Control and Cybernetics

vol. **31** (2002) No. 3

Proximal-based regularization methods and successive approximation of variational inequalities in Hilbert spaces

by

Alexander Kaplan and Rainer Tichatschke

Dept. of Mathematics, University of Trier, D-54286 Trier

Abstract: For variational inequalities with multi-valued, maximal monotone operators in Hilbert spaces we study proximal-based methods with an improvement of the data approximation after each (approximately performed) proximal iteration. The standard conditions on a distance functional of Bregman's type are weakened, depending on a "reserve of monotonicity" of the operator in the variational inequality, and the enlargement concept is used for approximating the operator. Weak convergence of the proximal iterates to a solution of the original problem is proved. The construction of the ϵ-enlargement of monotone operators is analyzed for some particular cases.

Keywords: variational inequalities, monotone operators, proximal point methods, regularization, Bregman function

1. Introduction

Let $(X, \|\cdot\|)$ be a Hilbert space with the topological dual X' and the duality pairing $\langle \cdot, \cdot \rangle$ between X and X'. The variational inequality

$$(P) \quad \text{find } x^* \in K \text{ such that}$$
$$\exists\, q \in \mathcal{Q}(x^*) : \ \langle q, x - x^* \rangle \geq 0 \ \ \forall x \in K,$$

is considered, where $K \subset X$ is a convex closed set and $\mathcal{Q} : X \to 2^{X'}$ is a maximal monotone operator. We generally suppose that (P) is solvable and denote by X^* its solution set.

The proximal point method (PPM), originally introduced by Martinet (1970) to solve convex variational problems and later on investigated in a more general setting by Rockafellar (1976), has initiated a lot of new algorithms for solving various classes of variational inequalities and related problems.

The exact proximal point method, applied to the variational inequality (P), *Given $x^1 \in K$ and a sequence $\{\chi_k\}$, $0 < \chi_k \leq \bar{\chi} < \infty$; with $x^k \in K$ from the previous step, define $x^{k+1} \in K$ such that*

$$\exists\ q(x^{k+1}) \in \mathcal{Q}(x^{k+1}):$$
$$\langle q(x^{k+1}) + \chi_k \nabla_1 D(x^{k+1}, x^k), x - x^{k+1}\rangle \geq 0 \quad \forall x \in K,$$

where $D(x,y) = \frac{1}{2}\|x-y\|^2$ and ∇_1 is the partial gradient w.r.t. x.

For different modifications of the PPM, also with other quadratic functionals D, we point out Kaplan and Tichatschke (1994, 2000b) and Kiwiel (1999), where numerous references can be found.

In the last decade, a new direction in the PPM's has been intensively developed, which is based on the use of *non*-quadratic "distance functionals" D. The main motivation for such proximal methods is the following:

- A non-quadratic proximal term permits us, for certain classes of problems, to preserve the main merits of the classical PPM (good stability of the auxiliary problems and convergence of the whole sequence of iterates to a solution of the original problem) and, at the same time, to guarantee that the iterates stay in the interior of the set K;
- the application of non-quadratic proximal techniques (as in Auslender, Teboulle and Ben-Tiba, 1999, Teboulle, 1992, Tseng and Bertsekas, 1993) to the dual of a smooth convex program leads to multiplier methods with twice or higher differentiable augmented Lagrangian functionals. Moreover, in Auslender, Teboulle and Ben-Tiba (1999) the Hessians of these functionals are bounded.

More motivation for the study of non-quadratic proximal methods can be found in Auslender and Haddou (1995), Burachik, Iusem and Svaiter (1997), Eckstein (1993), Polyak and Teboulle (1997). For infinite-dimensional convex optimization problems, non-quadratic proximal methods have been studied in Alber, Burachik and Iusem (1997), Butnariu and Iusem (1997), (2000), and for variational inequalities in Hilbert spaces – see Burachik and Iusem (1998).

In the present paper, we develop a uniform approach to the construction and convergence analysis of proximal like methods for solving variational inequalities in Hilbert spaces. The following *generalized proximal point method* (GPPM) is considered. Taking a linear monotone operator $\mathcal{B}: X \to X'$ such that the operator $(\mathcal{Q} - \mathcal{B})$ is still monotone, we choose a convex continuous functional $h: \bar{S} \to \mathbb{R}$ with $S \subset X$ so that

$$x \to \frac{1}{2}\langle \mathcal{B}x, x\rangle + h(x)$$

possesses properties like usually required for a Bregman function (with a zone S). For an approximation of the operator $\mathcal{Q}$, a family of operators $\{\mathcal{Q}^k\}$ with $\mathcal{Q} \subset \mathcal{Q}^k \subset \mathcal{Q}_{\epsilon_k}$ is used, where $\mathcal{Q}_\epsilon$ denotes the ϵ-enlargement of $\mathcal{Q}$ (Burachik, Iusem and Svaiter, 1997) and $\epsilon_k \to +0$. Precise conditions on the choice of the

At the k-th step of the GPPM, with $x^k \in K^{k-1} \cap S$ obtained at the previous step, the iterate $x^{k+1} \in K^k \cap \bar{S}$ is defined such that

$$\begin{aligned} &\exists\, q^k(x^{k+1}) \in \mathcal{Q}^k(x^{k+1}) : \\ &\langle q^k(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1}\rangle \\ &\geq -\delta_k \sqrt{\Gamma_1(x, x^{k+1})} \;\; \forall x \in K^k \cap \bar{S}. \end{aligned}$$

Here, $\{K^k\}$ is a sequence of convex, closed sets approximating K, the regularization parameter χ_k is as above, $\{\delta_k\}$ is a given non-negative sequence and

$$\Gamma_1(x, y) = \min\{\alpha\|x - y\|^2,\; \Gamma(x, y) + 1\}, \quad \alpha > 0\text{-const.},$$

with

$$\Gamma(x, y) = \frac{1}{2}\langle \mathcal{B}(x - y), x - y\rangle + h(x) - h(y) - \langle \nabla h(y), x - y\rangle$$

considered in $dom\Gamma = \bar{S} \times D(\nabla h)$ and used as a Lyapunov function. Conditions given below provide that $x^{k+1} \in K^k \cap S \cap D(\mathcal{Q}^k)$ (see Theorem 1).

In this paper, in comparison with the preceding publications dealing with non-quadratic regularization methods,

- the standard requirement of the strict monotonicity of the operator $\nabla_1 D(\cdot, y)$ (usually formulated as the strict convexity of Bregman's or an other function generating D) is weakened. This leads to an analogy of methods with weak regularization and regularization on a subspace (developed on the basis of the classical PPM in Kaplan and Tichatschke (1997, 2000b);
- a successive approximation of the set K is included.

With regard to the mentioned investigations for infinite-dimensional problems, here the class of operators $\mathcal{Q}$ is also extended (see the case D1 in Lemma 3 and Theorem 2, and Remark 7), an approximation of the operator $\mathcal{Q}$ is included and the auxiliary problems are supposed to be solved approximately.

The non-quadratic proximal method developed in Kaplan and Tichatschke (2002) is a partial variant of the GPPM with $\mathcal{Q}^k \equiv \mathcal{Q}$. In the present paper the operators $\mathcal{Q}^k$ are constructed as follows:

$$\mathcal{Q}^k = (\mathcal{Q} - \widehat{\mathcal{Q}})_{\epsilon_k} + \widehat{\mathcal{Q}}, \tag{1}$$

where $\widehat{\mathcal{Q}}$ is a continuous operator such that both $\mathcal{Q} - \widehat{\mathcal{Q}}$ and $\widehat{\mathcal{Q}} - \mathcal{B}$ are monotone. For a series of variational inequalities in mathematical physics, under an appropriate choice of $\widehat{\mathcal{Q}}$ the handling of $\mathcal{Q}^k$ is much simpler[1] than that of $\mathcal{Q}_{\epsilon_k}$.

[1]The corresponding theoretical results for an operator $\mathcal{Q}$ decomposed into the sum of a continuous monotone operator and the subdifferential of a convex positive homogeneous functional, as well as some examples on the calculation of the ϵ-enlargement of operators in

Moreover, the operator $\mathcal{Q}^k$ in (1) inherits all good continuity properties of the ϵ-enlargement and possesses also the property

$$\begin{aligned} &\langle q^k(x) - q(y), x - y\rangle \geq \langle \mathcal{B}(x-y), x-y\rangle - \epsilon_k \\ &\forall x \in D(\mathcal{Q}^k),\ y \in D(\mathcal{Q}),\ q^k(x) \in \mathcal{Q}^k(x),\ q(y) \in \mathcal{Q}(y). \end{aligned} \tag{2}$$

Just this fact permits to weaken the requirement of strict convexity of the regularizing functional h using "the reserve of monotonicity" $\mathcal{B}$ of the operator $\mathcal{Q}$. The analysis of quadratic proximal methods in Kaplan and Tichatschke (1994) and numerical experiments for control problems governed by PDE's (Rotin, 1999) and for Bingham problems (Schmitt, 1996) show that a significant acceleration of the numerical process can be expected on this way.

It should be noted that, for the operator $\mathcal{Q}_{\epsilon_k}$ in place of $\mathcal{Q}^k$, relation (2) is not valid, in general, if $\mathcal{B} \neq 0$.

Simultaneously, in this framework the class of operators $\mathcal{Q}$ is extended: Now, instead of the pseudo-monotonicity of $\mathcal{Q}$, the fulfillment of a weaker condition which we called Ψ-property (see after Lemma 2 in Section 4) is assumed.

The main contents of the paper are arranged as follows: In Section 2 conditions w.r.t. the successive approximation of Problem (P) and the regularizing functional are formulated, and the GPPM is specified. In Section 3 the solvability of the auxiliary problems is studied, and in Section 4 convergence of the GPPM is proved. As already mentioned, the Appendix contains results, simplifying the approximations of operators $\mathcal{Q}$ possessing a special structure.

2. Generalized proximal point method

In the sequel, we make use of the following notations: $S \subset X$ is an open convex set, its closure is denoted by $\bar{S}$; $\{K^k\} \subset X$, $K^k \supset K$, is a family of convex closed sets approximating K;

$$\mathcal{N}_K : y \to \begin{cases} \{z \in X' : \langle z, y - x\rangle \geq 0\ \forall x \in K\} & \text{if } y \in K \\ \emptyset & \text{otherwise} \end{cases}$$

is the normality operator for K; symbol $\rightharpoonup$ indicates weak convergence in X. With $\mathcal{B}$ and h as introduced in Section 1 (their properties will be specified below), we define the functional

$$\eta(x) = \begin{cases} \frac{1}{2}\langle \mathcal{B}x, x\rangle + h(x) & \text{if } x \in \bar{S} \\ +\infty & \text{otherwise.} \end{cases}$$

Now, the basic assumptions will be described.

ASSUMPTION 1 (on the successive approximation of Problem (P) and the choice of the controlling parameters):

(A1) For each k, the operator $\mathcal{Q} + \mathcal{N}_{K^k}$ is maximal monotone;

(A3) *(i) the operator* $(\mathcal{Q} - \mathcal{B})$ *is monotone,*
or, in case $\mathcal{Q}$ *is the subdifferential* ∂f *of a proper convex and lower semi-continuous (lsc) functional* f,

(ii) the functional $x \to f(x) - \langle \mathcal{B}x, x\rangle$ *is convex,*

where $\mathcal{B} : X \to X'$ *is a given linear, continuous and monotone operator with the symmetry property* $\langle \mathcal{B}x, y\rangle = \langle \mathcal{B}y, x\rangle$;

(A4) *any weak limit point of an arbitrary sequence* $\{v^k\}$, $v^k \in S \cap D(\mathcal{Q}^k) \cap K^k$, *belongs to* $K \cap D(\mathcal{Q})$;

(A5) *let* x^* *be some point belonging to* $X^* \cap \bar{S}$ *and* $q^*(x^*) \in \mathcal{Q}(x^*)$, *both obeying*
$\langle q^*(x^*), y - x^*\rangle \geq 0 \ \forall y \in K,$
and $\{\varphi_k\}$ *be a given nonnegative sequence. For an arbitrary sequence* $\{v^k\}$, $v^k \in S \cap D(\mathcal{Q}^k) \cap K^k$, *there exists a sequence* $\{w^k(v^k)\} \subset K \cap S$ *such that*
$\langle q^*(x^*), w^k(v^k) - v^k\rangle \leq c(\Gamma(x^*, v^k) + 1)\varphi_k \quad (c \geq 0 - \text{const.});$

(A6) *the non-negative sequences* $\{\varphi_k\}$ *(accuracy of approximation of* K *according to A6),* $\{\epsilon_k\}$ *(accuracy of approximation of* $\mathcal{Q}$*),* $\{\chi_k\}$ *(regularization parameter) and* $\{\delta_k\}$ *(exactness for solving the auxiliary problems) satisfy*

$$0 < \chi_k \leq 1, \quad \sum_{k=1}^{\infty} \frac{\varphi_k}{\chi_k} < \infty, \quad \sum_{k=1}^{\infty} \frac{\delta_k}{\chi_k} < \infty, \quad \sum_{k=1}^{\infty} \frac{\epsilon_k}{\chi_k} < \infty;$$

REMARK 1 *Of course, condition A3(ii) implies A3(i) considered for for* $\mathcal{Q} = \partial f$. *We separate A3(ii) bearing in mind different extensions of the subdifferential – and non-subdifferential – operators used below.*

Condition A5 seems to be rather unusual, especially due to the unknown element $q^*(x^*)$. *However, for a series of variational inequalities in mechanics and physics we have a helpful a priori information about* $q^*(x^*)$ *(see, for instance, the characterization of a solution for the problem on a steady movement of a fluid in a domain bounded by a semi-permeable membrane in Lions, 1969, Sect. 2.8.1, and for the contact problem between elastic bodies in Hlaváček et al., 1988, Theorem 2.1.1).*

In the general situation, since the constant $c \geq 0$ *in A5 is arbitrary, the estimate*

$$\|w^k(v^k) - v^k\| \leq c_1\varphi_k$$

is sufficient for the fulfillment of A5.

The assumption $K^k \supset K$ excludes, for instance, that K^k is obtained by applying usual discretization techniques to elliptic variational inequalities. But, the GPPM can be used in this case, too, by inserting a discretization procedure immediately into the algorithm for solving the auxiliary problems.

ASSUMPTION 2 (defining the regularizing functional h):

(B2) *h is Gâteaux-differentiable on* S*;*

(B3) *the functional* $x \to \frac{1}{2}\langle \mathcal{B}x, x\rangle + h(x)$ *is strictly convex on* $\bar{S}$*;*

(B4) $X^* \cap \bar{S} \neq \emptyset$*;*

(B5) *the set* $L_1(x, \delta) = \{y \in S : \Gamma(x, y) \leq \delta\}$ *is bounded for each* $x \in \bar{S}$ *and each* δ*;*

(B6) *if the sequences* $\{v^k\} \subset S$, $\{y^k\} \subset S$ *converge weakly to* v *and* $\lim_{k\to\infty} \Gamma(v^k, y^k) = 0$, *then* $\lim_{k\to\infty} \left[\Gamma(v, v^k) - \Gamma(v, y^k)\right] = 0$*;*

(B7) *if* $\{v^k\} \subset \bar{S}$ *is bounded,* $\{y^k\} \subset S$, $y^k \to \bar{y}$ *and* $\lim_{k\to\infty} \Gamma(v^k, y^k) = 0$, *then* $\lim_{k\to\infty} \|v^k - y^k\| = 0$;

(B8) *if* $\{v^k\} \subset S$, $\{y^k\} \subset S$, $v^k \rightharpoonup v$, $y^k \rightharpoonup y$ *and* $v \neq y$, *then*
$$\underline{\lim}_{k\to\infty} |\langle \nabla h(v^k) + \mathcal{B}v^k - \nabla h(y^k) - \mathcal{B}y^k, v - y\rangle| > 0;$$

(B9) $\forall z \in X'$, $\exists x \in S : \nabla h(x) + \mathcal{B}x = z$.

For the case $\mathcal{B} = \mathbf{0}$, the totality of conditions B1 - B9 is similar to the system of hypothesizes for Bregman functions in Burachik and Iusem (1998), only B7 is stronger than the corresponding assumption in the paper mentioned, where $v^k \rightharpoonup y$ stands in place of $\lim_{k\to\infty} \|v^k - y^k\| = 0$. In the cases D2 and D3 (see Lemma 2 and Theorem 2 below), this strengthening is not needed if $\mathcal{B}$ is a compact operator. At the same time, the use of B7 permits us to extend the class of operators $\mathcal{Q}$ by including the case D1.

If $\mathcal{B} = \mathbf{0}$, $X = \mathbb{R}^n$, conditions B1-B9 can be derived from the standard hypotheses for Bregman functions (see the analysis in Burachik and Iusem, 1998, Sect. 7).

The conditions B2 and B3 ensure that $\Gamma(x, y) > 0, \Gamma_1(x, y) > 0$ hold for $x \neq y$, and obviously $\Gamma(x, x) = 0, \Gamma_1(x, x) = 0$.

Remark 2 *It is quite clear that the conditions B1-B9 do not exclude the use of quadratic functionals* h*. In particular, the pair* $h = \frac{1}{2}\|\cdot\|^2$, $\mathcal{B} = \mathbf{0}$, *corresponding to the classical proximal point method, satisfies B1-B9. Thus, the notion "non-quadratic" (methods) means here, as well as in a series of preceding papers, "not only quadratic" and indicates the direction of the investigation.*

The consideration of an approximation of K by $\{K^k\}$ addresses, in particular, the situation when K is given in the form $K = K_1 \cap K_2$ and we choose h by taking into account the set K_1 only. In this case $K^k = K_1 \cap K_2^k$ is natural.

Let us give a simple example illustrating the choice of the functional h. Let $X = \mathbb{R}^n$, $K = \{x \in \mathbb{R}^n : x_j \geq 0,\ j = 1, \ldots, n_1;\ \sum_{j=n_1+1}^{n_2} j|x_j| \leq 1\}$ with $0 < n_1 < n_2 < n$,

$$\mathcal{Q} : x \to (\mathcal{A}(x_1, \ldots, x_{n_1}), x_{n_1+1} - 1, \ldots, x_n - 1),$$

where $\mathcal{A} : \mathbb{R}^{n_1} \to \mathbb{R}^{n_1}$ is an arbitrary continuous and monotone operator such that the corresponding Problem (P) is solvable. Then, considering the approximation

$$\sum_{j=n_1+1}^{n_2} j\sqrt{x_j^2+\tau_k} \le 1+\sqrt{\tau_k}\sum_{j=n_1+1}^{n_2} j\Big\},$$

where $\tau_k \to +0$, take $\mathcal{B} : x \to (0, \ldots, 0, x_{n_1+1}, \ldots, x_n)$. In this case it is easy to verify that the choice of $\{K^k\}$,

$$h(x) = \sum_{j=1}^{n_1} x_j \ln x_j - x_j (\text{with } 0 \times \ln 0 = 0 \text{ by convention})$$

and $S = \{x \in \mathbb{R}^n : x_j > 0, j = 1, \ldots, n_1\}$ satisfies the conditions A1–A4, B1–B9, while A6 can be fulfilled using an appropriate sequence

$$\{\tau_k\} : \sum_{k=1}^{\infty} \tau_k^{1/2}\chi_k^{-1} < \infty.$$

In the method under consideration we use a successive approximation of the operator $\mathcal{Q}$ based on concepts of the ϵ-subdifferential and the ϵ-enlargement of a maximal monotone operator. For a proper convex functional $f : X \to \bar{\mathbb{R}} \equiv \mathbb{R} \cup \{+\infty\}$, the ϵ-subdifferential $\partial_\epsilon f$ ($\epsilon \ge 0$) is defined as

$$\partial_\epsilon f(x) = \{u \in X' : \ f(y) - f(x) - \langle u, y - x\rangle \ge -\epsilon \ \ \forall y \in X\}.$$

In the sequel we suppose that f is lsc, and hence the subdifferential operator ∂f is maximal monotone.

According to Burachik, Iusem and Svaiter (1997) the ϵ-enlargement $\mathcal{T}_\epsilon$ ($\epsilon \ge 0$) of an arbitrary maximal monotone operator $\mathcal{T} : X \to 2^{X'}$ is defined as

$$\mathcal{T}_\epsilon(x) = \{u \in X' : \ \langle v - u, y - x\rangle \ge -\epsilon \ \ \forall y \in D(\mathcal{T}), v \in \mathcal{T}(y)\}.$$

If we identify operators with their graphs, for $\epsilon_1 \ge \epsilon_2 \ge 0$, $\epsilon \ge 0$ the relations

$$\mathcal{T}_{\epsilon_1} \supset \mathcal{T}_{\epsilon_2} \supset \mathcal{T}_0 = \mathcal{T}, \ \ \partial_{\epsilon_1} f \supset \partial_{\epsilon_2} f \supset \partial_0 f = \partial f \text{ and } \ (\partial f)_\epsilon \supset \partial_\epsilon f$$

are obvious, and simple examples show that the inclusion $(\partial f)_\epsilon \supset \partial_\epsilon f$ may be a strict one.

The *Brøndsted-Rockafellar property* of ϵ-subdifferentials (see Brøndsted and Rockafellar, 1965) was extended to the case of operators $\mathcal{T}_\epsilon$ in Burachik, Sagastizábal and Svaiter (1999b) as follows:

for any $x \in D(\mathcal{T}_\epsilon)$, $v \in \mathcal{T}_\epsilon(x)$ and each $\eta > 0$
there exist $y \in D(\mathcal{T})$ and $u \in \mathcal{T}(y)$ such that
$\|x - y\| \le \dfrac{\epsilon}{\eta}$ and$\|u - v\|_{X'} \le \eta$.

This permits one to estimate the closeness between $\mathcal{T}_\epsilon$ and $\mathcal{T}$.

Let us mention also the inclusion

$$D(\mathcal{T}) \subset D(\mathcal{T}_\epsilon) \subset \overline{D(\mathcal{T})}$$

and refer to Brøndsted and Rockafellar (1965), Burachik, Iusem and Svaiter (1997) for other properties of operators $\partial_\epsilon f$ and $\mathcal{T}_\epsilon$ and also for the motivation of their use in numerical methods.

In case condition A3(i) is valid, we choose a continuous operator $\widehat{\mathcal{Q}} : X \to X'$ such that both $(\mathcal{Q} - \widehat{\mathcal{Q}})$ and $(\widehat{\mathcal{Q}} - \mathcal{B})$ are monotone and define

$$\mathcal{Q}^k = (\mathcal{Q} - \widehat{\mathcal{Q}})_{\epsilon_k} + \widehat{\mathcal{Q}}.$$

But, if A3(ii) holds, then $\mathcal{Q}^k$ is constructed as

$$\mathcal{Q}^k = \partial_{\epsilon_k}(f - \widehat{f}) + \nabla \widehat{f},$$

with $\widehat{f}$ a functional continuous differentiable on X such that $(f - \widehat{f})$ and $\widehat{f}$ are convex functionals.

Then, $\mathcal{Q} \subset \mathcal{Q}^k \subset \mathcal{Q}_{\epsilon_k}$ follows immediately from the definitions of $\mathcal{Q}_{\epsilon_k}$ and $\mathcal{Q}^k$, and for $\mathcal{Q}^k = (\mathcal{Q} - \widehat{\mathcal{Q}})_{\epsilon_k} + \widehat{\mathcal{Q}}$ the relation (2) is obvious. But, in case $\mathcal{Q} = \partial f$ and $\mathcal{Q}^k = \partial_{\epsilon_k}(f - \widehat{f}) + \nabla \widehat{f}$, the inequality

$$\begin{aligned} &f(y) - f(x) - \langle q^k(x), y - x \rangle \geq -\epsilon_k \\ &\forall x \in D(\mathcal{Q}^k),\ y \in D(\mathcal{Q}),\ q^k(x) \in \mathcal{Q}^k(x) \end{aligned} \tag{3}$$

can be immediately concluded from the definition of the ϵ-subdifferential and the gradient inequality for convex functionals. Besides, according to Ekeland and Temam (1976), Sect. 1.6.3, $D(\partial_\epsilon(f - \widehat{f})) = dom(f - \widehat{f})$ is valid if $\epsilon > 0$. Therefore $D(\mathcal{Q}^k) = dom f$ has to be. Using A3(ii), we obtain

$$\begin{aligned} &f(x) - f(y) - \langle q(y), x - y \rangle \geq \langle \mathcal{B}(x - y), x - y \rangle \\ &\forall x \in dom f,\ y \in D(\mathcal{Q}),\ q(y) \in \mathcal{Q}(y), \end{aligned}$$

and together with (3) this yields (2), too.

Now, let us recall the method under consideration.

Generalized proximal point method (GPPM): Let $x^1 \in S$ be arbitrarily chosen, and at the $(k-1)$-th step let $x^k \in K^{k-1} \cap S$ be defined. In the k-th step solve

$(P^k_{\delta_k})$ find $x^{k+1} \in K^k \cap \bar{S} : \exists\, q^k(x^{k+1}) \in \mathcal{Q}^k(x^{k+1})$ with

$$\begin{aligned} &\langle q^k(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1} \rangle \\ &\geq -\delta_k \sqrt{\Gamma_1(x, x^{k+1})}\ \ \forall x \in K^k \cap \bar{S}. \end{aligned} \tag{4}$$

By $(\bar{P}^k_0)$ we denote Problem (4) with $\delta_k = 0$, while (P^k_0) means Problem (4) with $\mathcal{Q}^k = \mathcal{Q}$ and $\delta_k = 0$.

The criterion for the approximate calculation of the solution of Problem $(\bar{P}^k_0)$

extend the convergence results obtained here to related algorithms with more reasonable criteria. We have analyzed this question in Kaplan and Tichatschke (2002), where, as mentioned, a special case of GPPM with $\mathcal{Q}^k \equiv \mathcal{Q}$ was considered.

REMARK 3 *In fact, relations $\mathcal{Q} \subset \mathcal{Q}^k \subset \mathcal{Q}_{\epsilon_k}$, (2), (3), and the Brøndsted–Rockafellar property of $\mathcal{Q}_{\epsilon_k}$ determine the qualities of the approximation of $\mathcal{Q}$ needed for the convergence of GPPM.*

Although the choice $\widehat{\mathcal{Q}} = \mathcal{B}$ and $\widehat{f} \equiv 0$ provides these qualities for $\mathcal{Q}^k$, the use of $\widehat{\mathcal{Q}}$ different from $\mathcal{B}$, or $\widehat{f} \not\equiv 0$ may be preferable if the treatment of $\mathcal{Q}^k$ becomes simpler. In particular, this concerns some variational problems in mathematical physics, in which the non-differentiable part of the goal functional has a specific structure. We refer to the problem of linear elasticity with given friction (see Kaplan and Tichatschke, 1997, Sect. 4.3, 5.2), in which the non-differentiable term of the energy functional is convex and positive homogeneous, and taking the smooth term as $\widehat{f}$ one can use Lemma A.2 (see Appendix) to construct $\mathcal{Q}^k$.

For strongly monotone (with modulus α) operators $\mathcal{Q}$, the notion of the $(\alpha-\epsilon)$-enlargement

$$\mathcal{Q}_\alpha^\epsilon(x) = \{u \in X' : \langle v - u, y - x\rangle \geq \alpha\|y - x\|^2 - \epsilon \quad \forall y \in D(\mathcal{Q}), v \in \mathcal{Q}(y)\}$$

was introduced in Salmon, Nguyen and Strodiot (2000). One can easily see that

$$\mathcal{Q}_\alpha^\epsilon(x) = (\mathcal{Q} - \alpha I)_\epsilon(x) + \alpha I(x)$$

holds true, with $I : X \to X'$ the canonical isometry operator, i.e. $\mathcal{Q}_\alpha^\epsilon$, is a very particular case of the operators $(\mathcal{Q} - \widehat{\mathcal{Q}})_\epsilon + \widehat{\mathcal{Q}}$.

REMARK 4 *If f_1 and f_2 are two convex functionals and*

$$f_2(x) - \epsilon \leq f_1(x) \leq f_2(x) \quad \forall x \in X$$

holds true, then obviously, for each x, any subgradient of f_2 is an ϵ-subgradient of f_1.

Moreover, let the operator $\mathcal{Q}$ in the variational inequality (P) be split up into the sum of a continuous operator $\widehat{\mathcal{Q}}$ (with a reserve of monotonicity $\mathcal{B}$) and a subdifferential of a convex continuous functional φ. If we take a convex functional φ^k satisfying

$$\varphi^k(x) - \epsilon_k \leq \varphi(x) \leq \varphi^k(x) \quad \forall x \in X \tag{5}$$

and define $x^{k+1} \in K^k \cap \bar{S}$, $r^k(x^{k+1}) \in \partial\varphi^k(x^{k+1})$ such that

$$\langle \widehat{\mathcal{Q}}(x^{k+1}) + r^k(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1}\rangle$$

then $\widehat{\mathcal{Q}}(x^{k+1})+r^k(x^{k+1}) \in \mathcal{Q}^k(x^{k+1})$, *and the pair* x^{k+1}, $q^k(x^{k+1})$ *with* $q^k(x^{k+1}) = \widehat{\mathcal{Q}}(x^{k+1}) + r^k(x^{k+1})$ *solves Problem* $(P^k_{\delta_k})$.

For instance, in the problem of linear elasticity mentioned in Remark 3, such splitting is possible with $\widehat{\mathcal{Q}}$ *an affine operator and the functional*

$$\varphi(x) = \int_{\Gamma_c} \mu|x_t|\, d\Gamma$$

(for this problem, X *is the Sobolev space* $[H^1(\Omega)]^2$, Γ_c *is a known part of the boundary* Γ *of the domain* $\Omega \subset \mathbb{R}^2$, x_t *is a tangential component of a vector* $x \in X$ *and* μ *is a given positive constant).*

In that case, by taking

$$\varphi^k(x) = \int_{\Gamma_c} \mu\sqrt{|x_t|^2 + \tau_k}\, d\Gamma$$

with $0 < \tau_k \leq \left(\frac{\epsilon_k}{\mu \ mes\Gamma_c}\right)^2$, *the relations (5) are satisfied, the functional* φ^k *is differentiable (infinitely many times), and there are no serious troubles to calculate* r^k.

A similar situation arises for the problem of a laminar stationary flow of a Bingham fluid (see Glowinski, Lions and Trémolières, 1981) and for a series of other variational inequalities in mathematical physics.

3. Solvability of Problem $(P^k_{\delta_k})$

In this section we show existence and uniqueness of the solution of Problem (P^k_0), and the validity of the inclusion $x^{k+1} \in S$ for a solution of $(P^k_{\delta_k})$.

According to B1 and the definition of $\mathcal{B}$, the subdifferential operator $\partial\eta$ is maximal monotone. The conditions B1-B3 and B9 provide that $D(\partial\eta) = S$. Indeed, the inclusion $D(\partial\eta) \supset S$ follows from B2, and, assuming that $\partial\eta(x) \neq \emptyset$ holds for some $x \in \bar{S} \setminus S$, in view of B3 we obtain

$$\langle \nabla\eta(y) - \xi(x), y - x\rangle > 0 \quad \forall y \in S,\ \xi(x) \in \partial\eta(x).$$

But, for a fixed $\xi(x) \in \partial\eta(x)$, due to B9, there exists $y \in S$ such that $\nabla\eta(y) = \xi(x)$, in contradiction with the last inequality.

The conclusion $D(\partial\eta) = S$ means that $D(\nabla h) = S$, and both operators $\nabla\eta$ and ∇h are maximal monotone.

Thus, if Problem (P^k_0) is solvable, then it has a unique solution, here denoted by $\bar{x}^{k+1}$ (the strict monotonicity of $\mathcal{Q} + \chi_k\nabla h$ on $S \cap K^k \cap D(\mathcal{Q})$ follows immediately from A3 and B3), and $\bar{x}^{k+1} \in S$. Then, of course, the solution x^{k+1} of Problem $(P^k_{\delta_k})$ exists, and $D(\nabla h) = S$ provides $x^{k+1} \in S$.

Because the operator ∇h is maximal monotone and S is an open set, the maximal monotonicity of the operators $\mathcal{Q} + \chi_k\nabla h + \mathcal{N}_{K^k}$ and $x \to \mathcal{Q}(x) + \chi_k\nabla h(x) + \mathcal{N}_{K^k}(x) - \chi_k\nabla h(x^k)$ follows from A1, A2 and $\chi_k > 0$ (see A6)

Since $K^k \cap S \neq \emptyset$, the Moreau–Rockafellar theorem yields

$$\mathcal{N}_{K^k \cap \bar{S}} = \mathcal{N}_{K^k} + \mathcal{N}_{\bar{S}}.$$

Taking into account that $D(\nabla h) = S$, this permits us to transform Problem (P_0^k) into the inclusion

$$\begin{aligned} 0 &\in \mathcal{Q}(x) + \chi_k \nabla h(x) + \mathcal{N}_{K^k}(x) - \chi_k \nabla h(x^k) \\ &= \mathcal{Q}(x) - \chi_k \mathcal{B} x + \mathcal{N}_{K^k}(x) + \chi_k \mathcal{B} x^k + \chi_k(\nabla \eta(x) - \nabla \eta(x^k)), \end{aligned}$$

and with regard to A1, A3 and $0 < \chi_k \leq 1$, the operator

$$x \to \mathcal{Q}(x) - \chi_k \mathcal{B} x + \mathcal{N}_{K^k}(x) + \chi_k \mathcal{B} x^k$$

is maximal monotone (see Proposition 2.6 in Renaud and Cohen, 1997). Now, applying Lemma 5 by Burachik and Iusem (1998), one can conclude the solvability of Problem (P_0^k).

So, the following statement is proved:

THEOREM 1 *Let the conditions A1–A3, B1–B3, B8, B9 be valid and $\{\chi_k\}$ satisfy $0 < \chi_k \leq 1$. Then Problem (P_0^k) is uniquely solvable (for each k), the sequence $\{x^k\}$ is well defined and it is contained in S.*

REMARK 5 *Using instead of B9 the condition (see Iusem, 1995)*

$$\{v^k\} \subset S,\ v^k \to v \in \bar{S} \setminus S \implies \lim_{k \to \infty} \langle \nabla h(v^k), y - v^k \rangle = -\infty \ \forall y \in S,$$

the conclusion $D(\partial \eta) = S$ can be obtained from Lemma 1 of Burachik and Iusem (1998), and a result on solvability, like Theorem 2 in the paper mentioned, holds also true.

4. Convergence analysis

First we need the following assertion proved in Kaplan and Tichatschke (2000a).

LEMMA 1 *Let $C \subset X$ be a convex closed set, the operators $\mathcal{A} : X \to 2^{X'}$, $\mathcal{A} + \mathcal{N}_C$ be maximal monotone and $D(\mathcal{A}) \cap C$ be a convex set. Moreover, assume that the operator*

$$\mathcal{A}_C : v \to \begin{cases} \mathcal{A}(v) & \text{if } v \in C \\ \emptyset & \text{otherwise} \end{cases}$$

is locally hemi-bounded at each point $v \in D(\mathcal{A}) \cap C$ and that, for some $u \in D(\mathcal{A}) \cap C$ and each $v \in D(\mathcal{A}) \cap C$, there exists $\zeta(v) \in \mathcal{A}(v)$ satisfying

$$\langle \zeta(v), v - u \rangle \geq 0.$$

Then, with some $\zeta \in \mathcal{A}(u)$, the inequality

$$\langle \zeta, v - u \rangle \geq 0$$

REMARK 6 *Here, a weakened notion of the local hemi-boundedness is supposed. We call an operator $\mathcal{M} : X \to 2^{X'}$ locally hemi-bounded at a point v^0, if for each v, $v \neq v^0$, there exists a number $t_0(v^0, v) > 0$ such that the set*

$$\bigcup_{0<t\leq t_0(v^0,v)} \mathcal{M}(v^0 + t(v - v^0)) \text{ is bounded in } X'.$$

The standard notion supposes the boundedness of

$$\bigcup_{0\leq t\leq t_0(v^0,v)} \mathcal{M}(v^0 + t(v - v^0)).$$

This relaxation may be significant, see - for instance - the following example: $\mathcal{M} = \mathcal{N}_C$, with $C = \{v \in X = \mathbb{R}^n : \sum_{i=1}^n v_i^2 \leq 1\}$, $n > 1$.

LEMMA 2 *Let the sequence $\{x^k\}$, generated by the GPPM, belong to S and assume that, for some $x^* \in X^* \cap \bar{S}$, condition A5 is valid. Moreover, let the conditions A3, A6 and B1, B2, B5 be fulfilled. Then, the sequence $\{\Gamma(x^*, x^k)\}$ is convergent, $\{x^k\}$ is bounded, and*

$$\lim_{k\to\infty} \Gamma(x^{k+1}, x^k) = 0.$$

The proof of this statement repeats the proof of Lemma 2 in Kaplan and Tichatschke (2002) with one single alteration: The inequality

$$\langle q^k(x^{k+1}), x^* - x^{k+1}\rangle \leq \langle q^*(x^*), x^* - x^{k+1}\rangle + \epsilon_k,$$

which follows immediately from A3 and (2), is used in place of

$$\langle q(x^{k+1}), x^* - x^{k+1}\rangle \leq \langle q^*(x^*), x^* - x^{k+1}\rangle$$

(see (13) in the paper cited).

In the sequel, in particular, we deal with the case that, besides the usual property of maximal monotonicity, the operator $\mathcal{Q}$ is paramonotone and possesses the following

Ψ-property: *If $\{v^k\} \subset D(\mathcal{Q})$ converges weakly to $v \in D(\mathcal{Q}) \cap \bar{S} \cap K$ and*

$$\overline{\lim}_{k\to\infty}\langle w^k, v^k - v\rangle \leq 0$$

holds with $w^k \in \mathcal{Q}(v^k)$, then for each $y \in D(\mathcal{Q})$ there exists $w \in \mathcal{Q}(v)$ such that

$$\langle w, v - y\rangle \leq \overline{\lim}_{k\to\infty}\langle w^k, v^k - y\rangle.$$

An operator $\mathcal{Q}$ possesses this property, for instance, if

(i) $\mathcal{Q}$ is pseudomonotone in the sense of Brézis–Lions (see Lions, 1969, Sect. 2.2.4),

or (assuming the monotonicity of $\mathcal{Q}$) if

(ii) $\{v^k\} \subset D(\mathcal{Q})$, $v^k \rightharpoonup v \in D(\mathcal{Q}) \cap \bar{S} \cap K$ and $w^k \in \mathcal{Q}(v^k)$, $\overline{\lim}_{k\to\infty}\langle w^k, v^k -$

The first claim follows immediately from the definition of the pseudomonotonicity, and the second one from the relation

$$\langle w^k, v^k - y\rangle = \langle w, v - y\rangle + \langle w^k, v^k - v\rangle - \langle w - w^k, v - y\rangle$$
$$\geq \langle w, v - y\rangle + \langle w, v^k - v\rangle - \langle w - w^k, v - y\rangle,$$

which holds true with $w \in \mathcal{Q}(v)$ in view of the monotonicity of $\mathcal{Q}$.

In particular, in case of $X = \mathbb{R}^n$, a maximal monotone operator $\mathcal{Q}$ possesses the Ψ-property if $int D(\mathcal{Q}) \supset \bar{S} \cap K \cap D(\mathcal{Q})$.

Definition 1 *The operator $\mathcal{A} : X \to 2^{X'}$ is called paramonotone in a set $C \subset X$ if it is monotone and*

$$\langle z - z', v - v'\rangle = 0 \text{ with } v, v' \in C,\ z \in \mathcal{A}(v),\ z' \in \mathcal{A}(v')$$

implies $z \in \mathcal{A}(v'),\ z' \in \mathcal{A}(v)$.

We will use the following property of a paramonotone operator $\mathcal{A}$ in C (see Iusem, 1998):

Property (∗) *If x^* solves the variational inequality*

$$\langle \mathcal{A}(x), y - x\rangle \geq 0 \quad \forall y \in C \tag{6}$$

and for $\bar{x} \in C$ there exists $\bar{z} \in \mathcal{A}(\bar{x})$ with $\langle \bar{z}, x^* - \bar{x}\rangle \geq 0$, then $\bar{x}$ is also a solution of (6).

Remark 7 *Burachik and Iusem (1998) considered a non-quadratic proximal method for variational inequalities, where the operator is supposed to be maximal monotone, pseudomonotone and paramonotone.*

Lemma 3 *Let the assumptions of Lemma 2, as well as the conditions A1, A4 and B6, B7 be valid. Moreover, suppose that one of the following assumptions[2] is fulfilled:*

(D1) $S \supset \overline{D(\mathcal{Q}) \cap (\cup_{k \geq k_0} K^k)}$ for an arbitrary large k_0, ∇h is Lipschitz continuous on closed and bounded subsets of S and the conditions of Lemma 1 hold with $\mathcal{A} := \mathcal{Q},\ C := K \cap \bar{S}$;

(D2) $\mathcal{Q}$ is the subdifferential of a proper convex, lsc functional f, and f is continuous at some $x \in K$;

(D3) the operator $\mathcal{Q} : D(\mathcal{Q}) \to 2^{X'}$ possesses the Ψ-property, and $\mathcal{Q}$ is paramonotone in $\bar{S}$. Moreover[3], $\forall\ r > 0,\ \exists\ k(r),\ \epsilon(r) > 0 :$ $\mathcal{Q}$ is a bounded operator on

$$\Omega_r = \{x : \|x\| \leq r,\ dist\left(x, S \cap \left(\cup_{k \geq k(r)} K^k\right)\right) < \epsilon(r)\},$$

where $dist(x, V) = \inf_{v \in V} \|x - v\|$.

[2]For a motivation of the inclusion in D1, which prevents from the choice of a function h leading to interior point methods, see Eckstein (1993). In the case D2, condition A4 can be weakened assuming that each weak limit point of $\{v^k\}$ belongs to K (in place of $K \cap D(\mathcal{Q})$).

Then the sequence $\{x^k\}$, generated by the GPPM, is bounded and each weak limit point is a solution of Problem (P).

Proof. According to Lemma 2, the sequence $\{x^k\}$ is bounded, hence, there exists a weakly convergent subsequence $\{x^{j_k}\}$, $x^{j_k} \rightharpoonup \bar{x}$ as $k \to \infty$. In view of $\{x^k\} \subset S$, A4 and the convexity of S, the inclusion $\bar{x} \in \bar{S} \cap K \cap D(\mathcal{Q})$ ($\bar{x} \in \bar{S} \cap K$ in case D2) is valid.

Due to $\lim_{k\to\infty} \Gamma(x^{k+1}, x^k) = 0$, one can use condition B7 with $v^k := x^{j_k+1}$, $y^k := x^{j_k}$. This leads to

$$\lim_{k\to\infty} \|x^{j_k+1} - x^{j_k}\| = 0. \tag{7}$$

If D1 holds, then with regard to the boundedness of $\{x^k\}$, $\{x^k\}_{k\geq k_0} \subset \overline{D(\mathcal{Q})} \cap (\cup_{k\geq k_0} K^k)$, A6 and (7), the relation

$$\lim_{k\to\infty} \chi_{j_k} \langle \nabla h(x^{j_k+1}) - \nabla h(x^{j_k}), x - x^{j_k+1} \rangle = 0 \quad \forall x \in X \tag{8}$$

follows immediately. Now, take in (4) an arbitrary $x \in K \cap \bar{S}$ and replace $\langle q^k(x^{k+1}), x - x^{k+1} \rangle$ by $\langle q(x), x - x^{k+1} \rangle + \epsilon_k$ (this is possible in view of (2)). Then, passing to the limit in this altered inequality with $k := j_k$, $k \to \infty$, due to the boundedness of $\{x^k\}$, the definition of Γ_1, A6 and (8), we obtain

$$\langle q(x), x - \bar{x} \rangle \geq 0 \quad \forall x \in K \cap \bar{S}.$$

The conditions A1 and $S \supset \overline{D(\mathcal{Q}) \cap (\cup_{k\geq k_0} K^k)} \supset \overline{D(\mathcal{Q}) \cap K}$ guarantee the maximal monotonicity of the operator $\mathcal{Q} + \mathcal{N}_{K\cap\bar{S}}$ (in fact, $\mathcal{Q} + \mathcal{N}_{K\cap\bar{S}}$ coincides with $\mathcal{Q} + \mathcal{N}_K$). Thus, we are able to apply Lemma 1 with $C := K \cap \bar{S}$, $\mathcal{A} := \mathcal{Q}$, which ensures that

$$\exists\, q(\bar{x}) \in \mathcal{Q}(\bar{x}) : \ \langle q(\bar{x}), y - \bar{x} \rangle \geq 0 \quad \forall y \in K \cap \bar{S}.$$

This relation means that $-q(\bar{x}) \in \mathcal{N}_{K\cap\bar{S}}(\bar{x})$, hence

$$0 \in \mathcal{Q}(\bar{x}) + \mathcal{N}_{K\cap\bar{S}}(\bar{x}) = \mathcal{Q}(\bar{x}) + \mathcal{N}_K(\bar{x})$$

is valid, showing that $\bar{x} \in X^*$.

Now, take x^* as in A5. With regard to the symmetry of $\mathcal{B}$, a straightforward calculation gives

$$\begin{aligned} &-\langle \nabla h(x^k) - \nabla h(x^{k+1}), x^* - x^{k+1} \rangle = \Gamma(x^*, x^k) - \Gamma(x^*, x^{k+1}) \\ &- \Gamma(x^{k+1}, x^k) - \langle \mathcal{B}(x^{k+1} - x^k), x^* - x^{k+1} \rangle. \end{aligned} \tag{9}$$

Using Lemma 2, (7) and $0 < \chi_k \leq 1$, we infer from (9) that

$$\lim_{k\to\infty} \chi_{j_k} \langle \nabla h(x^{j_k+1}) - \nabla h(x^{j_k}), x^* - x^{j_k+1} \rangle = 0 \tag{10}$$

Suppose that D2 is valid. Then, relation (4) given with $x = x^*$ and $k := j_k$ implies, due to inequality[4] (3) that

$$\delta_{j_k}\sqrt{\Gamma_1(x^*, x^{j_k+1})} + \chi_{j_k}\langle \nabla h(x^{j_k+1}) - \nabla h(x^{j_k}), x^* - x^{j_k+1}\rangle \\ \geq f(x^{j_k+1}) - f(x^*) - \epsilon_{j_k}. \tag{11}$$

Taking the limit in (11) as $k \to \infty$, in view of Lemma 2, (10), A6 and the lower semi-continuity of f, one gets $f(\bar{x}) \leq f(x^*)$. Thus, $0 \in \partial(f(\bar{x}) + \delta(\bar{x}|K))$, $\delta(\cdot|K)$—the indicator function of K, and from the Moreau–Rockafellar theorem we have $0 \in \partial f(\bar{x}) + \mathcal{N}_K(\bar{x})$, proving that $\bar{x} \in X^*$.

Finally, let us consider the case D3. Using equality (9) with $k := j_k$ and $\bar{x}$ in place of x^*, from (7), Lemma 2 and condition B6 for $v^k := x^{j_k+1}$, $y^k := x^{j_k}$, we conclude that

$$\lim_{k\to\infty}\langle \nabla h(x^{j_k+1}) - \nabla h(x^{j_k}), \bar{x} - x^{j_k+1}\rangle = 0. \tag{12}$$

Due to the Brøndsted–Rockafellar property and the inclusion $\mathcal{Q}^l \subset \mathcal{Q}_{\epsilon_l}\ \forall\, l$, for each k there exist $\widetilde{x}^{j_k+1}$ and $q(\widetilde{x}^{j_k+1}) \in \mathcal{Q}(\widetilde{x}^{j_k+1})$ satisfying

$$\|x^{j_k+1} - \widetilde{x}^{j_k+1}\| \leq \sqrt{\epsilon_{j_k}}, \quad \|q^{j_k}(x^{j_k+1}) - q(\widetilde{x}^{j_k+1})\|_{X'} \leq \sqrt{\epsilon_{j_k}}. \tag{13}$$

Hence, $\widetilde{x}^{j_k+1} \to \bar{x}$ holds as $k \to \infty$. Taking r such that $\|x^{j_k+1}\| < r$, $\|\widetilde{x}^{j_k+1}\| < r$ for all k, choose $k(r)$ and $\epsilon(r) > 0$ according to D3. Because $x^{j_k+1} \in S \cap K^{j_k}$, using (13) and A6 , one can conclude that the inclusion $\widetilde{x}^{j_k+1} \in \Omega_r$ is valid for sufficiently large k ($k \geq k_0$). With regard to D3, this implies the boundedness of the sequence $\{q(\widetilde{x}^{j_k+1})\}$. Now, we use the identity

$$\langle q^{j_k}(x^{j_k+1}), x^{j_k+1} - x\rangle = \langle q(\widetilde{x}^{j_k+1}), \widetilde{x}^{j_k+1} - x\rangle \\ + \langle q(\widetilde{x}^{j_k+1}), x^{j_k+1} - \widetilde{x}^{j_k+1}\rangle + \langle q^{j_k}(x^{j_k+1}) - q(\widetilde{x}^{j_k+1}), x^{j_k+1} - x\rangle. \tag{14}$$

The boundedness of $\{x^{j_k+1}\}$ and $\{q(\widetilde{x}^{j_k+1})\}$ together with (13), A6 yield (for each x)

$$\lim_{k\to\infty}[\langle q(\widetilde{x}^{j_k+1}), x^{j_k+1} - \widetilde{x}^{j_k+1}\rangle \\ + \langle q^{j_k}(x^{j_k+1}) - q(\widetilde{x}^{j_k+1}), x^{j_k+1} - x\rangle] = 0. \tag{15}$$

But, from (4), (12), A6, the definition of Γ_1 as well as $\|x^{j_k+1}\| < r\ \forall\, k$, we get

$$\overline{\lim}_{k\to\infty}\langle q^{j_k}(x^{j_k+1}), x^{j_k+1} - \bar{x}\rangle \leq 0. \tag{16}$$

Combining (14), (15) (with $x = \bar{x}$) and (16), we derive

$$\overline{\lim}_{k\to\infty}\langle q(\widetilde{x}^{j_k+1}), \widetilde{x}^{j_k+1} - \bar{x}\rangle \leq 0.$$

[4] Here lies the reason for the use of $\mathcal{Q}^k = \partial_{\epsilon_k}(f - \widehat{f}) + \nabla\widehat{f}$ and A3(ii). In the case

Since the operator $\mathcal{Q}$ possesses the Ψ-property, the last relation provides existence of $q(\bar{x}) \in \mathcal{Q}(\bar{x})$ such that

$$\langle q(\bar{x}), \bar{x} - x^* \rangle \leq \overline{\lim}_{k\to\infty} \langle q(\tilde{x}^{j_k+1}), \tilde{x}^{j_k+1} - x^* \rangle. \tag{17}$$

Applying again (14) and (15) (now with $x = x^*$), one gets

$$\overline{\lim}_{k\to\infty} \langle q(\tilde{x}^{j_k+1}), \tilde{x}^{j_k+1} - x^* \rangle = \overline{\lim}_{k\to\infty} \langle q^{j_k}(x^{j_k+1}), x^{j_k+1} - x^* \rangle. \tag{18}$$

But, similarly to (16) (with (10) instead of (12)), we conclude that

$$\overline{\lim}_{k\to\infty} \langle q^{j_k}(x^{j_k+1}), x^{j_k+1} - x^* \rangle \leq 0. \tag{19}$$

In view of (17)–(19), the property (∗) of paramonotone operators permits us to claim that $\bar{x} \in X^*$.

If $\epsilon_k \equiv 0$, i.e. $\mathcal{Q}^k \equiv \mathcal{Q}$, formally (13) gives $\tilde{x}^{j_k+1} = x^{j_k+1}$, $q(\tilde{x}^{j_k+1}) = q^{j_k}(x^{j_k+1})$, and a straightforward analysis of the given proof shows that boundedness of $\mathcal{Q}$ on Ω_r is not needed. ∎

THEOREM 2 *Let the conditions A1-A4, A6 and B1-B9 be valid, and condition A5 hold for each $x^* \in X^* \cap \bar{S}$ (constant c in A5 may depend on x^*). Moreover, let one of the assumptions D1, D2 or D3 in Lemma 3 be fulfilled. Then the sequence $\{x^k\}$, generated by the GPPM, converges weakly to a solution of Problem (P).*

Proof. The existence of the sequence $\{x^k\}$ and the inclusion $\{x^k\} \subset S$ are guaranteed by Theorem 1. Denote

$$d_k(x) = \Gamma(x, x^k) - \frac{1}{2}\langle \mathcal{B}x, x \rangle - h(x).$$

According to Lemma 2, the sequence $\{\Gamma(x, x^k)\}$ converges for each $x \in X^* \cap \bar{S}$, hence, the sequence $\{d_k(x)\}$ possesses the same property. Boundedness of $\{x^k\}$ was proved in Lemma 2, and Lemma 3 yields that each weak limit point of $\{x^k\}$ belongs to $X^* \cap \bar{S}$. Assume that $\{x^{j_k}\}$ and $\{x^{i_k}\}$ converge weakly to $\bar{x}$, $\tilde{x}$, correspondingly. Then, it holds that $\bar{x}, \tilde{x} \in X^* \cap \bar{S}$. Let

$$l_1 = \lim_{k\to\infty} d_k(\bar{x}), \quad l_2 = \lim_{k\to\infty} d_k(\tilde{x}).$$

Obviously,

$$l_1 - l_2 = \lim_{k\to\infty} (d_k(\bar{x}) - d_k(\tilde{x})) = \lim_{k\to\infty} \langle \nabla h(x^k) + \mathcal{B}x^k, \tilde{x} - \bar{x} \rangle.$$

Considering the latter equality now for the subsequences $\{x^{j_k}\}$ and $\{x^{i_k}\}$, one can conclude that

$$\lim_{k\to\infty} \langle \nabla h(x^{j_k}) + \mathcal{B}x^{j_k} - \nabla h(x^{i_k}) - \mathcal{B}x^{i_k}, \tilde{x} - \bar{x} \rangle = 0. \tag{20}$$

A comparison of (20) and B8 (given with $v^k := x^{j_k}$, $y^k := x^{i_k}$) indicates that $\tilde{x} = \bar{x}$, proving uniqueness of the weak limit point of $\{x^k\}$. ∎

REMARK 8 *Theorems 1 and 2 remain true if condition B9 is replaced by any other condition guaranteeing that $\{x^k\}$ is well defined and $\{x^k\} \subset S$ (see, in*

Appendix

The verification of the inclusion $q \in \mathcal{Q}^k(x)$ for fixed q, x may be very difficult, just as the calculation of an element $q \in \mathcal{Q}^k(x)$ with certain properties. For implementable strategies, where the ϵ-enlargements or ϵ-subdifferentials are used in numerical methods, we refer to Burachik, Sagastizábal and Svaiter (1999a) and Kiwiel (1999).

In a series of variational inequalities in mathematical physics $\mathcal{Q}$ possesses the structure

$$\mathcal{Q} = \tilde{\mathcal{Q}} + \mathcal{F}, \tag{A.1}$$

with $\tilde{\mathcal{Q}} : X \to X'$ a continuous and monotone operator and $\mathcal{F}$ the subdifferential of a convex, positive homogeneous lsc. functional $f : X \to \bar{\mathbb{R}}$. In particular, for the contact problem with a given friction (see Hlaváček et al., 1988) and for the problem of a laminar stationary flow of a Bingham fluid (Glowinski, Lions and Trémolières, 1981), $\tilde{\mathcal{Q}}$ is an affine, monotone operator and $\mathcal{F}$ possesses the properties mentioned.

With the structure of the operator $\mathcal{Q}$ as in (A.1), we usually get a much more convenient construction for $\mathcal{Q}^k = \tilde{\mathcal{Q}} + \mathcal{F}_{\epsilon_k}$ than for $(\tilde{\mathcal{Q}} + \mathcal{F})_{\epsilon_k}$, moreover, $\mathcal{F}_\epsilon = \partial_\epsilon f$ holds for such $\mathcal{F}$. This is based on the following statements.

Lemma A.1 *Let j be a convex, positive homogeneous, lsc functional and $\mathcal{T} = \partial j$. Then, for each $\epsilon > 0$,*

$$\mathcal{T}_\epsilon(x) = \{u \in \mathcal{T}(0) : \langle u, x\rangle \geq -\epsilon + j(x)\}. \tag{A.2}$$

Proof. The maximal monotonicity of $\mathcal{T}$ is well-known, and

$$\mathcal{T}(\lambda x) = \mathcal{T}(x) \quad \forall \lambda > 0,\ x \in D(\mathcal{T}) \tag{A.3}$$

follows immediately from the definition of the subdifferential and the identity

$$j(\lambda x) = \lambda j(x) \quad \forall \lambda \geq 0,\ x \in dom\, j.$$

Due to the maximal monotonicity of $\mathcal{T}$ and (A.3), we have

$$\mathcal{T}(0) \supset \mathcal{T}(x) \text{ for any } x \in X. \tag{A.4}$$

If $v \notin \mathcal{T}(0)$, then $v \notin \mathcal{T}_\epsilon(x)$ holds for all $x \in X$ and $\epsilon > 0$. Indeed, in this case there exist $y \in D(\mathcal{T})$, $z \in \mathcal{T}(y)$ such that

$$\langle z - v, y - 0\rangle < 0.$$

Then, with arbitrarily chosen $x \in X$ and $\epsilon > 0$, we obtain and $\mathcal{T}(\lambda y) = \mathcal{T}(y)$ ensures that $z \in \mathcal{T}(\lambda y)$. Thus, $v \notin \mathcal{T}_\epsilon(x)$, and this proves

$$\mathcal{T}_\epsilon(x) \subset \mathcal{T}(0) \;\; \forall x \in X, \; \epsilon > 0. \tag{A.5}$$

Now, let $u(x) \in \mathcal{T}_\epsilon(x)$. By definition of the ϵ-enlargement, there must be

$$\langle z - u(x), y - x\rangle \geq -\epsilon \;\; \forall y \in D(\mathcal{T}), \; z \in \mathcal{T}(y),$$

hence,

$$\langle u(x), x\rangle \geq \langle z, x\rangle + \langle u(x) - z, y\rangle - \epsilon \;\; \forall y \in D(\mathcal{T}), \; z \in \mathcal{T}(y).$$

Because $0 \in D(\mathcal{T})$, the last inequality yields

$$\langle u(x), x\rangle \geq \langle z, x\rangle - \epsilon \;\; \forall z \in \mathcal{T}(0). \tag{A.6}$$

Combining (A.5) and (A.6) one can conclude that

$$\mathcal{T}_\epsilon(x) \subset \{u \in \mathcal{T}(0) : \langle u, x\rangle \geq -\epsilon + \sup_{w \in \mathcal{T}(0)} \langle w, x\rangle\}. \tag{A.7}$$

But, the inverse inclusion holds, too. Indeed, if $u \in \mathcal{T}(0)$, then

$$\langle z - u, y - 0\rangle \geq 0 \;\; \forall y \in D(\mathcal{T}), \; z \in \mathcal{T}(y) \tag{A.8}$$

follows from the monotonicity of $\mathcal{T}$. Assuming additionally that

$$\langle u, x\rangle \geq -\epsilon + \sup_{w \in \mathcal{T}(0)} \langle w, x\rangle, \tag{A.9}$$

we obtain by means of (A.4), (A.8) and (A.9) that

$$\begin{aligned}&\langle z - u, y - x\rangle = \langle z - u, y\rangle + \langle u, x\rangle - \langle z, x\rangle \\ &\geq \langle u, x\rangle - \sup_{w \in \mathcal{T}(0)} \langle w, x\rangle \geq -\epsilon \;\; \forall y \in D(\mathcal{T}), \; z \in \mathcal{T}(y).\end{aligned}$$

Thus, $u \in \mathcal{T}_\epsilon(x)$, and

$$\mathcal{T}_\epsilon(x) \supset \{u \in \mathcal{T}(0) : \; \langle u, x\rangle \geq -\epsilon + \sup_{w \in \mathcal{T}(0)} \langle w, x\rangle\}. \tag{A.10}$$

But, the proof of Proposition 4.1.1 in Ioffe and Tikhomirov (1975) leads immediately to

$$\sup_{w \in \mathcal{T}(0)} \langle w, x\rangle = j(x), \tag{A.11}$$

and (A.7), (A.10), (A.11) yield (A.2). ∎

REMARK A.1 *R.T. Rockafellar has drawn our attention to the fact that any maximal monotone operator $\mathcal{T}$ with property (A.3) is the subdifferential of some convex functional. Indeed, it follows immediately from the relations (A.7) and*

LEMMA A.2 *Let j be a convex, positive homogeneous, lsc functional. Then*

$$\partial_\epsilon j(x) = (\partial j)_\epsilon(x) \ \ \forall \epsilon > 0, \ x \in dom j. \tag{A.12}$$

Proof. The inclusion

$$\partial_\epsilon j(x) \subset \{u \in \partial j(0) : \ \langle u, x \rangle \geq -\epsilon + j(x)\} \ \ \forall \epsilon > 0 \tag{A.13}$$

is a straightforward consequence of $\partial_\epsilon j \subset (\partial j)_\epsilon$ and Lemma A.1. But for an arbitrary $u \in \partial j(0)$ the inequality

$$j(y) - \langle u, y \rangle \geq 0 \ \ \forall y \in X \tag{A.14}$$

holds true. Therefore, if

$$\langle u, x \rangle \geq -\epsilon + j(x), \tag{A.15}$$

then, summing up (A.14) and (A.15), we get

$$j(y) - j(x) - \langle u, y - x \rangle \geq -\epsilon \ \ \forall y \in X.$$

Hence, $u \in \partial_\epsilon j(x)$ and

$$\partial_\epsilon j(x) \supset \{u \in \partial j(0) : \ \langle u, x \rangle \geq -\epsilon + j(x)\}$$

can be concluded, proving (A.12). ■

Now, some examples are considered where the ϵ-subdifferential can be described in a constructive manner by means of Lemma A.2.

EXAMPLE A.1 Let $X = \{x \in H^1(0,1) : x(0) = 0\}, \ X' = X$,

$$\|x\| = \left(\int_0^1 \left(\frac{dx}{d\xi} \right)^2 d\xi \right)^{\frac{1}{2}} \text{ and } j(x) = \int_0^1 \left| \frac{dx}{d\xi} \right| d\xi.$$

Here, the choice of X and j corresponds to the problem of a laminar stationary flow of a Bingham fluid between two parallel plates (see Glowinski, Lions and Trémolières, 1981) if we take $K = \{x \in X : \ x(1) = 0\}$. According to the definition of ∂j, one gets

$$\partial j(0) = \left\{ u \in X : \ \int_0^1 \left| \frac{dy}{d\xi} \right| d\xi \geq \int_0^1 \frac{du}{d\xi} \frac{dy}{d\xi} d\xi \ \ \forall y \in X \right\}.$$

If $\left| \frac{du}{d\xi} \right| \leq 1$ a.e. on $(0,1)$, then the inequality

$$\int_0^1 \left| \frac{dy}{d\xi} \right| d\xi \geq \int_0^1 \frac{du}{d\xi} \frac{dy}{d\xi} d\xi \ \ \forall y \in X$$

is evident, hence, $u \in \partial j(0)$. Otherwise, denote $M = \{\xi \in (0,1) : |\frac{du}{d\xi}| > 1\}$ and define

$$v = \begin{cases} \dfrac{du}{d\xi} & \text{on } M, \\ 0 & \text{on } (0,1) \setminus M, \end{cases}$$

$$\widetilde{y}(\xi) = \int_0^\xi v d\xi.$$

This functional $\widetilde{y}$ is absolutely continuous, and using the formula of integration by parts, we obtain in a standard manner that v is a generalized derivative of $\widetilde{y}$ ($v \in L_2(0,1)$ is obvious). Therefore, $\widetilde{y} \in X$, and

$$j(\widetilde{y}) - \langle u, \widetilde{y} \rangle = \int_0^1 \left| \frac{d\widetilde{y}}{d\xi} \right| d\xi - \int_0^1 \frac{du}{d\xi} \frac{d\widetilde{y}}{d\xi} d\xi = \int_M \left[\left| \frac{du}{d\xi} \right| - \left(\frac{du}{d\xi} \right)^2 \right] d\xi < 0$$

proves that $u \notin \partial j(0)$. Thus,

$$\partial j(0) = \left\{ u \in X : \left| \frac{du}{d\xi} \right| \leq 1 \text{ a.e. on } (0,1) \right\}$$

and Lemmas A.1 and A.2 yield

$$\partial_\epsilon j(x) = \left\{ u \in X : \left| \frac{du}{d\xi} \right| \leq 1 \text{ a.e. on } (0,1), \right.$$
$$\left. \int_0^1 \frac{du}{d\xi} \frac{dx}{d\xi} d\xi \geq -\epsilon + \int_0^1 \left| \frac{dx}{d\xi} \right| d\xi \right\}. \quad \text{(A.16)}$$

The first condition in (A.16) implies

$$\int_0^1 \frac{du}{d\xi} \frac{dx}{d\xi} d\xi \geq - \int_0^1 \left| \frac{dx}{d\xi} \right| d\xi,$$

and if x fulfills $\int_0^1 \left| \frac{dx}{d\xi} \right| d\xi \leq \frac{\epsilon}{2}$, then the second condition is automatically valid. Hence,

$$\partial_\epsilon j(x) = \left\{ u \in X : \left| \frac{du}{d\xi} \right| \leq 1 \text{ a.e. on } (0,1) \right\}.$$

But in case $\int_0^1 |\frac{dx}{d\xi}| d\xi > \frac{\epsilon}{2}$, taking $u(\xi) = -\int_0^\xi \operatorname{sign} \frac{dx}{d\xi} d\xi$, one can see that the second condition in (A.16) is essential.

EXAMPLE A.2 Let $\Omega \subset \mathbb{R}^n$ be an open domain, $X = H_0^1(\Omega)$, $X' = H^{-1}(\Omega)$,

$$\|x\| = \left(\int \sum^n \left(\frac{dx}{\cdot\cdot} \right)^2 d\xi \right)^{\frac{1}{2}}, \quad j(x) = \|x\|.$$

It is well-known that the subdifferential of the norm at 0 coincides with the unit ball of the dual space. So, we obtain immediately from Lemmas A.1 and A.2 that

$$\partial_\epsilon j(x) = \{u \in H^{-1}(\Omega) : \|u\|_{H^{-1}} \leq 1, \langle u, x\rangle \geq \|x\| - \epsilon\}.$$

In case $\|u\|_{H^{-1}} \leq 1$, $\|x\| \leq \epsilon/2$, the inequality $\langle u, x\rangle \geq \|x\| - \epsilon$ is obvious, consequently,

$$\partial_\epsilon j(x) = \{u \in H^{-1}(\Omega) : \|u\|_{H^{-1}} \leq 1\} \text{ if } \|x\| \leq \frac{\epsilon}{2}.$$

But if $\|x\| > \epsilon/2$, the condition $\langle u, x\rangle \geq \|x\| - \epsilon$ is essential (take $u = -J(x)/\|x\|$ with $J : X \to X'$ the duality mapping).

The reader may convince himself on the following fact: Replacing the functional j in the examples above by a sum $j + \varphi$, with φ a convex, quadratic functional (for instance, $\varphi(x) = \|x\|^2$), a similar simple representation for the ϵ-subdifferential or of the operator $(\partial(j + \varphi))_\epsilon$ is not available. In particular, in order to verify that a given u belongs to $\partial_\epsilon(j + \varphi)(x)$, one has to calculate beforehand

$$\min\{j(y) + \varphi(y) - \langle u, y\rangle : y \in X\}.$$

In case $\varphi(x) = \|x\|^2$, this is an elliptic variational inequality of the second order.

REMARK A.2 *If* $X = \mathbb{R}^n$, $j(x) = \left(\sum_{i=1}^n x_i^2\right)^{1/2} = \|x\|$ *we obtain, quite similarly to Example A.2, that*

$$\partial_\epsilon j(x) = \{u \in \mathbb{R}^n : \|u\| \leq 1, \sum_{i=1}^n u_i x_i \geq \|x\| - \epsilon\}. \tag{A.17}$$

In this case, verification of the inclusion $u \in \partial_\epsilon j(x)$ *is quite trivial, and, for instance, calculation of an element* $u \in \partial_\epsilon j(x)$*, which minimizes the distance to a given point, requires solving a very simple convex programming problem, namely to minimize the distance from a point to the intersection of a ball and a halfspace. We suggest to compare the description of* $\partial_\epsilon j$ *in (A.17) and for* $j(x) = \|x\| + \|x\|^2$.

References

ALBER, Y., BURACHIK, R., and IUSEM, A. (1997) A proximal method for non-smooth convex optimization problems in Banach spaces. *Abstr. Appl. Anal.*, **2**, 97–120.

AUSLENDER, A. and HADDOU, M. (1995) An interior-proximal method for convex linearly constrained problems and its extension to variational in-

AUSLENDER, A., TEBOULLE, M., and BEN-TIBA, S. (1999) Interior proximal and multiplier methods based on second order homogeneous kernels. *Mathematics of Oper. Res.*, **24**, 645–668.

BRØNDSTED, A. and ROCKAFELLAR, R. (1965) On the subdifferentiability of convex functions. *Proc. Amer. Math. Soc.*, **16**, 605–611.

BURACHIK, R. and IUSEM, A. (1998) A generalized proximal point algorithm for the variational inequality problem in a Hilbert space. *SIAM J. Optim.*, **8**, 197–216.

BURACHIK, R., IUSEM, A., and SVAITER, B. (1997) Enlargements of maximal monotone operators with application to variational inequalities. *Set-Valued Analysis*, **5**, 159–180.

BURACHIK, R., SAGASTIZÁBAL, C., and SVAITER, B. (1999a) Bundle methods for maximal monotone operators. In: Thera, M. and Tichatschke, R., eds., *Ill-posed Variational Problems and Regularization Techniques*, LNEMS, **477**, 49–64. Springer.

BURACHIK, R., SAGASTIZÁBAL, C., and SVAITER, B. (1999B) ϵ-enlargement of maximal monotone operators: Theory and applications. In: Fukushima, M., ed., *Reformulation—Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, Kluwer, 25–43.

BUTNARIU, D. and IUSEM, A. (1997) On a proximal point method for convex optimization in Banach spaces. *Numer. Funct. Anal. and Optimiz.*, **18**, 723–744.

BUTNARIU, D. AND IUSEM, A. (2000) *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization.* Kluwer, Dordrecht-Boston-London.

ECKSTEIN, J. (1993) Nonlinear proximal point algorithms using Bregman functions, with application to convex programming. *Math. Oper. Res.*, **18**, 202–226.

EKELAND, I. and TEMAM, R. (1976) *Convex Analysis and Variational Problems.* North-Holland, Amsterdam.

GLOWINSKI, R., LIONS, J.L., and TRÉMOLIÈRES, R. (1981) *Numerical Analysis of Variational Inequalities.* North-Holland, Amsterdam.

HLAVÁČEK, I., HASLINGER, J., NEČAS, I., and LOVIŠEK, J. (1988) *Numerical Solution of Variational Inequalities.* Springer Verlag, Berlin-Heidelberg-New York.

IOFFE, A.D., and TIKHOMIROV, V.M. (1975) *Theory of Extremal Problems.* North Holland Publ. Co., Amsterdam.

IUSEM, A. (1995) On some properties of generalized proximal point methods for quadratic and linear programming. *JOTA*, **85**, 593–612.

IUSEM, A. (1998) On some properties of paramonotone operators. *J. of Conv. Analysis*, **5**, 269–278.

KAPLAN, A. and TICHATSCHKE, R. (2002) Convergence analysis of non-quadratic proximal methods for variational inequalities in Hilbert spaces. *J. of*

KAPLAN, A. and TICHATSCHKE, R. (2000a) Auxiliary problem principle and the approximation of variational inequalities with non-symmetric multi-valued operators. *CMS Conference Proc.*, **27**, 185–209.

KAPLAN, A. and TICHATSCHKE, R. (2000b) Proximal point approach and approximation of variational inequalities. *SIAM J. Control Optim.*, **39**, 1136–1159.

KAPLAN, A. and TICHATSCHKE, R. (1997) Prox-regularization and solution of ill-posed elliptic variational inequalities. *Applications of Mathematics*, **42**, 111–145.

KAPLAN, A. and TICHATSCHKE, R. (1994) *Stable Methods for Ill-Posed Variational Problems — Prox-Regularization of Elliptical Variational Inequalities and Semi-Infinite Optimization Problems.* Akademie Verlag, Berlin.

KIWIEL, K. (1999) A projection-proximal bundle method for convex nondifferentiable minimization. In: Théra, M. and Tichatschke, R., eds., *Ill-posed Variational Problems and Regularization Techniques*, LNEMS, **477**, 137–150. Springer.

LIONS, J.L. (1969) *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires.* Dunod, Gauthier-Villars, Paris.

MARTINET, B. (1970) Régularisation d'inéquations variationelles par approximations successives. *RIRO*, **4**, 154–159.

POLYAK, B.T. (1987) *Introduction to Optimization.* Optimization Software, Inc. Publ. Division, New York.

POLYAK, R. and TEBOULLE, M. (1997) Nonlinear rescaling and proximal-like methods in convex optimization. *Math. Programming*, **76**, 265–284.

RENAUD, A. and COHEN, G. (1997) An extension of the auxiliary problem principle to nonsymmetric auxiliary operators. *ESAIM: Control, Optimization and Calculus of Variations*, **2**, 281–306.

ROCKAFELLAR, R.T. (1970) On the maximality of sums of nonlinear monotone operators. *Trans. Amer. Math. Soc.*, **149**, 75–88.

ROCKAFELLAR, R.T. (1976) Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.*, **14**, 877–898.

ROTIN, S. (1999) Convergence of the proximal-point method for ill-posed control problems with partial differential equations. PhD Thesis, University of Trier.

SALMON, G., NGUYEN, V.H., and STRODIOT, J.J. (2000) A perturbed and inexact version of the auxiliary problem method for solving general variational inequalities with a multivalued operator. In: Nguyen, V. H., Strodiot, J. J. and Tossings, P., eds., *Optimization*, LNEMS, **481**, 396–418. Springer.

SCHMITT, H. (1997) On the regularized Bingham problem. In: Gritzmann, P., Horst, R., Sachs, E. and Tichatschke, R., eds., *Recent Advances in Optimization*, LNEMS, **452**, 298–314. Springer.

TEBOULLE, M. (1992) Entropic proximal mappings with application to non-

Tseng, P. and Bertsekas, D. (1993) On the convergence of the exponential multiplier method for convex programming. *Math. Programming*, **60**, 1–19.