ANALYSIS OF BOUNDED VARIATION PENALTY METHODS FOR ILL-POSED PROBLEMS

R. ACAR * AND C.R. VOGEL †

Abstract. This paper presents an abstract analysis of bounded variation (BV) methods for ill-posed operator equations $Au = z$. Let

$$T(u) \stackrel{\text{def}}{=} \|Au - z\|^2 + \alpha J(u),$$

where the penalty, or "regularization", parameter $\alpha > 0$ and the functional $J(u)$ is the BV norm or seminorm of u, also known as the total variation of u. Under mild restrictions on the operator A and the functional $J(u)$, it is shown that the functional $T(u)$ has a unique minimizer which is stable with respect to certain perturbations in the data z, the operator A, the parameter α, and the functional $J(u)$. In addition, convergence results are obtained which apply when these perturbations vanish and the regularization parameter is chosen appropriately.

Key words. total variation, bounded variation, regularization, compact operator equations, ill-posed problems

AMS(MOS) subject classifications. 49J10, 49J27, 49J45, 49K40, 65R30

1. Introduction.

Consider the equation

$$Au = z \tag{1.1}$$

where A is a linear operator from $L^p(\Omega)$ into a Hilbert space $\mathcal{Z}$ containing the data vector z. Of particular interest is the case where problem (1.1) is ill-posed, e.g., when A is compact. The data z and the operator A are assumed to be inexact, and approximate solutions to (1.1) are desired which minimize the undesirable effects of perturbations in z and A. Of practical interest are Fredholm first kind integral operators

$$Au(x) = \int_\Omega k(x,y)\, u(y)\, dy. \tag{1.2}$$

For example, certain blurring effects in image processing may be described by convolution operators, in which case $k(x,y) = k(x-y)$. See [9].

Problem (1.1) is ill-posed and discretizations of it are highly ill-conditioned. To deal with ill-posedness, one should apply methods which impose stability while retaining certain desired features of the solution. Historically, these have come to be known as "regularization" methods, since stability was typically obtained by imposing smoothness constraints on the approximate solutions. In many applications, particularly in image processing (see [9], [3]) and parameter identification (see [5]), a serious shortcoming of standard regularization methods is that they do not allow discontinuous solutions. This difficulty can be overcome

* Department of Mathematics, Eastern Montana College, Billings, MT 59101. e-mail: racar@viking.emcmt.edu

† Department of Mathematical Sciences, Montana State University, Bozeman, MT 59717, e-mail: vogel@math.montana.edu. Research was supported in part by the NSF under Grant DMS-9303222.

by achieving stability with the requirement that the solution be of bounded variation rather than smooth. For problem (1.1), this requirement may be enforced in several ways. One approach is to solve a constrained minimization problem like

$$\min_u J(u) \quad \text{subject to} \quad \|Au - z\|^2 = \sigma^2, \tag{1.3}$$

where σ^2 is an estimate of the size of the error in the data and $J(u)$ is the Bounded Variation (BV) norm or seminorm of u (see [4] for definitions and background). This is essentially the approach taken by Rudin and Osher et al [8], [9]. A closely related approach is taken by Dobson and Santosa [3], where the constraint in (1.3) is replaced by the operator equation (1.1). In the application considered in [3], discretizations of equation (1.1) are severely underdetermined. An earlier reference on the use of BV functions in a parameter identification setting (where a constraint on $J(u)$ is imposed instead) is the paper by Gutman ([5]).

Another closely related approach, which is taken by Santosa and Symes [10] and Vogel [13], is to solve the unconstrained minimization problem

$$\min_u \|Au - z\|^2 + \alpha J(u). \tag{1.4}$$

This can be viewed as a penalty method approach to solving the constrained minimization problem (1.3). Here the penalty parameter $\alpha > 0$ controls the tradeoff between goodness of fit to the data, as measured by $\|Au - z\|^2$, and the variability of the approximate solution, as measured by $J(u)$. This penalty approach is widely known in the inverse problems community as Tikhonov regularization, although the term "regularization" seems inappropriate here since discontinuous minimizers may be obtained.

As in [13], a slightly more general penalty functional than the BV seminorm will be considered. For sufficiently smooth u, define

$$J_\beta(u) = \int_\Omega \sqrt{|\nabla u|^2 + \beta}\, dx, \tag{1.5}$$

where $\beta \geq 0$. When $\beta = 0$, this reduces to the usual BV seminorm (the BV norm is given by $\|u\|_{BV} = \|u\|_{L^1(\Omega)} + J_0(u)$). $J_0(u)$ is also commonly referred to as the total variation of u. A variational definition of J_β is presented below which extends (1.5) to (nonsmooth) functions u. Taking $\beta > 0$ offers certain computational advantages, such as differentiability of the functional J_β when $\nabla u = 0$.

A number of important questions arise in the implementation of numerical methods to solve the minimization problem (1.4). For instance,

- Is problem (1.4) really well-posed?
- In what function space does the solution to (1.4) lie, and what norm is appropriate to measure convergence? These questions are of more than academic interest, since they should influence the choice of approximation schemes and the selection of stopping criteria. For instance, the analysis below shows that the choice of L^2 to measure convergence in an iterative solution of (1.4) may be inappropriate if the solution is a function of 2 or more (spatial) variables.
- What is the effect of taking small $\beta > 0$ in (1.5) rather than taking $\beta = 0$?

- As perturbations in the data z and the operator A vanish (say, as discrete approximations become more accurate), what conditions on the regularization parameter α are necessary in order to obtain convergence to an underlying exact solution (to an unperturbed problem)?

The goal of this paper is to provide qualitative answers to these questions. The analysis here is substantially different from that of Lions, Osher and Rudin presented in [7], which deals with artificial time-evolution algorithms to solve the Euler-Lagrange equations (obtained from first order necessary conditions) for a constrained problem similar to (1.3). This paper deals with properties of the minimization problems and not with the algorithms used to solve these problems.

This paper is organized as follows: Section 2 contains an overview of functions of bounded variation. Most of the results in this section are standard extensions to $L^p(\Omega)$ for $p > 1$ of results found in Giusti [4]. Included in this section is a variational definition of J_β and a discussion of important properties such as convexity, semicontinuity, and compactness which are associated with it. In Section 3, several abstract theorems are presented which guarantee the well-posedness of unconstrained minimization problems. Theorem 3.1 is a standard result yielding existence and uniqueness of solutions of problems of the form (1.4). Theorem 3.2 is an analogue of the standard result due to Tikhonov [12] concerning continuity of the inverse map for an injective continuous function restricted to a compact subset of a topological space. This theorem yields continuous dependence with respect to the data, the operator, and the parameter α in problem (1.4) with $J(u) = \|u\|_{BV}$. Section 4 deals with minimization problems in which the BV norm is replace by J_β as the penalty term. Section 5 deals with convergence of minimizers to an underlying exact solution as perturbations in the data and the operator vanish. Results in sections 3-5 (in particular, Theorems 3.1 and 5.1) are similar to those obtained by Seidman and Vogel [11] for ill-posed nonlinear operator equations in a reflexive Banach space setting. The stronger results (in particular, stability in Theorem 3.2) obtained in this paper rely on the linearity of the operator A and the specific function spaces that are dealt with.

2. Definitions and Preliminary Results. Let Ω be a bounded convex region in R^d, $d = 1, 2$, or 3, whose boundary $\partial\Omega$ is Lipschitz continuous. Let $|\mathbf{x}| = \sqrt{\sum_{i=1}^d x_i^2}$ denote the Euclidean norm on R^d. Denote the norm on the Banach spaces $L^p(\Omega)$ by $\|\cdot\|_{L^p(\Omega)}$, $1 \le p \le \infty$. Let $|\Omega|$ denote the (Lebesgue) measure of Ω, and unless otherwise specified, let χ_S denote the indicator function for a set $S \subset \Omega$.

As in [4], define the BV seminorm, or total variation,

$$J_0(u) \stackrel{\text{def}}{=} \sup_{\mathbf{v}\in\mathcal{V}} \int_\Omega (-u\, div\, \mathbf{v})\; dx, \tag{2.1}$$

where the set of test functions

$$\mathcal{V} = \{\mathbf{v} \in C_0^1(\Omega; R^d)\ :\ |\mathbf{v}(x)| \le 1 \text{ for all } x \in \Omega\}. \tag{2.2}$$

If $u \in C^1(\Omega)$, one can show using integration by parts that

$$J_0(u) = \int_\Omega |\nabla u|\, dx. \tag{2.3}$$

By a standard denseness argument, this also applies for u in the Sobolev space $W^{1,1}(\Omega)$. The space of functions of bounded variation on Ω is defined by

$$BV(\Omega) = \{u \in L^1(\Omega) : J_0(u) < \infty\}. \tag{2.4}$$

The BV norm is given by

$$\|u\|_{BV} = \|u\|_{L^1(\Omega)} + J_0(u). \tag{2.5}$$

$BV(\Omega)$ is complete, and hence a Banach space, with respect to this norm. The Sobolev space $W^{1,1}(\Omega)$ is a proper subset of $BV(\Omega)$, as is shown by the example in [4, p. 4]. Note that for Ω bounded, $L^p(\Omega) \subset L^1(\Omega)$ for $p > 1$. From the definition, $BV(\Omega) \subset L^1(\Omega)$. It is shown below that $BV(\Omega) \subset L^p(\Omega)$ for $1 \le p \le d/(d-1)$.

Next, define an extension of (1.5) which is analogous to (2.1). Identifying the convex functional $f(\mathbf{x}) = \sqrt{|\mathbf{x}|^2 + \beta}$ with its second conjugate, or Fenchel transform (see [2, p. 289]),

$$\sqrt{|\mathbf{x}|^2 + \beta} = \sup\{\mathbf{x} \cdot \mathbf{y} + \sqrt{\beta\,(1 - |\mathbf{y}|^2)} \ : \ \mathbf{y} \in R^d, \ |\mathbf{y}| \le 1\}, \tag{2.6}$$

the supremum being attained for $\mathbf{y} = \mathbf{x}/\sqrt{|\mathbf{x}|^2 + \beta}$. Motivated by this and (2.1), define

$$J_\beta(u) \stackrel{\text{def}}{=} sup_{\mathbf{v} \in \mathcal{V}} \int_\Omega \left(-u \, div \, \mathbf{v} + \sqrt{\beta(1 - |\mathbf{v}(x)|^2)} \right) \, dx. \tag{2.7}$$

Note that for $\beta > 0$, J_β is not a seminorm.

Theorem 2.1. *If $u \in W^{1,1}(\Omega)$, then (1.5) holds.*

Proof. Since $C^1(\Omega)$ is dense in $W^{1,1}(\Omega)$, it suffices to show (1.5) for $u \in C^1(\Omega)$. In this case, for any $\mathbf{v} \in \mathcal{V}$, Green's Theorem (integration by parts) gives

$$\begin{aligned} \int_\Omega \left(-u \, div \, \mathbf{v} + \sqrt{\beta(1 - |\mathbf{v}|^2)} \right) \, dx &= \int_\Omega \left(\nabla u \cdot \mathbf{v} + \sqrt{\beta(1 - |\mathbf{v}|^2)} \right) \, dx \\ &\le \int_\Omega \sqrt{|\nabla u|^2 + \beta} \, dx. \end{aligned} \tag{2.8}$$

The inequality above follows from (2.6). Consequently, $J_\beta(u) \le \int_\Omega \sqrt{|\nabla u|^2 + \beta} \, dx$. To show the reverse inequality, take $\overline{\mathbf{v}} = \nabla u / \sqrt{|\nabla u|^2 + \beta}$, and observe that

$$\int_\Omega \left(\nabla u \cdot \overline{\mathbf{v}} + \sqrt{\beta(1 - |\overline{\mathbf{v}}|^2)} \right) \, dx = \int_\Omega \sqrt{|\nabla u|^2 + \beta} \, dx,$$

and $\overline{\mathbf{v}} \in C(\Omega; R^d)$ with $|\overline{\mathbf{v}}(x)| < 1$ for all $x \in \Omega$. By multiplying $\overline{\mathbf{v}}$ by a suitable characteristic function compactly supported in Ω and then mollifying, one can obtain $\mathbf{v} \in \mathcal{V} \cap C_0^\infty(\Omega)$ for which the left hand side of (2.8) is arbitrarily close to $\int_\Omega \sqrt{|\nabla u|^2 + \beta} \, dx$. □

The next Theorem shows that both J_0 and J_β have $BV(\Omega)$ for their effective domain, and that J_0 is the pointwise limit of J_β.

Theorem 2.2. *(i) For any $\beta > 0$ and $u \in L^1(\Omega)$, $J_0(u) < \infty$ if and only if $J_\beta(u) < \infty$; (ii) For any $u \in BV(\Omega)$,*

$$\lim_{\beta \to 0} J_\beta(u) = J_0(u). \tag{2.9}$$

Proof. For any $\mathbf{v} \in \mathcal{V}$ and $u \in L^1(\Omega)$,

$$\begin{aligned}\int_\Omega (-u\, div\, \mathbf{v})\, dx &\leq \int_\Omega \left(-u\, div\, \mathbf{v} + \sqrt{\beta(1-|\mathbf{v}|^2)}\right) dx \\ &\leq \int_\Omega \left(-u\, div\, \mathbf{v} + \sqrt{\beta}\right) dx.\end{aligned}$$

Taking the sup over $\mathbf{v} \in \mathcal{V}$,

$$J_0(u) \leq J_\beta(u) \leq J_0(u) + \sqrt{\beta}\, |\Omega|. \tag{2.10}$$

The results follow from the boundedness of Ω. $\square$

THEOREM 2.3. *For any $\beta \geq 0$, J_β is weakly lower semicontinuous with respect to the L^p topology for $1 \leq p < \infty$.*

Proof. Let $u_n \rightharpoonup \overline{u}$ (weak convergence in $L^p(\Omega)$). For any $\mathbf{v} \in \mathcal{V}$, $div\, \mathbf{v} \in C(\Omega)$, and hence,

$$\begin{aligned}\int_\Omega \left((-\overline{u}\, div\, \mathbf{v}) + \sqrt{\beta(1-|\mathbf{v}|^2)}\right) dx &= \lim_{n\to\infty} \int_\Omega \left((-u_n\, div\, \mathbf{v}) + \sqrt{\beta(1-|\mathbf{v}|^2)}\right) dx \\ &= \liminf_{n\to\infty} \int_\Omega \left((-u_n\, div\, \mathbf{v}) + \sqrt{\beta(1-|\mathbf{v}|^2)}\right) dx \\ &\leq \liminf_{n\to\infty} J_\beta(u_n).\end{aligned}$$

Taking the supremum over $\mathbf{v} \in \mathcal{V}$ gives $J_\beta(\overline{u}) \leq \liminf_{n\to\infty} J_\beta(u_n)$. $\square$

THEOREM 2.4. *For any $\beta \geq 0$, J_β is convex.*

Proof. Let $0 \leq \gamma \leq 1$ and $u_1, u_2 \in L^p(\Omega)$. For any $\mathbf{v} \in \mathcal{V}$,

$$\begin{aligned}&\int_\Omega \left(-(\gamma u_1 + (1-\gamma)u_2) div\, \mathbf{v} + \sqrt{\beta(1-|\mathbf{v}|^2)}\right) dx \\ =\ & \gamma \int_\Omega \left((-u_1 div\, \mathbf{v}) + \sqrt{\beta(1-|\mathbf{v}|^2)}\right) dx + (1-\gamma) \int_\Omega \left((-u_2 div\, \mathbf{v}) + \sqrt{\beta(1-|\mathbf{v}|^2)}\right) dx \\ \leq\ & \gamma J_\beta(u_1) + (1-\gamma) J_\beta(u_2).\end{aligned}$$

Taking the supremum in the top line over $\mathbf{v} \in \mathcal{V}$ gives the convexity of J_β. $\square$

A set of functions $\mathcal{S}$ is defined to be BV-bounded if there exists a constant $B > 0$ for which $\|u\|_{BV} \leq B$ for all $u \in \mathcal{S}$. The relative compactness of BV-bounded sets in $L^p(\Omega)$ follows from the next lemma (see [1] and [4, p. 14]):

LEMMA 2.5. *If $u \in BV(\Omega)$, then there exists a sequence $\{u_n\}$ in $C^\infty(\Omega)$ such that $\lim \|u_n - u\|_{L^p(\Omega)} = 0$ and $\lim J_0(u_n) = J_0(u)$.*

THEOREM 2.6. *Let $\mathcal{S}$ be a BV-bounded set of functions. Then $\mathcal{S}$ is relatively compact in $L^p(\Omega)$ for $1 \leq p < \frac{d}{d-1}$. $\mathcal{S}$ is bounded, and hence relatively weakly compact for dimensions $d \geq 2$, in $L^p(\Omega)$ for $p = \frac{d}{d-1}$.*

Proof. See [4, p. 17 and p. 24]. Note that $d/(d-1)$ is the Sobolev conjugate of 1 in dimension d, the Sobolev conjugate of p, where $1 \leq p < d$, being defined by $1/p^* = 1/p - 1/d$. For $1 \leq p < d/(d-1)$, the Rellich-Kondrachov compact embedding theorem holds. A sequence u_n in $\mathcal{S}$ may then be approximated by a sequence of functions $\tilde{u}_n$ in $C^\infty(\Omega)$, themselves uniformly bounded in $BV(\Omega)$ and in $L^p(\Omega)$, so that their sequence must have a subsequence converging in $L^p(\Omega)$ to some u. By semicontinuity of J_0 and Lemma 2.5, $u \in BV(\Omega)$ and is the limit (in L^p) of a subsequence extracted from u_n.

For $p = d/(d-1)$, one can similarly use Lemma 2.5 to extend to BV-functions the Poincaré-Wirtinger inequality: if

$$\mu = \frac{1}{|\Omega|}\int_\Omega u(x)\,dx,$$

then there exists C such that

$$\|u-\mu\|_{L^p(\Omega)} \le C\,J_0(u-\mu) = C\,J_0(u). \tag{2.11}$$

Hence, if, say, $\|u\|_{BV} \le M$, then $J_0(u-\mu)$ is also bounded by M, and, by the Poincaré-Wirtinger inequality, $\|u-\mu\|_{L^p} \le CM$. Consequently,

$$\begin{aligned} \|u\|_{L^p(\Omega)} &\le \|\mu\chi_\Omega\|_{L^p(\Omega)} + \|u-\mu\|_{L^p(\Omega)} \\ &\le |\mu|\,|\Omega|^{1/p} + CM \\ &\le \|u\|_{L^1(\Omega)}|\Omega|^{1/p-1} + CM \\ &\le (|\Omega|^{1/p-1} + C)M \\ &= (|\Omega|^{-1/d} + C)M. \end{aligned}$$

Relative weak compactness in dimensions $d \ge 2$ follows from the Banach-Alaoglu Theorem [6]. □

The following example shows that the above result is sharp.

Example 2.1. *Let $\Omega = \{\mathbf{x} \in R^d : |\mathbf{x}| < 2\}$ and $u_n = n^{d-1}\chi_n$, where*

$$\chi_n(\mathbf{x}) = \begin{cases} 1, & \text{if } |\mathbf{x}| \le \frac{1}{n}, \\ 0, & \text{otherwise.} \end{cases}$$

Let ω_d denote the volume of the unit ball in R^d. Then

$$\|u_n\|_{L^p(\Omega)} = \begin{cases} n^{d-1}\left(\int_\Omega \chi_n\,dx\right)^{1/p} = n^{d-1-d/p}\omega_d^{1/p}, & \text{if } 1 \le p < \infty, \\ n^{d-1}, & \text{if } p = \infty \end{cases} \tag{2.12}$$

Hence the sequence $\{u_n\}$ is unbounded in $L^p(\Omega)$ whenever $p > \frac{d}{d-1}$. Similarly, if $p = \frac{d}{d-1}$, $d > 1$, and $m > n$, then

$$\begin{aligned} \|u_n - u_m\|_{L^p(\Omega)} &\ge n^{d-1}\|\chi_n - \chi_m\|_{L^p(\Omega)} \\ &= \omega_d^{1/p}\left(1 - \left(\frac{n}{m}\right)^{d-1}\right). \end{aligned}$$

On the other hand, if $d = 1$ and $m > n$, then $\|u_n - u_m\|_{L^\infty(\Omega)} = 1$. In either case, the sequence $\{u_n\}$ is bounded but not Cauchy in $L^p(\Omega)$ for $p = \frac{d}{d-1}$.

Now let σ_d denote the area of the unit sphere S_{d-1} in R^d. From (2.12) and (2.5),

$$\|u_n\|_{BV} = \frac{1}{n}\omega_d + \sigma_d. \tag{2.13}$$

Hence, the sequence is BV-bounded but has no convergent subsequence in $L^p(\Omega)$ whenever $p \geq \frac{d}{d-1}$.

Recall that a functional J is strictly convex if

$$J(\gamma u_1 + (1-\gamma)u_2) < \gamma J(u_1) + (1-\gamma)J(u_2), \tag{2.14}$$

whenever $u_1 \neq u_2$ and $0 < \gamma < 1$. The following example shows that J_β fails to be strictly convex on $BV(\Omega)$.

EXAMPLE 2.2. *Take $\Omega = (0,1)$, $u_1 = \chi_{[a,b]}$, $u_2 = \chi_{[c,d]}$, where $0 < a < b < c < d < 1$. For any $\beta \geq 0$, a direct computation shows that $J_\beta(u_1) = J_\beta(u_2) = J_\beta((u_1+u_2)/2) = 2+\sqrt{\beta}$. Since $u_1 \neq u_2$, J_β cannot be strictly convex.*

3. Well-Posedness of Minimization Problems. A problem is said to be well-posed in the sense of Hadamard if (i) it has a solution, (ii) the solution is unique, and (iii) the solution is stable. Let T be a functional defined on $L^p(\Omega)$ with values in the extended reals. Theorems 3.1 and 3.2 below guarantee the well-posedness of the unconstrained minimization problem

$$\min_{u \in L^p(\Omega)} T(u). \tag{3.1}$$

These Theorems are followed by some illustrative examples pertaining to problem (1.4).

In order to use the compactness results of Section 2 while still dealing with unconstrained minimization problems, we introduce the following property: Define T to be BV-coercive if

$$T(u) \to +\infty \quad \text{whenever} \quad \|u\|_{BV} \to +\infty. \tag{3.2}$$

Note that "lower level sets" $\{u \in L^p(\Omega) : T(u) \leq a\}$, where $a \geq 0$, are BV-bounded.

THEOREM 3.1 (EXISTENCE AND UNIQUENESS OF MINIMIZERS). *Suppose that T is BV-coercive. If $1 \leq p < \frac{d}{d-1}$ and T is lower semicontinuous, then problem (3.1) has a solution. If in addition $p = \frac{d}{d-1}$, dimension $d \geq 2$, and T is* weakly *lower semicontinuous, then a solution also exists. In either case, the solution is unique if T is strictly convex.*

Proof. The following argument is standard (see [2]): Let u_n be a minimizing sequence for T; in other words,

$$T(u_n) \to \inf_{u \in L^p(\Omega)} T(u) \stackrel{\text{def}}{=} T_{\min}. \tag{3.3}$$

By hypothesis (3.2), the u_n's are BV-bounded. As a consequence of Theorem 2.6, there exists a subsequence u_{n_j} which converges to some $\bar{u} \in L^p(\Omega)$. Convergence is weak if $p = \frac{d}{d-1}$. By the (weak) lower semicontinuity of T,

$$T(\bar{u}) \leq \liminf T(u_{n_j}) = T_{min}.$$

Uniqueness of minimizers follows immediately from strict convexity. □

Next consider a sequence of perturbed problems

$$\min_{u \in L^p(\Omega)} T_n(u). \tag{3.4}$$

THEOREM 3.2 (STABILITY OF MINIMIZERS). *Assume that* $1 \le p < \frac{d}{d-1}$ *and that* T *and each of the* T_n*'s are BV-coercive, lower semicontinuous, and have a unique minimizer. Assume in addition*

(i) Uniform BV-Coercivity: For any sequence $v_n \in L^p(\Omega)$,

$$\lim T_n(v_n) = +\infty \quad \textit{whenever} \quad \lim \|v_n\|_{BV} = +\infty. \tag{3.5}$$

(ii) Consistency: $T_n \to T$ *uniformly on BV-bounded sets, i.e., given* $B > 0$ *and* $\epsilon > 0$, *there exists* N *such that*

$$|T_n(u) - T(u)| < \epsilon \quad \textit{whenever} \quad n \ge N,\ \|u\|_{BV} \le B. \tag{3.6}$$

Then problem (3.1) is stable with respect to the perturbations (3.4), i.e., if $\bar{u}$ *minimizes* T *and* u_n *minimizes* T_n*, then*

$$\|u_n - \bar{u}\|_{L^p(\Omega)} \to 0. \tag{3.7}$$

If $p = d/(d-1)$, $d \ge 2$, *and one replaces the lower semicontinuity assumption on* T *and each* T_n *by weak lower semicontinuity, then convergence is weak:*

$$u_n - \bar{u} \rightharpoonup 0. \tag{3.8}$$

Proof. Note that $T_n(u_n) \le T_n(\bar{u})$. From this and (3.6),

$$\liminf T_n(u_n) \le \limsup T_n(u_n) \le T(\bar{u}) < \infty, \tag{3.9}$$

and hence by (3.5), the u_n's are BV-bounded. Now suppose (3.7) (or (3.8) if $p = d/(d-1)$) does not hold. By Theorem 2.6 there exists a subsequence u_{n_j} which converges in $L^p(\Omega)$ (weak L^p) to some $\hat{u} \ne \bar{u}$. By the (weak) lower semicontinuity of T, (3.9), and (3.6),

$$\begin{aligned} T(\hat{u}) &\le \liminf T(u_{n_j}) \\ &= \lim (T(u_{n_j}) - T_{n_j}(u_{n_j})) + \liminf T_{n_j}(u_{n_j}) \\ &\le T(\bar{u}). \end{aligned}$$

But this contradicts the uniqueness of the minimizer $\bar{u}$ of T. ☐

EXAMPLE 3.1 (EXISTENCE-UNIQUENESS). *Consider the problem of minimizing*

$$T(u) = \|Au - z\|^2_{\mathcal{Z}} + \alpha \|u\|_{BV} \tag{3.10}$$

for $u \in L^p(\Omega)$, *where the restrictions on* p *in Theorem 3.1 apply. Here* $\alpha > 0$ *and* $z \in \mathcal{Z}$ *are fixed, and* $A : L^p(\Omega) \to \mathcal{Z}$ *is bounded and linear. Then*

$$\|u\|_{BV} \le \frac{1}{\alpha} T(u), \tag{3.11}$$

and hence, the coercivity condition (3.2) holds. Weak lower semicontinuity of T *follows from the boundedness of* A, *the weak lower semicontinuity of the norms on Banach spaces, and*

Theorem 2.3. By Theorem 2.4, the linearity of A, and convexity of norms, T is convex. By Theorem 3.1 a minimizer exists. T is strictly convex if A is injective, in which case the minimizer is unique.

The following examples deal with stability. In the next three examples, assume again that the restrictions on p of Theorem 3.1 apply.

EXAMPLE 3.2 (PERTURBATIONS IN THE DATA z). *Let*

$$T_n(u) \stackrel{\text{def}}{=} \|Au - z_n\|_{\mathcal{Z}}^2 + \alpha \|u\|_{BV}, \tag{3.12}$$

where $z_n = z + \eta_n$ *and* $\|\eta_n\|_{\mathcal{Z}} \to 0$ *as* $n \to \infty$. *Then*

$$\begin{aligned} |T_n(u) - T(u)| &= \left| \|\eta_n\|_{\mathcal{Z}}^2 + 2\langle Au - z, \eta_n \rangle_{\mathcal{Z}} \right| \\ &\le \|\eta_n\|_{\mathcal{Z}} \left(\|\eta_n\|_{\mathcal{Z}} + 2\|A\|\, \|u\|_{L^p(\Omega)} + 2\|z\|_{\mathcal{Z}}. \right) \end{aligned}$$

Here $\langle \cdot, \cdot \rangle_{\mathcal{Z}}$ *denotes the inner product on the Hilbert space* $\mathcal{Z}$, *and the above inequality follows from Cauchy-Schwarz. Note that if u is BV-bounded, then it is norm bounded in* $L^p(\Omega)$ *by Theorem 2.6, and hence (3.6) holds. (3.5) holds because for each n,*

$$\|u\|_{BV} \le \frac{T_n(u)}{\alpha}. \tag{3.13}$$

EXAMPLE 3.3 (PERTURBATIONS IN THE PENALTY FUNCTIONAL). *Take*

$$T_n(u) \stackrel{\text{def}}{=} \|Au - z\|_{\mathcal{Z}}^2 + \alpha \left(\|u\|_{L^1(\Omega)} + J_{\beta_n}(u) \right), \tag{3.14}$$

where $\beta_n \to 0$. *In this case,*

$$|T_n(u) - T(u)| = \alpha |J_{\beta_n}(u) - J_0(u)| \le \alpha \sqrt{\beta_n}\, |\Omega|. \tag{3.15}$$

The above inequality follows from (2.10). This verifies (3.6). Similarly, (3.5) holds because

$$\frac{1}{\alpha} T_n(u) \ge \|u\|_{L^1(\Omega)} + J_\beta(u) \ge \|u\|_{L^1(\Omega)} + J_0(u) = \|u\|_{BV}. \tag{3.16}$$

EXAMPLE 3.4 (PERTURBATIONS IN THE PENALTY PARAMETER α). *Let*

$$T_n(u) = \|Au - z\|_{\mathcal{Z}}^2 + \alpha_n \|u\|_{BV}, \tag{3.17}$$

where the α_n*'s are bounded below by* $\alpha_{min} > 0$ *and converge to* α. *Stability follows from the facts that*

$$\|u\|_{BV} \le \frac{T_n(u)}{\alpha_{min}}$$

and

$$|T_n(u) - T(u)| \le |\alpha_n - \alpha|\; \|u\|_{BV}.$$

EXAMPLE 3.5 (PERTURBATIONS OF THE OPERATOR A). *Assume $1 \leq p < \frac{d}{d-1}$, and let*

$$T_n(u) \stackrel{\text{def}}{=} \|A_n u - z\|_{\mathcal{Z}}^2 + \alpha \|u\|_{BV}, \tag{3.18}$$

where the A_n's converge strongly (i.e., pointwise) in $L^p(\Omega)$ to A. Note that strong operator convergence is a reasonable assumption. It holds for consistent Galerkin approximations, e.g., Finite Element approximations as the mesh spacing $h \to 0$. Then

$$\begin{aligned} |T_n(u) - T(u)| &= |\, \|A_n u\|_{\mathcal{Z}}^2 - \|Au\|_{\mathcal{Z}}^2 - 2\langle (A_n - A)u, z\rangle_{\mathcal{Z}}| \\ &\leq (\|A_n u\|_{\mathcal{Z}} + \|Au\|_{\mathcal{Z}} + 2\|z\|_{\mathcal{Z}})\, \|(A_n - A)u\|_{\mathcal{Z}}. \end{aligned}$$

Note that pointwise convergence of bounded linear operators becomes uniform on compact sets. Since BV-boundedness implies relative compactness in $L^p(\Omega)$, (3.6) holds. Uniform coercivity (3.5) again holds because of (3.13).

4. Other Penalty Terms. In this section, the BV norm in the penalty term is replaced by the BV seminorm J_0, or more generally, by J_β. Consider the following functional defined on $L^p(\Omega)$,

$$T(u) = \|Au - z\|_{\mathcal{Z}}^2 + \alpha J_\beta(u), \tag{4.1}$$

again taking on values in the extended reals. From a computational standpoint, for positive β the penalty functional $J_\beta(u)$ is Gateaux differentiable with respect to u, and hence much easier to deal with than $\|u\|_{BV}$. However, the analysis becomes much more complicated. Certain conditions on A are clearly needed to guarantee BV-coercivity. For example from (1.5), T cannot be BV-coercive if A annihilates constant functions. Conversely,

LEMMA 4.1. *Assume that $1 \leq p \leq d/(d-1)$, and that A does not annihilate constant functions. Equivalently, since A is linear, assume*

$$A\,\chi_\Omega \neq 0. \tag{4.2}$$

Then T in (4.1) is BV-coercive.

Proof. From the inequalities (2.10), it suffices to consider the case of $\beta = 0$. Any $u \in BV(\Omega)$ has a decomposition

$$u = v + w, \tag{4.3}$$

where

$$w = \left(\frac{\int_\Omega u\, dx}{|\Omega|} \right) \chi_\Omega, \qquad \int_\Omega v\, dx = 0. \tag{4.4}$$

By (2.11) and Hölder's inequality, there exists a positive constant C such that for any p such that $1 \leq p \leq d/(d-1) \stackrel{\text{def}}{=} q$,

$$\begin{aligned} \|v\|_{L^p(\Omega)} &\leq |\Omega|^{1/p - 1/q} \|v\|_{L^q(\Omega)} \\ &\leq (|\Omega| + 1)^{1-1/q} C J_0(v) \\ &= C_1 J_0(v), \end{aligned} \tag{4.5}$$

where $C_1 \stackrel{\text{def}}{=} (|\Omega|+1)^{1/d} C$. Using (4.5) and the decomposition (4.3),

$$\|u\|_{BV} \leq \|w\|_{L^1(\Omega)} + (C_1+1) J_0(v). \tag{4.6}$$

From (4.2), there exists $C_2 > 0$ such that

$$\|Aw\|_{\mathcal{Z}} = C_2 \|w\|_{L^1(\Omega)}. \tag{4.7}$$

On the other hand, the decomposition (4.3) yields

$$\begin{aligned} T(u) &= \|(Av - z) + Aw\|_{\mathcal{Z}}^2 + \alpha J_0(v) \\ &\geq (\|Av - z\|_{\mathcal{Z}} - \|Aw\|_{\mathcal{Z}})^2 + \alpha J_0(v) \\ &\geq \|Aw\|_{\mathcal{Z}} (\|Aw\|_{\mathcal{Z}} - 2\|Av - z\|_{\mathcal{Z}}) + \alpha J_0(v). \end{aligned} \tag{4.8}$$

But by (4.5),

$$\|Av - z\|_{\mathcal{Z}} \leq \|A\|\, C_1\, J_0(v) + \|z\|_{\mathcal{Z}}. \tag{4.9}$$

Combining (4.8) with this and (4.7) yields

$$T(u) \geq C_2 \|w\|_{L^1(\Omega)} \Big(C_2 \|w\|_{L^1(\Omega)} - 2(\|A\|\, C_1\, J_0(v) + \|z\|_{\mathcal{Z}}) \Big) + \alpha J_0(v). \tag{4.10}$$

Now if

$$C_2 \|w\|_{L^1(\Omega)} - 2(\|A\|\, C_1\, J_0(v) + \|z\|_{\mathcal{Z}}) \geq 1, \tag{4.11}$$

then from (4.10)

$$\|w\|_{L^1(\Omega)} \leq \frac{1}{C_2} T(u). \tag{4.12}$$

As a consequence of this and

$$J_0(v) \leq \frac{1}{\alpha} T(u), \tag{4.13}$$

one obtains from (4.6)

$$\|u\|_{BV} \leq \left(\frac{1}{C_2} + \frac{C_1 + 1}{\alpha} \right) T(u). \tag{4.14}$$

But if (4.11) does not hold, then

$$\|w\|_{L^1(\Omega)} < \frac{1 + 2(\|A\|\, C_1\, J_0(v) + \|z\|_{\mathcal{Z}})}{C_2} \tag{4.15}$$

and hence from (4.6) and (4.13),

$$\|u\|_{BV} - \frac{1 + 2\|z\|_{\mathcal{Z}}}{C_2} \leq \left(\frac{2\,\|A\| C_1}{C_2} + C_1 + 1 \right) \frac{1}{\alpha} T(u). \tag{4.16}$$

From (4.14) and (4.16), one obtains BV-coercivity. □

One now obtains the following from Theorem 3.1.

THEOREM 4.2. *Suppose p satisfies the restrictions of Theorem 3.1, $\beta \geq 0$, and A is bounded linear and satisfies (4.2). Then the functional T in (4.1) has a minimizer.*

The following example illustrates that a condition stronger than (4.2) may be necessary to guarantee uniqueness of minimizers of T in (4.1).

EXAMPLE 4.1. *Define $A : L^1(-2,2) \to R^2$ by*

$$[Au]_1 = \int_{-2}^{-1} u(x)\,dx, \quad [Au]_2 = \int_{1}^{2} u(x)\,dx.$$

Let $\mathbf{z} = [z_1, z_2]^T = [-1, 1]^T \in R^2$. Define

$$T_\beta(u) = \sum_{i=1}^{2}([Au]_i - z_i)^2 + J_\beta(u). \tag{4.17}$$

For any $\beta > 0$, the unique minimizer of T_β over $L^1(\Omega)$ is

$$u(x) = \begin{cases} -1, & \text{if } x \leq -1, \\ x, & \text{if } -1 < x < 1, \\ 1, & \text{if } x \geq 1. \end{cases} \tag{4.18}$$

On the other hand, for $\beta = 0$ one obtains a minimizer by defining u on the subinterval $-1 < x < 1$ to be any monotonic increasing function taking on values between -1 and 1.

This next theorem addresses the stability of minimizers to functionals of the form (4.1). Consider perturbed functionals

$$T_n(u) = \|A_n u - z_n\|_{\mathcal{Z}}^2 + \alpha J_\beta(u). \tag{4.19}$$

THEOREM 4.3. *Assume $1 \leq p < \frac{d}{d-1}$, $\|z_n - z\|_{\mathcal{Z}} \to 0$, the A_n's are each bounded linear and converge pointwise to A, and for each n,*

$$\|A_n\, \chi_\Omega\|_{\mathcal{Z}} \geq \gamma > 0. \tag{4.20}$$

Also assume each T_n has a unique minimizer u_n and that T has a unique minimizer $\bar{u}$. Then

$$\|u_n - \bar{u}\|_{L^p(\Omega)} \to 0. \tag{4.21}$$

Proof. It suffices to show that conditions (i) and (ii) of Theorem 3.2 hold. For condition (i) (uniform BV-coercivity), put $u_n = v_n + w_n$ as in (4.3)–(4.4), and repeat the proof of Lemma 4.1. Since $\|A_n w_n\|_{\mathcal{Z}} \geq \gamma \|w_n\|_{L^1(\Omega)}$, letting M be an upper bound on $\|A\|$ and each $\|A_n\|$ (such a bound exists by the Banach-Steinhaus theorem, also known as the uniform boundedness principle), and m be an upper bound on $\|z\|_{\mathcal{Z}}$ and each $\|z_n\|_{\mathcal{Z}}$, one obtains

$$T_n(u_n) \geq \gamma \|w_n\|_{L^1(\Omega)} \Big(\gamma \|w_n\|_{L^1(\Omega)} - 2(M\, C_1\, J_0(v_n) + m)\Big) + \alpha J_0(v_n). \tag{4.22}$$

This yields uniform coercivity as in the proof of Lemma 4.1.

Condition (ii) (consistency) follows as in Example 3.2 and Example 3.5. □

5. Convergence of Minimizers. Assume an exact problem

$$Au = z \tag{5.1}$$

which has a unique solution $u_{exact} \in BV(\Omega)$. Assume a sequence of perturbed problems

$$A_n u = z_n \tag{5.2}$$

having approximate solutions u_n (not necessarily unique) obtained by minimizing the functionals

$$T_n(u) = \|A_n u - z_n\|_{\mathcal{Z}}^2 + \alpha_n \|u\|_{BV}. \tag{5.3}$$

The following theorem provides conditions which guarantee convergence of the u_n's to u_{exact}.

THEOREM 5.1. *Let $1 \le p \le \frac{d}{d-1}$. Suppose $\|z_n - z\|_{\mathcal{Z}} \to 0$, $A_n \to A$ pointwise in $L^p(\Omega)$, and $\alpha_n \to 0$ at a rate for which $\|A_n u_{exact} - z_n\|^2/\alpha_n$ remains bounded. Then $u_n \to u_{exact}$ strongly in $L^p(\Omega)$ if $1 \le p < d/(d-1)$. Convergence is weak in $L^p(\Omega)$ if $p = d/(d-1)$.*

Proof. Note that

$$\begin{aligned} \|A_n u_n - z_n\|_{\mathcal{Z}}^2 &\le T_n(u_n) \\ &\le T_n(u_{exact}) \\ &= \|A_n u_{exact} - z_n\|_{\mathcal{Z}}^2 + \alpha_n \|u_{exact}\|_{BV}. \end{aligned}$$

Thus from the assumption that $\|A_n u_{exact} - z_n\|^2/\alpha_n$ remains bounded and the fact that $\alpha_n \to 0$,

$$\|A_n u_n - z_n\|_{\mathcal{Z}}^2 \to 0. \tag{5.4}$$

Similarly,

$$\begin{aligned} \|u_n\|_{BV} &\le \frac{T_n(u_n)}{\alpha_n} \le \frac{T_n(u_{exact})}{\alpha_n} \\ &= \frac{\|A_n u_{exact} - z_n\|^2}{\alpha_n} + \|u_{exact}\|_{BV}, \end{aligned}$$

and hence, the u_n's are BV-bounded. Suppose they do not converge strongly (weakly, if $p = d/(d-1)$) to u_{exact}. By Theorem 2.6 there is a subsequence u_{n_j} which converges strongly (weakly, respectively) in $L^p(\Omega)$ to some $\hat{u} \neq u_{exact}$. For any $v \in \mathcal{Z}$,

$$\begin{aligned} |\langle A\hat{u} - z, v\rangle_{\mathcal{Z}}| \le\ & |\langle A(\hat{u} - u_{n_j}), v\rangle_{\mathcal{Z}}| + |\langle (A - A_{n_j}) u_{n_j}, v\rangle_{\mathcal{Z}}| \\ & + |\langle A_{n_j} u_{n_j} - z_{n_j}, v\rangle_{\mathcal{Z}}| + |\langle z_{n_j} - z, v\rangle_{\mathcal{Z}}|. \end{aligned} \tag{5.5}$$

The third and fourth terms on the right hand side vanish as $j \to \infty$ because of (5.4) and the assumption $z_n \to z$. The second term also vanishes, since

$$|\langle (A - A_{n_j}) u_{n_j}, v\rangle_{\mathcal{Z}}| \le \|u_{n_j}\|_{L^p(\Omega)} \|(A^* - A_{n_j}^*) v\|_{L^{p}(\Omega)} \to 0$$

by the pointwise convergence of the A_n's (and hence, their adjoints) and the norm boundedness of the u_n's in $L^p(\Omega)$. The first term vanishes as well, taking adjoints and using the (weak) convergence of u_{n_j} to $\hat{u}$. Consequently, $\langle A\hat{u} - z, v\rangle_{\mathcal{Z}} = 0$ for any $v \in \mathcal{Z}$, and hence, $A\hat{u} = z$. But this violates the uniqueness of the solution u_{exact} of (5.1). □

As in the previous section, one can consider instead the functional

$$T_n(u) = \|A_n u - z_n\|_{\mathcal{Z}}^2 + \alpha_n J_\beta(u) \tag{5.6}$$

and obtain the same results as in the previous theorem.

THEOREM 5.2. *In Theorem 5.1, replace T_n by (5.6), and make the same assumptions on A_n, α_n, z_n and p. Assume furthermore that $|A_n\chi_\Omega| \geq \gamma > 0$. Then the conclusions of Theorem 5.1 follow.*

Proof. From the inequalities (2.10) one can assume $\beta = 0$. As in the proof of Theorem 5.1, one obtains that $\|A_n u_n - z_n\|^2 \leq \|A_n u_{\text{exact}} - z_n\|^2 + \alpha_n J_0(u_{\text{exact}})$, which implies (5.4). On the other hand, putting $u_n = v_n + w_n$ and referring again to the proofs of Lemma 4.1 and Theorem 4.3, the present assumptions also imply that (4.22) holds. As in Lemma 4.1, this implies that the u_n are uniformly BV-bounded. The last part of the proof is then the same as that of Theorem 5.1. □

REFERENCES

[1] G. Anzellotti and M. Giaquinta, *Funzioni BV e tracce*, Rend. Sem. Mat. Padova 60 (1978), 1–21.

[2] K. Deimling, *Nonlinear Functional Analysis*, Springer-Verlag, 1980.

[3] D. Dobson and F. Santosa, *An image enhancement technique for electrical impedance tomography*, IMA Preprint Series No. 1145, June 1993, submitted to Inverse Problems.

[4] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, 1984.

[5] S. Gutman, *Identification of discontinuous parameters in flow equations*, SIAM J. Control Optim., vol. 28 (1990), pp. 1049–1060.

[6] V. Hutson and J. S. Pym, *Applications of Functional Analysis and Operator Theory*, Academic Press, 1980.

[7] P.L. Lions, S. Osher, and L. Rudin, *Denoising and Deblurring Algorithms with Constrained Nonlinear PDE's*, submitted to SIAM J. Numer. Analysis.

[8] L.I. Rudin, S. Osher, and E. Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*, Physica D, vol. 60 (1992), pp. 259-268.

[9] L.I. Rudin, S. Osher, and C. Fu, *Total Variation Based Restoration of Noisy, Blurred Images*, preprint, Cognitech, Inc., 2800 - 28th St., Suite 101, Santa Monica, CA 90405, submitted to SIAM J. Numer. Analysis.

[10] F. Santosa and W. Symes, *Reconstruction of blocky impedance profiles from normal-incidence reflection seismographs which are band-limited and miscalibrated*, Wave Motion, vol. 10 (1988), pp. 209-230.

[11] T. I. Seidman and C. R. Vogel, *Well–posedness and convergence of some regularization methods for nonlinear ill–posed problems*, Inverse Problems, vol. 5 (1989), pp. 227-238.

[12] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, Wiley, 1977.

[13] C. R. Vogel, *Total Variation regularization for ill-posed problems*, Department of Mathematical Sciences Technical Report, April 1993, Montana State University.