Analysis of Regularized Total Variation Penalty Methods for Denoising

David Dobson* and Otmar Scherzer†
Department of Mathematics
Texas A & M University
College Station, Texas 77843-3368, USA

Abstract

The problem of recovering images with sharp edges by total variation denoising has recently received considerable attention in image processing. Numerical difficulties in implementing this nonlinear filter technique are partly due to the fact that it involves the stable evaluations of unbounded operators. To overcome that difficulty we propose to approximate the evaluation of the unbounded operator by a stable approximation. A convergence analysis for this regularized approach is presented.

1 Introduction

We consider the penalized least squares problem

$$\min_{u\in L^2(\Omega)} \frac{1}{2}\|u-z\|^2_{L^2(\Omega)} + \alpha J_\beta(u) \tag{1.1}$$

where

$$J_\beta(u) := \int_\Omega \sqrt{|\nabla u|^2+\beta^2}. \tag{1.2}$$

*The work of the first author is partially supported by US AFOSR grant number F49620-95-1-0497.

†The work of the second author is supported by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung, Grant J01088-TEC; O.S. is on leave from Institut für Mathematik, Universität Linz, Altenberger Str. 69, A–4040 Linz, Austria

Following the original work of Rudin, Osher, and Fatemi [10], this problem has received considerable attention recently in image processing as a means of recovering images u with sharp edges from noisy measurements z [1, 2, 11, 5].

Compared to problems involving standard Tikhonov regularization, minimization of (1.1) is quite different to handle since the functional (1.2) involves evaluations of the unbounded operator ∇u. Groetsch (see [7, 8, 9]) proposed a concept (based on a Tikhonov-like method) for approximating values of unbounded operators. The evaluation of a closed, densely defined, unbounded, linear operator $L : D(L) \subseteq H_1 \to H_2$ from a Hilbert space H_1 into a Hilbert space H_2 is performed by approximating Lu (which of course only exists if $u \in D(L)$) by

$$L(I + \gamma L^* L)^{-1} u. \tag{1.3}$$

In this 'stabilized' operator γ denotes a regularization parameter and L^* denotes the adjoint operator of L (with respect to the Hilbert spaces H_1 and H_2). A typical example of an unbounded operator is 'Differentiation'

$$\begin{array}{rcl} L : D(L) \subseteq H^1(\Omega) \subseteq L^2(\Omega) & \to & \left(L^2(\Omega)\right)^d . \\ u & \to & \nabla u \end{array} \tag{1.4}$$

Introducing Groetsch's concept to approximately minimize (1.1), ∇u might be replaced by a 'stable' approximation

$$L_\gamma u := L(I + \gamma L^* L)^{-1} u, \tag{1.5}$$

where L^* denotes the L^2–adjoint of L. This leads then to the minimization problem

$$\min_{u \in L^2(\Omega)} \frac{1}{2} \|u - z\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u) \tag{1.6}$$

where

$$J_{\beta\gamma}(u) := \int_\Omega \sqrt{|L_\gamma u|^2 + \beta^2}. \tag{1.7}$$

Of course one could think of many other different schemes to replace $u \to \nabla u$ by a bounded operator, for example operators based on difference quotients, finite element methods, etc. Actually the general results of this paper will not be based on the specific features of the approximation scheme, but on its approximation properties. Each general result will be followed by an example of an approximation scheme which satisfies the general assumptions. To consider the minimization of (1.6) reflects computational relevance for problem (1.1), since in practice the minimization of (1.1) is often done by approximately evaluating the gradient with a finite difference scheme.

In the following section, we show that solutions to problem (1.6) are stable with respect to perturbations in the data and converge to the "true solution" (uncorrupted by noise) as the noise level is reduced, provided that α, β, and γ are taken to zero at appropriate rates. In Section 3, we show that a fixed point scheme for solving the nonlinear problem (1.6) is convergent for appropriate values of α, β, and γ. Finally in Section 4 we illustrate some of these ideas with some simple numerical experiments.

2 Convergence and Stability

In this section some general results concerning minimization of the functional (1.6) are established. The first result assures that minimization of (1.6) is a well–defined problem, i.e., that the minimum of (1.6) is actually attained. In the proofs presented below we make use of the Fenchel transform (see e.g. [3]) of the convex functional $f(x) = \sqrt{|x|^2 + \beta^2}$, which is

$$\sqrt{|x|^2 + \beta^2} = \sup\{x \cdot y + \beta\sqrt{1 - |y|^2} : y \in \mathbb{R}^d, |y| \leq 1\} \tag{2.1}$$

The Fenchel transform has been already used successfully by Acar and Vogel [1], Dobson and Vogel [5], Chambolle and Lions [2] for the convergence analysis of bounded variation penalty methods.

Throughout this paper the following will be assumed to hold:

- Ω is a bounded domain in $\mathbb{R}^d$ with "sufficiently smooth" boundary,
- L_γ is a bounded linear operator from $L^2(\Omega)$ into $(L^2(\Omega))^d$, for $\gamma > 0$. Note that for now we do not assume the specific form (1.5) for L_γ.

Theorem 2.1 (Existence of a minimizer) *Problem (1.6) attains a minimum. Moreover the solution of the minimization problem is unique.*

Proof: If (1.6) does not attain a minimum, there exists a sequence $\{u_n\}$ such that

$$\frac{1}{2}\|u_n - z\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u_n) \to q := \inf_{u \in L^2(\Omega)} \left\{ \frac{1}{2}\|u - z\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u) \right\} \tag{2.2}$$

but

$$q < \frac{1}{2}\|u - z\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u) \quad \text{for all } u \in L^2(\Omega).$$

Since $\{u_n\}$ is bounded in $L^2(\Omega)$, it has a weakly convergent subsequence $u_m \rightharpoonup \tilde{z}$ in $L^2(\Omega)$. Since L_γ is a bounded linear operator it is also weakly continuous, i.e., $L_\gamma u_m \rightharpoonup L_\gamma \tilde{z}$ in $L^2(\Omega)$. Due to the weak lower semicontinuity of the norm

$$\frac{1}{2}\|\tilde{z} - z\|^2_{L^2(\Omega)} \le \frac{1}{2} \liminf \|u_m - z\|^2_{L^2(\Omega)}. \tag{2.3}$$

Therefore, there exists a subsequence $\{u_l\}$ of $\{u_m\}$ such that

$$\lim_{l \to \infty} \|u_l - z\|^2_{L^2(\Omega)} = \liminf_{m \in \mathbb{N}} \|u_m - z\|^2_{L^2(\Omega)}.$$

By Fenchel duality

$$\int_\Omega \sqrt{|L_\gamma \tilde{z}|^2 + \beta^2} \le \sup_{v \in \mathcal{V}} \int_\Omega \tilde{z} L_\gamma^* v + \beta\sqrt{1 - |v|^2} \equiv J_{\beta\gamma}(\tilde{z})$$

where

$$v \in \mathcal{V} := \{v \in L^2(\Omega; \mathbb{R}^d) : \text{essential supremum of } |v| \le 1\}. \tag{2.4}$$

For all $v \in \mathcal{V}$

$$\begin{aligned} \textstyle\int_\Omega \tilde{z} L_\gamma^* v + \beta\sqrt{1 - |v|^2} &= \textstyle\lim_{l \to \infty} \int_\Omega u_l L_\gamma^* v + \beta\sqrt{1 - |v|^2} \\ &\le \textstyle\limsup_{l \to \infty} \int_\Omega \sqrt{|L_\gamma u_l|^2 + \beta^2} \end{aligned}$$

Consequently,

$$\int_\Omega \sqrt{|L_\gamma \tilde{z}|^2 + \beta^2} \le \limsup_{l \to \infty} \int_\Omega \sqrt{|L_\gamma u_l|^2 + \beta^2}. \tag{2.5}$$

Combination of (2.3) and (2.5) yields

$$\frac{1}{2}\|\tilde{z}-z\|^2_{L^2(\Omega)} + \alpha \int_\Omega \sqrt{|L_\gamma \tilde{z}|^2 + \beta^2} \leq \limsup_{l\to\infty} \left\{ \frac{1}{2}\|u_l - z\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u_l) \right\} \leq q,$$

which contradicts (2.2).

It remains to prove uniqueness of the minimizer of (1.6). Since $\frac{1}{2}\|u-z\|^2_{L^2(\Omega)}$ is strictly convex and $\alpha J_{\beta\gamma}(u)$ is convex, as we will see below, their sum is strictly convex. This implies that the minimum of (1.6) is unique. To prove the convexity of $\alpha J_{\beta\gamma}(u)$ we use again Fenchel duality: For all $v \in \mathcal{V}$

$$\begin{aligned} \int_\Omega (\lambda u_1 + (1-\lambda)u_2) L_\gamma^* v + \beta\sqrt{1-|v|^2} &= \lambda \int_\Omega u_1 L_\gamma^* v + \beta \sqrt{1-|v|^2} + \\ &\quad (1-\lambda)\int_\Omega u_2 L_\gamma^* v + \beta\sqrt{1-|v|^2} \\ &\leq \lambda \int_\Omega \sqrt{|L_\gamma u_1|^2 + \beta^2} + (1-\lambda)\int_\Omega \sqrt{|L_\gamma u_2|^2+\beta^2} \end{aligned}$$

and consequently

$$\int_\Omega \sqrt{|L_\gamma(\lambda u_1 + (1-\lambda)u_2)|^2 + \beta^2} \leq \lambda \int_\Omega \sqrt{|L_\gamma u_1|^2+\beta^2} + (1-\lambda)\int_\Omega \sqrt{|L_\gamma u_2|^2 + \beta^2}.$$

□

Theorem 2.2 (Stability of minimizers) *The minimization of (1.6) is stable with respect to to perturbations in the data, i.e., if $z_k \to z$ in $L^2(\Omega)$ and u_k denotes the solution to (1.6) with z replaced by z_k, then $u_k \to u$ in $L^2(\Omega)$ and $J_{\beta\gamma}(u_k) \to J_{\beta\gamma}(u)$.*

Proof: Let $z_k \to z$ (with respect to $\|.\|_{L^2(\Omega)}$). Then u_k satisfies

$$\frac{1}{2}\|u_k - z_k\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u_k) \leq \frac{1}{2}\|u - z_k\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u) \tag{2.6}$$

for every $u \in L^2(\Omega)$. Thus $\{u_k\}$ is bounded in $L^2(\Omega)$, and therefore has a weakly convergent subsequence $u_m \rightharpoonup \hat{z}$. By the weak lower semicontinuity (2.5) of $J_{\beta\gamma}$,

$$J_{\beta\gamma}(\hat{z}) \leq \limsup J_{\beta\gamma}(u_m). \tag{2.7}$$

By the weak lower semicontinuity of $\|.\|_{L^2(\Omega)}$ we have

$$\|\hat{z}-z\|_{L^2(\Omega)} \leq \liminf \|u_m - z_m\|_{L^2(\Omega)}$$

and therefore by (2.6)

$$\begin{aligned}
\|\hat{z}-z\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(\hat{z}) &\leq \liminf\{\|u_m - z_m\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u_m)\} \\
&\leq \limsup\{\|u_m - z_m\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u_m)\} \\
&\leq \lim_{m\to\infty}\{\|u - z_m\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u)\} \\
&= \|u - z\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u)
\end{aligned}$$

for all $u \in L^2(\Omega)$. This implies that $\hat{z}$ is a minimizer of (1.6) and that

$$\lim_{m\to\infty} \left\{\|u_m - z_m\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u_m)\right\} = \|\hat{z}-z\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(\hat{z}). \tag{2.8}$$

If $\{u_m\}$ did not converge strongly to $\hat{z}$ with respect to $\|.\|_{L^2(\Omega)}$, then

$$c := \limsup \|u_m - z\|_{L^2(\Omega)} > \|\hat{z}-z\|_{L^2(\Omega)}$$

and there exists a subsequence $\{u_n\}$ of $\{u_m\}$ satisfying

$$u_n \rightharpoonup \hat{z}, \quad J_{\beta\gamma}(\hat{z}) \leq \lim_{n\to\infty} J_{\beta\gamma}(u_n), \quad \|u_n - \hat{z}\|_{L^2(\Omega)} \to c. \tag{2.9}$$

As a consequence of (2.8) we obtain

$$\alpha \lim_{n\to\infty} J_{\beta\gamma}(u_n) = \alpha J_{\beta\gamma}(\hat{z}) + \|\hat{z}-z\|^2_{L^2(\Omega)} - c^2 < \alpha J_{\beta\gamma}(\hat{z})$$

which is a contradiction to (2.9). This shows $u_n \to \hat{z}$ in $L^2(\Omega)$ and $\lim_{n\to\infty} J_{\beta\gamma}(u_n) = J_{\beta\gamma}(\hat{z})$. Since each sequence $\{u_m\}$ has a convergent subsequence $\{u_n\}$ which converges to $\hat{z}$ (note that the minimizer of (1.6) is unique) the sequence is itself convergent; and analogously it can be verified $\lim_{n\to\infty} J_{\beta\gamma}(u_m) = J_{\beta\gamma}(u_*)$. □

In the following we denote by δ an estimate for the measurement error, i.e.,

$$\|z_\delta - u_*\|_{L^2(\Omega)} \leq \delta, \tag{2.10}$$

where u_* is the "true solution" of the denoising problem, uncorrupted by measurement noise. We assume that $u_* \in BV(\Omega)$, so

$$\int_\Omega |\nabla u_*| := \sup_{v \in \hat{V}} \int_\Omega u_* \, (\nabla \cdot v) \quad < \infty, \tag{2.11}$$

where

$$v \in \hat{V} := \{v \in C_0^\infty(\Omega; \mathbb{R}^d) : |v| \leq 1\}.$$

Theorem 2.3 (Convergence) *Let $\alpha = \alpha(\delta)$, $\gamma = \gamma(\delta)$, $\beta = \beta(\delta)$ be such that as $\delta \to 0$,*

$$\alpha \to 0, \quad \frac{\delta^2}{\alpha} \to 0, \quad \gamma \to 0, \quad \beta \to 0. \tag{2.12}$$

Assume

$$\int_\Omega |L_\gamma u_*| - \int_\Omega |\nabla u_*| \to 0, \tag{2.13}$$

and that

$$||L_\gamma^* v - (-\nabla . v)||_{L^2(\Omega)} \to 0, \tag{2.14}$$

as $\gamma \to 0$, for all $v \in \hat{V}$. Then $u_\delta := u_{\alpha,\beta,\gamma,\delta}$, the unique solution of (1.6) where z is replaced by z^δ, is convergent to u_* with respect to bounded variation, i.e.,

$$\|u_\delta - u_*\|_{L^2(\Omega)} \to 0, \quad J_{\beta\gamma}(u_\delta) \to J(u_*); \tag{2.15}$$

here we use the notation $J(u_*) = \int_\Omega |\nabla u_*|$.

Proof: u_δ satisfies

$$\begin{aligned} \tfrac{1}{2}\|u_\delta - z_\delta\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u_\delta) &\leq \tfrac{1}{2}\|u_* - z_\delta\|^2_{L^2(\Omega)} + \alpha J_{\beta\gamma}(u_*) \\ &\leq \tfrac{1}{2}\delta^2 + \alpha J_{\beta\gamma}(u_*). \end{aligned} \tag{2.16}$$

From this, for $\delta \to 0$, $\alpha \to 0$,

$$\|u_\delta - u_*\|_{L^2(\Omega)} \leq \|u_\delta - z_\delta\|_{L^2(\Omega)} + \|z_\delta - u_*\|_{L^2(\Omega)} \to 0. \tag{2.17}$$

Moreover, it follows from (2.16)

$$J_{\beta\gamma}(u_\delta) - J(u_*) \leq \frac{1}{2}\frac{\delta^2}{\alpha} + (J_{\beta\gamma}(u_*) - J(u_*)). \tag{2.18}$$

This together with

$$\lim_{\delta\to 0} J_{\beta\gamma}(u_*) - J(u_*) = \lim_{\delta\to 0}\int_\Omega (\sqrt{|L_\gamma u_*|^2} - \sqrt{|\nabla u_*|^2}) \to 0 \text{ (by assumption)}$$

yields

$$\limsup_{\delta\to 0} J_{\beta\gamma}(u_\delta) \le J(u_*). \tag{2.19}$$

Since by (2.14) for all $v \in \hat{V}$

$$\begin{aligned}
\int_\Omega u_* \nabla . v &= \lim_{\delta\to 0}\int_\Omega u_\delta \nabla . v \\
&= \lim_{\delta\to 0}\int_\Omega u_\delta (L_\gamma^* v - (-\nabla .)v) - \lim_{\delta\to 0}\int_\Omega u_\delta L_\gamma^* v \\
&= -\lim_{\delta\to 0}\int_\Omega L_\gamma u_\delta . v + \beta\sqrt{1-|v|^2} \\
&\le \limsup_{\delta\to 0} J_{\beta\gamma}(u_\delta).
\end{aligned}$$

Consequently (by 2.11) $J(u_*) \le \limsup_{\delta\to 0} J_{\beta\gamma}(u^\delta)$ and by (2.19)

$$J(u_*) = \limsup_{\delta\to 0} J_{\beta\gamma}(u^\delta);$$

Since for any $\delta_n \to 0$, $J_{\beta\gamma}(u^{\delta_n})$ has a subsequence, which converges to $J(u_*)$, the sequence is itself convergent. □

For functions of bounded variation (2.15) is a typical convergence result. Usually one does not get stronger convergence results (like $J(u_\delta - u_*) \to 0$) when dealing with functions of bounded variations (cf. [6]).

Example 2.4 *Here we verify that conditions (2.13)–(2.14) hold for certain finite difference approximations and for the approximation (1.5). In the case of the finite difference schemes, we assume $\Omega = [0,1]$ for simplicity.*

Difference schemes: All calculations presented here are carried for the **forward difference scheme**. *Similar calculations can be done for the* **backward difference**

scheme *and the central difference scheme. Let the* **forward difference operator** *be defined by*

$$L_h^d u := \begin{cases} \dfrac{u(x+h)-u(x)}{h} & \text{for } x \in \Omega_h^+, \\ 0 & \text{for } x \notin \Omega_h^+, \end{cases}$$

where $\Omega_h^+ = [0, 1-h]$. *The adjoint operator of* L_h^d *is given by*

$$L_h^{d*} u := \begin{cases} \dfrac{u(x-h)-u(x)}{h} & \text{for } x \in \Omega_h^-, \\ 0 & \text{for } x \notin \Omega_h^-, \end{cases}$$

where $\Omega_h^- = [h, 1]$.

L_h^d *and* L_h^{d*} *satisfy*

1. L_h^d *is a bounded linear operator and*

$$\|L_h^d\|_{L^2(\Omega)\to L^2(\Omega)} \leq \frac{2}{h}.$$

2. *For every* $v \in \hat{V}$,

$$\begin{aligned} \|L_h^{d*} v - (-v')\|_{L^2(\Omega)} &\leq \sqrt{\int_h^1 \left(\tfrac{v(x-h)-v(x)}{h} - (-v'(x))\right)^2 dx + \int_0^h (v'(x))^2\, dx} \\ &\leq \sqrt{C_1^2 h + C_2^2 h} \\ &\leq O(\sqrt{h}) \to 0 \text{ as } h \to 0; \end{aligned}$$

   here $C_1 = \sup_{t\in[0,1]} |v''(t)|$, $C_2 = \sup_{t\in[0,1]} |v'(t)|$. *Consequently (2.14) holds.*

3. *Since*

$$\begin{aligned} \left| \int_\Omega |L_h^d u_*| - |u_*{}'| \right| &\leq \int_\Omega |L_h^d u_* - u_*{}'| \\ &\leq \int_{[0,1-h[} |L_h^d u_* - u_*{}'| + \int_{[1-h,1[} |u_*{}'| \\ &\leq \sqrt{\int_{[0,1-h[} (L_h^d u_* - u_*{}')^2} + \sqrt{\int_{[1-h,1[} |u_*{}'|^2}. \end{aligned}$$

   The right hand side converges to 0 (as $h \to 0$*) if e.g.,* $u_* \in C^1[0,1]$. *Thus (2.13) is satisfied, if* $u_* \in C^1[0,1]$.

A mean value finite difference scheme: All calculations presented here are carried for the **mean value forward difference scheme** *but all calculations can be done in a similar way for the* **mean value backward difference scheme** *and the* **mean value central difference scheme**. *Let the* **mean value forward difference operator** *be defined by*

$$L_h^m u := \begin{cases} \dfrac{v(x+h)-v(x)}{h} & \textit{for } x \in \Omega_h^+, \\ \qquad 0 & \textit{for } x \notin \Omega_h^+, \end{cases} \tag{2.20}$$

where

$$v(x) = v_i := \frac{1}{h}\int_{(i-1)h}^{ih} v(t)\,dt \textit{ for } x \in [(i-1)h, ih]\,, \quad i = 1,...,n, \quad h = \frac{1}{n}.$$

Since $L_h^m = L_h^d \circ R_h^m$, *where*

$$R_h^m v(x) := \frac{1}{h}\int_{(i-1)h}^{ih} v(t)\,dt \textit{ for } x \in [(i-1)h, ih], \tag{2.21}$$

the adjoint can be represented in the following way

$$L_h^{m*} = R_h^{m*} \circ L_h^{d*}.$$

Due to the fact that

$$\begin{aligned} \int_0^1 \int_{[x]}^{[x]+h} v(t)\,dt\; w(x)\,dx &= \sum_{j=1}^{n} \int_{(j-1)h}^{jh} \int_{(j-1)h}^{jh} v(t)\,dt\; w(x)\,dx \\ &= \int_0^1 v(t) \int_{[x]}^{[x]+h} w(x)\,dx\,dt; \end{aligned}$$

here $[x] := (i-1)h$ *for* $(i-1)h \le x < ih$, $i = 1,...,n$. *This shows that* R_h^m *is selfadjoint with respect to the inner product in* $L^2[0,1]$. *Therefore,* $L_h^{m*} = R_n^m \circ L_h^{d*}$, *i.e.,*

$$L_h^{m*} v := \begin{cases} \frac{u(x-h)-u(x)}{h} & \textit{for } x \in \Omega_h^-, \\ \qquad 0 & \textit{for } x \notin \Omega_h^-, \end{cases}$$

$$u(x) = u_i := \frac{1}{h}\int_{(i-1)h}^{ih} v(t)\,dt \textit{ for } x \in [(i-1)h, ih]\,, \quad i = 1,...,n, \quad h = \frac{1}{n}.$$

L_h^m *and* L_h^{m*} *satisfy*

1. L_h^m *is a bounded linear operator and*

$$\|L_h^m\|_{L^2(\Omega)\to L^2(\Omega)} \le \frac{2}{h}$$

to see this note that

$$\|R_h^m v\|_{L^2(\Omega)}^2 = \frac{1}{h^2}\sum_{j=1}^{n}\int_{(j-1)h}^{jh}\left(\int_{(j-1)h}^{jh} v(t)\,dt\right)^2 \le \sum_{j=1}^{n}\int_{(j-1)h}^{jh} v^2(t)\,dt.$$

2. *For every* $v \in \hat{V}$,

$$\begin{array}{rcl} \|L_h^{m*}v - (-v')\|_{L^2(\Omega)} & \le & \sqrt{\int_h^1 \left(\frac{R_h^m v(x-h) - R_h^m v(x)}{h} - (-v')(x)\right)^2 dx + \int_0^h (v'(x))^2\,dx} \\ & \le & \sqrt{C_1^2 h + C_2^2 h} \\ & \le & O(\sqrt{h}) \to 0 \text{ as } h \to 0; \end{array}$$

here $C_1 = \sup_{t\in[0,1]}|v''(t)|$, $C_2 = \sup_{t\in[0,1]}|v'(t)|$. *Consequently (2.14) holds.*

3. *Since*

$$\begin{array}{rcl} \left|\int_\Omega |L_h^m u_*| - |{u_*}'|\right| & \le & \int_\Omega |L_h^m u_* - {u_*}'| \\ & \le & \int_{[0,1-h[} |L_h^m u_* - {u_*}'| + \int_{[1-h,1[} |{u_*}'| \\ & \le & \sqrt{\int_{[0,1-h[} (L_h^m u_* - {u_*}')^2 + \sqrt{\int_{[1-h,1[} |{u_*}'|^2}}, \end{array}$$

one can prove analogously as in 2. that for $u_* \in C^2[0,1]$ *the right hand side converges to* 0 *(as* $h \to 0$*). Thus (2.13) is satisfied, if* $u_* \in C^2[0,1]$.

The approximation (1.5): Let L *be the differential operator*

$$\begin{array}{rcl} L : D(L) := H_0^1(\Omega) \subseteq L^2(\Omega) & \to & (L^2(\Omega))^d. \\ u & \to & \nabla u \end{array}$$

The adjoint of L *is defined by*

$$\begin{array}{rcl} L^* : D(L^*) := \{v \in (L^2(\Omega))^d : \nabla\cdot v \in L^2(\Omega)\} & \to & L^2(\Omega). \\ v & \to & -\nabla\cdot v \end{array}$$

Then the operators L_γ defined in (1.5) and its adjoint

$$L_\gamma^* = (I + \gamma L^* L)^{-1} L^*$$

satisfy:

1. $\|L_\gamma\|_{L^2(\Omega)\to L^2(\Omega)} \le \gamma^{-1}\sqrt{\textit{meas}\ \Omega}$. *This follows from elliptic estimates.*
2. *For every $v \in \hat{V}$, also $v \in D(L^*)$ and therefore $w = L^* v \in L^2(\Omega)$. By spectral estimates (see eg. [12])*

$$\begin{aligned}
\|L_\gamma^* v - (-v')\|_{L^2(\Omega)}^2 &= \|(I + \gamma L^* L)^{-1} w - w\|_{L^2(\Omega)}^2 \\
&= \int_0^\infty \frac{\gamma^2 \lambda^2}{(1+\gamma\lambda)^2} d\|E_\lambda w\|_{L^2(\Omega)}^2 \\
&\to 0 \ \textit{as}\ \gamma \to 0.
\end{aligned}$$

   which shows (2.14); in the last estimate E_λ denotes the spectral representation of the operator $L^* L$.

3. *Analogously to above it can be shown that for $u_* \in D(L)$*

$$\int_\Omega |L_\gamma u_*| - |u_*'| \le \sqrt{\int_\Omega (L_\gamma u_* - u_*)^2} \to 0.$$

   Thus (2.13) is satisfied as long as $u \in D(L) = H_0^1(\Omega)$.

Similar estimates hold for different boundary conditions.

3 A Fixed Point Iteration for Minimizing the Penalty Functional

In this section we establish a *convergence* result for a fixed point iteration to minimize the functional (1.6). The basic fixed point scheme we consider was introduced by Vogel and Oman [11] in the case of the non-regularized penalty functional (1.2). Our analysis modifies their scheme to the case of regularized penalty functionals (1.7), where L_γ is any operator satisfying certain properties. A convergence analysis for the non-regularized iteration was carried out in [5], but with somewhat weaker results than can be obtained for regularized formulations. Chambolle and Lions [2] used different techniques for proving convergence of an iterative relaxation technique for solving the unregularized least squares problem (1.1).

The minimizer of (1.6) satisfies the first order necessary optimality condition:

$$u + \alpha L_\gamma^* \left(\frac{L_\gamma u}{\sqrt{|L_\gamma u|^2 + \beta^2}} \right) = z. \tag{3.1}$$

In terms of

$$L_{\alpha\beta\gamma}(u)v := \frac{\alpha}{\beta} L_\gamma^* \left(\frac{1}{\sqrt{\left|\frac{L_\gamma u}{\beta}\right|^2 + 1}} L_\gamma v \right). \tag{3.2}$$

the fixed point equation (3.1) can be rewritten as

$$(I + L_{\alpha\beta\gamma}(u))u = z. \tag{3.3}$$

In this section the fixed point iteration

$$u^{n+1} = F(u^n) := (I + L_{\alpha\beta\gamma}(u^n))^{-1} z \tag{3.4}$$

is considered for the solution of (3.3).

In order to prove local convergence of (3.4) some Gateaux derivatives are calculated. Let $q(t) := \alpha\sqrt{t^2 + \beta^2}$. Then

$$q'(t) = \frac{\alpha}{\beta} \frac{t}{\left(1 + \left(\frac{t}{\beta}\right)^2\right)^{(1/2)}}, \quad q''(t) = \frac{\alpha}{\beta} \frac{1}{\left(1 + \left(\frac{t}{\beta}\right)^2\right)^{(3/2)}}, \quad q'''(t) = -3\frac{\alpha}{\beta^3} \frac{t}{\left(1 + \left(\frac{t}{\beta}\right)^2\right)^{(5/2)}}.$$

Hence the first, second, and third Gateaux derivatives of

$$f(u) := \frac{1}{2}\|u - z\|^2 + \alpha J_{\beta\gamma}(u)$$

in direction v are

$$\begin{aligned}\langle f'(u), v\rangle &:= \frac{d\tilde{f}}{d\tau}(\tau; u, v)\Big|_{\tau=0} \\ &= \int_\Omega (u-z)v \;+\; \frac{\alpha}{\beta}\int_\Omega \frac{L_\gamma u . L_\gamma v}{\left(\left|\frac{L_\gamma u}{\beta}\right|^2 + 1\right)^{(1/2)}} \\ &=: \langle (1 + L_{\alpha\beta\gamma}(u))u - z, v\rangle\,,\end{aligned} \tag{3.5}$$

$$\begin{aligned}\langle f''(u)v, v\rangle &:= \frac{d^2\tilde{f}}{d\tau^2}(\tau; u, v)\Big|_{\tau=0} \\ &= \int_\Omega v^2 \;+\; \frac{\alpha}{\beta}\int_\Omega \frac{|L_\gamma v|^2}{\left(\left|\frac{L_\gamma u}{\beta}\right|^2 + 1\right)^{(3/2)}} \\ &=: \langle ((1 + M_{\alpha\beta\gamma}(u))v, v\rangle\,,\end{aligned} \tag{3.6}$$

and

$$\begin{aligned}\langle f'''(u)(v,v), v\rangle &:= \frac{d^3\tilde{f}}{d\tau^3}(\tau; u, v)\Big|_{\tau=0} \\ &= -\frac{3\alpha}{\beta^2}\int_\Omega \frac{|L_\gamma v|^2 \frac{L_\gamma u}{\beta} . L_\gamma v}{\left(\left|\frac{L_\gamma u}{\beta}\right|^2 + 1\right)^{(5/2)}};\end{aligned} \tag{3.7}$$

here we used $\tilde{f}(\tau; u, v) = f(u + \tau v)$.

Thus the gradient and Hessian of f are given, respectively, by

$$g(u) := (I + L_{\alpha\beta\gamma}(u))u - z \tag{3.8}$$

and

$$H(u) := I + M_{\alpha\beta\gamma}(u) = I + L_{\alpha\beta\gamma}(u) + L'_{\alpha\beta\gamma}(u)u, \tag{3.9}$$

where

$$\langle L_{\alpha\beta\gamma}(u)v, v\rangle = \frac{\alpha}{\beta}\int_\Omega \frac{1}{\left(\left|\frac{L_\gamma u}{\beta}\right|^2 + 1\right)^{(1/2)}} |L_\gamma v|^2, \tag{3.10}$$

$$\left\langle (L'_{\alpha\beta\gamma}(u)u)v, v \right\rangle = -\frac{\alpha}{\beta^3} \int_\Omega \frac{|L_\gamma u|^2}{\left(\left|\frac{L_\gamma u}{\beta}\right|^2 + 1\right)^{(3/2)}} |L_\gamma v|^2. \tag{3.11}$$

The fixed point iteration can be written

$$\begin{aligned} u^{n+1} &= (I + L_{\alpha\beta\gamma}(u^n))^{-1} z \\ &= u^n - (I + L_{\alpha\beta\gamma}(u^n))^{-1} g(u^n), \quad n = 0, 1, 2, ... \end{aligned} \tag{3.12}$$

Thus

$$g(u^n) = -(I + L_{\alpha\beta\gamma}(u^n)) d^n, \quad d^n := u^{n+1} - u^n. \tag{3.13}$$

The following properties are immediate consequences of (3.12)–(3.13) and the fact that $L_{\alpha\beta\gamma}(u^n)$ is positive semidefinite:

Lemma 3.1 *For $n = 0, 1, 2, 3, ...$,*

$$\langle g(u^n), d^n \rangle_{L^2(\Omega)} \le -\|d^n\|^2_{L^2(\Omega)}, \quad \|d^n\|_{L^2(\Omega)} \le \|g(u^n)\|_{L^2(\Omega)}. \tag{3.14}$$

For $n = 1, 2, 3, ...$,

$$\begin{aligned} \|u^n\|_{L^2(\Omega)} &\le \|z\|_{L^2(\Omega)}, \\ \|g(u^n)\|_{L^2(\Omega)} &\le (1 + \|L_{\alpha\beta\gamma}(u^n)\|_{L^2(\Omega)\to L^2(\Omega)}) \|d^n\|_{L^2(\Omega)}. \end{aligned} \tag{3.15}$$

Property (3.14) implies that d^n is a descent direction for f at u^n, while (3.15) shows (for fixed α, β, γ) that the iterates are bounded. In the following a proof of convergence of (3.4) is presented – there is a special emphasis on the constants depending on parameters α, β, γ to make evident their interplay, which is of considerable interest for the numerical implementation of the iterative scheme.

Lemma 3.2 *If $\|L_\gamma\|_{L^2(\Omega)\to L^3(\Omega)}$ is bounded, then*

$$f(u^{n+1}) - f(u^n) \le \frac{1}{2} \langle g(u^n), d^n \rangle + r(d^n), \tag{3.16}$$

where

$$\|r(d^n)\|_{L^2(\Omega)} \le C_{\alpha\beta\gamma} \|d^n\|^3_{L^2(\Omega)}. \tag{3.17}$$

and

$$C_{\alpha\beta\gamma} = \frac{1}{4} \left(\frac{4}{5}\right)^{(5/2)} \frac{\alpha}{\beta^2} \|L_\gamma\|^3_{L^2(\Omega)\to L^3(\Omega)}.$$

Proof: By Taylor's Theorem

$$
\begin{aligned}
f(u^{n+1}) - f(u^n) &= \langle g(u^n), d^n \rangle + \frac{1}{2} \langle H(u^n) d^n, d^n \rangle + r(d^n) \\
&= \langle g(u^n), d^n \rangle + \frac{1}{2} \langle (I + L_{\alpha\beta\gamma}(u^n)) d^n, d^n \rangle + \\
&\quad \frac{1}{2} \left\langle (L'_{\alpha\beta\gamma}(u^n) u^n) d^n, d^n \right\rangle + r(d^n).
\end{aligned}
$$

Due to the fact that $L'_{\alpha\beta\gamma}(u^n)u^n$ is negative semidefinite

$$
f(u^{n+1}) - f(u^n) \leq \frac{1}{2} \langle g(u^n), d^n \rangle + r(d^n).
$$

Since by Taylor series expansion

$$
\|r(d^n)\|_{L^2(\Omega)} \leq \frac{1}{6} \max_{t \in [0,1]} \left| \frac{d^3 \tilde{f}}{d\tau^3}(\tau, u^n + t d^n, d^n)|_{\tau=0} \right|
$$

and since

$$
\frac{\left| \frac{L_\gamma u}{\beta} \right|}{\left(\left| \frac{L_\gamma u}{\beta} \right|^2 + 1 \right)^{(5/2)}} \leq \frac{1}{2} \left(\frac{4}{5} \right)^{(5/2)}
$$

it follows from (3.7) that

$$
\|r(d^n)\|_{L^2(\Omega)} \leq \frac{1}{4} \left(\frac{4}{5} \right)^{(5/2)} \frac{\alpha}{\beta^2} \int_\Omega |L_\gamma d^n|^3 = C_{\alpha\beta\gamma} \|d^n\|^3_{L^2(\Omega)},
$$

which proves the lemma. □

The crucial assumption in the proof of Lemma 3.2 is that $\|L_\gamma\|_{L^2(\Omega) \to L^3(\Omega)}$ is bounded; it guarantees that the third derivative of f is well-defined. The following examples shows that indeed approximation schemes can be constructed which satisfy this assumption and the general assumptions of Sections 2.

Example 3.3 *We consider the* **mean value forward difference scheme** *of Example 2.4, but this time for the sake of completeness of the paper its multidimensional analog is considered. For simplicity we assume that the domain* Ω *is the unit interval in* $\mathbb{R}^d$*. This* *unit interval is divided into $m = n^d$, $n = 1/h$ equal, separate parts. Then a continuous mapping*

$$\begin{aligned} \hat{I} : L^2(\Omega) &\rightarrow L^2(\Omega), \\ u &\rightarrow v \end{aligned}$$

where $v = v_i$ in the i-th subinterval Ω_i, and v_i is the mean value of u in the i-th subinterval, i.e.,

$$v_i = \frac{1}{h^d} \int_{\Omega_i} u(x)\, dx.$$

One direction of the gradient of u is approximated by

$$R_h(u) = \frac{v(x+\hat{h}) - v(x)}{h}, \text{ where } \hat{h} = (0, ...0, h, 0, ..., 0)^t$$

For $x \in \Omega_i$

$$\begin{aligned} |v(x)| &= |v_i| \\ &= h^{-d} \left| \textstyle\int_{\Omega_i} u(x) dx \right| \\ &\leq h^{-d/2} \sqrt{\textstyle\int_{\Omega_i} u^2(x) dx} \\ &\leq h^{-d/2} \|u\|_{L^2(\Omega)}. \end{aligned}$$

This especially means

$$|v(x)|_{L^\infty(\Omega)} \leq h^{-d/2} \|u\|_{L^2(\Omega)}$$

and moreover

$$|R_h u(x)|_{L^\infty(\Omega_h^+)} \leq 2h^{-d/2-1} \|u\|_{L^2(\Omega)},$$

where

$$\Omega_h^+ = \{x : x + he_i \in \Omega, \text{ for the orthonormal basis } e_i = (0, ..., 0, 1, 0...0)^t\}.$$

As approximation for the gradient ∇u we use

$$L_h(u)(x) = \begin{cases} \left(\frac{v(x+he_i) - v(x)}{h} \right)_{i=1,\ldots,d} & x \in \Omega_h^+ \\ 0 & x \in \Omega - \Omega_h^+ \end{cases} \tag{3.18}$$

and consequently

$$|L_h u(x)|_{L^\infty(\Omega)} \leq 2\sqrt{d} h^{-d/2-1} \|u\|_{L^2(\Omega)},$$

from which the boundedness of $\|L_h\|_{L^2(\Omega) \rightarrow L^3(\Omega)}$ can be easily deduced.

With the previous lemma the main result of this section can be proven:

Theorem 3.4 *Let α, β, γ satisfy*

$$C^0_{\alpha\beta\gamma} := C_{\alpha\beta\gamma}\left(2\|z\|_{L^2(\Omega)} + \alpha\|L_\gamma\|_{L^2(\Omega)\to L^2(\Omega)}\sqrt{\textit{meas }\Omega}\right) \leq \psi, \tag{3.19}$$

where $0 < \psi < \frac{1}{2}$ *and*

$$\frac{1}{2} - C^1_{\alpha\beta\gamma} - 2\|z\|_{L^2(\Omega)}C_{\alpha\beta\gamma} > 0, \quad C^1_{\alpha\beta\gamma} = \frac{1}{3^{(3/2)}}\frac{\alpha}{\beta^2}\|L_\gamma\|^2_{L^2(\Omega)\to L^2(\Omega)}. \tag{3.20}$$

Then the sequence $u^{n+1} = F(u^n)$, where (F is as in (3.4), with $L_{\alpha,\beta,\gamma}$ as defined in (3.3)) is strongly convergent (with respect to $\|.\|_{L^2(\Omega)}$) to the minimizer of (1.6).

Proof: Using

$$\frac{\left|\frac{L_\gamma u^n}{\beta}\right|}{\sqrt{\left|\frac{L_\gamma u^n}{\beta}\right|^2 + 1}} \leq 1$$

in (3.2) yields

$$\|L_{\alpha\beta\gamma}(u^n)u^n\|_{L^2(\Omega)} \leq \alpha\|L_\gamma\|_{L^2(\Omega)\to L^2(\Omega)}\sqrt{\text{meas }\Omega}. \tag{3.21}$$

From (3.21), (3.14) and (3.15) follows for $n = 1, 2, 3...$

$$\begin{aligned}
\|d^n\|_{L^2(\Omega)} &\leq \|g(u^n)\|_{L^2(\Omega)} \\
&\leq \|(I + L_{\alpha\beta\gamma}(u^n))u^n - z\|_{L^2(\Omega)} \\
&\leq \|u^n\|_{L^2(\Omega)} + \|L_{\alpha\beta\gamma}(u^n)u^n\|_{L^2(\Omega)} + \|z\|_{L^2(\Omega)} \\
&\leq 2\|z\|_{L^2(\Omega)} + \alpha\|L_\gamma\|_{L^2(\Omega)\to L^2(\Omega)}\sqrt{\text{meas }\Omega},
\end{aligned} \tag{3.22}$$

and consequently from (3.19)

$$\begin{aligned}
C_{\alpha\beta\gamma}\|d^n\|_{L^2(\Omega)} &\leq C_{\alpha\beta\gamma}\left(2\|z\|_{L^2(\Omega)} + \alpha\|L_\gamma\|_{L^2(\Omega)\to L^2(\Omega)}\sqrt{\text{meas }\Omega}\right) \\
&= C^0_{\alpha\beta\gamma} \\
&\leq \psi.
\end{aligned} \tag{3.23}$$

From (3.14) and Lemma 3.2 it follows

$$f(u^{n+1}) - f(u^n) \leq -\frac{1}{2}\|d^n\|^2_{L^2(\Omega)} + C_{\alpha\beta\gamma}\|d^n\|^3_{L^2(\Omega)}. \tag{3.24}$$

Using (3.23) we find

$$f(u^{n+1}) - f(u^n) \le -\left(\frac{1}{2} - \psi\right) \|d^n\|^2_{L^2(\Omega)}. \tag{3.25}$$

Since $f(u^n)$ is positive and monotonically decreasing, it is also convergent; moreover, from (3.25), $\|d^n\|_{L^2(\Omega)} \to 0$, and consequently from (3.15),

$$\|g(u^n)\|_{L^2(\Omega)} \to 0. \tag{3.26}$$

Due to the boundedness of $\{u^n\}$ in $L^2(\Omega)$ (cf. (3.15)), it has a weakly convergent subsequence:

$$u^k \rightharpoonup \hat{z}, \quad \text{weakly in } L^2(\Omega). \tag{3.27}$$

In the sequel we verify that

$$f(\hat{z}) = \lim_{k\to\infty} f(u^k). \tag{3.28}$$

First we prove

$$f(\hat{z}) \le \lim_{k\to\infty} f(u^k). \tag{3.29}$$

From the weak convergence of $u^k \rightharpoonup \hat{z}$ in $L^2(\Omega)$ and Fenchel duality it follows that for $v \in \mathcal{V}$ (cf. (2.4)),

$$\int_\Omega \hat{z} L_\gamma^* v + \beta\sqrt{1 - |v|^2} = \lim_{k\to\infty} \int_\Omega u^k L_\gamma^* v + \beta\sqrt{1-|v|^2} \le \lim_{k\to\infty} \int_\Omega \sqrt{|L_\gamma u^k|^2 + \beta^2}.$$

Since the right hand side of the last inequality is independent of $v \in \mathcal{V}$

$$\int_\Omega \sqrt{|L_\gamma \hat{z}|^2 + \beta^2} = \sup_{v\in\mathcal{V}} \int_\Omega \hat{z} L_\gamma^* v + \beta\sqrt{1-|v|^2} \le \liminf_{k\to\infty} \int_\Omega \sqrt{|L_\gamma u^k|^2 + \beta^2}. \tag{3.30}$$

From the weak lower semicontinuity of $\|.\|_{L^2(\Omega)}$ it follows

$$\frac{1}{2}\|\hat{z} - z\|^2_{L^2(\Omega)} \le \liminf_{k\in\mathbb{N}} \frac{1}{2}\|u^k - z\|^2_{L^2(\Omega)}. \tag{3.31}$$

Summing up (3.30) and (3.31) yields (3.29).

To verify

$$f(\hat{z}) \ge \lim_{k\to\infty} f(u^k), \tag{3.32}$$

the Taylor series expansion of the functional f around u^k is used (cf. Proof of Lemma 3.2):

$$\begin{aligned} &= \limsup_{k\to\infty} \Big(- \left\langle g(u^k), e^k \right\rangle - \frac{1}{2} \left\langle (I + L_{\alpha\beta\gamma}(u^k))e^k, e^k \right\rangle - \\ &\quad \frac{1}{2} \left\langle (L'_{\alpha\beta\gamma}(u^k)u^k)e^k, e^k \right\rangle - r(e^k) \Big), \end{aligned} \tag{3.33}$$

where $e^k = \hat{z} - u^k$.
From (3.10) and (3.11)

$$- \left\langle L_{\alpha\beta\gamma}(u^k)e^k, e^k \right\rangle \leq 0, \tag{3.34}$$

$$\left\{ \begin{aligned} -\frac{1}{2} \left\langle (L'_{\alpha\beta\gamma}(u^k)u^k)e^k, e^k \right\rangle &= \frac{1}{2}\frac{\alpha}{\beta^2} \int\limits_\Omega \frac{\left|\frac{L_\gamma(u^k)}{\beta}\right|}{\left(\left|\frac{L_\gamma(u^k)}{\beta}\right|^2 + 1\right)^{(3/2)}} |L_\gamma e^k|^2 \\ &\leq \frac{1}{3^{(3/2)}}\frac{\alpha}{\beta^2} \|L_\gamma\|^2_{L^2(\Omega)\to L^2(\Omega)} \|e_k\|^2_{L^2(\Omega)} \\ &=: C^1_{\alpha\beta\gamma} \|e_k\|^2_{L^2(\Omega)}. \end{aligned} \right. \tag{3.35}$$

Analogously to Lemma 3.2 one can verify by Taylor series expansion

$$\|r(e^k)\|_{L^2(\Omega)} \leq C_{\alpha\beta\gamma} \|e^k\|^3_{L^2(\Omega)}. \tag{3.36}$$

From (3.16) it follows

$$\|e_k\|_{L^2(\Omega)} \leq \|\hat{z}\|_{L^2(\Omega)} + \|u^k\|_{L^2(\Omega)} \leq \liminf_{k\in\mathbb{N}} \|u^k\|_{L^2(\Omega)} + \|u^k\|_{L^2(\Omega)} \leq 2\|z\|_{L^2(\Omega)}$$

and consequently from (3.26), (3.29), and (3.33)–(3.36)

$$0 \leq \lim_{k\to\infty} f(u^k) - f(\hat{z}) \leq \lim_{k\to\infty} -\|e^k\|^2 \left(\frac{1}{2} - C^1_{\alpha\beta\gamma} - 2\|z\|_{L^2(\Omega)} C_{\alpha\beta\gamma} \right).$$

Since the right hand side of the last inequality is negative the only choice

$$\lim_{k\to\infty} f(u^k) = f(\hat{z}) \text{ and } u^k \to \hat{z} \text{ in } L^2(\Omega)$$

is left.

To prove that $\hat{z}$ is a minimizer of (1.6), it is proved below that $g(u^k) \to g(\hat{z})$ in $L^2(\Omega)$. Since $\|g(u^k)\|_{L^2(\Omega)} \to 0$ this shows then that $g(\hat{z}) = 0$. Since f is strictly convex, there is only one point which satisfies $g(z) = 0$, and this must be the minimizer of (1.6). To prove that $g(u^k) \to g(\hat{z})$ in $L^2(\Omega)$ it suffices to prove $L_{\alpha\beta\gamma}(u^k)u^k \to L_{\alpha\beta\gamma}(\hat{z})\hat{z}$:

$$\begin{aligned} \|L_{\alpha\beta\gamma}(u^k)u^k - L_{\alpha\beta\gamma}(\hat{z})\hat{z}\|_{L^2(\Omega)} \quad \leq \quad & \left\| \frac{\alpha}{\beta} L_\gamma^* \left(\frac{L_\gamma(u^k) - L_\gamma(\hat{z})}{\sqrt{\left|\frac{L_\gamma(u^k)}{\beta}\right|^2 + 1}} \right) \right\|_{L^2(\Omega)} + \\ & \left\| \frac{\alpha}{\beta} L_\gamma^* L_\gamma \left(\frac{1}{\sqrt{\left|\frac{L_\gamma(u^k)}{\beta}\right|^2 + 1}} - \frac{1}{\sqrt{\left|\frac{L_\gamma(\hat{z})}{\beta}\right|^2 + 1}} \right) \right\|_{L^2(\Omega)} \end{aligned}$$

Both terms on the right hand side of this inequality tend to 0 for $u^k \to \hat{z}$ in $L^2(\Omega)$. Since every subsequence of $\{u^n\}$ has a convergent subsequence, which is convergent to the minimizer of (1.6), $\{u^n\}$ is itself convergent, which completes the proof. □

The assumptions (3.19) and (3.20) can always be fulfilled, by making α small, $\|L_\gamma\|$ small, or β large. Since dealing with an unstable problem, it seems curious that for $\alpha = 0$ the iterative scheme (3.4) is convergent. But note that for $\alpha = 0$ the functional f reduces to $\|u - z\|^2_{L^2(\Omega)}$, which is quadratic and the solution is obvious: $u = z$.

So far it is an open question how to choose α, β, γ in an optimal way to get both

- Optimal convergence of the minimizers of (1.6)
- Fast convergence of the iterative scheme (3.4)

4 Some Numerical Experiments

In this section we illustrate some of the ideas analyzed in this paper, by means of some simple numerical experiments. It should be emphasized that we are not attempting here to explore the efficacy of total variation denoising techniques, as this is beyond the scope of the paper and has been studied elsewhere. Our goal is merely to illustrate the role of the parameters α, β, h in the convergence of the regularized fixed point scheme analyzed in Section 3.

Consider the mean value forward difference scheme on the domain $\Omega = [0,1]$ described in Example 2.4. As mentioned previously, the convergence results in Sections 2 and 3 are valid for this approximation scheme. Numerically, the "regularized" derivative operators L_h^m defined in (2.20) can be realized by representing functions $u \in L^2(\Omega)$ as piecewise constant over n uniform subintervals on $[0,1]$. The discretization parameter $h = 1/n$ then serves the role of the regularization parameter γ in the analysis. In other words, the finite difference operator L_h^m is bounded for all $h > 0$, but its operator norm grows without bound as $h \to 0$.

Applying the fixed point iteration scheme (3.4) with the regularized derivative operators L_h^m, we wish to examine the convergence behavior as the parameters α, β, and h are varied. Recall that Theorem 3.4 guarantees strong convergence of the iterates in $L^2(\Omega)$ for α small, h large, or β large.

For all of our experiments we take the same "noisy" function $z \in L^2(\Omega)$, shown in Figure 4.1. This function was obtained by perturbing a piecewise constant function on a 64-interval grid with normally distributed random noise of standard deviation 0.05. It should be noted that the noise level in this particular example is relatively low.

Applying the regularized fixed point iteration with parameters $\alpha = 12 \times 10^{-4}$, $\beta = 10^{-4}$, and $h = 1/512$, a typical reconstruction u is shown in Figure 4.2. This reconstruction was obtained after 10 iterations of the fixed point scheme. Roughly speaking, choosing larger α results in reconstructions with smaller total variation, and hence less "detail"; larger β makes the penalty functional $J_\beta(u)$ more like arclength, and can result in reconstructions with slightly rounded corners; and larger h decreases the resolution of the reconstruction.

In Figure 4.3 the convergence behavior of the scheme is examined as $\beta = 10^{-4}$ and $h = 1/256$ are held fixed, and α takes on the values $(3, 6, 12, 24) \times 10^{-4}$. Theorem 3.4 guarantees strong L^2 convergence for α sufficiently small. The numerical results indicate that the convergence *rate* actually improves with decreasing α. A similar conclusion was obtained in [5], where convergence results are also obtained for large α.

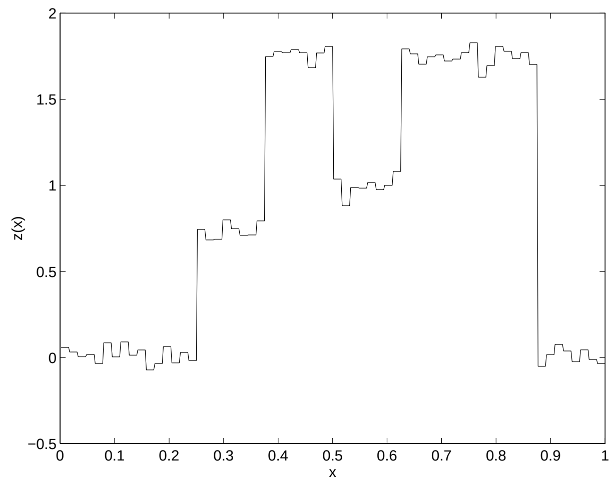

Figure 4.1: "Noisy" function z used in experiments.

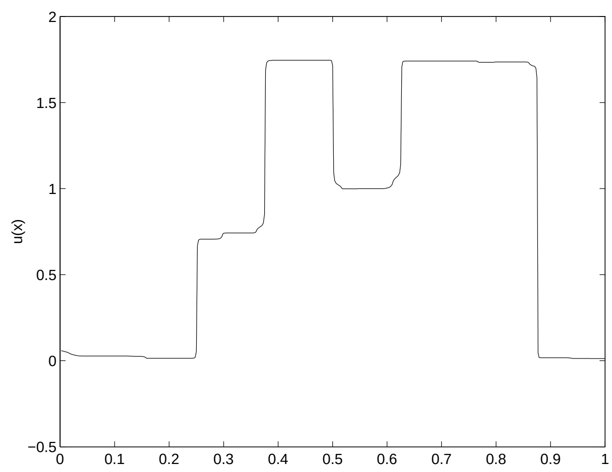

Figure 4.2: "Reconstruction" obtained by applying the regularized fixed point iteration to z.

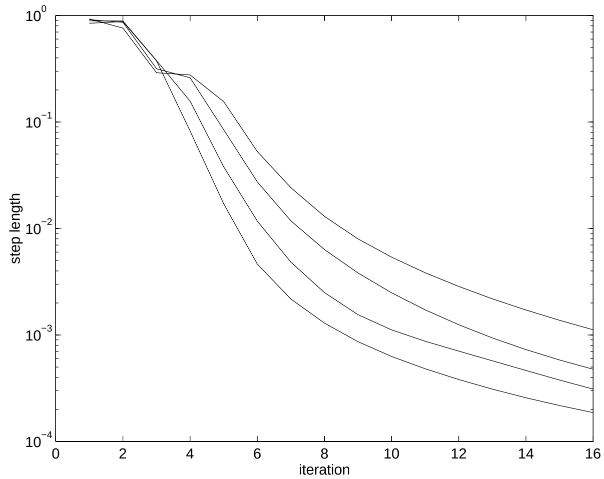

Figure 4.3: Convergence as measured by L^2-norm of difference in successive iterates, as α is varied over four values. Convergence is faster for smaller α.

Convergence behavior appears to be much less sensitive to β. In several experiments β was varied from 1 to 10^{-5} with relatively little effect on the rate of convergence, although larger β were usually somewhat faster.

Finally, we examine convergence as $\alpha = 12 \times 10^{-4}$ and $\beta = 10^{-4}$ are held fixed, and h takes on the four values 1/64, 1/128, 1/256 and 1/512. As shown in Figure 4.4, convergence is slower for smaller "regularization" parameters h. As indicated by the results in Section 3, the difficulty is caused by the growing operator norms of the regularized derivative operators L_h^m as $h \to 0$, and is a reflection of the unboundedness of the operators in the underlying infinite-dimensional problem.

References

[1] R. Acar and C.R. Vogel, Analysis of total variation penalty methods, *Inverse Problems* **10**(1994), 1217 – 1229.

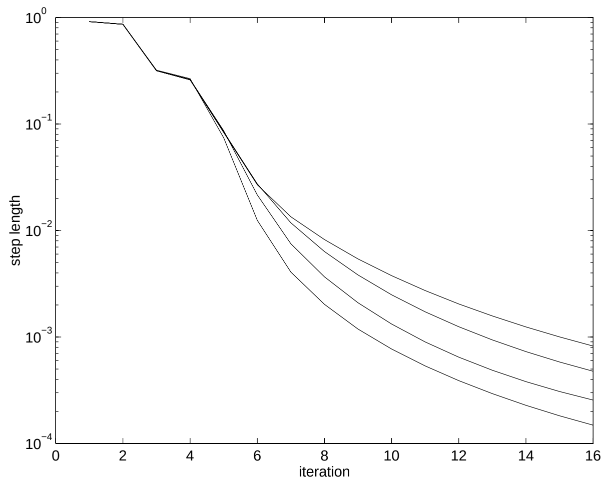

Figure 4.4: Convergence as measured by L^2-norm of difference in successive iterates as h is varied over four values. Convergence slows as h is decreased.

[2] A. CHAMBOLLE AND P.-L. LIONS, Image recovery via total variation minimization and related problems, preprint

[3] K. DEIMLING, Nonlinear Functional Analysis, *Springer Verlag* (1980)

[4] D. DOBSON AND F. SANTOSA, Recovery of blocky images from noisy and blurred data, *SIAM J. Appl. Math.* **56** (1996), to appear.

[5] D. DOBSON AND C.R. VOGEL, Convergence of an iterative method for total variation denoising, *SIAM J. Numer. Anal.*, to appear.

[6] L.C. EVANS AND R.F. GARIEPY, Measure Theory and Fine Properties of Functions, CRC Press, Boca Raton (1992).

[7] C.W. GROETSCH, Spectral methods for linear inverse problems with unbounded operators, *Journal of Approximation Theory* **70** (1992), 16–28

[8] C.W. GROETSCH AND M. HANKE, A general framework for regularized evaluation of unstable operators, preprint, University of Cincinnati, OH

[9] C.W. Groetsch and O. Scherzer The optimal order of convergence for stable evaluation of differential operators, *Electronic Journal of Differential Equations*, **4**(1994), 1–10.

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

[11] C. R. Vogel and M. E. Oman, Iterative Methods for Total Variation Denoising, *SIAM J. Sci. Comput.***17** (1996).

[12] J. Weidmann Linear Operators in Hilbert Spaces, Springer Verlag, New York (1980).