



# The residual method for regularizing ill-posed problems

Markus Grasmair<sup>a</sup>, Markus Haltmeier<sup>b,\*</sup>, Otmar Scherzer<sup>a,c</sup>

<sup>a</sup> Computational Science Center, University of Vienna, Nordbergstraße 15, Vienna, Austria

<sup>b</sup> Statistical Inverse Problems Group, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, Göttingen, Germany

<sup>c</sup> Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Austria

## ARTICLE INFO

### Keywords:

Ill-posed problems  
Regularization  
Residual method  
Sparsity  
Stability  
Convergence rates

## ABSTRACT

Although the *residual method*, or *constrained regularization*, is frequently used in applications, a detailed study of its properties is still missing. This sharply contrasts the progress of the theory of Tikhonov regularization, where a series of new results for regularization in Banach spaces has been published in the recent years. The present paper intends to bridge the gap between the existing theories as far as possible. We develop a stability and convergence theory for the residual method in general topological spaces. In addition, we prove convergence rates in terms of (generalized) Bregman distances, which can also be applied to non-convex regularization functionals.

We provide three examples that show the applicability of our theory. The first example is the regularized solution of linear operator equations on  $L^p$ -spaces, where we show that the results of Tikhonov regularization generalize unchanged to the residual method. As a second example, we consider the problem of density estimation from a finite number of sampling points, using the Wasserstein distance as a fidelity term and an entropy measure as regularization term. It is shown that the densities obtained in this way depend continuously on the location of the sampled points and that the underlying density can be recovered as the number of sampling points tends to infinity. Finally, we apply our theory to compressed sensing. Here, we show the well-posedness of the method and derive convergence rates both for convex and non-convex regularization under rather weak conditions.

© 2011 Elsevier Inc. Open access under [CC BY-NC-ND license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

## 1. Introduction

We study the solution of ill-posed operator equations

$$\mathbf{F}(x) = y, \quad (1)$$

where  $\mathbf{F} : X \rightarrow Y$  is an operator between the topological spaces  $X$  and  $Y$ , and  $y \in Y$  are given, noisy data, which are assumed to be close to some unknown, noise-free data  $y^\dagger \in \text{ran}(\mathbf{F})$ . If the operator  $\mathbf{F}$  is not continuously invertible, then (1) may not have a solution and, if a solution exists, arbitrarily small perturbations of the data may lead to unacceptable results.

If  $Y$  is a Banach space and the given data are known to satisfy an estimate  $\|y^\dagger - y\| \leq \beta$ , one strategy for defining an approximate solution of (1) is to solve the *constrained minimization problem*

$$\mathcal{R}(x) \rightarrow \min \quad \text{subject to } \|\mathbf{F}(x) - y\| \leq \beta. \quad (2)$$

\* Corresponding author.

E-mail addresses: [markus.grasmair@univie.ac.at](mailto:markus.grasmair@univie.ac.at) (M. Grasmair), [markus.haltmeier@mpibpc.mpg.de](mailto:markus.haltmeier@mpibpc.mpg.de) (M. Haltmeier), [otmar.scherzer@univie.ac.at](mailto:otmar.scherzer@univie.ac.at) (O. Scherzer).

Here, the *regularization term*  $\mathcal{R} : X \rightarrow [0, +\infty]$  is intended to enforce certain regularity properties of the approximate solution and to stabilize the process of solving (1). In [39,55], this strategy is called the *residual method*. It is closely related to *Tikhonov regularization*

$$\mathcal{T}(x) := \|\mathbf{F}(x) - y\|^2 + \alpha\mathcal{R}(x) \rightarrow \min, \quad (3)$$

where  $\alpha > 0$  is a regularization parameter. In the case that the operator  $\mathbf{F}$  is linear and  $\mathcal{R}$  is convex, (2) and (3) are basically equivalent, if  $\alpha$  is chosen according to Morozov's discrepancy principle (see [39, Chap. 3]).

While the theory of Tikhonov regularization has received much attention in the literature (see for instance [1,14,21,22,33,37,45,49,52,56,58]), the same cannot be said about the residual method. The existing results are mainly concerned with the existence theory of (2) and with the question of convergence, which asks whether solutions of (2) converge to a solution of (1) as  $\|y - y^\dagger\| \leq \beta \rightarrow 0$ . These problems have been treated in very general settings in [38,51] (see also [34,54,55]). Convergence rates have been derived in [6] for linear equations in Hilbert spaces and later generalized in [34] to non-linear equations in Banach spaces. Convergence rates have also been derived in [7,9,32] for the reconstruction of sparse sequences.

The problem of stability, however, that is, continuous dependence of the solution of (2) on the input data  $y$  and the presumed noise level  $\beta$ , has been hardly considered at all. One reason for the lack of results is that, in contrast to Tikhonov regularization, stability simply does not hold for general non-linear operator equations. But even for the linear case, where we indeed prove stability, so far stability theorems are non-existent in the literature. Though some results have been derived in [34], they only cover a very weak form of stability, which states that the solutions of (2) with perturbed data stay close to the solution with unperturbed data, if one additionally increases the regularization parameter  $\beta$  in the perturbed problem by a sufficient amount.

The present paper tries to generalize the existent theory on the residual method as far as possible. We assume that  $X$  and  $Y$  are mere topological spaces and consider the minimization of  $\mathcal{R}(x)$  subject to the constraint  $\mathcal{S}(\mathbf{F}(x), y) \leq \beta$ . Here  $\mathcal{S}$  is some distance like functional taking over the role of the norm in (2). In addition, we discuss the case where the operator  $\mathbf{F}$  is not known exactly. This subsumes errors due to the modeling process as well as discretizations of the problem necessary for its numerical solution. We provide different criteria that ensure stability (Lemma 3.6, Theorem 3.9 and Proposition 4.3) and convergence (Theorem 3.10 and Proposition 4.3) of the residual method. In particular, our conditions also include certain non-linear operators (see Example 4.6).

Section 5 is concerned with the derivation of convergence rates, i.e., quantitative estimates between solutions of (2) and the exact data  $y^\dagger$ . Using notions of abstract convexity, we define a generalized Bregman distance that allows us to state and prove rates on arbitrary topological spaces (see Theorem 5.4). In Section 6 we apply our general results to the case of sparse  $\ell^p$ -regularization with  $p \in (0, 2)$ . We prove the well-posedness of the method and derive convergence rates with respect to the norm in a fairly general setting. In the case of convex regularization, that is,  $p \geq 1$ , we derive a convergence rate of order  $\mathcal{O}(\beta^{1/p})$ . In the non-convex case  $0 < p < 1$ , we show that the rate  $\mathcal{O}(\beta)$  holds.

## 2. Definitions and mathematical preliminaries

Throughout the paper,  $X$  and  $Y$  denote sets. Moreover,  $\mathcal{R} : X \rightarrow [0, +\infty]$  is a functional on  $X$ , and  $\mathcal{S} : Y \times Y \rightarrow [0, +\infty]$  is a functional on  $Y \times Y$  such that  $\mathcal{S}(y, z) = 0$  if and only if  $y = z$ .

### 2.1. The residual method

For given mapping  $\mathbf{F} : X \rightarrow Y$ , given data  $y \in Y$ , and fixed parameter  $\beta \geq 0$ , we consider the constrained minimization problem

$$\mathcal{R}(x) \rightarrow \min \quad \text{subject to } \mathcal{S}(\mathbf{F}(x), y) \leq \beta. \quad (4)$$

For the analysis of the residual method (4) it is convenient to introduce the following notation.

The *feasible set*  $\Phi(\mathbf{F}, y, \beta)$ , the *value*  $v(\mathbf{F}, y, \beta)$ , and the *set of solutions*  $\Sigma(\mathbf{F}, y, \beta)$  of (4) are defined by

$$\begin{aligned} \Phi(\mathbf{F}, y, \beta) &:= \{x \in X : \mathcal{S}(\mathbf{F}(x), y) \leq \beta\}, \\ v(\mathbf{F}, y, \beta) &:= \inf\{\mathcal{R}(x) : x \in \Phi(\mathbf{F}, y, \beta)\}, \\ \Sigma(\mathbf{F}, y, \beta) &:= \{x \in \Phi(\mathbf{F}, y, \beta) : \mathcal{R}(x) = v(\mathbf{F}, y, \beta)\}. \end{aligned}$$

In particular,  $\Phi(\mathbf{F}, y, 0)$  consist of all solutions of the equation  $\mathbf{F}(x) = y$ . The elements of  $\Sigma(\mathbf{F}, y, 0)$  are therefore referred to as  *$\mathcal{R}$ -minimizing solutions of  $\mathbf{F}(x) = y$* .

In addition, for  $t \geq 0$ , we set

$$\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t) := \Phi(\mathbf{F}, y, \beta) \cap \{x \in X : \mathcal{R}(x) \leq t\}. \quad (5)$$

An immediate consequence of the above definitions is the identity

$$\Sigma(\mathbf{F}, y, \beta) = \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, v(\mathbf{F}, y, \beta)). \quad (6)$$

**Remark 2.1.** We do not assume a priori that a solution of the minimization problem (4) exists. Only in the next section shall we deduce the existence of solutions under a compactness assumption on the sets  $\Phi_{\mathcal{R}}(\mathbf{F}, \mathbf{y}, \beta, t)$ , see Theorem 3.1.

**Lemma 2.2.** The sets  $\Phi_{\mathcal{R}}(\mathbf{F}, \mathbf{y}, \beta, t)$  defined in (5) satisfy

$$\Phi_{\mathcal{R}}(\mathbf{F}, \mathbf{y}, \beta, t) \subset \Phi_{\mathcal{R}}(\mathbf{F}, \mathbf{y}, \beta + \gamma, t + \varepsilon) \tag{7}$$

for every  $\gamma, \varepsilon \geq 0$ , and

$$\Phi_{\mathcal{R}}(\mathbf{F}, \mathbf{y}, \beta, t) = \bigcap_{\gamma, \varepsilon > 0} \Phi_{\mathcal{R}}(\mathbf{F}, \mathbf{y}, \beta + \gamma, t + \varepsilon). \tag{8}$$

**Proof.** The inclusion (7) follows immediately from the definition of  $\Phi_{\mathcal{R}}$ . For the proof of (8) note that  $x \in \bigcap_{\gamma, \varepsilon > 0} \Phi_{\mathcal{R}}(\mathbf{F}, \mathbf{y}, \beta + \gamma, t + \varepsilon)$  if and only if  $\mathcal{S}(\mathbf{F}(x), \mathbf{y}) \leq \beta + \gamma$  for all  $\gamma > 0$  and  $\mathcal{R}(x) \leq t + \varepsilon$  for all  $\varepsilon > 0$ . This, however, is the case if and only if  $\mathcal{S}(\mathbf{F}(x), \mathbf{y}) \leq \beta$  and  $\mathcal{R}(x) \leq t$ , which means that  $x \in \Phi_{\mathcal{R}}(\mathbf{F}, \mathbf{y}, \beta, t)$ .  $\square$

Further properties of the value  $v$  and the sets  $\Phi, \Phi_{\mathcal{R}}$  and  $\Sigma$  are summarized in Appendix A.

### 2.2. Convergence of sets of solutions

In the next section we study convergence and stability of the residual method, that is, the behavior of the set of solutions  $\Sigma(\mathbf{F}_k, \mathbf{y}_k, \beta_k)$  for  $\beta_k \rightarrow \beta, \mathbf{y}_k \rightarrow \mathbf{y}$ , and  $\mathbf{F}_k \rightarrow \mathbf{F}$ . In [21,50], where convergence and stability of Tikhonov regularization have been investigated, the stability results are of the following form: for every sequence  $(\mathbf{y}_k)_{k \in \mathbb{N}} \rightarrow \mathbf{y}$  and every sequence of minimizers  $x_k \in \arg \min \{ \|\mathbf{F}(x) - \mathbf{y}_k\|^2 + \alpha \mathcal{R}(x) \}$  there exists a subsequence of  $(x_k)_{k \in \mathbb{N}}$  that converges to a minimizer of  $\|\mathbf{F}(x) - \mathbf{y}\|^2 + \alpha \mathcal{R}(x)$ . In this paper we prove similar results for the residual method but with a different notation using a type of convergence of sets (see, for example, [41, Section 29]).

**Definition 2.3.** Let  $\tau$  be a topology on  $X$  and let  $(\Sigma_k)_{k \in \mathbb{N}}$  be a sequence of subsets of  $X$ .

(a) The upper limit of  $(\Sigma_k)_{k \in \mathbb{N}}$  is defined as

$$\tau - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k := \bigcap_{k \in \mathbb{N}} \left( \tau - \text{cl} \bigcup_{k' \geq k} \Sigma_{k'} \right),$$

where  $\tau - \text{cl}$  denotes the closure with respect to  $\tau$ .

(b) An element  $x \in X$  is contained in the lower limit of the sequence  $(\Sigma_k)_{k \in \mathbb{N}}$ , in short

$$x \in \tau - \text{Lim inf}_{k \rightarrow \infty} \Sigma_k,$$

if for every neighborhood  $N$  of  $x$  there exists  $k_0 \in \mathbb{N}$  such that  $N \cap \Sigma_k \neq \emptyset$  for every  $k \geq k_0$ .

(c) If the lower limit and the upper limit of  $(\Sigma_k)_{k \in \mathbb{N}}$  coincide, we define

$$\tau - \text{lim}_{k \rightarrow \infty} \Sigma_k := \tau - \text{Lim inf}_{k \rightarrow \infty} \Sigma_k = \tau - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k$$

as the limit of the sequence  $(\Sigma_k)_{k \in \mathbb{N}}$ .

**Remark 2.4.** As a direct consequence of Definition 2.3, an element  $x$  is contained in the upper limit  $\tau - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k$ , if and only if for every neighborhood  $N$  of  $x$  and every  $k_0 \in \mathbb{N}$  there exists  $k \geq k_0$  with  $N \cap \Sigma_k \neq \emptyset$ .

If  $X$  satisfies the first axiom of countability, then  $x \in \tau - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k$ , if and only if there exists a subsequence  $(\Sigma_{k_j})_{j \in \mathbb{N}}$  of  $(\Sigma_k)_{k \in \mathbb{N}}$  and a sequence of elements  $x_j \in \Sigma_{k_j}$  such that  $x_j \rightarrow \tau x$  (see [41, Section 29.IV]). Note that in particular every metric space satisfies the first axiom of countability.

The following proposition clarifies the relation between the stability and convergence results in [21,50] and the results in the present paper.

**Proposition 2.5.** Let  $(\Sigma_k)_{k \in \mathbb{N}}$  be a sequence of nonempty subsets of  $X$ , and assume that there exists a compact set  $K$  such that  $\Sigma_k \subset K$  for all  $k \in \mathbb{N}$ . Then  $\tau - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k$  is non-empty.

If, in addition,  $X$  satisfies the first axiom of countability, then every sequence of elements  $x_k \in \Sigma_k$  has a subsequence converging to some element  $x \in \tau - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k$ .

**Proof.** By assumption, the sets  $S_k := \tau - \text{cl} \bigcup_{k' \geq k} \Sigma_{k'}$  form a decreasing family of non-empty, compact sets. Thus also their intersection  $\bigcap_{k \in \mathbb{N}} S_k = \tau - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k$  is non-empty (see [40, Theorem 5.1]).

Now assume that  $X$  satisfies the first axiom of countability. Then in particular every compact set is sequentially compact (see [40, Theorem 5.5]). Let now  $x_k \in \Sigma_k$  for every  $k \in \mathbb{N}$ . Then  $(x_k)_{k \in \mathbb{N}}$  is a sequence in the compact set  $K$  and therefore has a subsequence  $(x_{k_j})_{j \in \mathbb{N}}$  converging to some element  $x \in K$ . From Remark 2.4 it follows that  $x \in \tau - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k$ , which shows the assertion.  $\square$

### 2.3. Convergence of the data

In addition to the convergence of subsets  $\Sigma_k$  of  $X$ , it is necessary to define a notion of convergence on the set  $Y$  that is compatible with the distance measure  $\mathcal{S}$ .

**Definition 2.6.** The sequence  $(y_k)_{k \in \mathbb{N}} \subset Y$  converges  $\mathcal{S}$ -uniformly to  $y \in Y$ , if

$$\sup\{|\mathcal{S}(z, y_k) - \mathcal{S}(z, y)| : z \in Y\} \rightarrow 0.$$

The sequence of mappings  $\mathbf{F}_k : X \rightarrow Y$  converges locally  $\mathcal{S}$ -uniformly to  $\mathbf{F} : X \rightarrow Y$ , if

$$\sup\{|\mathcal{S}(\mathbf{F}_k(x), y) - \mathcal{S}(\mathbf{F}(x), y)| : y \in Y, x \in X, \mathcal{R}(x) \leq t\} \rightarrow 0$$

for every  $t \geq 0$ .

**Remark 2.7.** The  $\mathcal{S}$ -uniform convergence on  $Y$  is induced by the extended metric

$$\mathcal{S}_1(y_1, y_2) := \sup\{|\mathcal{S}(z, y_1) - \mathcal{S}(z, y_2)| : z \in Y\}.$$

If the distance measure  $\mathcal{S}$  itself equals a metric, then  $\mathcal{S}_1$  coincides with  $\mathcal{S}$ . Similarly, local  $\mathcal{S}$ -uniform convergence of a sequence of mappings  $\mathbf{F}_k$  equals the uniform convergence of  $\mathbf{F}_k$  on  $\mathcal{R}$ -bounded sets with respect to the extended metric

$$\mathcal{S}_2(y_1, y_2) := \sup\{|\mathcal{S}(y_1, z) - \mathcal{S}(y_2, z)| : z \in Y\}.$$

## 3. Well-posedness of the residual method

In the following we investigate the existence of minimizers, and the stability and the convergence of the residual method. Throughout the whole section we assume that  $\tau$  is a topology on  $X$ ,  $\mathbf{F} : X \rightarrow Y$  is a mapping,  $y \in Y$  are given data and  $\beta \geq 0$  is a fixed parameter.

### 3.1. Existence

We first investigate under which conditions  $\Sigma(\mathbf{F}, y, \beta)$ , the set of solutions of (4), is not empty.

**Theorem 3.1** (Existence). Assume that  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t)$  is  $\tau$ -compact for every  $t \geq 0$  and non-empty for some  $t_0 \geq 0$ . Then Problem (4) has a solution.

**Proof.** Eq. (6) and Lemma 2.2 imply the identity

$$\Sigma(\mathbf{F}, y, \beta) = \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, v(\mathbf{F}, y, \beta)) = \bigcap_{\varepsilon > 0} \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, v(\mathbf{F}, y, \beta) + \varepsilon).$$

Because  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t_0) \neq \emptyset$ , the value of (4) satisfies  $u(\mathbf{F}, y, \beta) \leq t_0 < \infty$  and therefore  $\emptyset \neq \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, v(\mathbf{F}, y, \beta) + \varepsilon)$  for every  $\varepsilon > 0$ . Consequently,  $\Sigma(\mathbf{F}, y, \beta)$  is the intersection of a decreasing family of non-empty  $\tau$ -compact sets and thus non-empty (see [40, Theorem 5.1]).  $\square$

Recall that a mapping  $\mathcal{F} : X \rightarrow [0, +\infty]$  is lower semi-continuous, if its lower level sets  $\{x \in X : \mathcal{F}(x) \leq t\}$  are closed for every  $t \geq 0$ . Moreover, the mapping  $\mathcal{F}$  is coercive, if its lower level sets are pre-compact, see [4]. (In a Banach space one often calls a functional coercive, if it is unbounded on unbounded sets. The notion used here is equivalent if the Banach space is reflexive and  $\tau$  is the weak topology.) In particular, the mapping  $\mathcal{F}$  is lower semi-continuous and coercive, if and only if its lower level sets are compact.

**Proposition 3.2.** Assume that  $\mathcal{R}$  and  $x \mapsto \mathcal{S}(\mathbf{F}(x), y)$  are lower semi-continuous and one of them, or their sum, is coercive. Then  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t)$  is  $\tau$ -compact for every  $t \geq 0$ . If additionally  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t_0)$  is non-empty for some  $t_0 \geq 0$ , then Problem (4) has a solution.

**Proof.** If  $\mathcal{R}$  and  $x \mapsto \mathcal{S}(\mathbf{F}(x), y)$  are lower semi-continuous and one of them is coercive, then

$$\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t) = \{x : \mathcal{S}(\mathbf{F}(x), y) \leq \beta\} \cap \{x : \mathcal{R}(x) \leq t\}$$

is the intersection of a closed and a  $\tau$ -compact set and therefore itself  $\tau$ -compact. In case that only the sum  $x \mapsto \mathcal{S}(\mathbf{F}(x), y) + \mathcal{R}(x)$  is coercive, the set

$$\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t) = \{x : \mathcal{S}(\mathbf{F}(x), y) \leq \beta\} \cap \{x : \mathcal{R}(x) \leq t\} \subset \{x : \mathcal{S}(\mathbf{F}(x), y) + \mathcal{R}(x) \leq \beta + t\}$$

is a closed set contained in a  $\tau$ -compact set and therefore again  $\tau$ -compact.  $\square$

The lower semi-continuity of  $x \mapsto \mathcal{S}(\mathbf{F}(x), y)$  certainly holds if  $\mathbf{F}$  is continuous and  $\mathcal{S}$  is lower semi-continuous with respect to the first component (for some given topology on  $Y$ ). It is, however, also possible to obtain lower semi-continuity, if  $\mathbf{F}$  is not continuous but the functional  $\mathcal{S}$  satisfies a stronger condition:

**Proposition 3.3.** *Let  $\tau'$  be a topology on  $Y$  such that  $z \mapsto \mathcal{S}(z, y)$  is lower semi-continuous and coercive, and assume that  $\mathbf{F} : X \rightarrow Y$  has a closed graph. Then the functional  $x \mapsto \mathcal{S}(\mathbf{F}(x), y)$  is lower semi-continuous.*

**Proof.** Because  $\mathbf{F}$  has a closed graph, the pre-image under  $\mathbf{F}$  of every compact set is closed (see [38, Theorem 4]). This shows that

$$\{x \in X : \mathcal{S}(\mathbf{F}(x), y) \leq \beta\} = \mathbf{F}^{-1}(\{z \in Y : \mathcal{S}(z, y) \leq \beta\})$$

is closed for every  $\beta$ , that is, the mapping  $x \mapsto \mathcal{S}(\mathbf{F}(x), y)$  is lower semi-continuous.  $\square$

### 3.2. Stability

Stability is concerned with the continuous dependence of the solutions of (4) of the input data, that is, the element  $y$ , the parameter  $\beta$ , and, possibly, the operator  $\mathbf{F}$ . Given sequences  $\beta_k \rightarrow \beta$ ,  $y_k \rightarrow y$ , and  $\mathbf{F}_k \rightarrow \mathbf{F}$ , we ask whether the sequence of sets  $\Sigma(\mathbf{F}_k, y_k, \beta_k)$  converges to  $\Sigma(\mathbf{F}, y, \beta)$ . As already indicated in Section 2, we will make use of the upper convergence of sets introduced in Definition 2.3. The topology, however, with respect to which the results are formulated, is stronger than  $\tau$ .

**Definition 3.4.** The topology  $\tau_{\mathcal{R}}$  on  $X$  is generated by all sets of the form  $U \cap \{x \in X : \mathcal{R}(x) > s\}$  with  $s \in \mathbb{R}$  and  $U \in \tau$  and all sets of the form  $U \cap \{x \in X : \mathcal{R}(x) < t\}$  with  $t \in \mathbb{R} \cup \{\infty\}$  and  $U \in \tau$ . (Hence  $\tau_{\mathcal{R}}$  consists of all unions of finite intersections of sets of the form  $U \cap \{x \in X : \mathcal{R}(x) > s\}$  or  $U \cap \{x \in X : \mathcal{R}(x) < t\}$ .)

Note that a sequence  $(x_k)_{k \in \mathbb{N}} \subset X$  converges to  $x$  with respect to  $\tau_{\mathcal{R}}$ , if and only if  $(x_k)_{k \in \mathbb{N}}$  converges to  $x$  with respect to  $\tau$  and satisfies  $\mathcal{R}(x_k) \rightarrow \mathcal{R}(x)$  for  $k \rightarrow \infty$ .

For the stability results we make the following assumption:

#### Assumption 3.5

1. Let  $\beta \geq 0$ , let  $y \in Y$ , and let  $\mathbf{F} : X \rightarrow Y$  be a mapping.
2. Let  $(\beta_k)_{k \in \mathbb{N}}$  be a sequence of nonnegative numbers, let  $(y_k)_{k \in \mathbb{N}}$  be a sequence in  $Y$ , and let  $(\mathbf{F}_k)_{k \in \mathbb{N}}$  be a sequence of mappings  $\mathbf{F}_k : X \rightarrow Y$ .
3. The sequence  $(\beta_k)_{k \in \mathbb{N}}$  converges to  $\beta$ , the sequence  $(y_k)_{k \in \mathbb{N}}$  converges  $\mathcal{S}$ -uniformly to  $y$ , and  $(\mathbf{F}_k)_{k \in \mathbb{N}}$  converges locally  $\mathcal{S}$ -uniformly to  $\mathbf{F}$ .
4. The sets  $\Phi_{\mathcal{R}}(\mathbf{F}_k, w, \gamma, t)$  and  $\Phi_{\mathcal{R}}(\mathbf{F}, w, \gamma, t)$  are compact for all  $w, \gamma, t$ , and  $k$ . Moreover, for every  $w, \gamma, k$  there exist some  $t_0$  such that  $\Phi_{\mathcal{R}}(\mathbf{F}_k, w, \gamma, t_0)$  and  $\Phi_{\mathcal{R}}(\mathbf{F}, w, \gamma, t_0)$  are nonempty.

The following lemma is the key result to prove stability of the residual method.

**Lemma 3.6.** *Let Assumption 3.5 hold and assume that*

$$\limsup_{k \rightarrow \infty} v(\mathbf{F}_k, y_k, \beta_k) \leq v(\mathbf{F}, y, \beta) < \infty. \tag{9}$$

Then,

$$\emptyset \neq \tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma(\mathbf{F}_k, y_k, \beta_k) \subset \Sigma(\mathbf{F}, y, \beta). \tag{10}$$

If, additionally, the set  $\Sigma(\mathbf{F}, y, \beta)$  consists of a single element  $x_{\beta}$ , then

$$\{x_{\beta}\} = \tau_{\mathcal{R}} - \lim_{k \rightarrow \infty} \Sigma(\mathbf{F}_k, y_k, \beta_k). \tag{11}$$

**Proof.** In order to simplify the notation, we define

$$\begin{aligned} \Phi_k(t) &:= \Phi_{\mathcal{R}}(\mathbf{F}_k, y_k, \beta_k, t), & \Phi(t) &:= \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t), \\ v_k &:= v(\mathbf{F}_k, y_k, \beta_k), & v &:= v(\mathbf{F}, y, \beta), \\ \Sigma_k &:= \Sigma(\mathbf{F}_k, y_k, \beta_k), & \Sigma &:= \Sigma(\mathbf{F}, y, \beta). \end{aligned}$$

Moreover we define the set  $T := \tau - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k$ . Because the topology  $\tau_{\mathcal{R}}$  is finer than  $\tau$ , it follows that  $\tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k \subset T$ . We proceed by showing that  $\emptyset \neq T \subset \Sigma$  and  $T \subset \tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k$ , which then gives the assertion (10).

The inequality (9) implies that for every  $\varepsilon > 0$  there exists some  $k_0 \in \mathbb{N}$  such that  $v_k \leq v + \varepsilon$  for all  $k \geq k_0$ . Since  $\beta_k \rightarrow \beta$ , we may additionally assume that  $\beta_k \leq \beta + \varepsilon$ . Lemma A.1 implies, after possibly enlarging  $k_0$ ,

$$\Phi_k(v_k) \subset \Phi_{\mathcal{R}}(\mathbf{F}_k, y_k, \beta + \varepsilon, v_k) \subset \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta + 2\varepsilon, v_k) \subset \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta + 2\varepsilon, v + \varepsilon) \tag{12}$$

for all  $k \geq k_0$ . Thus,

$$T = \tau - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k = \bigcap_{k \in \mathbb{N}} \left( \tau - \text{cl} \bigcup_{k' \geq k} \Sigma_{k'} \right) = \bigcap_{k \geq k_0} \left( \tau - \text{cl} \bigcup_{k' \geq k} \Phi_{k'}(v_{k'}) \right) \subset \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta + 2\varepsilon, v + \varepsilon). \tag{13}$$

The sets  $\tau - \text{cl} \bigcup_{k' \geq k} \Sigma_{k'}$  are closed and non-empty and, by assumption, the set  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta + 2\varepsilon, v + \varepsilon)$  is compact. Thus  $T$  is the intersection of a decreasing family of non-empty compact sets and therefore non-empty. Moreover, because (13) holds for every  $\varepsilon > 0$ , we have

$$\emptyset \neq T \subset \bigcap_{\varepsilon > 0} \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta + 2\varepsilon, v + \varepsilon) = \Phi(v) = \Sigma. \tag{14}$$

Next we show the inclusion  $T \subset \tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k$ . To that end, we first prove that

$$v = \lim_k v_k. \tag{15}$$

Recall that Theorem 3.1 implies that  $\Phi_k(v_k) = \Sigma_k$  is non-empty. Therefore, (12) implies that also  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta + 2\varepsilon, v_k)$  is non-empty, which in turn shows that  $v_k \geq u(\mathbf{F}, y, \beta + 2\varepsilon)$  for all  $k$  large enough. Consequently,

$$\text{Lim inf}_{k \rightarrow \infty} v_k \geq v(\mathbf{F}, y, \beta + 2\varepsilon) \tag{16}$$

for all  $\varepsilon > 0$ . From Lemma A.2 we obtain that  $v = \sup_{\varepsilon > 0} u(\mathbf{F}, y, \beta + 2\varepsilon)$ . Together with (16) and (9) this shows (15).

Let now  $x \in T$ , let  $N$  be a neighborhood of  $x$  with respect to  $\tau$ , let  $\delta > 0$  and  $k_0 \in \mathbb{N}$ . Since  $T \subset \Sigma$  (see (14)), it follows that  $\mathcal{R}(x) = v$ . Thus it follows from (15) that there exists  $k_1 \geq k_0$  such that

$$|v_k - \mathcal{R}(x)| < \delta$$

for all  $k \geq k_1$ . In particular,

$$\Sigma_k \subset \{\tilde{x} \in X : \mathcal{R}(x) - \delta < \mathcal{R}(\tilde{x}) < \mathcal{R}(x) + \delta\} \tag{17}$$

for all  $k \geq k_1$ . Remark 2.4 implies that there exists  $k_2 \geq k_1$  such that

$$N \cap \Sigma_{k_2} \neq \emptyset. \tag{18}$$

Now recall that the sets  $N \cap \{\tilde{x} \in X : \mathcal{R}(x) - \delta < \mathcal{R}(\tilde{x}) < \mathcal{R}(x) + \delta\}$  form a basis of neighborhoods of  $x$  for the topology  $\tau_{\mathcal{R}}$ . Therefore (17) and (18), and the characterization of the upper limit of sets given in Remark 2.4 imply that  $x \in \tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma_k$ . Thus the inclusion (10) follows.

If the set  $\Sigma(\mathbf{F}, y, \beta)$  consists of a single element  $x_{\beta}$ , then the first part of the assertion implies that for every subsequence  $(k_j)_{j \in \mathbb{N}}$  we have

$$\tau_{\mathcal{R}} - \text{Lim sup}_{j \rightarrow \infty} \Sigma(\mathbf{F}_{k_j}, y_{k_j}, \beta_{k_j}) = \{x_{\beta}\}.$$

Thus the assertion follows from Lemma A.3.  $\square$

The crucial condition in Lemma 3.6 is the inequality (9). Indeed, one can easily construct examples, where this condition fails and the solution of Problem (4) is unstable, see Example 3.7. What happens in this example is that the upper limit  $\tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma(\mathbf{F}, y_k, \beta)$  consists of local minima of  $\mathcal{R}$  on  $\Phi(\mathbf{F}, y, \beta)$  that fail to be global minima of  $\mathcal{R}$  restricted to  $\Phi(\mathbf{F}, y, \beta)$ .

**Example 3.7.** Consider the function  $\mathbf{F} : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^3 - x^2$  (illustrated in Fig. 1) and the regularization functional  $\mathcal{R}(x) = x^2$ . Let  $y > 0$  and choose  $\beta = y$ . Then

$$\arg \min\{\mathcal{R}(x) : |\mathbf{F}(x) - y| \leq \beta\} = \arg \min\{x^2 : |x^3 - x^2 - y| \leq y\} = 0. \tag{19}$$

Now let  $y_k > y$ . Then

$$\arg \min\{\mathcal{R}(x) : |\mathbf{F}(x) - y_k| \leq \beta\} = \arg \min\{x^2 : |x^3 - x^2 - y_k| \leq y\} = x_k,$$

where  $x_k$  is the unique solution of the equation  $\mathbf{F}(x) = y_k - y$ . Thus, if the sequence  $(y_k)_{k \in \mathbb{N}}$  converges to  $y$  from above, we have  $x_k > 1$  for all  $k$  and  $\lim_{k \rightarrow \infty} x_k = 1$ . According to (19), however, the solution of the limit problem equals zero.

The following two theorems are central results of this paper. They answer the question to which extent we obtain stability results for the residual method similar to the ones known for Tikhonov regularization.

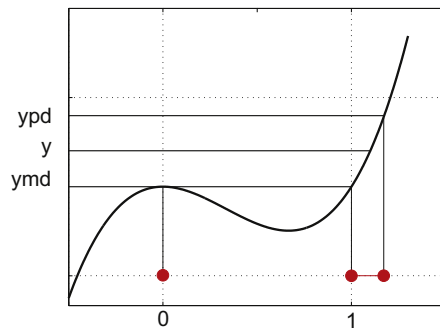


Fig. 1. The nonlinear function  $F$  from Example 3.7. The feasible set  $\Phi(\mathbf{F}, y, \beta) = \{x \in \mathbb{R} : |\mathbf{F}(x) - y| \leq \beta\}$  consists of an interval and the isolated point  $\{0\}$ .

**Theorem 3.8** (Approximate Stability). *Let Assumption 3.5 hold. Then there exists a sequence  $\varepsilon_k \rightarrow 0$  such that*

$$\emptyset \neq \tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma(\mathbf{F}_k, y_k, \beta_k + \varepsilon_k) \subset \Sigma(\mathbf{F}, y, \beta).$$

**Proof.** Define

$$\varepsilon_k := \inf\{\varepsilon > 0 : \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, v(\mathbf{F}, y, \beta)) \subset \Phi_{\mathcal{R}}(\mathbf{F}_k, y_k, \beta_k + \varepsilon, v(\mathbf{F}, y, \beta))\}.$$

Lemma A.1 and the assumption that  $\beta_k \rightarrow \beta$  imply that  $\varepsilon_k \rightarrow 0$ . Since by assumption

$$\emptyset \neq \Sigma(\mathbf{F}, y, \beta) = \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, v(\mathbf{F}, y, \beta)) \subset \Phi_{\mathcal{R}}(\mathbf{F}_k, y_k, \beta_k + \varepsilon_k, v(\mathbf{F}, y, \beta)),$$

we obtain that  $v(\mathbf{F}_k, y_k, \beta_k + \varepsilon_k) \leq v(\mathbf{F}, y, \beta)$ . Thus the assertion follows from Lemma 3.6.  $\square$

Theorem 3.8, is a stability result in the same spirit as the one derived in [34]. While it does not assert that, in the general setting described by Assumption 3.5, the residual method is stable in the sense that the solutions depend continuously on the input data, it does state that the solutions of the perturbed problems stay close to the solution of the original problem, if one allows the regularization parameter  $\beta$  to increase slightly. Apart from the more general, topological setting, the main difference to [34, Lemma 2.2] is the additional inclusion of operator errors into the result.

The next theorem provides a true stability theorem, including both data as well as operator perturbations.

**Theorem 3.9** (Stability). *Let Assumption 3.5 hold with  $\beta > 0$  and assume that the inclusion*

$$\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t) \subset \bigcap_{\delta > 0} \left( \tau - \text{cl} \bigcup_{\varepsilon > 0} \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta - \varepsilon, t + \delta) \right) \tag{20}$$

holds for every  $t \geq 0$ . Then,

$$\emptyset \neq \tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma(\mathbf{F}_k, y_k, \beta_k) \subset \Sigma(\mathbf{F}, y, \beta). \tag{21}$$

If, additionally, the set  $\Sigma(\mathbf{F}, y, \beta)$  consists of a single element  $x_{\beta}$ , then

$$\{x_{\beta}\} = \tau_{\mathcal{R}} - \lim_{k \rightarrow \infty} \Sigma(\mathbf{F}_k, y_k, \beta_k).$$

**Proof.** The convergence of  $(\beta_k)_{k \in \mathbb{N}}$  to  $\beta$  and Lemma A.1 imply that for every  $\varepsilon > 0$  and  $t \in \mathbb{R}$  there exists  $k_0 \in \mathbb{N}$  such that

$$\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta - \varepsilon, t) \subset \Phi_{\mathcal{R}}(\mathbf{F}_k, y_k, \beta_k, t)$$

for all  $k \geq k_0$ . Consequently,

$$\limsup_{k \rightarrow \infty} v(\mathbf{F}_k, y_k, \beta_k) = \limsup_{k \rightarrow \infty} (\inf\{t : \Phi_{\mathcal{R}}(\mathbf{F}_k, y_k, \beta_k, t) \neq \emptyset\}) \leq \inf_{\varepsilon > 0} (\inf\{t : \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta - \varepsilon, t) \neq \emptyset\}).$$

From (20) we obtain that

$$\inf_{\varepsilon > 0} (\inf\{t : \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta - \varepsilon, t) \neq \emptyset\}) \leq \inf\{t : \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t) \neq \emptyset\} = v(\mathbf{F}, y, \beta).$$

This shows (9). Now (21) follows from Lemma 3.6.  $\square$

For Theorem 3.9 to hold, the mapping  $x \rightarrow S(\mathbf{F}(x), y)$  has to satisfy the additional regularity property (20). This property requires that every  $x \in X$  for which  $\mathbf{F}(x) \neq y$  can be approximated by elements  $\tilde{x}$  with  $S(\mathbf{F}(\tilde{x}), y) < S(\mathbf{F}(x), y)$  and



$\mathcal{R}(\bar{x}) \leq \mathcal{R}(x) + \beta$ . That is, the function  $x \mapsto \mathcal{S}(\mathbf{F}(x), y)$  does not have local minima in the sets  $\{x \in X : \mathcal{R}(x) < t\}$ . As will be shown in the following Section 4, this property is naturally satisfied for linear operators on Banach spaces.

### 3.3. Convergence

The following theorem states the solutions obtained with the residual method indeed converge to the  $\mathcal{R}$ -minimizing solution of the equation  $\mathbf{F}(x) = y$ , if the noise level decreases to zero. Recall that the set of all  $\mathcal{R}$ -minimizing solution of the equation  $F(x) = y$  is given by  $\Sigma(\mathbf{F}, y, 0)$ .

**Theorem 3.10** (Convergence). *Let  $y \in Y$  be such that there exists  $x \in X$  with  $\mathbf{F}(x) = y$  and  $\mathcal{R}(x) < \infty$  and assume that  $\Phi_{\mathcal{R}}(\mathbf{F}, w, \gamma, t)$  is  $\tau$ -compact for all  $w \in Y$  and  $\gamma, t \geq 0$ . If  $(y_k)_{k \in \mathbb{N}}$  converges  $\mathcal{S}$ -uniformly to  $y$  and satisfies  $\mathcal{S}(y, y_k) \leq \beta_k \rightarrow 0$ , then*

$$\limsup_{k \rightarrow \infty} v(\mathbf{F}, y_k, \beta_k) \leq v(\mathbf{F}, y, 0) < \infty. \quad (22)$$

In particular,

$$\emptyset \neq \tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma(\mathbf{F}, y_k, \beta_k) \subset \Sigma(\mathbf{F}, y, 0). \quad (23)$$

If, additionally, the  $\mathcal{R}$ -minimizing solution  $x^\dagger$  is unique, then

$$\{x^\dagger\} = \tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma(\mathbf{F}, y_k, \beta_k). \quad (24)$$

**Proof.** By assumption  $\mathcal{S}(y, y_k) \leq \beta_k$ , which implies that  $v(\mathbf{F}, y_k, \beta_k) \leq \mathcal{R}(x')$  for all  $x' \in \Phi(\mathbf{F}, y, 0)$ . This proves (22). Now (23) and (24) follow from Lemma 3.6.  $\square$

## 4. Linear spaces

Now we assume that  $X$  and  $Y$  are subsets of topological vector spaces. Then the linear structures allows us to introduce more tangible conditions implying stability of the residual method.

For the following we assume that  $\mathbf{F} : X \rightarrow Y$  and  $y \in Y$  are fixed.

**Assumption 4.1.** Assume that the following hold:

1. The set  $X$  is a convex subset of a topological vector space, and  $Y$  is a topological vector space.
2. The mapping  $x \mapsto \mathcal{S}(\mathbf{F}(x_0), y)$  is semi-strictly quasi-convex. That is, for all  $x_0, x_1 \in X$  with  $\mathcal{S}(\mathbf{F}(x_0), y), \mathcal{S}(\mathbf{F}(x_1), y) < \infty$ , and all  $0 < \lambda < 1$  we have

$$\mathcal{S}(\mathbf{F}(\lambda x_0 + (1 - \lambda)x_1), y) \leq \max\{\mathcal{S}(\mathbf{F}(x_0), y), \mathcal{S}(\mathbf{F}(x_1), y)\}.$$

Moreover, the inequality is strict whenever  $\mathcal{S}(\mathbf{F}(x_0), y) \neq \mathcal{S}(\mathbf{F}(x_1), y)$ .

3. For every  $\beta \geq 0$  there exists  $x \in X$  with  $\mathcal{S}(\mathbf{F}(x), y) \leq \beta$  and  $\mathcal{R}(x) < \infty$ .
4. The domain  $\text{dom}\mathcal{R} = \{x \in X : \mathcal{R}(x) < +\infty\}$  of  $\mathcal{R}$  is convex and for every  $x_0, x_1 \in \text{dom}\mathcal{R}$  the restriction of  $\mathcal{R}$  to

$$L = \{\lambda x_0 + (1 - \lambda)x_1 : 0 \leq \lambda \leq 1\}$$

is continuous.

We now show that Assumption 4.1 implies the main condition of the stability result Theorem 3.9, the inclusion (20):

**Lemma 4.2.** Assume that Assumption 4.1 holds. Then (20) is satisfied.

**Proof.** Let  $x_0 \in \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t)$  for some  $\beta > 0$ . We have to show that for every neighborhood  $N \subset X$  of  $x_0$  and every  $\delta > 0$  there exist  $\varepsilon > 0$  and  $x' \in N$  such that  $x' \in \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta - \varepsilon, t + \delta)$ .

Item 3 in Assumption 4.1 implies the existence of some  $x_1 \in X$  satisfying the inequalities  $\mathcal{S}(\mathbf{F}(x_1), y) \leq \beta/2$  and  $\mathcal{R}(x_1) < \infty$ . Since we have  $\mathcal{S}(\mathbf{F}(x_1), y) < \beta$  and  $\mathcal{S}(\mathbf{F}(x_0), y) \leq \beta$ , we obtain from Item 2 that  $\mathcal{S}(\mathbf{F}(x), y) < \beta$  for every  $x \in L := \{\lambda x_0 + (1 - \lambda)x_1 : 0 \leq \lambda < 1\}$ . Since  $x_0, x_1 \in \text{dom}\mathcal{R}$ , it follows from Item 4 that  $\mathcal{R}$  is continuous on  $L$ . Consequently  $\lim_{\lambda \rightarrow 1} \mathcal{R}(\lambda x_0 + (1 - \lambda)x_1) = \mathcal{R}(x_0) \leq t$ . In particular, there exists  $\lambda_0 < 1$  such that  $\mathcal{R}(\lambda x_0 + (1 - \lambda)x_1) \leq t + \delta$  for all  $1 > \lambda > \lambda_0$ . Since  $X$  is a topological vector space (Item 1), it follows that  $x' := \lambda x_0 + (1 - \lambda)x_1 \in N$  for some  $1 > \lambda > \lambda_0$ . This shows the assertion with  $\varepsilon := \beta - \mathcal{S}(\mathbf{F}(x'), y) > 0$ .  $\square$

Lemma 4.2 allows us to apply the stability result Theorem 3.9, which shows that Assumption 4.1 implies the continuous dependence of the solutions of (4) on the data  $y$  and the regularization parameter  $\beta$ .



**Proposition 4.3** (Stability & Convergence). *Let Assumption 4.1 hold and assume that the sets  $\Phi_{\mathcal{R}}(\mathbf{F}, w, \gamma, t)$  are compact for every  $\gamma \in \mathbb{R}$ ,  $t \in \mathbb{R}$ , and  $w \in Y$ . Assume moreover that  $(y_k)_{k \in \mathbb{N}}$  converges  $\mathcal{S}$ -uniformly to  $y \in Y$ , and that  $\beta_k \rightarrow \beta$ . If  $\beta = 0$ , assume in addition that  $\mathcal{S}(y, y_k) \leq \beta_k$ . Then*

$$\emptyset \neq \tau_{\mathcal{R}} - \limsup_{k \rightarrow \infty} \Sigma(\mathbf{F}, y_k, \beta_k) \subset \Sigma(\mathbf{F}, y, \beta).$$

If, additionally, the set  $\Sigma(\mathbf{F}, y, \beta)$  consists of a single element  $x_{\beta}$ , then

$$\{x_{\beta}\} = \tau_{\mathcal{R}} - \limsup_{k \rightarrow \infty} \Sigma(\mathbf{F}, y_k, \beta_k).$$

**Proof.** If  $\beta = 0$ , the assertion follows from Theorem 3.10. In the case  $\beta > 0$ , Lemma 4.2 implies that (20) holds. Thus, the assertion follows from Theorem 3.9. Note that the non-emptiness of the sets  $\Phi_{\mathcal{R}}(\mathbf{F}, w, \gamma, t)$  for some  $t$  follows from Item 3 in Assumption 4.1.  $\square$

**Proposition 4.4** (Stability). *Let Assumption 4.1 hold. Assume that  $(y_k)_{k \in \mathbb{N}}$  converges  $\mathcal{S}$ -uniformly to  $y \in Y$ , the mappings  $\mathbf{F}_k : X \rightarrow Y$  converge locally  $\mathcal{S}$ -uniformly to  $\mathbf{F} : X \rightarrow Y$  (see Definition 2.6), and  $\beta_k \rightarrow \beta > 0$ . Assume that the sets  $\Phi_{\mathcal{R}}(\mathbf{F}_k, w, \gamma, t)$  and  $\Phi_{\mathcal{R}}(\mathbf{F}, w, \gamma, t)$  are compact for every  $\gamma \geq 0, t \in \mathbb{R}$ , and  $w \in Y$ . Then*

$$\emptyset \neq \tau_{\mathcal{R}} - \limsup_{k \rightarrow \infty} \Sigma(\mathbf{F}_k, y_k, \beta_k) \subset \Sigma(\mathbf{F}, y, \beta).$$

If, additionally, the set  $\Sigma(\mathbf{F}, y, \beta)$  consists of a single element  $x_{\beta}$ , then

$$\{x_{\beta}\} = \tau_{\mathcal{R}} - \lim_{k \rightarrow \infty} \Sigma(\mathbf{F}_k, y_k, \beta_k).$$

**Proof.** Again, Lemma 4.2 shows that (20) holds. Moreover, the non-emptiness of the sets  $\Phi_{\mathcal{R}}(\mathbf{F}, w, \gamma, t)$  and  $\Phi_{\mathcal{R}}(\mathbf{F}_k, w, \gamma, t)$  (at least for  $k$  sufficiently large) for some  $t$  follows from Item 3 in Assumption 4.1 and the local  $\mathcal{S}$ -uniform convergence of the mappings  $\mathbf{F}_k$  to  $\mathbf{F}$ . Thus the assertion follows from Theorem 3.9.  $\square$

Item 2 in Assumption 4.1 is concerned with the interplay of the functional  $\mathbf{F}$  and the distance measure  $\mathcal{S}$ . The next two examples consider two situations, where this part of the assumption holds. Example 4.5 considers linear operators  $\mathbf{F}$  and convex distance measures  $\mathcal{S}$ . Example 4.6 introduces a class of non-linear operators on Hilbert spaces, where Item 2 is satisfied if the distance measure equals the squared Hilbert space norm.

**Example 4.5.** Assume that  $\mathbf{F} : X \rightarrow Y$  is linear and  $\mathcal{S}$  is convex in its first component. Then Item 2 in Assumption 4.1 is satisfied. Indeed, in such a situation,

$$\mathcal{S}(\mathbf{F}(\lambda x_0 + (1 - \lambda)x_1), y) = \mathcal{S}(\lambda \mathbf{F}(x_0) + (1 - \lambda)\mathbf{F}(x_1), y) \leq \lambda \mathcal{S}(\mathbf{F}(x_0), y) + (1 - \lambda)\mathcal{S}(\mathbf{F}(x_1), y) \leq \max\{\mathcal{S}(\mathbf{F}(x_0), y), \mathcal{S}(\mathbf{F}(x_1), y)\}.$$

If moreover,  $\mathcal{S}(\mathbf{F}(x_0), y) \neq \mathcal{S}(\mathbf{F}(x_1), y)$  and  $0 < \lambda < 1$ , then the last inequality is strict.

**Example 4.6.** Assume that  $Y$  is a Hilbert space,  $\mathcal{S}(y, z) = \|y - z\|^2$ , and  $\mathbf{F} : X \rightarrow Y$  is two times Gâteaux differentiable. Then Item 2 in Assumption 4.1 holds if for all  $x_0 \neq x_1 \in X$  the mapping

$$t \mapsto \varphi(t; x_0, x_1) := \|\mathbf{F}(x_0 + tx_1) - y\|^2$$

has no local maxima. This condition holds, if the inequality  $\partial_t^2 \varphi(0; x_0, x_1) > 0$  is satisfied whenever  $\partial_t \varphi(0; x_0, x_1) = 0$ . The computation of the derivative of  $\varphi(\cdot; x_0, x_1)$  at zero yields that

$$\partial_t \varphi(0; x_0, x_1) = 2 \langle \mathbf{F}'(x_0)(x_1), \mathbf{F}(x_0) \rangle$$

and

$$\partial_t^2 \varphi(0; x_0, x_1) = 2 \langle \mathbf{F}''(x_0)(x_1; x_1), \mathbf{F}(x_0) \rangle + 2 \|\mathbf{F}'(x_0)x_1\|^2.$$

Consequently, Item 2 in Assumption 4.1 is satisfied if, for every  $x_0, x_1 \in X$  with  $x_1 \neq 0$ , the equality  $\langle \mathbf{F}'(x_0)(x_1), \mathbf{F}(x_0) \rangle = 0$  implies that

$$\langle \mathbf{F}''(x_0)(x_1; x_1), \mathbf{F}(x_0) \rangle + \|\mathbf{F}'(x_0)(x_1)\|^2 > 0.$$

#### 4.1. Regularization on $L^p$ -spaces

Let  $p \in (1, \infty)$  and set  $X = L^p(\Omega, \mu)$  for some  $\sigma$ -finite measure space  $(\Omega, \mu)$ . Assume that  $Y$  is a Banach space and  $\mathbf{F} : X \rightarrow Y$  is a bounded linear operator with dense range. Let  $\mathcal{R}(x) = \|x\|_p^p$  and  $\mathcal{S}(w, y) = \|w - y\|$ . We thus consider the minimization problem

$$\|x\|_p^p \rightarrow \min \quad \text{subject to } \|Fx - y\| \leq \beta.$$

We now show that in this situation the assumptions of Proposition 4.3 are satisfied. To that end, let  $\tau$  be the weak topology on  $L^p(\Omega, \mu)$ . As  $L^p(\Omega, \mu)$  is reflexive, the level sets  $\{x \in X : \mathcal{R}(x) \leq t\}$  are weakly compact. Moreover, the mapping  $x \mapsto \|Fx - y\|$  is weakly lower semi-continuous. Thus all the sets  $\Phi_{\mathcal{R}}(\mathbf{F}, w, \gamma, t)$  are weakly compact. Example 4.5 shows that Item 2 in Assumption 4.1 holds. Item 3 follows from the density of the range of  $\mathbf{F}$ . Finally, Item 4 holds, because  $\mathcal{R}$  is norm continuous and convex.

Now assume that  $y_k \rightarrow y$  and  $\beta_k \rightarrow \beta$ . If  $\beta = 0$  assume in addition that  $\|y_k - y\| \leq \beta_k$ . The strict convexity of  $\mathcal{R}$  and convexity of the mappings  $x \mapsto \|Fx - y_k\|$  imply that each set  $\Sigma(\mathbf{F}, y_k, \beta_k)$  consists of a single element  $x_k$ . Similarly,  $\Sigma(\mathbf{F}, y, \beta)$  consists of a single element  $x^\dagger$ . From Proposition 4.3 we now obtain that  $(x_k)_{k \in \mathbb{N}}$  weakly converges to  $x^\dagger$  and  $\|x_k\|_p^p \rightarrow \|x^\dagger\|_p^p$ . Thus, in fact, the sequence  $(x_k)_{k \in \mathbb{N}}$  strongly converges to  $x^\dagger$  (see [44, Corollary 5.2.19]).

Let  $\beta > 0$  and assume that  $\mathbf{F}_k : X \rightarrow Y$  is a sequence of bounded linear operators converging to  $\mathbf{F}$  with respect to the strong topology on  $L(X, Y)$ , that is,  $\sup\{\|\mathbf{F}_k x - \mathbf{F}x\| : \|x\| \leq 1\} \rightarrow 0$ . Let again  $\beta_k \rightarrow \beta$  and  $y_k \rightarrow y$ , and denote by  $x_k$  the single element in  $\Sigma(\mathbf{F}_k, y_k, \beta_k)$ . Applying Proposition 4.4, we again obtain that  $x_k \rightarrow x^\dagger$ .

**Remark 4.7.** The above results rely heavily on the assumption that  $p > 1$ , which implies that the space  $L^p(\Omega, \mu)$  is reflexive. In the case  $X = L^1(\Omega, \mu)$ , the level sets  $\{x \in X : \|x\|_1 \leq t\}$  fail to be weakly compact, and thus even the existence of a solution of Problem (4) need not hold.

**Remark 4.8.** The assertions concerning stability and convergence with respect to the norm topology remain valid, if  $X$  is any uniformly convex Banach space and  $\mathcal{R}$  the norm on  $X$  to some power  $p > 1$ . Also in this case, weak convergence and convergence of norms imply the strong convergence of a sequence [44, Theorem 5.2.18]. More generally, this property is called the Radon–Riesz property [44, p. 453]. Spaces satisfying this property are also called Efimov–Stechkin spaces in [55].

#### 4.2. Regularization of probability measures

Let  $(\Omega, d)$  be a separable, complete metric space with distance  $d$  and denote by  $\mathcal{P}(\Omega)$  the space of probability measures on the Borel sets of  $\Omega$ . That is,  $\mathcal{P}(\Omega)$  consists of all positive Borel measures  $\mu$  on  $\Omega$  that satisfy  $\mu(\Omega) = 1$ . For  $p \geq 1$  the  $p$ -Wasserstein distance on  $\mathcal{P}(\Omega)$  is defined as

$$W_p(\mu, \nu) := \left( \inf \left\{ \int d(x, y)^p d\xi : \xi \in \mathcal{P}(\Omega \times \Omega), \pi_{\#}^1 \xi = \mu, \pi_{\#}^2 \xi = \nu \right\} \right)^{1/p}.$$

Here  $\pi_{\#}^i \xi$  denotes the push forward of the measure  $\xi$  by means of the  $i$ th projection. In other words,  $\pi_{\#}^1 \xi(U) = \xi(U \times \Omega)$  and  $\pi_{\#}^2 \xi(U) = \xi(\Omega \times U)$  for every Borel set  $U \subset \Omega$ .

Recall that the narrow topology on  $\mathcal{P}(\Omega)$  is induced by the action of elements of  $\mathcal{P}(\Omega)$  on continuous functions  $u \in C(\Omega)$ . That is, a sequence  $(\mu_k)_{k \in \mathbb{N}} \subset \mathcal{P}(\Omega)$  converges narrowly to  $\mu \in \mathcal{P}(\Omega)$ , if

$$\int_{\Omega} u d\mu_k = \int_{\Omega} u d\mu \quad \text{for all } u \in C(\Omega).$$

**Lemma 4.9.** Let  $p \geq 1$ . Then the Wasserstein distance satisfies, for every  $\mu_1, \mu_2, \nu \in \mathcal{P}(\Omega)$  and  $0 \leq \lambda \leq 1$ , the inequality

$$W_p(\lambda\mu_1 + (1 - \lambda)\mu_2, \nu)^p \leq \lambda W_p(\mu_1, \nu)^p + (1 - \lambda)W_p(\mu_2, \nu)^p. \tag{25}$$

Moreover it is lower semi-continuous with respect to the narrow topology.

**Proof.** The lower semi-continuity of  $W_p$  has, for instance, been shown in [27]. In order to show the inequality (25), let  $\xi_1, \xi_2 \in \mathcal{P}(\Omega \times \Omega)$  be two measures that realize the infimum in the definition of  $W_p(\mu_1, \nu)$  and  $W_p(\mu_2, \nu)$ , respectively. Then  $\pi_{\#}^1(\lambda\xi_1 + (1 - \lambda)\xi_2) = \lambda\mu_1 + (1 - \lambda)\mu_2$  and  $\pi_{\#}^2(\lambda\xi_1 + (1 - \lambda)\xi_2) = \nu$ , which implies that the measure  $\lambda\xi_1 + (1 - \lambda)\xi_2$  is admissible for measuring the distance between  $\lambda\mu_1 + (1 - \lambda)\mu_2$  and  $\nu$ . Therefore

$$\begin{aligned} W_p(\lambda\mu_1 + (1 - \lambda)\mu_2, \nu)^p &= \inf \left\{ \int d(x, y)^p d\xi : \pi_{\#}^1 \xi = \lambda\mu_1 + (1 - \lambda)\mu_2, \pi_{\#}^2 \xi = \nu \right\} \leq \int d(x, y)^p d(\lambda\xi_1 + (1 - \lambda)\xi_2) \\ &= \lambda W_p(\mu_1, \nu)^p + (1 - \lambda)W_p(\mu_2, \nu)^p, \end{aligned}$$

which proves the assertion.  $\square$

Since  $\mathcal{P}(\Omega)$  is a convex subset of the space  $\mathcal{M}(\Omega)$  of all finite Radon measures on  $\Omega$ , and the narrow topology on  $\mathcal{P}(\Omega)$  is the restriction of the weak\* topology on  $\mathcal{M}(\Omega)$  considered as the dual of  $C_b(\Omega)$ , the space of bounded continuous functions on  $\Omega$ , it is possible to apply the results of this section also to the situation where  $Y = \mathcal{P}(\Omega)$  and  $\mathcal{S} = W_p$ . As an easy example, we consider the problem of density estimation from a finite number of measurements.

**Example 4.10.** Let  $\Omega \subset \mathbb{R}^n$  be an open domain. Given a finite number of measurements  $\{y_1, \dots, y_k\} \subset \Omega$ , the task of density estimation is the problem of finding a simple density function  $u$  on  $\Omega$  in such a way that the measurements look like a typical sample of the distribution defined by  $u$ . Interpreting the measurements as a normalized sum of delta peaks, that is, equating  $\{y_1, \dots, y_k\}$  with the measure  $\mathbf{y} := \frac{1}{k} \sum_i \delta(y_i) \in \mathcal{P}(\Omega)$ , we can easily translate the problem into the setting of this paper.

We set  $X := \{u \in L^1(\Omega) : u \geq 0 \text{ and } \|u\|_1 = 1\}$ , which is a convex and closed subset of  $L^1(\Omega)$ ,  $Y := \mathcal{P}(\Omega)$ , and consider the embedding  $\mathbf{F} : X \rightarrow \mathcal{P}(\Omega)$ ,  $u \mapsto u\mathcal{L}^n$ . Then  $\mathbf{F}$  is continuous with respect to the weak topology on  $X$  and the narrow topology on  $\mathcal{P}(\Omega)$ . We now consider the distance measure  $S = W_p$  for some  $p \geq 1$  and the Euclidean distance  $d$  on  $\Omega$ . Then Lemma 4.9 implies that, for every  $\mu \in \mathcal{P}(\Omega)$ , the mapping  $u \mapsto W_p(Fu, \mu)$  is weakly lower semi-continuous.

There are several possibilities for choosing a regularization functional on  $X$ . If  $\Omega$  is bounded (or at least  $\mathcal{L}^n(\Omega) < \infty$ ), one can, for instance, use the Boltzmann–Shannon entropy defined by

$$\mathcal{R}(u) := \int_{\Omega} u \log(u) dx \quad \text{for } u \in X.$$

Then the theorems of De la Vallée Poussin and Dunford–Pettis (see [24, Theorems 2.29, 2.54]) show that the lower level sets of  $\mathcal{R}$  are weakly pre-compact in  $L^1(\Omega)$ . Moreover, the functional  $\mathcal{R}$  is convex and therefore weakly lower semi-continuous (see [24, Theorem 5.14]). Using Proposition 3.2, we therefore obtain that the compactness required in Assumption 3.5 holds. Also, Lemma 4.9 shows that Item 2 in Assumption 4.1 holds. Items 1 and 4 are trivially satisfied. Finally, Item 3 follows from the density of  $\text{dom} R$  in  $X$  and the density (with respect to the narrow topology) of  $\text{ran } \mathbf{F}$  in  $\mathcal{P}(\Omega)$ . In addition, it has been shown in [59] that the weak convergence of a sequence  $(u_k)_{k \in \mathbb{N}} \subset L^1(\Omega)$  to  $u \in L^1(\Omega)$  together with the convergence  $\mathcal{R}(u_k) \rightarrow \mathcal{R}(u)$  imply that  $\|u_k - u\|_1 \rightarrow 0$ . Thus the topology  $\tau_{\mathcal{R}}$  coincides with the strong topology on  $X$ .

Proposition 4.3 therefore implies that the residual method is a stable and convergent regularization method with respect to the strong topology on  $X$ . More precisely, given a sample  $\mathbf{y} = \frac{1}{k} \sum_i \delta(y_i)$ , the density estimate  $u$  depends continuously on the positions  $y_i$  of the measurements and on the regularization parameter  $\beta$ . In addition, if the number of measurements increases, then the Wasserstein distance between the sample and the true probability converges almost surely to zero. Thus also the reconstructed density converges to the true underlying density, provided the regularization parameters decrease to zero slowly enough.

### 5. Convergence rates

In this section we derive quantitative estimates (convergence rates) for the difference between regularized solutions  $x_{\beta} \in \Sigma(\mathbf{F}, \mathbf{y}, \beta)$  and the exact solution of the equation  $\mathbf{F}(x^{\dagger}) = \mathbf{y}^{\dagger}$ .

For Tikhonov regularization, convergence rates have been derived in [3,6,23,36,46,47] in terms of the Bregman distance. However, its classical definition,

$$D_{\xi}(x, x^{\dagger}) = \mathcal{R}(x) - \mathcal{R}(x^{\dagger}) + \langle \xi, x^{\dagger} - x \rangle_{X^*, X}, \tag{26}$$

where  $\xi \in \partial \mathcal{R}(x^{\dagger}) \subset X^*$ , requires the space  $X$  to be linear and the functional  $\mathcal{R}$  to be convex, as the (standard) subdifferential  $\partial \mathcal{R}(x^{\dagger})$  is only defined for convex functionals. In the sequel we will extend the notion of subdifferentials and Bregman distances to work for arbitrary functionals  $\mathcal{R}$  on arbitrary sets  $X$ . To that end, we make use of a generalized notion of convexity, which is not based on the duality between a Banach space  $X$  and its dual  $X^*$  but on more general pairings (see [53]). The same notion has recently been used in [29] for the derivation of convergence rates for non-convex regularization functionals.

**Definition 5.1** (Generalized Bregman Distance). Let  $W$  be a set of functions  $w : X \rightarrow \mathbb{R}$ , let  $\mathcal{R} : X \rightarrow \mathbb{R} \cup \{+\infty\}$  be a functional and let  $x^{\dagger} \in X$ .

(a) The functional  $\mathcal{R}$  is convex at  $x^{\dagger}$  with respect to  $W$ , if

$$\mathcal{R}(x^{\dagger}) = \mathcal{R}^{**}(x^{\dagger}) := \sup_{w \in W} \left( \inf_{x \in X} (\mathcal{R}(x) - w(x) + w(x^{\dagger})) \right). \tag{27}$$

(b) Let  $\mathcal{R}$  be convex at  $x^{\dagger}$  with respect to  $W$ . The subdifferential at  $x^{\dagger}$  with respect  $W$  is defined as

$$\partial_w \mathcal{R}(x^{\dagger}) := \{w \in W : \mathcal{R}(x) \geq \mathcal{R}(x^{\dagger}) + w(x) - w(x^{\dagger}) \text{ for all } x \in X\}.$$

(c) Let  $\mathcal{R}$  be convex at  $x^{\dagger}$  with respect to  $W$ . For  $w \in \partial_w \mathcal{R}(x^{\dagger})$  and  $x \in X$ , the Bregman distance between  $x^{\dagger}$  and  $x$  with respect to  $w$  is defined as

$$D_w(x, x^{\dagger}) := \mathcal{R}(x) - \mathcal{R}(x^{\dagger}) - w(x) + w(x^{\dagger}). \tag{28}$$

**Remark 5.2.** Let  $X$  be a Banach space and set  $W = X^*$ . Then a functional  $\mathcal{R} : X \rightarrow \mathbb{R} \cup \{+\infty\}$  is convex with respect to  $W$ , if and only if it is lower semi-continuous and convex in the classical sense. Moreover, at every  $x^{\dagger} \in X$ , the subdifferential with respect  $W$  coincides with the classical subdifferential  $\partial \mathcal{R}(x^{\dagger}) \subset X^*$ . Finally, the standard Bregman distance, defined by (26), coincides with the Bregman distance obtained by means of Definition 5.1.

In the following, let  $W$  be a given family of real valued functions on  $X$ . Convergence rates in Bregman distance with respect to  $W$  will be derived under the following assumption:

**Assumption 5.3**

1. There exists a monotonically increasing function  $\psi : [0, \infty) \rightarrow [0, \infty)$  such that

$$S(y_1, y_2) \leq \psi(S(y_1, y_3) + S(y_2, y_3)) \quad \text{for all } y_1, y_2, y_3 \in Y. \quad (29)$$

2. For some given point  $x^\dagger \in X$ , the functional  $\mathcal{R} : X \rightarrow \mathbb{R} \cup \{+\infty\}$  is convex at  $x^\dagger$  with respect to  $W$ .

3. There exist  $w \in \partial_w \mathcal{R}(x^\dagger)$  and constants  $\gamma_1 \in [0, 1)$ ,  $\gamma_2 \geq 0$  such that

$$w(x^\dagger) - w(x) \leq \gamma_1 D_w(x, x^\dagger) + \gamma_2 S(\mathbf{F}(x), \mathbf{F}(x^\dagger)) \quad (30)$$

for every  $x \in \Phi_{\mathcal{R}}(\mathbf{F}, \mathbf{F}(x^\dagger), \psi(2\beta), \mathcal{R}(x^\dagger))$ .

In a Banach space setting, the *source inequality* (30) has already been used in [36,50] to derive convergence rates for Tikhonov regularization with convex functionals and in [34] for multiparameter regularization. Eq. (29) is an alternate for the missing triangle inequality in the non-metric case.

**Theorem 5.4 (Convergence Rates).** Let Assumption 5.3 hold and let  $y \in Y$  satisfy  $S(\mathbf{F}(x^\dagger), y) \leq \beta$ . Then, the estimate

$$D_w(x_\beta, x^\dagger) \leq \frac{\gamma_2}{1 - \gamma_1} \psi(\beta + S(\mathbf{F}(x^\dagger), y)) \quad (31)$$

holds for all  $x_\beta \in \Sigma(\mathbf{F}, y, \beta)$ .

**Proof.** Let  $x_\beta \in \Sigma(\mathbf{F}, y, \beta)$ . This, together with (29) and the assumption that  $S(\mathbf{F}(x^\dagger), y) \leq \beta$ , implies that

$$S(\mathbf{F}(x_\beta), \mathbf{F}(x^\dagger)) \leq \psi(S(\mathbf{F}(x_\beta), y) + S(\mathbf{F}(x^\dagger), y)) \leq \psi(2\beta).$$

Together with (30) it follows that

$$D_w(x_\beta, x^\dagger) = \mathcal{R}(x_\beta) - \mathcal{R}(x^\dagger) - w(x_\beta) + w(x^\dagger) \leq \mathcal{R}(x_\beta) - \mathcal{R}(x^\dagger) + \gamma_1 D_w(x_\beta, x^\dagger) + \gamma_2 S(\mathbf{F}(x_\beta), \mathbf{F}(x^\dagger)).$$

The assumption  $\gamma_1 \in [0, 1)$  implies the inequality

$$D_w(x_\beta, x^\dagger) \leq \frac{1}{1 - \gamma_1} (\mathcal{R}(x_\beta) - \mathcal{R}(x^\dagger)) + \frac{\gamma_2}{1 - \gamma_1} S(\mathbf{F}(x_\beta), \mathbf{F}(x^\dagger)). \quad (32)$$

Since  $S(\mathbf{F}(x^\dagger), y) \leq \beta$ , we have  $\mathcal{R}(x_\beta) \leq \mathcal{R}(x^\dagger)$ . Therefore (32) and (29) imply

$$D_w(x_\beta, x^\dagger) \leq \frac{\gamma_2}{1 - \gamma_1} S(\mathbf{F}(x_\beta), \mathbf{F}(x^\dagger)) \leq \frac{\gamma_2}{1 - \gamma_1} \psi(\beta + S(\mathbf{F}(x^\dagger), y)),$$

which concludes the proof.  $\square$

**Remark 5.5.** Typically, convergence rates are formulated in a setting which slightly differs from the one of Theorem 5.4, see [6,21,36,50]. There one assumes the existence of an  $\mathcal{R}$ -minimizing solution  $x^\dagger \in X$  of the equation  $\mathbf{F}(x^\dagger) = y^\dagger$ , for some exact data  $y^\dagger \in \text{ran}(\mathbf{F})$ . Instead of  $y^\dagger$ , only noisy data  $y \in Y$  and the error bound  $S(y^\dagger, y) \leq \beta$  are given.

For this setting, (31) implies the rate

$$D_w(x_\beta, x^\dagger) \leq \frac{\gamma_2}{1 - \gamma_1} \psi(2\beta) = \mathcal{O}(\psi(2\beta)) \quad \text{as } \beta \rightarrow 0,$$

where  $x_\beta \in \Sigma(\mathbf{F}, y, \beta)$  denotes any regularized solution.

**Remark 5.6.** The inequality (30) is equivalent to the existence of  $\eta_1, \eta_2 > 0$  such that

$$w(x^\dagger) - w(x) \leq \eta_1 (\mathcal{R}(x) - \mathcal{R}(x^\dagger)) + \eta_2 S(\mathbf{F}(x), \mathbf{F}(x^\dagger)). \quad (33)$$

Indeed, we obtain (33) from (30) by setting  $\eta_1 := \gamma_1/(1 - \gamma_1)$  and  $\eta_2 := \gamma_2/(1 - \gamma_1)$ . Conversely, (33) implies (30) by taking  $\gamma_1 := \eta_1/(1 + \eta_1)$  and  $\gamma_2 := \eta_2/(1 + \eta_1)$ .

### 5.1. Convergence rates in Banach spaces

In the following, assume that  $X$  and  $Y$  are Banach spaces with norms  $\|\cdot\|$  and  $\|\cdot\|$ , and assume that  $\mathcal{R}$  is a convex and lower semi-continuous functional on  $X$ . We set  $S(y, z) := \|y - z\|$  and let  $D_\zeta$  with  $\zeta \in \partial \mathcal{R}(x^\dagger)$  denote the classical Bregman distance (see Remark 5.2).

If  $x^\dagger$  satisfies the inequality

$$\langle \xi, x^\dagger - x \rangle \leq \gamma_1 D_\xi(x, x^\dagger) + \gamma_2 \|\mathbf{F}(x) - \mathbf{F}(x^\dagger)\| \tag{34}$$

and  $y$  are given data with  $\|\mathbf{F}(x^\dagger) - y\| \leq \beta$ , then **Theorem 5.4** implies the convergence rate  $D_\xi(x_\beta, x^\dagger) = \mathcal{O}(\beta)$ . In the special case where  $X$  is a Hilbert space and  $\mathcal{R}(x) = \|x\|^2$  we have  $D_\xi(x, x^\dagger) = \|x - x^\dagger\|^2$ , which implies the convergence rate  $\|x - x^\dagger\| = \mathcal{O}(\beta^{1/2})$  with respect to the norm. In **Proposition 5.8** we show that the same convergence rate holds on any 2-convex space. For  $r$ -convex Banach spaces with  $r > 2$ , we derive the rate  $\mathcal{O}(\beta^{1/r})$ .

**Definition 5.7.** The Banach space  $X$  is called  $r$ -convex (or is said to have modulus of convexity of power type  $r$ ), if there exists a constant  $C > 0$  such that

$$\inf\{1 - \|(x + y)/2\| : \|x\| = \|y\| = 1, \|x - y\| \geq \epsilon\} \geq C\epsilon^r$$

for all  $\epsilon \in [0, 2]$ .

Note that every Hilbert space is 2-convex and that there is no Banach space (with  $\dim(X) \geq 2$ ) that is  $r$ -convex for some  $r < 2$  (see [42, pp. 63ff]).

**Proposition 5.8** (Convergence rates in the norm). Let  $X$  be an  $r$ -convex Banach space with  $r \geq 2$  and let  $\mathcal{R}(x) := \|x\|^r/r$ . Assume that there exists  $x^\dagger \in X$ , a subgradient  $\xi \in \partial\mathcal{R}(x^\dagger)$ , and constants  $\gamma_1 \in [0, 1)$ ,  $\gamma_2 \geq 0$ ,  $\beta_0 > 0$  such that (34) holds for every  $x \in \Phi_{\mathcal{R}}(\mathbf{F}, \mathbf{F}(x^\dagger), 2\beta_0, \mathcal{R}(x^\dagger))$ .

Then there exists a constant  $c > 0$  such that

$$\|x_\beta - x^\dagger\| \leq c(\beta + \|\mathbf{F}(x^\dagger) - y\|)^{1/r} \tag{35}$$

for all  $\beta \in [0, \beta_0]$ , all  $y \in Y$  with  $\|\mathbf{F}(x^\dagger) - y\| \leq \beta$ , and all  $x_\beta \in \Sigma(\mathbf{F}, y, \beta)$ .

**Proof.** Let  $J_r : X \rightarrow 2^{X^*}$  denote the duality mapping with respect to the weight function  $s \mapsto s^{r-1}$ . In [60, Eq. (2.17)] it is shown that there exists a constant  $K > 0$  such that

$$\|x^\dagger + z\|^r \geq \|x^\dagger\|^r + r \langle j_r(x^\dagger), z \rangle_{X^*, X} + K\|z\|^r \tag{36}$$

for all  $x^\dagger, z \in X$  and  $j_r(x^\dagger) \in J_r(x^\dagger)$ . By Asplund’s theorem [13, Chap. 1, Theorem 4.4], the duality mapping  $J_r$  equals the subgradient of  $\mathcal{R} = \|\cdot\|^r/r$ . Therefore, by taking  $z = x - x^\dagger$  and  $j_r(x^\dagger) = \xi$ , inequality (36) implies

$$D_\xi(x, x^\dagger) \geq \frac{K}{r} \|x - x^\dagger\|^r \quad \text{for all } x^\dagger, x \in X \text{ and } \xi \in \partial\mathcal{R}(x^\dagger). \tag{37}$$

Consequently, (35) follows from **Theorem 5.4**.  $\square$

Exact values for the constant  $K$  in (37) (and thus for the constant  $c$  in (35)) can be derived from [60]. Bregman distances satisfying (37) are called  $r$ -coercive in [35]. This  $r$ -coercivity has already been applied in [2] for the minimization of Tikhonov functionals in Banach spaces.

**Example 5.9.** The spaces  $X = L^p(\Omega, \mu)$  for  $p \in (1, 2]$  and some  $\sigma$ -finite measure space  $(\Omega, \mu)$  are examples of 2-convex Banach spaces (see [42, p. 81, Remarks following Theorem 1.f.1.]). Consequently we obtain for these spaces the convergence rate  $\mathcal{O}(\beta^{1/2})$ . The spaces  $X = L^p(\Omega, \mu)$  for  $p > 2$  are only  $p$ -convex, leading to the rate  $\mathcal{O}(\beta^{1/p})$  in those spaces.

**Remark 5.10.** The book [50, pp. 70ff] clarifies the relation between (34) and the source conditions used to derive convergence rates for convex functionals on Banach spaces. In particular, it is shown that, if  $\mathbf{F}$  and  $\mathcal{R}$  are Gâteaux differentiable at  $x^\dagger$  and there exist  $\gamma > 0$  and  $\omega \in Y^*$  such that  $\gamma\|\omega\| < 1$  and

$$\xi = \mathbf{F}'(x^\dagger)^* \omega \in \partial\mathcal{R}(x^\dagger), \tag{38}$$

$$\|\mathbf{F}(x) - \mathbf{F}(x^\dagger) - \mathbf{F}'(x^\dagger)(x - x^\dagger)\| \leq \gamma D_\xi(x, x^\dagger) \tag{39}$$

for every  $x \in X$ , then (34) holds on  $X$ . (Here  $\mathbf{F}'(x^\dagger)^* : Y^* \rightarrow X^*$  is the adjoint of  $\mathbf{F}'(x^\dagger)$ .) Conversely, if  $\xi \in \partial\mathcal{R}(x^\dagger)$  satisfies (34), then (38) holds for every  $x \in X$ .

In the particular case that  $\mathbf{F} : X \rightarrow Y$  is linear and bounded, the inequality (39) is trivially satisfied with  $\gamma = 0$ . Thus, (34) is equivalent to the sourcewise representability of the subgradient,  $\xi \in \partial\mathcal{R}(x^\dagger) \cap \text{ran}(\mathbf{F}^*)$ .

## 6. Sparse regularization

Let  $\mathcal{A}$  be an at most countable index set, define

$$\ell^2(\mathcal{A}) := \left\{ x = (x_\lambda)_{\lambda \in \mathcal{A}} \subset \mathbb{R} : \sum_{\lambda \in \mathcal{A}} |x_\lambda|^2 < \infty \right\},$$

and assume that  $\mathbf{F} : X := \ell^2(\mathcal{A}) \rightarrow Y$  is a bounded linear operator with dense range in the Hilbert space  $Y$ . We consider for  $p \in (0, 2)$  the minimization problem

$$\mathcal{R}_p(x) := \|x\|_{\ell^p(\mathcal{A})}^p := \sum_{\lambda \in \mathcal{A}} |x_\lambda|^p \rightarrow \min \quad \text{subject to } \|\mathbf{F}x - y\|^2 \leq \beta. \quad (40)$$

For  $p > 1$ , the subdifferential  $\partial \mathcal{R}_p(x^\dagger)$  is at most single valued and is identified with its single element. (The subdifferential may be empty since we consider  $\mathcal{R}_p$  as functions on  $\ell^2(\mathcal{A})$ .)

**Remark 6.1** (Compressed Sensing). In a finite dimensional setting with  $p = 1$ , the minimization problem (40) has received a lot of attention during the last years under the name of *compressed sensing* (see [7,8,10,16–18,20,26,57]). Under some assumptions, the solution of (40) with  $y = \mathbf{F}x^\dagger$  and  $\beta = 0$  has been shown to recover  $x^\dagger$  *exactly* provided the set  $\{\lambda \in \mathcal{A} : x_\lambda^\dagger \neq 0\}$  has sufficiently small cardinality (that is, it is sufficiently sparse). Results for  $p < 1$  can be found in [11,15,25,48].

In this section we prove well-posedness of (40) and derive convergence rates in a possibly infinite dimensional setting. This inverse problems point of view has so far only been treated for the case  $p = 1$  (see [32]). The more general setting has only been considered for Tikhonov regularization

$$\|\mathbf{F}x - y\|^2 + \alpha \mathcal{R}_p(x) \rightarrow \min$$

(see [12,14,28,31,43,61]).

### 6.1. Well-posedness

In the following,  $\tau$  denotes the weak topology on  $\ell^2(\mathcal{A})$ , and  $\tau_p := \tau_{\mathcal{R}_p}$  denotes the topology as in Definition 3.4. Then a sequence  $(x_k)_{k \in \mathbb{N}} \subset \ell^2(\mathcal{A})$  converges to  $x \in \ell^2(\mathcal{A})$  with respect to  $\tau_p$  if and only if  $x_k \rightarrow x$  and  $\mathcal{R}_p(x_k) \rightarrow \mathcal{R}_p(x)$ .

**Proposition 6.2** (Well-Posedness). *Let  $\mathbf{F} : \ell^2(\mathcal{A}) \rightarrow Y$  be a bounded linear operator with dense range. Then constrained  $\ell^p$  regularization with  $0 < p < 2$  is well-posed:*

1. *Existence: For every  $\beta > 0$  and  $y \in Y$ , the set of regularized solutions  $\Sigma(\mathbf{F}, y, \beta)$  is non-empty.*
2. *Stability: Let  $(\beta_k)$  and  $(y_k)$  be sequences with  $\beta_k \rightarrow \beta > 0$  and  $y_k \rightarrow y \in Y$ . Then  $\emptyset \neq \tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma(\mathbf{F}, y_k, \beta_k) \subset \Sigma(\mathbf{F}, y, \beta)$ .*
3. *Convergence: Let  $\|y_k - y\| \leq \beta_k \rightarrow 0$  and assume that the equation  $\mathbf{F}x = y$  has a solution in  $\ell^p(\mathcal{A})$ . Then we have*

$$\emptyset \neq \tau_{\mathcal{R}} - \text{Lim sup}_{k \rightarrow \infty} \Sigma(\mathbf{F}, y_k, \beta_k) \subset \Sigma(\mathbf{F}, y, 0).$$

Moreover, if the equation  $\mathbf{F}x = y$  has a unique  $\mathcal{R}_p$ -minimizing solution  $x^\dagger$ , then we have  $\tau_p - \text{Lim sup}_{k \rightarrow \infty} \Sigma(\mathbf{F}, y_k, \beta_k) = \{x^\dagger\}$ .

**Proof.** In order to prove the existence of minimizers, we apply Theorem 3.1 by showing that  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t)$  is compact with respect to the weak topology on  $\ell^2(\mathcal{A})$  for every  $t > 0$  and is nonempty for some  $t$ . Because  $\mathbf{F}$  has dense range, the set

$$\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t) = \{x \in \ell^2(\mathcal{A}) : \mathcal{R}_p(x) \leq t, \|\mathbf{F}(x) - y\|^2 \leq \beta\}$$

is non-empty for  $t$  large enough.

It remains to show that the sets  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t)$  are weakly compact on  $\ell^2(\mathcal{A})$  for every positive  $t$ . The functional  $\mathcal{R}_p(x) = \sum_{\lambda \in \mathcal{A}} |x_\lambda|^p$  is weakly lower semi-continuous (on  $\ell^2(\mathcal{A})$ ) as the sum of non-negative and weakly continuous functionals (see [19]). Moreover, the mapping  $\mathbf{F}$  is weakly continuous, and therefore  $x \mapsto \|\mathbf{F}x - y\|^2$  is weakly lower semi-continuous, too. The estimate  $\mathcal{R}_p(x) \geq \|x\|_{\ell^2(\mathcal{A})}^p$  (see [31, Eq. (5)]) shows that  $\mathcal{R}_p$  is weakly coercive. Therefore the sets  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t)$  are weakly compact for all  $t > 0$ , see Proposition 3.2.

Taking into account Example 4.5, it follows that  $\mathcal{R}_p$ ,  $\mathcal{S}$ , and  $\mathbf{F}$  satisfy Assumption 4.1. Consequently, Items 2 and 3 follow from Proposition 4.3.  $\square$

**Remark 6.3.** In the case  $p > 1$ , the functional  $\mathcal{R}_p$  is strictly convex, and therefore the  $\mathcal{R}_p$ -minimizing solution  $x^\dagger$  of  $\mathbf{F}x = y$  is unique. Consequently the equality

$$\tau_p - \text{Lim sup}_{k \rightarrow \infty} \Sigma(\mathbf{F}, y_k, \beta_k) = \{x^\dagger\}$$

holds for every  $y$  in the range of the operator  $\mathbf{F}$ .

**Remark 6.4.** For the convex case  $p \geq 1$ , it is shown in [31, Lemma 2] that the  $\tau_p$  convergence of a sequence  $x_k$  already implies  $\mathcal{R}_p(x_k - x) \rightarrow 0$ . In particular, the topology  $\tau_p$  is stronger than the topology induced by  $\|\cdot\|_{\ell^2(\mathcal{A})}$ . A similar result for  $0 < p < 1$  has been derived in [30].

6.2. Convergence rates

In the following, we derive two types of convergence rates results with respect the  $\ell^2$ -norm: the convergence rate  $\mathcal{O}(\beta^{1/2})$  (for  $p \in (1, 2)$ ), and the convergence rate  $\mathcal{O}(\beta^{\min(1, 1/p)})$  (for every  $p \in (0, 2)$ ) for sparse sequences—here and in the following,  $x^\dagger \in \ell^2(A)$  is called sparse, if

$$\text{supp}(x^\dagger) := \{\lambda \in A : x^\dagger_\lambda \neq 0\}$$

is finite. The convergence rates results for constrained  $\ell^p$  regularization, derived in this section, are summarized in Table 1.

For  $p \geq 1$ , the same type of results (Propositions 6.5, 6.7) has also been obtained for  $\ell^p$ -Tikhonov regularization in [31,50]. The results for the non-convex case,  $p \in (0, 1)$ , are based on [30], where the same rate for non-convex Tikhonov regularization with a priori parameter choice has been derived (see also [29]). Similar, but weaker, results have been already been derived in [5,28,61] in the context of Tikhonov regularization. In [61], the conditions for the convergence rates result for non-convex regularization are basically the same as in Proposition 6.7, but only a rate of order  $\mathcal{O}(\beta^{1/2})$  has been obtained. In [5,28], a linear convergence rate  $\mathcal{O}(\beta)$  is proven, but with a considerably stronger range condition: each standard basis vector  $e_\lambda$ ,  $\lambda \in A$ , has to satisfy  $e_\lambda \in \text{ran} \mathbf{F}^*$ .

**Proposition 6.5.** *Let  $1 < p < 2$ ,  $x^\dagger = (x^\dagger_\lambda)_{\lambda \in A} \in \ell^2(A)$ , and let  $\mathbf{F} : \ell^2(A) \rightarrow Y$  be a bounded linear operator. Moreover, assume that there exists  $\omega \in Y$  with  $\partial \mathcal{R}_p(x^\dagger) = \mathbf{F}^* \omega$ . Then the set  $\Sigma(\mathbf{F}, y, \beta) = \{x_\beta\}$  consists of a single element and there exists a constant  $d_p > 0$  only depending on  $p$ , such that*

$$\|x_\beta - x^\dagger\|_{\ell^2(A)}^2 \leq \frac{d_p \|\omega\|}{3 + 2\mathcal{R}_p(x^\dagger)} (\beta + \|\mathbf{F}x^\dagger - y\|) \tag{41}$$

for all  $\beta > 0$  and  $y \in Y$  with  $\|\mathbf{F}(x^\dagger) - y\| \leq \beta$ .

**Proof.** The assumption  $\partial \mathcal{R}_p(x^\dagger) = \mathbf{F}^* \omega$  then implies that (30) is satisfied with  $W = X^*$ ,  $\gamma_1 = 0$  and  $\gamma_2 = \|\omega\|$ . Theorem 5.4 therefore implies the inequality

$$\text{supp}\{D_{\partial \mathcal{R}_p(x^\dagger)}(x_\beta, x^\dagger) : x_\beta \in \Sigma(\mathbf{F}, y, \beta)\} \leq \|\omega\|(\beta + \|\mathbf{F}x^\dagger - y\|). \tag{42}$$

From [31, Lemma 10] we obtain the inequality

$$\|x - x^\dagger\|_{\ell^2(A)}^2 \leq \frac{d_p}{3 + 2\mathcal{R}_p(x^\dagger) + \mathcal{R}_p(x)} D_{\partial \mathcal{R}_p(x^\dagger)}(x, x^\dagger) \tag{43}$$

for all  $x \in \text{dom}(\mathcal{R}_p)$ . Now, (41) follows from (42) and (43).  $\square$

**Remark 6.6.** Since  $\ell^p(A)$  is 2-convex (see [42]) and continuously embedded in  $\ell^2(A)$ , Proposition 5.8 provides an alternative estimate for  $x_\beta - x^\dagger$  in terms of the stronger distance  $\|\cdot\|_{\ell^p(A)}$ . The prefactor in (35), however, is constant, whereas the prefactor in (41) tends to 0 as  $\mathcal{R}_p(x^\dagger)$  increases. Thus the two estimates are somehow independent from each other.

**Proposition 6.7 (Sparse Regularization).** *Let  $p \in (0, 2)$ , let  $x^\dagger = (x^\dagger_\lambda)_{\lambda \in A} \in \ell^2(A)$  be sparse, and let  $\mathbf{F} : \ell^2(A) \rightarrow Y$  be bounded linear. Assume that one of the following conditions holds:*

- We have  $p \in (1, 2)$ , there exists  $\omega \in Y$  with  $\partial \mathcal{R}_p(x^\dagger) = \mathbf{F}^* \omega$ , and  $\mathbf{F}$  is injective on  $V = \{x \in \ell^2(A) : \text{supp}(x) \subset \text{supp}(x^\dagger)\}$ .
- We have  $p = 1$ , there exist  $\xi = (\xi_\lambda)_{\lambda \in A} \in \partial \mathcal{R}_1(x^\dagger)$  and  $\omega \in Y$  with  $\xi = \mathbf{F}^* \omega$ , and  $\mathbf{F}$  is injective on  $V = \{x \in \ell^2(A) : \text{supp}(x) \subset \{\lambda \in A : |\xi_\lambda| = 1\}\}$ .
- We have  $p \in (0, 1)$ ,  $x^\dagger$  is the unique  $\mathcal{R}_p$ -minimizing solution of  $\mathbf{F}x = \mathbf{F}x^\dagger$ , and  $\mathbf{F}$  is injective on  $V = \{x \in \ell^2(A) : \text{supp}(x) \subset \text{supp}(x^\dagger)\}$ .

**Table 1**  
Convergence rates for constrained  $\ell^p$  regularization.

Rate	Norm	Premises (besides $\text{ran}(\mathbf{F}^*) \cap \partial \mathcal{R}_p \neq \emptyset$ )	Results
$\beta^{1/2}$	$\ \cdot\ _{\ell^2}$	$p \in (1, 2)$	Proposition 6.5
$\beta^{1/2}$	$\ \cdot\ _{\ell^p}$	$p \in (1, 2)$	Remark 6.6
$\beta^{1/p}$	$\ \cdot\ _{\ell^2}$	$p \in (1, 2)$ , sparsity, injectivity on $V$	Proposition 6.7
$\beta$	$\ \cdot\ _{\ell^2}$	$p \in (0, 1)$ , uniqueness of $x^\dagger$ , sparsity, injectivity on $V$	Proposition 6.7



Then

$$\sup\{\|x_\beta - x^\dagger\|_{\ell^2(\mathcal{A})} : x_\beta \in \Sigma(\mathbf{F}, y, \beta), \|\mathbf{F}x^\dagger - y\| \leq \beta\} = \mathcal{O}(\beta^{\min\{1, 1/p\}}) \quad \text{as } \beta \rightarrow 0.$$

**Proof.** Assume first that  $p \in (1, 2)$ . Define  $W := \{w(x) := -c\|x - \bar{x}\|^p : \bar{x} \in X, c > 0\}$ . Then the functional  $\mathcal{R}_p$  is convex at  $x^\dagger$  with respect to  $W$ . Moreover it has been shown in [31, Proof of Theorem 14] that there exists  $w(x) = -c\|x - x^\dagger\|^p \in \partial_W w(x^\dagger) \subset W$  such that for some  $\eta_1, \eta_2 > 0$  the inequality

$$-w(x) = c\|x - x^\dagger\|^p \leq \eta_1(\mathcal{R}_p(x) - \mathcal{R}_p(x^\dagger)) + \eta_2\|\mathbf{F}(x - x^\dagger)\| \quad (44)$$

holds on  $\Sigma(2\beta, y^\dagger, \mathbf{F})$  for  $\beta$  small enough. Using Remark 5.6, Theorem 5.4 therefore implies the rate

$$\sup\{D_w(x_\beta, x^\dagger) : x_\beta \in \Sigma(\mathbf{F}, y, \beta), \|\mathbf{F}x^\dagger - y\| \leq \beta\} = \mathcal{O}(\beta) \quad \text{as } \beta \rightarrow 0.$$

The assertion then follows from the fact that the norm on  $\ell^2(\mathcal{A})$  can be bounded by the Bregman distance  $D_w$ .

The proofs for  $p = 1$  and  $p \in (0, 1)$  are similar; the required estimate (44) has been shown for  $p = 1$  in [31, Proof of Theorem 15] and for  $p \in (0, 1)$  in [30, Eq. (7)].  $\square$

## 7. Conclusion

Due to modeling, computing, and measurement errors, the solution of an ill-posed equation  $\mathbf{F}(x) = y$ , even if it exists, typically yields unacceptable results. The residual method replaces the exact solution by the set  $\Sigma(\mathbf{F}, y, \beta) = \arg \min\{\mathcal{R}(x) : S(\mathbf{F}(x), y) \leq \beta\}$ , where  $\mathcal{R}$  is a stabilizing functional and  $S$  denotes a distance measure between  $\mathbf{F}(x)$  and  $y$ . This paper shows that in a very general setting  $\Sigma(\mathbf{F}, y, \beta)$  is stable with respect to perturbations of the data  $y$  and the operator  $\mathbf{F}$  (Lemma 3.6 and Theorem 3.9), and the regularized solutions converge to  $\mathcal{R}$ -minimizing solutions of  $\mathbf{F}(x) = y$  as  $\beta \rightarrow 0$  (Theorem 3.10). In particular the stability issue has hardly been considered so far in the literature.

In the case where  $\mathbf{F}$  acts between linear spaces  $X$  and  $Y$ , stability and convergence have been shown under a list of reasonable properties (see Assumption 4.1). These assumptions are satisfied for bounded linear operators, but also for a certain class of nonlinear operators (Example 4.6). If  $Y$  is reflexive,  $X$  satisfies the Radon–Riesz property,  $\mathbf{F}$  is a closed linear operator, and  $\mathcal{R}$  and  $S$  are given by powers of the norms on  $X$  and  $Y$ , the set  $\Sigma(\mathbf{F}, y, \beta)$  consists of a single element  $x_\beta$ . This element is shown to converge strongly to the minimal norm solution  $x^\dagger$  as  $\beta \rightarrow 0$ . In this special situation, norm convergence has also been shown in [39, Theorem 3.4.1].

In Section 5 we have derived quantitative estimates (convergence rates) for the difference between  $x^\dagger$  and minimizers  $x_\beta \in \Sigma(\mathbf{F}, y, \beta)$  in terms of a (generalized) Bregman distance. All these estimates hold provided  $S(\mathbf{F}(x^\dagger), y) \leq \beta$  and a source inequality introduced in [36] is satisfied. For linear operators, the required source inequality follows from a source wise representation of a subgradient of  $\mathcal{R}$  at  $x^\dagger$ . This carries on the result of [6] for constrained regularization. In the special case that  $X$  is an  $r$ -convex Banach space with  $r \geq 2$  and  $\mathcal{R}$  is the  $r$ th power of the norm on  $X$ , we have obtained convergence rates  $\mathcal{O}(\beta^{1/r})$  with respect to the norm. The spaces  $X = L^p(\Omega)$  for  $p \in (1, 2]$  are examples of 2-convex Banach spaces, leading to the rate  $\mathcal{O}(\sqrt{\beta})$  in those spaces.

As an application for our rather general results we have investigated sparse  $\ell^p$  regularization with  $p \in (0, 2)$ . We have shown well-posedness in both the convex ( $p \geq 1$ ) and the non-convex case ( $p < 1$ ). In addition, we have studied the reconstruction of sparse sequence. There we have derived the improved convergence rates  $\mathcal{O}(\beta^{1/p})$  for the convex and  $\mathcal{O}(\beta)$  for the non-convex case.

## Acknowledgement

This work has been supported by the Austrian Science Fund (FWF), projects 9203-N12 and project S10505-N20.

## Appendix A. Auxiliary results

**Lemma A.1.** Assume that  $(y_k)_{k \in \mathbb{N}}$  converges  $S$ -uniformly to  $y \in Y$  and the mappings  $\mathbf{F}_k : X \rightarrow Y$  converge locally  $S$ -uniformly to  $\mathbf{F} : X \rightarrow Y$ .

Then, for every  $\beta > 0$ ,  $t > 0$  and  $\varepsilon > 0$ , there exists some  $k_0 \in \mathbb{N}$  such that

$$\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta - \varepsilon, t') \subset \Phi_{\mathcal{R}}(\mathbf{F}_k, y_k, \beta, t') \subset \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta + \varepsilon, t') \quad (\text{A.1})$$

for every  $t' \leq t$  and  $k \geq k_0$ .

**Proof.** Since  $y_k \rightarrow y$   $S$ -uniformly and  $\mathbf{F}_k \rightarrow \mathbf{F}$  locally  $S$ -uniformly, there exists  $k_0 \in \mathbb{N}$  such that

$$\begin{aligned} |\mathcal{S}(\mathbf{F}_k(x), y_k) - \mathcal{S}(\mathbf{F}_k(x), y)| &\leq \varepsilon/2, \\ |\mathcal{S}(\mathbf{F}_k(x), y) - \mathcal{S}(\mathbf{F}(x), y)| &\leq \varepsilon/2, \end{aligned} \quad (\text{A.2})$$

for all  $x \in X$  with  $\mathcal{R}(x) \leq t$  and  $k \geq k_0$ .

Now let  $t' \leq t$  and let  $x \in \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta - \varepsilon, t')$ . Then (A.2) implies that

$$|\mathcal{S}(\mathbf{F}_k(x), y_k) - \mathcal{S}(\mathbf{F}(x), y)| \leq |\mathcal{S}(\mathbf{F}_k(x), y_k) - \mathcal{S}(\mathbf{F}_k(x), y)| + |\mathcal{S}(\mathbf{F}_k(x), y) - \mathcal{S}(\mathbf{F}(x), y)| \leq \varepsilon,$$

and thus

$$\mathcal{S}(\mathbf{F}_k(x), y_k) \leq \mathcal{S}(\mathbf{F}(x), y) + \varepsilon \leq \beta,$$

that is,  $x \in \Phi_{\mathcal{R}}(\mathbf{F}_k, y_k, \beta, t')$ , which proves the first inclusion in (A.1). The second inclusion is shown in a similar manner.  $\square$

The following lemma states that the value of the minimization problem (4) behaves well as the parameter  $\beta$  decreases.

**Lemma A.2.** Assume that  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \gamma, t)$  is  $\tau$ -compact for every  $\gamma$  and every  $t$ . Then the value  $v$  of the constraint optimization problem (4) is right continuous in the first variable, that is,

$$v(\mathbf{F}, y, \beta) = \lim_{\varepsilon \rightarrow 0^+} v(\mathbf{F}, y, \beta + \varepsilon) = \sup_{\varepsilon > 0} v(\mathbf{F}, y, \beta + \varepsilon). \quad (\text{A.3})$$

**Proof.** Since  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, t) \subset \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta + \varepsilon, t)$ , it follows that  $v(\mathbf{F}, y, \beta) \geq v(\mathbf{F}, y, \beta + \varepsilon)$  for every  $\varepsilon > 0$ , and therefore  $v(\mathbf{F}, y, \beta) \geq \sup_{\varepsilon > 0} v(\mathbf{F}, y, \beta + \varepsilon)$ .

In order to show the converse inequality, let  $\delta > 0$ . Then the definition of  $v(\mathbf{F}, y, \beta)$  implies that  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, v(\mathbf{F}, y, \beta) - \delta) = \emptyset$ . Since (cf. Lemma 2.2)

$$\emptyset = \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta, v(\mathbf{F}, y, \beta) - \delta) = \bigcap_{\varepsilon > 0} \Phi_{\mathcal{R}}(\mathbf{F}, y, \beta + \varepsilon, v(\mathbf{F}, y, \beta) - \delta) \quad (\text{A.4})$$

and the right hand side of (A.4) is a decreasing family of compact sets. It follows that already  $\Phi_{\mathcal{R}}(\mathbf{F}, y, \beta + \varepsilon_0, v(\mathbf{F}, y, \beta) - \delta) = \emptyset$  for some  $\varepsilon_0 > 0$ , and thus

$$\sup_{\varepsilon > 0} v(\mathbf{F}, y, \beta + \varepsilon) \geq v(\mathbf{F}, y, \beta + \varepsilon_0) \geq v(\mathbf{F}, y, \beta) - \delta.$$

Since  $\delta$  was arbitrary, this shows the assertion.  $\square$

**Lemma A.3.** Let  $(\Sigma_k)_{k \in \mathbb{N}}$  be a sequence of subsets of  $X$ . Then  $U = \tau - \text{Lim}_{k \rightarrow \infty} \Sigma_k$  if and only if every subsequence  $(\Sigma_{k_j})_{j \in \mathbb{N}}$  satisfies

$$U = \tau - \text{Lim}_{j \rightarrow \infty} \sup \Sigma_{k_j}.$$

**Proof.** See [41, Section 29.V].  $\square$

## References

- [1] R. Acar, C.R. Vogel, Analysis of bounded variation penalty methods for ill-posed problems, *Inverse Prob.* 10 (6) (1994) 1217–1229.
- [2] T. Bonesky, K.S. Kazimierski, P. Maass, F. Schöpfer, T. Schuster, Minimization of Tikhonov functionals in Banach spaces, *Abstr. Appl. Anal.* 19 (2008). Art. ID 192679.
- [3] R.I. Boţ, B. Hofmann, An extension of the variational approach for obtaining convergence rates in regularization of nonlinear ill-posed problems, *J. Integral Equ. Appl.* 22 (2010) 369–392.
- [4] A. Braides,  $\Gamma$ -convergence for beginners, *Oxford Lecture Series in Mathematics and its Applications*, vol. 22, Oxford University Press, Oxford, 2002.
- [5] K. Bredies, D. Lorenz, Regularization with non-convex separable constraints, *Inverse Prob.* 25 (8) (2009) 085011 (14p).
- [6] M. Burger, S. Osher, Convergence rates of convex variational regularization, *Inverse Prob.* 20 (5) (2004) 1411–1421.
- [7] E.J. Candès, The restricted isometry property and its implications for compressed sensing, *C.R. Math. Acad. Sci. Paris* 346 (9–10) (2008) 589–592.
- [8] E.J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inf. Theory* 52 (2) (2006) 489–509.
- [9] E.J. Candès, J.K. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, *Commun. Pure Appl. Math.* 59 (8) (2006) 1207–1223.
- [10] E.J. Candès, T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans. Inf. Theory* 52 (12) (2006).
- [11] R. Chartrand, Exact reconstructions of sparse signals via nonconvex minimization, *IEEE Signal Process. Lett.* 14 (2007) 707–710.
- [12] C. Chau, P.L. Combettes, J.-C. Pesquet, V.R. Wajs, A variational formulation for frame-based inverse problems, *Inverse Prob.* 23 (4) (2007) 1495–1518.
- [13] I. Cioranescu, *Geometry of Banach Spaces, Duality mappings and nonlinear problems*, *Mathematics and its Applications*, vol. 62, Kluwer, Dordrecht, 1990.
- [14] I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Commun. Pure Appl. Math.* 57 (11) (2004) 1413–1457.
- [15] M.E. Davies, R. Gribonval, Restricted isometry constants where  $\ell_p$  sparse recovery can fail for  $0 < p \leq 1$ , *IEEE Trans. Inf. Theory* 55 (5) (2009) 2203–2214.
- [16] D.L. Donoho, For most large underdetermined systems of equations, the minimal  $\ell_1$ -norm near-solution approximates the sparsest near-solution, *Commun. Pure Appl. Math.* 59 (7) (2006) 907–934.

- [17] D.L. Donoho, For most large underdetermined systems of linear equations the minimal  $\ell_1$ -norm solution is also the sparsest solution, *Commun. Pure Appl. Math.* 59 (6) (2006) 797–829.
- [18] D.L. Donoho, M. Elad, V.N. Temlyakov, Stable recovery of sparse overcomplete representations in the presence of noise, *IEEE Trans. Inf. Theory* 52 (1) (2006) 6–18.
- [19] I. Ekeland, R. Temam, *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, 1976.
- [20] M. Elad, *Sparse and Redundant Representations*, Springer, New York, 2010. From theory to applications in signal and image processing. With a foreword by Alfred M. Bruckstein.
- [21] H.W. Engl, M. Hanke, A. Neubauer, *Regularization of inverse problems*, *Mathematics and its Applications*, vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.
- [22] H.W. Engl, K. Kunisch, A. Neubauer, Convergence rates for Tikhonov regularisation of nonlinear ill-posed problems, *Inverse Prob.* 5 (3) (1989) 523–540.
- [23] J. Flemming, B. Hofmann, A new approach to source conditions in regularization with general residual term, *Numer. Funct. Anal. Optim.* 31 (3) (2010) 245–284.
- [24] I. Fonseca, G. Leoni, *Modern methods in the calculus of variations:  $L^p$  spaces*, *Springer Monographs in Mathematics*, Springer, New York, 2007.
- [25] S. Foucart, M. Lai, Sparsest solutions of underdetermined linear systems via  $\ell^q$ -minimization for  $0 < q \leq 1$ , *Appl. Comput. Harmon. Anal.* 26 (3) (2009) 395–407.
- [26] J.J. Fuchs, Recovery of exact sparse representations in the presence of bounded noise, *IEEE Trans. Inf. Theory* 51 (10) (2005) 3601–3608.
- [27] C.R. Givens, R.M. Shortt, A class of Wasserstein metrics for probability distributions, *Michigan Math. J.* 31 (2) (1984) 231–240.
- [28] M. Grasmair, Well-posedness and convergence rates for sparse regularization with sublinear  $l^p$  penalty term, *Inverse Prob. Imaging* 3 (3) (2009) 383–387.
- [29] M. Grasmair, Generalized Bregman distances and convergence rates for non-convex regularization methods, *Inverse Prob.* 26 (11) (2010) 115014.
- [30] M. Grasmair, Non-convex sparse regularisation, *J. Math. Anal. Appl.* 365 (1) (2010) 19–28.
- [31] M. Grasmair, M. Haltmeier, O. Scherzer, Sparse regularization with  $l^p$  penalty term, *Inverse Prob.* 24 (5) (2008) 055020.
- [32] M. Grasmair, M. Haltmeier, O. Scherzer, Necessary and sufficient conditions for linear convergence of  $\ell^1$ -regularization, *Commun. Pure Appl. Math.* 64 (2) (2011) 161–182.
- [33] C.W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, Boston, 1984.
- [34] T. Hein, Convergence rates for multi-parameter regularization in Banach spaces, *Int. J. Pure Appl. Math.* 43 (4) (2008) 593–614.
- [35] T. Hein, B. Hofmann, Approximate source conditions for nonlinear ill-posed problems – chances and limitations, *Inverse Prob.* 25 (2009) 035003.
- [36] B. Hofmann, B. Kaltenbacher, C. Pöschl, O. Scherzer, A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators, *Inverse Prob.* 23 (3) (2007) 987–1010.
- [37] B. Hofmann, M. Yamamoto, Convergence rates for Tikhonov regularization based on range inclusions, *Inverse Prob.* 21 (3) (2005) 805–820.
- [38] V.K. Ivanov, Ill-posed problems in topological spaces, *Sibirsk. Mat. Ž.* 10 (1969) 1065–1074. In Russian.
- [39] V.K. Ivanov, V.V. Vasin, V.P. Tanana, *Theory of linear ill-posed problems and its applications*, second ed., *Inverse and Ill-posed Problems Series*, VSP, Utrecht, 2002. Translated and revised from the 1978 Russian original.
- [40] J.L. Kelley, *General Topology*, D. Van Nostrand Company, Toronto-New York-London, 1955.
- [41] K. Kuratowski, *Topology*, in: J. Jaworowski (Ed.), *Revised and Augmented*, Translated from the French, New ed., vol. I, Academic Press, New York, 1966.
- [42] J. Lindenstrauss, L. Tzafriri, *Classical Banach spaces. II*, *Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas]*, vol. 97, Springer Verlag, Berlin, 1979. Function spaces.
- [43] D. Lorenz, Convergence rates and source conditions for Tikhonov regularization with sparsity constraints, *J. Inverse Ill-Posed Prob.* 16 (5) (2008) 463–478.
- [44] R.E. Megginson, *An Introduction to Banach Space Theory*, *Graduate Texts in Mathematics*, vol. 183, Springer Verlag, New York, 1998.
- [45] V.A. Morozov, *Regularization Methods for Ill-Posed Problems*, CRC Press, Boca Raton, 1993.
- [46] E. Resmerita, Regularization of ill-posed problems in Banach spaces: convergence rates, *Inverse Prob.* 21 (4) (2005) 1303–1314.
- [47] E. Resmerita, O. Scherzer, Error estimates for non-quadratic regularization and the relation to enhancement, *Inverse Prob.* 22 (3) (2006) 801–814.
- [48] R. Saab, R. Chartrand, O. Yilmaz, Stable sparse approximations via nonconvex optimization, in: *33rd International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2008.
- [49] O. Scherzer, H.W. Engl, K. Kunisch, Optimal a posteriori parameter choice for Tikhonov regularization for solving nonlinear ill-posed problems, *SIAM J. Numer. Anal.* 30 (6) (1993) 1796–1838.
- [50] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, F. Lenzen, *Variational methods in imaging*, *Applied Mathematical Sciences*, vol. 167, Springer, New York, 2009.
- [51] T.I. Seidman, Convergent approximation methods for ill-posed problems. I. General theory, *Polish Academy of Sciences. Institute of Applied Cybernetics. Control Cybernet.* 10 (1–2) (1981) 31–49.
- [52] T.I. Seidman, C.R. Vogel, Well posedness and convergence of some regularisation methods for non-linear ill posed problems, *Inverse Prob.* 5 (2) (1989) 227–238.
- [53] I. Singer, *Abstract convex analysis*, *Canadian Mathematical Society Series of Monographs and Advanced Texts*, John Wiley & Sons, Inc., New York, 1997. With a foreword by A.M. Rubinov, A Wiley-Interscience Publication.
- [54] V.P. Tanana, On a criterion for the convergence of the residual method, *Dokl. Akad. Nauk* 343 (1) (1995) 22–24.
- [55] V.P. Tanana, *Methods for solution of nonlinear operator equations*, *Inverse and Ill-posed Problems Series*, VSP, Utrecht, 1997.
- [56] A.N. Tikhonov, V.Y. Arsenin, *Solutions of Ill-Posed Problems*, John Wiley & Sons, Washington, DC, 1977.
- [57] J.A. Tropp, Just relax: convex programming methods for identifying sparse signals in noise, *IEEE Trans. Inf. Theory* 52 (3) (2006) 1030–1051.
- [58] V.V. Vasin, Some tendencies in the Tikhonov regularization of ill-posed problems, *J. Inverse Ill-Posed Prob.* 14 (8) (2006) 813–840.
- [59] A. Visintin, Strong convergence results related to strict convexity, *Comm. Partial Diff. Equ.* 9 (5) (1984) 439–466.
- [60] Z.B. Xu, G.F. Roach, Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces, *J. Math. Anal. Appl.* 157 (1) (1991) 189–210.
- [61] C.A. Zarzer, On Tikhonov regularization with non-convex sparsity constraints, *Inverse Prob.* 25 (2009) 025006.