Discretization of variational regularization in Banach spaces

Christiane Pöschl, Elena Resmerita and Otmar Scherzer

1 Department of Information and Communication Technologies, Universitat Pompeu Fabra
C/Tanger 122-140, 08018 Barcelona, Spain
2 Department of Industrial Mathematics, Johannes Kepler University, Altenbergerstraße 69, A-4040 Linz, Austria
3 Computational Science Center, University of Vienna, Nordbergstraße 15, A-1090 Vienna, Austria
4 Radon Institute of Computational and Applied Mathematics, Altenbergerstraße 69, A-4040 Linz, Austria

E-mail: Christiane.Poeschl@univie.ac.at, elena.resmerita@jku.at and otmar.scherzer@univie.ac.at

Received 14 April 2010, in final form 13 August 2010
Published 9 September 2010
Online at stacks.iop.org/IP/26/105017

Abstract
Consider a nonlinear ill-posed operator equation \( F(u) = y \), where \( F \) is defined on a Banach space \( X \). In this paper we analyze finite-dimensional variational regularization, which takes into account operator approximations and noisy data. As shown in the literature, depending on the setting, convergence of the regularized solutions of the finite-dimensional problems can be with respect to the strong or just a weak topology. In this paper our contribution is twofold. First, we derive convergence rates in terms of Bregman distances in the convex regularization setting under appropriate sourcewise representation of a solution of the equation. Secondly, for particular regularization realizations in nonseparable Banach spaces, we discuss the finite-dimensional approximations of the spaces and the type of convergence, which is needed for the convergence analysis. These considerations lay the fundament for efficient numerical implementation. In particular, we emphasize on the space \( X \) of finite total variation functions and analyze in detail the cases when \( X \) is the space of the functions of finite bounded deformation and the \( L^\infty \)-space. The latter two settings are of interest in numerous problems arising in optimal control, machine learning and engineering.
1. Introduction

Let $F : X \rightarrow Y$ be a nonlinear operator with the domain $\mathcal{D}(F)$, where $X$ is a Banach space and $Y$ is a Hilbert space. We would like to approximate the solutions of the ill-posed equation

$$F(u) = y$$

via variational regularization.

Let $R : X \rightarrow [0, +\infty]$ be a penalty functional with the nonempty domain $\mathcal{D}(R)$. An element $\bar{u} \in \mathcal{D}(F) \cap \mathcal{D}(R)$ is called an $R$-minimizing solution of (1) if it solves the constraint optimization problem

$$\min R(u) \text{ subject to } F(u) = y.$$  

We assume that the noisy data $y^\delta$ are given such that

$$\|y^\delta - y\| \leq \delta.$$  

In order to solve equation (1) numerically, the space $X$ has to be approximated by a sequence of finite-dimensional subspaces $X_n$. The situation when the spaces $X$ and $Y$ are Hilbert and the regularization is quadratic has been analyzed in [30, 32] (linear problems) and [31, 33] (nonlinear problems). However, recent advances in variational regularization theory deal with general Banach spaces. Although convergence of regularization methods in the general setting has been established (see, e.g., [40, section 2], [22, 35]), the literature on the analysis of discretization, and therefore numerical implementation, is not quite developed. The focus of this work is on problems modeled in nonseparable Banach spaces. In comparison with the Hilbert space theory, a significant complication is due to the fact that such a space cannot be approximated by a nested sequence of finite-dimensional subspaces, with respect to the norm topology. We have in mind the space of bounded variation functions, $BV$, the space of bounded deformation functions, BD, and the space of essentially bounded functions $L^\infty$ which are all nonseparable. However, as mentioned in [3, page 121] in the context of $BV$, ‘the norm topology is too strong for many applications’, and in particular when considering finite-dimensional approximations. The reader is referred to [40, section 4] for the convergence results of finite-dimensional regularization in general topological spaces and to, e.g., [8, 9, 17] for the $BV$ case. More precisely, convergence was shown with respect to a weak topology of the space and with respect to the penalty function (see section 2 for details). As explained in [6], Bregman distances are appropriate ‘measures’ for quantifying the convergence rate. Thus, deriving convergence rates (in the sense of Bregman distances) for variational regularization by discretization under appropriate sourcewise representation of a solution of the equation is our first concern. As stated in [40, section 4], [17] and [8], it is natural to consider the convergence of finite-dimensional discretizations with respect to a topology which is weaker than the norm topology. Instead of the norm topology on $X$ (as in the separable Hilbert space or in the separable Banach space setting), one uses a metric on $X$, which is induced from the metric of a superspace of $X$. As we show, in some examples, the convergence in the sense of such a metric is equivalent to the approximation properties of the finite-dimensional subspaces with respect to the weak topology. Thus, our second contribution is the study of the convergence type for finite-dimensional regularization. We discuss the case when $X$ is the space of finite total variation functions (see [8, 9, 17]) from the perspective described above and analyze in detail the case when $X$ is the space of the functions of finite bounded deformation and the $L^\infty$-space. Needless to say that the $BV$-space and the BD-space are successfully employed in the theory of image processing and of elasto-plastic body deformations. Modeling regularization in the $L^\infty$-space has not been widely studied in the inverse problem community. However, it is
relevant in optimal control, signal processing and machine learning—see, e.g., [14, 16, 20, 21, 
24, 41].

The paper is organized as follows. Section 2 specifies the assumptions, discusses well-
posedness and convergence of the discretized regularization method in the case when the 
Banach space \( X \) is nonseparable, not even requiring that the regularization term is a convex 
function. Also, convergence rates with respect to Bregman distances are obtained in the 
convex regularization setting, under a standard source condition. The finite-dimensional 
approximation of the solutions of the equation \( F(x) = y \) in separable Banach spaces \( X \) is briefly discussed in section 3. Section 4 studies in detail the discretization of several 
nonseparable Banach spaces, such as the space of bounded variation functions, the space of 
bounded deformation functions and the space of essentially bounded functions. The inverse 
ground water filtration problem is analyzed in the natural setting of \( L^\infty(\Omega) \) in section 5.

2. The case of nonseparable Banach spaces

2.1. Main assumptions

Let \( X \) be a not necessarily separable Banach space which can be embedded into a separable 
Banach space \( Z \). Let \( \tau \) be a topology on \( X \) which is weaker than the norm topology on \( X \). 
Therefore, we refer to \( \tau \) as the weak topology. In addition let \( \mathcal{R} : X \to [0, +\infty] \) be a proper 
functional.

We define a metric on the space \( X \) by the norm of \( Z \) induced on \( X \) and a pseudometric 
generated by the function \( \mathcal{R} \):

\[
d(u, v) = \|u - v\|_Z + |\mathcal{R}(u) - \mathcal{R}(v)|. \tag{4}
\]

We shall denote by \( \tau_d \) the topology generated by this metric. Relating to the above discussion, 
this is an intermediary topology between the norm topology on \( X \) and the weak topology \( \tau \).

The following elementary result gives us a motivation for approximating a nonseparable 
Banach space in a topology that is weaker than the norm topology.

**Proposition 2.1.** A Banach space \( X \) is separable if and only if there exists a nested (increasing) 
sequence of finite-dimensional subspaces \( \{X_n\} \) such that

\[
\bigcup_{n\in\mathbb{N}} X_n = X,
\]

where the closure is considered with respect to the norm topology of \( X \).

Consider a sequence of nested, finite-dimensional subspaces \( \{X_n\} \) which satisfies

\[
\bigcup_{n\in\mathbb{N}} X_n = X. \tag{5}
\]

That is, \( X \) is the closure of the reunion of the subspaces \( X_n \) with respect to the topology of the 
metric \( d \). This property holds for many nonseparable Banach spaces—see several examples 
in section 4.

We are given the approximation operators \( F_m \) of \( F \), which have the same domain \( \mathcal{D}(F) \) as 
\( F \). We assume that the operators \( F, F_m \) satisfy the following approximation property:

\[
\|F(u) - F_m(u)\| \leq \rho_m \quad \text{for all} \quad u \in \mathcal{D}(F) \cap \mathcal{D}(\mathcal{R}). \tag{6}
\]

Here, the constant \( \rho_m \) should only depend on \( m \) and satisfy

\[
\lim_{m\to\infty} \rho_m = 0.
\]

Denote

\[
D_n := X_n \cap \mathcal{D}(F) \cap \mathcal{D}(\mathcal{R}), \quad n \in \mathbb{N},
\]
and assume that the sets $D_n$ are nonempty. We are interested in approximating the $R$-minimizing solutions of equation (1) by solutions $u_{m,n} \in D_n$ of the problem

$$
\min \{ \| F_m(u) - y^\delta \|^2 + \alpha R(u) \} \quad \text{subject to} \quad u \in D_n.
$$

In order to pursue the analysis, we make several (standard) assumptions on the spaces $X, Y$, the operator $F$, the functional $R$ (see also [22, 35, 36]), and on the approximations $X_n, F_m$ as well.

**Assumption 2.2.**

(i) The Banach space $X$ is associated with a topology $\tau$ such that
   - the topology $\tau_d$ is finer than the topology $\tau$;
   - the norm topology is finer than the topology $\tau_d$.

(ii) The domain $D(F)$ is $\tau$-closed.

(iii) The operator $F : D(F) \subseteq X \rightarrow Y$ is sequentially $\tau$-weakly closed. That is, $\{u_k\} \subseteq D(F)$, $u_k \xrightarrow{\tau} u$ and $F(u_k) \xrightarrow{\tau} v$ imply $u \in D(F)$ and $v = F(u)$. Moreover, the operator $F$ is continuous from $D(F)$ to $(Y, \| \cdot \|)$.

(iv) For every $m \in \mathbb{N}$, the operator $F_m$ is sequentially $\tau$-weakly closed.

(v) The function $R$ is sequentially $\tau$—lower semi-continuous.

(vi) For every $M > 0$, $\alpha > 0$ and every $m, n \in \mathbb{N}$, the sets

$$
\{ u \in X : \| F(u) \|^2 + \alpha R(u) \leq M \}
$$

and

$$
\{ u \in X_n : \| F_m(u) \|^2 + \alpha R(u) \leq M \}
$$

are $\tau$-sequentially relatively compact.

(vii) For every $u \in X$, there exist some $u_n \in X_n$ such that $d(u_n, u) \rightarrow 0$ as $n \rightarrow \infty$.

### 2.2. Well-posedness of the discretized problem

We emphasize again one of the main ideas in this work: when discretizing a nonseparable Banach space, one could work with topologies which are weaker than the original norm topology and which might be more natural than the latter. A well-posedness result for problem (7) can be stated now in this theoretical setting.

**Proposition 2.3.** Let $m, n \in \mathbb{N}$ and $\alpha, \delta > 0$ be fixed. Moreover, let assumptions 2.2 and (6), (3) be satisfied.

Then, for every $y^\delta \in Y$, there exists at least one minimizer $u$ of (7).

Moreover, the minimizers of (7) are stable with respect to the data $y^\delta$ in the following sense: if $\{\gamma_k\}_{k \in \mathbb{N}}$ converges strongly to $y^\delta$, then every sequence $\{u_k\}_{k \in \mathbb{N}}$ of the minimizers of (7) where $y^\delta$ is replaced by $\gamma_k$ has a subsequence $\{u_l\}_{l \in \mathbb{N}}$ which converges with respect to the topology $\tau$ to a minimizer $\tilde{u}$ of (7) and for which $\{R(u_l)\}_{l \in \mathbb{N}}$ converges to $R(\tilde{u})$, as $l \rightarrow \infty$.

The proof is analogous to the one for theorem 3.23 in [36], which in turn generalizes a proof in [15] from Hilbert to Banach spaces. With a different notation, the result and its proof can also be found in [40]. The same applies to the convergence result stated below—see [40, sections 2–4]. However, for the sake of completeness and connection to the main results in subsection 2.3, we have included the statement and the proof here.

**Theorem 2.4.** Let assumptions 2.2 be satisfied. Moreover, assume that

(i) equation (1) has an $R$-minimizing solution $\bar{u}$ in the interior of $D(R) \cap D(F)$, considered in the norm topology;

(ii) $v_0 \in D(F)$ for $n$ sufficiently large, where $v_0 \in X_n$ and $d(v_0, \bar{u}) \rightarrow 0$ as $n \rightarrow \infty$;
(iii) the parameter \( \alpha = \alpha(m, n, \delta) \) is such that \( \alpha \to 0, \frac{\delta^2}{\alpha} \to 0, \frac{\rho_m}{\alpha} \to 0 \) and
\[
\frac{\| F(v_n) - y \|}{\sqrt{\alpha}} \to 0
\]
as \( \delta \to 0 \) and \( m, n \to \infty \).

If (6) and (3) also hold, then every sequence \( \{u_k\} \), with \( u_k := u_{m_k,n_k}^{\delta_k} \) and \( \alpha_k := \alpha(m_k, n_k, \delta_k) \) where \( \delta_k \to 0, n_k \to \infty, m_k \to \infty \) as \( k \to \infty \) and \( u_k \) is a solution of (7), has a subsequence \( \{u_{j_k}\} \) which converges with respect to the topology \( \tau \) to an \( \mathcal{R} \)-minimizing solution \( \tilde{u} \) of equation (1) and for which \( \{\mathcal{R}(u_k)\}_{k \in \mathbb{N}} \) converges to \( \mathcal{R}(\tilde{u}) \), as \( l \to \infty \). Moreover, if \( \tilde{u} \) is the unique solution of (1), then the entire sequence \( \{u_k\} \) converges to \( \tilde{u} \) in the sense of \( \tau \) and \( \mathcal{R} \) as above.

**Proof.** From (4) and assumption 2.2, item 7, it follows that \( \mathcal{R}(v_n) \to \mathcal{R}(\tilde{u}) \) for \( n \to \infty \). From the definition of \( u_{m,n}^{\alpha,\delta} \), the estimate (3) and (6), it follows that
\[
\begin{align*}
\| F_m(u_{m,n}^{\alpha,\delta}) - y^\delta \|^2 + \alpha \mathcal{R}(u_{m,n}^{\alpha,\delta}) &\leq \| F_m(v_n) - y^\delta \|^2 + \alpha \mathcal{R}(v_n) \\
&\leq (\| F_m(v_n) - F(v_n) \| + \| F(v_n) - F(\tilde{u}) \| + \| F(\tilde{u}) - y^\delta \|)^2 + \alpha \mathcal{R}(v_n) \\
&\leq (\rho_m + \| F(v_n) - y \| + \delta)^2 + \alpha \mathcal{R}(v_n).
\end{align*}
\]

Therefore,
\[
\mathcal{R}(u_{m,n}^{\alpha,\delta}) \leq \frac{(\rho_m + \| F(v_n) - y \| + \delta)^2}{\alpha} + \mathcal{R}(v_n).
\]

Assumption (iii) now guarantees that
\[
\limsup_{n \to \infty} \mathcal{R}(u_{m,n}^{\alpha,\delta}) \leq \mathcal{R}(\tilde{u}).
\]

Observe that \( \lim_{n \to \infty} \| F(v_n) - F(\tilde{u}) \| = 0 \) since \( F \) is continuous at \( \tilde{u} \) (compare assumption 2.2, item 3 and (i), (ii)). By (iii), the quantity \( (\rho_m + \| F(v_n) - F(\tilde{u}) \| + \delta)^2 \) in (10) converges to zero as \( \delta \to 0 \) and \( m, n \to \infty \), and then
\[
\lim \| F_m(u_{m,n}^{\alpha,\delta}) - y^\delta \| = 0.
\]

Therefore, by applying again (6) and (3), we also have that
\[
F(u_{m,n}^{\alpha,\delta}) \to y,
\]
with respect to the norm of \( Y \), as \( \delta, \alpha \to 0 \) and \( m, n \to \infty \). Consider \( \alpha_k := \alpha(m_k, n_k, \delta_k) \) and \( u_k := u_{m_k,n_k}^{\alpha_k} \). Note that \( (\| F_m(u_k) \|^2 + \alpha_k \mathcal{R}(u_k)) \) is bounded. Hence, by (6) and by the compactness hypothesis in assumption 2.2, there exists a subsequence \( \{u_{j_k}\}_{k \in \mathbb{N}} \) which is \( \tau \)-convergent to some \( \tilde{u} \in D_{\alpha} \). Due to the lower semicontinuity of \( \mathcal{R} \) and (11), we get
\[
\mathcal{R}(\tilde{u}) \leq \liminf_{j \to \infty} \mathcal{R}(u_{j_k}) \leq \limsup_{j \to \infty} \mathcal{R}(u_{j_k}) \leq \mathcal{R}(\tilde{u}).
\]

We also have that \( F(\tilde{u}) = y \), as \( j \to \infty \) due to (12) and item 4 in assumption 2.2. Therefore, \( \tilde{u} \) is an \( \mathcal{R} \) minimizing solution of equation (1) and
\[
\mathcal{R}(\tilde{u}) = \lim_{j \to \infty} \mathcal{R}(u_j).
\]

If the solution \( \tilde{u} \) is unique, the only limit point of \( \{u_k\} \) with respect to \( \tau \) and \( \mathcal{R} \) is \( \tilde{u} \). \( \square \)

**Remark 2.5.** In some situations, convergence of a sequence \( \{u_k\} \) to \( u \) with respect to the topology \( \tau \) and such that \( \mathcal{R}(u_k) \to \mathcal{R}(u) \) implies \( d(u_k, u) \to 0 \) as \( k \to \infty \). This happens for \( BV \) and \( BD \) which are embedded into \( Z = L^1 \), and \( \tau \) is chosen as the weak* topology, \( \mathcal{R} \) is the total variation and total deformation seminorm, respectively—see section 4.
2.3. Convergence rates

In this section, we establish error estimates for the approximation method we analyze, with respect to Bregman distances. To this end, assume throughout this section that the regularization functional $R$ is convex.

Recall that the Bregman distance with respect to a possibly non-smooth convex functional $R$ is defined by

$$D_R(v, u) = \{ D_\xi R(v, u) : \xi \in \partial R(u) \neq \emptyset \}, \quad u, v \in D(R),$$

where

$$D_\xi R(v, u) = R(v) - R(u) - \langle \xi, v - u \rangle.$$

More information about Bregman distances and their role in optimization and inverse problems can be found in [34]. Error estimates for variational or iterative regularization of (1) by means of a non-quadratic penalty have been shown in [6, 7, 18, 22, 34, 35]. The Bregman distance $D_R$ associated with $R$ was naturally chosen as the measure of discrepancy between the error estimates.

We assume Fréchet differentiability of the operator $F$ around $\bar{u}$ which is considered to be in the interior of $D(F) \cap D(R)$; moreover, assume that its extension to the space $Z$ is also Fréchet differentiable around $\bar{u}$. In fact, our study is based on the following sourcewise representation:

there exists $\omega \in Y$ such that

$$\xi = F'(\bar{u})^* \omega \in \partial R(\bar{u}), \quad (13)$$

and on the following nonlinearity condition (see also [35]):

there exist $\varepsilon > 0$ and $c > 0$ such that

$$\| F(u) - F(\bar{u}) - F'(\bar{u})(u - \bar{u}) \| \leq c D_R(u, \bar{u}), \quad (14)$$

is satisfied for all $u \in D(F) \cap U_{\varepsilon}(\bar{u})$ with respect to the above subgradient $\xi$ and such that

$$c \| \omega \| < 1. \quad (15)$$

Here

$$D_R(u, \bar{u}) = R(u) - R(\bar{u}) - \langle \xi, u - \bar{u} \rangle,$$

where $\xi \in \partial R(\bar{u})$ satisfies (13). Let us denote by

$$\gamma_n := \| F'(\bar{u})(v_n - \bar{u}) \|, \quad (16)$$

$$\lambda_n := D_R(v_n, \bar{u}). \quad (17)$$

Here $\{v_n\}$ is a sequence as in theorem 2.4.

In the following we derive a relation between the Bregman distance and the metric $d$ at $\bar{u}$. Observe that

$$D_R(v_n, \bar{u}) = R(v_n) - R(\bar{u}) - \langle \omega, F'(\bar{u})(v_n - \bar{u}) \rangle.$$

One has $R(v_n) - R(\bar{u}) \to 0$ since $d(v_n, \bar{u}) \to 0$, as $n \to \infty$ (see (4)). Moreover, $F'(\bar{u})(v_n - \bar{u}) \to 0$. As a consequence, $\lim_{n \to \infty} \lambda_n = 0$. Thus, convergence with respect to the metric $d$ is stronger than convergence with respect to the related Bregman distance.

**Theorem 2.6.** Suppose that assumption 2.2, the assumptions in the previous result, and inequalities (3), (13) and (14) hold. Moreover, assume that $\rho_m = O(\delta + \lambda_n + \gamma_n)$, with $\lambda_n, \gamma_n$ given by (17) and (16). If $\alpha \sim \max\{\delta, \lambda_n, \gamma_n\}$, then

$$D_R(u^m, \bar{u}) = O(\delta + \lambda_n + \gamma_n). \quad (18)$$

6
Proof. We have
\[\|F_m(u_{m,n}^{\alpha,\delta}) - y^\delta\|^2 + \alpha R(u_{m,n}^{\alpha,\delta})\]
\[\leq \|F_m(v_n) - y^\delta\|^2 + \alpha R(v_n)\]
\[\leq \|F_m(v_n) - F(v_n)\| + \|F(v_n) - F(\bar{u})\| + \|F(\bar{u}) - y^\delta\|^2 + \alpha R(v_n)\]
\[\leq (\rho_m + \|F(v_n) - F(\bar{u}) - F'(\bar{u})(v_n - \bar{u})\| + \|F'(\bar{u})(v_n - \bar{u})\| + \delta)^2 + \alpha R(v_n)\]
\[\leq \rho_m + c_\lambda_n + y_n + \delta^2 + \alpha R(v_n).\]

Denote \(\beta_n := (\rho_m + c_\lambda_n + y_n + \delta)^2\). (19)

We use (14) and get
\[\|F_m(u_{m,n}^{\alpha,\delta}) - y^\delta\|^2 + \alpha D_{R}(u_{m,n}^{\alpha,\delta}, \bar{u})\]
\[\leq \beta_n + \alpha R(v_n) - \alpha R(\bar{u}) - \alpha(\xi, u_{m,n}^{\alpha,\delta} - \bar{u})\]
\[= \beta_n + \alpha D_{R}(v_n, \bar{u}) - \alpha(\xi, u_{m,n}^{\alpha,\delta} - v_n)\]
\[= \beta_n + \alpha \lambda_n - \alpha(\omega, F'(\bar{u})(u_{m,n}^{\alpha,\delta} - v_n))\]
\[= \beta_n + \alpha \lambda_n - \alpha(\omega, F'(\bar{u})(u_{m,n}^{\alpha,\delta} - \bar{u})) + \alpha(\omega, F'(\bar{u})(v_n - \bar{u}))\]
\[\leq \beta_n + \alpha \lambda_n + \alpha c_\omega D_{R}(u_{m,n}^{\alpha,\delta}, \bar{u}) + \alpha \|\omega\| F(u_{m,n}^{\alpha,\delta}) - F(\bar{u})\] + \(\alpha \|\omega\| y_n\).

Therefore,
\[\|F_m(u_{m,n}^{\alpha,\delta}) - y^\delta\|^2 + \alpha(1 - c_\omega) D_{R}(u_{m,n}^{\alpha,\delta}, \bar{u}) \leq \beta_n + \alpha \lambda_n + \alpha \|\omega\| (\xi_n + y_n),\] (20)

where \(\xi_n = \|F_m(u_{m,n}^{\alpha,\delta}) - y\|.\) Due to (15), the term \(\alpha(1 - c_\omega) D_{R}(u_{m,n}^{\alpha,\delta}, \bar{u})\) is nonnegative.

Therefore,
\[\|F_m(u_{m,n}^{\alpha,\delta}) - y^\delta\|^2 \leq \beta_n + \alpha \lambda_n + \alpha \|\omega\| (\xi_n + y_n).\] (21)

Using (3) we have
\[\xi_n^2 \leq (\|F_m(u_{m,n}^{\alpha,\delta}) - y^\delta\| + \|y^\delta - y\|)^2 \leq 2\|F_m(u_{m,n}^{\alpha,\delta}) - y^\delta\|^2 + 2\delta^2.\]

This together with inequality (21) implies
\[\xi_n^2 \leq 2\beta_n + 2\alpha \lambda_n + 2\alpha \|\omega\| \xi_n + 2\alpha \|\omega\| y_n + 2\delta^2,\]

which yields
\[\xi_n \leq 2\alpha \|\omega\| + (2\alpha^2 \|\omega\|^2 + 2\delta^2 + 2\beta_n + 2\alpha \lambda_n + 2\alpha \|\omega\| y_n)^{1/2}.\] (22)

From (20), it follows that
\[\alpha(1 - c_\omega) D_{R}(u_{m,n}^{\alpha,\delta}, \bar{u}) \leq \beta_n + \alpha \lambda_n + \alpha \|\omega\| \xi_n + \alpha \|\omega\| y_n,\]

with \(\xi_n\) estimated above. Using (19) and taking \(\alpha \sim \max\{\delta, \lambda_n, y_n\}\) yield the above convergence rate. \(\square\)

3. The case of separable Banach spaces

Let \(X\) be a separable Banach space. Consider a nested sequence of finite-dimensional subspaces \(X_n, n \in \mathbb{N}\), such that
\[\bigcup_{n \in \mathbb{N}} X_n = X,\]

where the closure is considered with respect to the norm topology of \(X\). By letting \(Z = X\), a weak topology \(\tau\) on \(X\), a regularization functional \(\mathcal{R} : X \rightarrow [0, +\infty]\) and using the assumptions
employed for the results in section 2, one obtains stability and convergence results similar to proposition 2.3 and theorem 2.4. Moreover, if $\mathcal{R}$ is convex, then convergence rates can also be established.

**Remark 3.1.** In some situations, convergence of a sequence $\{u_k\}$ to $u$ with respect to the topology $\tau$ and such that $\mathcal{R}(u_k) \to \mathcal{R}(u)$ implies $\|u_k - u\| \to 0$, as $k \to \infty$. This property is known as the Radon–Riesz or Kadec–Klee or $H$-property—see for instance [29, section 2.5]. Locally uniformly convex reflexive Banach spaces, in particular Hilbert spaces, enjoy such a property when $\tau$ is chosen as the weak topology on $X$ and $\mathcal{R} = \|\cdot\|^p$, with $p \in (1, +\infty)$. The same is for $L^p$-spaces and $W^{m,p}$ spaces, but also for $L^1$ when $\tau$ is the weak topology and $\mathcal{R}$ is the Shannon entropy (see [5]).

### 3.1. Sparsity regularization

Let $\{\phi_i\}$ be an orthonormal basis of $L^2(\Omega)$. Denote by $X$ the Banach space $\ell^2$ which is identified with the functions with bounded $\ell^2$ coefficients. Let $X_n$ be the linear span of the first $n$ modes. For sparsity regularization one usually takes $\mathcal{R}(u) = \sum_i |u_i|$, where $u = \sum_i u_i \phi_i$. Thus, we consider the regularization method of minimizing the functional

$$u \to \|F(u) - y_\delta\|^2 + \alpha \sum_i |u_i|.$$ 

In this case, the topology $\tau$ on $X$ is the weak topology of $\ell^2$. The regularization results also apply in this setting.

Another example of the regularization term which promotes sparsity is

$$\mathcal{R}(u) = \sum_i |u_i|^p, \quad p \in (0, 1). \quad (23)$$

This setting with $X = \ell^2$ and $\tau$ taken as the weak topology of $\ell^2$ is also covered by the regularization theory analyzed in this work. Moreover, convergence of a sequence $\{u_k\}$ to $u$ with respect to the topology $\tau$ and such that $\mathcal{R}(u_k) \to \mathcal{R}(u)$ implies convergence of $\{u_k\}$ to $u$ relative to the quasinorm (23), as $k \to \infty$, cf [19].

### 4. Particular nonseparable spaces

#### 4.1. The bounded variation function space

Recall that, for a bounded Lipschitz domain $\Omega \subset \mathbb{R}^N$ and for a given $N \in \mathbb{N}$, the space $BV(\Omega)$ of $L^1(\Omega)$-functions of bounded variation mapping $\Omega$ into $\mathbb{R}$ can be defined as the set of functions $w \in L^1(\Omega)$ such that the total variation of $w$ is finite, that is,

$$\int_{\Omega} |Dw|_p = \sup \left\{ \int_{\Omega} w(x) \psi(x) \, dx : \psi \in C^1(\Omega), |\psi(x)|_{p'} \leq 1 \text{ for all } x \in \Omega \right\} < \infty.$$ 

Here, $|\cdot|_{p'}$ denotes the $l_{p'}$ vector norm, where $p' = p/(p - 1)$ is the conjugate exponent to $p$. In particular we are interested in the cases $p = 1, 2$, where $l_{p'} = l_{\infty}, l_2$. The case $p = 2$ corresponds to *isotropic* total variation.

Let us recall several properties of the space $BV(\Omega)$.

- It is the dual of a separable Banach space (see [3, remark 3.12]) when provided with the norm

$$\|u\|_{BV} = \|u\|_{L^1} + \int_{\Omega} |Du|_p.$$
The space $\text{BV}(\Omega)$ is continuously embedded in $L^r(\Omega)$, where $1 \leq r \leq \frac{N}{N-1}$.

We consider the setting $X = \text{BV}(\Omega)$, with $\tau$ being the weak* topology on $\text{BV}(\Omega)$, and $Z = L^1(\Omega)$.

The functional $\mathcal{R}$ is the total variation seminorm. Consider

$$d(u, v) = \|u - v\|_{L^1(\Omega)} + \left| \int_{\Omega} |Du|_p - \int_{\Omega} |Dv|_p \right|.$$  

The metric $d$ is also used in [8], which gives the so-called strict convergence, according to [3]. Moreover, the strict convergence of a sequence $\{u_k\}$ to $u$ is equivalent to the convergence with respect to the topology $\tau$ together with $\int_{\Omega} |Du_k|_1 \to \int_{\Omega} |Du|_1$ as $k \to \infty$ since the weak* convergence of a sequence $\{u_k\}$ to $u$ in $\text{BV}(\Omega)$ is equivalent to the boundedness of $\{\|u_k\|_{\text{BV}}\}$ together with the convergence of $\{u_k\}$ to $u$ in $L^1(\Omega)$—see, e.g., proposition 3.13 in [3].

The choice of the vector norm in the definition of the bounded variation norm is of special importance for approximating the BV space by subspaces consisting of piecewise constant functions. Assume that $\{\Omega_j\}$ is a decomposition of $\Omega$, and consider the following finite-dimensional spaces:

$$X_n = \left\{ u_n = \sum_{j=1}^{n} u^j \chi_{\Omega_j} : u^j \in \mathbb{R}, 1 \leq j \leq n \right\}.$$  

When considering a partition of $\Omega$ into uniform parallelepipeds, we can only guarantee the density assumption (5) when considering the $L_1$-norm in the definition of BV—see [9]. If one wants an isotropic behavior of the regularization term, one has to consider the $L_2$-norm in the definition of the BV-seminorm. The problem is that in the case of uniform parallelepipeds (for instance pixels in imaging), there is no convergence with respect to this isotropic BV-seminorm. One has to consider a more general partition of the domain $\Omega$, that allows us to approximate level lines with any direction. One idea is to use an irregular triangulation. These observations have been made by [8] and [9]; here we only state their main result, concerning anisotropic and isotropic total variation.

**Theorem 4.1 ([8, 9]).** Let $\Omega \subset \mathbb{R}^n$ be a polygonal domain and let $h_0 > 0$.

- Given $u \in \text{BV}(\Omega)$; then there exists a family $\{t_h : 0 < h \leq h_0\}$ of triangulations of $\Omega$ such that the mesh size of $t_h$ is at most $h$, and functions $u_h \in A_h^0$, where $A_h^0$ denotes the space of piecewise constant functions corresponding to the triangulation $t_h$ such that

$$\|u - u_h\|_{L^1} + \left| \int_{\Omega} |Du|_2 - \int_{\Omega} |Du_h|_2 \right| \to 0 \quad \text{as} \quad h \to 0.$$

- Given $u \in \text{BV}(\Omega)$. Let $\{\Omega_j\}$ be a decomposition of $\Omega$, into parallelepipeds, such that the maximal length of a parallelepiped is smaller then $h$. Then there exist functions $u_h \in V_h^n$ where $V_h^n$ denotes the space of piecewise constant functions corresponding to the partition $\{\Omega_j\}$, such that

$$\|u - u_h\|_{L^1} + \left| \int_{\Omega} |Du|_1 - \int_{\Omega} |Du_h|_1 \right| \to 0 \quad \text{as} \quad h \to 0.$$  

Another type of approximation of $X$ is the ‘$J_{\mu}$-approximation property’ employed in [17] which, adapted to our notation, reads as follows.

Given $u \in X$, there exists a sequence $\{u_n\} \subset X_n$ such that $\|u - u_n\|_Z \to 0$ and $\Phi_{\alpha}(u_n) \to \Phi_{\alpha}(u)$ as $n \to \infty$, for any $\alpha \geq 0$, where $\Phi_{\alpha}$ is given by

$$\Phi_{\alpha}(u) = \|F(u) - y^*\|^2 + \alpha \mathcal{R}(u).$$
In [17] one aims at approximating the minimizers of $\Phi_\alpha (\text{which depends on } \alpha)$ in $X$ by the minimizers of $\Phi_\alpha$ in $X_n$ for a fixed $\alpha > 0$, while our aim is to approximate the solutions of the operator equation by the minimizers of $\Phi_\alpha$ in $X_n$ when the regularization parameter $\alpha$ depends on the dimension $n$.

Also, note that the strong convergence in the BV-space is shown in [26, 27] for regularization with a functional $J$ which is of Shannon entropy type.

We summarize the results of this section in an example.

**Example 4.2.** Let $X = \text{BV}(\Omega)$ with the weak* topology $\tau$. Moreover, let $Z = L^1(\Omega)$ and let $d$ be the metric

$$d(u,v) = \|u - v\|_{L^1(\Omega)} + \left| \int_{\Omega} |Du|_p - \int_{\Omega} |Dv|_p \right|,$$

with $p = 1, 2$, depending on the decomposition of $\Omega$.

Let $F_\alpha$ and $F : D(F) \subseteq \text{BV}(\Omega) \to L^2(\hat{\Omega})$ satisfy the related conditions in assumption 2.2.

Then according to theorem 4.1, for every $u \in \text{BV}(\Omega)$, there exists an approximating sequence of piecewise constant functions. Consequently, minimization of the discretized regularized problem is well-posed, stable and convergent (cf proposition 2.3). The piecewise constant regularizers $\{u_{m,n}^\alpha\}$ approximate the $\mathcal{R}$-minimizing solution $\bar{u}$ in the sense of the metric (24).

### 4.2. The bounded deformation function space

In the following let $\Omega = (0, 1)^N$ be the open unit cube. We choose the simple geometry not to be forced to take into account the approximations of $\Omega$, or irregular meshes, for the finite element method considered below.

The space $\text{BD}(\Omega)$ [39] of the functions with *bounded deformation* in an open set $\Omega \subset \mathbb{R}^N$ is defined as the set of functions $u = (u^1, \ldots, u^N) \in L^1(\Omega; \mathbb{R}^N)$ such that the symmetric distributional derivative

$$E_{ij} u := \frac{1}{2} (D_i u^j + D_j u^i)$$

is a (matrix-valued) measure with finite total variation in $\Omega$:

$$\text{BD}(\Omega) := \{ u \in L^1(\Omega; \mathbb{R}^N), E_{ij}(u) \in M_1(\Omega), i, j = 1, \ldots, N \},$$

where $M_1(\Omega)$ denotes the space of bounded measures. $\text{BD}(\Omega)$ is a nonseparable Banach space provided with the norm

$$\|u\|_{\text{BD}} = \|u\|_{L^1(\Omega; \mathbb{R}^N)} + \sum_{i,j=1}^N \int_{\Omega} |E_{ij}(u)|.$$ 

This space is strictly larger than the space of bounded variation functions $\text{BV}(\Omega; \mathbb{R}^N)$. It was introduced in [37] and [28] and has been widely considered in the literature (see [38, 39]) in connection with the mathematical theory of plasticity. Several interesting properties of the BD space are as follows: it is the dual of a separable Banach space (see [39]); the space $\text{BD}(\Omega)$ is continuously embedded in $L^p(\Omega; \mathbb{R}^N)$, where $1 \leq p \leq \frac{N}{N-1}$. In addition the space $\text{BD}(\Omega; \mathbb{R}^N)$ is not separable — if it would be, the space $\text{BV}(\Omega; \mathbb{R}^N)$, which is the subspace of $\text{BD}(\Omega; \mathbb{R}^N)$ where all components $u^j, j = 2, \ldots, N$, vanish, would be as well. However, this is not true as already stated in the previous example.

Let us return to the finite-dimensional regularization framework we investigate in this work.
Consider the setting $X = \text{BD}(\Omega)$, $\tau$ the weak* topology on $\text{BD}(\Omega)$ and $Z = L^1(\Omega; \mathbb{R}^N)$. We associate with $n \in \mathbb{N}$ the discretization size $h_n := \frac{1}{n}$, use multi-indices $\alpha := (\alpha_1, \ldots, \alpha_N)$ and set $A := \{0, \ldots, n\}^N$.

We consider the finite-dimensional product space of piecewise linear splines:

$$X_n := \left\{ u_n = \sum_{\alpha \in A} \sum_{k=1}^N u_\alpha^k \Delta \left( \frac{x - \xi_\alpha}{h_n} \right) : u_\alpha^k \in \mathbb{R} \right\}$$

where $\xi_\alpha \in h_n \{0, 1, \ldots, n\}^N$, and $\Delta$ is the following function:

$$\Delta(x) := \prod_{i=1}^N \max(0, 1 - |x^i|).$$

This finite element discretization has already been used in [10] for the numerical minimization of variational energies. For $x \in (0, h_n)^N$ the derivative of $\Delta((x - \xi_\alpha)/h_n)$ in the direction $e_j$ is given by

$$D_j \Delta \left( \frac{x - \xi_\alpha}{h_n} \right) = \text{sign}(\xi_\alpha^j - x^j) \frac{1}{h_n} \prod_{i \neq j} (\max(0, 1 - |x^i - \xi_\alpha^i/h_n|)).$$

For every $\alpha$ with $\alpha_i < n$ and for every $k, 1 \leq k \leq N$ define $A_{\alpha}^k := \{\beta \in A_\alpha, \beta_k = \alpha_k\}$. Additionally define $\Omega_\alpha$ as the $N$-dimensional cube, spanned by the vectors $[\xi_\alpha + e_k h_n]_{k=1, \ldots, N}$, and $A_\alpha = \bigcup_{k=1}^N A_{\alpha}^k$ (see figure 1).

Moreover, for $\beta \in A_{\alpha}$, we have

$$\int_{\Omega_n} D_j \Delta \left( \frac{x - \xi_\beta}{h_n} \right) dx = \begin{cases} -\left(\frac{h_n}{2}\right)^{N-1} & \text{if } \beta^j = \alpha^j \\ \left(\frac{h_n}{2}\right)^{N-1} & \text{if } \beta^j = \alpha^j + 1. \end{cases}$$

(25)

In the following we prove the main result on a pseudometric.

**Theorem 4.3.** We assume that $h_n \to 0$ when $n \to \infty$. Then for every $u \in \text{BD}(\Omega, \mathbb{R}^N) \cap L^1(\Omega, \mathbb{R}^N)$, $1 \leq r < \infty$, we can find a sequence $\{u_n\}$, with $u_n \in X_n$, such that

$$\lim_{n \to \infty} \int_{\Omega} |u - u_n|^r dx = 0 \quad \text{and} \quad \lim_{n \to \infty} \int_{\Omega} |E u_n| = \int_{\Omega} |E u|.$$

Setting

$$d(u_1, u_2) = \|u_1 - u_2\|_{L^1(\Omega; \mathbb{R}^N)} + \int_{\Omega} |E u_1| - \int_{\Omega} |E u_2|,$$

(26)

we obtain $\lim_{n \to \infty} d(u_n, u) = 0$. 

\[11\]
In order to prove this theorem, we need some additional facts on BD, given by the following lemmas.

**Lemma 4.4.** For every \( n \in \mathbb{N} \), the inclusion \( X_n \subset \text{BD}(\Omega, \mathbb{R}^N) \) holds and for each \( u = (u^1, \ldots, u^N) \in X_n \),

\[
\int_{\Omega} |E u| = \sum_{\alpha} \sum_{k=1}^{N} \left| \sum_{\beta \in A^k_\alpha} (u_{\beta+e_k}^k - u_\beta^k) \right| \left( \frac{h_n}{2} \right)^{N-1}
\]

\[
+ \sum_{\alpha} \sum_{k \neq l}^{N} \left| \sum_{\beta \in A^k_\alpha} (u_{\beta+e_k}^k - u_\beta^k) + \sum_{\beta \in A^l_\alpha} (u_{\beta+e_l}^l - u_\beta^l) \right| \left( \frac{h_n}{2} \right)^{N-1}
\]

**Proof.** From the definition of \( u \in X_m \), we obtain

\[
D_l u^k(x) = \sum_{\alpha} u_{\alpha}^k D_l \Delta \left( \frac{x - \xi_{\alpha}}{h_n} \right).
\]

Moreover, since \( \Omega = \bigcup_{\alpha \in \Lambda} \Omega_\alpha \), we can split up the integral. When integrating \( D_l u^k + D_k u^l \) over \( \Omega_\alpha \), we only have to sum over those \( \beta \) for which \( \Delta \left( \frac{x - \xi_{\beta}}{h_n} \right) \neq 0 \). Hence we only sum over \( \beta \in A^k_\alpha \) in the inner sum:

\[
\int_{\Omega_\alpha} |D_l u^k + D_k u^l| = \sum_{\alpha} \int_{\Omega_\alpha} |D_l u^k + D_k u^l|
\]

\[
= \sum_{\alpha} \int_{\Omega_\alpha} \left| \sum_{\beta \in A^k_\alpha} u_{\beta}^k D_l \Delta \left( \frac{x - \xi_{\beta}}{h_n} \right) + u_{\beta}^l D_k \Delta \left( \frac{x - \xi_{\beta}}{h_n} \right) \right|.
\]

Next we use (25) and obtain

\[
\int_{\Omega_\alpha} |D_l u^k + D_k u^l| = \left| \sum_{\beta \in A^k_\alpha} (u_{\beta+e_k}^k - u_\beta^k) + \sum_{\beta \in A^l_\alpha} (u_{\beta+e_l}^l - u_\beta^l) \right| \left( \frac{h_n}{2} \right)^{N-1}.
\]

The lemma follows from summing over all \( \alpha \) and all \( l, k = 1, \ldots, N \). \( \square \)

**Lemma 4.5.**

1. If \( \{u_n\} \subset \text{BD}(\Omega) \) and \( u_n \rightharpoonup u \) in \( (L^1(\Omega))^N \), then

\[
\int_{\Omega} |E_{ij} u| \leq \liminf_{n \to \infty} \int_{\Omega} |E_{ij} u_n|.
\]

2. For every \( u \in \text{BD}(\Omega) \cap L^r(\Omega) \), \( 1 \leq r < \infty \), there exists a sequence \( \{u_n\} \subset C^\infty(\overline{\Omega}) \) such that

\[
\lim_{n \to \infty} \int_{\Omega} |u - u_n| r dx = 0 \quad \text{and} \quad \lim_{n \to \infty} \int_{\Omega} |E_{ij} u_n| = \int_{\Omega} |E_{ij} u|.
\]

**Proof.** 1. Follows from the standard properties of convex measures.

2. See [38, theoreme 3.2, chapitre II]. \( \square \)

Now we are ready for the proof of theorem 4.3.
Proof. The $L^p$-convergence of the piecewise polynomial functions $u_n$ can be found in the book of Ciarlet [12]. Due to lemma 4.5, we can assume that $u \in C^\infty(\Omega)$. Set the coefficient of $u_n$ as $(u_n)^k_\alpha := u^k(\xi_\alpha)$; then we have

$$u_n = \sum_{k=1}^{N} \sum_{\alpha} u^k(\xi_\alpha) \Delta \left( \frac{x - \xi_\alpha}{h_n} \right).$$

From lemma 4.4, it follows that

$$\int_{\Omega} |E_{ij} u_n| = \int_{\Omega} \left| D_i u_n^k + D_k u_n^l \right| = \left\| \sum_{\beta \in \Lambda^l_\alpha} (u^k(\xi_{\beta+e_l}) - u^k(\xi_\beta)) + \sum_{\beta \in \Lambda^l_\alpha} (u^l(\xi_{\beta+e_l}) - u^l(\xi_\beta)) \right\|(h_n)^{N-1}. \quad (27)$$

Next we use the mean value theorem: for $m, n \in \{l, k\}, m \neq n$, we can find points $\eta_{\beta,m,n}$ between $\xi_\beta$ and $\xi_{\beta+e_l}$ such that

$$u^k(\xi_{\beta+e_l}) - u^k(\xi_\beta) = h_n D_i u^k(\eta_{\beta,m,k}),$$

$$u^l(\xi_{\beta+e_l}) - u^l(\xi_\beta) = h_n D_l u^l(\eta_{\beta,k,l});$$

hence, from (27) it follows that

$$\int_{\Omega} |E_{ij} u_n| = \left\| \sum_{\beta \in \Lambda^l_\alpha} D_i u^k(\eta_{\beta,l,k}) + \sum_{\beta \in \Lambda^l_\alpha} D_l u^l(\eta_{\beta,l,k}) \right\|(h_n)^{N-1}. \quad |\Omega| \left( \frac{1}{2} \right)^{N-1}. \quad (28)$$

Summing over all $\alpha$ and $l, k = 1, \ldots, N$, and taking the limit $h_n \to 0$, we have that

$$\lim_{n \to \infty} \int_{\Omega} |E_{ij} u_n| = \int_{\Omega} |E_{ij} u|. \quad \square$$

We summarize the results of this section as follows.

Example 4.6. Let $X = \text{BD}(\Omega)$ and $\tau$ the weak* topology. Moreover, let $Z = L^1(\Omega, \mathbb{R}^N)$ and let $d$ be the metric

$$d(u, v) = \|u - v\|_{L^1(\Omega, \mathbb{R}^N)} + \int_{\Omega} |E u| - \int_{\Omega} |E v|.$$ 

Let $F_m$ and $F : D(F) \subseteq \text{BD}(\Omega) \to L^2(\hat{\Omega})$ satisfy the related conditions in assumption 2.2. Take $h = 1/n$. Then according to theorem 4.3, for every $u \in \text{BD}(\Omega, \mathbb{R}^N) \cap L^r(\hat{\Omega}, \mathbb{R}^N)$, $1 \leq r < \infty$, there exists an approximating sequence of piecewise linear functions in the sense of the metric $d$. Consequently, the minimization of the discretized regularized problem is well-posed, stable and convergent according to proposition 2.3. The piecewise linear regularizers $\{u^{a,\delta}_{m,n}\}$ approximate the $R$-minimizing solution $\overline{u}$ in the sense $\lim_{m,n} u^{a,\delta}_{m,n} = \overline{u}$ in the weak star topology and $\lim_{m,n} \int_{\Omega} |E u^{a,\delta}_{m,n}| = \int_{\Omega} |E \overline{u}|$.

4.3. The $L^\infty$-space

In this section we analyze the following regularization method:

$$u \rightarrow \|F(u) - y^d\|^2 + \alpha \|u\|_{\infty}. \quad (29)$$
In the following we show that there exist finite-dimensional subspaces \( \{X_n\} \) of \( L^\infty(\Omega) \) satisfying equality (5). More precisely, there exist finite-dimensional subspaces \( \{X_n\} \) of \( L^\infty(\Omega) \) such that for any \( u \in L^\infty(\Omega) \), one can find \( u_n \in X_n, n \in \mathbb{N}, \) with

\[
\lim_{n \to \infty} (\|u_n - u\|_p + \|u_n\|_\infty - \|u\|_\infty) = 0, \quad p \in [1, +\infty]. \tag{29}
\]

It is known that \( L^\infty(\Omega) \) is not separable, while \( L^p(\Omega) \), for every \( p \in [1, +\infty) \), is separable. However, every function in \( L^\infty(\Omega) \) can be approximated uniformly, and thus, in the \( L^\infty(\Omega) \)-norm by a sequence of simple functions. The proof of this statement is constructive. However, it provides a nonlinear approximation, in the sense that the piecewise constant functions which approximate the \( L^\infty(\Omega) \)-function do not yield linear subspaces—see, e.g., [13, section 3.2]. This does not fit the theoretical framework we consider here. An alternative is to consider the approximations of \( L^\infty(\Omega) \)-functions by piecewise constant functions in a weaker topology, as shown in the following.

For the sake of simplicity, let \( \Omega = (0, 1)^N \). Assume that \( \{\Omega_j\} \) is a decomposition of \( \Omega \) in parallelepipeds with the maximal diagonal length \( h_n \) as in [9], and consider the following finite-dimensional subspaces of \( L^\infty(\Omega) \):

\[
X_n = \left\{ u_n = \sum_{j=1}^n u_j \chi_{\Omega_j} : u_j \in \mathbb{R}, 1 \leq j \leq n \right\}.
\]

The following results are essential in proving the main statement of this section.

Denote by \( J_\varepsilon \ast u \) the mollification of \( u \), for every \( \varepsilon > 0 \).

**Theorem 4.7.** [p 36 [2]] Let \( u \) be a function which is defined on \( \mathbb{R}^N \) and vanishes identically outside \( \Omega \).

If \( u \in L^p(\Omega) \), then \( J_\varepsilon \ast u \in C^\infty(\mathbb{R}^N) \), in fact \( J_\varepsilon \ast u \in C^\infty(\overline{\Omega}) \) and \( J_\varepsilon \ast u \in L^p(\Omega) \), for every \( p \in [1, +\infty) \). Also,

\[
\lim_{\varepsilon \to 0_+} \|J_\varepsilon \ast u - u\|_p = 0, \tag{30}
\]

\[
\|J_\varepsilon \ast u\|_p \leq \|u\|_p \quad \text{for all } \varepsilon > 0. \tag{31}
\]

**Lemma 4.8.** For every \( n \in \mathbb{N} \) one has \( X_n \subseteq L^\infty(\Omega) \), and for each \( u_n \in X_n \)

\[
\|u_n\|_\infty = \max_{1 \leq j \leq n} |u_j|.
\]

**Proof.** Follows immediately from the definition of \( u_n \), taking into account that \( \mu(\Omega_j) > 0 \), for \( 1 \leq j \leq n \).

**Lemma 4.9.** If \( \{u_n\} \subseteq L^\infty(\Omega) \) and \( u \in L^\infty(\Omega) \) are such that \( \|u - u_n\|_p \to 0 \) as \( n \to \infty \), for some \( p \in [1, +\infty) \) and \( \|u_n\|_\infty \leq \|u\|_\infty \) for every \( n \in \mathbb{N} \), then \( \{u_n\} \) converges weakly* to \( u \) and \( \|u_n\|_\infty \to \|u\|_\infty \) as \( n \to \infty \).

**Proof.** Since \( \|u_n\|_\infty \) is bounded, there exists a subsequence \( \{u_k\} \) which converges weakly* to some \( v \in L^\infty(\Omega) \), cf Alaoglu–Bourbaki theorem, [23, p 70]. By the definition of weak* convergence, one obtains that \( \{u_k\} \) converges also weakly, with respect to \( L^p(\Omega) \) to some \( v \). Therefore \( u = v \). In fact, every subsequence of \( \{u_n\} \) converges weakly* to \( u \), which yields a weak* convergence of the entire sequence \( \{u_n\} \) to \( u \). In addition, the weak* lower semicontinuity of the \( L^\infty \)-norm implies \( \|u\|_\infty \leq \liminf_{n \to \infty} \|u_n\|_\infty \). Thus the assertions are proved.
Remark 4.12. One can define the subspaces as in (29) due to lemma 4.10, the conclusion follows immediately.

Lemma 4.10. For every $u \in C^\infty(\bar{\Omega})$, there exists a sequence $\{u_n\}$ with $u_n \in X_n$ such that

$$\lim_{n \to \infty} (\|u_n - u\|_{L^p} + \|u_n\|_\infty - \|u\|_\infty) = 0,$$

$p \in [1, +\infty)$.

Proof. Let $\xi_j$ be the gravity centers of the parallelepipeds $\Omega_j$. Define

$$u_n = \sum_{j=1}^n u(\xi_j)\chi_{\Omega_j}.$$

Then $\|u_n - u\|_{L^p} \to 0$—see, for instance, the book of Ciarlet [12]. Moreover, by using lemma 4.8,

$$\|u_n\|_\infty = \max_{1 \leq j \leq n} |u(\xi_j)| \leq \max_{x \in \bar{\Omega}} |u(x)| = \|u\|_\infty$$

for all $n \in \mathbb{N}$, where the last equality holds due to the continuity of $u$ on $\bar{\Omega}$. Thus, lemma 4.9 applies and yields $\|u_n\|_\infty - \|u\|_\infty \to 0$ as $n \to \infty$. □

Theorem 4.11. Assume that $h_n \to 0$ when $n \to \infty$. Then for every $u \in L^\infty(\bar{\Omega})$, one can find $u_n \in X_n$ such that (29) holds.

Proof. Let $u \in L^\infty(\bar{\Omega})$ and $p \in [1, +\infty)$. By theorem 4.7, there exists $\{u_j\} \subset C^\infty(\mathbb{R}^N)$, in fact in $C^\infty(\bar{\Omega})$, with $\{u_j\} \subset L^p(\bar{\Omega})$ such that

$$\lim_{j \to \infty} \|u_j - u\|_p = 0,$$

and $\|u_j\|_p \leq \|u\|_p$ for all $j \in \mathbb{N}$.

By letting $p \to +\infty$ in the last inequality and using theorem 2.8 on p 25 in [1], one also has

$$\|u_j\|_\infty \leq \|u\|_\infty$$

for all $j \in \mathbb{N}$.

Lemma 4.9 applies and yields $\|u_j\|_\infty \to \|u\|_\infty$ as $j \to \infty$.

By consequence, every function in $L^\infty(\bar{\Omega})$ can be approximated by functions from $C^\infty(\bar{\Omega})$ in the sense stated in (29). Since every function $v \in C^\infty(\bar{\Omega})$ can be approximated by $u_n \in X_n$ as in (29) due to lemma 4.10, the conclusion follows immediately. □

Remark 4.12. One can define the subspaces $X_n$ in the previous theorem also by means of piecewise polynomial functions of degree no bigger than 1 in each variable, which are continuous on $\bar{\Omega}$. Moreover, one can employ $n$-simplices instead of parallelepipeds and piecewise linear functions, according to remark 3.8 in [9].

We summarize the results of this section as follows.

Example 4.13. Let $X = L^\infty(\bar{\Omega})$ and $\tau$ the weak* topology. Moreover, let $Z = L^p(\bar{\Omega})$, $p \in (1, +\infty)$ and let $d$ be the metric:

$$d(u, v) = \|u - v\|_{L^p(\bar{\Omega})} + \|u\|_\infty - \|v\|_\infty. \quad (32)$$

Let $F_m$ and $F : D(F) \subseteq BV(\bar{\Omega}) \to L^2(\bar{\Omega})$ satisfy the related conditions in assumption 2.2.

Take $h = 1/n$. Then according to theorem 4.11, for every $u \in L^\infty(\bar{\Omega})$, there exists a sequence of piecewise constant functions which approximates $u$ in the sense of metric $d$.

Consequently, the minimization of the discretized regularized problem is well-posed, stable and convergent according to proposition 2.3. The piecewise constant regularizers $\{u_{m,\alpha}\}$ approximate the $R$-minimizing solution $\tilde{u}$ in the sense $\lim u_{m,\alpha} = \tilde{u}$ in the weak star topology and $\lim \|u_{m,\alpha}\|_\infty = \|\tilde{u}\|_\infty$. 

15
The convergence described in the previous sentence is usually weaker than the convergence with respect to the metric given by (32). However, in special situations it implies strong convergence with respect to any $L^p$-norm with $p \in [1, +\infty)$, cf theorem 3 in [21]. The fact that the two types of convergence in discussion are not equivalent in general is shown by the following counterexample.

Consider the Rademacher functions $f_n : [0, 1] \to \{-1, 1\}$:

$$f_n(t) = (-1)^{i+1} \quad \text{if} \quad x \in \left[\frac{i-1}{2^n}, \frac{i}{2^n}\right), \quad 1 \leq i \leq 2^n.$$ 

This sequence converges weakly star to zero in $L^\infty([0, 1])$, but not in the $L^1([0, 1])$-norm. Moreover, consider $g_n : [0, 2] \to \mathbb{R}$,

$$g_n(t) = f_n(t), \quad \text{if} \quad t \in [0, 1],$$

and $g_n(t) = 1$ for $t \in [1, 2]$. Then $\{g_n\}$ converges weakly star to $\chi_{[1, 2]}$, the characteristic function of $[1, 2]$ in $L^\infty([0, 2])$, but not in the $L^1([0, 2])$-norm. Note that $\|g_n\|_\infty = \|\chi_{[1,2]}\|_\infty = 1$. Thus, $\{g_n\}$ is a sequence in $L^\infty([0, 2])$ which converges weakly star to $\chi_{[1,2]}$ and for which $\lim_{n \to \infty} \|g_n\|_\infty = \|\chi_{[1,2]}\|_\infty$, but $\lim_{n \to \infty} (\|g_n - \chi_{[1,2]}\|_L^1 + \|g_n\|_\infty - \|\chi_{[1,2]}\|_\infty) \neq 0$.

**Remark 4.14.** Given a direct problem formulated in $L^\infty$, is it worth to formulate the inverse problem in $L^\infty$ or is it more appropriate to formulate the problem in $L^2$, where we can get $L^2$-approximations? $L^2$-approximations are quite advantageous because in the $L^\infty$ case, one might not even get convergence with respect to the $L^1$-norm, as demonstrated by the above counterexample. However, it is sometimes desirable that the regularized solutions are guaranteed to belong also to the $L^\infty$-space. Moreover, the extremal behavior of the solution can be evaluated by $L^\infty$-regularization since the convergence of the $L^\infty$-norm of the regularized solutions to the $L^\infty$-norm of the true solution is achieved. Also, in some inverse problems high interest is given to estimating a linear functional of the solution rather than the solution—according to the mollifier idea in [4], which corresponds to the weak-star approximation of our $L^\infty$-regularization results.

5. The inverse ground water filtration problem

We consider the problem of recovering the diffusion coefficient in

$$-(au_x)_x = f \quad \text{in} \quad \Omega,$$

$$u(0) = 0 = u(1),$$

with $f \in L^2[0, 1]$. The operator $F$ is defined as the parameter-to-solution mapping:

$$F : D(F) := \{a \in L^\infty[0, 1] : a(x) \geq c > 0, \text{a.e.}\} \to L^2[0, 1],$$

$$a \mapsto F(a) := u(a),$$

where $u(a)$ is the unique solution of the above equation and $c$ is a constant. More details about this ill-posed problem can be found in [25] and [11, chapter 1]. Note that $D(F)$ is a subset of the interior of the nonnegative cone of $L^\infty[0, 1]$. In fact $L^\infty$ is the natural function space when formulating this problem. However, to the best of our knowledge, the previous literature dealing with the problem from the regularization viewpoint has usually considered $D(F)$ in $H^1$ which is embedded in $L^\infty$, mainly due to the Hilbert space setting which is enforced by using $H^1$. The operator $F$ is Fréchet differentiable from $L^\infty$ to $L^2$—see [25]. Note that
\(D(F)\) is closed and convex with respect to \(L^2\), so it is weakly closed in \(L^2\). This implies that \(D(F)\) is weakly\(^*\) closed in \(L^\infty\). Similarly one can argue that the operator \(F\) is sequentially weakly\(^*\)–weakly closed.

In the following we are going to employ the approximation operators \(F_m\) as in [31]. Let \(Y_m\) be the space of linear splines on a uniform grid of \(m + 1\) points in \([0, 1]\), which vanish at 0 and 1. By using the variational formulation, let \(u_m(a) \in Y_m\) be the unique solution of

\[
(a(u_m)x, v_x)_{L^2} = (f, v)_{L^2} \quad \text{for all} \quad v \in Y_m.
\]

The operators \(F_m\) are defined as

\[
F_m : D(F) := \{ a \in L^\infty[0, 1] : a(x) \geq c > 0, \text{a.e.} \} \rightarrow L^2[0, 1],
\]

\[
a \mapsto F_m(a) := u_m(a).
\]

Then (6) holds for \(\rho_m = m^{-2}\), cf [12, theorems 3.2.2, 3.2.5]:

\[
\|F_m(a) - F(a)\|_{L^2} = \|u_m(a) - u(a)\|_{L^2} = O(\|a\|_{L^\infty} \cdot m^{-2}).
\]

Then by choosing the discretization of \(L^\infty[0, 1]\) as in section 4, one obtains the convergence results of the discretized regularization method as in example 4.13.

Acknowledgments

The authors thank Professor J B Cooper (Kepler University, Linz) for providing the counterexample in section 4. They are grateful to S Pereverzev (Radon Institute, Linz) for helpful discussions, and to A Rieder (Karlsruhe University) and V Vasin (Institute of Mathematics and Mechanics, Ekaterinburg) for pointing out interesting references. The authors also thank a referee for their constructive comments. Moreover, they acknowledge the support by the Austrian Science Fund (FWF) within the National Research Networks Industrial Geometry, project 9203-N12, and Photoacoustic Imaging in Biology and Medicine, project S10505-N20 (CP and OS) and by an Elise Richter scholarship, project V82-N18 (ER).

References

[1] Adams R 1975 *Sobolev Spaces* (New York: Academic)
[2] Adams R and Fournier J 2003 *Sobolev Spaces* (New York: Academic)
[3] Ambrosio L, Fusco N and Pallara D 2000 *Functions of Bounded Variation and Free Discontinuity Problems* (Oxford Mathematical Monographs) (New York: Clarendon, Oxford University Press)
[4] Anderssen R S 1986 The linear functional strategy for improperly posed problems Inverse Problems ed J Cannon and U Hormung (Basel: Birkhäuser) pp 11–30
[5] Borwein J and Lewis A 1991 Convergence of best entropy estimates SIAM J. Optim. 1 191–205
[6] Burger M and Osher S 2004 Convergence rates of convex variational regularization Inverse Problems 20 1411–21
[7] Burger M, Resmerita E and He L 2007 Error estimation for Bregman iterations and inverse scale space methods in image restoration Computing 81 109–35 (Special Issue on Industrial Geometry)
[8] Bělý P and Luskin M 2003 Approximation by piecewise constant functions in a BV metric Math. Models Methods Appl. Sci. 13 373–93 (Dedicated to Jim Douglas, Jr., on the occasion of his 75th birthday)
[9] Casas E, Kunisch K and Pola C 1999 Regularization by functions of bounded variation and applications to image enhancement Appl. Math. Optim. 40 229–57
[10] Chambolle A 2004 An result for special functions with bounded deformation J. Math. Pures Appl. 83 929–54
[11] Chavent G 2009 *Non Linear Least Squares For Inverse Problems*. Scientific Computation (Berlin: Springer)
[12] Ciarlet P 2002 *The Finite Element Method for Elliptic Problems* (SIAM Classics in Applied Mathematics Number 40) (Philadelphia: SIAM)
[13] DeVore R A 1998 Nonlinear approximation Acta Numer. 7 51–150
[14] Duchi J and Singer Y 2009 Efficient online and batch learning using forward backward splitting. J. Mach. Learn. Res. 10 2899–934

[15] Engl H, Neubauer A and Kunisch K 1989 Convergence rates for tikhonov regularisation of non-linear ill-posed problems Inverse Problems 5 523–40

[16] Fabre C, Puel J-P and Zuazua E 1993 Contrôlabilité approchée de l’équation de la chaleur linéaire avec des contrôles de norme $l^\infty$ minimale. (Approximate controllability for the linear heat equation with controls of minimal $l^\infty$ norm) C. R. Acad. Sci., Paris I 316 679–84

[17] Fitzpatrick B G and Keeling S L 1997 On approximation in total variation penalization for image reconstruction and inverse problems Numer. Funct. Anal. Optim. 18 941–58

[18] Frick K and Scherzer O 2010 Regularization of ill-posed linear equations by the non-stationary augmented Lagrangian method J. Integral Equ. Appl. 22 215–57

[19] Grasmair M, Haltmeier M and Scherzer O 2008 Sparse regularization with $F^p$ penalty term Inverse Problems 24 055020

[20] Gugat M 2002 Analytic solutions of $L^\infty$ optimal control problems for the wave equation J. Optim. Theory Appl. 114 397–421

[21] Gugat M and Leugering G 2002 Regularization of $L^\infty$-optimal control problems for distributed parameter systems Comput. Optim. Appl. 22 151–92

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

[23] Holmes R B 1975 Geometric Functional Analysis and Its Applications (Berlin: Springer)

[24] Holohan A M and Meinsma G 1999 On inverse problems in 2-norm and infinity-norm controller synthesis Proc. 38th CDC Conf. (Phoenix, Arizona) pp 3627–32

[25] Ito K and Kunisch K 1994 On the injectivity and linearization of the coefficient-to-solution mapping for elliptic boundary value problems J. Math. Anal. Appl. 188 1040–66

[26] Leonov A S 2005 Regularization of ill-posed problems in Sobolev space $w^{1,1}$ J. Inverse Ill-Posed Problems 13 595–619

[27] Leonov A S 2007 On the total-variation convergence of regularizing algorithms for ill-posed problems Zh. Vychisl. Mat. Mat. Fiz. 47 732–47

[28] Matthias H, Strang G and Christiansen E 1979 The saddle point of a differential program Energy Methods in Finite Element Analysis (Chichester: Wiley) pp 309–18

[29] Megginson R E 1998 An Introduction to Banach Space Theory (New York: Springer)

[30] Neubauer A 1989 Tikhonov regularization for non-linear ill-posed problems: optimal convergence rates and finite-dimensional approximation Inverse Problems 5 541–57

[31] Neubauer A and Scherzer O 1990 Finite-dimensional approximation of Tikhonov regularized solutions of nonlinear ill-posed problems Numer. Funct. Anal. Optim. 11 85–99

[32] Plato R and Vanikko G 1990 On the regularization of projection methods for solving ill-posed problems Numer. Math. 63 79–89

[33] Qinian J 1999 Applications of the modified discrepancy principle to Tikhonov regularization of nonlinear ill-posed problems SIAM J. Numer. Anal. 36 475–90

[34] Resmerita E 2005 Regularization of ill-posed problems in Banach spaces: Convergence rates Inverse Problems 21 1303–14

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

[36] Scherzer O, Grasmair M, Grossauer H, Haltmeier M and Lenzen F 2008 Variational Methods in Imaging (Applied Mathematical Sciences) (Berlin: Springer)

[37] Suquet P-M 1978 Existence et régularité des solutions des équations de la plasticité C. R. Acad. Sci. Paris A-B 286 A1201–4

[38] Temam R 1983 Problèmes Mathématiques en Plasticité, volume 12 of Méthodes Mathématiques de l’Informatique (Mathematical Methods of Information Science) (Monteouge: Gauthier-Villars)

[39] Temam R and Strang G 1980 Functions of bounded deformation Arch. Ration. Mech. Anal. 75 7–21

[40] Tikhonov A N, Leonov A S and Yagola A G 1998 Nonlinear Ill-Posed Problems (London: Chapman & Hall)

[41] Zou H and Yuan M 2008 The F-infinity-norm support vector machine Stat. Sin. 18 379–98