Journal of Nonlinear and Convex Analysis
Volume 2, Number 3, 2001, 305–332

Yokohama Publishers
ISSN 1880-5221 ONLINE JOURNAL

A GENERAL VIEW ON PROXIMAL POINT METHODS TO VARIATIONAL INEQUALITIES IN HILBERT SPACES - ITERATIVE REGULARIZATION AND APPROXIMATION

A. KAPLAN AND R. TICHATSCHKE

Abstract. A general approach for analyzing convergence of proximal-like methods for variational inequalities with set-valued maximal monotone operators is developed. This approach is oriented to methods coupling successive approximation of the variational inequality with the proximal point algorithm as well as to related methods using regularization on a subspace and/or weak regularization. The convergence results are proved under mild assumptions with respect to the original variational inequality and admit, in particular, the use of the ϵ-enlargement of an operator. Also conditions providing linear and superlinear convergence are established. As an application, the proximal-based variant of the elliptic regularization method is considered.

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'. We consider the variational inequality

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

where $K \subset X$ is a convex closed set and $\mathcal{Q} : X \to 2^{X'}$ is a maximal monotone operator.

The proximal point method, originally introduced by Martinet [48] to solve convex variational problems and later on investigated in a more general setting by Rockafellar [57], has initiated a lot of new algorithms for solving various classes of variational inequalities and related problems.
One can observe some main directions in the development of this technique:

- Modifications of the standard methods for convex optimization to provide a more qualified convergence of a generated minimizing sequence and a better stability of the auxiliary problems (cf. [2], [3], [5], [30], [50], [56], [63]);
- Decomposition and splitting methods for variational inequalities (cf. [17], [21], [24], [29], [64] - [66]);
- Stable successive approximation and/or discretization of ill-posed variational inequalities, especially problems in mathematical physics, as well as

2000 *Mathematics Subject Classification.* 65J20, 65K10, 90C25, 90C48.
Key words and phrases. Variational inequalities, monotone operators, proximal point methods, regularization.

generalizations of the principle of iterative regularization (multi-step proximal regularization, regularization on particular subspaces, weak regularization) (cf. [1], [30], [31], [33], [34], [41], [51], [58], [59], [60]);
- Proximal bundle methods for non-smooth convex optimization problems and variational inequalities (cf. [10], [36], [37], [42], [43], [49]);
- Methods with non-quadratic proximal regularization based on the use of Bregman functions (cf. [6], [15], [16], [19], [38]) or logarithmic-quadratic kernels (cf. [7] - [9], [54], [67]);
- Hybrid algorithms on the basis of proximal methods (cf. [39], [40], [61], [62]).

The literature on this subject is vast, and we are able to mention only a small part of publications.

The basic results of Rockafellar [57] on convergence of the proximal point method for solving variational inequalities with maximal monotone operators were generalized in [47] concerning the rate of convergence, and in [26] a similar analysis was performed for methods using the proximal regularization on a subspace. More precisely, in these papers the methods were studied for the equivalent problem of finding a zero of a maximal monotone operator and under the assumption that the proximal iterations are performed inexactly. However, an approximation of the problem data was not considered.

The general scheme for the convergence analysis of proximal methods, including a successive approximation of the variational inequality, has been developed in [33] and [34]. It concerns iterative regularization methods, when an approximation of K and $\mathcal{Q}$ is usually improved after each proximal step, as well as multi-step regularization methods, in which proximal iterations for each approximated problem are repeated as long as they remain "efficient" according to a special criterion (see, for instance, [33]). This scheme covers not only methods using the classical proximal mapping but also proximal-like methods with regularization on a subspace and regularization in a weaker norm [30]. However, a certain "unconventional" information about the variational inequality is needed, in particular, an upper bound r for the norm of some solution, and bounds for the $\mathcal{Q}$-image of the set $K \cap S_r,\ S_r = \{x \in X : \|x\| \leq r\}$.

The goal of the present paper is a uniform analysis of (proximal-like) iterative regularization methods under mild assumptions w.r.t. the original problem. This analysis is performed in a general algorithmic framework, called here as the *generalized proximal point method* (GPP-method). The main convergence result for this method (cf. Theorem 1) does not use any hard available information about Problem (P), see Assumption 1 in Section 2. At the same time, in comparison with Theorem 1 in [33] and Theorem 3.7 in [34], the conditions on the regularizing functional are weakened, in particular, it does not need to be quadratic. Also the requirements on the exactness of the data approximation are less restrictive. Moreover, an ϵ-enlargement of the operator $\mathcal{Q}$, with $\epsilon \to 0$, can be used as an approximation of $\mathcal{Q}$, whereas in [33], [34] the approximating operators are supposed to be maximal monotone.

We do not consider here an adaptation of the general framework to special problems and algorithms: this will be the object of a forthcoming paper. The conditions w.r.t. data approximation are mainly aimed at standard discretization techniques for variational inequalities in mathematical physics and convex semi-infinite programs, and therefore, the problems of elasticity theory, fluid mechanics and control problems with PDE present an appropriate field for such an adaptation (see [27], [31], [32], [58], [60]).

The paper is organized as follows. In Section 2 the GPP-method is described and general assumptions concerning Problem (P) and the data approximation are discussed. Section 3 contains some preliminary results and the proof of convergence of the method. Conditions providing linear and superlinear convergence as well as a sort of "finite " convergence (see Remark 5) are established in Section 4. An extension of the main results, which admits the use of the ϵ-enlargement of the operator $\mathcal{Q}$, is reported in Section 5. In Appendix 1 we analyze the choice of a regularizing functional which leads to a proximal-based modification of the elliptic regularization method, and Appendix 2 contains an example showing how some conditions w.r.t. successive approximation of Problem (P) can be carried out.

2. Generalized proximal point method

We make use of the following basic assumption concerning Problem (P).

Assumption 1. (i) *$D(\mathcal{Q}) \cap K$ is a non-empty convex set and the operator*

$$\mathcal{Q}_K : y \to \begin{cases} \mathcal{Q}(y) & if\ y \in K \\ \emptyset & otherwise \end{cases}$$

is locally hemi-bounded at each point of $D(\mathcal{Q}) \cap K$;

(ii) *the operator $\mathcal{Q} + \mathcal{N}_K$ is maximal monotone, where*

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

is the normality operator for K;

(iii) *Problem (P) is solvable.*

Assumptions 1(i), (ii) are, in fact, mild conditions which ensure the implication:

$$\text{if } x^* \in K \text{ and } \forall x \in K,\ \exists\, q \in \mathcal{Q}(x) : \ \langle q, x - x^*\rangle \geq 0, \quad \text{then } x^* \text{ solves } (P)$$

(see Lemma 1 below). This implication is essential for the theoretical and numerical analysis of variational inequalities.

The studied solution scheme includes successive approximations of K by a family $\{K^k\}$ of convex closed sets and of $\mathcal{Q}$ by a family $\{\mathcal{Q}^k\}$ of monotone operators.

Let $h : X \to \bar{I\!R} \equiv I\!R \cup \{+\infty\}$ be a convex functional, Gàteaux-differentiable on the set $\hat{K} \supset K \cup (\cup_{k=1}^{\infty} K^k)$, $\hat{K} \subset X$.

The choice of $\{K^k\}, \{\mathcal{Q}^k\}$ and h will be specified in Assumption 2 below.

Using controlling sequences $\{\delta_k\}$ and $\{\chi_k\}$, such that

$$\delta_k \geq 0, \lim_{k\to\infty} \delta_k = 0 \text{ and } 0 < \chi_k \leq \bar{\chi} < \infty, \tag{2.1}$$

the following generalized proximal point method is considered.

GPP-method: *Let $x^1 \in \hat{K}$ be arbitrarily chosen and x^k be known, solve*

$$(\mathbf{P^k}) \quad find \ \ x^{k+1} \in K^k, \ q^k(x^{k+1}) \in \mathcal{Q}^k(x^{k+1}) :$$

$$\langle q^k(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1}\rangle \geq -\delta_k\|x - x^{k+1}\| \quad \forall x \in K^k. \tag{2.2}$$

Remark 1. *All the results of this paper remain true if we replace inequality (2.2) in Problem (P^k) by the inclusion*

$$e^k + q^k(x^{k+1}) + \chi_k \Big(\nabla h(x^{k+1}) - \nabla h(x^k)\Big) \in -\mathcal{N}_{K^k}(x^{k+1})$$

with $e^k \in X'$, $\|e^k\|_{X'} \leq \delta_k$.
Indeed, the definition of the normality operator yields

$$\langle e^k + q^k(x^{k+1}) + \chi_k \Big(\nabla h(x^{k+1}) - \nabla h(x^k)\Big), x - x^{k+1}\rangle \geq 0 \ \forall x \in K^k$$

and regarding $\|e^k\|_{X'} \leq \delta_k$, (2.2) follows immediately.
This inclusion (in case $h(x) = \frac{1}{2}\|x\|^2$, $K^k \equiv K$, $\mathcal{Q}^k \equiv \mathcal{Q}$) was used in a series of papers starting with [57]. *For an analysis of different criteria in proximal-like methods we refer to* [20].

In the sequel we employ the following notations: X^* is the solution set of Problem (P); for any $x \in X^*$ the set $\hat{\Lambda}(x) = \{q \in \mathcal{Q}(x) : \langle q, y - x\rangle \geq 0 \ \forall y \in K\}$ is defined; $q^*(x)$ is an element of $\hat{\Lambda}(x)$; with any u, C, D from X (resp., from X') the distances

$$dist(u, D) = \inf_{w\in D} \|u - w\|, \ dist(C, D) = \sup_{u\in C} dist(u, D)$$

(resp., $dist_{X'}(u, D) = \inf_{w\in D} \|u - w\|_{X'}$) are used; the symbol "$\rightharpoonup$" denotes weak convergence in X or X'.

Assumption 2. (i) *h is a convex functional on X and the mapping ∇h is Lipschitz continuous on $\hat{K}$ with a Lipschitz constant l_h;*

(ii) *for each k it holds $K^k \cap D(\mathcal{Q}^k) \neq \emptyset$ and*

$$\langle q^k(x) - q^k(y), x - y\rangle \geq \langle \mathcal{B}(x - y), x - y\rangle$$
$$\forall x, y \in K^k \cap D(\mathcal{Q}^k), \ \forall q^k(\cdot) \in \mathcal{Q}^k(\cdot),$$

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

(iii) *with given constants $\tilde{\chi} > 0$, $m > 0$, the inequality*

$$\frac{1}{2}\tilde{\chi}\langle \mathcal{B}(x - y), x - y\rangle + h(x) - h(y) - \langle \nabla h(y), x - y\rangle \geq m\|x - y\|^2$$

is valid for all $x, y \in \hat{K}$, and $\bar{\chi}$ is chosen in (2.1) such that $2\bar{\chi}\tilde{\chi} < 1$;

(iv) *for all k, the operators $\mathcal{Q}^k + \mathcal{N}_{K^k} + \chi_k\nabla h$ are maximal monotone;*

(v) *for each $w \in D(\mathcal{Q}) \cap K$, there exist a sequence $\{w^k\}$, $w^k \in D(\mathcal{Q}^k) \cap K^k$, and a compact set $\Lambda(w) \subset \mathcal{Q}(w)$ such that*

$$w^k \rightharpoonup w \text{ as } k \to \infty, \quad \lim_{k\to\infty} \inf_{\zeta \in \mathcal{Q}^k(w^k)} dist_{X'}(\zeta, \Lambda(w)) = 0$$

(in general, $\Lambda(w)$ depends on $\{w^k\}$);

(vi) *with given non-negative constants c_1, c_2 and sequences $\{\varphi_k\}, \{\sigma_k\}$, satisfying*

$$\sum_{k=1}^{\infty} \frac{\varphi_k}{\chi_k} < \infty, \quad \sum_{k=1}^{\infty} \frac{\sigma_k}{\chi_k} < \infty, \tag{2.3}$$

for some solution x^* of Problem (P) there exist sequences $\{w^k\}$, $w^k \in D(\mathcal{Q}^k) \cap K^k$, and $\{q^k(w^k)\}$, $q^k(w^k) \in \mathcal{Q}^k(w^k)$, and an element $q^*(x^*) \in \hat{\Lambda}(x^*)$ such that

$$\|w^k - x^*\| \le c_1 \varphi_k, \quad \|q^k(w^k) - q^*(x^*)\|_{X'} \le c_2 \sigma_k \tag{2.4}$$

holds for sufficiently large k;

(vii) *with $x^*, q^*(x^*)$ and $\{\varphi_k\}$ as in (2-vi) and some constant $c_3 \ge 0$, for any sequence $\{v^k\}, v^k \in K^k \cap D(\mathcal{Q}^k)$, there exists a sequence $\{z^k(v^k)\} \subset K$ such that the relation*

$$\langle q^*(x^*), z^k(v^k) - v^k \rangle \le c_3(\|v^k - x^*\|^2 + 1)\varphi_k$$

is valid;

(viii) *each weak limit point of an arbitrary sequence $\{v^k\}$, $v^k \in K^k \cap D(\mathcal{Q}^k)$, belongs to $K \cap D(\mathcal{Q})$.*

Referring to the separate conditions in Assumptions 1, 2 etc., we write (1-*i*), (1-*ii*), ... and (2-*i*), (2-*ii*), ..., respectively.

Let us discuss some notions and conditions in Assumptions 1 and 2.

- Local hemi-boundedness of an operator $\mathcal{M}$ at a point x^0 means: for each $x,\ x \ne x^0$, there exists a number $t_0(x^0, x) > 0$ such that the set

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

  Throughout this paper we use a weakened notion of local hemi-boundedness: the standard notion supposes boundedness of

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

  The simple example $\mathcal{M} = \mathcal{N}_C$, where $C = \{x \in I\!R^2 : \ x_1^2 + x_2^2 \le 1\}$, shows that this relaxation may be very essential.
- Due to the monotonicity of $\mathcal{Q}^k$, Assumption (2-*ii*) is evidently fulfilled with $\mathcal{B} = \mathbf{0}$ if $K^k \cap D(\mathcal{Q}^k) \ne \emptyset\ \forall k$. In case $\mathcal{B} = \mathbf{0}$, according to (2-*iii*) the functional h has to be strongly convex, which corresponds to the classical proximal point method. But, if the operators $\mathcal{Q}^k$ possess a certain "reserve of monotonicity" (for instance, $\mathcal{Q}^k$ is uniformly strong monotone on a subspace

of X or on a space X with a weaker norm[1]), then the choice of an appropriate operator $\mathcal{B}$ allows to weaken the mentioned requirement for h.
On this way, proximal methods with weak regularization and regularization on a subspace have been developed for elliptic variational inequalities in [31], [60] and for optimal control problems with PDE's in [27], [32], [58].

- With $\mathcal{Q}$ a maximal monotone operator and K a convex closed set, the operator $\mathcal{Q} + \mathcal{N}_K$ is maximal monotone if, for instance, $intD(\mathcal{Q}) \cap K \neq \emptyset$ or, equivalently, $\mathcal{Q}$ is locally bounded[2] at some $x \in K \cap clD(\mathcal{Q})$ (see [55]). Under (2-*i*), Assumption (2-*iv*) is valid if each operator $\mathcal{Q}^k$ is maximal monotone and locally bounded at some $x \in K^k \cap clD(\mathcal{Q}^k)$ or (regarding also (2-*ii*)) if $D(\mathcal{Q}^k) \supset K^k$ and $\mathcal{Q}^k$ is hemicontinuous on K^k (this follows from the Theorems 1 and 3 in [55]). Other conditions, ensuring (1-*ii*) and (2-*iv*), can be derived from results about the maximality of the sum of two monotone operators in [55], [4], [12]. Together with (2-*ii*) and (2-*iii*), Assumption (2-*iv*) provides the solvability of the regularized problems (with $\delta_k = 0$).
- Concerning the Assumptions (2-*v*) – (2-*vii*) on the successive data approximation, of course, one can suppose a Hausdorff-type convergence of $\mathcal{Q}^k$, K^k to $\mathcal{Q}$, K, respectively, and in this way the long list of conditions can be shorten. But, from the numerical point of view this would be not realistic. We intend to consider first of all variational inequalities in mathematical physics and a certain approximation technique used for that class of problems.
  The assumption (2-*vii*) is obviously fulfilled if $K^k \subset K$ for all k. In Appendix II an example is analyzed showing how the single-valued operators $\mathcal{Q}^k$, observing (2-*v*), (2-*vi*), can be chosen. In this example $\mathcal{Q}$ is a multi-valued, non-symmetric operator.
- For the convergence results described in Section 3 we don't need to know the values of the constants c_1, c_2. Only existence of these constants is essential. This is very important: In particular, using finite element methods for solving problems in mathematical physics, the calculation of c_1, c_2 requires certain estimates for x^*, whereas for a lot of these problems the existence of c_1, c_2 follows from regularity results of the solutions.

The GPP-method can be considered as a particular case of the *proximal auxiliary problem method* in [35]. However, in [35] the conditions on the successive data approximation ($K^k \subset K$, and $\mathcal{Q}^k \equiv \mathcal{Q}$ or $\mathcal{Q}^k$ is a single-valued operator) are much more restrictive, and there are no results on the rate of convergence.

3. Convergence results

We start the study of convergence of the GPP-method with some preliminary results.

Lemma 1. *Let $C \subset X$ be a convex closed set, the operators $\mathcal{A}_0 : X \to 2^{X'}$, $\mathcal{A}_0 + \mathcal{N}_C$ be maximal monotone and $D(\mathcal{A}_0) \cap C$ be a convex set. Moreover, assume that the*

[1] of course, these properties of $\mathcal{Q}^k$ depend mainly on $\mathcal{Q}$

[2] i.e. $\mathcal{Q}$ maps some neighborhood of x into a bounded set.

operator

$$\mathcal{A}_C : v \to \begin{cases} \mathcal{A}_0(v) & if\ v \in C \\ \emptyset & otherwise \end{cases}$$

is locally hemi-bounded at each point $v \in D(\mathcal{A}_0) \cap C$ and that, for some $u \in D(\mathcal{A}_0) \cap C$ and each $v \in D(\mathcal{A}_0) \cap C$, there exists $\eta(v) \in \mathcal{A}_0(v)$ satisfying

$$\langle \eta(v), v - u \rangle \geq 0. \tag{3.1}$$

Then, with some $\eta \in \mathcal{A}_0(u)$, the inequality

$$\langle \eta, v - u \rangle \geq 0 \tag{3.2}$$

holds for all $v \in C$.

Proof. In view of the maximal monotonicity of $\mathcal{A}_0$ and $\mathcal{A}_0 + \mathcal{N}_C$, the operators $\mathcal{A} : v \to \mathcal{A}_0(v) + \mathcal{I}(v-u)$ and $\mathcal{A}_1 = \mathcal{A} + \mathcal{N}_C$ (with $\mathcal{I} : X \to X'$ the canonical isometry) are also maximal monotone. Moreover, they are strongly monotone. Therefore, there exists $w \in D(\mathcal{A}_0) \cap C$ such that $0 \in \mathcal{A}(w) + \mathcal{N}_C(w)$, and due to the definition of the normality operator, this yields

$$\langle \eta(w), v - w \rangle \geq 0 \quad \forall v \in C, \tag{3.3}$$

with some $\eta(w) \in \mathcal{A}(w)$.
If $w = u$, then, of course, $\eta(w) \in \mathcal{A}_0(w)$, hence, the conclusion of the lemma is valid. Otherwise, we use the relation

$$\langle \bar{\eta}(v), v - u \rangle \geq 0 \quad \forall v \in D(\mathcal{A}_0) \cap C, \tag{3.4}$$

which follows from (3.1) taking $\bar{\eta}(v) = \eta(v) + \mathcal{I}(v - u) \in \mathcal{A}(v)$.
Let $w_\lambda = u + \lambda(w - u)$ for $\lambda \in (0, 1]$. Obviously, $w_\lambda \in D(\mathcal{A}_0) \cap C$, and according to (3.4) there exists $\bar{\eta}(w_\lambda) \in \mathcal{A}(w_\lambda)$ ensuring

$$\langle \bar{\eta}(w_\lambda), w - u \rangle \geq 0.$$

Because the operator $\mathcal{A}_C$ is locally hemi-bounded at u, the set $\{\bar{\eta}(w_\lambda) : \lambda \in (0, \lambda_0]\}$ is bounded in V' for a sufficiently small $\lambda_0 > 0$. Hence, if λ tends to 0 in an appropriate manner, the corresponding sequence $\{\bar{\eta}(w_\lambda)\}$ converges weakly in V' to some $\bar{\eta}$. Taking into account that $\lim_{\lambda \to 0} \|w_\lambda - u\| = 0$ and that $\mathcal{A}$ is maximal monotone, one can conclude that $\bar{\eta} \in \mathcal{A}(u)$ and

$$0 \leq \lim \langle \bar{\eta}(w_\lambda), w - u \rangle = \langle \bar{\eta}, w - u \rangle.$$

Combining this inequality and inequality (3.3) given with $v = u$, we obtain

$$\langle \bar{\eta} - \eta(w), u - w \rangle \leq 0,$$

but that contradicts the strong monotonicity of $\mathcal{A}$. □

Remark 2. *Due to the Assumptions (2-ii), (2-iii), the convexity of K^k and relation (2.1), the operators $\mathcal{Q}^k + \mathcal{N}_{K^k} + \chi_k \nabla h$ are strongly monotone. Moreover, according to (2-iv), they are maximal monotone. Hence, for each k, Problem (P^k) with $\delta_k = 0$ is uniquely solvable.*

With $x^* \in X^*$ as in (2-vi), define

$$\begin{aligned}\Gamma(x^*, x) = &\tilde{\chi}\langle \mathcal{B}(x - x^*), x - x^*\rangle \\ &+ h(x^*) - h(x) - \langle \nabla h(x), x^* - x\rangle.\end{aligned} \tag{3.5}$$

The function Γ plays the role of a Ljapunov function in the further analysis.

Lemma 2. *Let the Assumptions (1-iii), (2-i) - (2-iv), (2-vi), (2-vii) and relation (2.1) be fulfilled. Moreover, let* $\sum_{k=1}^{\infty} \frac{\delta_k}{\chi_k} < \infty$.
Then the sequence $\{x^k\}$*, generated by the GPP-method is bounded,* $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$ *and the sequence* $\{\Gamma(x^*, x^k)\}$ *converges.*

Proof. In the sequel we make use of the following inequalities, which are valid for arbitrary $a, b, x \in X, p \in X'$ and a number $\nu > 0$:

$$\langle p, a\rangle \le \frac{1}{2\nu}\|p\|_{X'}^2 + \frac{\nu}{2}\|a\|^2, \tag{3.6}$$

$$\langle \mathcal{B}(a-b), a-b\rangle \le (1+\nu)\langle \mathcal{B}(a-x), a-x\rangle + \frac{1+\nu}{\nu}\langle \mathcal{B}(b-x), b-x\rangle. \tag{3.7}$$

For a fixed k, assuming that w^k and $q^k(w^k)$ satisfy (2.4) and

$$\varphi_k < 1,\ \sigma_k < 1,\ \delta_k < 1 \text{ and } \frac{c_3\varphi_k + \sigma_k/2 + \delta_k}{\chi_k m} < 1, \tag{3.8}$$

we estimate

$$\Gamma(x^*, x^{k+1}) - \Gamma(x^*, x^k) = s_1 + \chi_k^{-1}s_2 + s_3 + s_4,$$

with

$$\begin{aligned}
s_1 &= h(x^k) - h(x^{k+1}) + \langle \nabla h(x^k), x^{k+1} - x^k\rangle,\\
s_2 &= \chi_k\langle \nabla h(x^k) - \nabla h(x^{k+1}), w^k - x^{k+1}\rangle,\\
s_3 &= \langle \nabla h(x^k) - \nabla h(x^{k+1}), x^* - w^k\rangle\\
s_4 &= \tilde{\chi}\langle \mathcal{B}(x^{k+1} - x^*), x^{k+1} - x^*\rangle - \tilde{\chi}\langle \mathcal{B}(x^k - x^*), x^k - x^*\rangle.
\end{aligned}$$

From the definition of x^{k+1} and the inclusion $w^k \in K^k$, the inequality

$$\langle q^k(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), w^k - x^{k+1}\rangle \ge -\delta_k\|w^k - x^{k+1}\|$$

follows immediately. Together with (2-ii) and (3.7) this yields

$$\begin{aligned}
s_2 &\le \langle q^k(x^{k+1}), w^k - x^{k+1}\rangle + \delta_k\|w^k - x^{k+1}\|\\
&\le \langle q^k(w^k), w^k - x^{k+1}\rangle - \langle \mathcal{B}(w^k - x^{k+1}), w^k - x^{k+1}\rangle + \delta_k\|w^k - x^{k+1}\|\\
&\le \langle q^k(w^k), w^k - x^{k+1}\rangle - \frac{1}{1+\nu}\langle \mathcal{B}(x^{k+1} - x^*), x^{k+1} - x^*\rangle\\
&\quad + \frac{1}{\nu}\langle \mathcal{B}(w^k - x^*), w^k - x^*\rangle + \delta_k\|w^k - x^{k+1}\|,
\end{aligned}$$

with an arbitrary $\nu > 0$. Taking $\nu = \hat{\nu} = (2\tilde{\chi}\bar{\chi})^{-1} - 1$ and $z^{k+1} = z^{k+1}(x^{k+1})$, where $z^{k+1}(x^{k+1})$ corresponds to (2-vii), one can continue:

$$\begin{aligned}
s_2 \le &\langle q^k(w^k) - q^*(x^*), w^k - x^*\rangle + \langle q^k(w^k) - q^*(x^*), x^* - x^{k+1}\rangle\\
&+ \langle q^*(x^*), w^k - x^*\rangle + \langle q^*(x^*), z^{k+1} - x^{k+1}\rangle
\end{aligned}$$

$$+ \langle q^*(x^*), x^* - z^{k+1}\rangle - 2\tilde{\chi}\bar{\chi}\langle \mathcal{B}(x^{k+1} - x^*), x^{k+1} - x^*\rangle$$

$$+ \frac{1}{\hat{\nu}}\langle \mathcal{B}(w^k - x^*), w^k - x^*\rangle + \delta_k\|w^k - x^{k+1}\|. \tag{3.9}$$

But, applying inequality (3.6), we obtain for the third term in (3.9):

$$\langle q^k(w^k) - q^*(x^*), x^* - x^{k+1}\rangle \le \frac{1}{2\theta_k}\|q^k(w^k) - q^*(x^*)\|^2_{X'} + \frac{\theta_k}{2}\|x^* - x^{k+1}\|^2,$$

with an arbitrary $\theta_k > 0$, and for the last term in (3.9):

$$\begin{aligned} \delta_k\|w^k - x^{k+1}\| &\le \delta_k\|w^k - x^*\| + \delta_k\|x^{k+1} - x^*\| \\ &\le \delta_k\|w^k - x^*\| + \delta_k\left[\frac{1}{4} + \|x^* - x^{k+1}\|^2\right]. \end{aligned}$$

By the definition of x^*,

$$\langle q^*(x^*), x^* - z^{k+1}\rangle \le 0$$

is valid, and after simple calculations one gets

$$s_2 \le \hat{s}_2 - 2\tilde{\chi}\bar{\chi}\langle \mathcal{B}(x^{k+1} - x^*), x^{k+1} - x^*\rangle, \tag{3.10}$$

where

$$\begin{aligned} \hat{s}_2 = {}& \|q^k(w^k) - q^*(x^*)\|_{X'}\|w^k - x^*\| + \|q^*(x^*)\|_{X'}\|w^k - x^*\| \\ & + \langle q^*(x^*), z^{k+1} - x^{k+1}\rangle + \frac{\|\mathcal{B}\|}{\hat{\nu}}\|w^k - x^*\|^2 + \delta_k\|w^k - x^*\| \\ & + \frac{1}{2\theta_k}\|q^k(w^k) - q^*(x^*)\|^2_{X'} + \frac{\theta_k}{2}\|x^* - x^{k+1}\|^2 \\ & + \delta_k\left[\frac{1}{4} + \|x^* - x^{k+1}\|^2\right]. \end{aligned} \tag{3.11}$$

Taking into account that $0 < \chi_k \le \bar{\chi}$, from (3.10) we have

$$\begin{aligned} &\chi_k^{-1}s_2 + s_4 \\ &\quad \le \chi_k^{-1}\hat{s}_2 - 2\chi_k^{-1}\tilde{\chi}\bar{\chi}\langle \mathcal{B}(x^{k+1} - x^*), x^{k+1} - x^*\rangle \\ &\qquad + \tilde{\chi}\langle \mathcal{B}(x^{k+1} - x^*), x^{k+1} - x^*\rangle - \tilde{\chi}\langle \mathcal{B}(x^k - x^*), x^k - x^*\rangle \\ &\quad \le \chi_k^{-1}\hat{s}_2 - \tilde{\chi}\left[\langle \mathcal{B}(x^{k+1} - x^*), x^{k+1} - x^*\rangle + \langle \mathcal{B}(x^k - x^*), x^k - x^*\rangle\right] \end{aligned}$$

and (3.7) with $\nu = 1, a = x^{k+1}, b = x^k$ and $x = x^*$ yields

$$\chi_k^{-1}s_2 + s_4 \le \chi_k^{-1}\hat{s}_2 - \frac{\tilde{\chi}}{2}\langle \mathcal{B}(x^{k+1} - x^k), x^{k+1} - x^k\rangle. \tag{3.12}$$

Together with (2-*iii*) this provides

$$s_1 + \chi_k^{-1}s_2 + s_4 \le \chi_k^{-1}\hat{s}_2 - m\|x^{k+1} - x^k\|^2.$$

At the same time, for s_3 we obtain from (2-*i*) and (3.6) (given with $\nu = m^{-1}l_h^2\varphi_k^{-1}$) that

$$s_3 \le \frac{m}{2}\ \varphi_k\|x^{k+1} - x^k\|^2 + \frac{m^{-1}}{2}\ l_h^2\varphi_k^{-1}\|x^* - w^k\|^2, \tag{3.13}$$

hence,

$$
\begin{aligned}
& s_1 + \chi_k^{-1} s_2 + s_3 + s_4 \le \chi_k^{-1} \hat{s}_2 \\
& \qquad + \left(\frac{m}{2}\varphi_k - m\right) \|x^{k+1} - x^k\|^2 + \frac{m^{-1}}{2}\, l_h^2 \varphi_k^{-1} \|x^* - w^k\|^2.
\end{aligned} \tag{3.14}
$$

If $\sigma_k > 0$, then due to (2-vi), (2-vii) and

$$\varphi_k < 1,\ \sigma_k < 1,\ \delta_k < 1,\ 0 < \chi_k \le \bar{\chi},$$

the relations (3.11) (taken with $\theta_k = \sigma_k$) and (3.14) lead to

$$
\begin{aligned}
& s_1 + \chi_k^{-1} s_2 + s_3 + s_4 \le \chi_k^{-1} [d_1 \varphi_k + d_2 \sigma_k + d_3 \delta_k] \\
& \qquad + \chi_k^{-1} [c_3 \varphi_k + \sigma_k/2 + \delta_k] \|x^* - x^{k+1}\|^2 - \frac{m}{2} \|x^{k+1} - x^k\|^2,
\end{aligned} \tag{3.15}
$$

where

$$
\begin{aligned}
d_1 &= c_1 \left[c_2 + \|q^*(x^*)\|_{X'} + \|\mathcal{B}\| \frac{c_1}{\hat{\nu}} + 1 + \frac{l_h^2 \bar{\chi}}{2m} c_1 \right] + c_3, \\
d_2 &= \frac{1}{2} c_2^2, \quad d_3 = c_1 + \frac{1}{4} \quad \text{and} \quad \hat{\nu} = (2\tilde{\chi}\bar{\chi})^{-1} - 1.
\end{aligned}
$$

In case $\sigma_k = 0$, estimate (3.15) is also true because (2.4) ensures that

$$\langle q^k(w^k) - q^*(x^*), x^* - x^{k+1} \rangle = 0.$$

Now, using the inequality

$$\|x^{k+1} - x^*\|^2 \le \frac{1}{m} \Gamma(x^*, x^{k+1}), \tag{3.16}$$

which follows immediately from (2-iii), we derive from (3.15), (3.8) that

$$
\begin{aligned}
& \left(1 - \frac{c_3 \varphi_k + \sigma_k/2 + \delta_k}{\chi_k m}\right) \Gamma(x^*, x^{k+1}) - \Gamma(x^*, x^k) \\
& \qquad \le \frac{d_1 \varphi_k + d_2 \sigma_k + d_3 \delta_k}{\chi_k} - \frac{m}{2} \|x^{k+1} - x^k\|^2,
\end{aligned} \tag{3.17}
$$

hence,

$$
\begin{aligned}
\Gamma(x^*, x^{k+1}) &\le \left(1 - \frac{c_3 \varphi_k + \sigma_k/2 + \delta_k}{\chi_k m}\right)^{-1} \Gamma(x^*, x^k) \\
& \quad + \left(1 - \frac{c_3 \varphi_k + \sigma_k/2 + \delta_k}{\chi_k m}\right)^{-1} \frac{d_1 \varphi_k + d_2 \sigma_k + d_3 \delta_k}{\chi_k}.
\end{aligned}
$$

In view of

$$\left(1 - \frac{c_3 \varphi_k + \sigma_k/2 + \delta_k}{m \chi_k}\right)^{-1} = 1 + \frac{\chi_k^{-1} (c_3 \varphi_k + \sigma_k/2 + \delta_k)}{m - \chi_k^{-1} (c_3 \varphi_k + \sigma_k/2 + \delta_k)}$$

and the conditions (2.1),(2.3) and $\sum_{k=1}^{\infty} \frac{\delta_k}{\chi_k} < \infty$, Lemma 2.2.2 in [53] permits to conclude that the sequence $\{\Gamma(x^*, x^k)\}$ converges. Taking into account (3.16), this provides the boundedness of $\{x^k\}$, and (3.17) yields $\lim_{k \to \infty} \|x^{k+1} - x^k\| = 0$. $\square$

Remark 3. *Let the conditions of Lemma 2 be fulfilled and (2.4),(3.8) be valid for each k. If we are able to estimate $\|x^* - x^1\|$ and the constants d_1, d_2, d_3 and c_3 from above, then there is no problem to calculate from (3.16) and (3.17) a radius r such that*

$$\{x^k\} \subset \{x \in X : \|x - x^*\| \le r\}.$$

The key for such an estimation is: if

$$0 \le t_{i+1} \le (1+\gamma_i)t_i + \beta_i$$

holds for each i and $\gamma_i \ge 0, \beta_i \ge 0$, then a simple induction yields

$$t_i \le (t_1 + \sum_{j=1}^{i-1} \beta_j) \prod_{j=1}^{i-1} (1+\gamma_j).$$

In the particular case that $\mathcal{Q}^k \equiv \mathcal{Q}$ and $K^k \supset K\ \forall k$, one can take $d_1 = c_3,\ d_2 = 0$ and $d_3 = 1/4$.

Lemma 3. *Let the Assumptions (1-i), (1-ii), (2-i), (2-ii), (2-v) and (2-viii) be fulfilled. Moreover, suppose that the sequence $\{x^k\}$ generated by the GPP-method is bounded and $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$.*
Then each weak limit point of $\{x^k\}$ is a solution of Problem (P).

Proof. Let $\bar{x}$ be an arbitrary weak limit point of $\{x^k\}$ and assume that $\{x^k\}_{k\in\mathfrak{K}}$ converges weakly to $\bar{x}$. Due to (2-*viii*), $\bar{x}$ belongs to $K \cap D(\mathcal{Q})$, and

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0 \text{ yields } x^{k+1} \rightharpoonup \bar{x}$$

if $k \in \mathfrak{K}, k \to \infty$.
According to (2-*v*), for each $y \in D(\mathcal{Q}) \cap K$ one can choose the subsequences $\mathfrak{K}' \subset \mathfrak{K}$ and $\{y^k\}_{k\in\mathfrak{K}'}, y^k \in D(\mathcal{Q}^k) \cap K^k$, such that $y^k \rightharpoonup y$ for $k \in \mathfrak{K}', k \to \infty$ and

$$\lim_{k\in\mathfrak{K}'} \|q^k(y^k) - q(y)\|_{X'} = 0 \tag{3.18}$$

holds true with some $q^k(y^k) \in \mathcal{Q}^k(y^k)$ and $q(y) \in \Lambda(y) \subset \mathcal{Q}(y)$ (for $\Lambda(y)$ see (2-*v*)).
By definition of x^{k+1}, for $k \in \mathfrak{K}'$ and suitably chosen $q^k(x^{k+1}) \in \mathcal{Q}^k(x^{k+1})$ the inequality

$$\langle q^k(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), y^k - x^{k+1}\rangle \ge -\delta_k\|y^k - x^{k+1}\|$$

is valid, and the monotonicity of $\mathcal{Q}^k$ yields

$$\langle q^k(y^k) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), y^k - x^{k+1}\rangle \ge -\delta_k\|y^k - x^{k+1}\|. \tag{3.19}$$

From the properties

$$y^k \rightharpoonup y, x^{k+1} \rightharpoonup \bar{x}\ (k \in \mathfrak{K}'),\ \lim_{k\to\infty} \|x^{k+1} - x^k\| = 0,$$

Assumption (2-*i*) and relations (2.1), (3.18), passing to the limit in (3.19) for $k \in \mathfrak{K}'$, we obtain

$$\langle q(y), y - \bar{x}\rangle \ge 0.$$

Together with Assumptions (1-*i*), (1-*ii*) this permits to apply Lemma 1, which ensures that, for some $q(\bar{x}) \in \mathcal{Q}(\bar{x})$, the inequality

$$\langle q(\bar{x}), y - \bar{x}\rangle \ge 0\ \forall y \in K$$

is valid, hence $\bar{x} \in X^*$. □

Theorem 1. *Let the Assumptions 1 and 2 be fulfilled and* $\sum_{k=1}^{\infty} \frac{\delta_k}{\chi_k} < \infty$. *Then the following conclusions are true:*

(i) *Problem* (P^k) *is solvable for each* k, *the sequence* $\{x^k\}$ *generated by the GPP-method is bounded, and each weak limit point of* $\{x^k\}$ *is a solution of Problem* (P);

(ii) *if, in addition, (2-vi), (2-vii) are valid for each* $x \in X^*$ (c_1, c_2, c_3 *and* $\{z^k(v^k)\}$ *may depend on* x) *and*

$$p^k \rightharpoonup p \text{ in } X,\ p^k \in K^k \implies \nabla h(p^k) \rightharpoonup \nabla h(p) \text{ in } X', \tag{3.20}$$

then the whole sequence $\{x^k\}$ *converges weakly to a solution* x^* *of Problem* (P);

(iii) *if, moreover,*

$$\lim_{k\to\infty} h(x^k) = h(x^*) \tag{3.21}$$

holds with x^* *as in (ii), then* $\{x^k\}$ *converges strongly to* x^*.

Proof. Conclusion (i) follows immediately from the Lemmata 2, 3 and Remark 1. To prove (ii), suppose that $\{x^k\}_{k\in\mathfrak{K}_1}$, $\{x^k\}_{k\in\mathfrak{K}_2}$ are two subsequences converging weakly to $\bar{x}$, $\tilde{x}$, respectively. Then, according to (i), $\bar{x}$, $\tilde{x}$ belong to X^*, and because (2-vi), (2-vii) are valid for each $x \in X^*$, Lemma 2 ensures that the sequences $\{\Gamma(\bar{x}, x^k)\}_{k\in I\!N}$, $\{\Gamma(\tilde{x}, x^k)\}_{k\in I\!N}$ are convergent.
Due to $\bar{x} \in X^*$, the symmetry of the operator $\mathcal{B}$ (Assumption (2-ii)) and (2-iii), we obtain for $x \in \hat{K}$

$$\begin{aligned}
&\Gamma(\bar{x}, x) - \Gamma(\tilde{x}, x)\\
&\quad = (h(\bar{x}) - h(\tilde{x}) - \langle \nabla h(\tilde{x}), \bar{x} - \tilde{x}\rangle) + \langle \nabla h(\tilde{x}) - \nabla h(x), \bar{x} - \tilde{x}\rangle\\
&\qquad + \tilde{\chi}\langle \mathcal{B}(\bar{x} - \tilde{x}), \bar{x} - \tilde{x}\rangle + 2\tilde{\chi}\langle \mathcal{B}(\bar{x} - \tilde{x}), \tilde{x} - x\rangle\\
&\quad \geq m\|\bar{x} - \tilde{x}\|^2 + \langle \nabla h(\tilde{x}) - \nabla h(x), \bar{x} - \tilde{x}\rangle\\
&\qquad + 2\tilde{\chi}\langle \mathcal{B}(\bar{x} - \tilde{x}), \tilde{x} - x\rangle.
\end{aligned} \tag{3.22}$$

Inserting $x = x^k$ in (3.22) and passing to the limit for $k \in \mathfrak{K}_2$, one can infer from (3.20) that

$$\bar{\gamma} - \tilde{\gamma} \geq m\|\bar{x} - \tilde{x}\|^2,$$

where $\bar{\gamma} = \lim_{k\to\infty} \Gamma(\bar{x}, x^k)$, $\tilde{\gamma} = \lim_{k\to\infty} \Gamma(\tilde{x}, x^k)$.
Obviously, in the same way the "symmetric" inequality

$$\tilde{\gamma} - \bar{\gamma} \geq m\|\bar{x} - \tilde{x}\|^2$$

can be concluded, and therefore $\bar{x} = \tilde{x}$ is valid, proving the uniqueness of the weak limit point for $\{x^k\}$.
Denote this limit point by x^*. With $\{w^k\}, \{q^k(w^k)\}$ and $q^*(x^*)$ chosen according to (2-vi), and $q^k(x^{k+1})$ as in Problem (P^k), Assumption (2-ii) yields

$$\begin{aligned}
&\langle \mathcal{B}(x^{k+1} - x^*), x^{k+1} - x^*\rangle\\
&\quad = \langle \mathcal{B}(x^{k+1} - w^k), x^{k+1} - w^k\rangle - \langle \mathcal{B}(x^* - w^k), x^{k+1} - x^*\rangle
\end{aligned}$$

$$
\begin{aligned}
&- \langle \mathcal{B}(x^{k+1} - w^k), x^* - w^k\rangle \\
&\le \langle q^k(x^{k+1}) - q^k(w^k), x^{k+1} - w^k\rangle - \langle \mathcal{B}(x^{k+1} - x^*), x^* - w^k\rangle \\
&- \langle \mathcal{B}(x^{k+1} - w^k), x^* - w^k\rangle.
\end{aligned} \tag{3.23}
$$

To estimate the term $\langle q^k(x^{k+1}), x^{k+1} - w^k\rangle$, we use Problem (P^k). Together with (3.23) this gives

$$
\begin{aligned}
&\langle \mathcal{B}(x^{k+1} - x^*), x^{k+1} - x^*\rangle \\
&\quad \le \langle q^k(w^k) - q^*(x^*), w^k - x^{k+1}\rangle \\
&\quad + \langle q^*(x^*), w^k - x^*\rangle + \langle q^*(x^*), x^* - x^{k+1}\rangle \\
&\quad + \chi_k\langle \nabla h(x^{k+1}) - \nabla h(x^k), w^k - x^{k+1}\rangle \\
&\quad + \langle \mathcal{B}x^* + \mathcal{B}w^k - 2\mathcal{B}x^{k+1}, x^* - w^k\rangle + \delta_k\|w^k - x^{k+1}\|.
\end{aligned} \tag{3.24}
$$

Now, due to (2.1), (2-i), (2-ii) and (2-vi) and taking into account that $x^k \rightharpoonup x^*$, $\|x^k - x^{k+1}\| \to 0$, the relation

$$
\lim_{k\to\infty} \langle \mathcal{B}(x^{k+1} - x^*), x^{k+1} - x^*\rangle = 0 \tag{3.25}
$$

can be easily deduced from (3.24). If, additionally, (3.21) is valid, then using (2-iii) with $x = x^{k+1}, y = x^*$, the relations (3.21) and (3.25) permit to conclude that $\lim_{k\to\infty} \|x^k - x^*\| = 0$. □

From (3.25) it follows immediately that the strong convergence of x^k to x^* is also guaranteed if (3.21) is replaced by the condition that there exists a symmetric linear compact operator $\mathcal{B}^1 : X \to X'$ such that the sum $\mathcal{B} + \mathcal{B}^1$ is positive definite.

Remark 4. *Theorem 1 remains true if we replace (3.20) by the following weaker assumption:*

$$
p^k \subset K^k, \quad r^k \subset K^k, \quad p^k \rightharpoonup p, \quad r^k \rightharpoonup r, \quad p \neq r
$$

imply

$$
\underline{\lim}_{k\to\infty} \left| \langle \nabla h(p^k) + 2\tilde{\chi}\mathcal{B}p^k - \nabla h(r^k) - 2\tilde{\chi}\mathcal{B}r^k, p - r\rangle \right| > 0.
$$

Now, let $\mathcal{Q}^0 : X \to X'$ be a single-valued monotone operator and $f : X \to \bar{I\!R}$ be a proper convex lsc functional. The suggested scheme can be directly extended to the variational inequality

$$
\begin{aligned}
(\tilde{\mathbf{P}}) \quad & find \quad x^* \in X \text{ such that} \\
& \langle \mathcal{Q}^0(x^*), x - x^*\rangle + f(x) - f(x^*) \ge 0 \quad \forall x \in X
\end{aligned}
$$

by means of the following reformulation of $(\tilde{P})$:

$$
\begin{aligned}
& find \quad x^* \in dom f \text{ such that} \\
& \quad \exists\, g \in \partial f(x^*) : \ \langle \mathcal{Q}^0(x^*) + g, x - x^*\rangle \ge 0 \ \forall x \in dom f.
\end{aligned}
$$

However, assuming that $\mathcal{Q}^0$ is Lipschitz-continuous on $D(\mathcal{Q}^0)$, one can deduce more convenient requirements for the data approximation if we deal with the original form of Problem $(\tilde{P})$ and modify the GGP-method as follows:

$$
(\tilde{\mathbf{P}}^{\mathbf{k}}) \quad find \quad x^{k+1} \in X \text{ such that}
$$

$$\langle \mathcal{Q}^0(x^{k+1}) + \chi_k \left(\nabla h(x^{k+1}) - \nabla h(x^k)\right), x - x^{k+1}\rangle$$
$$+ f^k(x) - f^k(x^{k+1}) \geq -\delta_k \|x - x^{k+1}\| \qquad \forall x \in X$$

(here f^k is a convex lsc approximation of f).
In that case the statement of Theorem 1 remains true (with minor alterations in the proofs), replacing (2-*ii*), (2-*v*) and (2-*vi*), correspondingly, by:

- (2-*ii*)*: for each k and the operator $\mathcal{B}$ as in (2-*ii*) it holds $D(\mathcal{Q}^0) \cap D(\partial f^k) \neq \emptyset$ and
$$\langle \mathcal{Q}^0(x) - \mathcal{Q}^0(y), x - y\rangle$$
$$+ f^k(x) - f^k(y) - \langle g^k(y), x - y\rangle \geq \langle \mathcal{B}(x-y), x - y\rangle$$
$$\forall x, y \in D(\mathcal{Q}^0) \cap D(\partial f^k),\ \forall g^k(y) \in \partial f^k(y);$$
- (2-*v*)*: $f^k \geq f$, and for each $w \in D(\mathcal{Q}^0) \cap dom f$ there exists a sequence $\{w^k\}$, $w^k \in D(\mathcal{Q}^0) \cap dom f^k$ such that
$$\lim_{k\to\infty} \|w^k - w\| = 0 \quad \text{and} \quad \lim_{k\to\infty} f^k(w^k) = f(w);$$
- (2-*vi*)*: with given positive constants c_1, c_2 and non-negative sequences $\{\varphi_k\}, \{\sigma_k\}$ satisfying (2.3), for some solution x^* of Problem $(\tilde{P})$ there exists a sequence $\{w^k\}$, $w^k \in D(\mathcal{Q}^0) \cap dom f^k$ such that
$$\|w^k - x^*\| \leq c_1 \varphi_k \quad \text{and} \quad f^k(w^k) - f(x^*) \leq c_2 \sigma_k.$$
- Moreover, the Assumptions (2-*vii*) and (2-*viii*) are skipped in this case.

4. Rate of convergence

Now, we consider the partial case of GPP-method with $h(x) = \frac{1}{2}\|x\|^2$. This corresponds to the usual proximal point method coupled with a successive approximation of the operator $\mathcal{Q}$ and the set K. Taking in that case $\mathcal{B} = \mathbf{0}$, Assumptions (2-*ii*) and (2-*iii*) are automatically fulfilled with $m = 1/2$ and with an arbitrary $\tilde{\chi} \in (0, 1/2\bar{\chi})$.

Throughout this section we suppose that the Assumptions 1 and 2 in Section 2 hold true, $\sum_{k=1}^{\infty} \frac{\delta_k}{\chi_k} < \infty$ and that $x^1 \in \hat{K}$ is fixed. Then, according to Lemma 2, the sequence $\{x^k\}$ generated by the GPP-method is bounded. Let r be chosen such that $\{x^k\} \subset S_r$. From statement (i) of Theorem 1 and the weak closedness of S_r it follows that $X^* \cap S_r \neq \emptyset$.

Denote $G = D(\mathcal{Q}) \cap K \cap S_r$, $\ G^k = D(\mathcal{Q}^k) \cap K^k \cap S_r$.

Assumption 3. (i) $\sup_{x\in G}\ \sup_{q\in\mathcal{Q}(x)} \|q\|_{X'} \leq c_4 < \infty$;

(ii) *given the constants c_5 and c_6, for each triplet $k, y^k \in G^k, q^k \in \mathcal{Q}^k(y^k)$ there exist $v^k \in G$ and $q(v^k) \in \mathcal{Q}(v^k)$ such that*

$$\|y^k - v^k\| \leq c_5 \varphi_k, \quad \|q^k - q(v^k)\| \leq c_6 \sigma_k, \tag{4.1}$$

with φ_k, σ_k as in (2-vi);

(iii) *for each k and each $x^* \in X^* \cap S_r$, there exists $w^k \in D(\mathcal{Q}^k) \cap K^k$ such that*

$$\|w^k - x^*\| \leq c_7 \varphi_k. \tag{4.2}$$

Assumption (3-*ii*) is evidently fulfilled with $c_5 = c_6 = 0$ in the case that $\mathcal{Q}^k \equiv \mathcal{Q}$ and $K^k \subset K\ \forall k$. If, moreover, $K^k \equiv K\ \forall k$, then (3-*i*), (3-*iii*) are superfluous for the further consideration.

Let $\delta \in (0, 2r)$ and $l > 0$ be fixed, and denote

$$X_\delta = \{x \in G : dist(x, X^* \cap S_r) \leq \delta\}.$$

Assumption 4. *For each $x \in X_\delta$ and each $q \in \mathcal{Q}(x)$ the inequality*

$$\inf_{v \in X^* \cap S_r} \langle q, x - v\rangle \geq d_0 \|x - x^*(x)\|^l \tag{4.3}$$

holds, with $x^*(x) = \arg\min_{w \in X^* \cap S_r} \|x - w\|$ and d_0 a positive constant.

Lemma 4. *Let Assumption 4 be fulfilled. Then, for each $x \in G$ and each $q \in \mathcal{Q}(x)$ the inequality*

$$\inf_{v \in X^* \cap S_r} \langle q, x - v\rangle \geq d \|x - x^*(x)\|^l \tag{4.4}$$

is true with $d = \left(\frac{\delta}{2r}\right)^l d_0$.

Proof. Consider the nontrivial case that $G \backslash X_\delta \neq \emptyset$, and let $x \in G\backslash X_\delta, v \in X^* \cap S_r$ be chosen arbitrarily. Due to the convexity of $D(\mathcal{Q}) \cap K$, the set $\{\lambda x + (1-\lambda)v : 0 < \lambda < 1\}$ belongs to $D(\mathcal{Q}) \cap K \cap int S_r$. Thus, there exists $\bar{\lambda} = \lambda(x, v) \in (0, 1)$ such that $\tilde{x} = \bar{\lambda} x + (1 - \bar{\lambda}) v \in \partial X_\delta$ (∂X_δ is the boundary of X_δ). Obviously, $\tilde{x} - v = \bar{\lambda}(x - v)$, $\frac{1-\bar{\lambda}}{\bar{\lambda}}(\tilde{x} - v) = x - \tilde{x}$, and in view of the monotonicity of $\mathcal{Q}$, we obtain

$$\frac{1 - \bar{\lambda}}{\bar{\lambda}} \langle q(x) - q(\tilde{x}), \tilde{x} - v\rangle = \langle q(x) - q(\tilde{x}), x - \tilde{x}\rangle \geq 0$$

for any $q(x) \in \mathcal{Q}(x), q(\tilde{x}) \in \mathcal{Q}(\tilde{x})$.
Hence, Assumption 4 yields

$$\langle q(x), \tilde{x} - v\rangle \geq \langle q(\tilde{x}), \tilde{x} - v\rangle \geq d_0 \|\tilde{x} - x^*(\tilde{x})\|^l,$$

and

$$\langle q(x), x - v\rangle \geq \frac{d_0}{\bar{\lambda}} \|\tilde{x} - x^*(\tilde{x})\|^l > d_0 \|\tilde{x} - x^*(\tilde{x})\|^l.$$

But $\|\tilde{x} - x^*(\tilde{x})\| = \delta$ and $\|x - x^*(x)\| \leq 2r$, therefore

$$\langle q(x), x - v\rangle > \left(\frac{\delta}{2r}\right)^l d_0 \|x - x^*(x)\|^l.$$

Because v is an arbitrary point in $X^* \cap S_r$, this leads to (4.4). But if $x \in X_\delta$, then (4.4) follows immediately from (4.3) and $\delta < 2r$. □

Let be

$$c_8 \geq 2 \max\{c_4(c_5 + c_7) + 4drc_5 + 2r\bar{\chi}c_7,\ 2r + c_7,\ c_6(2r + c_7)\}$$

and

$$c_9 \geq \frac{2}{d} \max\{c_6(2r + c_7),\ 2r + c_7,\ 2rc_7\bar{\chi} + c_4(c_5 + c_7) + dc_5\},$$

and denote $\rho_k = dist(x^k, X^* \cap S_r)$.

Theorem 2. *Let the Assumptions 1, 2 and 3 be fulfilled and*

$$\varphi_k < 1 \ \forall k, \ \sum_{k=1}^{\infty} \frac{\delta_k}{\chi_k} < \infty. \tag{4.5}$$

(i) *If Assumption 4 with $l = 2$ is valid as well as*

$$c_8 \left(\frac{\varphi_k + \sigma_k + \delta_k}{\chi_k} \right) \leq \left[\left(1 + \frac{2d}{\chi_k} \right)^{1/2} - 1 \right] \prod_{i=1}^{k-1} \left(1 + \frac{2d}{\chi_i} \right)^{-1/2} \rho_1^2 \ \ \forall k, \tag{4.6}$$

then for each k

$$\rho_k^2 \leq \prod_{i=1}^{k-1} \left(1 + \frac{2d}{\chi_i} \right)^{-1/2} \rho_1^2. \tag{4.7}$$

(ii) *If Assumption 4 with $l = 1$ is valid, then there exists k_0 such that*

$$\rho_{k+1} \leq c_9(\varphi_k + \sigma_k + \delta_k) \tag{4.8}$$

holds true for $k \geq k_0$.

Proof. Here we often use, without a special reminder, that certain points belong to S_r. The symbol $(\cdot,\cdot)$ denotes the inner product in X.

For each k, take $v^{k+1} \in G, q(v^{k+1}) \in \mathcal{Q}(v^{k+1})$ and $w^k \in K^k \cap D(\mathcal{Q}^k)$ such that

$$\|x^{k+1} - v^{k+1}\| \leq c_5 \varphi_k, \ \ \|q^k(x^{k+1}) - q(v^{k+1})\| \leq c_6 \sigma_k \tag{4.9}$$

and

$$\|w^k - x^*(x^k)\| \leq c_7 \varphi_k, \tag{4.10}$$

with $q^k(x^{k+1})$ as in Problem (P^k). This is possible due to the Assumptions (3-*ii*) and (3-*iii*).

Using the definition of x^{k+1} and (4.10), we obtain

$$\begin{aligned}
&\|x^{k+1} - x^*(x^k)\|^2 - \|x^k - x^*(x^k)\|^2 \\
&\quad = \|x^{k+1} - w^k\|^2 - \|x^k - w^k\|^2 + 2(w^k - x^*(x^k), x^{k+1} - x^k) \\
&\quad \leq 4rc_7\varphi_k - \|x^{k+1} - x^k\|^2 + 2(x^{k+1} - x^k, x^{k+1} - w^k) \\
&\quad \leq 4rc_7\varphi_k - \|x^{k+1} - x^k\|^2 + \frac{2}{\chi_k} \langle q^k(x^{k+1}), w^k - x^{k+1} \rangle + \frac{2\delta_k}{\chi_k} \|w^k - x^{k+1}\| \\
&\quad \leq -\|x^{k+1} - x^k\|^2 + \frac{2}{\chi_k} \langle q^k(x^{k+1}), w^k - x^{k+1} \rangle + 4rc_7\bar{\chi}\frac{\varphi_k}{\chi_k} + 2(2r + c_7)\frac{\delta_k}{\chi_k}.
\end{aligned} \tag{4.11}$$

But, regarding that $\varphi_k < 1$,

$$\begin{aligned}
&\langle q^k(x^{k+1}), w^k - x^{k+1} \rangle \\
&\quad = \langle q^k(x^{k+1}) - q(v^{k+1}), w^k - x^{k+1} \rangle + \langle q(v^{k+1}), w^k - x^*(x^k) \rangle \\
&\quad\ + \langle q(v^{k+1}), x^*(x^k) - v^{k+1} \rangle + \langle q(v^{k+1}), v^{k+1} - x^{k+1} \rangle
\end{aligned}$$

$$\leq c_6(2r+c_7)\sigma_k + c_4(c_5+c_7)\varphi_k + \langle q(v^{k+1}), x^*(x^k) - v^{k+1}\rangle. \tag{4.12}$$

Due to Assumption 4 and Lemma 4,

$$\langle q(v^{k+1}), x^*(x^k) - v^{k+1}\rangle \leq -d\|v^{k+1} - x^*(v^{k+1})\|^l, \tag{4.13}$$

and together with (4.12) this yields

$$\begin{aligned} &\langle q^k(x^{k+1}), w^k - x^{k+1}\rangle \\ &\quad \leq -d\|v^{k+1} - x^*(v^{k+1})\|^l + c_6(2r+c_7)\sigma_k + c_4(c_5+c_7)\varphi_k. \end{aligned} \tag{4.14}$$

Now, we prove statement (i), taking $l = 2$ in (4.14). In this case, in view of

$$\|v^{k+1} - x^*(v^{k+1})\|^2 \geq \|x^{k+1} - x^*(v^{k+1})\|^2 - 4rc_5\varphi_k,$$

inequality (4.14) gives

$$\begin{aligned} \langle q^k(x^{k+1}), w^k - x^{k+1}\rangle &\leq -d\|x^{k+1} - x^*(v^{k+1})\|^2 \\ &\quad + c_6(2r+c_7)\sigma_k + \left[c_4(c_5+c_7) + 4drc_5\right]\varphi_k. \end{aligned} \tag{4.15}$$

Inserting (4.15) into (4.11), we obtain

$$\begin{aligned} &\|x^{k+1} - x^*(x^k)\|^2 - \|x^k - x^*(x^k)\|^2 \\ &\quad \leq -\frac{2d}{\chi_k}\|x^{k+1} - x^*(v^{k+1})\|^2 + c_8\left(\frac{\varphi_k + \sigma_k + \delta_k}{\chi_k}\right). \end{aligned} \tag{4.16}$$

Hence,

$$\left(1 + \frac{2d}{\chi_k}\right)\rho_{k+1}^2 - \rho_k^2 \leq c_8\left(\frac{\varphi_k + \sigma_k + \delta_k}{\chi_k}\right), \tag{4.17}$$

and a straightforward induction yields estimate (4.7).

To prove statement (ii), which corresponds to $l = 1$ in Assumption 4 and in (4.14), we obtain from (4.14) and the inequality

$$\|v^{k+1} - x^*(v^{k+1})\| \geq \|x^{k+1} - x^*(v^{k+1})\| - c_5\varphi_k$$

that

$$\begin{aligned} &\langle q^k(x^{k+1}), w^k - x^{k+1}\rangle \\ &\quad \leq -d\|x^{k+1} - x^*(v^{k+1})\| + c_6(2r+c_7)\sigma_k + \left[c_4(c_5+c_7) + dc_5\right]\varphi_k, \end{aligned}$$

and together with (4.11) this provides

$$\begin{aligned} &\|x^{k+1} - x^*(x^k)\|^2 - \|x^k - x^*(x^k)\|^2 \\ &\quad \leq -\|x^{k+1} - x^k\|^2 - \frac{2d}{\chi_k}\|x^{k+1} - x^*(v^{k+1})\| + dc_9\left(\frac{\varphi_k + \sigma_k + \delta_k}{\chi_k}\right). \end{aligned} \tag{4.18}$$

Thus,

$$\begin{aligned} &\frac{2d}{\chi_k}\|x^{k+1} - x^*(v^{k+1})\| \\ &\quad \leq \|x^k - x^*(x^{k+1})\|^2 - \|x^{k+1} - x^*(x^{k+1})\|^2 - \|x^{k+1} - x^k\|^2 \\ &\qquad + dc_9\left(\frac{\varphi_k + \sigma_k + \delta_k}{\chi_k}\right) \end{aligned}$$

$$\leq 2\|x^k - x^{k+1}\| \cdot \|x^{k+1} - x^*(x^{k+1})\| + dc_9 \left(\frac{\varphi_k + \sigma_k + \delta_k}{\chi_k} \right). \tag{4.19}$$

But, from Lemma 2, $\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0$ holds, therefore, there exists k_0 such that

$$\|x^{k+1} - x^k\| \leq \frac{d}{2\bar{\chi}} \quad \text{is valid for } k \geq k_0,$$

and (4.19) leads to

$$\frac{d}{\chi_k}\rho_{k+1} \leq dc_9 \frac{\varphi_k + \sigma_k + \delta_k}{\chi_k},$$

proving (4.8). □

Remark 5. *Estimate (4.7) shows linear convergence of $\{x^k\}$ to X^* with the factor $\left(1 + 2d\bar{\chi}^{-1}\right)^{-1/4}$, and the convergence is superlinear if $\chi_k \to 0$. If Assumption (2-v) is valid with $w^k \to 0$ (instead of $w^k \rightharpoonup 0$), then combining the proofs of Theorem 2 above and Theorem 15.3 in* [30], *one can establish linear convergence of the iterates to an element $x \in X^*$. But in that case the conditions for the choice of the controlling parameters will be harder.*
Estimate (4.8) means, in particular, that in the "ideal" situation (without data approximation and with performing exact proximal iterations) a solution of Problem (P) can be obtained in a finite number of steps.
For a comparison of Assumption 4 with related conditions in [57], [47] *see* [34].

5. The use of the ϵ-enlargement of the operator $\mathcal{Q}$

For given $\epsilon \geq 0$, the ϵ-enlargement of the operator $\mathcal{Q}$ is defined as follows [13]:

$$\mathcal{Q}_\epsilon(z) = \{w \in X' : \langle w - v, z - x\rangle \geq -\epsilon \;\; \forall x \in D(\mathcal{Q}),\; \forall v \in \mathcal{Q}(x)\}. \tag{5.1}$$

Obviously, $\mathcal{Q}_\epsilon \supset \mathcal{Q}\; \forall \epsilon > 0$, and $\mathcal{Q}_\epsilon$ possesses, in general, better continuity properties than $\mathcal{Q}$.

An extension of the GPP-method with $\mathcal{Q}^k \equiv \mathcal{Q}_{\epsilon_k}$ requires only the following modifications in some conditions of Assumption 2:

- (2-*ii*)': for each k, it holds $K^k \cap D(\mathcal{Q}) \neq \emptyset$ and

$$\begin{aligned} &\langle q^k(y) - q(x), y - x\rangle \geq \langle \mathcal{B}(y-x), y - x\rangle - a\epsilon_k, \\ &\forall y \in K^k \cap D(\mathcal{Q}_{\epsilon_k}),\; \forall x \in K^k \cap D(\mathcal{Q}), \\ &\forall q^k(y) \in \mathcal{Q}_{\epsilon_k}(y),\; \forall q(x) \in \mathcal{Q}(x), \end{aligned} \tag{5.2}$$

  where $a > 0$ and $\mathcal{B}$ is a linear continuous, symmetric and monotone operator;
- (2-*iv*)': for all k, the operators $\mathcal{Q} + \mathcal{N}_{K^k} + \chi_k \nabla h$ are maximal monotone and strongly monotone;
- (2-*v*)', (2-*vi*)' are obtained substituting $\mathcal{Q}$ for $\mathcal{Q}^k$ in (2-*v*), (2-*vi*).

If the operator $\mathcal{Q} - \mathcal{B}$ is monotone, then a quite similar modification of Assumption 2 (replace $\mathcal{Q}_{\epsilon_k}$ by $\mathcal{Q}_1^k$ in (2-*ii*)') permits to consider the GPP-method, where

$$\mathcal{Q}^k = (\mathcal{Q} - \mathcal{B})_{\epsilon_k} + \mathcal{B} \equiv \mathcal{Q}_1^k$$

is chosen. Moreover, (2-ii)' is valid for these operators if $K^k \cap D(\mathcal{Q}) \neq \emptyset \ \forall k$.
It should be noted that the operators $\mathcal{Q}_{\epsilon_k}$ and $\mathcal{Q}_1^k$ are, in general, not monotone, and that the inclusions

$$\mathcal{Q}_{\epsilon_k}(x) \supset \mathcal{Q}_1^k(x) \supset \mathcal{Q}(x)$$

hold true for any $x \in X$.

In the both cases mentioned, all the statements remain true if the condition

$$\sum_{k=1}^{\infty} \frac{\epsilon_k}{\chi_k} < \infty$$

is inserted in Lemma 2 and in the Theorems 1, 2, whereas in Lemma 3 $\epsilon_k \to 0$ is required.

The alterations in the proofs are rather straightforward:

- the solvability of Problem (P^k) follows from (2-iv)' and from the inclusion $\mathcal{Q}_{\epsilon_k}(x) \supset \mathcal{Q}(x)$ or $\mathcal{Q}_1^k(x) \supset \mathcal{Q}(x)$;
- everywhere $q^k(w^k)$ is replaced by $q(w^k)$;
- in Lemma 3, $y^k \in D(\mathcal{Q}) \cap K^k$ has to be chosen;
- the definition of $\mathcal{Q}_{\epsilon_k}$ or $\mathcal{Q}_1^k$ is used instead of the monotonicity of the "old" $\mathcal{Q}^k$;
- Assumptions (2-ii)', (2-v)' and (2-vi)' are used instead of (2-ii), (2-v) and (2-vi), respectively.

Concerning the concept of the ϵ-enlargement of a maximal monotone operator, see [13], [14]. In particular, operators $\mathcal{Q}_{\epsilon_k}$ possess the Brøndsted-Rockafellar property ([13], Theorem 3.7), which can be useful to verify Assumption (3-ii).

Obviously, (5.1) provides (5.2) with $\mathcal{B} = \mathbf{0}$. Let us give an example illustrating the fulfillment of (5.2) with a non-trivial operator $\mathcal{B}$.

Example 1. Let $X' = X$, $\mathcal{T} : X \to X$ be a linear continuous, symmetric and monotone operator, $\mathcal{T}X = X^1$ and

$$\exists m > 0 : (\mathcal{T}x, x) \geq m\|x\|^2 \ \forall x \in X^1. \tag{5.3}$$

According to the definition of the ϵ-enlargement,

$$\mathcal{T}_\epsilon(z) = \{w \in X : (w - \mathcal{T}x, z - x) \geq -\epsilon \ \forall x \in X\},$$

and Corollary 3.8 in [14] yields $\mathcal{T}_\epsilon(z) \subset X^1$.
Minimizing $(w - \mathcal{T}x, z - x)$ w.r.t. x, we obtain $\bar{x} = \frac{1}{2}\left(\tilde{\mathcal{T}}^{-1}w + \tilde{\mathcal{T}}^{-1}\mathcal{T}z\right)$, where $\tilde{\mathcal{T}} = \mathcal{T}|_{X^1}$. With this minimizer $\bar{x}$ one gets

$$(w - \mathcal{T}\bar{x}, z - \bar{x}) = \frac{1}{4}\left(w - \mathcal{T}z, \tilde{\mathcal{T}}^{-1}(\mathcal{T}z - w)\right).$$

Thus, due to the definition of $\mathcal{T}_\epsilon$, the element w has to satisfy

$$\left(\tilde{\mathcal{T}}^{-1}(w - \mathcal{T}z), \mathcal{T}z - w\right) \geq -4\epsilon.$$

Taking $w_0 = w - \mathcal{T}z$, we obtain for all $x \in X$

$$
\begin{aligned}
(w - \mathcal{T}x, z - x) &= \frac{1}{2}(\mathcal{T}z - \mathcal{T}x, z - x) \\
&\quad + \frac{1}{2}\left(\mathcal{T}z + w_0 - \mathcal{T}x, z + \tilde{\mathcal{T}}^{-1}w_0 - x\right) - \frac{1}{2}\left(w_0, \tilde{\mathcal{T}}^{-1}w_0\right) \\
&\geq \frac{1}{2}(\mathcal{T}z - \mathcal{T}x, z - x) - 2\epsilon.
\end{aligned}
$$

Hence, in the case of $\mathcal{Q} = \mathcal{T}$, relation (5.2) is valid with $\mathcal{B} = \frac{1}{2}\mathcal{T}$ (or $\mathcal{B} : x \to \frac{m}{2}\mathcal{P}x,\ \mathcal{P} : X \to X^1$ the orthoprojector) and $a = 2$.

But, as distinct from the operator $\mathcal{Q}_1^k$, it is not clear whether the relation (5.2) (even though relaxed by inserting a coefficient $c \in (0,1)$ in the term $\langle \mathcal{B}(y-x), y-x\rangle$) holds for an arbitrary nonlinear operator $\mathcal{Q}$ satisfying in $D(\mathcal{Q})$

$$\langle q(y) - q(x), y - x\rangle \geq \langle \mathcal{B}(y-x), y-x\rangle \quad \forall q(\cdot) \in \mathcal{Q}(\cdot).$$

Nevertheless, if $\mathcal{Q}$ is the subdifferential of a convex function $f : X \to \bar{I\!R}$ and

$$f(y) - f(x) - \langle q(x), y - x\rangle \geq \langle \mathcal{B}(y-x), y-x\rangle \tag{5.4}$$

holds for all $q(x) \in \mathcal{Q}(x),\ x \in K^k \cap D(\mathcal{Q}),\ y \in K^k \cap dom f$, then (5.2) is true with the ϵ-subdifferential $\partial_{\epsilon_k} f$ in place of $\mathcal{Q}_{\epsilon_k}$ (let us remind that $\partial_{\epsilon_k} f \subset (\partial f)_{\epsilon_k}$ and the inclusion may be a strict one [13]).
Indeed, $D(\partial_{\epsilon_k} f) = dom f$ is valid for $\epsilon > 0$ (cf. [22]), and (5.2) follows immediately from (5.4) and the inequality

$$f(x) - f(y) - \langle q_\epsilon(y), x - y\rangle \geq -\epsilon \quad \forall q_\epsilon(y) \in \partial_\epsilon f(y), \forall x \in X,\ \forall y \in dom f$$

(considered with $\epsilon = \epsilon_k$), which defines the ϵ-subdifferential.
The convergence results for the GPP-method with $\mathcal{Q}^k = \partial f_{\epsilon_k}$ are quite analogous.

6. Appendix I

Now, we analyze a special choice of the function h, which leads to a proximal-based variant of the elliptic regularization method. Starting with the papers of Lions [44] and Olejnik [52], elliptic regularization is an efficient tool for the theoretical and numerical treatment of parabolic and degenerate elliptic problems. The main idea is that the original problem is approximated by a family of non-degenerate elliptic problems. We observe this approach in an example of a parabolic variational inequality .
The following notation is used here: $\Omega \subset I\!R^n$ is an open set with a "sufficiently smooth" boundary $\Gamma,\ \Sigma = \Omega \times]0,T[;\ \mathcal{H} = L^2(0,T;L^2(\Omega)),\ \mathcal{V} = L^2(0,T;H_0^1(\Omega))$, where $H_0^1(\Omega)$ is endowed with the norm $\|z\|_{H_0^1(\Omega)} = \left(\int_\Omega \sum_{i=1}^n \left(\frac{\partial z}{\partial y_i}\right)^2 dy\right)^{1/2}$, and $L^2(0,T;W)$ denotes the space of measurable on $]0,T[$ functions $v : t \to v(t) \in W$ such that

$$\|v\|_{L^2(0,T;W)} := \left(\int_0^T \|v(t)\|_W^2 dt\right)^{1/2} < \infty;$$

$\mathcal{V}' = L^2(0,T;H^{-1}(\Omega))$ is the dual space of $\mathcal{V}$;

$$a(t;\varphi,\psi) = \sum_{i,j=1}^{n} \int_\Omega a_{ij}(y,t)\frac{\partial\varphi}{\partial y_i}\frac{\partial\psi}{\partial y_j}dy + \int_\Omega a_0(y,t)\varphi\psi dy,$$

where $a_0,\ a_{ij} \in L^\infty(\Sigma)$ and the relations

$$\sum_{i,j=1}^{n} a_{ij}(y,t)\xi_i\xi_j \geq \alpha|\xi|^2 \quad \forall \xi \in I\!R^n, \tag{6.1}$$

$$a_0(y,t) \geq \alpha_0 \tag{6.2}$$

(with constants $\alpha > 0,\ \alpha_0 \geq 0$) are fulfilled *a.e.* in Σ.

We identify the space $\mathcal{H}$ with its dual. Introducing the linear unbounded (in $\mathcal{V}$) operator $\Lambda = \frac{d}{dt}$ with the domain

$$D(\Lambda) = \left\{v :\ v \in \mathcal{V}, \frac{dv}{dt} \in \mathcal{V}',\ v(0) = 0\right\},$$

one can consider $D(\Lambda)$ as a Hilbert space endowed with the graph-norm $\|v\|_{D(\Lambda)} = \left(\|v\|^2 + \|\Lambda u\|^2_{\mathcal{V}'}\right)^{1/2}$. Obviously,

$$D(\Lambda) \subset \mathcal{V} \subset \mathcal{H} \subset \mathcal{V}' \subset D(\Lambda)',$$

and each space is dense in the next one.
Then, with $\langle\cdot,\cdot\rangle$ the duality pairing between $D(\Lambda)$ and $D(\Lambda)'$, the linear, bounded and maximal monotone operator $\mathcal{A} : D(\Lambda) \to D(\Lambda)'$ is defined by

$$\langle \mathcal{A}u, v\rangle = \int_0^T a(t;u(t),v(t))dt \quad \forall u,v \in D(\Lambda).$$

Assuming that $f \in \mathcal{V}'$ and a convex closed set $\mathcal{K} \subset D(\Lambda)$ are given, the parabolic variational inequality

$$\begin{array}{l} find \quad u \in \mathcal{K} \ such\ that \\ \langle \mathcal{A}u, v-u\rangle + \langle \Lambda u, v-u\rangle \geq \langle f, v-u\rangle \quad \forall v \in \mathcal{K} \end{array} \tag{6.3}$$

is considered.
The operator Λ is closed and $\Lambda \geq 0$ on $D(\Lambda)$ (cf. Lions [45], Sect. 3.1.1). Therefore, Lemma 3.1.1 in [45] guarantees that Λ is maximal monotone, and because $D(\mathcal{A}) = D(\Lambda)$, the sum $\mathcal{A} + \Lambda$ is maximal monotone on $D(\Lambda)$, too. The solution of (6.3) (if it exists) is unique, and for conditions providing solvability, see [45], Sect. 2.9.6. Henceforth we suppose that (6.3) is solvable.

Following Lions, the elliptic regularization of (6.3) takes the form:

$$\begin{array}{l} find \quad u_\epsilon \in \mathcal{K} \ such\ that \\ \langle \epsilon\Lambda^* J^{-1}\Lambda u_\epsilon + \mathcal{A}u_\epsilon + \Lambda u_\epsilon, v - u_\epsilon\rangle \geq \langle f, v-u_\epsilon\rangle \quad \forall v \in \mathcal{K}, \end{array} \tag{6.4}$$

where $\epsilon > 0, \epsilon \to 0$.
Here $\Lambda^* = -\frac{d}{dt}$ with the domain

$$D(\Lambda^*) = \left\{ v : \ v \in \mathcal{V}, \frac{dv}{dt} \in \mathcal{V}', \ v(T) = 0 \right\}$$

is the conjugate operator of Λ, and $J : v \to -\Delta_y v$ is the duality mapping between $\mathcal{V}$ and $\mathcal{V}'$.

Remark 6. *More precisely, in* [45] *the elliptic regularization method is described for parabolic equations. In* [46], *considering the operator* $\Lambda = \frac{d}{dt}$ *with the domain*[3]

$$D(\Lambda) = \left\{ v : \ v \in \mathcal{V}, \frac{dv}{dt} \in \mathcal{H}, \ v(0) = 0 \right\},$$

the regularizing operator $-\frac{d^2}{dt^2}$ *stands in place of* $\Lambda^* J^{-1} \Lambda$. *In that case the ellipticity of the regularized operator is evident.*

The formal analogy of (6.4), based on the principle of proximal regularization, is:

$$\begin{aligned} &\textit{find} \quad u^{k+1} \in \mathcal{K} \textit{ such that} \\ &\qquad \langle \chi_k \Lambda^* J^{-1} \Lambda (u^{k+1} - u^k) + \mathcal{A} u^{k+1} + \Lambda u^{k+1}, v - u^{k+1} \rangle \geq \langle f, v - u^{k+1} \rangle \\ &\qquad \forall v \in \mathcal{K} \end{aligned} \tag{6.5}$$

$(0 < \chi_k \leq \bar{\chi} < \infty)$.
To our knowledge, this iteration process was never studied before. It corresponds to the GPP-method with

$$X = D(\Lambda), \ \mathcal{Q}^k \equiv \mathcal{Q} : u \to \mathcal{A}u + \Lambda u - f, \ K^k \equiv \mathcal{K}, \ \delta_k \equiv 0$$

and $h : \nabla h(u) = \Lambda^* J^{-1} \Lambda u$ in (2.2). Moreover, the operator $\mathcal{M}_k : D(\Lambda) \to D(\Lambda)'$ defined by

$$\chi_k \langle J^{-1} \Lambda u, \Lambda v \rangle + \langle \Lambda u, v \rangle = \langle \mathcal{M}_k u, v \rangle \quad \forall u, v \in D(\Lambda)$$

is obviously linear, bounded and monotone. Hence $\mathcal{M}_k + \Lambda$ is maximal monotone in $D(\Lambda)$, and Assumption 1 is valid.
With the linear and bounded operator

$$\mathcal{B} : \ \langle \mathcal{B}\varphi, \psi \rangle = \alpha \int_0^T \int_\Omega \sum_{i=1}^n \frac{\partial \varphi}{\partial y_i} \frac{\partial \psi}{\partial y_i} dy dt$$

we obtain for $u, v \in D(\Lambda)$ that

$$\langle \mathcal{B}(u - v), u - v \rangle + \langle \Lambda^* J^{-1} \Lambda (u - v), u - v \rangle \geq \min\{\alpha, 1\} \|u - v\|^2_{D(\Lambda)}.$$

Using the well-known properties of convex functions, this yields

$$\langle \mathcal{B}(u - v), u - v \rangle + h(v) - h(u) - \langle \nabla h(u), v - u \rangle \geq \frac{1}{2} \min\{\alpha, 1\} \|u - v\|^2_{D(\Lambda)},$$

and in view of (6.1), (6.2), the Assumptions (2-*ii*), (2-*iii*) on the choice of h are fulfilled with this operator $\mathcal{B}$ and $\tilde{\chi} = 2, \ m = \frac{1}{2} \min\{\alpha, 1\}$. Assumption (2-*i*) follows immediately from the fact that $\Lambda^* J^{-1} \Lambda : D(\Lambda) \to D(\Lambda)'$ is a linear, monotone and

[3] The use of this domain seems to be less natural.

bounded operator.
Now, as it follows from Theorem 1, the above mentioned properties of the function h and the original problem (6.3) are sufficient to ensure weak convergence of the GPP-method (6.5) as well as of a more realistic variant with an approximation of K (according to the Assumptions (2-v) - (2-$viii$)) and $\sum_{k=1}^{\infty} \delta_k \chi_k^{-1} < \infty$.

Taking into account that $\Lambda^* J^{-1} \Lambda$ is a bounded and closed operator, condition (3.20) is fulfilled, at least if $K^k \equiv \mathcal{K}$. This is some hint for the case, when $\mathcal{A}$ is a degenerate elliptic operator and the uniqueness of a solution of an (original) problem is not guaranteed (of course, in this situation we need an additional regularizing operator, for instance $-\Delta_y$, to provide non-degenerate ellipticity of the Problems (P^k)).

7. Appendix II

Here, we describe an example illustrating the choice of single-valued operators $\mathcal{Q}^k$ in accordance with the Assumptions (2-v), (2-vi). The approximation of the set K in this example will be performed quite artificially, modeling a situation which arises under a successive discretization of variational inequalities in mathematical physics. Usually, such a discretization is carried out by means of the finite element method on a sequence of triangulations, and for an arbitrary element $w \in K$ one can guarantee only that there exists a sequence $\{w^k\}$, $w^k \in K^k$, with

$$\lim_{k \to \infty} \|w - w^k\| = 0$$

(cf. condition (a) below). Certainly, it does not suffice to provide an approximation of a solution x^* as in Assumption (2-vi). But, a helpful circumstance is that, according to the regularity theorems (see [11], [25], [23], [28]), solutions of many variational inequalities in mathematical physics possess better regularity properties than arbitrary elements from K. Then, error estimates for finite element methods [18] show typically better approximations of the solutions by elements from K^k:

$$\forall x \in X^* \ \exists \ w^k \in K^k : \quad \|x - w^k\| \le a(x) h_k^{\alpha},$$

with h_k a triangulation parameter and $\alpha > 0$ independent of x (this quality is reflected in condition (b) below).

Taking into account the mentioned properties of finite element methods, the approach used in the example can be extended to approximate some problems of linear elasticity (where the operator $\mathcal{Q}$ is multi-valued due to the friction), fluid mechanics, image reconstruction etc.. Concerning the proximal point method for some of these problems, see [31], [33], [60]).

Example 2. Let $X = I\!R^2$, $\mathcal{Q} : x \to \binom{z(x_1) - x_2}{x_1}$, where

$$z(x_1) = \begin{cases} -1 & if \quad x_1 < -1, \\ [-1, 0] & if \quad x_1 = -1, \\ 0 & if \quad x_1 \in (-1, 1), \\ [0, 1] & if \quad x_1 = 1, \\ 1 & if \quad x_1 > 1, \end{cases}$$

and we consider two cases:

$$K = K_1 = \{x \in I\!R^2 : \; |x_1| \le 2, \; x_2 \ge 0\}$$

and

$$K = K_2 = \{x \in I\!R^2 : x_1 \ge 1, \; x_2 \ge 0\}.$$

It is not difficult to see that

$$X^* = \{(x_1, 0) : 0 \le x_1 \le 1\}, \quad \hat{\Lambda}(x) = \{\begin{pmatrix} 0 \\ x_1 \end{pmatrix}\} \;\text{ if }\; K = K_1,$$

and

$$X^* = \{(1,0)\}, \quad \hat{\Lambda}(1,0) = \{\begin{pmatrix} a \\ 1 \end{pmatrix} : a \in [0,1]\} \;\text{ if }\; K = K_2.$$

The operator $\mathcal{Q}$ is maximal monotone, because

$$\mathcal{Q}(x) = \{\partial_{x_1} f(x_1, x_2), \partial_{x_2}(-f(x_1, x_2))\}^T,$$

where

$$f(x_1, x_2) = \max\{0, -x_1 - 1\} + \max\{0, x_1 - 1\} - x_1 x_2$$

is a convex-concave function and $\partial_{x_1}, \partial_{x_2}$ denote the partial subdifferentials. But, this operator is neither strictly monotone (take $x^1 = (-\frac{1}{2}, -1), \; x^2 = (\frac{1}{2}, -1)$), nor symmetric.

We choose the maximal monotone operators

$$\mathcal{Q}^k : x \to \begin{pmatrix} z^k(x_1) - x_2 \\ x_1 \end{pmatrix},$$

with

$$z^k(x_1) = \frac{1}{2}\left[\frac{1 + x_1}{\sqrt{(1 + x_1)^2 + k^{-6}}} + \frac{x_1 - 1}{\sqrt{(x_1 - 1)^2 + k^{-6}}}\right].$$

The idea of such an approximation is obvious: the operators $\mathcal{Q}^k$ can be described as

$$\mathcal{Q}^k(x) = \{\frac{\partial}{\partial_{x_1}} f^k(x_1, x_2), -\frac{\partial}{\partial_{x_2}} f^k(x_1, x_2)\}^T,$$

where convex-concave and differentiable functions

$$f^k(x) = \frac{1}{2}\left(\sqrt{(1 + x_1)^2 + k^{-6}} + \sqrt{(x_1 - 1)^2 + k^{-6}}\right) - x_1 x_2 - 1,$$

converge to f uniformly on X.

Taking $w = (w_1, w_2), \; w^k = (w_1^k, w_2^k)$, any $q(w) \in \mathcal{Q}(w)$ has the form

$$q(w) = \begin{pmatrix} \xi(w_1) - w_2 \\ w_1 \end{pmatrix},$$

with $\xi(w_1) \in z(w_1)$, thus

$$\|\mathcal{Q}^k(w^k) - q(w)\|^2 = \left(z^k(w_1^k) - \xi(w_1) - w_2^k + w_2\right)^2 + \left(w_1^k - w_1\right)^2$$

and

$$\|\mathcal{Q}^k(w^k) - q(w)\| \le |z^k(w_1^k) - \xi(w_1)| + |w_1^k - w_1| + |w_2^k - w_2|.$$

A straightforward calculation shows that:

- $\lim_{k\to\infty} \|\mathcal{Q}^k(w^k) - q(w)\| = 0$ if $w_1 \neq \pm 1$ and $\{w^k\}$ is an arbitrary sequence converging to w;
- $\|\mathcal{Q}^k(w^k) - \binom{\xi(w_1)-w_2}{w_1}\| \leq 3k^{-2}$
  - if $w_1 = 1,\ \xi(w_1) = 1$ and $w^k = (1 + k^{-2}, w_2 + k^{-2})$,
  - or $0 \leq w_1 \leq 1,\ \xi(w_1) = 0,\ k \geq 2$ and $w^k = (w_1 - k^{-2}, w_2 + k^{-2})$,
  - or $w_1 = -1,\ \xi(w_1) = -1$ and $w^k = (-1 - k^{-2}, w_2 + k^{-2})$.

This permits to conclude immediately that the Assumptions (2-v), (2-vi) hold true with

$$0 < \underline{\chi} \leq \chi_k \leq \bar{\chi},\ \varphi_k = k^{-2},\ \sigma_k = k^{-2} \ \text{ and } \ c_1 = \sqrt{2},\ c_2 = 3$$

in (2.3), (2.4), if the family $\{K^k\} \subset I\!R^2$ satisfies the following conditions:

(a) $\forall w \in K,\ \forall k\ \exists w^k \in K^k:\ \lim_{k\to\infty} \|w^k - w\| = 0$;

(b) in case $K = K_1$, the pairs $\binom{-1-k^{-2}}{k^{-2}}$ and $\binom{1+k^{-2}}{k^{-2}}$ belong to K^k for sufficiently large k;
in case $K = K_2$, the inclusion $\binom{1+k^{-2}}{k^{-2}} \in K^k$ is fulfilled for sufficiently large k.

Moreover, for $K = K_1$, taking into account the convexity of K^k, (a) and (b) provides the fulfillment of Assumption (2-vi) for each $x \in X^*$ (if $K = K_2$, X^* is a singleton). Obviously, conditions (a) and (b) are satisfied if

$$K^k \supset \{x \in I\!R^2 : |x_1| \leq 2 - \alpha_k,\ x_2 \geq k^{-2}\} \ \text{ with } \ \alpha_k \to +0,$$

and

$$K^k \supset \{x \in I\!R^2 : x_1 \geq 1 + k^{-2},\ x_2 \geq k^{-2}\}$$

for $K = K_1$ and $K = K_2$, respectively.

References

[1] L. Abbe, *Two logarithmic barrier methods for convex semi-infinite problems*, Proceedings of the Conference on Semi-infinite Programming, Alicante (M.A. Goberna and M.A. Lopez, eds.), 2000 (to appear).

[2] P. Alart and B. Lemaire, *Penalization in non-classical convex programming via variational convergence*, Math. Programming **51** (1991), 307–331.

[3] A.S. Antipin, *On a method for convex programs using a symmetrical modification of the Lagrange function*, Ekonomika i Mat. Metody **12** (1976), 1164–1173.

[4] H. Attouch, J. B. Baillon, and M. Théra, *Sum of maximal monotone operators revisited: The variational sum*, J. of Convex Analysis **1** (1994), 1–29.

[5] A. Auslender, J.P. Crouzeix, and P. Fedit, *Penalty proximal methods in convex programming*, J. Opt. Theory Appl. **55** (1987), 1–21.

[6] A. Auslender and M. Haddou, *An interior-proximal method for convex linearly constrained problems and its extension to variational inequalities*, Math. Programming **71** (1995), 77–100.

[7] A. Auslender, M. Teboulle, and S. Ben-Tiba, *Interior proximal and multiplier methods based on second order homogeneous kernels*, Mathematics of Oper. Res. **24** (1999).

[8] ———, *A logarithmic-quadratic proximal method for variational inequalities*, Computational Optimization and Applications **12** (1999), 31–40.

[9] A. Ben-Tal and M. Zibulevski, *Penalty-barrier methods for convex programming problems*, SIAM J. Optim. **7** (1997), 347–366.

[10] J. Bonnans, J. Ch. Gilbert, C. Lemaréchal, and C. Sagastizabal, *A family of variable metric proximal methods*, Math. Programming **68** (1995), 15–47.

[11] H. Brezis, *Problèmes Unilatéraux*, J. Math. Pures Appl. **72** (1971), 1–168.

[12] F. E. Browder, *Nonlinear maximal monotone operators in Banach spaces*, Math. Annalen **175** (1968), 89–113.

[13] R.S. Burachik, A.N. Iusem, and B.F. Svaiter, *Enlargement of monotone operators with applications to variational inequalities*, Set-Valued Analysis **5** (1997), 159–180.

[14] R.S. Burachik and B.F. Svaiter, *ϵ-enlargement of maximal monotone operators in Banach spaces*, Set-Valued Analysis **7** (1999), 117–132.

[15] Y. Censor and S. A. Zenios, *Proximal minimization algorithm with D-functions*, J. Optim. Theory Appl. **73** (1992), 451–464.

[16] G. Chen and M. Teboulle, *Convergence analysis of a proximal-like minimization algorithm using Bregman functions*, SIAM J. Optim. **3** (1993), 538–543.

[17] ———, *A proximal-based decomposition method for convex minimization problems*, Math. Programming **64** (1994), 81–101.

[18] Ph. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.

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

[20] ———, *Approximate iterations in Bregman-function-based proximal algorithms*, Math. Programming **83** (1998), 113–123.

[21] J. Eckstein and D. Bertsekas, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Math. Programming **55** (1992), 293–318.

[22] I. Ekeland and R. Temam, *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, 1976.

[23] A. Friedman, *Variational Principles and Free-boundary Problems*, J. Wiley & Sons, Chichester - New York - Brisbane - Toronto - Singapure, 1982.

[24] M. Fukushima, *The primal Douglas-Rachford splitting algorithm for a class of monotone mappings with applications to the traffic equilibrium problem*, Math. Programming **72** (1996), 1–15.

[25] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer, New York - Berlin - Heidelberg - Tokyo, 1984.

[26] Cu D. Ha, *A generalization of the proximal point algorithm*, SIAM J. Control Optim. **28** (1990), 503–512.

[27] R. Hettich, A. Kaplan, and R. Tichatschke, *Regularized penalty methods for ill-posed optimal control with elliptic equations, Parts I and II*, Control and Cybernetics **26** (1997), 5–27, 29–43.

[28] I. Hlaváček, J. Haslinger, I. Nečas, and J. Lovišek, *Numerical Solution of Variational Inequalities*, Springer Verlag, Berlin-Heidelberg-New York, 1988.

[29] S. Ibaraki, M. Fukushima, and T. Ibaraki, *Primal-dual proximal point algorithm for linearly constrained convex programming problems*, Comput. Opt. Appl. **1** (1992), 207–226.

[30] A. Kaplan and R. Tichatschke, *Stable Methods for Ill-posed Variational Problems - Prox-regularization of Elliptic Variational Inequalities and Semi-infinite Optimization Problems*, Akademie Verlag Berlin, 1994.

[31] ———, *Prox-regularization and solution of ill-posed elliptic variational inequalities*, Applications of Mathematics **42** (1997), 111–145.

[32] ———, *Proximal penalty methods for ill-posed parabolic optimal control problems*, Control and Estimation of Distributed Parameter Systems (ISNM, Birkhäuser) (W. Desch, F. Kappel, and K. Kunisch, eds.), vol. 126, 1998, pp. 169–182.

[33] ———, *Stable solution of variational inequalities with composed monotone operators*, Ill-posed Variational Problems and Regularization Techniques (Lecture Notes in Economics and Mathematical Systems, Springer) (M. Théra and R. Tichatschke, eds.), vol. 477, 1999, pp. 111–136.

[34] ———, *Proximal point approach and approximation of variational inequalities*, SIAM J. Control Optim. **39** (2000), 1136–1159.

[35] ———, *Auxiliary problem principle and the approximation of variational inequalities with non-symmetric multi-valued operators*, CMS Conference Proc. **27** (2000), 185–209.

[36] K. Kiwiel, *An ellipsoid trust region bundle method for nonsmooth convex minimization*, SIAM J. Control Optim. **27** (1989), 737–757.

[37] ———, *Proximal level bundle methods for convex nondifferentiable optimization, saddle point problems and variational inequalities*, Math. Programming **69 B** (1995), 89–109.

[38] ———, *Proximal minimization methods with generalized Bregman functions*, SIAM J. Control Optim. **35** (1997), 1142–1168.

[39] I.V. Konnov, *A class of combined iterative methods for solving variational inequalities*, J. Opt. Theory Appl. **94** (1997), 677–693.

[40] ———, *A combined relaxation method for variational inequalities with nonlinear constraints*, Math. Programming **80** (1998), 239–252.

[41] B. Lemaire, *Coupling optimization methods and variational convergence*, ISNM **84** (1988), 163–179.

[42] C. Lemaréchal and C.A. Sagastizábal, *An approach to variable metric bundle methods*, Lecture Notes in Control and Information Sciences **197** (1994), 144–162.

[43] ———, *Variable metric bundle methods: from conceptual to implementable forms*, Mathematical Programming **76** (1997), 393–410.

[44] J. L. Lions, *Equations Différentielles Opérationelles dans les Espaces de Hilbert*, C.I.M.E. Varenna, Juillet, 1963.

[45] ———, *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires*, Dunod, Gauthier-Villars, Paris, 1969.

[46] J. L. Lions and E. Magenes, *Problèmes aux Limites Non Homogènes et Applications*, V.1. Dunod, Paris, 1968.

[47] E.J. Luque, *Asymptotic convergence analysis of the proximal point algorithms*, SIAM J. Control Optim. **22** (1984), 277–293.

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

[49] R. Mifflin, *A quasi-second-order proximal bundle algorithm*, Math. Programming **73** (1996), 51–72.

[50] K. Mouallif and P. Tossings, *Une méthode de pénalisation exponentielle associée à une régularisation proximale*, Bull. Soc. Roy. Sc. de Liège **56** (1987), 181–192.

[51] R. V. Namm, *Stable methods for ill-posed variational inequalities in mechanics*, Recent Advances in Optimization (Lecture Notes in Economics and Mathematical Systems, Springer) (P. Gritzmann, R. Horst, E. Sachs, and R. Tichatschke, eds.), vol. 452, 1997, pp. 214–228.

[52] O.A. Olejnik, *On second order linear equations with nonnegative characteristic form*, Matem. Sb. **69** (1966), 111–140.

[53] B. T. Polyak, *Introduction to Optimization*, Optimization Software, Inc. Publ. Division, New York, 1987.

[54] R.A. Polyak and M. Teboulle, *Nonlinear rescaling and proximal-like methods in convex optimization*, Math. Programming **76** (1997), 265–284.

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

[56] ———, *Augmented Lagrange multiplier functions and applications of the proximal point algorithm in convex programming*, Math. Oper. Res. **1** (1976), 97–116.

[57] ———, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim. **14** (1976), 877–898.

[58] S. Rotin, *Convergence of the proximal-point method for ill-posed control problems with partial differential equations*, Ph.D. thesis, University of Trier, 1999.

[59] H. Schmitt, *Normschwächere prox-regularisierungen*, Ph.D. thesis, University of Trier, 1996.

[60] ———, *On the regularized Bingham problem*, Recent Advances in Optimization (Lecture Notes in Economics and Mathem. Systems, Springer, Berlin) (P. Gritzmann, R. Horst, E. Sachs, and R. Tichatschke, eds.), vol. 452, 1997, pp. 298–314.

[61] M.V. Solodov and B.F. Svaiter, *A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator*, Set-Valued Analysis **7** (1999), 323–345.

[62] ———, *A hybrid projection-proximal point algorithm*, Journal of Convex Analysis **6** (1999), 59–70.

[63] P. Tossings, *Algorithme du point proximal perturbé et applications*, Prépublication, Inst. de Mathématique, Univ. de Liège **90-015** (1990).
[64] P. Tseng, *Further applications of a splitting algorithm to decomposition in variational inequalities and convex programming*, Math. Programming **48** (1990), 249–263.
[65] ———, *Applications of a splitting algorithm to decomposition in convex programming and variational inequalities*, SIAM J. Control Optim. **29** (1991), 119–138.
[66] ———, *A modified forward-backward splitting method for maximal monotone mappings*, SIAM J. Control Optim. **38** (2000), 431–446.
[67] P. Tseng and D. Bertsekas, *On the convergence of the exponential multiplier method for convex programming*, Math. Programming **60** (1993), 1–19.

Manuscript received May 22, 2001

A. Kaplan
Department of Mathematics, University of Trier, D-54286 Trier
E-mail address: a.a.kaplan@tiscalinet.de

R. Tichatschke
Department of Mathematics, University of Trier, D-54286 Trier
E-mail address: tichat@uni-trier.de