IOP PUBLISHING INVERSE PROBLEMS

Inverse Problems **27** (2011) 085001 (11pp) doi:10.1088/0266-5611/27/8/085001

Convergence rates in constrained Tikhonov regularization: equivalence of projected source conditions and variational inequalities

Jens Flemming and Bernd Hofmann

Department of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany

E-mail: jens.flemming@mathematik.tu-chemnitz.de and hofmannb@mathematik.tu-chemnitz.de

Received 21 March 2011, in final form 13 May 2011
Published 4 July 2011
Online at stacks.iop.org/IP/27/085001

Abstract

In this paper, we enlighten the role of variational inequalities for obtaining convergence rates in Tikhonov regularization of nonlinear ill-posed problems with convex penalty functionals under convexity constraints in Banach spaces. Variational inequalities are able to cover solution smoothness and the structure of nonlinearity in a uniform manner, not only for unconstrained but, as we indicate, also for constrained Tikhonov regularization. In this context, we extend the concept of projected source conditions already known in Hilbert spaces to Banach spaces, and we show in the main theorem that such projected source conditions are to some extent equivalent to certain variational inequalities. The derived variational inequalities immediately yield convergence rates measured by Bregman distances.

1. Introduction

After the millennium, there was a substantial progress in regularization theory including convergence rate results for linear and nonlinear ill-posed operator equations in Banach spaces (see, e.g., [1, 4, 6, 13, 23–25]). The extension of the theory from Hilbert spaces to Banach spaces was strongly motivated by a wide field of inverse problems in natural sciences, engineering and finance, which were modeled by the operator equations under consideration. Owing to the seminal paper [6] the concept of Bregman distances could be established as a powerful tool for measuring the regularization error in the Banach space setting. The paper [13] introduced variational inequalities into the theory that cover solution smoothness and the structure of nonlinearity in a uniform manner. From those variational inequalities which are formulated based on dual pairings, Bregman distances and norms of differences of operator values, one can immediately derive convergence rates in Tikhonov regularization with convex penalty functionals also for non-smooth operators. In [10] it was shown that the classical concept of source conditions and the variational inequality concept coincide for the special case of linear operators in Hilbert spaces. The discussions in [3, 11, 13, 15] concerning the interplay of source conditions and such variational inequalities, however, do not explicitly meet the practically important case that *a priori* information restricts the admissible solutions to a convex set.

In a Hilbert space setting, Tikhonov regularization of ill-posed operator equations under convexity constraints was originally studied in a systematic way for the linear case in [21] and extended to nonlinear problems and general Hölder source conditions in [7, 22]. It was outlined there that source conditions required for obtaining convergence rates can be constructed by applying metric projectors to classical source conditions. This is mostly a reformulation of another type of variational inequality formulated for inner products and holding for all elements of the convex set of admissible solutions. For the analysis of such projected source conditions, we also refer to [8, 16, 19, 20, 26, 27].

In this paper, we extend the results on projected source conditions to the concept of Tikhonov-type regularization in Banach spaces in light of the variational inequality approach from [13].

The paper is organized as follows. After preparing the setting in the next section we present our main result in section 3. The normal cone condition occurring there is reinterpreted as a projected source condition in section 4. For linear forward operators, the main theorem is proven in section 5 and for nonlinear ones in section 6.

2. Problem setting and basic assumptions

We consider the problem of solving ill-posed operator equations

$$F(x) = y \tag{2.1}$$

as mathematical models of inverse problems, where $F : \mathcal{D}(F) \subseteq X \to Y$ is, in general, the nonlinear forward operator possessing the domain $\mathcal{D}(F)$ and mapping between Banach spaces X and Y with dual spaces X^* and Y^*, respectively. For simplicity, we denote by $\|\cdot\|$ the norms in both spaces X and Y. In X and Y, we consider in addition to the norm convergence the associated weak convergence. That means in X

$$x_k \rightharpoonup x \quad \Longleftrightarrow \quad \langle \xi, x_k\rangle \to \langle \xi, x\rangle \qquad \text{for all} \quad \xi \in X^*$$

for the dual pairing $\langle\cdot,\cdot\rangle$ with respect to X^* and X. The weak convergence in Y is defined in an analog manner.

Ill-posedness of (2.1) means that for exact right-hand sides $y = y^0 \in F(\mathcal{D}(F))$, the solutions of the operator equation need not be uniquely determined and small perturbations on the right-hand side may lead to arbitrarily large errors in the solution. For such problems, regularization methods are required in order to obtain stable approximate solutions. Here we assume that perturbed data y^δ are available instead of y satisfying the inequality

$$\|y^\delta - y\| \leqslant \delta \tag{2.2}$$

with noise level $\delta \geqslant 0$ and that *a priori* information can be exploited which allows us to restrict the set of admissible solutions to some non-empty subset

$$C \subseteq \mathcal{D}(F), \quad C \text{ convex},$$

of the domain of F. The set C defines the constraints of the problem.

Our focus is on Tikhonov-type regularization with the penalty functional $\Omega : X \to [0, \infty]$ and regularization parameter $\alpha \geqslant 0$, where the regularized solutions x_α^δ are the solutions of the extremal problem

$$T_\alpha^\delta(x) := \|F(x) - y^\delta\|^p + \alpha\,\Omega(x) \to \min, \qquad \text{subject to} \quad x \in C, \tag{2.3}$$

with a prescribed norm exponent $1 < p < \infty$. In this context, we refer to

$$\mathcal{D}(\Omega) := \{x \in X : \Omega(x) < \infty\}$$

as the domain of Ω and set $\mathcal{D} := \mathcal{D}(\Omega) \cap C \neq \emptyset$.

Throughout this paper, we make the following assumptions.

Assumption 2.1.

(a) $F : \mathcal{D}(F) \subseteq X \to Y$ *is weakly sequentially continuous and* $\mathcal{D}(F)$ *is weakly sequentially closed, i.e.*

$$x_k \rightharpoonup x \quad \text{and} \quad x_k \in \mathcal{D}(F) \quad \Longrightarrow \quad x \in \mathcal{D}(F) \quad \text{and} \quad F(x_k) \rightharpoonup F(x).$$

(b) The functional Ω *is convex and weakly sequentially lower semi-continuous.*

(c) For every $\alpha > 0$ *and* $c \geqslant 0$, *the level sets*

$$\mathcal{M}_\alpha(c) := \left\{x \in \mathcal{D} : T_\alpha^0(x) \leqslant c\right\}$$

are weakly sequentially pre-compact in the following sense: provided that $\mathcal{M}_\alpha(c)$ *is non-empty, then every sequence from* $\mathcal{M}_\alpha(c)$ *has a subsequence, which is weakly convergent to some element from X.*

(d) Let $x^\dagger \in \mathcal{D} = \mathcal{D}(\Omega) \cap C$ *satisfy*

$$\Omega(x^\dagger) = \min\{\Omega(x) : F(x) = y,\ x \in \mathcal{D}\} < \infty,$$

i.e. $x^\dagger$ *is an* Ω*-minimizing solution to (2.1) in C. For that* $x^\dagger$, *the subdifferential*

$$\partial\Omega(x^\dagger) := \{\xi \in X^* : \Omega(x) \geqslant \Omega(x^\dagger) + \langle \xi, x - x^\dagger\rangle \text{ for all } x \in X\} \subseteq X^*$$

at $x^\dagger$ *is assumed to be non-empty, and moreover a bounded linear operator* $F'[x^\dagger] : X \to Y$ *is assumed to exist such that*

$$F'[x^\dagger](x - x^\dagger) = \lim_{t \to +0} \frac{1}{t}(F(x^\dagger + t(x - x^\dagger)) - F(x^\dagger)) \text{ for all } x \in C.$$

Remark 2.2. In principle, assumption 2.1 above and the assumptions in [13] are comparable. Hence, for all $\alpha > 0$ and all y^δ, the existence of regularized solutions x_α^δ minimizing T_α^δ can be concluded from theorem 3.1 in [13]. Moreover, from [13, theorem 3.4], we find the existence of Ω-minimizing solutions $x^\dagger$ required in our assumption 2.1 (*d*) if the operator equation (2.1) has a solution in C and if C is closed and therefore due to the convexity also weakly closed. In this context, note that in [13, assumption 4.1] conditions were formulated that had to hold on level sets $\mathcal{M}_{\alpha_{\max}}(\rho)$ for sufficiently large $\rho > 0$. Throughout this paper, we will avoid such conditions by making associate assumptions for all elements from C. We conclude this remark by mentioning that the operator $F'[x^\dagger]$ in assumption 2.1 (*d*) is a Gâteaux derivative if $x^\dagger$ is an inner point of C.

As obvious for regularization theory in Banach spaces, we will use Bregman distances with respect to Ω of two elements x and $\tilde{x}$ from $\mathcal{D}(\Omega)$ and associated with some $\tilde{\xi} \in \partial\Omega(\tilde{x})$ which are defined as

$$B_{\tilde{\xi}}(x, \tilde{x}) := \Omega(x) - \Omega(\tilde{x}) - \langle \tilde{\xi}, x - \tilde{x}\rangle.$$

The Bregman distance can be defined only for $\tilde{x}$ with $\partial\Omega(\tilde{x}) \neq \emptyset$.

We measure the accuracy of approximations $x_\alpha^\delta \in \mathcal{D}$ to solutions $x^\dagger \in \mathcal{D}$ of (2.1) in the form $B_{\xi^\dagger}\left(x_\alpha^\delta, x^\dagger\right)$ for some $\xi^\dagger \in \partial\Omega(x^\dagger)$. Following the ideas from [13] with the extensions in [15] and [3], then one can prove the following *convergence rate* result.

Proposition 2.3. *Let $\varphi : [0, \infty) \to [0, \infty)$ be a concave and increasing function, twice differentiable on $(0, \infty)$ with $\lim_{t\to+0} \varphi(t) = 0$, and assume that there are $\beta_1 \in [0, 1)$, $\beta_2 \geqslant 0$ and $\xi^\dagger \in \partial\Omega(x^\dagger)$ such that the Ω-minimizing solution $x^\dagger$ of (2.1) satisfies the variational inequality*

$$\langle -\xi^\dagger, x - x^\dagger\rangle \leqslant \beta_1 B_{\xi^\dagger}(x, x^\dagger) + \beta_2\, \varphi(\|F(x) - F(x^\dagger)\|) \qquad \text{for all} \quad x \in C. \tag{2.4}$$

Then

$$B_{\xi^\dagger}\left(x^\delta_{\alpha(\delta)}, x^\dagger\right) = \mathcal{O}(\varphi(\delta)) \qquad \text{as} \quad \delta \to 0 \tag{2.5}$$

for an appropriate parameter choice $\alpha = \alpha(\delta)$.

Proof. Since $x^\delta_\alpha \in C$ for all $\alpha > 0$ and all $\delta \geqslant 0$, the assertion can be proven in the same way as theorem 4.4 in [13] for the case $\varphi(t) = t$ and theorem 4.3 in [3] for general φ. One only has to replace the sets $\mathcal{M}_{\alpha_{\max}}(\rho)$ there by C. We also refer to both papers with respect to the explicit structure of the required parameter choice. □

In the following, we formulate sufficient and for the special case $\varphi(t) = t$ also necessary conditions under which a variational inequality (2.4) holds true. With those studies, we complement the results presented in [13, remark 4.2 and 4.3], [24, section 3.2], [15, section 4–6] and [3, section 5] on the interplay of *source conditions, nonlinearity conditions and variational inequalities* by extending them to *convexity constraints* expressed by the set C.

In a Hilbert space setting under constraints imposed by a convex and closed set C and for nonlinear Tikhonov regularization with the penalty functional $\Omega(x) = \|x - \bar{x}\|^2$, the traditional source conditions

$$x^\dagger - \bar{x} = F'[x^\dagger]^* w, \qquad w \in Y,$$

(cf, e.g., [9, chapter 10]) have to be replaced by *the projected source conditions*

$$x^\dagger = P_C(\bar{x} + F'[x^\dagger]^* w), \qquad w \in Y, \tag{2.6}$$

in order to obtain comparable convergence rates, where $P_C : X \to X$ denotes the metric projector onto the set C which is well defined in the Hilbert space X. By means of the inner product $\langle\cdot,\cdot\rangle$ in the Hilbert space X, this condition can be rewritten as

$$\langle F'[x^\dagger]^* w + \bar{x} - x^\dagger, x - x^\dagger\rangle \leqslant 0 \text{ for all } x \in C \tag{2.7}$$

or when using the *normal cone*

$$N_C(\tilde{x}) := \{z \in X : \langle z, x - \tilde{x}\rangle \leqslant 0 \text{ for all } x \in C\} \tag{2.8}$$

of C at the point $\tilde{x} \in C$ alternatively as

$$F'[x^\dagger]^* w + \bar{x} - x^\dagger \in N_C(x^\dagger) \tag{2.9}$$

(cf, e.g., [7, 22]).

In the Banach space setting and for general convex penalty functionals Ω, source conditions for the unconstrained case attain the form

$$\xi^\dagger = F'[x^\dagger]^* \eta, \qquad \eta \in Y^*, \tag{2.10}$$

for some $\xi^\dagger \in \partial\Omega(x^\dagger)$ (cf, e.g., [13, 24] and [3, 12, 15]). An extension of the concept of projected source conditions to the Banach space setting can be based on the extended analog

$$F'[x^\dagger]^* \eta - \xi^\dagger \in N_C(x^\dagger) \tag{2.11}$$

of condition (2.9), where the associated normal cone $N_C(x^\dagger)$ of C at $x^\dagger \in C$ is defined as

$$N_C(x^\dagger) := \{\xi \in X^* : \langle \xi, x - x^\dagger\rangle \leqslant 0 \text{ for all } x \in C\} \tag{2.12}$$

exploiting the dual pairing $\langle\cdot,\cdot\rangle$ with respect to X^* and X.

3. Main result

Let C be convex and let assumption 2.1 concerning F, Ω, their domains $\mathcal{D}(F)$, $\mathcal{D}(\Omega)$, and the element $x^\dagger$ be satisfied. Then we can formulate our main result connecting the projected source conditions (2.11) with the variational inequalities (2.4).

Theorem 3.1. *If there are $\xi^\dagger \in \partial\Omega(x^\dagger)$ and $\eta \in Y^*$ such that (2.11) with the normal cone (2.12) is valid and if the nonlinearity condition*

$$\|F'[x^\dagger](x - x^\dagger)\| \leqslant \varphi(\|F(x) - F(x^\dagger)\|) \qquad \text{for all} \quad x \in C \tag{3.1}$$

is satisfied for some concave and increasing function $\varphi : [0,\infty) \to [0,\infty)$ with $\lim_{t\to+0}\varphi(t) = 0$, then the variational inequality (2.4) is fulfilled for all $\beta_1 \in [0,1)$ and all $\beta_2 \geqslant \|\eta\|$.

If $N_{\mathcal{D}(\Omega)}(x^\dagger) \cap (-N_C(x^\dagger)) = \{0\}$ and if at least one of the sets $\mathcal{D}(\Omega)$ or C has interior points, then we also have a converse result: if there are $\xi^\dagger \in \partial\Omega(x^\dagger)$ and constants $\beta_1 \in [0,1)$, $\beta_2 \geqslant 0$, such that

$$\langle -\xi^\dagger, x - x^\dagger\rangle \leqslant \beta_1 B_{\xi^\dagger}(x, x^\dagger) + \beta_2\|F(x) - F(x^\dagger)\| \qquad \text{for all} \quad x \in C, \tag{3.2}$$

then elements $\tilde{\xi}^\dagger \in \partial\Omega(x^\dagger)$ and $\eta \in Y^*$ exist with $\|\eta\| \leqslant \beta_2$ such that

$$F'[x^\dagger]^*\eta - \tilde{\xi}^\dagger \in N_C(x^\dagger)$$

with $N_C(x^\dagger)$ from (2.12).

The proof of theorem 3.1 will be given in section 6 in a rather simple way after presenting some preparations and auxiliary results in the subsequent sections. Note that the nonlinearity condition (3.1) can be replaced by alternative conditions; see remark 6.2.

Combining the first part of theorem 3.1 with proposition 2.3, we obtain convergence rates from a projected source condition (2.11) in Banach spaces. In fact, if there are $\xi^\dagger \in \partial\Omega(x^\dagger)$ and $\eta \in Y^*$ such that

$$F'[x^\dagger]^*\eta - \xi^\dagger \in N_C(x^\dagger)$$

with $N_C(x^\dagger)$ from (2.12) and if

$$\|F'[x^\dagger](x - x^\dagger)\| \leqslant \varphi(\|F(x) - F(x^\dagger)\|) \qquad \text{for all} \quad x \in C,$$

then theorem 3.1 yields a variational inequality (2.4), and based on this inequality proposition 2.3 provides the convergence rate

$$B_{\xi^\dagger}\left(x^\delta_{\alpha(\delta)}, x^\dagger\right) = \mathcal{O}(\varphi(\delta)) \qquad \text{as} \quad \delta \to 0$$

for an appropriate parameter choice $\alpha = \alpha(\delta)$.

4. Projected source conditions

In this section, we will show that the conditions of the form (2.11) that play a prominent role in our theory can be referred to as projected source conditions generalizing (2.6) to Banach spaces and sets of constraints C that fulfill the following assumption.

Assumption 4.1.

(a) *X is a reflexive, strictly convex and smooth Banach space.*
(b) *C is a non-empty, closed and convex subset of X.*

Under assumption 4.1, the *duality mapping* $J : X \to X^*$ for the Banach space X (as an analog to the Riesz isomorphism in the Hilbert space) defined as

$$J(x) := \{\xi \in X^* : \langle \xi, x \rangle = \|x\|^2 = \|\xi\|^2\} \subseteq X^* \tag{4.1}$$

is a well defined, bijective and unitary mapping, i.e. the set $J(x)$ is a singleton for all $x \in X$, and we have $J^{-1} = J_*$ (see, e.g., [28, proposition 47.19]), where J_* denotes the duality mapping on X^*. Moreover, the *metric projector* $P_C : X \to X$ is well defined and determines for all $\tilde{x} \in X$ a unique element

$$P_C(\tilde{x}) := \operatorname{argmin}_{x \in C} \|x - \tilde{x}\| \tag{4.2}$$

(see, e.g., [2, chapter 3, section 3.2]). Following [18], we have the equivalence

$$x^\dagger = P_C(\tilde{x}) \quad \Longleftrightarrow \quad 0 \in -J(\tilde{x} - x^\dagger) + N_C(x^\dagger).$$

Hence, under assumption 4.1, condition (2.11) can be rewritten as

$$\begin{aligned} F'[x^\dagger]^*\eta - \xi^\dagger \in N_C(x^\dagger) &\Longleftrightarrow J(J_*(F'[x^\dagger]^*\eta - \xi^\dagger)) \in N_C(x^\dagger) \\ &\Longleftrightarrow 0 \in -J(x^\dagger + J_*(F'[x^\dagger]^*\eta - \xi^\dagger) - x^\dagger) + N_C(x^\dagger) \\ &\Longleftrightarrow x^\dagger = P_C(x^\dagger + J_*(F'[x^\dagger]^*\eta - \xi^\dagger)). \end{aligned}$$

As this chain of equivalences shows, (2.11) attains the alternative form

$$x^\dagger = P_C(x^\dagger + J_*(F'[x^\dagger]^*\eta - \xi^\dagger)), \tag{4.3}$$

which is a projective version of (2.10) under the convexity constraints expressed by the set C.

Example 4.2. If X, Y are Hilbert spaces and C is a non-empty closed and convex subset of X, then assumption 4.1 is satisfied, and we can verify the projected source condition (4.3) for the most prominent version of Tikhonov regularization in Hilbert spaces. We identify $X^* = X$ by using the Riesz isomorphism J, and set $p = 2$ and $\Omega(x) := \frac{1}{2}\|x - \bar{x}\|^2$ with fixed $\bar{x} \in X$ in (2.3). That is, we consider the minimization problem

$$\|F(x) - y^\delta\|^2 + \frac{\alpha}{2}\|x - \bar{x}\|^2 \to \min, \qquad \text{subject to} \quad x \in C.$$

For this specific setting, the subdifferential of Ω at $x^\dagger$ is $\partial\Omega(x^\dagger) = \{\xi^\dagger\}$ with $\xi^\dagger = x^\dagger - \bar{x}$, and therefore with the Riesz isomorphism J and by setting $w := \eta$, the projected source condition (4.3) reduces to (2.6).

5. The case of linear forward operators

We first consider bounded linear operators $A = F : X \to Y$ with adjoint $A^* : Y^* \to X^*$. Theorem 3.1 will turn out to be an extension of the results derived in this section.

The proofs below are essentially based on separation of convex sets. Therefore, we state the following separation theorem, which is an immediate consequence of [5, theorem 2.1.2].

Lemma 5.1. *Let $E_1, E_2 \subseteq X \times \mathbb{R}$ be convex sets. If one of them has non-empty interior and the interior does not intersect with the other set, then $\xi \in X^*$ and $\tau \in \mathbb{R}$ exist with $(\xi, \tau) \neq (0, 0)$ such that*

$$\sup_{(x,t)\in E_1} (\langle \xi, x \rangle + \tau t) \leqslant \inf_{(x,t)\in E_2} (\langle \xi, x \rangle + \tau t).$$

With the help of this lemma, we now derive a projected source condition (2.11) from the variational inequality (3.2) with $\beta_1 = 1$ in the case of a linear operator $A = F$.

Theorem 5.2. *Assume* $N_{\mathcal{D}(\Omega)}(x^\dagger) \cap (-N_C(x^\dagger)) = \{0\}$ *and that at least one of the sets* $\mathcal{D}(\Omega)$ *or* C *has interior points. If there is some* $\beta_2 \geqslant 0$, *such that*

$$0 \leqslant \Omega(x) - \Omega(x^\dagger) + \beta_2 \|A(x - x^\dagger)\| \qquad \text{for all} \quad x \in C, \tag{5.1}$$

then there exist $\xi^\dagger \in \partial\Omega(x^\dagger)$ *and* $\eta \in Y^*$ *with* $\|\eta\| \leqslant \beta_2$ *such that*

$$A^*\eta - \xi^\dagger \in N_C(x^\dagger). \tag{5.2}$$

Proof. We apply lemma 5.1 to the sets

$$E_1 := \{(x, t) \in X \times \mathbb{R} : x \in C,\ t \leqslant \Omega(x^\dagger) - \Omega(x)\},$$
$$E_2 := \{(x, t) \in X \times \mathbb{R} : t \geqslant \beta_2 \|A(x - x^\dagger)\|\}.$$

To see that the assumptions of that lemma are satisfied, first note that $\operatorname{int} E_2 \neq \emptyset$. Further, we have $E_1 \cap (\operatorname{int} E_2) = \emptyset$ if we can show that $E_1 \cap E_2$ is a subset of the boundary of E_2. So let $(x, t) \in E_1 \cap E_2$ and set $(x_n, t_n) := \left(x, t - \frac{1}{n}\right)$ for $n \in \mathbb{N}$. Then $(x_n, t_n) \to (x, t)$ and using the definition of E_1 and inequality (5.1), we obtain

$$t_n = t - \frac{1}{n} < t \leqslant \Omega(x^\dagger) - \Omega(x) \leqslant \beta_2 \|A(x - x^\dagger)\| = \beta_2 \|A(x_n - x^\dagger)\|,$$

that is, $(x_n, t_n) \notin E_2$. In other words, (x, t) is indeed a boundary point of E_2.

Together with $(x^\dagger, 0) \in E_1 \cap E_2$, lemma 5.1 provides $\xi \in X^*$ and $\tau \in \mathbb{R}$ such that

$$\langle \xi, x - x^\dagger \rangle + \tau t \leqslant 0 \qquad \text{for all} \quad (x, t) \in E_1, \tag{5.3}$$

$$\langle \xi, x - x^\dagger \rangle + \tau t \geqslant 0 \qquad \text{for all} \quad (x, t) \in E_2. \tag{5.4}$$

If $\tau < 0$, then (5.4) implies $\langle -\frac{1}{\tau}\xi, x - x^\dagger \rangle \geqslant t$ for all $t \geqslant \beta_2 \|A(x - x^\dagger)\|$ and all $x \in X$, which is obviously not possible. In the case $\tau = 0$, inequality (5.4) gives $\langle \xi, x - x^\dagger \rangle \geqslant 0$ for all $x \in X$, and therefore $\xi = 0$. But this contradicts $(\xi, \tau) \neq (0, 0)$. Thus, $\tau > 0$ has to be true.

From (5.4) with $t := \beta_2 \|A(x - x^\dagger)\|$, we obtain for all $x \in X$

$$\left\langle -\tfrac{1}{\tau}\xi, x - x^\dagger \right\rangle \leqslant \beta_2 \|A(x - x^\dagger)\|.$$

Hence there is some $\eta \in Y^*$ such that $\frac{1}{\tau}\xi = A^*\eta$ and $\|\eta\| \leqslant \beta_2$ (see [24, lemma 8.21 and its proof]). Inequality (5.3) with $t := \Omega(x^\dagger) - \Omega(x)$ now yields $\Omega(x^\dagger) - \Omega(x) \leqslant \langle -A^*\eta, x - x^\dagger \rangle$ for all $x \in C$.

To obtain a subgradient of Ω at $x^\dagger$, we apply lemma 5.1 to the sets

$$\tilde{E}_1 := \{(x, t) \in X \times \mathbb{R} : t \leqslant \Omega(x^\dagger) - \Omega(x)\},$$
$$\tilde{E}_2 := \{(x, t) \in X \times \mathbb{R} : x \in C,\ t \geqslant \langle -A^*\eta, x - x^\dagger \rangle\}.$$

The assumptions of that lemma can be verified in a similar way as discussed for E_1, E_2 since $\operatorname{int} \tilde{E}_1 \neq \emptyset$ if $\operatorname{int} \mathcal{D}(\Omega) \neq \emptyset$ and $\operatorname{int} \tilde{E}_2 \neq \emptyset$ if $\operatorname{int} C \neq \emptyset$. With $(x^\dagger, 0) \in \tilde{E}_1 \cap \tilde{E}_2$, lemma 5.1 yields $\tilde{\xi} \in X^*$ and $\tilde{\tau} \in \mathbb{R}$ such that

$$\langle \tilde{\xi}, x - x^\dagger \rangle + \tilde{\tau} t \leqslant 0 \qquad \text{for all} \quad (x, t) \in \tilde{E}_1, \tag{5.5}$$

$$\langle \tilde{\xi}, x - x^\dagger \rangle + \tilde{\tau} t \geqslant 0 \qquad \text{for all} \quad (x, t) \in \tilde{E}_2. \tag{5.6}$$

If $\tilde{\tau} < 0$, then (5.6) implies $\left\langle -\frac{1}{\tilde{\tau}}\tilde{\xi}, x - x^\dagger \right\rangle \geqslant t$ for all $t \geqslant \langle -A^*\eta, x - x^\dagger \rangle$ and all $x \in C$, which is obviously not possible $\left(\text{because } \left\langle -\frac{1}{\tilde{\tau}}\tilde{\xi}, x - x^\dagger \right\rangle < \infty\right)$. In the case $\tilde{\tau} = 0$, inequality (5.5) gives $\langle \tilde{\xi}, x - x^\dagger \rangle \leqslant 0$ for all $x \in \mathcal{D}(\Omega)$ and (5.6) gives $\langle \tilde{\xi}, x - x^\dagger \rangle \geqslant 0$ for all $x \in C$. Thus $\xi \in N_{\mathcal{D}(\Omega)}(x^\dagger) \cap (-N_C(x^\dagger))$, which implies $\tilde{\xi} = 0$ by assumption. This contradicts $(\tilde{\xi}, \tilde{\tau}) \neq (0, 0)$. Therefore, $\tilde{\tau} > 0$ is true.

With $t := \Omega(x^\dagger) - \Omega(x)$ from (5.5) we obtain $\left\langle \frac{1}{\tau}\tilde{\xi}, x - x^\dagger \right\rangle \leqslant \Omega(x) - \Omega(x^\dagger)$ for all $x \in X$, that is, $\xi^\dagger := \frac{1}{\tau}\tilde{\xi} \in \partial\Omega(x^\dagger)$. Eventually, (5.6) with $t := \langle -A^*\eta, x - x^\dagger\rangle$ yields $\langle -\xi^\dagger, x - x^\dagger\rangle \leqslant \langle -A^*\eta, x - x^\dagger\rangle$ for all $x \in C$. Thus, we found $\xi^\dagger \in \partial\Omega(x^\dagger)$ and $\eta \in Y^*$ such that $A^*\eta - \xi^\dagger \in N_C(x^\dagger)$. □

As a by-product of this theorem, we can show that the constant β_1 in the variational inequality (3.2) plays only a minor role if $A = F$ is linear. In fact, if a variational inequality holds for one $\beta_1 \in [0, 1)$, then it holds for all $\beta_1 \in [0, 1)$.

Corollary 5.3. *Assume that $N_{\mathcal{D}(\Omega)}(x^\dagger) \cap (-N_C(x^\dagger)) = \{0\}$ and that at least one of the sets $\mathcal{D}(\Omega)$ or C has interior points. Further let $\beta_1 \in [0, 1)$ and $\beta_2 \geqslant 0$. Then*

$$0 \leqslant \Omega(x) - \Omega(x^\dagger) + \beta_2\|A(x - x^\dagger)\| \qquad \text{for all} \quad x \in C,$$

if and only if $\xi^\dagger \in \partial\Omega(x^\dagger)$ exists such that

$$\langle -\xi^\dagger, x - x^\dagger\rangle \leqslant \beta_1 B_{\xi^\dagger}(x, x^\dagger) + \beta_2\|A(x - x^\dagger)\| \qquad \text{for all} \quad x \in C.$$

Proof. Let the first inequality in the corollary be satisfied. Then theorem 5.2 provides $\xi^\dagger \in \partial\Omega(x^\dagger)$ and $\eta \in Y^*$ with $\|\eta\| \leqslant \beta_2$ such that $A^*\eta - \xi^\dagger \in N_C(x^\dagger)$. From the definition of $N_C(x^\dagger)$, we see that $\langle \xi^\dagger, x - x^\dagger\rangle + \langle A^*(-\eta), x - x^\dagger\rangle \geqslant 0$ and therefore $\langle \xi^\dagger, x - x^\dagger\rangle + \beta_2\|A(x - x^\dagger)\| \geqslant 0$ for all $x \in C$, which together with $B_{\xi^\dagger}(x, x^\dagger) \geqslant 0$ implies the second inequality in the corollary.

The reverse direction is trivially true since

$$\begin{aligned}\Omega(x) - \Omega(x^\dagger) + \beta_2\|A(x - x^\dagger)\| &= \langle \xi^\dagger, x - x^\dagger\rangle + B_{\xi^\dagger}(x, x^\dagger) + \beta_2\|A(x - x^\dagger)\| \\ &\geqslant \langle \xi^\dagger, x - x^\dagger\rangle + \beta_1 B_{\xi^\dagger}(x, x^\dagger) + \beta_2\|A(x - x^\dagger)\| \geqslant 0\end{aligned}$$

for all $x \in C$. □

The proof of the corollary implicitly proves the first part of theorem 3.1 in the case of a linear operator $A = F$. For nonlinear F, the proof is given in the next section.

6. Proof of the main theorem

Now we are ready to prove theorem 3.1 as our main result. At first we show how to reduce a variational inequality with the nonlinear operator F to a variational inequality with the bounded linear operator $F'[x^\dagger]$ introduced in assumption 2.1 (*d*).

Lemma 6.1. *If there are $\xi^\dagger \in \partial\Omega(x^\dagger)$, $\beta_1 \in [0, 1)$ and $\beta_2 \geqslant 0$ such that*

$$\langle -\xi^\dagger, x - x^\dagger\rangle \leqslant \beta_1 B_{\xi^\dagger}(x, x^\dagger) + \beta_2\|F(x) - F(x^\dagger)\| \qquad \text{for all} \quad x \in C, \tag{6.1}$$

then

$$\langle -\xi^\dagger, x - x^\dagger\rangle \leqslant \beta_1 B_{\xi^\dagger}(x, x^\dagger) + \beta_2\|F'[x^\dagger](x - x^\dagger)\| \qquad \text{for all} \quad x \in C. \tag{6.2}$$

Proof. For fixed $x \in C$ and all $t \in (0, 1]$, the given variational inequality (6.1) and the convexity of Ω imply

$$\begin{aligned}\frac{\beta_2}{t}\|F(x^\dagger + t(x - x^\dagger)) - F(x^\dagger)\| &\geqslant \frac{1}{t}\langle -\xi^\dagger, t(x - x^\dagger)\rangle - \beta_1 B_{\xi^\dagger}(x^\dagger + t(x - x^\dagger), x^\dagger) \\ &= (1 - \beta_1)\langle -\xi^\dagger, x - x^\dagger\rangle - \frac{\beta_1}{t}(\Omega((1 - t)x^\dagger + tx) - \Omega(x^\dagger)) \\ &\geqslant (1 - \beta_1)\langle -\xi^\dagger, x - x^\dagger\rangle - \beta_1(\Omega(x) - \Omega(x^\dagger)) \\ &= \langle -\xi^\dagger, x - x^\dagger\rangle - \beta_1 B_{\xi^\dagger}(x, x^\dagger).\end{aligned}$$

If we let $t \to +0$ we derive under assumption 2.1 (*d*)

$$\beta_2 \|F'[x^\dagger](x - x^\dagger)\| = \lim_{t\to+0} \frac{\beta_2}{t} \|F(x^\dagger + t(x - x^\dagger)) - F(x^\dagger)\|$$
$$\geqslant \langle -\xi^\dagger, x - x^\dagger\rangle - \beta_1 B_{\xi^\dagger}(x, x^\dagger)$$

for all $x \in C$. Therefore the variational inequality (6.2) is valid. □

Remark 6.2. We should mention here that the opposite direction in the lemma, from (6.2) back to (6.1), is in general not true even if a modified value of the constant β_2 can be accepted. Additional assumptions on the structure of nonlinearity of F have to be satisfied for that direction. The simplest form of such structural conditions is

$$\|F'[x^\dagger](x - x^\dagger)\| \leqslant c_1 \|F(x) - F(x^\dagger)\| \qquad \text{for all} \quad x \in C \tag{6.3}$$

with some constant $c_1 > 0$, which coincides with the nonlinearity condition (3.1) in theorem 3.1 if $\varphi(t) = c_1 t$. For example, such a condition is an implication of the *tangential cone condition*

$$\|F(x) - F(x^\dagger) - F'[x^\dagger](x - x^\dagger)\| \leqslant c_2 \|F(x) - F(x^\dagger)\| \qquad \text{for all} \quad x \in C$$

with some constant $c_2 > 0$, which plays a prominent role for obtaining convergence rates of iterative regularization methods; see [17] and [1].

The concept of a *degree of nonlinearity* presented for Hilbert spaces in [14] was extended to Banach spaces in [3, 12]. From the latter two papers, it becomes clear that a condition of type (6.3) is more powerful with respect to convergence rates than the weaker condition

$$\|F(x) - F(x^\dagger) - F'[x^\dagger](x - x^\dagger)\| \leqslant c_3 B_{\xi^\dagger}(x, x^\dagger) \qquad \text{for all} \quad x \in C \tag{6.4}$$

with some constant $c_3 \geqslant 0$. Nonlinearity conditions of type (6.4) were exploited in the papers [13] and [23]. If we replace (6.3) by (6.4), then a variational inequality (2.4) with $\varphi(t) = ct$, $c \geqslant 0$, can be proven, too, if there are $\xi^\dagger \in \partial\Omega(x^\dagger)$ and $\eta \in Y^*$ such that (2.11) is valid because of

$$\|F'[x^\dagger](x - x^\dagger)\| \leqslant \|F(x) - F(x^\dagger) - F'[x^\dagger](x - x^\dagger)\| + \|F(x) - F(x^\dagger)\|$$
$$\leqslant c_3 B_{\xi^\dagger}(x, x^\dagger) + \|F(x) - F(x^\dagger)\|,$$

but we have $\beta_1 \in [0, 1)$ only under the smallness condition $c_3\|\eta\| < 1$.

However, if we diminish the nonlinearity condition from (6.3) to (3.1) with some increasing and strictly concave function $\varphi : [0, \infty) \to [0, \infty)$ satisfying $\lim_{t\to+0} \varphi(t) = 0$, then we obtain only a weaker variational inequality (2.4) if there are $\xi^\dagger \in \partial\Omega(x^\dagger)$ and $\eta \in Y^*$ such that (2.11) is valid, but again without a smallness condition.

The proof of theorem 3.1 is now quite simple.

Proof of theorem 3.1. Assume that there are $\xi^\dagger \in \partial\Omega(x^\dagger)$ and $\eta \in Y^*$ such that the projected source condition (2.11) is satisfied. Then from the definition of $N_C(x^\dagger)$, we see that $\langle\xi^\dagger, x - x^\dagger\rangle + \langle F'[x^\dagger]^*(-\eta), x - x^\dagger\rangle \geqslant 0$ and therefore $\langle\xi^\dagger, x - x^\dagger\rangle + \|\eta\|\|F'[x^\dagger](x - x^\dagger)\| \geqslant 0$ for all $x \in C$, which together with $B_{\xi^\dagger}(x, x^\dagger) \geqslant 0$ and $\beta_1 \in [0, 1)$ implies

$$\langle -\xi^\dagger, x - x^\dagger\rangle \leqslant \beta_1 B_{\xi^\dagger}(x, x^\dagger) + \|\eta\|\|F'[x^\dagger](x - x^\dagger)\| \qquad \text{for all} \quad x \in C.$$

Taking into account the structural assumption (3.1), we derive that the variational inequality (2.4) is fulfilled for all $\beta_1 \in [0, 1)$ and all $\beta_2 \geqslant \|\eta\|$. Hence, the first assertion of theorem 3.1 is proven.

Now assume that the special case (3.2) of (2.4) with $\varphi(t) = t$ is satisfied for constants $\beta_1 \in [0, 1)$ and $\beta_2 \geqslant 0$. Then we have by lemma 6.1 that

$$\langle -\xi^\dagger, x - x^\dagger \rangle \leqslant \beta_1 B_{\xi^\dagger}(x, x^\dagger) + \beta_2 \|F'[x^\dagger](x - x^\dagger)\| \qquad \text{for all} \quad x \in C$$

and therefore

$$\begin{aligned}\Omega(x) - \Omega(x^\dagger) + \beta_2 \|F'[x^\dagger](x - x^\dagger)\| \\ &= \langle \xi^\dagger, x - x^\dagger \rangle + B_{\xi^\dagger}(x, x^\dagger) + \beta_2 \|F'[x^\dagger](x - x^\dagger)\| \\ &\geqslant \langle \xi^\dagger, x - x^\dagger \rangle + \beta_1 B_{\xi^\dagger}(x, x^\dagger) + \beta_2 \|F'[x^\dagger](x - x^\dagger)\| \geqslant 0\end{aligned}$$

for all $x \in C$. Thus, the second assertion of theorem 3.1 follows from theorem 5.2. □

Acknowledgment

This research was supported by the DFG under grant HO 1454/8-1.

References

[1] Bakushinsky A B and Kokurin M Yu 2004 *Iterative Methods for Approximate Solution of Inverse Problems* (Dordrecht: Springer)

[2] Barbu V and Precupanu T 1986 *Convexity and Optimization in Banach Spaces* (*Mathematics and its Applications (East European Series)* vol 10) 2nd English edn (Bucharest: Editura Academiei, D Reidel Publishing)

[3] Boţ R I and Hofmann B 2010 An extension of the variational inequality approach for obtaining convergence rates in regularization of nonlinear ill-posed problems *J. Integral Equ. Appl.* **22** 369–92

[4] Bonesky T, Kazimierski K S, Maass P, Schöpfer F and Schuster T 2008 Minimization of Tikhonov functionals in Banach spaces *Abstr. Appl. Anal.* **2008** 192679

[5] Boţ R I, Grad S-M and Wanka G 2009 *Duality in Vector Opimization* (*Vector Optimization*) (Berlin: Springer)

[6] Burger M and Osher S 2004 Convergence rates of convex variational regularization *Inverse Problems* **20** 1411–21

[7] Chavent G and Kunisch K 1994 Convergence of Tikhonov regularization for constrained ill-posed inverse problems *Inverse Problems* **10** 63–76

[8] Eggermont P P B, LaRiccia V N and Nashed M Z 2009 On weakly bounded noise in ill-posed problems *Inverse Problems* **25** 115018

[9] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (*Mathematics and its Applications* vol 375) (Dordrecht: Kluwer)

[10] Flemming J 2011 Solution smoothness of ill-posed equations in Hilbert spaces: four concepts and their cross connections *Appl. Anal.* **90** at press (doi:10.1080/00036811.2011.563736)

[11] Flemming J and Hofmann B 2010 A new approach to source conditions in regularization with general residual term *Numer. Funct. Anal. Optim.* **31** 254–84

[12] Hein T and Hofmann B 2009 Approximate source conditions for nonlinear ill-posed problems—chances and limitations *Inverse Problems* **25** 035003

[13] Hofmann B, Kaltenbacher B, Pöschl C and Scherzer O 2007 A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators *Inverse Problems* **23** 987–1010

[14] Hofmann B and Scherzer O 1994 Factors influencing the ill-posedness of nonlinear problems *Inverse Problems* **10** 1277–97

[15] Hofmann B and Yamamoto M 2010 On the interplay of source conditions and variational inequalities for nonlinear ill-posed problems *Appl. Anal.* **89** 1705–27

[16] Jačimovič M and Krnič I 1999 On some classes of regularization methods for minimization problem of quadratic functional on a half-space *Hokkaido Math. J.* **28** 57–69

[17] Kaltenbacher B, Neubauer A and Scherzer O 2008 *Iterative Regularization Methods for Nonlinear Ill-Posed Problems* (*Radon Series on Computational and Applied Mathematics* vol 6) (Berlin: Walter de Gruyter)

[18] Kien B T 2002 The normalized duality mapping and two related characteristic properties of a uniformly convex Banch space *Acta Math. Vietnamica* **27** 53–67

[19] Krnič I and Potapov M M 2001 Projective sourcewise representability of normal solutions to linear equations on convex sets *Comput. Math. Math. Phys.* **41** 1251–8

[20] Morozov V A 1997 Generalized 'source-likeness' and the rate of convergence of regularized solutions *Fundam. Prikl. Mat.* **3** 171–7

[21] Neubauer A 1988 Tikhonov-regularization of ill-posed linear operator equations on closed convex sets *J. Approx. Theory* **53** 304–20

[22] Neubauer A 1989 Tikhonov regularisation for nonlinear ill-posed problems: optimal convergence rates and finite-dimensional approximation *Inverse Problems* **5** 541–57

[23] Resmerita E and Scherzer O 2006 Error estimates for non-quadratic regularization and the relation to enhancement *Inverse Problems* **22** 801–14

[24] Scherzer O, Grasmair M, Grossauer H, Haltmeier M and Lenzen F 2009 *Variational Methods in Imaging* (*Applied Mathematical Sciences* vol 167) (New York: Springer)

[25] Schöpfer F, Louis A K and Schuster T 2006 Nonlinear iterative methods for linear ill-posed problems in Banach spaces *Inverse Problems* **22** 311–29

[26] Tikhonov A N, Goncharsky A V, Stepanov V V and Yagola A G 1995 *Numerical Methods for the Solution of Ill-Posed Problems* (Dordrecht: Kluwer) Translated from the 1990 Russian original

[27] Vasil'ev F P 1981 *Metody Resheniya Ekstremalnykh Zadach (Methods for Solving Extremal Problems)* (Moscow: Nauka)

[28] Zeidler E 1985 *Nonlinear Functional Analysis and its Applications: III. Variational Methods and Optimization* (New York: Springer)