ON THE USE OF DUAL NORMS IN BOUNDED VARIATION TYPE REGULARIZATION

Andreas Obereder
Department of Computer Science, University of Innsbruck
Technikerstr. 25, A–6020 Innsbruck, Austria
andreas.obereder@uibk.ac.at

Stanley Osher
Mathematics Department,
Los Angeles, CA 90095-1555, USA
sjo@math.ucla.edu

Otmar Scherzer
Department of Computer Science, University of Innsbruck
Technikerstr. 25, A–6020 Innsbruck, Austria
otmar.scherzer@uibk.ac.at

Abstract Recently Y. Meyer gave a characterization of the minimizer of the Rudin-Osher-Fatemi functional in terms of the G-norm. In this work we generalize this result to regularization models with second order derivatives of bounded variation. This requires us to define generalized G-norms. We present some numerical experiments to support the theoretical considerations.

Keywords: Bounded variation regularization, contact problems, Rudin-Osher-Fatemi functional.

1. Introduction

In this paper we are concerned with minimization of *bounded variation type regularization functionals* of the form

$$\mathcal{F}(u) := \frac{1}{2}\int (u-f)^2 + \alpha p(u) \quad (\alpha > 0)\,,$$

with

$$p(u) = \|D^k u\| \text{ for } k = 1,2 \text{ and } p(u) = \int |\Delta u|\ .$$

Here $\|D^k u\|$ denotes the *total variation* semi-norm of the $(k-1)$-th derivative of u and $\int |\Delta u|$ denotes the *variation measure* of Δu. The results of this paper can be generalized to higher order derivatives (i.e., for functionals with regularization terms $\|D^k u\|, k = 3, \ldots$) but it is omitted due to the notational complexity.

The special case $k = 1$ is the *Rudin-Osher-Fatemi* (ROF) functional [31] (see also [27, 28]) - the minimizer is called *bounded variation regularized* solution. Since the invention of the ROF-model several results for characterizing properties of the minimizer have been derived. Moreover, in special situations of data f the minimizer could be calculated analytically: Strong & Chan [35] characterized the minimizer of the ROF-model for 1 dimensional data and for spherically symmetric data f (see also Ring [30]). Nikolova [23, 25] analyzed the ROF-model in a discretized setting; in the latter paper also higher order derivatives of bounded variation have been used. Osher & Esedoglu [26] analyzed generalized ROF-models. Y. Meyer [21] gave a characterization of properties of the minimizer of the ROF-functional in terms of the G-norm. These results will be generalized to characterize minimizers of the functional $\mathcal{F}$, i.e., involving regularization functionals with second order derivatives of bounded variation. Motivated from the taut-string algorithm commonly used in statistics (cf. Mammen & Geer [20], Davies & Kovac [10], and Dümbgen & Kovac [12]) and Y. Meyer's [21] characterization of the minimizer we are able to reformulate bounded variation regularization as a bilateral contact problem. The well-known (undesirable) effect of stair casing of the bounded variation regularized solution can be limited by smoothing in contact zones with the tube.

In a discrete setting, for analyzing one dimensional data f, there exist various ways for calculating minimizers of the ROF-model: Mammen & Geer [20] showed that the *taut-string* algorithm, commonly used in statistics, minimizes the ROF-model. Brox & Mrázek & Steidl & Weickert & Welk [7, 22, 34] proved that wavelet thresholding based on the Haar wavelet is equivalent to minimizing the discretized ROF-model, which in turn is *equivalent* to solving the discretized total variation flow equation

$$\frac{\partial u}{\partial t} = \left(\frac{u_x}{|u_x|} \right)_x$$

at time α. Note that the ROF model can be interpreted as a fully implicit time step of the total variation flow equation with step length α. In higher space dimensions properties of the total variation flow equation have been derived by Belletini & Caselles & Novaga [6], Andreu & Ballester & Caselles & Diaz & Mazón [2–5] and Alter & Caselles & Chambolle [1]. Note however, that the equivalence relations do not hold in higher space dimensions.

The outline of this work is as follows: In Section 2 we recall some basic facts on G-norms and bounded variation regularization. In Section 3 we recall tube methods. Finally in Section 4 we present some numerical experiments.

2. Higher Order G-Norms

In this section we introduce generalized G-norms. We give a quite general definition, although in the subsequent sections (for notational convenience) only the cases $k = 1, 2$ and $s = 2$ (see below) are used. This section is central to prove G-norm properties of minimizers of regularization functionals and thus presented in great generality.

For $k = 1, 2, \ldots$ and $s \in [1, \infty]$ we denote by

$$\mathcal{G}^{k,s}(\alpha) := \{v = (\nabla\cdot)^k \vec{v} : \| |\vec{v}|_{l^s} \|_{L^\infty} \leq \alpha\} = \{v : \|v\|_{G^{k,s}} \leq \alpha\} ,$$

where $\vec{v} : \mathbb{R}^n \to \mathbb{R}^{n \times k}$. Here

$$(\nabla\cdot)^k \vec{v} = \sum_{\substack{i_l = 1, \ldots, n \\ l = 1, \ldots, k}} \frac{\partial^k v_{i_1, \ldots, i_k}}{\partial x_{i_1} \ldots \partial x_{i_k}}$$

denotes the k-th divergence and

$$|\vec{v}|_{l^s} = \left(\sum_{\substack{i_l = 1, \ldots, n \\ l = 1, \ldots, k}} |v_{i_1, \ldots, i_k}|^s \right)^{1/s} .$$

Moreover

$$\nabla^k v = \left[\frac{\partial^k v}{\partial x_{i_1} \ldots \partial x_{i_k}} \right]_{\substack{i_l = 1, \ldots, n \\ l = 1, \ldots, k}}$$

denotes the k-th derivative.

If not specified otherwise we denote by $|\cdot| = |\cdot|_{l^2}$ the Euclidean norm (respectively Frobenius norm for matrices and tensors). We call

$$\|v\|_{G^{k,s}} := \inf\{\| |\vec{v}|_{l^s} \|_{L^\infty} : v = (\nabla\cdot)^k \vec{v}\}$$

generalized G-norm. For $k = 1$ and $s = 2$ the generalized G-norm corresponds to the classical definition (see [21]). We denote by

$$G := \{v : \|v\|_{G^{k,s}} < \infty\}$$

the set of all distributions v which have finite G-norm. Analogously to Meyer [21] (see also [33]) it can be shown that for every $v \in G$, there exists $\vec{v}$ satisfying $v = (\nabla\cdot)^k \vec{v}$ with $\|v\|_{G^{k,s}} = \||\vec{v}|_{l^s}\|_{L^\infty} < \infty$.

In the following, for notational convenience, we omit in the case $k = 1$ the first index in the pair of subscripts, i.e., $G^s := G^{1,s}$. For $s \in [1,\infty]$ let $s_* \in [1,\infty]$ satisfy $1 = 1/s + 1/s_*$.

In the following we make use of results from functional analysis and measure theory. Appropriate references, where the necessary results can be found are Yosida [36], Zeidler [37], and Evans & Gariepy [13].

Below we show that G is Banachspace, by showing that it is a dual of a Sobolev space. In the case $k = 1$ and $s = 2$ this result is stated in [21].

THEOREM 1 *The set G associated with the norm $\|\cdot\|_{G^{k,s}}$ is a Banach space, which is the dual of the Sobolev space*

$$\tilde{W}^{k,1} := \overline{C_0^\infty}\,,$$

where the closure is taken with respect to the norm

$$\|w\|_{W^{k,1,s_*}} := \int |\nabla^k w|_{l^{s_*}}\ .$$

By "G is the dual space of $\tilde{W}^{k,1}$" we mean not only that the sets are identical but also the associated norms are identical, i.e., $\|\cdot\|_{G^{k,s}} = \|\cdot\|_{(W^{k,1})^*}$.

Proof: On the linear space C_0^∞, $\|\cdot\|_{W^{k,1,s_*}}$ is a norm. Therefore the completion, $\tilde{W}^{k,1}$ is a Banach space (see e.g. [36]).

In the following we denote by $(\tilde{W}^{k,1})^*$ the dual of $\tilde{W}^{k,1}$ and by

$$\begin{aligned}\nabla^k : \tilde{W}^{k,1} &\to L^1(\mathbb{R}^n, \mathbb{R}^{n\times k})\ .\\ u &\to \nabla^k u\end{aligned}$$

With the space of absolutely integrable functions from $\mathbb{R}^n$ into $\mathbb{R}^{n\times k}$, $L^1(\mathbb{R}^n, \mathbb{R}^{n\times k})$, we use the product space norm $\int |\vec{v}|_{l^{s_*}}$. The operator is injective. With respect to the specified norms the operator ∇^k is an isometrical isomorphism between $\tilde{W}^{k,1}$ and the range of ∇^k. Thus the range of ∇^k is closed.

Denoting by $(\nabla^k)^*$ the dual operator and by X^* the dual space of a Banach space X, it follows from Banach's closed range theorem (see e.g. [37, p777]) and the fact that the dual of L^1 is isometrically isomorph to L^∞ (in signs $\hat{=}$) that the adjoint operator

$$\begin{aligned}(\nabla^k)^* : L^\infty(\mathbb{R}^n, \mathbb{R}^{n\times k}) \hat{=} \Big(L^1(\mathbb{R}^n, \mathbb{R}^{n\times k})\Big)^* &\to (\tilde{W}^{k,1})^*\ .\\ \vec{w} &\to (\nabla^k)^* \vec{w}\end{aligned}$$

satisfies that the range of $(\nabla^k)^*$ equals the complement of the kernel of ∇^k. Since ∇^k is injective the kernel is trivial and thus the complement is the whole space $(\tilde{W}^{k,1})^*$.

In particular, since $(\nabla^k)^* = (\nabla\cdot)^k$, this shows that any element of $(\tilde{W}^{k,1})^*$ can be written as $(\nabla\cdot)^k\vec{w}$.

From the definition of the space G we have the characterization

$$G \hat{=} \frac{L^\infty(\mathbb{R}^n, \mathbb{R}^{n\times k})}{\mathcal{N}((\nabla\cdot)^k)} \; .$$

Here $\mathcal{N}$ denotes the kernel of a linear operator and $\frac{X}{Y}$ denotes the factorization space of X with respect to Y. Using that $(\nabla\cdot)^k$ is the adjoint of ∇^k we see that

$$\frac{L^\infty(\mathbb{R}^n, \mathbb{R}^{n\times k})}{\mathcal{N}((\nabla\cdot)^k)} = \frac{L^\infty(\mathbb{R}^n, \mathbb{R}^{n\times k})}{\mathcal{N}((\nabla^k)^*)} \; .$$

Using that L^∞ and L^1 are isometrically isomorph and that $\mathcal{N}((\nabla^k)^*) = (\text{Range}(\nabla^k))^\perp$, we get

$$\frac{L^\infty(\mathbb{R}^n, \mathbb{R}^{n\times k})}{\mathcal{N}((\nabla^k)^*)} \hat{=} \frac{(L^1(\mathbb{R}^n, \mathbb{R}^{n\times k}))^*}{(\text{Range}(\nabla^k))^\perp} \; .$$

Since $\text{Range}(\nabla^k)$ is closed in $L^1(\mathbb{R}^n, \mathbb{R}^{n\times k})$, we have

$$\frac{(L^1(\mathbb{R}^n, \mathbb{R}^{n\times k}))^*}{(\text{Range}(\nabla^k))^\perp} \hat{=} \text{Range}((\nabla^k)^*) = (\tilde{W}^{k,1})^* \; .$$

Combination of the identities gives the desired result. □

The above proof applies to any space which is constructed as the completion of C_0^∞. The characterization with the G-norm relies on the dual operator. For the k-th derivative it is the k-th divergence operator. The adjoint of the Laplacian is the Laplacian and so on.

We note that $\tilde{W}^{k,1}$ is independent of $1 \leq s_* \leq \infty$. Moreover, $\tilde{W}^{k,1}$ is *not* the standard Sobolev space $W^{k,1}$, where in its definition the closure of C_0^∞ is taken with respect to the norm

$$\|w\|_{W^{k,1}} = \sum_{l=0}^{k} \| |\nabla^l w|_{l^2} \|_{L^1} \; .$$

In fact from the Gagliardo-Nirenberg-Sobolev inequality (see e.g. [13]) it follows that

$$\| |\nabla^{k-1} w|_{l^r} \|_{L^{p_n}} \leq C \int |\nabla^k w|_{l^r} \text{ for every } w \in \tilde{W}^{k,1} \, , \tag{1}$$

where

$$p_n := \frac{n}{n-1} \text{ for space dimension } n \geq 2 \text{ and } p_n := \infty \text{ for } n = 1\ . \qquad (2)$$

LEMMA 2 *Assume that there exists* $\alpha > 0$ *such that for every* $v \in C_0^\infty$

$$\left| \int wv \right| \leq \alpha \int |\nabla^k v|_{l^{s*}} \qquad (3)$$

holds, then $\|w\|_{G^{k,s}} \leq \alpha$*.*

Proof: The linear operator

$$L : C_0^\infty \to \mathbb{R}, \quad v \to \int wv$$

can be extended to a linear bounded operator on $\tilde{W}^{k,1}$. Note that by (3) for a sequence $\{v_n\}_{n\in\mathbb{N}}$ converging to v, $\{Lv_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence and thus convergent with limit Lv.

Therefore, from (3) it follows that $w \in (\tilde{W}^{k,1})^*$, with dual norm (which equals the G norm) is less than α. □

Let

$$\mathcal{G}^\Delta(\alpha) := \{v = \Delta\tilde{v} : \|\tilde{v}\|_{L^\infty} \leq \alpha\} = \{v : \|v\|_{G^\Delta} \leq \alpha\}\ .$$

We define

$$\|v\|_{G^\Delta} := \inf\{\|\tilde{v}\|_{L^\infty} : v = \Delta\tilde{v}\}\ .$$

The proof of the following theorem is analogous to the proof of Theorem 1 and thus omitted.

THEOREM 3

$$G^\Delta := \{v : \|v\|_{G^\Delta} < \infty\}$$

associated with the norm $\|\cdot\|_{G^\Delta}$ *is a Banach space, which is the dual of*

$$W^\Delta := \overline{C_0^\infty}\ ,$$

where the closure is taken with respect to the norm

$$\|w\|_{W^\Delta} := \int |\Delta w|\ .$$

Moreover,

$$\|\cdot\|_{G^\Delta} = \|\cdot\|_{(W^\Delta)^*}\ .$$

The proof of the following lemma is analogous to the proof of Lemma 2 and thus omitted.

LEMMA 4 *Assume that there exists* $\alpha > 0$ *such that for every* $v \in C_0^\infty$

$$\left| \int wv \right| \leq \alpha \int |\Delta v| \tag{4}$$

holds, then $\|w\|_{G^\Delta} \leq \alpha$.

In the following we highlight some properties of functions of bounded variation. For more background on this subject we refer to Evans & Gariepy [13].

DEFINITION 5 *The space of* functions of bounded variation *(*BV*) consists of functions* $u \in L^{p_n}$ *satisfying*

$$\|Du\|_{s_*} := \sup \left\{ \int u(\nabla\cdot)\vec{\varphi} \; : \vec{\varphi} \in C_0^1(\mathbb{R}^n;\mathbb{R}^n)\,, |\vec{\varphi}(x)|_{l^s} \leq 1 \right\} < \infty \,.$$

The standard definition of BV requires $u \in L^1$ (cf. [13]). Actually the assumption $u \in L^{p_n}$ is less restrictive, since any function $u \in L^1$ satisfying $\|Du\|_{s_*} < \infty$ is in L^{p_n} (which follows from the Gagliardo-Nirenberg-Sobolev inequality). Note that for $u \in \tilde{W}^{1,1}$, $\|Du\|_{s_*} = \int |\nabla u|_{l^{s_*}}$.

In this paper we further consider functions with derivatives of bounded variation and functions of bounded Laplacian. To avoid notational difficulties we just consider the case $s = s_* = 2$. The generalization to the case $s \neq 2$ is obvious.

DEFINITION 6

- *We define the set of* functions with derivatives of bounded variation *(*$\mathtt{BV}^k$*) as functions* $u \in L^2$ *satisfying*

$$\|D^k u\| := \sup \left\{ \int u(\nabla\cdot)^k \vec{\varphi} \; : \vec{\varphi} \in C_0^k(\mathbb{R}^n; \underbrace{(\mathbb{R}^n \times \mathbb{R}^n \ldots \mathbb{R}^n)}_{k\times})\,, |\vec{\varphi}(x)| \leq 1 \right\} < \infty \,.$$

- *The space of* functions of bounded Laplacian *(*$\mathtt{BV}^\Delta$*) consists of functions* $u \in L^2$ *satisfying*

$$\|\Delta u\| := \sup \left\{ \int u\Delta\varphi : \varphi \in C_0^2(\mathbb{R}^n;\mathbb{R})\,, |\varphi(x)| \leq 1 \right\} < \infty \,.$$

Note that for $u \in C_0^\infty$, $\|\Delta u\| = \int |\Delta u|$.

We consider $\mathtt{BV}^k$ and $\mathtt{BV}^\Delta$ as subsets of L^2. This is not standard, but simplifies the notation considerably. We could actually proceed iteratively (analogously to Definition 5) and define $\mathtt{BV}^k$ as a subset of W^{k-1,p_n} which can be embedded in L^q with $p_n \leq q \leq \frac{n}{n-k}$ for $1 < k < n$ and in L^∞ if $k \geq n$. We avoid distinguishing between the different cases, by defining $\mathtt{BV}^k$ and $\mathtt{BV}^\Delta$ as subsets of L^2. The space $\mathtt{BV}^2$ is commonly denoted as the space of bounded Hessian (see e.g. Demengel [11]).

We have the following Lemma:

LEMMA 7 *Assume* $w \in L^2$.

1 *Let* $\|w\|_{G^{2,2}} \leq \alpha$*, then for any* $h \in \mathtt{BV}^2$

$$\left| \int wh \right| \leq \alpha \|D^2 h\| \ . \tag{5}$$

2 *Let* $\|w\|_{G^\Delta} \leq \alpha$*, then for any* $h \in \mathtt{BV}^\Delta$

$$\left| \int wh \right| \leq \alpha \|\Delta h\| \ . \tag{6}$$

Proof: The proof of the second item is analogous and thus omitted. For $h \in C_0^\infty$ and $w = (\nabla \cdot)^2 \vec{w}$ satisfying $\||\vec{w}|\|_{L^\infty} = \|w\|_{G^{2,2}}$ it follows that

$$\left| \int wh \right| = \left| \int (\nabla \cdot)^2 \vec{w} h \right| = \left| \int \vec{w} \nabla^2 h \right| \leq \|w\|_{G^{2,2}} \|D^2 h\| \ . \tag{7}$$

Let $h \in \mathtt{BV}^2$, then there exists a sequence $\{h_l\}_{l \in \mathbb{N}}$ in C_0^∞ such that $h_l \to h$ in L^2 and $\|D^2 h_l\| \to \|D^2 h\|$. Consequently, from (7) it follows that

$$\begin{aligned} \left| \int wh \right| &= \lim_{l \to \infty} \left| \int w h_l \right| \\ &\leq \liminf_{l \to \infty} \|w\|_{G^{2,2}} \|D^2 h_l\| \\ &= \|w\|_{G^{2,2}} \|D^2 h\| \ . \end{aligned}$$

□

3. Tube Methods

In this section we recall some basic facts on the minimizer of the ROF-functional:

THEOREM 8 *Let* $f \in L^2$*. Then the minimizer* u_α *of the generalized ROF functional*

$$\frac{1}{2} \int (u-f)^2 + \alpha \|Du\|_{s_*}$$

satisfies:

1 $u_\alpha \in L^2 \cap \mathtt{BV}$;

2 $u_\alpha \equiv 0$ *if and only if* $\|f\|_{G^s} \leq \alpha$;

3 *If* $\|f\|_{G^s} > \alpha$, u_α *is characterized by*

   (a) $\|u_\alpha - f\|_{G^s} = \alpha$ *and*

   (b) $\int (f - u_\alpha) u_\alpha = \alpha \|Du_\alpha\|_{s_*}$.

For $s = 2$ this results has been given in [21] and for general s it has been presented in [29].

Let Φ be measurable and satisfy $\Delta\Phi = f$ with $F_f := \nabla\Phi \in L^\infty_{\text{loc}}$. All along this paper we have been considering data filtering on $\mathbb{R}^n$. If we consider data smoothing on a bounded, smooth domain Ω, the existence of a solution of Laplace's equation $\Delta\Phi = f$ with Neumann boundary data is guaranteed if $\int f = 0$. For $\mathbb{R}^n$ we assume the existence of a solution of this equation, which imposes further requirements on the data f. By definition $\|\rho - f\|_{G^s} \leq \alpha$ if and only if $\rho - f = (\nabla\cdot)\vec{v}$ and $\||\vec{v}|_{l^s}\|_{L^\infty} \leq \alpha$. This is equivalent to

$$\rho = (\nabla\cdot)(\vec{v} + F_f) \text{ and } \||\vec{v}|_{l^s}\|_{L^\infty} \leq \alpha \; .$$

Or in other words, ρ is the divergence of a vector valued function $\vec{\rho}$ which is in a tube around the "primitive" of f (to be precise, we solve Laplace's equation and differentiate). The tube is a subset of $\mathbb{R}^{2n}$ around the vector valued function F_f. We recall that u_α is the divergence of a vector valued function $\vec{u}_\alpha$ and the distance between $\vec{u}_\alpha$ and F_f is less than α, i.e., $\||F_f - \vec{u}_\alpha|_{l^s}\|_{L^\infty} \leq \alpha$. Note that the tube geometry varies with s and has an impact on the solution (cf. [26]). For $s = 2$ the tube has a cylindrical shape and for $s = 1$ or ∞ the tube is a slot.

The following geometric interpretations of the bounded variation regularized solutions u_α are immediate: the associated vector field $\vec{u}_\alpha$ does *not* have contact with the tube if and only if $\|f\|_{G^s} \leq \alpha$. For more background on the concept of tube methods we refer to [15, 33].

3.1 Higher order derivatives of functions of bounded variation

To our knowledge Chambolle & Lions [8] first studied $\mathtt{BV}$-models with second order derivatives for denoising. Their approach consists in minimization of the functional

$$\mathcal{F}_{C-L}(u_1, u_2) := \frac{1}{2} \int (u_1 + u_2 - f)^2 + \beta \|Du_1\| + \alpha \|D^2 u_2\|$$

with $0 < \alpha, \beta$. The asymptotic model, for $\beta \to +\infty$, for *denoising* has been introduced in [32]: the noisy function f is approximated by the minimizer of the functional

$$\mathcal{F}_D(u) := \frac{1}{2} \int (u - f)^2 + \alpha \|D^2 u\| \ . \tag{8}$$

The motivation for studying this type of regularization arises from *nondestructive evaluation* to recover discontinuities of a derivative of a potential u in impedance problems. The discontinuities of u are locations of *material defects* (see e.g. Isakov [17, 18]). Later on second order models for denoising have been considered by Chan & Marquina & Mulet [9] and Lysaker & Lundervold & Tai [19]. It can also be used for segmentation of *low contrast data* [16].

3.2 Characterization of minimizers of regularization functionals with higher order derivatives

In the following we summarize some basic characterization for the minimizer of the functional $\mathcal{F}_D$ and $\mathcal{F}_\Delta$. The following results can be generalized in a straight forward manner to higher order derivatives: the following proofs of the results require only elementary calculations and references to the general lemmas in Section 2, which can be generalized for higher order derivatives. However, for the sake of simplicity of notation we restrict attention to the case $k = 2$ and $s = 2$. The case $s \neq 2$ can be treated by following the proofs in [29].

THEOREM 9 *Assume $f \in L^2$. Then,*

1 *the functional $\mathcal{F}_D$ attains a unique minimizer $u_\alpha \in L^2$ with $\nabla u_\alpha \in L^{p_n}$.*

2 *$\|f\|_{G^{2,2}} \leq \alpha$ if and only if u_α is zero.*

Proof:

1 The existence of minimizer of the functional $\mathcal{F}_D$ follows from its weak lower semi continuity and coercivity in L^2. The weak lower semi continuity of $\mathcal{F}_D$ follows from the weak lower semi continuity of $\int(u - f)^2$ in L^2 and the weak lower semi continuity of the Radon measure $\|D^2 u\|$. The functional $\int(u - f)^2$ is strictly convex and the regularization functional $\|D^2 u\|$ is convex. Thus $\mathcal{F}_D$ is strictly convex and attains a unique minimizer. The assertion $\nabla u \in L^{p_n}$ follows from the Gagliardo-Nirenberg-Sobolev inequality.

2 We have $u \in \mathtt{BV}^2$ if and only if $\mathcal{F}_D(u) < +\infty$. From the definition of a minimizer u_α of $\mathcal{F}_D$ it follows that for every $h \in \mathtt{BV}^2$, $\varepsilon \neq 0$

$$\begin{aligned}
&\frac{1}{2}\int (u_\alpha - f)^2 \;+\; \alpha\|D^2 u_\alpha\| \\
\leq&\frac{1}{2}\int (u_\alpha + \varepsilon h - f)^2 \;+\; \alpha\|D^2(u_\alpha + \varepsilon h)\| \\
\leq&\frac{1}{2}\int (u_\alpha - f)^2 + \varepsilon \int (u_\alpha - f)h + \frac{\varepsilon^2}{2}\int h^2 \\
&+\; \alpha\left(\|D^2 u_\alpha\| + |\varepsilon|\|D^2 h\|\right)
\end{aligned} \tag{9}$$

Consequently, it follows by dividing the terms in the inequality by $|\varepsilon|$ and taking $\varepsilon \to 0^\pm$ afterward that

$$\left|\int (u_\alpha - f)h\right| \leq \alpha\|D^2 h\| \text{ for } h \in \mathtt{BV}^2\ . \tag{10}$$

If $u_\alpha \equiv 0$, then from (10) it follows that for every $h \in \mathtt{BV}^2$

$$\left|\int fh\right| \leq \alpha\|D^2 h\|\ . \tag{11}$$

If (11) holds, then for every $h \in \mathtt{BV}^2$ we have

$$\begin{aligned}
\mathcal{F}_D(h) - \mathcal{F}_D(0) =& \frac{1}{2}\int ((h-f)^2 - f^2) + \alpha\|D^2 h\| \\
\geq& \int -fh + \alpha\|D^2 h\| \\
\geq& 0\ .
\end{aligned}$$

Or in other words $u_\alpha \equiv 0$ is a global minimizer. That is, we have shown that $u_\alpha \equiv 0$ if and only if (11) holds for every $h \in \mathtt{BV}^2$.

From (11), the assumption $f \in L^2$, and Lemma 2 it follows that $\|f\|_{G^{2,2}} \leq \alpha$. Conversely, if $\|f\|_{G^{2,2}} \leq \alpha$, then from Lemma 7 it follows that for any $h \in \mathtt{BV}^2$

$$\left|\int fh\right| \leq \|D^2 h\| \cdot \|f\|_{G^{2,2}} \leq \alpha\|D^2 h\|\ .$$

That is, we have shown that (11) holds for every $h \in \mathtt{BV}^2$ if and only if $\|f\|_{G^{2,2}} \leq \alpha$ and referring back to the above equivalence relation the assertion follows.

□

THEOREM 10 *Let $f \in L^2$ satisfy $\|f\|_{G^{2,2}} > \alpha$. Then $u = u_\alpha$ minimizes $\mathcal{F}_D$ if and only if*

1 $u \in \mathtt{BV}^2$,

2

$$\|u - f\|_{G^{2,2}} = \alpha \,, \tag{12}$$

3 and

$$-\int (u-f)u = \alpha \|D^2 u\| \;. \tag{13}$$

Proof: From the assumption $\|f\|_{G^{2,2}} > \alpha$ it follows from Theorem 9 that $u_\alpha \neq 0$.

From the definition of a minimizer u_α of $\mathcal{F}_D$ it follows that for every $0 \neq |\varepsilon| < 1$

$$\frac{1}{2}\int (u_\alpha - f)^2 \;+\; \alpha\|D^2 u_\alpha\| \leq \frac{1}{2}\int ((1+\varepsilon)u_\alpha - f)^2 \;+\; \alpha(1+\varepsilon)\|D^2 u_\alpha\| \;,$$

showing that

$$-\varepsilon \int (u_\alpha - f)u_\alpha - \frac{\varepsilon^2}{2}\int u_\alpha^2 \leq \alpha\varepsilon \|D^2 u_\alpha\| \;.$$

Dividing the inequality by $|\varepsilon|$ and taking $\varepsilon \to 0\pm$ shows (13). Since $\|D^2 u_\alpha\| \neq 0$, it follows from (10) that $\|u_\alpha - f\|_{G^{2,2}} = \alpha$. To prove the converse direction we note that for $u \in \mathtt{BV}^2$ satisfying $\|u - f\|_{G^{2,2}} = \alpha$ it follows from Lemma 7 that for any function $h \in \mathtt{BV}^2$

$$\|D^2(u+h)\| \geq -\frac{1}{\alpha}\int (u+h)(u-f) \;. \tag{14}$$

From (14), and (13) it follows that for any function $h \in \mathtt{BV}^2$

$$\begin{aligned} &\frac{1}{2}\int (u+h-f)^2 \;+\; \alpha\|D^2(u+h)\| \\ \geq&\frac{1}{2}\int (u-f)^2 \;+\; \int h(u-f) - \int (u+h)(u-f) \\ =&\frac{1}{2}\int (u-f)^2 + \alpha\|D^2 u\| \;. \end{aligned}$$

This shows that u is a global minimizer. □

In the following we consider the problem of minimization of the functional

$$\mathcal{F}_\Delta := \frac{1}{2}\int (u-f)^2 + \alpha \int |\Delta u| \;.$$

Analogously as for $\mathcal{F}_D$ it can be proven that the functional $\mathcal{F}_\Delta$ attains a unique minimizer in $\mathtt{BV}_\Delta$. For $f \in L^2$ it follows from the general results above that $\|f\|_{G^\Delta} \leq \alpha$ if and only if u_α is zero.

The minimizer u_α of $\mathcal{F}_\Delta$ is in the tube G^Δ. Geometrically this means that the second primitive of u_α (to be precise the second primitive is the solution of Laplace's equation $\Delta u = u_\alpha$) is in a tube of radius α around the second primitive of f.

4. Numerical results

In this section we present some numerical results for minimizing the functionals $\mathcal{F}_{ROF}$ and $\mathcal{F}_\Delta$ or $\mathcal{F}_D$, respectively.

We concentrate on minimization of the functionals for one dimensional input data f. For higher dimensional data the tube properties cannot be visualized easily and are not as illustrative.

The first examples is discrete bounded variation minimization for analyzing one dimensional data. There the Rudin-Osher-Fatemi functional is discretized with a one-sided difference operator. The according optimality condition has been solved with a fixed point iteration. In the case of 1-dimensional data one might alternatively use the taut-string algorithm (cf. [20]) for calculating the BV-minimizer. For one dimensional discrete data, a simple method to minimize the ROF model has been proposed in [24]. In Figure 1 (left row) we have plotted synthetic data f with different noise levels, and the discrete BV-minimizer. In the case of noise free data (top figure) significant stair casing occurs in the inclined part of the function.

The stair casing effect is inherent to the minimizer of the ROF-model (see [30]). At least in the discrete setting the BV-minimizer can be exactly calculated with the taut-string algorithm [20]. The taut-string algorithm consists in integration of the discrete data, constructing a tube around the (discrete) primitive, and finding a string of minimal length in the tube. The finite difference quotient of the taut-string is the BV-minimizer. Since the primitive of the piecewise linear function $\hat{f}$ according to the sample data f is quadratic in the inclined region, here the taut-string approximates the inclined region by a piecewise linear function, resulting in a significant stair-casing. In the inclined region the stair casing can be removed by taking into account the information that the primitive is quadratic in the inclined region, and therefore the BV-minimizer must be linear. Thus in these regions we can use linear approximations and prevent some stair casing. Locally quadratic regions of the discrete BV-minimizer are detected when two consecutive samples have contact with the tube. In this case we approximate the BV-minimizer by a piecewise linear function instead of a piecewise constant. This visual correction has been performed in the right row of Figure 1. The bottom of Figure 1 provides a zoom into the BV-minimizer

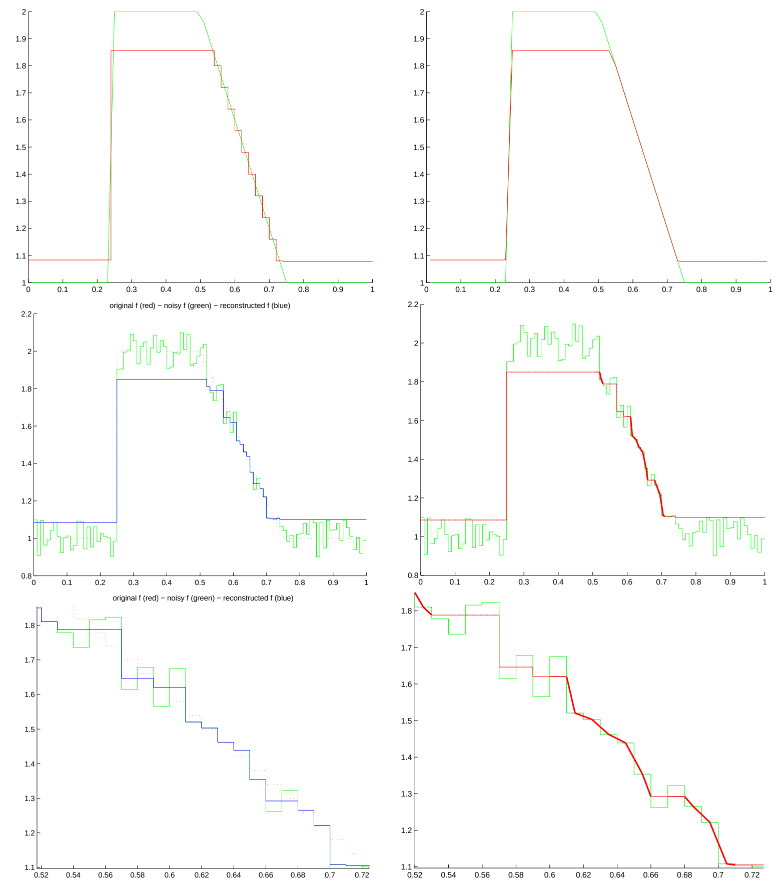

Figure 1. *Left:* Bounded variation regularization. *Right:* Smoothing in contact zones, using the characterization of minimizers with the G-norm. *Top:* Exact Data. *Middle* Noisy Data. *Bottom:* Zoom in the bounded variation regularized solution and the smoothing in contact zones.

and the correction. In the noise free case the inclined part of the reconstruction is perfect linear.

In the next example (cf. Figure 2) we present some experiments for minimization of $\mathcal{F}_\Delta$. From the analytical results we know that $u_\alpha \in \mathcal{G}_\Delta(\alpha)$. The numerical results show that the reconstruction has contact at locations, where the data f has discontinuities. Discontinuities in the derivative show up significantly in the tube, but do not have contact with the tube boundary. In comparison bounded variation regularization does not reveal a separation in multiple scales for discontinuities in the function and derivative (cf. Figure 3). The bottom left image of Figure 2 has four pumps, which indicate the discontinuities (large pumps with contact to the tube) and small pumps for the discontinuities of the derivatives. This behavior is also clearly visible for the noisy data (cf. bottom right image of Figure 2). For purely denoising this model has the disadvantage that the reconstruction reveals a “ramp-shape” effect and to an imprecise edge localization, but on the other hand also discontinuities in the first derivative can be recognized. Bounded variation regularization has either contact with the tube or is linear in between. Note that for bounded variation regularization the tube is around the first primitive, while it is around the second in the second order model.

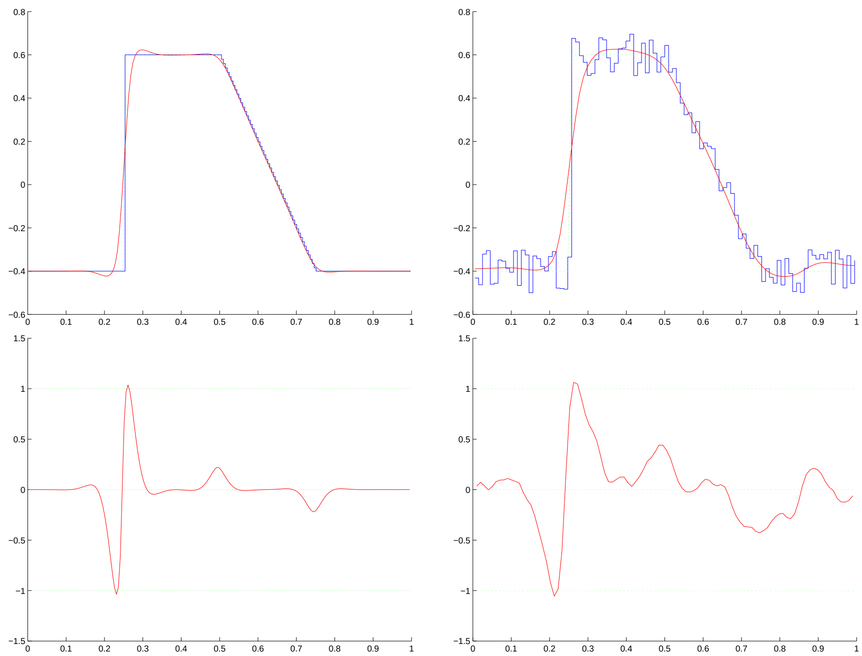

Figure 2. *Top:* Synthetic data and minimizer of $\mathcal{F}_\Delta$ *Bottom:* Second primitive of the minimizer reveals features at multiple scales.

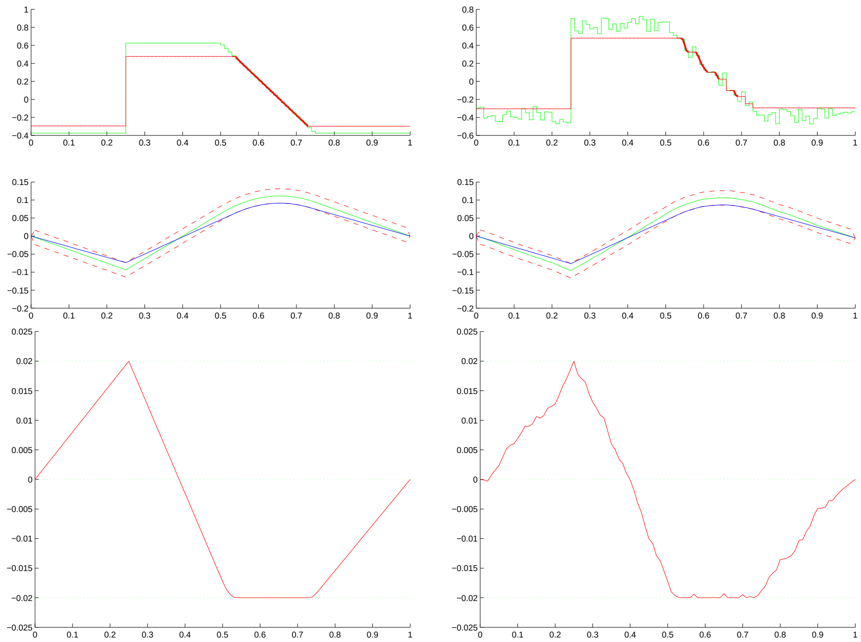

Figure 3. *Top:* Data f and u_α (BV minimizer). *Middle:* The first primitive of u_α is in a tube around the primitive of f. *Bottom* Zoom, to visualize the primitive of $u_\alpha - f$.

Acknowledgment

The work of A.O. and O.S. has been supported by the FWF (Österreichischer Fonds zur Förderung der wissenschaftlichen Forschung), grant Y-123 INF-N04 and P-15617-N04. The work of S.O. was supported by NSF grants ACI-0321917 and DMS 0312222. O.S. would like to thank Markus Grasmair for stimulating discussions on dual spaces.

References

[1] F. Alter, V. Caselles, and A. Chambolle. Evolution of convex sets in the plane by the minimizing total variation flow. *Preprint*, pages 1–29, 2003.

[2] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Minimizing total variation flow. *C. R. Acad. Sci. Paris Sér. I Math.*, 331:867–872, 2000.

[3] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. The Dirichlet problem for the total variation flow. *J. Funct. Anal.*, 180:347–403, 2001.

[4] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Minimizing total variation flow. *Differential Integral Equations*, 14:321–360, 2001.

[5] F. Andreu, V. Caselles, J. I. Diaz, and J. M. Mazón. Some qualitative properties for the total variation flow. *J. Funct. Anal.*, 188:516–547, 2002.

[6] G. Bellettini, V. Caselles, and M. Novaga. The total variation flow in $\mathbb{R}^N$. *J. Differential Equations*, 184:475–525, 2002.

[7] T. Brox, M. Welk, G. Steidl, and J. Weickert. Equivalence results for TV diffusion and TV regularisation. In *[14]*, 2003.

[8] A. Chambolle and P.L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.*, 76:167–188, 1997.

[9] T. Chan, A. Marquina, and P. Mulet. High-order total variation-based image restoration. *SIAM J. Sci. Comput.*, 22:503–516 (electronic), 2000.

[10] P. L. Davies and A. Kovac. Local extremes, runs, strings and multiresolution. *Ann. Statist.*, 29:1–65, 2001. With discussion and rejoinder by the authors.

[11] F. Demengel. Fonctions a Hessien borné. *Ann. Inst. Fourier (Grenoble)*, 34:155–190, 1984.

[12] L. Dümbgen and A. Kovac. Extensions of smoothing via taut strings. report, Universität Duisburg Essen, 2004.

[13] L.C. Evans and R.F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC–Press, Boca Raton, 1992.

[14] L.D. Griffin and M. Lillholm, editors. *Scale-Space Methods in Computer Vision*. Lecture Notes in Computer Science Vol. 2695, Springer Verlag, 2003. Proceedings of the 4th International Conference, Scale-Space 2003, Isle of Skye, UK, June 2003.

[15] W. Hinterberger, M. Hintermüller, K. Kunisch, M. von Oehsen and O. Scherzer. Tube Methods for BV regularization, *J. of Math. Imaging and Vision*, 19:223–238, 2003.

[16] W. Hinterberger and O. Scherzer. Variational methods on the space of functions of bounded Hessian for convexification and denoising. 2003. submitted.

[17] V. Isakov. *Inverse Source Problems*. American Mathematical Society, Providence, Rhode Island, 1990.

[18] V. Isakov. *Inverse Problems for Partial Differential Equations*. Springer, New York, 1998. Applied Mathematical Sciences (127).

[19] M. Lysaker, A. Lundervold and X. Tai. *Noise removal using fourth order partial differential equations with applications to medical magnetic resonance imaging in space and time*. IEEE Trans. on Image Processing, v12, pp1579-1590, 2003

[20] E. Mammen and S. van de Geer. Locally adaptive regression splines. *Ann. Statist.*, 25:387–413, 1997.

[21] Y. Meyer. *Oscillating patterns in image processing and nonlinear evolution equations*, volume 22 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2001.

[22] P. Mrázek, J. Weickert, and G. Steidl. Correspondences between wavelet shrinkage and nonlinear diffusion. In *[14]*, pages 101–116, 2003.

[23] M. Nikolova. Local strong homogeneity of a regularized estimator. *SIAM J. Appl. Math.*, 61:633–658, 2000.

[24] M. Nikolova. A variational approach to remove outliers and impulse noise. *J. of Math. Imaging and Vision*, 20:99–120, 2004.

[25] M. Nikolova. Weakly constraint minimization. Application to the estimation of images and signals involving constant regions. *J. of Math. Imaging and Vision*, 21:155–175, 2004.

[26] S. Osher and S. Esedoglu. Decomposition of images by the anistropic Rudin-Osher-Fatemi model. *Comm. Pure Appl. Math.*, to appear, 2004.

[27] S. Osher and R. Fedkiw. *Level set methods and dynamic implicit surfaces*, volume 153 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2003.

[28] S. Osher and N. Paragios. *Geometric level set methods in Imaging, Vision, and Graphics*. Springer-Verlag, New York, 2003.

[29] S. Osher and O. Scherzer. G-norm properties of bounded variation regularization. *Comm. Math. Sci*, 2:237-254, 2004.

[30] W. Ring. Structural properties of solutions of total variation regularization problems. *M2AN Math. Model. Numer. Anal.*, 34:799-810, 2000.

[31] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[32] O. Scherzer. Denoising with higher order derivatives of bounded variation and an application to parameter estimation. *Computing*, 60:1–27, 1998.

[33] O. Scherzer. Taut-string algorithm and regularization programs with G-norm data fit. *J. of Math. Imaging and Vision*, to appear.

[34] G. Steidl, J. Weickert, T. Brox, P. Mrázek, and M. Welk. On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDes. *SIAM J. Num. Anal*. in press.

[35] D. Strong and T. F. Chan. Exact solutions to the total variation regularization problem. Technical report, University of California, Los Angeles, 1996. CAM 96 - 41.

[36] K. Yosida. *Functional analysis*. Springer-Verlag, New York, 1995.

[37] E. Zeidler. *Nonlinear Functional Analysis and its Applications I*. Springer–Verlag, New York, 1993. corrected printing.