# On the interplay of source conditions and variational inequalities for nonlinear ill-posed problems

Bernd Hofmann *and Masahiro Yamamoto†

Updated version – April 12, 2010

*Dedicated to Professor Lothar von Wolfersdorf on the occasion of his* $75^{\text{th}}$ *birthday*

**Abstract**

In the last years convergence rates results for Tikhonov regularization of nonlinear ill-posed problems in Banach spaces have been published, where the classical concept of source conditions was replaced with variational inequalities holding on some level sets. This essentially advanced the analysis of non-smooth situations with respect to forward operators and solutions. In fact, such variational inequalities combine both structural conditions on the nonlinearity of the operator and smoothness properties of the solution. Varying exponents in the variational inequalities correspond to different levels of convergence rates. In this paper, we discuss the range of occurring exponents in the Banach space setting. To lighten the cross-connections between generalized source conditions, degree of nonlinearity of the forward operator and associated variational inequalities we study the Hilbert space situation and even prove some converse result for linear operators. Finally, we outline some aspects for the interplay of variational regularization and conditional stability estimates for partial differential equations. As an example, we apply the theory to a specific parameter identification problem for a parabolic equation.

**MSC2000 subject classification:** 47J06, 65J20, 47J20, 47A52

**Keywords:** Nonlinear ill-posed problems, regularization, variational inequalities, source conditions, degree of nonlinearity, convergence rates, conditional stability, inverse P.D.E. problems.

*Department of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany. Email: hofmannb@mathematik.tu-chemnitz.de. Corresponding author.

†Graduate School of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba, Tokyo, 153-8914 Japan. Email: myama@ms.u-tokyo.ac.jp.

# 1 Introduction

After turn of the millennium there seems to be a substantial progress in regularization theory for the stable approximate solution of ill-posed inverse problems. On the one hand, partially motivated by specific applications in imaging and by a growing interest in sparsity of solutions as well as in new types of stabilizing terms in variational regularization, the Banach space treatment of linear and nonlinear operator equations and occurring difficulties in this context came into the focus of recent papers and books. On the other hand, Bregman distances for measuring the regularization error and variational inequalities for replacing the standard form of source conditions offer now good prospects for proving convergence rates results also for non-smooth situations with respect to solution and forward operator. For recent results we refer to the monograph [29] and in an exemplary manner to the papers [3, 4, 10, 12, 13, 14, 19, 21, 22, 26, 27, 28, 30, 31, 33] as well as to the thesis [24].

This paper is devoted to the utility of variational inequalities combining both structural conditions on the nonlinearity of the operator and smoothness properties of the solution. Varying exponents in the variational inequalities correspond to different levels of convergence rates. We are going to discuss the range of occurring exponents in the Banach space setting and the interplay of general source conditions and variational inequalities in Banach and Hilbert spaces. The paper is organized as follows: In Section 2 we describe the Tikhonov type regularization for the stable approximate solution of nonlinear ill-posed operator equations in a Banach space setting under basic assumptions which follow the corresponding assumptions of the papers [13, 14]. As in the previous papers the focus is again on level sets for the Tikhonov sum functional, and the majority of conditions under consideration have to hold on such sets. Section 3 summarizes propositions on convergence and convergence rates under variational inequalities. Moreover, we recall the concept of a degree of nonlinearity for characterizing the local structural nonlinearity conditions in the solution point. The range of occurring exponents in the variational inequalities is discussed in Section 4. Here the forward operator and the stabilizing functional are assumed to be Gâteaux differentiable. We distinguish three typical cases of exponents and make assertions for all of them. In Section 5 the concept of the degree of nonlinearity will be modified, since weaker norms with respect to the first order Taylor remainder extend the applicability of the theory to a wider class of problems. Some open questions cannot be answered currently for the general Banach space setting. Therefore we restrict our considerations in the concluding Section 6 to Hilbert spaces situations. Under that restriction we are able to formulate assertions on the interplay of variational inequalities and Hölder source conditions with fractional exponents including some converse result for the subcase of linear operators. For inverse problems in partial differential equations conditional stability estimates are frequently more appropriate than estimates for the Taylor remainder. Therefore, we outline some cross-connections between variational regularization and conditional stability estimates for partial differential equations. Estimates of this type are currently in the focus of numerous papers with respect to various methods such as Carleman estimates. We refer, e.g., to the monograph [16]. In our concluding section, as an example we apply the theory to a specific parameter identification problem for a parabolic equation.

## 2 Problem, notation, and basic assumptions

We are going to study ill-posed operator equations

$$F(u) = v \tag{2.1}$$

expressing inverse problems with an in general nonlinear forward operator $F : \mathcal{D}(F) \subseteq U \to V$ possessing the domain $\mathcal{D}(F)$ and mapping between normed real linear spaces $U$ and $V$ with norms $\|\cdot\|_U$ and $\|\cdot\|_V$, respectively. Based on noisy data $v^\delta$ of the exact right-hand side $v = v^0 \in F(\mathcal{D}(F))$ with

$$\|v^\delta - v\|_V \le \delta \tag{2.2}$$

and noise level $\delta \ge 0$ we consider stable approximate solutions $u_\alpha^\delta$ as minimizers over $U$ of the Tikhonov type functional

$$T_\alpha^{v^\delta}(u) := \|F(u) - v^\delta\|_V^p + \alpha\,\Omega(u) \tag{2.3}$$

with a prescribed norm exponent

$$1 < p < \infty$$

and a regularization parameter $\alpha > 0$. In this context, let $\Omega : U \to [0, +\infty]$ be a stabilizing functional with

$$\mathcal{D}(\Omega) := \{u \in U : \Omega(u) \neq +\infty\} \neq \emptyset$$

and set $T_\alpha^{v^\delta}(u) = \infty$ if $u \notin \mathcal{D}(F)$. For studies on residual terms $\mathcal{S}(F(u), v^\delta)$ in (2.3) replacing $\|F(u) - v^\delta\|_V^p$ we refer to [24] and [9], where the latter reference also makes assertions on $p < 1$ in the norm case.

Throughout this paper we make the following assumptions:

**Assumption 2.1**

1. *$U$ and $V$ are reflexive Banach spaces with duals $U^*$ and $V^*$, respectively. In $U$ and $V$ we consider in addition to the norm convergence the associated weak convergence. That means in $U$*

$$u_k \rightharpoonup u \quad \Longleftrightarrow \quad \langle f, u_k\rangle_{U^*,U} \to \langle f, u\rangle_{U^*,U} \;\; \forall f \in U^*$$

   *for the dual pairing $\langle\cdot,\cdot\rangle_{U^*,U}$ with respect to $U^*$ and $U$. The weak convergence in $V$ is defined in an analog manner.*

2. *$F : \mathcal{D}(F) \subseteq U \to V$ is weakly-weakly sequentially continuous and $\mathcal{D}(F)$ is weakly sequentially closed, i.e.,*

$$u_k \rightharpoonup u \quad \text{in} \quad U \quad \text{with} \quad u_k \in \mathcal{D}(F) \quad \Longrightarrow \quad u \in \mathcal{D}(F) \quad \text{and} \quad F(u_k) \rightharpoonup F(u) \quad \text{in} \quad V.$$

3. *The functional $\Omega$ is convex and weakly sequentially lower semi-continuous.*

4. *The domain $\mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(\Omega)$ is non-empty.*

5. *For every* $\alpha > 0$, $c \geq 0$, *and for the exact right-hand side* $v = v^0$ *of (2.1), the sets*

$$\mathcal{M}_\alpha^v(c) := \{u \in \mathcal{D} : T_\alpha^v(u) \leq c\} \tag{2.4}$$

*are weakly sequentially pre-compact in the following sense: every sequence* $\{u_k\}_{k=1}^\infty$ *in* $\mathcal{M}_\alpha^v(c)$ *has a subsequence, which is weakly convergent in* $U$ *to some element from* $U$.

Under the stated assumptions existence and stability of regularized solutions $u_\alpha^\delta$ can be shown (cf. [14, §3] and [29, Theorems 3.22 and 3.23]).

In the Banach space theory of Tikhonov type regularization methods, regularization errors are frequently measured, for the convex functional $\Omega$ with subdifferential $\partial\Omega$, by means of Bregman distances

$$D_\xi(\tilde{u}, u) := \Omega(\tilde{u}) - \Omega(u) - \langle \xi, \tilde{u} - u \rangle_{U^*,U} \,, \quad \tilde{u} \in \mathcal{D}(\Omega) \subseteq U \,,$$

at $u \in \mathcal{D}(\Omega) \subseteq U$ and $\xi \in \partial\Omega(u) \subseteq U^*$. The set

$$\mathcal{D}_B(\Omega) := \{u \in \mathcal{D}(\Omega) : \partial\Omega(u) \neq \emptyset\}$$

is called Bregman domain. For more details see, e.g., [29, Lemmas 3.16 and 3.17].

An element $u^\dagger \in \mathcal{D}$ is called an $\Omega$-*minimizing solution* to (2.1) if

$$\Omega(u^\dagger) = \min\{\Omega(u) : F(u) = v,\ u \in \mathcal{D}\} < \infty \,.$$

Such $\Omega$-minimizing solutions exist under Assumption 2.1 if (2.1) has a solution $u \in \mathcal{D}$ (see [29, Theorem 3.25]), and by [29, Theorem 3.26]

# 3 Convergence, convergence rates, the degree of nonlinearity, and variational inequalities

As the following proposition shows, all regularized solutions associated with data possessing a sufficiently small noise level $\delta$ belong to a common weakly pre-compact level set of type $\mathcal{M}_\alpha^v(c)$ whenever the regularization parameters $\alpha = \alpha(\delta)$ are chosen such that weak convergence to $\Omega$-minimizing solutions $u^\dagger$ is enforced.

**Proposition 3.1** *Consider an a priori choice* $\alpha = \alpha(\delta) > 0$, $0 < \delta < \infty$, *for the regularization parameter in (2.3) depending on the noise level* $\delta$ *such that*

$$\alpha(\delta) \to 0 \quad \text{and} \quad \frac{\delta^p}{\alpha(\delta)} \to 0, \quad \text{as} \quad \delta \to 0. \tag{3.1}$$

*Provided that (2.1) has a solution* $u \in \mathcal{D}$ *then under Assumption 2.1 every sequence* $\{u_n\}_{n=1}^\infty := \{u_{\alpha(\delta_n)}^{\delta_n}\}_{n=1}^\infty$ *of regularized solutions corresponding to a sequence* $\{v^{\delta_n}\}_{n=1}^\infty$ *of data with* $\lim_{n\to\infty} \delta_n = 0$ *has a subsequence* $\{u_{n_k}\}_{k=1}^\infty$, *which is weakly convergent in* $U$, *i.e.* $u_{n_k} \rightharpoonup u^\dagger$, *and its limit* $u^\dagger$ *is an* $\Omega$-*minimizing solution of (2.1) with* $\Omega(u^\dagger) = \lim_{k\to\infty} \Omega(u_{n_k})$.

*For given $\alpha_{max} > 0$ let $u^\dagger$ denote an $\Omega$-minimizing solution of (2.1). If we set*

$$\rho := 2^{p-1}\alpha_{max}(1+\Omega(u^\dagger))\,, \tag{3.2}$$

*then we have $u^\dagger \in \mathcal{M}^v_{\alpha_{max}}(\rho)$ and there exists some $\delta_{max} > 0$ such that*

$$u^\delta_{\alpha(\delta)} \in \mathcal{M}^v_{\alpha_{max}}(\rho) \qquad \textit{for all} \qquad 0 \le \delta \le \delta_{max}\,. \tag{3.3}$$

**Proof:** The first part of the proposition concerning convergence replicates only the result of [29, Theorem 3.26] and we refer to the proof ibidem. The second part can be proven as follows: Owing to (3.1) there exists some $\delta_{max} > 0$ such that $\alpha(\delta) \le \alpha_{max}$ and $\frac{\delta^p}{\alpha(\delta)} \le \frac{1}{2}$ for all $0 < \delta \le \delta_{max}$. Then for such $\delta$, by writing for simplicity $\alpha$ instead of $\alpha(\delta)$, we have with $(a+b)^p \le 2^{p-1}(a^p+b^p)$ $(a,b\ge 0,\ p>1)$ the estimate

$$T^v_{\alpha_{max}}(u^\delta_\alpha) \le 2^{p-1}\left[\left\|F(u^\delta_\alpha)-v^\delta\right\|^p_V + \delta^p + \alpha_{max}\Omega(u^\delta_\alpha)\right]$$

$$= 2^{p-1}\left[\left\|F(u^\delta_\alpha)-v^\delta\right\|^p_V + \alpha\Omega(u^\delta_\alpha) + (\alpha_{max}-\alpha)\Omega(u^\delta_\alpha) + \delta^p\right]$$

$$\le 2^{p-1}\left[T^{v^\delta}_\alpha(u^\dagger) + (\alpha_{max}-\alpha)\Omega(u^\delta_\alpha) + \delta^p\right] \le 2^{p-1}\left[\delta^p + \alpha\Omega(u^\dagger) + (\alpha_{max}-\alpha)\Omega(u^\delta_\alpha) + \delta^p\right].$$

From $T^{v^\delta}_\alpha(u^\delta_\alpha) \le T^{v^\delta}_\alpha(u^\dagger)$ $(\alpha > 0)$ we obtain $\Omega(u^\delta_\alpha) \le \frac{\delta^p}{\alpha} + \Omega(u^\dagger)$ and with $\frac{\alpha_{max}}{\alpha} \ge 1$, $\frac{\delta^p}{\alpha} \le \frac{1}{2}$ this yields

$$T^v_{\alpha_{max}}(u^\delta_\alpha) \le 2^{p-1}\left[\delta^p + \alpha_{max}\frac{\delta^p}{\alpha} + \alpha_{max}\Omega(u^\dagger)\right] \le 2^{p-1}\left[2\alpha_{max}\frac{\delta^p}{\alpha} + \alpha_{max}\Omega(u^\dagger)\right] \le \rho$$

and hence proves (3.3). Evidently, it holds $T^v_{\alpha_{max}}(u^\dagger) = \alpha_{max}\Omega(u^\dagger) \le 2^{p-1}\alpha_{max}\Omega(u^\dagger)$ for all $p > 1$, which implies $u^\dagger \in \mathcal{M}^v_{\alpha_{max}}(\rho)$ and completes the proof. □

Proposition 3.1 makes only assertions on weakly convergent sequences of regularized solutions. However, the convergence rates results presented below will imply the strong convergence of such sequences. For further results on strong convergence we refer, for example, to Proposition 3.32 in [29].

For the analysis of nonlinear problems both the smoothness of $\Omega$-minimizing solutions $u^\dagger$ and the smoothness of the forward operator $F$ in a neighbourhood of $u^\dagger$ are essential ingredients. In this context, the term 'smoothness' has to be considered in a very general sense. With respect to the operator we recall the concept of a degree of nonlinearity from [13, Definition 2.5] which represents a Banach space update of Definition 1 from [15].

**Definition 3.2** *Let $c_1, c_2 \ge 0$ and $c_1 + c_2 > 0$. We define $F$ to be nonlinear of degree $(c_1, c_2)$ for the Bregman distance $D_\xi(\cdot, u^\dagger)$ of $\Omega$ at a solution $u^\dagger \in \mathcal{D}_B(\Omega) \subseteq U$ of (2.1) with $\xi \in \partial\Omega(u^\dagger) \subseteq U^*$ if there is a constant $K > 0$ such that*

$$\left\|F(u)-F(u^\dagger)-F'(u^\dagger)(u-u^\dagger)\right\|_V \le K\left\|F(u)-F(u^\dagger)\right\|^{c_1}_V D_\xi(u,u^\dagger)^{c_2} \tag{3.4}$$

*for all $u \in \mathcal{M}^v_{\alpha_{max}}(\rho)$.*

In recent publications the distinguished role of variational inequalities

$$\langle \xi, u^\dagger - u \rangle_{U^*,U} \le \beta_1 D_\xi(u,u^\dagger) + \beta_2 \|F(u) - F(u^\dagger)\|_V^\kappa \qquad \text{for all} \quad u \in \mathcal{M}^v_{\alpha_{max}}(\rho) \quad (3.5)$$

with some $\xi \in \partial\Omega(u^\dagger)$, two multipliers $0 \le \beta_1 < 1$, $\beta_2 \ge 0$ and an exponent $\kappa > 0$ for obtaining convergence rates was elaborated. The subsequent proposition outlines the chances of such variational inequalities for ensuring convergence rates in Tikhonov type regularization. Here we summarize convergence rates results from [13], [14], and [29, Section 3.2].

**Proposition 3.3** *Assume that $F, \Omega, \mathcal{D}, U$ and $V$ satisfy the Assumption 2.1 and that there is an $\Omega$-minimizing solution from the Bregman domain $u^\dagger \in \mathcal{D}_B(\Omega)$. If there exist an element $\xi \in \partial\Omega(u^\dagger)$ and constants $0 \le \beta_1 < 1$, $\beta_2 \ge 0$, and $0 < \kappa \le 1$ such that the variational inequality (3.5) holds with $\rho$ from (3.2), then we have the convergence rate*

$$D_\xi(u^\delta_{\alpha(\delta)}, u^\dagger) = \mathcal{O}(\delta^\kappa) \quad \text{as} \quad \delta \to 0 \qquad (3.6)$$

*for an a priori parameter choice $\alpha(\delta) \asymp \delta^{p-\kappa}$.*

**Proof:** We write again for simplicity $\alpha$ instead of $\alpha(\delta)$ and note that the parameter choice rule $\alpha \asymp \delta^{p-\kappa}$ satisfies the condition (3.1) with the consequence that Proposition 3.1 is applicable. Then by using $T^{v^\delta}_\alpha(u^\delta_\alpha) \le T^{v^\delta}_\alpha(u^\dagger)$, (2.2), and the definition of the Bregman distance we can estimate as follows:

$$\|F(u^\delta_\alpha) - v^\delta\|_V^p + \alpha D_\xi(u^\delta_\alpha, u^\dagger) \le \delta^p + \alpha\left(\Omega(u^\dagger) - \Omega(u^\delta_\alpha) + D_\xi(u^\delta_\alpha, u^\dagger)\right) . \qquad (3.7)$$

Moreover, by exploiting the inequality $(a+b)^\kappa \le a^\kappa + b^\kappa$ $(a, b > 0,\ 0 < \kappa \le 1)$ because of (3.3) we obtain from the variational inequality (3.5) that

$$\begin{aligned} \Omega(u^\dagger) - \Omega(u^\delta_\alpha) + D_\xi(u^\delta_\alpha, u^\dagger) &= -\langle \xi, u^\delta_\alpha - u^\dagger \rangle_{U^*,U} \\ &\le \beta_1\, D_\xi(u^\delta_\alpha, u^\dagger) + \beta_2 \|F(u^\delta_\alpha) - F(u^\dagger)\|_V^\kappa \\ &\le \beta_1\, D_\xi(u^\delta_\alpha, u^\dagger) + \beta_2 \left(\|F(u^\delta_\alpha) - v^\delta\|_V^\kappa + \delta^\kappa\right) . \end{aligned}$$

Therefore from (3.7) it follows that

$$\|F(u^\delta_\alpha) - v^\delta\|_V^p + \alpha D_\xi(u^\delta_\alpha, u^\dagger) \le \delta^p + \alpha\left(\beta_1 D_\xi(u^\delta_\alpha, u^\dagger) + \beta_2\left(\|F(u^\delta_\alpha) - v^\delta\|_V^\kappa + \delta^\kappa\right)\right). \quad (3.8)$$

Using the variant

$$a\,b \le \varepsilon\, a^{p_1} + \frac{b^{p_2}}{(\varepsilon\, p_1)^{p_2/p_1} p_2} \qquad (a, b \ge 0,\ \ \varepsilon > 0\ \ p_1, p_2 > 1 \ \text{ with } \ \frac{1}{p_1} + \frac{1}{p_1} = 1) \qquad (3.9)$$

of Young's inequality twice with $\varepsilon := 1$, $p_1 := p/\kappa$, $p_2 := p/(p-\kappa)$ and $b := \alpha\beta_2$, on the one hand with $a := \|F(u^\delta_\alpha) - u^\dagger\|_V^\kappa$ and on the other hand with $a := \delta^\kappa$, the inequality

$$\alpha D_\xi(u^\delta_\alpha, u^\dagger) \le 2\delta^p + \alpha\beta_1 D_\xi(u^\delta_\alpha, u^\dagger) + \frac{2(p-\kappa)}{(p/\kappa)^{\kappa/(p-\kappa)}\, p}(\alpha\beta_2)^{p/(p-\kappa)}$$

follows from (3.8). Because of $0 \le \beta_1 < 1$ this provides us with the estimate

$$D_\xi(u_\alpha^\delta, u^\dagger) \le \frac{2\delta^p + \frac{2(p-\kappa)}{(p/\kappa)^{\kappa/(p-\kappa)}\, p}\,(\alpha\,\beta_2)^{p/(p-\kappa)}}{\alpha\,(1-\beta_1)} \tag{3.10}$$

for sufficiently small $\delta > 0$, which yields (3.6) for the a priori parameter choice $\alpha \asymp \delta^{p-\kappa}$ and proves the proposition. As a by-product from formula (3.10) we obtain for the case $\delta = 0$ of noiseless data the corresponding estimate

$$D_\xi(u_\alpha^0, u^\dagger) \le \hat{C}\,\alpha^{\frac{\kappa}{p-\kappa}} \tag{3.11}$$

with some constant $\hat{C} > 0$. □

**Remark 3.4** The Proposition 3.3 shows the *formidable capability of variational inequalities* (3.5) *for obtaining convergence rates* without any additional requirements on the solution smoothness and on the nonlinearity structure of the forward operator. In this sense, the validity of such variational inequality (3.5) on the associated level set embodies an advantageous combination of properties on $u^\dagger$ and $F$ in a neighbourhood of $u^\dagger$. Necessary and sufficient conditions for (3.5) are given in the literature only in a fragmented manner, mostly expressing the interplay with classical source conditions.

In the next section we discuss the limited variability of exponents $\kappa > 0$ in (3.5).

# 4 A case distinction for the exponent in the variational inequality

We specify the general Assumption 2.1 to Assumption 4.1 by additional requirements for local use in this section.

**Assumption 4.1**

1. *$F, \Omega, \mathcal{D}, U$ and $V$ satisfy the Assumption 2.1.*
2. *Let $u^\dagger \in \mathcal{D}$ be an $\Omega$-minimizing solution of (2.1).*
3. *The operator $F$ is Gâteaux differentiable in $u^\dagger$ with the Gâteaux derivative $F'(u^\dagger) \in \mathcal{L}(U, V)$.*
4. *The functional $\Omega$ is Gâteaux differentiable in $u^\dagger$ with the Gâteaux derivative $\xi = \Omega'(u^\dagger) \in U^*$, i.e., the subdifferential $\partial\Omega(u^\dagger) = \{\xi\}$ is a singleton.*

**Remark 4.2** The Gâteaux differentiability of $F$ and $\Omega$ in $u^\dagger$ implies that $u^\dagger$ belongs to core($\mathcal{D}$), the algebraic interior of $\mathcal{D}$. Not in all cases the algebraic interior core($\mathcal{D}$) and its subset int($\mathcal{D}$), the set of inner points, do coincide. In such cases $u^\dagger$ need not be an inner point of the domains $\mathcal{D}(F)$ and $\mathcal{D}(\Omega)$.

**Case $\kappa > 1$:**

The following proposition shows that exponents $\kappa > 1$ in the variational inequality (3.5) under Assumption 4.1 in principle cannot occur.

**Proposition 4.3** *Under the Assumption 4.1 the variational inequality (3.5) cannot hold with $\xi = \Omega'(u^\dagger) \neq 0$ and multipliers $\beta_1, \beta_2 \geq 0$ whenever $\kappa > 1$.*

**Proof:** To prove the proposition we assume that the variational inequality (3.5) holds for $\xi = \Omega'(u^\dagger) \neq 0$ and some $\kappa > 1$ with multipliers $\beta_1, \beta_2 \geq 0$ and for all $u \in \mathcal{M}^v_{\alpha_{max}}(\rho)$. Then there is an element $u_\xi \in U$ with $\langle \xi, u_\xi \rangle_{U^*,U} > 0$ and some $t_0 > 0$ such that because of $u^\dagger \in \mathrm{core}(\mathcal{D})$ we have $u^\dagger - t u_\xi \in \mathcal{M}^v_{\alpha_{max}}(\rho)$ for all $0 \leq t \leq t_0$. Hence we have for all $0 < t \leq t_0$

$$0 < \langle \xi, t u_\xi \rangle_{U^*,U} \leq \beta_1 D_\xi(u^\dagger - t u_\xi, u^\dagger) \,+\, \beta_2 \big\| F(u^\dagger - t u_\xi) - F(u^\dagger) \big\|_V^\kappa$$

and dividing by $t > 0$

$$\big\langle \xi, u_\xi \big\rangle_{U^*,U} \leq \beta_1 \left[ \frac{\Omega(u^\dagger - t u_\xi) - \Omega(u^\dagger)}{t} + \big\langle \xi, u_\xi \big\rangle_{U^*,U} \right] + \beta_2 \left\| \frac{F(u^\dagger - t u_\xi) - F(u^\dagger)}{t} \right\|_V^\kappa t^{\kappa - 1}. \tag{4.1}$$

The left-hand side of inequality (4.1) is a positive constant. The right-hand side, however, tends to zero as $t \to 0$, since we have the limit conditions

$$\lim_{t\to 0} \Big\| \frac{\Omega(u^\dagger - t u_\xi) - \Omega(u^\dagger)}{t} + \big\langle \xi, u_\xi \big\rangle_{U^*,U} \Big\| = 0 \quad \text{and} \quad \lim_{t\to 0} \left\| \frac{F(u^\dagger - t u_\xi) - F(u^\dagger)}{t} \right\|_V = \|F'(u^\dagger) u_\xi\|_V < \infty\,,$$

because of the Gâteaux-differentiability of $F$ and $\Omega$ in $u^\dagger$. This contradicts the assumption and proves the proposition. □

**Case $\kappa = 1$:**

As the next proposition shows the variational inequality (3.5) is closely connected with the source condition $\xi \in \mathcal{R}(F'(u^\dagger)^*)$, where $\mathcal{R}(A)$ denotes the range of a linear operator $A$. The assertion a) of Proposition 4.4 repeats the Proposition 3.38 from [29], but reflects in contrast to the original the fact that the proof ibidem does not need the condition $\beta_1 < 1$. Note that the proof given there is similar to the proof of Proposition 4.3 presented above. On the other hand, for the assertion b) of Proposition 4.4 and its proof we refer to Proposition 3.35 in [29].

**Proposition 4.4** *Under the Assumption 4.1 the following two assertions hold:*

*a) The validity of a variational inequality*

$$\big\langle \xi, u^\dagger - u \big\rangle_{U^*,U} \leq \beta_1 D_\xi(u, u^\dagger) \,+\, \beta_2 \big\| F(u) - F(u^\dagger) \big\|_V \quad \textit{for all} \quad u \in \mathcal{M}^v_{\alpha_{max}}(\rho) \tag{4.2}$$

*for $\xi = \Omega'(u^\dagger)$ and two multipliers $\beta_1, \beta_2 \geq 0$ implies the source condition*

$$\xi \;=\; F'(u^\dagger)^*\, w, \qquad w \in V^*. \tag{4.3}$$

*b) Let $F$ be nonlinear of degree $(0,1)$ for the Bregman distance $D_\xi(\cdot, u^\dagger)$ of $\Omega$ at $u^\dagger$, i.e., we have*

$$\left\|F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)\right\|_V \le K\, D_\xi(u, u^\dagger) \tag{4.4}$$

*for a constant $K > 0$ and all $u \in \mathcal{M}^v_{\alpha_{max}}(\rho)$. Then the source condition (4.3) together with the smallness condition*

$$K \,\|w\|_{V^*} < 1 \tag{4.5}$$

*imply the validity of a variational inequality (4.2) with $\xi = \Omega'(u^\dagger)$ and multipliers $0 \le \beta_1 = K\|w\|_{V^*} < 1,\ \beta_2 = \|w\|_{V^*} \ge 0$.*

**Case $0 < \kappa \le 1$:**

The following proposition extends the result b) of Proposition 4.4 to a wider class of degrees of nonlinearity. The particular case $\kappa = 1$ discussed above occurs here only for the complementary situation $c_1 > 0$.

**Proposition 4.5** *Under the Assumption 4.1 let $F$ be nonlinear of degree $(c_1, c_2)$ with $0 < c_1 \le 1$, $0 \le c_2 < 1$, $c_1 + c_2 \le 1$ for the Bregman distance $D_\xi(\cdot, u^\dagger)$ of $\Omega$ at $u^\dagger$, i.e., we have*

$$\left\|F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)\right\|_V \le K\, \|F(u) - F(u^\dagger)\|_V^{c_1}\, D_\xi(u, u^\dagger)^{c_2} \tag{4.6}$$

*for a constant $K > 0$ and all $u \in \mathcal{M}^v_{\alpha_{max}}(\rho)$. Then the source condition (4.3) without any additional condition implies the validity of a variational inequality (3.5) with*

$$\kappa = \frac{c_1}{1 - c_2}\,, \tag{4.7}$$

*$\xi = \Omega'(u^\dagger)$ and multipliers $0 \le \beta_1 < 1,\ \beta_2 \ge 0$.*

**Proof:** We can estimate for $u \in \mathcal{M}^v_{\alpha_{max}}(\rho)$

$$\left\langle \xi, u^\dagger - u\right\rangle_{U^*,U} = \left\langle F'(u^\dagger)^*\, w, u^\dagger - u\right\rangle_{U^*,U} = \left\langle w, F'(u^\dagger)(u^\dagger - u)\right\rangle_{V^*,V} \le \|w\|_{V^*}\|F'(u^\dagger)(u^\dagger - u)\|_V$$

$$\le \|w\|_{V^*}\left\|F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)\right\|_V + \|w\|_{V^*}\left\|F(u) - F(u^\dagger)\right\|_V$$

$$\le K\|w\|_{V^*}\left\|F(u) - F(u^\dagger)\right\|_V^{c_1} D_\xi(u, u^\dagger)^{c_2} + \|w\|_{V^*}\left\|F(u) - F(u^\dagger)\right\|_V.$$

Taking into account that $\left\|F(u) - F(u^\dagger)\right\|_V \le \rho^{1/p}$ for $u \in \mathcal{M}^v_{\alpha_{max}}(\rho)$ this implies for the case $c_2 = 0$ and $0 < c_1 \le 1$ the variational inequality (3.5) with $\beta_1 = 0$, $\beta_2 = \|w\|_{V^*}(K + \rho^{\frac{1-c_1}{p}})$ and $\kappa = c_1$. On the other hand, for $0 < c_2 < 1$ and $0 < c_1 \le 1$ the variant (3.9) of Young's inequality with $p_1 := \frac{1}{c_2}$, $p_2 := \frac{1}{1-c_2}$, $\varepsilon := c_2$, $a := D_\xi(u, u^\dagger)^{c_2}$ and $b := K\|w\|_{V^*}\left\|F(u) - F(u^\dagger)\right\|_V^{c_1}$ yields here

$$K\|w\|_{V^*}\left\|F(u) - F(u^\dagger)\right\|_V^{c_1} D_\xi(u, u^\dagger)^{c_2} \le c_2\, D_\xi(u, u^\dagger) + (1 - c_2)(K\|w\|_{V^*})^{\frac{1}{1-c_2}}\left\|F(u) - F(u^\dagger)\right\|_V^{\frac{c_1}{1-c_2}}.$$

and hence the validity of a variational inequality (3.5) with $\kappa = \frac{c_1}{1-c_2}$ and multipliers

$$0 \le \beta_1 = c_2 < 1,\ \ \beta_2 = \rho^{\frac{1-\kappa}{p}}\|w\|_{V^*} + (1 - c_2) K^{\frac{1}{1-c_2}}\|w\|_{V^*}^{\frac{1}{1-c_2}}.$$

This proves the proposition. □

Note that essential ingredients for Proposition 4.5 and its proof have already been presented in [13, Lemma 3.1]. The proposition shows that the variational inequality (3.5) holds with the maximum exponent $\kappa = 1$ if either $c_1$ itself is maximal, i.e., $c_1 = 1$, or its defect in the case $0 < c_1 < 1$ can be compensated by $c_2 > 0$ whenever we have $c_1 + c_2 = 1$.

We mention here three questions, which cannot be answered in the moment for the used general Banach space setting under consideration in this paper:
I. Are there alternative sufficient conditions for obtaining a variational inequality (3.5) with exponents $0 < \kappa < 1$ when $\xi$ fails to satisfy a source condition (4.3)?
II. What combinations of $c_1$ and $c_2$ in the degree of nonlinearity do really occur?
III. Are the degrees of nonlinearity $(c_1, c_2)$ with $c_1 + c_2 > 1$ of interest?
The questions, however, will be partially answered in the subsequent Section 5 for the standard Tikhonov regularization in a Hilbert space setting

**Remark 4.6** Taking into account that only exponents $0 < \kappa \le 1$ make sense in general, we close this section with some remark on the pathological case $p < 1$. It is well-known that the case $p = 1$ in the Tikhonov functional (2.3) is singular and leads to the so-called *exact penalization* (see [4, 11, 14]). Even if the numerical difficulties (non-convexity) should discourage any attempt to use $0 < p < 1$ in variational regularization with norm powers as residual term, it is of interest whether convergence rates also can be obtained in that case. It seems to be no problem to extend the assertions on existence and stability to that case. However, the convergence condition $\frac{\delta^p}{\alpha(\delta)} \to 0$ as $\delta \to 0$ shows that the decay of $\alpha(\delta) \to 0$ has to become slow if $p$ is very small. For obtaining convergence rates this corresponds with the a priori parameter choice $\alpha(\delta) \asymp \delta^{p-\kappa}$ in Proposition 3.3, which only works for

$$0 < \kappa < p \qquad \text{whenever} \qquad 0 < p < 1\,. \tag{4.8}$$

As one can easily check by the proof the assertion $D_\xi(u^\delta_{\alpha(\delta)}, u^\dagger) = \mathcal{O}(\delta^\kappa)$ of Proposition 3.3 can be extended to the case (4.8) for the same $\alpha$-choice. The important consequence is that an exponent $0 < p < 1$ for $\kappa \ge p$ artificially bounds the occurring convergence rate to $D_\xi(u^\delta_{\alpha(\delta)}, u^\dagger) = \mathcal{O}(\delta^\nu)$ which then only holds for $0 < \nu < p$. In this context we make the important note that the validity of a variational inequality (3.5) with some exponent $0 < \kappa \le 1$ implies the validity of such variational inequalities also for all positive exponents smaller than $\kappa$, where when indicated the corresponding level sets have to be adapted.

# 5 The degree of nonlinearity with weaker norms

In the Banach space setting the condition (4.6) characterizing the degree of nonlinearity is in general difficult to verify for specific nonlinear inverse problems. As was shown in [14] sometimes the situation tends to the better if weaker norms are introduced in addition to the stronger ones (cf. [14, Remark 4.2]). Therefore we relax the condition by imposing a weaker norm on the first order Taylor remainder on the left hand side of (4.6). In this context, we complement the Assumption 4.1 as follows:

**Assumption 5.1**

1. *Let hold the Assumption 4.1.*

2. *Let $\tilde{V}$ be a reflexive Banach space with dual $\tilde{V}^*$, where $V$ is densely and continuously embedded in $\tilde{V}$ such that we have with a constant $C > 0$*

$$\|v\|_{\tilde{V}} \le C\,\|v\|_V \qquad \text{for all} \quad v \in V\,.$$

3. *For the $\Omega$-minimizing solution $u^\dagger \in \mathcal{D}$ with $\xi = \Omega'(u^\dagger)$ let exist $c_1 > 0$ and $0 \le c_2 < 1$ such that we have with some constant $K > 0$*

$$\left\|F(u) - F(u^\dagger) - F'(u^\dagger)(u-u^\dagger)\right\|_{\tilde{V}} \;\le\; K\,\left\|F(u) - F(u^\dagger)\right\|_V^{c_1}\, D_\xi(u,u^\dagger)^{\,c_2} \tag{5.1}$$

*for all $u \in \mathcal{M}^v_{\alpha_{max}}(\rho)$.*

4. *For that $\xi$ there exists an element $\tilde{w} \in \tilde{V}^*$ such that*

$$\left\langle \xi, u^\dagger - u\right\rangle_{U^*,U} \le \left|\left\langle \tilde{w}, F'(u^\dagger)(u-u^\dagger)\right\rangle_{\tilde{V}^*,\tilde{V}}\right| \tag{5.2}$$

*for all $u \in \mathcal{M}^v_{\alpha_{max}}(\rho)$.*

**Proposition 5.2** *Under the Assumption 5.1 a variational inequality*

$$\left\langle \xi, u^\dagger - u\right\rangle_{U^*,U} \le \beta_1 D_\xi(u,u^\dagger) \;+\; \beta_2\left\|F(u) - F(u^\dagger)\right\|_V^\kappa$$

*holds for all $u \in \mathcal{M}^v_{\alpha_{max}}(\rho)$ with $\kappa = \frac{c_1}{1-c_2}$ and for some $0 \le \beta_1 < 1,\ \beta_2 \ge 0$. Hence Proposition 3.3 applies and yields*

$$D_\xi(u^\delta_{\alpha(\delta)}, u^\dagger) \;=\; \mathcal{O}\left(\delta^\kappa\right) \quad \text{as} \quad \delta \to 0$$

*for an a priori parameter choice $\alpha(\delta) \asymp \delta^{p-\kappa}$.*

**Proof:** The proof is completely analogous to that of Proposition 4.5. Only the constants are different and they are just without meaning in the present case $c_1 > 0$. Namely, we can estimate for $u \in \mathcal{M}^v_{\alpha_{max}}(\rho)$

$$\left\langle \xi, u^\dagger - u\right\rangle_{U^*,U} \le \left|\left\langle \tilde{w}, F'(u^\dagger)(u-u^\dagger)\right\rangle_{\tilde{V}^*,\tilde{V}}\right|$$

$$\le \|\tilde{w}\|_{\tilde{V}^*}\|F'(u^\dagger)(u^\dagger - u)\|_{\tilde{V}} \le C\,\|\tilde{w}\|_{\tilde{V}^*}\|F'(u^\dagger)(u^\dagger - u)\|_V$$

$$\le C\,\|\tilde{w}\|_{\tilde{V}^*}\left\|F(u) - F(u^\dagger) - F'(u^\dagger)(u-u^\dagger)\right\|_V + C\,\|\tilde{w}\|_{\tilde{V}^*}\left\|F(u) - F(u^\dagger)\right\|_V$$

$$\le C\,K\|\tilde{w}\|_{\tilde{V}^*}\left\|F(u) - F(u^\dagger)\right\|_V^{c_1}\, D_\xi(u,u^\dagger)^{\,c_2} + C\,\|\tilde{w}\|_{\tilde{V}^*}\left\|F(u) - F(u^\dagger)\right\|_V.$$

Then Young's inequality yields the variational inequality as in the former proof. □

**Remark 5.3** Proposition 5.2 extends the main result of [14], where the convergence rate $D_\xi(u^\delta_{\alpha(\delta)}, u^\dagger) = \mathcal{O}(\delta)$ was shown under similar assumptions. Precisely, in [14] the nonlinearity condition (5.1) was focused on exponents $c_1 = 0$, $c_2 = 1$ and (5.2) had to be complemented by a smallness condition $K\|\tilde{w}\|_{\tilde{V}^*} < 1$. Note that as in [14] the condition (5.2) of our Proposition 5.2 is satisfied if there is an element $\tilde{w} \in \tilde{V}^*$ such that the equation

$$\langle \xi, u^\dagger - u \rangle_{U^*,U} = \langle \tilde{w}, F'(u^\dagger)(u - u^\dagger) \rangle_{\tilde{V}^*,\tilde{V}} \tag{5.3}$$

holds for all $u \in \mathcal{M}^v_{\alpha_{max}}(\rho)$. This is easy to verify if $V = L^2$ and $\tilde{V} = L^{2-\varepsilon}$ $(0 < \varepsilon < 1)$. Then we have $\tilde{V}^* = L^\mu$ with $\mu = \frac{2-\varepsilon}{1-\varepsilon} > 2$ and (5.3) attains the form

$$\langle \xi, u^\dagger - u \rangle_U = \langle \tilde{w}, F'(u^\dagger)(u - u^\dagger) \rangle_{V^*,V}$$

(cf. [14, Remark 4.2 and §6]).

# 6 Extended results for Hilbert space situations

In Assumption 6.1 we specify now the requirements expressing the setting of this section.

**Assumption 6.1**

1. *Set $p := 2$ and let $U, V$ be Hilbert spaces, where we identify $U$ and $U^*$ as well as $V$ and $V^*$ by using the Riesz isomorphism. By the adjoint $A^*$ of a linear operator $A$ in this section we always mean the Hilbert space adjoint.*
2. *The operator $F$, $\mathcal{D}(F)$, $u^\dagger$ and $\xi$ are chosen such that they satisfy together with $U, V$ and $\Omega$ the Assumption 4.1.*

3a. *Let $\Omega(u) := \|u - u^*\|^2_U$ with fixed reference element $u^* \in U$ and $\mathcal{D}(\Omega) = U$.*

3b. *Let $B : \mathcal{D}(B) \subset U \to U$ be an unbounded injective, positive definite, self-adjoint linear operator with domain $\mathcal{D}(B)$ dense in $U$. Furthermore let $\tilde{C} > 0$ be a constant such that $\|u\|_{\tilde{U}} := \|Bu\|_u \ge \tilde{C}\,\|u\|_U$ $(u \in \tilde{U})$, and $\tilde{U}$ is a Hilbert space with norm $\|\cdot\|_{\tilde{U}}$ stronger than $\|\cdot\|_U$. Moreover, let $\Omega(u) := \|Bu\|_U = \|u\|_{\tilde{U}}$ with $\mathcal{D}(\Omega) = \tilde{U}$.*

**Remark 6.2 (Case 3a)** Under the Assumption 6.1 in the case 3a the $\Omega$-minimizing solutions and the classical $u^*$-minimum solutions (cf. [7, 8]) coincide. Moreover, we have $\mathcal{D} = \mathcal{D}(F)$ and for $\xi$ and $D_\xi(\tilde{u}, u)$ the simple structure

$$\xi = 2\,(u^\dagger - u^*) \qquad \text{and} \qquad D_\xi(\tilde{u}, u) = \|\tilde{u} - u\|^2_U \tag{6.1}$$

with Bregman domain $\mathcal{D}_B(\Omega) = U$. Regularized solutions $u^\delta_\alpha$ are minimizers over $U$ of the classical Tikhonov functional of Hilbert space type

$$T^{v^\delta}_\alpha(u) := \|F(u) - v^\delta\|^2_V + \alpha\,\|u - u^*\|^2_U$$

comprehensively studied in [7, Chapter 10].

**Remark 6.3 (Case 3b)** Under the Assumption 6.1 in the case 3b we have $\mathcal{D}(\Omega) = \tilde{U}$ and if the limit $\langle \xi, \hat{u} \rangle_U = \Omega'(u^\dagger)\,\hat{u} = \lim_{t\to 0} \frac{\Omega(u^\dagger + t\hat{u} - \Omega(u^\dagger)}{t}$ exists it has the form $\langle \xi, \hat{u} \rangle_U = 2\,\langle Bu^\dagger, B\hat{u} \rangle_U = \langle B^2 u^\dagger, \hat{u} \rangle_U$. If and only if $u^\dagger \in \mathcal{D}(B^2) = \{u \in U : \|B^2 u\|_U < \infty\}$ (a proper subset of $\tilde{U}$), the symbol $\Omega'(u^\dagger)$ characterizes the Gâteaux derivative of $\Omega$ at the point $u^\dagger$ characterized as a bounded linear functional defined on the whole space $U$. Just then the subdifferential $\partial\Omega(u^\dagger)$ is nonempty and a singleton $\partial\Omega(u^\dagger) = \{\xi\}$ with

$$\xi \;=\; 2\,B^2 u^\dagger \qquad \text{and} \qquad D_\xi(\tilde{u}, u) \;=\; \|B(\tilde{u} - u)\|_U^2 \;\geq\; \tilde{C}^2\,\|\tilde{u} - u\|_U^2\,, \tag{6.2}$$

where the Bregman domain is $\mathcal{D}_B(\Omega) = \mathcal{D}(B^2)$. Regularized solutions $u_\alpha^\delta$ are minimizers over $\tilde{U}$ of the functional

$$T_\alpha^{v^\delta}(u) := \|F(u) - v^\delta\|_V^2 + \alpha\,\|Bu\|_U^2\,.$$

Note that for all $0 < \mu \leq 1$ as a consequence of (6.2) a convergence rate

$$D_\xi(u_{\alpha(\delta)}^\delta, u^\dagger) = \mathcal{O}\left(\delta^\mu\right) \qquad \text{implies} \qquad \|u_{\alpha(\delta)}^\delta - u^\dagger\| = \mathcal{O}\left(\delta^{\frac{\mu}{2}}\right)\,.$$

## 6.1 The Hilbert space situation for variational inequalities

In this subsection under Assumption 6.1 we refer to Case 3a. To focus on the distinguished character of the Hilbert space setting we will specify the Definition 3.2 as follows:

**Definition 6.4** *Let $c_1, c_2 \geq 0$ and $c_1 + c_2 \geq 0$. We define $F$ to be nonlinear of degree $(c_1, c_2)$ at a solution $u^\dagger \in \mathcal{D}(F)$ of (2.1) if there is a constant $K > 0$ such that*

$$\left\|F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)\right\|_V \;\leq\; K\,\left\|F(u) - F(u^\dagger)\right\|_V^{c_1}\,\|u - u^\dagger\|_U^{2c_2} \tag{6.3}$$

*for all $u \in \mathcal{M}_{\alpha_{max}}^v(\rho)$.*

**Remark 6.5** In this Hilbert space setting we can formulate conditions for admissible pairs $(c_1, c_2)$ in formula (6.3) of Definition 6.4 and study the smoothness background of such degrees of nonlinearity.

A sufficient condition for the classical case $c_1 = 0$, $c_2 = 1$, assumed for example in [29, Section 3.2]), is the Lipschitz continuity

$$\|F'(u) - F'(u^\dagger)\|_{\mathcal{L}(U,V)} \leq L\|u - u^\dagger\|_U$$

of $F'$ for all $u$ in a neighbourhood of $u^\dagger$. On the other hand, the case $c_1 = 1$, $c_2 = 0$ characterized by a tangential cone condition is frequently discussed in the theory of iterative regularization (cf. [7, Chapter 11] and [18]). In [13] the focus is on the case $c_1 > 0$, $0 < c_1 + c_2 \leq 1$, but it is well-known that numerous applications of ill-posed nonlinear inverse problems occur, where $c_1 = 1$, $c_2 = 1/2$ can be shown, i.e. $1 < c_1 + c_2 \leq 2$. We conjecture that the conditions

$$0 \leq c_1, c_2 \leq 1, \quad 0 < c_1 + 2c_2 \leq 2 \tag{6.4}$$

characterize all really occurring situations apart from singular cases.

As already mentioned in [15] the pairs $(c_1, c_2)$ of the degree of nonlinearity are not necessarily uniquely determined. Namely, under a local Lipschitz condition

$$\|F(u) - F(u^\dagger)\|_V \;\le\; C\,\|u - u^\dagger\|_U \tag{6.5}$$

for all $u$ in a neighbourhood of $u^\dagger$ a degree $(c_1, c_2)$ evidently implies the degree $(0, c_1/2+c_2)$. Then $c_1 + 2c_2 > 2$ would lead to some $\varepsilon > 0$ such that

$$\left\|F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)\right\|_V \;\le\; K\,\|u - u^\dagger\|_U^{2+\varepsilon}$$

for all $u$ from appropriate level sets. If the operator $F$ is continuously twice differentiable in a neighbourhood of $u^\dagger$ with bilinear operators $F''(u) : U \times U \to V$ using the integral representation of the second order Taylor remainder this would imply $\|F''(u^\dagger)(h,h)\|_V = 0$ for all $h \in U$ indicating a singular case.

Using similar arguments as in the proof of Proposition 4.3 we can easily see that for $F'(u^\dagger) \neq 0$ an inequality

$$\left\|F'(u^\dagger)(u - u^\dagger)\right\|_V \;\le\; K\,\left\|F(u) - F(u^\dagger)\right\|_V^{c_1}$$

cannot hold for all $u \in \mathcal{M}_{\alpha_{max}}(\rho)$ whenever $c_1 > 1$. Then because of $c_1 - 1 > 0$ an ansatz $u := u^\dagger + th$ with $h \neq 0$ and $\|h\|_U$ sufficiently small would after division by $t > 0$ lead to

$$\left\|F'(u^\dagger)h\right\|_V \;\le\; K\,\left\|F'(u^\dagger)h\right\|_V \lim_{t\to 0}\left\|F(u) - F(u^\dagger)\right\|_V^{c_1-1} = 0$$

in the limit case for $t \to 0$.

**Proposition 6.6** *Under the Assumption 6.1 in one of the version 3a or 3b let the operator $F$ mapping between the Hilbert spaces $U$ and $V$ be nonlinear of degree $(c_1, c_2)$ at $u^\dagger$ with $c_1 > 0$ and let $\xi$ satisfy the general source condition*

$$\xi = (F'(u^\dagger)^* F'(u^\dagger))^{\eta/2} w, \qquad 0 < \eta < 1, \;\; w \in U. \tag{6.6}$$

*Then we have the variational inequality (3.5) with exponent*

$$\kappa \;=\; \min\left\{ \frac{2\eta c_1}{1 + \eta(1 - 2c_2)}\,,\; \frac{2\eta}{1+\eta} \right\} \tag{6.7}$$

*for all $u \in \mathcal{M}_{\alpha_{max}}(\rho)$ and multipliers $0 \le \beta_1 < 1$, $\beta_2 \ge 0$.*

**Proof:** Under the general source condition (6.6) we can estimate for all $u \in \mathcal{M}_{\alpha_{max}}(\rho)$ with the interpolation inequality [7, formula (2.49), p. 47]

$$\left\langle \xi, u^\dagger - u \right\rangle_U \le \left\langle w, (F'(u^\dagger)^* F'(u^\dagger))^{\eta/2}(u^\dagger - u)\right\rangle_U$$

$$\le \|w\|_U \|(F'(u^\dagger)^* F'(u^\dagger))^{\eta/2}(u^\dagger - u)\|_U^{\eta} \|u^\dagger - u\|_U^{1-\eta} = \|w\|_U \|F'(u^\dagger)(u^\dagger - u)\|_V^{\eta} \|u^\dagger - u\|_U^{1-\eta},$$

where $\langle\cdot,\cdot\rangle_U$ denotes the inner product in the Hilbert space $U$. Now we use the degree of nonlinearity in order to estimate the term $\|F'(u^\dagger)(u^\dagger - u)\|_V^{\eta}$ from above for $u \in \mathcal{M}_{\alpha_{max}}(\rho)$. Owing to $0 < \eta < 1$ we have

$$\|F'(u^\dagger)(u^\dagger - u)\|_V^{\eta} \le \|F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)\|_V^{\eta} + \|F(u^\dagger) - F(u)\|_V^{\eta}$$

and hence with some constants $K_1, K_2 > 0$

$$\langle \xi, u^\dagger - u \rangle_U \le \|w\|_U \left( \|F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)\|_V^\eta + \|F(u^\dagger) - F(u)\|_V^\eta \right) \|u^\dagger - u\|_U^{1-\eta}$$
$$\le K_1 \|F(u^\dagger) - F(u)\|_V^{c_1 \eta} \|u^\dagger - u\|_U^{1-\eta+2c_2\eta} + K_2 \|F(u^\dagger) - F(u)\|_V^\eta \|u^\dagger - u\|_U^{1-\eta}.$$

Applying again Young's inequality (3.9) twice with $\varepsilon := 1/4$ such that terms $\frac{1}{4}\|u^\dagger - u\|_U^2$ occur in a sum with powers of $\|F(u^\dagger) - F(u)\|_V$ we obtain with some constants $C_1, C_2 > 0$

$$\langle \xi, u^\dagger - u \rangle_U \le \frac{1}{2}\|u^\dagger - u\|_U^2 + C_1 \|F(u^\dagger) - F(u)\|_V^{\frac{2c_1\eta}{1+\eta(1-2c_2)}} + C_2 \|F(u^\dagger) - F(u)\|_V^{\frac{2\eta}{1+\eta}}.$$

Taking into account that there is a constant $\bar{K} > 0$ such that $\|F(u^\dagger) - F(u)\|_V \le \bar{K}$ for all $u \in \mathcal{M}_{\alpha_{max}}(\rho)$ we have the variational inequality

$$\langle \xi, u^\dagger - u \rangle_U \le \frac{1}{2}\|u^\dagger - u\|_U^2 + \beta_2 \|F(u^\dagger) - F(u)\|_V^\kappa$$

with $\kappa$ from (6.7) for all such $u$ and some constant $\beta_2 > 0$. This completes the proof. □

**Remark 6.7** An exponent $\kappa = \frac{2\eta}{1+\eta}$ in Proposition 6.6 indicates order optimal convergence rates with respect to the general source condition (6.6). This is the case if the condition

$$1 + \eta(1 - 2c_2 - c_1) \le c_1 \tag{6.8}$$

already occurring in [15] is satisfied. Note that the condition (6.8) holds for $0 < \eta < 1$ only if either $c_1 = 1$ or for $0 < c_1 < 1$ if $c_1 + c_2 > 1$ and $\eta$ is large enough. In the case $c_1 = 0$, $c_2 = 1$ no convergence rate result based on low order general source conditions (6.6) is known (see also [17]). We conjecture that such results really cannot be formulated.

The authors have no general answer to the question whether one can formulate converse assertions concluding in the Hilbert space setting from a variational inequality (3.5) with exponents $0 < \kappa < 1$ and nonlinear forward operator $F$ to Hölder source conditions of type (6.6). However, for the subcase of a continuous linear operator

$$F := A \in \mathcal{L}(U, V) \tag{6.9}$$

we can prove a converse result in the following proposition (cf. also [6, Section 3] with respect to approximate source conditions). Since the conditions $\xi \in \mathcal{R}(F'(u^\dagger)^*)$ and $\xi \in \mathcal{R}((F'(u^\dagger)^* F'(u^\dagger))^{1/2})$ are equivalent this result complements for the subcase the assertion a) of our Proposition 4.4 and of Proposition 3.38 in [29] which just handle the case $\kappa = 1$. We should mention here that for linear operators (6.9) no structural condition (degree of nonlinearity) is required and Proposition 6.6 always yields the implication from

$$\xi = (A^* A)^{\eta/2} w, \qquad 0 < \eta < 1, \quad w \in U, \tag{6.10}$$

to a variational inequality

$$\langle \xi, u^\dagger - u \rangle_U \le \beta_1 \|u^\dagger - u\|_U^2 + \beta_2 \|A(u^\dagger - u)\|_V^\kappa \tag{6.11}$$

with exponent

$$\kappa = \frac{2\eta}{1+\eta} \in (0, 1) \tag{6.12}$$

for all $u \in \mathcal{M}_{\alpha_{max}}(\rho)$ and multipliers $0 \le \beta_1 < 1$, $\beta_2 \ge 0$.

**Proposition 6.8** *If for linear forward operators (6.9) a variational inequality (6.11) holds for all $u \in \mathcal{M}_{\alpha_{max}}(\rho)$ and multipliers $0 \le \beta_1 < 1$, $\beta_2 \ge 0$, then under Assumption 6.1 a general source condition (6.10) is valid for all $\eta < \frac{\kappa}{2-\kappa}$.*

**Proof:** Under Assumption 6.1 we obtain from (6.11) and (3.11) for noiseless data and an a priori parameter choice $\alpha(\delta) \asymp \delta^{2-\kappa}$ the estimate

$$\|u_\alpha^0 - u^\dagger\|_U \le \hat{C}\, \alpha^{\frac{\kappa}{2(2-\kappa)}}.$$

This allows us to apply the converse result of [20] for linear Tikhonov regularization which provides a Hölder source condition (6.10) for all exponents $\eta > 0$ satisfying the inequality $\eta < \frac{\kappa}{2-\kappa}$. This completes the proof. □

## 6.2 Hilbert space regularization and conditional stability

In this subsection under Assumption 6.1 we refer to Case 3b. For parameter identification problems in partial differential equations (cf., e.g., [2, 16]) nonlinear forward operators $F$ occur, for which the required Taylor remainder $\left\|F(u) - F(u^\dagger) - F'(u^\dagger)(u-u^\dagger)\right\|_V$ is difficult to handle and the variational inequality approach may fail. However, let hold for all $R > 0$ a *conditional stability estimate* of the form

$$\|u_1 - u_2\|_U \;\le\; K\,\|F(u_1) - F(u_2)\|_V^\kappa, \quad \text{if} \quad u_i \in \mathcal{D}(F) \cap \tilde{U}, \;\; \|u_i\|_{\tilde{U}} \le R \;\; (i = 1,2) \tag{6.13}$$

with some $0 < \kappa \le 1$ and a constant $K = K(R) > 0$ which may depend on the radius $R$.

To explain the cross-connections between conditional stability and the degree of nonlinearity one should note that if $F$ is nonlinear of degree $(0, c_2)$, $0 < c_2 \le 1$, at the point $u^\dagger$ in the sense of Definition 6.4, then (6.13) implies a degree $(2c_2\,\kappa, 0)$. In other words, the conditional stability with some Hölder rate converts the degree of nonlinearity to $(c_1, 0)$ such that $c_1 > 0$. In [5], a rate of the condition stability is related with the convergence rate of regularized solutions. That is:

**Proposition 6.9** *Under Assumption 2.1 let $F, \Omega, \mathcal{D}, U$ and $V$ satisfy the Hilbert space specification expressed by items 1, 2, and 3b of Assumption 6.1, i.e., regularized solutions $u_\alpha^\delta$ are minimizers over $\mathcal{D} = \mathcal{D}(F) \cap \tilde{U}$ of the functional*

$$T_\alpha^{v^\delta}(u) := \|F(u) - v^\delta\|_V^2 + \alpha\,\|u\|_{\tilde{U}}^2\,. \tag{6.14}$$

*Moreover, for all $R > 0$ let hold a conditional stability estimate of the form (6.13) with some $0 < \kappa \le 1$ and a constant $K = K(R) > 0$. Then for a solution $u^\dagger \in \mathcal{D}$ of equation (2.1) we obtain the convergence rate*

$$\|u_{\alpha(\delta)}^\delta - u^\dagger\|_U \;=\; \mathcal{O}\left(\delta^\kappa\right) \quad \text{as} \quad \delta \to 0\,. \tag{6.15}$$

*with an a priori parameter choice $\underline{c}\delta^2 \le \alpha(\delta) \le \overline{c}\delta^2$ with constants $0 < \underline{c} \le \overline{c} < \infty$.*

**Proof:** For the proof we follow ideas of [5]. As a minimizer of (6.14) the element $u_\alpha^\delta$ satisfies the inequalities

$$\|F(u_\alpha^\delta) - v^\delta\|_V^2 \le \|F(u_\alpha^\delta) - v^\delta\|_V^2 + \alpha\,\|u_\alpha^\delta\|_{\tilde U}^2 \le \|F(u^\dagger) - v^\delta\|_V^2 + \alpha\,\|u^\dagger\|_{\tilde U}^2 .$$

Hence we have

$$\|F(u_\alpha^\delta) - v^\delta\|_V^2 \le \delta^2 + \alpha\,\|u^\dagger\|_{\tilde U}^2$$

and with $\alpha(\delta) \sim \delta^2$

$$\|F(u_{\alpha(\delta)}^\delta) - v^\delta\|_V \;=\; \mathcal{O}\,(\delta) \quad \text{as} \quad \delta \to 0\,. \tag{6.16}$$

In a similar manner one obtains $\alpha\|u_\alpha^\delta\|_{\tilde U}^2 \le \delta^2 + \alpha\,\|u^\dagger\|_{\tilde U}^2$ and $\|u_{\alpha(\delta)}^\delta\|_{\tilde U} \le R := \sqrt{\frac{1}{\underline{c}} + \|u^\dagger\|_{\tilde U}^2}$. Because we also have $\|u^\dagger\|_{\tilde U} \le R$ the estimate (6.13) is applicable and yields with some constants $K = K(R)$

$$\|u_{\alpha(\delta)}^\delta - u^\dagger\|_U \le K\,\left\|F(u_{\alpha(\delta)}^\delta) - F(u^\dagger)\right\|_V^\kappa \le \left\|F(u_{\alpha(\delta)}^\delta) - v^\delta\right\|_V^\kappa + \left\|F(u^\dagger) - v^\delta\right\|_V^\kappa .$$

Taking into account (6.16) and $\left\|F(u^\dagger) - v^\delta\right\|_V^\kappa \le \delta^\kappa$ this proves the proposition. □

**Example 6.10** We give an example that shows the applicability of Proposition 6.9 for inverse problems in partial differential equations:

Let $G \subset \mathbb{R}^n$ be a bounded domain with smooth boundary $\partial G$ and set $L := -\Delta$ with $\mathcal{D}(L) = H^2(G) \cap H_0^1(G)$. Here and henceforth $H^2(G)$, $H_0^1(G)$, $H^\theta(0,T)$, $\theta > 0$, etc. denote usual Sobolev spaces. We consider

$$\begin{cases} \partial_t y = -Ly + u(t)y(x,t), & x \in G,\ t > 0, \\ y(x,0) = a(x), & x \in G, \\ y|_{\partial G \times (0,T)} = 0. \end{cases} \tag{6.17}$$

The equation models a reaction and diffusion process whose reaction rate depends on $t$ with factor $u(t)$. We are interested in an *inverse problem* of determining such a function $u$ forming a time factor $u(t)$, $0 < t < T$, by available observation data. The observations are represented by interior mean data. In this context, let $G_0 \subset G$ be a subdomain such that $\overline{G_0} \subset G$. Then we observe noisy data of the function

$$v(t) := \int_{G_0} y(x,t)dx, \qquad 0 < t < T,$$

which expresses the exact right-hand in equation (2.1) (see [25] as for similar inverse problems). Both $u$ and $v$ are functions over the time interval $(0,T)$ under consideration.

We consider (6.17) as an ordinary differential equation in $t$ in a Hilbert space $L^2(G)$ with the norm $\|\cdot\|$. We set $y(t); = y(\cdot,t)$ mapping from $(0,T)$ to $L^2(G)$. Then we can write (6.17) as

$$\begin{cases} y'(t) = -Ly + u(t)y(t), & t > 0, \\ y(0) = a, & x \in G. \end{cases} \tag{6.18}$$

We treat (6.18) with respect to the semigroup theory (cf., e.g., [23]) and note that $y(t) \in \mathcal{D}(L),\ t > 0$, refers to the boundary condition in (6.17), where we assume

$$a \in C_0^\infty(G), \quad a \not\equiv 0, \quad a(x) \geq 0,\ x \in G. \tag{6.19}$$

Moreover, let hold the substitution

$$\varphi(\eta) := \int_{G_0} \eta(x)dx, \qquad \eta \in L^2(G).$$

We will show that the forward operator $F : u \mapsto v$ is mapping from $L^2(0,T)$ into $H^1(0,T)$ is well-defined and satisfies there a stability estimate . First let $\|u\|_{L^2(0,T)} \leq R$. Henceforth $C_j > 0$ denote constants which are dependent on $R$, $a$, but independent of the choice of $u$. We have

$$y(t) = e^{-tL}a + \int_0^t u(s)e^{-(t-s)L}y(s)ds, \quad t > 0. \tag{6.20}$$

Therefore

$$Ly(t) = e^{-tL}La + \int_0^t u(s)e^{-(t-s)L}Ly(s)ds, \quad t > 0.$$

Then

$$\|Ly(t)\| \leq C_1 + \int_0^t C_1|u(s)|\|Ly(s)\|ds, \quad t > 0.$$

Now Gronwall's inequality yields

$$\|Ly(t)\| \leq C_1 \exp\left(C_1 \int_0^t |u(s)|ds\right) \leq C_1 e^{C_1 R\sqrt{T}} \equiv C_2, \quad 0 \leq t \leq T. \tag{6.21}$$

By (6.20), we obtain

$$y'(t) = -e^{-tL}La + u(t)y(t) - \int_0^t u(s)e^{-(t-s)L}Ly(s)ds, \quad t > 0. \tag{6.22}$$

Therefore, using (6.21), we have

$$\|y'(t)\| \leq C_3 + C_3\|u(t)y(t)\| + \int_0^t C_3|u(s)|\|Ly(s)\|ds \leq C_3 + C_3|u(t)| + \int_0^t C_3|u(s)|ds,$$

so that

$$\int_0^t \|y'(\eta)\|^2 d\eta \leq C_4 \int_0^t |u(\eta)|^2 d\eta + C_4, \quad 0 < t < T.$$

Hence

$$\|\varphi(y'(u))\|_{L^2(0,T)} + \|\varphi(y(u))\|_{L^2(0,T)} \leq C_5 \quad \text{if} \quad \|u\|_{L^2(0,T)} \leq R. \tag{6.23}$$

Consequently $F : L^2(0,T) \longrightarrow H^1(0,T)$ is well-defined.

Now let $u_1, u_2 \in L^2(0,T)$ and let $\|u_1\|_{L^2(0,T)}, \|u_2\|_{L^2(0,T)} \leq R$, and set $z_1 = y(u_1), z_2 = y(u_2)$. Then

$$(z_1 - z_2)(t) = \int_0^t (u_1 - u_2)(s)e^{-(t-s)L}z_1(s)ds + \int_0^t u_2(s)e^{-(t-s)L}(z_1 - z_2)(s)ds, \quad 0 < t < T. \tag{6.24}$$

By (6.21), we obtain

$$\|L(z_1-z_2)(t)\| \le C_1\int_0^t |(u_1-u_2)(s)|\|Lz_1(s)\|ds + C_1\int_0^t |u_2(s)|\|L(z_1-z_2)(s)\|ds$$
$$\le C_6\int_0^t |(u_1-u_2)(s)|ds + C_6\int_0^t |u_2(s)|\|L(z_1-z_2)(s)\|ds.$$

Gronwall's inequality yields

$$\|L(z_1-z_2)(t)\| \le C_6\left(\int_0^t |(u_1-u_2)(s)|^2ds\right)^{\frac{1}{2}}$$
$$+\int_0^t \exp\left(C_6\int_0^t |u_2(s)|ds\right)\times C_6|u_2(s)|\times C_6\left(\int_0^t |(u_1-u_2)(\xi)|^2d\xi\right)^{\frac{1}{2}} ds,$$

that is,

$$\|L(z_1-z_2)(t)\| \le C_7\left(\int_0^t |(u_1-u_2)(s)|^2ds\right)^{\frac{1}{2}}, \quad 0\le t\le T. \tag{6.25}$$

Next we estimate

$$(z_1-z_2)'(t) = (u_1-u_2)(t)z_1(t) + u_2(t)(z_1-z_2)(t)$$
$$-\ \int_0^t (u_1-u_2)(s)e^{-(t-s)A}Lz_1(s)ds - \int_0^t u_2(s)e^{-(t-s)L}L(z_1-z_2)(s)ds.$$

Here

$$\varphi(z_1)(t)\times(u_1-u_2)(t) = \int_0^t (u_1-u_2)(s)\varphi(e^{-(t-s)L}Lz_1(s))ds \tag{6.26}$$
$$+\ \ \varphi((z_1-z_2)')(t) - u_2(t)(\varphi(z_1-z_2))(t) + \int_0^t u_2(s)\varphi(e^{-(t-s)L}L(z_1-z_2)(s))ds.$$

By (6.19) and the maximum principle, we have

$$\varphi(z_1)(t)\ge\varepsilon, \quad 0\le t\le T,$$

where $\varepsilon>0$ depends on $u_1$. Consequently

$$(u_1-u_2)(t) = \int_0^t (u_1-u_2)(s)\varphi(z_1)(t)^{-1}\varphi(e^{-(t-s)L}Lz_1(s))ds + \varphi(z_1)(t)^{-1}\varphi((z_1-z_2)')(t)$$
$$-\varphi(z_1)(t)^{-1}u_2(t)\varphi(z_1-z_2)(t) + \int_0^t u_2(s)\varphi(z_1)(t)^{-1}\varphi(e^{-(t-s)L}L(z_1-z_2)(s))ds.$$

By (6.21) and (6.25), we have

$$|(u_1-u_2)(t)| \le C_8\int_0^t |(u_1-u_2)(s)|ds + C_8(|\varphi((z_1-z_2)')(t)| + |\varphi(z_1-z_2)(t)|)$$
$$+C_8\left(\int_0^t |u_2(s)|^2ds\right)^{\frac{1}{2}}\left(\int_0^t \|L(z_1-z_2)(s)\|^2ds\right)^{\frac{1}{2}}.$$

Hence, in view of (6.25), we obtain

$$|(u_1-u_2)(t)| \le C_9(|\varphi((z_1-z_2)')(t)| + |\varphi(z_1-z_2)(t)|)$$

$$+C_9 \int_0^t |(u_1 - u_2)(s)| ds + C_6 \left( \int_0^t \left( \int_0^s |(u_1 - u_2)(\xi)|^2 d\xi \right) ds \right)^{\frac{1}{2}}.$$

Squaring the both sides, integrating over $(0, t)$ and setting $p(t) = \int_0^t |(u_1 - u_2)(s)|^2 ds$, we have

$$p(t) \le C_{10} \int_0^t (|\varphi((z_1 - z_2)')(t)|^2 + |\varphi(z_1 - z_2)(t)|^2) ds + C_{10} \int_0^t \left( \int_0^s |(u_1 - u_2)(\xi)|^2 d\xi \right) ds$$
$$+C_{10} \int_0^t \left( \int_0^s \left( \int_0^\xi |(u_1 - u_2)(\eta)|^2 d\eta \right) d\xi \right) ds$$
$$\le C_{10} \int_0^t (|\varphi((z_1 - z_2)')(t)|^2 + |\varphi(z_1 - z_2)(t)|^2) ds + C_{10} \int_0^t p(s) ds + C_{10} T \int_0^t p(s) ds.$$

At the last inequality we used

$$\int_0^t \left( \int_0^s \left( \int_0^\xi |(u_1 - u_2)(\eta)|^2 d\eta \right) d\xi \right) ds \le \int_0^t \left( \int_0^s p(\xi) d\xi \right) ds$$
$$\le \int_0^t \left( \int_0^s p(s) d\xi \right) ds \le t \int_0^t p(s) ds.$$

Again Gronwall's inequality yields

$$p(t) \le C_{11} \|\varphi(z_1 - z_2)\|^2_{H^1(0,T)}, \quad 0 < t < T.$$

Therefore we have proved the *basic inequality* for the considered inverse problem with respect to the parabolic initial-boundary value problem (6.17):

$$\|u_1 - u_2\|_{L^2(0,T)} \le C_{11} \|F(u_1) - F(u_2)\|_{H^1(0,T)} \quad \text{if} \quad \|u_1\|_{L^2(0,T)}, \|u_2\|_{L^2(0,T)} \le R. \tag{6.27}$$

Note that we have weak continuity of $F$ as a by-product of the calculations above. Namely, if $u_k$ converge to $u$ weakly in $L^2(0,T)$, then by (6.24) and (6.26) we can directly verify that $F(u_k)$ converge to $F(u)$ weakly in $H^1(0,T)$.

Exploiting the result (6.27) we now come back to Proposition 6.9 by setting

$$U := L^2(0,T), \quad \tilde{U} := H^\theta(0,T), \ \frac{1}{2} < \theta < 1, \quad V := L^2(0,T), \quad \kappa := \frac{\theta}{\theta + 1}. \tag{6.28}$$

We are going to prove that in such a case a stability estimate (6.13) can be verified.

**Proposition 6.11** *Under the setting (6.28) a conditional stability estimate (6.13) holds true, and Proposition 6.9 is appllicable, i.e., the convergence rate (6.15) occurs for the required choice of the regularization parameter.*

**Proof:** We use an equivalent representation of the norm of $H^\theta(0,T)$ by Slobodeckij (see, e.g., [1, Theorem VII.7.48, p. 214]). We rewrite (6.22) as

$$y'(t) = -e^{-tL} La + u(t) y(t) - \int_0^t u(s) e^{-(t-s)L} L y(s) ds \equiv I_1(t) + I_2(t) + I_3(t).$$

First we have

$$\varphi(I_2(t)) = u(t)\int_{G_0} y(x,t)dx.$$

For $\theta > \frac{1}{2}$ we have that $H^\theta(0,T) \subset C[0,T]$. Then by (6.23) one can write

$$\int_0^T\int_0^T \frac{|u(t)\varphi y(t) - u(s)\varphi y(s)|^2}{|t-s|^{1+2\theta}}dsdt \le 2\int_0^T\int_0^T \frac{|u(t)-u(s)|^2|\varphi y(t)|^2}{|t-s|^{1+2\theta}}dsdt$$

$$+2\int_0^T\int_0^T \frac{|u(s)|^2|\varphi y(t) - \varphi y(s)|^2}{|t-s|^{1+2\theta}}dsdt \le 2\|\varphi y\|^2_{C[0,T]}\|u\|^2_{H^\theta(0,T)} + 2\|u\|^2_{C[0,T]}\|\varphi y\|^2_{H^\theta(0,T)}$$

$$\le 2\|\varphi y\|^2_{C[0,T]}R^2 + 2\|u\|^2_{H^\theta(0,T)}\|\varphi y\|^2_{H^\theta(0,T)} \le 2C_5^2R^2 + 2R^2C_5^2.$$

Hence

$$\|\varphi(I_2)\|_{H^\theta(0,T)} \le C_{12}.$$

Next we estimate:

$$\|\varphi(I_1)\|^2_{H^\theta(0,T)} \le \|\varphi(I_1)\|^2_{H^1(0,T)} \le \int_0^T \left|\int_{G_0} e^{-tL}Ladx\right|^2 dt + \int_0^T \left|\int_{G_0} e^{-tL}L^2adx\right|^2 dt$$

$$\le C_{13}(\|La\|^2 + \|L^2a\|^2).$$

By (6.20), we have

$$L^2y(t) = e^{-tL}L^2a + \int_0^t u(s)e^{-(t-s)L}L^2y(s)ds.$$

Therefore, similarly to (6.21), Gronwall's inequality implies $\|L^2y(t)\| \le C_2$. Then we proceed to estimation of $I_3$:

$$\varphi(I_3)(t) = \int_{G_0}\left(\int_0^t u(s)e^{-(t-s)L}Ly(s)ds\right)dx = \int_0^t u(\xi)\left(\int_{G_0} e^{-(t-\xi)L}Ly(\xi)dx\right)d\xi.$$

It is sufficient to assume $t \ge s$:

$$(t-s)^{-1-2\theta}\left|\int_0^t u(\xi)\left(\int_{G_0} e^{-(t-\xi)L}Ly(\xi)dx\right)d\xi - \int_0^s u(\xi)\left(\int_{G_0} e^{-(s-\xi)L}Ly(\xi)dx\right)d\xi\right|^2$$

$$= (t-s)^{-1-2\theta}\left|\int_0^s u(\xi)\left(\int_{G_0} e^{-(s-\xi)L}(e^{-(t-s)L}-1)Ly(\xi)dx\right)d\xi\right.$$

$$\left. + \int_s^t u(\xi)\left(\int_{G_0} e^{-(t-\xi)L}Ly(\xi)dx\right)d\xi\right|^2$$

$$\le 2(t-s)^{-1-2\theta}\left|\int_0^s u(\xi)\left(\int_{G_0} e^{-(s-\xi)L}(e^{-(t-s)L}-1)Ly(\xi)dx\right)d\xi\right|^2$$

$$+2(t-s)^{-1-2\theta}\left|\int_s^t u(\xi)\left(\int_{G_0} e^{-(t-\xi)L}Ly(\xi)dx\right)d\xi\right|^2 \equiv 2J_1 + 2J_2.$$

First, noting that $\|(e^{-(t-s)L}-1)Ly(\xi)\| \le C'_{14}(t-s)\|L^2a\|$ (cf. [23]), we see that

$$\int_0^T\int_0^t J_1 dsdt \le \int_0^T\int_0^t (t-s)^{-1-2\theta}\left|\int_0^s u(\xi)\left(\int_{G_0} e^{-(s-\xi)L}(e^{-(t-s)L}-1)Ly(\xi)dx\right)d\xi\right|^2 dsdt$$

$$\le C_{14}\int_0^T\int_0^t\left(\int_0^s |u(\xi)|^2 d\xi\right)(t-s)^{-1-2\theta}\left(\int_0^s \|e^{-(s-\xi)L}(e^{-(t-s)L}-1)Ly(\xi)\|^2 d\xi\right)dsdt$$

$$\le C_{14}\int_0^T\int_0^t \|u\|^2_{L^2(0,T)}(t-s)^{-1-2\theta}\left(\int_0^s (t-s)^2\|L^2y(\xi)\|^2 d\xi\right)dsdt$$

$$\le C_{15}\int_0^T\left(\int_0^t (t-s)^{1-2\theta}ds\right)dt \le C_{16}$$

by the Cauchy-Schwarz inequality and $\theta < 1$. As a consequence of (6.21) and of $H^\theta(0,T)\subset C[0,T]$ we have

$$\int_0^T\int_0^t J_2 dsdt = \int_0^T\int_0^t (t-s)^{-1-2\theta}\left|\int_s^t u(\xi)\left(\int_{G_0} e^{-(t-\xi)L}Ly(\xi)dx\right)d\xi\right|^2 dsdt$$

$$\le C_{17}\int_0^T\int_0^t (t-s)^{-1-2\theta}\left(\int_s^t |u(\xi)|^2 d\xi\right)\left(\int_s^t \|e^{-(t-\xi)L}Ly(\xi)\|^2 d\xi\right)dsdt$$

$$\le C_{17}\int_0^T\int_0^t (t-s)^{-1-2\theta}(t-s)^2\|u\|^2_{H^\theta(0,T)}dsdt \le C_{18}.$$

Hence $\|\varphi y'\|_{H^\theta(0,T)} \le C_{19}$, that is, $\|\varphi y(u)\|_{H^{1+\theta}(0,T)} \le C_{19}$. The interpolation inequality implies

$$\|F(u_1)-F(u_2)\|_{H^1(0,T)} \le C_{20}\|F(u_1)-F(u_2)\|^{\frac{1}{\theta+1}}_{H^{\theta+1}(0,T)}\|F(u_1)-F(u_2)\|^{\frac{\theta}{\theta+1}}_{L^2(0,T)}$$

$$\le C_{20}(2C_{19})^{\frac{1}{\theta+1}}\|F(u_1)-F(u_2)\|^{\frac{\theta}{\theta+1}}_{L^2(0,T)}.$$

Substituting this into (6.27), we obtain

$$\|u-u^\dagger\|_{L^2(0,T)} \le C_{11}C_{20}(2C_{19})^{\frac{1}{\theta+1}}\|F(u)-F(u^\dagger)\|_V^{\frac{\theta}{\theta+1}}$$

for $\|u\|_{U_0}, \|u^\dagger\|_{U_0} \le R$. This proves the proposition. □

For classes of identification problems in partial differential equations it may be necessary to replace the power-type norm differences $\|F(u_1)-F(u_2)\|_V^\kappa$ occurring in conditional stability estimates (6.13) with $\omega(\|F(u_1)-F(u_2)\|_V)$ using more general (mostly concave) index functions $\omega$. For obtaining logarithmic rates such a generalization is useful with $\omega(\eta) = \left(\frac{1}{\log\frac{1}{\eta}}\right)^\kappa$ and some $\kappa > 0$, for example for *parabolic problems backward in time* where the forward operator $F : u(\cdot,0) \mapsto u(\cdot,T)$ is defined by solutions $u$ of the equation

$$\partial_t u = \Delta u(x,t) + f(x,t,u), \quad x\in G, \quad t > 0$$

with boundary condition $u|_{\partial G\times(0,T)} = 0$. Due to the occurrence of nonlinear terms $f$ in the equation Carleman type inequalities are required for finding conditional stability estimates. For such proof and related results we refer to [32] and references therein.

## Acknowlegdgements

The paper was started during a research stay of the first author at the *Graduate School of Mathematical Sciences of the University of Tokyo* in February/March 2009 and has been completed in Linz/Austria during the *Mini Special Semester on Inverse Problems*, May 18 - July 15, 2009, organized by RICAM, Austrian Academy of Sciences. B. Hofmann thanks both hosts for kind hospitality and allowance. The research of B. Hofmann was also supported by Deutsche Forschungsgemeinschaft (DFG) under Grant HO1454/7-2. M. Yamamoto was partly supported by Grants 20654011 and 21340021 from Japan Society for the Promotion of Science. Moreover, both authors thank Radu Ioan Boţ and Jens Flemming (TU Chemnitz) for fruitful discussions.


## References


[1] Adams, R.A. (1975): *Sobolev Spaces*. New York: Academic Press.

[2] Banks, H.T.; Kunisch, K. (1989): *Estimation Techniques for Distributed Parameter Systems*. Boston, MA: Birkhäuser.

[3] Bonesky, T.; Kazimierski, K.S.; Maass, P.; Schöpfer, F.; Schuster, T. (2008): Minimization of Tikhonov functionals in Banach spaces. *Abstract and Applied Analysis* **2008**. Article ID 192679 (19 pp), DOI:10.1155/2008/192679.

[4] Burger, M.; Osher, S. (2004): Convergence rates of convex variational regularization. *Inverse Problems* **20**, 1411–1421.

[5] Cheng, J.; Yamamoto, M. (2000): One new strategy for a priori choice of regularizing parameters in Tikhonov's regularization. *Inverse Problems* **16**, L31–L38.

[6] Düvelmeyer, D.; Hofmann, B.; Yamamoto, M. (2007): Range inclusions and approximate source conditions with general benchmark functions. *Numerical Functional Analysis and Optimization* **28**, 1245–1261.

[7] Engl, H. W.; Hanke, M.; Neubauer, A. (1996): *Regularization of Inverse Problems*. Dordrecht: Kluwer.

[8] Engl, H.W.; Kunisch, K.; Neubauer, A. (1989): Convergence rates for Tikhonov regularization of nonlinear ill-posed problems. *Inverse Problems* **5**, 523–540.

[9] Flemming, J.; Hofmann, B. (2010): A new approach to source conditions in regularization with general residual term. *Numer. Funct. Anal. Optim.* **31**. To appear. Published electronically in preliminary form as arXiv:0906.3438v1.

[10] Grasmair, M.; Haltmeier, M.; Scherzer, O. (2008): sparse regularization with $l^q$ penalty term. *Inverse Problems* **24**, 055020 (13pp).

[11] Hein, T. (2008): Convergence rates for regularization of ill-posed problems in Banach spaces by approximate source conditions. *Inverse Problems* **24**, 045007 (10pp).

[12] Hein, T. (2009): Tikhonov regularization in Banach spaces - improved convergence rates results. *Inverse Problems* **25**, 035002 (18pp).

[13] Hein, T.; Hofmann, B. (2009): Approximate source conditions for nonlinear ill-posed problems – chances and limitations. *Inverse Problems* **25**, 035003 (16pp).

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

[15] Hofmann, B.; Scherzer, O. (1994): Factors influencing the ill-posedness of nonlinear problems. *Inverse Problems* **10**, 1277–1297.

[16] Isakov, V. (2006): *Inverse Problems for Partial Differential Equations* (2nd edition). New York, NY: Springer.

[17] Kaltenbacher, B. (2008): A note on logarithmic convergence rates for nonlinear Tikhonov regularization. *J. Inv. Ill-Posed Problems* **16**, 79–88.

[18] Kaltenbacher, B.; Neubauer, A.; Scherzer, O. (2008): *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*. Berlin: Walter de Gruyter.

[19] Lorenz, D.A.; Trede, D. (2008): Optimal convergence rates for Tikhonov regularization in Besov scales. *Inverse Problems* **24**, 055010 (14pp).

[20] Neubauer, A. (1997): On converse and saturation results for Tikhonov regularization of linear ill-posed problems. *SIAM J. Numer. Anal.* **34**, 517–527.

[21] Neubauer, A. (2009): On enhanced convergence rates for Tikhonov regularization of nonlinear ill-posed problems in Banach spaces. *Inverse Problems* **25**, 065009 (10 pp).

[22] Neubauer, A.; Hein, T.; Hofmann, B.; Kindermann, S.; Tautenhahn, U. (2010): Improved and extended results for enhanced convergence rates of Tikhonov regularization in Banach spaces. *Appl. Anal.* **89**. To appear.

[23] Pazy, A. (1983): *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Berlin: Springer-Verlag.

[24] Pöschl, C. (2008): *Tikhonov Regularization with General Residual Term*. Dissertation. Innsbruck: Leopold Franzens Universität.

[25] Prilepko, A.I.; Orlovsky, D.G.; Vasin, I.A. (2000): *Methods for Solving Inverse Problems in Mathematical Physics*. New York: Marcel Dekker.

[26] Ramlau, R. (2008): Regularization properties of Tikhonov regularization with sparsity constraints. *Electronic Transactions on Numerical Analysis* **30**, 54–74.

[27] Resmerita, E. (2005): Regularization of ill-posed problems in Banach spaces: convergence rates. *Inverse Problems* **21**, 1303–1314.

[28] Resmerita, E.; Scherzer, O. (2006): Error estimates for non-quadratic regularization and the relation to enhancement. *Inverse Problems* **22**, 801–814.

[29] Scherzer, O.; Grasmair, M.; Grossauer, H.; Haltmeiner, M.; Lenzen, F. (2009): *Variational Methods in Imaging*. New York: Springer.

[30] Schöpfer, F.; Louis, A.K.; Schuster, T. (2006): Nonlinear iterative methods for linear ill-posed problems in Banach spaces. *Inverse Problems* **22**, 311–329.

[31] Schöpfer, F.; Schuster, T. (2009): Fast regularizing sequential subspace optimization in Banach spaces. *Inverse Problems* **25**, 015013 (22pp).

[32] Yamamoto, M. (2009): Carleman estimates for parabolic equations and applications. *Inverse Problems* **25**, 123013 (75pp).

[33] Zarzer, C.A. (2009): On Tikhonov regularization with non-convex sparsity constraints. *Inverse Problems* **25**, 025006 (13pp).