^k)^T d \le v \quad \text{and} \tag{47}$$

$$-\tilde{\beta}_{g,p}^k + (p_g^k)^T d \le v, \tag{48}$$

where $\tilde{\beta}_{f,p}^k = \max\{|f(x_k) - \tilde{f}_p^k|,\ \gamma_f(\tilde{s}_f^k)^2\}$ and $\tilde{\beta}_{g,p}^k = \max\{|\tilde{g}_p^k|,\ \gamma_g(\tilde{s}_g^k)^2\}$, and the aggregate distance measures are defined by $(s_f^1 = s_g^1 = 0)$

$$\tilde{s}_f^k = \textstyle\sum_{j\in J_f^k} \tilde{\lambda}_j^k s_j^k + \tilde{\lambda}_p^k s_f^k \qquad \tilde{s}_g^k = \displaystyle\sum_{j\in J_g^k} \tilde{\mu}_j^k s_j^k + \tilde{\mu}_p^k s_g^k \tag{49}$$

$$s_f^{k+1} = \tilde{s}_f^k + \|x_{k+1} - x_k\| \qquad s_g^{k+1} = \tilde{s}_g^k + \|x_{k+1} - x_k\|. \tag{50}$$

However there is one drawback: at the beginning of iteration k the vectors p_f^k and p_g^k are still unknown. This can be avoided by employing the information of the previous iteration as follows. Due to formulas (36) and (37) we define

$$f_p^{k+1} = \tilde{f}_p^k + (p_f^k)^T(x_{k+1} - x_k) \quad \text{and} \tag{51}$$

$$g_p^{k+1} = \tilde{g}_p^k + (p_g^k)^T(x_{k+1} - x_k). \tag{52}$$

At the first iteration let $x_1 \in K$ be a feasible starting point supplied by the user, then we initialize our algorithm by $y_1 = x_1$, $p_f^0 = \xi_1^f \in \partial f(y_1)$, $p_g^0 = \xi_1^g \in \partial g(y_1)$, $f_p^1 = f_1^1 = f(y_1)$, $g_p^1 = g_1^1 = g(y_1)$ and $J_f^1 = J_g^1 = \{1\}$. At iteration k we replace the unknown vectors $\tilde{f}_p^k$, $\tilde{g}_p^k$, p_f^k and p_g^k by the previously generated f_p^k, g_p^k, p_f^{k-1} and p_g^{k-1}, respectively, and define

$$\begin{aligned} \beta_{f,p}^k &= \max\{|f(x_k) - f_p^k|,\ \gamma_f(s_f^k)^2\} \qquad \text{and} \\ \beta_{g,p}^k &= \max\{|g_p^k|,\ \gamma_f(s_g^k)^2\}. \end{aligned} \tag{53}$$

This leads to the following aggregate modification of (BP)

$$\begin{cases} \text{minimize} & v + \frac{u_k}{2}\|d\|^2 \\ \text{subject to} & -\beta_{f,j}^k + (\xi_j^f)^T d \le v \quad \text{for all} \quad j \in J_f^k \\ & -\beta_{f,p}^k + (p_f^{k-1})^T d \le v \\ & -\beta_{g,j}^k + (\xi_j^g)^T d \le v \quad \text{for all} \quad j \in J_g^k \\ \text{and} & -\beta_{g,p}^k + (p_g^{k-1})^T d \le v \end{cases} \tag{ABP}$$

and via dualization we find multipliers λ_p^k, μ_p^k, λ_j^k for $j \in J_f^k$ and μ_j^k for $j \in J_g^k$ that solve the problem

$$\begin{cases} \text{minimize} & \frac{1}{2u_k}\|\sum_{j\in J_f^k}\lambda_j\xi_j^f + \lambda_p p_f^{k-1} + \sum_{j\in J_g^k}\mu_j\xi_j^g + \mu_p p_g^{k-1}\|^2 \\ & +\sum_{j\in J_f^k}\lambda_j\beta_{f,j}^k + \lambda_p\beta_{f,p}^k + \sum_{j\in J_g^k}\mu_j\beta_{g,j}^k + \mu_p\beta_{g,p}^k \\ \text{subject to} & \sum_{j\in J_f^k}\lambda_j + \lambda_p + \sum_{j\in J_g^k}\mu_j + \mu_p = 1 \\ \text{and} & \lambda_j, \lambda_p, \mu_j, \mu_p \ge 0. \end{cases} \tag{ADP}$$

Suppose now that $(\lambda_p^k, \mu_p^k, \lambda_j^k, \mu_j^k)$ is the solution of the problem (ADP). Then we can similarly denote $\lambda_f^k = \lambda_p^k + \sum_{j\in J_f^k}\lambda_j^k$ and $\mu_g^k = \mu_p^k + \sum_{j\in J_g^k}\mu_j^k$ and define the scaled multipliers for all $j \in J_f^k$ and $j \in J_f^k$ by

$$\tilde{\lambda}_j^k = \begin{cases} \lambda_j^k/\lambda_f^k, & \text{if } \lambda_f^k > 0 \\ 1/(|J_f^k|+1), & \text{if } \lambda_f^k = 0 \end{cases} \qquad \tilde{\lambda}_p^k = \begin{cases} \lambda_p^k/\lambda_f^k, & \text{if } \lambda_f^k > 0 \\ 1/(|J_f^k|+1), & \text{if } \lambda_f^k = 0 \end{cases}$$

and

$$\tilde{\mu}_j^k = \begin{cases} \mu_j^k/\mu_g^k, & \text{if } \mu_g^k > 0 \\ 1/(|J_g^k|+1), & \text{if } \mu_g^k = 0 \end{cases} \qquad \tilde{\mu}_p^k = \begin{cases} \mu_p^k/\mu_g^k, & \text{if } \mu_f^k > 0 \\ 1/(|J_g^k|+1), & \text{if } \mu_g^k = 0 \end{cases}$$

and the aggregate subgradients by

$$(p_f^k, \tilde{f}_p^k) = \sum_{j \in J_f^k} \tilde{\lambda}_j^k (\xi_j^f, f_j^k) + \tilde{\lambda}_p^k (p_f^{k-1}, f_p^k) \qquad \text{and} \tag{54}$$

$$(p_g^k, \tilde{g}_p^k) = \sum_{j \in J_g^k} \tilde{\mu}_j^k (\xi_j^g, g_j^k) + \tilde{\mu}_p^k (p_g^{k-1}, g_p^k). \tag{55}$$

Finally we denote by

$$p_k = \sum_{j \in J_f^k} \lambda_j^k \xi_j^f + \lambda_p^k p_f^{k-1} + \sum_{j \in J_g^k} \mu_j^k \xi_j^g + \mu_p^k p_g^{k-1} \qquad \text{and} \tag{56}$$

$$\tilde{\beta}_p^k = \lambda_f^k \tilde{\beta}_{f,p}^k + \mu_g^k \tilde{\beta}_{g,p}^k. \tag{57}$$

THEOREM 4. *Problems (ABP) and (ADP) are equivalent, and they have unique solutions* (d_k, v_k) *and* $(\lambda_p^k, \mu_p^k, \lambda_j^k, \mu_j^k)$, *respectively, such that*

$$d_k = -\frac{1}{u_k} p_k, \tag{58}$$

$$v_k = -u_k \|d_k\|^2 - \tilde{\beta}_p^k. \tag{59}$$

For the proof we refer to [15].

In theory this aggregation strategy allows us to choose the index sets J_f^k and J_g^k quite freely. In practice this choice still has a strong effect on the trade-off between efficiency and amount of work per iteration. To strike a balance we use a user-supplied bound $M_J \geq 2$ on the number of indices.

4.3. Line search

Due to the trust region idea the proximal bundle type methods in convex case do not require any uncertain and lot of function evaluation demanding line search operation. In nonconvex case we cannot avoid line search in order to guarantee the convergence.

We assume that $m_L \in (0, \frac{1}{2})$, $m_R \in (m_L, 1)$ and $\bar{t} \in (0, 1]$ are fixed line search parameters. First we shall search for the largest number $t_L^k \in [0, 1]$ such that

(a) $f(x_k + t_L^k d_k) \leq f(x_k) + m_L t_L^k v_k$,

(b) $g(x_k + t_L^k d_k) \leq 0$,

(c) $t_L^k \geq \bar{t}$.

If such a parameter exists we take a long serious step: $x_{k+1} = x_k + t_L^k d_k$ and $y_{k+1} = x_{k+1}$. Otherwise, if requirements (a) and (b) hold but $0 < t_L^k < \bar{t}$ then we take a short serious step: $x_{k+1} = x_k + t_L^k d_k$ and $y_{k+1} = x_k + t_R^k d_k$, and if $t_L^k = 0$ we take a null step: $x_{k+1} = x_k$ and $y_{k+1} = x_k + t_R^k d_k$, where $t_R^k > t_L^k$ is such that

(d) $-\beta_{f,k+1}^{k+1} + (\xi_{k+1}^f)^T d_k \geq m_R v_k$.

In long serious steps there occurs a significant decrease in the value of the objective function. Thus there is no need for detecting discontinuities in the gradient of f, and so we set $\xi_{k+1}^f \in \partial f(x_{k+1})$. In short serious steps and null steps there exists discontinuity in the gradient of f. Then the requirement (d) ensures that x_k and y_{k+1} lie on the opposite sides of this discontinuity and the new subgradient $\xi_{k+1}^f \in \partial f(y_{k+1})$ will force a remarkable modification of the next search direction finding problem. In what follows we are using the line search algorithm presented in in [15], which finds step sizes t_L^k and t_R^k such that requirements (a)–(d) hold.

4.4. Weight updating

The last but not least important open question is the choice of the weight u_k. The simplest strategy might be to keep it fixed $u_k \equiv u_{fix}$. This, however, leads to several difficulties. Due to Theorem 4. we observe the following:

(i) If u_{fix} is very large, we shall have small $|v_k|$ and $||d_k||$, almost all steps are serious and we have slow descent.

(ii) If u_{fix} is very small, we shall have large $|v_k|$ and $||d_k||$, and each serious step will be followed by many null steps.

Therefore, we keep u_k as a variable and update it when necessary. In this we use the safeguarded quadratic interpolation technique due to [9].

4.5. Proximal Bundle Algorithm

Step 0: (Initialization) Select a starting point $x_1 \in K$, a final accuracy tolerance $\varepsilon_s > 0$, the maximum number of stored subgradients $M_J \geq 2$, an initial weight $u_1 > 0$ and line search parameters $m_L \in (0, \frac{1}{2})$, $m_R \in (m_L, 1)$ and $\bar{t} \in (0, 1]$. Choose the distance measure parameters $\gamma_f > 0$ and $\gamma_g > 0$ ($\gamma_f = 0$ if f is convex; $\gamma_g = 0$ if g is convex). Set the iteration counter $k = 1$ and initialize the following variables: $y_1 = x_1$, $p_f^0 = \xi_1^f \in \partial f(y_1)$, $p_g^0 = \xi_1^g \in \partial g(y_1)$, $f_p^1 = f_1^1 = f(y_1)$, $g_p^1 = g_1^1 = g(y_1)$, $s_f^1 = s_g^1 = s_1^1 = 0$ and $J_f^1 = J_g^1 = \{1\}$.

Step 1: (Direction finding). Find multipliers λ_p^k, μ_p^k, λ_j^k for $j \in J_f^k$ and μ_j^k for $j \in J_g^k$ by solving the problem (ADP). Calculate multipliers λ_f^k, μ_g^k, $\tilde{\lambda}_p^k$, $\tilde{\mu}_p^k$, $\tilde{\lambda}_j^k$ and $\tilde{\mu}_j^k$ for $j \in J_f^k$ and $j \in J_g^k$ and set

$$(p_f^k, \tilde{f}_p^k, \tilde{s}_f^k) = \sum_{j \in J_f^k} \tilde{\lambda}_j^k (\xi_j^f, f_j^k, s_j^k) + \tilde{\lambda}_p^k (p_f^{k-1}, f_p^k, s_f^k), \tag{60}$$

$$\begin{aligned}
(p_g^k, \tilde{g}_p^k, \tilde{s}_g^k) &= \sum_{j \in J_g^k} \tilde{\mu}_j^k(\xi_j^g, g_j^k, s_j^k) + \tilde{\mu}_p^k(p_g^{k-1}, g_p^k, s_g^k), && (61)\\
p_k &= \lambda_f^k p_f^k + \mu_p^k p_g^k, && (62)\\
\tilde{\beta}_{f,p}^k &= \max\{|f(x_k) - \tilde{f}_p^k|,\ \gamma_f(\tilde{s}_f^k)^2\}, && (63)\\
\tilde{\beta}_{g,p}^k &= \max\{|\tilde{g}_p^k|,\ \gamma_g(\tilde{s}_g^k)^2\}, && (64)\\
\tilde{\beta}_p^k &= \lambda_f^k \tilde{\beta}_{f,p}^k + \mu_g^k \tilde{\beta}_{g,p}^k. && (65)
\end{aligned}$$

Set $d_k = -\frac{1}{u_k} p_k$.

Step 2: (Stopping criterion). Set

$$w_k = \frac{1}{2}\|p_k\|^2 + \tilde{\beta}_p^k.$$

If $w_k \le \varepsilon_s$ then STOP.

Step 3: (Line search). Find step sizes $t_L^k \in [0,1]$ and $t_R^k \in [t_L^k, 1]$ by the line search algorithm of [15]. Set $x_{k+1} = x_k + t_L^k d_k$ and $y_{k+1} = x_k + t_R^k d_k$.

Step 4: (Linearization updating). Calculate the linearization values

$$\begin{aligned}
f_j^{k+1} &= f_j^k + t_L^k(\xi_j^f)^T d_k, \quad \text{for} \quad j \in J_f^k, && (66)\\
g_j^{k+1} &= g_j^k + t_L^k(\xi_j^g)^T d_k, \quad \text{for} \quad j \in J_g^k, && (67)\\
s_j^{k+1} &= s_j^k + t_L^k\|d_k\|, \quad \text{for} \quad j \in J_f^k \cup J_g^k, && (68)\\
f_p^{k+1} &= \tilde{f}_p^k + t_L^k(p_f^k)^T d_k, && (69)\\
g_p^{k+1} &= \tilde{g}_p^k + t_L^k(p_g^k)^T d_k, && (70)\\
s_f^{k+1} &= \tilde{s}_f^k + t_L^k\|d_k\|, && (71)\\
s_g^{k+1} &= \tilde{s}_g^k + t_L^k\|d_k\|. && (72)
\end{aligned}$$

Evaluate $\xi_{k+1}^f \in \partial f(y_{k+1})$ and $\xi_{k+1}^g \in \partial g(y_{k+1})$ and set

$$\begin{aligned}
f_{k+1}^{k+1} &= f(y_{k+1}) + (t_L^k - t_R^k)(\xi_{k+1}^f)^T d_k, && (73)\\
g_{k+1}^{k+1} &= g(y_{k+1}) + (t_L^k - t_R^k)(\xi_{k+1}^g)^T d_k, && (74)\\
s_{k+1}^{k+1} &= (t_R^k - t_L^k)\|d_k\|. && (75)
\end{aligned}$$

Step 5: (Weight updating). Select u_{k+1} by the weight updating algorithm of [9].

Step 6: (Updating). Set $J_f^{k+1} = J_f^k \cup \{k+1\}$ and $J_g^{k+1} = J_g^k \cup \{k+1\}$. If $|J_f^{k+1}| > M_J$, then $J_f^{k+1} = J_f^{k+1} \setminus \{\min\ j \mid j \in J_f^{k+1}\}$. If $|J_g^{k+1}| > M_J$, then $J_g^{k+1} = J_g^{k+1} \setminus \{\min\ j \mid j \in J_g^{k+1}\}$. Increase k by 1 and go to Step 1.

If we assume that f is weakly semismooth, i.e. the directional derivative $f'(x;d)$ exists for all x and d, and $f'(x;d) = \lim_{t\downarrow 0}(\xi^f)^T d$, where $\xi^f \in \partial f(x+td)$, then the convergence of the above algorithm can proved in the same way as in [8].

5. Numerical results

In this section we shall present one example of hemivariational inequalities in nonsmooth mechanics of solids, namely a nonmonotone contact problem. Its approximation and numerical realization will be based on the methods presented in the previous sections.

Let us study a two–dimensional undeformed elastic body represented by a bounded domain $\Omega \subset {}^2$ and let Γ be the boundary of Ω. We shall assume that the body is subjected to internal and external forces which cause the deformation of the body. We denote by $n = (n_i)_{i=1}^2$ the outward unit normal vector to Γ, $\sigma = (\sigma_{ij})_{i,j=1}^2$ the stress tensor, $\varepsilon = (\varepsilon_{ij})_{i,j=1}^2$ the strain tensor, $S = (S_i = \sigma_{ij}n_j)_{i=1}^2$ the boundary force and $u = (u_i)_{i=1}^2$ the displacement. The boundary Γ is divided into three nonoverlapping open sets Γ_1, Γ_2 and Γ_3 such that $\Gamma = \overline{\Gamma}_1 \cup \overline{\Gamma}_2 \cup \overline{\Gamma}_3$. Let us suppose that the domain Ω is as in Fig. 1, i.e.

$$\Omega = \{(x_1, x_2) \in {}^2 : x_1 \in (a,b) \text{ and } x_2 \in (\alpha(x_1), \gamma)\}, \tag{76}$$

where a, b, γ are positive constants and $\alpha \in C^{1,1}([a,b])$, $\alpha \geq 0$. We assume that on Γ_1 the displacements are given by

$$u(x) = U(x) \quad \text{on } \Gamma_1. \tag{77}$$

For simplicity, we consider the homogenous boundary value $U = 0$ on Γ_1. On the other hand on Γ_2 the boundary forces are given by

$$S(x) = g^1(x) \quad \text{on } \Gamma_2. \tag{78}$$

Furthermore, we assume that Γ_3, which is a graph of the function α, i.e. $\Gamma_3 = \{(x_1, x_2) \in {}^2 : x_1 \in (a,b) \text{ and } x_2 = \alpha(x_1)\}$, can be in contact with a

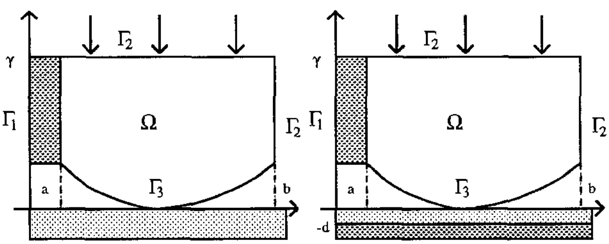

Fig. 1. Contact problems of a linear elastic body

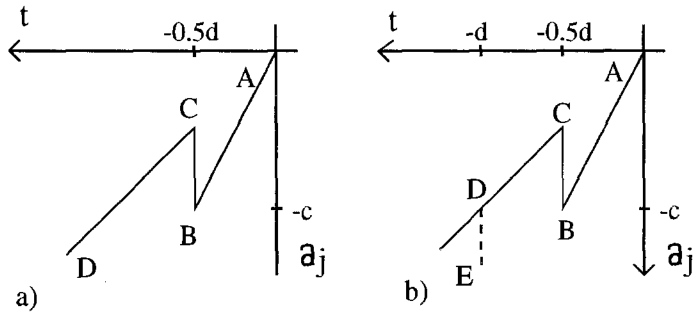

Fig. 2. Nonmonotone contact laws

nonmonotone frictionless foundation. Thus the body obeys on Γ_3 the following system of boundary conditions:

$$\begin{aligned} -S_2(x_1) &\in \partial j(u_2(x_1,\alpha(x_1))+\alpha(x_1)) && \text{a.e. } x\in(a,b) \\ S_1(x_1,\alpha(x_1)) &= 0 && \text{a.e. } x\in(a,b), \end{aligned} \tag{79}$$

where j is a locally Lipschitz continuous function from to satisfying (3) and (4) (see in Fig. 2a the generalized gradient of j). We shall also consider a contact problem with a rigid foundation and a nonmonotone layer of thickness d above it (see Fig. 1). Now the conditions (79) are replaced by (see Fig. 2b)

$$\begin{aligned} -S_2(x_1) &\in \partial j(u_2(x_1,\alpha(x_1))+\alpha(x_1)) && \text{a.e. } x\in(a,b) \\ u_2(x_1,\alpha(x_1)) &\geq -\alpha(x_1)-d && \text{a.e. } x\in(a,b) \\ S_1(x_1,\alpha(x_1)) &= 0 && \text{a.e. } x\in(a,b). \end{aligned} \tag{80}$$

We assume that the linearized strain tensor ε obeys the Hooke's law of the form

$$\sigma_{ij}=c_{ijkl}\varepsilon_{kl}(u), \quad \text{where } \varepsilon_{kl}(u)=\frac{1}{2}\left(\frac{\partial u_k}{\partial x_l}+\frac{\partial u_l}{\partial x_k}\right) \tag{81}$$

and the elasticity coefficients c_{ijkl} satisfy the usual symmetry and elasticity conditions in Ω. Then the equilibrium state of Ω is described by means of the following system:

$$\sigma_{ij,j}+g_i^2=0 \quad \text{in } \Omega, \quad i=1,2, \tag{82}$$

where g^2 is a body force. We shall assume that $g^2=0$. The weak formulation of the contact problem in the case of (79) reads as follows:

$$\left\{\begin{array}{l} \text{find } u\in V \text{ and } \mathcal{X}(u)\in L^2(\Gamma_3) \text{ such that} \\ a(u,v)+\int_{\Gamma_3}\mathcal{X}(x)v_2(x)\,ds=(g^1,v)_{0,\Gamma_2} \quad \forall v\in V \\ \text{and } \mathcal{X}(x)\in\partial j(u_2(x_1,\alpha(x_1))+\alpha(x_1)) \quad \text{a.e. } x_1\in(a,b), \end{array}\right. \tag{P}$$

where V is the space of admissable displacements defined by

$$V = \{v = (v_1, v_2) \in (H^1(\Omega))^2 : v_i = 0 \text{ on } \Gamma_1,\ i = 1,2\}, \tag{83}$$

a is the bilinear form from $V \times V$ to defined by $a(u,v) = \int_\Omega c_{ijkl}\varepsilon_{kl}(u)\, \varepsilon_{ij}(v)\, dx$ for all $u, v \in V$ and $(\cdot,\cdot)_{0,\Gamma_2}$ is the $(L^2(\Gamma_2))^2$-norm. In the case of the rigid foundation with a nonmonotone layer, i.e. (80) the weak form is the following:

$$\begin{cases} \text{find } u \in K \text{ and } \mathcal{X}(u) \in L^2(\Gamma_3) \text{ such that} \\ a(u, v-u) + \int_{\Gamma_3} \mathcal{X}(x)(v_2(x) - u_2(x))\, ds \geq (g^1, v-u)_{0,\Gamma_2}\ \forall v \in K \\ \text{and } \mathcal{X}(x) \in \partial j(u_2(x_1, \alpha(x_1)) + \alpha(x_1)) \quad \text{a.e. } x_1 \in (a,b), \end{cases} \tag{CP}$$

where K is a nonempty, closed, convex constraint set defined by

$$K = \{v \in V : v_2(x_1, \alpha(x_1)) \geq -\alpha(x_1) - d \text{ a.e. in } (a,b)\}. \tag{84}$$

Let us consider the approximation of the problems (P) and (CP) (since h is fixed we shall skip it in many places). First we define Ω_h a polygonal approximation of Ω. Let $a \equiv a_0 < a_1 < \ldots < a_m \equiv b$ be a partition of $[a,b]$ and α_h be a piecewise linear function such that $\alpha_h(a_i) = \alpha(a_i)$ for all $i = 0, 1, \ldots, m$. Then we set that

$$\Omega_h = \{(x_1, x_2) \in {}^2 : x_1 \in (a,b) \text{ and } x_2 \in (\alpha_h(x_1), \gamma)\}; \tag{85}$$

$$\Gamma_{3h} = \{(x_1, x_2) \in {}^2 : x_1 \in (a,b) \text{ and } x_2 = \alpha_h(x_1)\}. \tag{86}$$

For V_h we choose the space of piecewise linear functions over a regular triangulation $\mathcal{T}_h$ of $\overline{\Omega}_h$ such that the whole segment $\{(x_1, x_2) : x_1 \in [a_{i-1}, a_i]$ and $x_2 = \alpha_h(x_1)\}$ is the whole side of some triangle $T \in \mathcal{T}_h$

$$V_h = \{v \in (C(\overline{\Omega}_h))^2 : v|_T \in (P_1(T))^2\ \forall T \in \mathcal{T}_h,\ v = 0 \text{ on } \Gamma_1\}. \tag{87}$$

Let $\{\varphi^j\}_{j=1}^n$ be the Courant basis functions of V_h. The approximation of the constraint set K is defined by

$$K_h = \{v \in V_h : v_2(a_i, \alpha_h(a_i)) \geq -\alpha_h(a_i) - d \text{ for all } i = 1, \ldots, m\}. \tag{88}$$

For approximating the integral $\int_{\Gamma_{3h}} \mathcal{X}(x) v_2(x)\, ds$ we use the following numerical integration formula

$$\int_{\Gamma_{3h}} f(x)\, ds \approx \sum_{i=0}^{m-1} |(a_i, \alpha_h(a_i))(a_{i+1}, \alpha_h(a_{i+1}))| \\ \frac{1}{2}(f(a_i, \alpha_h(a_i)) + f(a_{i+1}, \alpha_h(a_{i+1}))), \tag{89}$$

where $|(a_i, \alpha_h(a_i))(a_{i+1}, \alpha_h(a_{i+1}))|$ is the length of the line segment from $(a_i, \alpha_h(a_i))$ to $(a_{i+1}, \alpha_h(a_{i+1}))$. It is easy to see that this formula is related to the well-known *trapezoidal* rule. Thus the nodal points and the weights of (10) are

$$x^i = (a_i, \alpha_h(a_i)) \quad i = 1, \dots, m; \tag{90}$$

$$\begin{aligned} c_i &= \frac{1}{2}(|(a_{i-1}, \alpha_h(a_{i-1}))(a_i, \alpha_h(a_i))| \\ &\quad + |(a_i, \alpha_h(a_i))(a_{i+1}, \alpha_h(a_{i+1}))|) \quad i = 1, \dots, m-1; \\ c_m &= \frac{1}{2}|(a_{m-1}, \alpha_h(a_{m-1}))(a_m, \alpha_h(a_m))|. \end{aligned} \tag{91}$$

In order to write the discrete problems in the matrix form we need to define an $m \times n$ matrix $\mathcal{P}$ corresponding to the linear mapping P_h introduced in Section 2. Indeed: due to the choice of the formula (89) it reads as follows:

$$(\mathcal{P}\mathbf{v})_i = \left\{ \begin{array}{l} v_j, \text{ if } v_j \text{ corresponds to the vertical displacement} \\ \quad \text{of the nodal point } (a_i, \alpha_h(a_i)) \in \overline{\Gamma}_{3h} \end{array} \right. \tag{92}$$

Then, the approximation problem $(\mathrm{P})_h$ is defined by

$$\left\{ \begin{array}{l} \text{find } \mathbf{u} \in {}^n \text{ and } \mathbf{s} \in {}^m \text{ such that} \\ \mathbf{v}^T \mathbf{A} \mathbf{u} + (\mathcal{P}\mathbf{v})^T \mathbf{s} = \mathbf{v}^T \mathbf{g} \quad \forall \mathbf{v} \in {}^n \\ \text{and } s_i \in c_i \partial j((\mathcal{P}\mathbf{u})_i + \alpha_h(a_i)) \quad i = 1, \dots, m, \end{array} \right. \tag{P$)_h$}$$

and, consequently, $(\mathrm{CP})_h$

$$\left\{ \begin{array}{l} \text{find } \mathbf{u} \in \mathbf{K} \text{ and } \mathbf{s} \in {}^m \text{ such that} \\ (\mathbf{v} - \mathbf{u})^T \mathbf{A} \mathbf{u} + (\mathcal{P}\mathbf{v} - \mathcal{P}\mathbf{u})^T \mathbf{s} \geq (\mathbf{v} - \mathbf{u})^T \mathbf{g} \quad \forall \mathbf{v} \in \mathbf{K} \\ \text{and } s_i \in c_i \partial j((\mathcal{P}\mathbf{u})_i + \alpha_h(a_i)) \quad i = 1, \dots, m, \end{array} \right. \tag{CP$)_h$}$$

where the $n \times n$-matrix $\mathbf{A}$, the ^{n}–vector $\mathbf{g}$ and the constraint set $\mathbf{K}$ is defined as in Section 3.

In the numerical realization we transform first $(\mathrm{P})_h$ and $(\mathrm{CP})_h$ to the problems of minimizing the corresponding potential function, i.e.

$$\arg\min_{\mathbf{v} \in {}^n} \mathbf{L}(\mathbf{v}) \quad \text{or} \quad \arg\min_{\mathbf{v} \in \mathbf{K}} \mathbf{L}(\mathbf{v}), \tag{93}$$

where $\mathbf{L} : {}^n \to$ is defined by

$$\mathbf{L}(\mathbf{v}) = \frac{1}{2}\mathbf{v}^T \mathbf{A} \mathbf{v} - \mathbf{v}^T \mathbf{g} + \mathbf{J}(\mathbf{v}), \quad \mathbf{J}(\mathbf{v}) = \sum_{i=1}^{m} c_i j((\mathcal{P}\mathbf{v})_i). \tag{94}$$

Then we make use of the fact that the nonlinear behaviour of the contact problem is prescribed only in the nodes of $\overline{\Gamma}_{3h}$: We list the components of $\mathbf{u}$ representing the nodal displacement on $\overline{\Gamma}_{3h}$ first. Then we can decompose $\mathbf{u}$ as follows: $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2)$. We eliminate $\mathbf{u}_2$ from $(\mathrm{P})_h$,$(\mathrm{CP})_h$ and, consequently, from $\mathbf{L}$. Thus the size of our problem, i.e. the number of the unknowns is

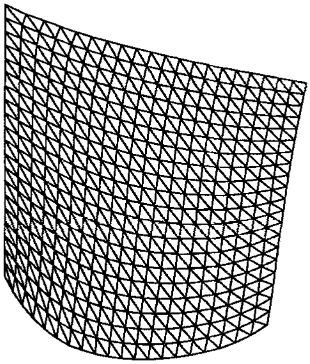

Fig. 3. Solution 1

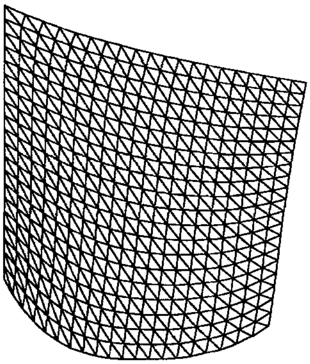

Fig. 4. Solution 2

reduced from $2m(m+1)$ to $2m$ (the number of the subintervals in x– and y–directions is m).

In the numerical calculations we used the plane stress model and the following values: $a = 0$, $b = \gamma = 1$, $\alpha(x_1) = 0.25 + x_1(x_1 - 1)$, $d = 0.04$, $c = 5$ and $g^1 \equiv -0.15$. We solved the both problems $(\mathrm{P})_h$ and $(\mathrm{CP})_h$ using $m = 10, 20, 40, 60$ (see Fig. 3 and Fig. 4 for the solutions with $m = 20$). Fig. 5 show the displacements of the nodal points on Γ_{3h} of the solution of $(\mathrm{P})_h$ (solution 1) and the solution of $(\mathrm{CP})_h$ (solution 2). We can see that some of those nodal points of the both solutions are on the branches AB and CD of Fig. 2 a),b) implying that the nonmonotone contact law really affects. The flat part of the solution 2 in Fig. 5 points out that some of the nodes on Γ_{3h} are contact with the rigid foundation (the branch DE of Fig. 2b) affects). Therefore the displacements of the solution 2 were smaller. Tables I and II show $(L^2(\Gamma_{3h}))^2$–error, the needed CPU-time in seconds, the iterations number and the value of the minimized function $\mathbf{L}$. The results showed very good convergence in $(L^2(\Gamma_{3h}))^2$–norm (as an exact solution is used the

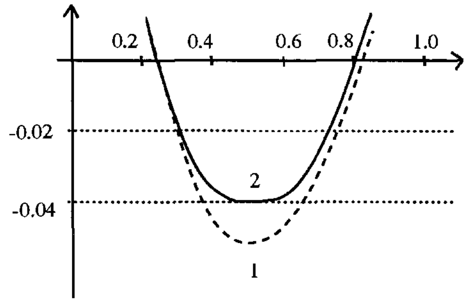

Fig. 5. Displacements of solutions 1 and 2 on Γ_{3h}

TABLE I. Solution 1

m	$(L^2(\Gamma_{3h}))^2$–error	CPU/sec	iterations	value of **L**
10	7.7e-3	2.7	44	-4.71e-3
20	4.1e-3	6.8	61	-4.84e-3
40	6.5e-4	46	114	-4.93e-3
60	*	480	146	-4.96e-3

solution obtained with $m = 60$). The main drawback was that the CPU-time increased quite rapidly when the number of the unknowns incresed, although the iteration numbers behaved quite reasonable. This suggests that the implementation of the nonsmooth optimizer which takes into account the special structure of the considered problem may improve the efficiency of our numerical approach. If we compare the results without elimination (as $m = 20$ the iteration number was 600 and CPU-time 380 sec) we see that our elimination stragegy was justified. Finally we would like to remind that these numerical results have pointed out that the used nonsmooth, nonconvex optimizer is very effective and reliable when the number of the unknowns is hundreds (the total number without elimination was thousands) and it can also handle problems with few thousands unknowns.

References

1. Aubin J.–P. and Clarke F.H. (1979), Shadow Prices and and Duality for a Class of Optimal Control Problems, *SIAM J. Control Optim.*, 17, pp. 567–586
2. Ciarlet, P.G. (1978), *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam
3. Clarke F.H. (1983), *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York
4. Glowinski R., Lions J.L. and Trémolièrs R. (1981), *Numerical Analysis of Variational Inequalities*, North–Holland, Amsterdam

TABLE II. Solution 2

m	$(L^2(\Gamma_{3h}))^2$–error	CPU/sec	iterations	value of $\mathbf{L}$
10	1.0e-2	2.6	45	-4.77e-3
20	5.2e-3	6.3	62	-4.91e-3
40	1.5e-4	45	121	-5.00e-3
60	*	400	121	-5.03e-3

5. Glowinski R. (1984), *Numerical Methods for Nonlinear Variational Problems*, Springer–Verlag, New York
6. Hlaváček I, Haslinger J., Nečas J. and Lovišek J. (1988), *Solutions of Variational Inequalities in Mechanics*, Springer–Verlag, New York
7. Kelley, J. E. (1960), The Cutting Plane Method for Solving Convex Programs, *Journal for SIAM*, **8**, pp. 703–712
8. Kiwiel, K. C. (1985), *Methods of Descent for Nondifferentiable Optimization*, Lecture Notes in Mathematics 1133, Springer–Verlag, Berlin–Heidelberg
9. Kiwiel, K. C. (1990), Proximity Control in Bundle Methods for Convex Nondifferentiable Optimization *Mathematical Programming*, **46**, pp. 105–122
10. Miettinen M. and Haslinger J. (1992), Approximation of Optimal Control Problems of Hemivariational Inequalities, *Numer. Funct. Anal. and Optimiz.*, **13** (1&2), pp.43–68
11. Miettinen M. and Haslinger J. (1993), Approximation of Nonmonotone Multivalued Differential Inclusions, University of Jyvskyl, Department of Mathematics, Laboratory of Scientific Computing, Report 4/93, to appear in IMA J. Num. Anal.
12. Miettinen M. and Haslinger J. (1994), Finite Element Approximation of Vector–Valued Hemivariational Problems, University of Jyvskyl, Department of Mathematics, Laboratory of Scientific Computing, Report 94, submitted
13. Miettinen M. (1994), On Constrained Hemivariational Inequalities and Their Approximation, to appear in Applicable Analysis
14. Miettinen M. (1993), *Approximation of Hemivariational Inequalities and Optimal Control Problems*, PhD Thesis, University of Jyvskyl, Department of Mathematics, Report 59
15. Mäkelä M.M. and Neittanmäki P. (1992), *Nonsmooth Optimization. Analysis and Algorithms with Applications to Optimal Control*, World Scientific, Singapore
16. M. M. Mäkelä (1993), Issues of Implementing a Fortran Subroutine Package NSOLIB for Nonsmooth Optimization, University of Jyvskyl, Department of Mathematics, Laboratory of Scientific Computing, Report 5/93
17. M. M. Mäkelä and T. Männikkö (1994), Numerical Solution of Nonsmooth Optimal Control Problems with an Application to the Continuous Casting Process, *Advances in Mathematical Sciences and Applications*, 4, pp. 491–515
18. Naniewicz Z. and Panagiotopoulos P.D. (1994), *Mathematical Theory of Hemivariational Inequalities and Applications*, Marcel Dekker, New York
19. Naniewicz Z. (1994), Hemivariational Inequalities with Functions Fulfilling Directional Growth Condition, to appear in Applicable Analysis
20. Panagiotopoulos P.D. (1985), *Inequality Problems in Mechanics and Applications. Convex and Nonconvex Energy Functions.*, Birkhäuser, Basel
21. Panagiotopoulos P.D. (1991), Coercive and Semicoercive Hemivariational Inequalities, *Nonlin. Anal.*,**16**, pp. 209–231
22. Panagiotopoulos P.D. (1993), *Hemivariational Inequalities. Applications in Mechanics and Engineering*, Springer–Verlag, New York

23. Schramm, H. and Zowe, J. (1992), A Version of the Bundle Idea for Minimizing a Nonsmooth Function: Conceptual Idea, Convergence Analysis, Numerical Results, *SIAM Journal on Optimization*, **2**, pp. 121–152