WOLFGANG RING

Structural properties of solutions to total variation regularization problems

<http://www.numdam.org/item?id=M2AN_2000__34_4_799_0>

STRUCTURAL PROPERTIES OF SOLUTIONS TO TOTAL VARIATION REGULARIZATION PROBLEMS

WOLFGANG RING[1]

Abstract. In dimension one it is proved that the solution to a total variation-regularized least-squares problem is always a function which is "constant almost everywhere", provided that the data are in a certain sense outside the range of the operator to be inverted. A similar, but weaker result is derived in dimension two.

Mathematics Subject Classification. 26B30, 46N10, 47A52, 49J52, 49N45, 65K10.

1. INTRODUCTION

Since the work by Rudin, Osher and Fatemi [15] published in 1992, regularization by total variation functionals has received considerable attention in image and signal processing. See the articles of Chambolle and Lions [3], Acar and Vogel [1], Dobson and Santosa [6], Vogel and Oman [17], Chavent and Kunisch [4], Ito and Kunisch [11], Dobson and Scherzer [5], Nashed and Scherzer [13] and the references cited therein.

In this paper, we consider the least-squares problem

$$\min_{u \in BV(\Omega)} \|Ku - z\|^2 + \alpha |\nabla u|(\Omega) \tag{1.1}$$

where $|\nabla u|(\Omega)$ stands for the total variation of the (distributional) gradient of u and $BV(\Omega)$ denotes the space of functions of bounded variation on Ω. The variational problem (1.1) should be considered as a stable approximation to the possibly ill-posed problem

$$Ku = z.$$

It has been observed by several authors that "blocky" structures or piecewise constant functions can be reconstructed very well by total variation regularization (see [6]). On the other hand, solutions to total variation-regularized problems are usually rather "blocky" even if the exact solution u_0, corresponding to exact, noise-free data z_0, is smooth (see [3,12], and the recent work of Nikolova [14]).

We illustrate the situation in two numerical examples. Here our goal is not the best possible reconstruction of the ideal solution. We choose configurations of noise and regularization parameter which display structural properties of the solution as good as possible.

[1] Institut für Mathematik, Universität Graz, Heinrichstrasse 36, 8010 Graz, Austria. e-mail: wolfgang.ring@kfunigraz.ac.at
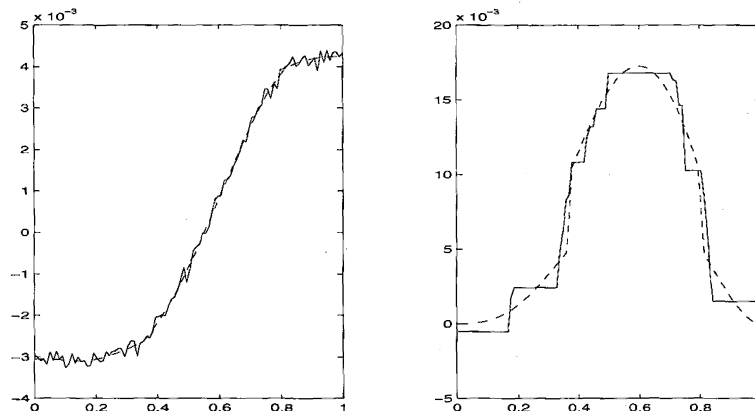
FIGURE 1. Unperturbed and perturbed data and solution to (1.1).

We first consider a one-dimensional problem of type (1.1). Here K is an integral operator of the first kind,

$$Ku(x) = \int_0^x u(\xi)\,d\xi - \int_0^1 (1-\xi)u(\xi)\,d\xi.$$

The plot on the left hand side in Figure 1 shows exact data z_0 together with noise-corrupted data z. The solution to (1.1) and the exact solution u_0 with $Ku_0 = z_0$ are plotted on the right hand side as solid and dashed lines respectively. It is apparent, that the solution $\bar{u}$ to (1.1) is constant over large subintervals and that there are jump discontinuities at a number of points.

Similar effects can be observed in dimension two. Figure 2 shows the regularized inversion of noise corrupted data $z \in L^2([0,1] \times [0,1])$ when K is the compact Fredholm operator

$$Kf(\mathbf{x}) = \int_\Omega e^{-2|\mathbf{x}-\boldsymbol{\xi}|^2} f(\boldsymbol{\xi})\,d\boldsymbol{\xi}$$

with $\Omega = [0,1] \times [0,1]$. In this case, the solution is the union of disjoint patches where on each patch the gradient of the solution is pointing parallel to one of the coordinate axes.

In this paper, we give some theoretical explanations for this behavior. In dimension one, we can characterize the structure of the solution in presence of data noise quite thoroughly. In this situation, we prove that the derivative of the solution is zero almost everywhere. In dimension two, we do not obtain such a strong result but we can derive some structural properties of solutions which correspond to the above mentioned observation. Specifically, if we use the most simple (anisotropic) definition of the norm on $BV(\Omega)$, we can prove that generically the gradient of the solution is parallel to one of the coordinate directions. In both one- and two-dimensional situations, we focus our attention on two important special cases: the problem of denoising noise-corrupted data and, the problem of deblurring data which are blurred by some smoothing filter. We also mention that similar results were derived recently by M. Nikolova [14] for finite dimensional approximations of the functional (1.1) using quite different methods.

In Section 2 we introduce the spaces $BV(\Omega)$ and $\mathcal{M}_n(\Omega)$ and we recall some structural properties of functions of bounded variation and of Radon measures. Here we use mainly the Lebesgue decomposition of a measure into its absolutely continuous and singular parts. In Section 3, we formulate the variational problem and its functional analytic setting, we prove existence of a solution, and we derive an optimality system for (1.1). In Sections 4 and 5 we study structural properties of solutions to (1.1) in dimensions one and two respectively.
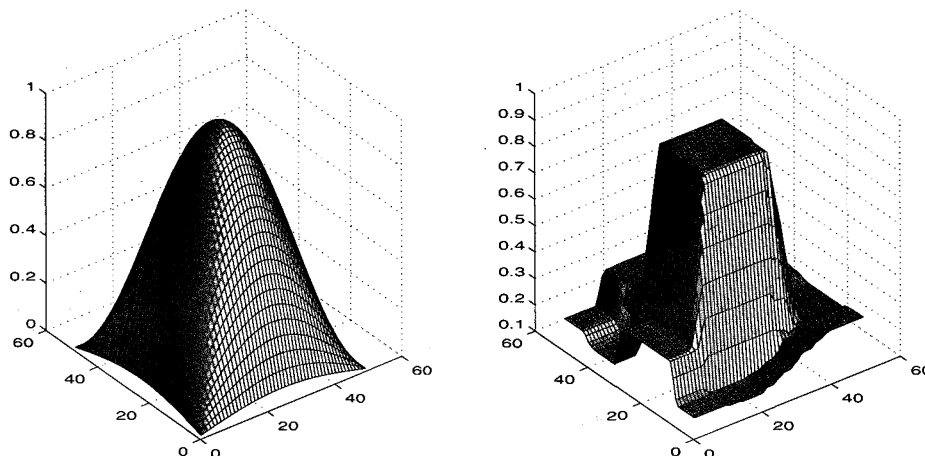
FIGURE 2. Ideal solution and solution to (1.1) for a two-dimensional problem.

2. THE SPACES $BV(\Omega)$ AND $\mathcal{M}_n(\Omega)$

Let Ω be a bounded, open, simply connected domain in $\mathbb{R}^n$ with $n \geq 1$. We consider the space

$$BV(\Omega) = \{f \in L^1(\Omega) : \frac{\partial f}{\partial x_i} \in \mathcal{M}(\Omega) \text{ for } i = 1, \ldots, n\} \tag{2.1}$$

where $\frac{\partial}{\partial x_i}$ denotes the distributional derivative with respect to x_i and $\mathcal{M}(\Omega)$ is the vector space of all Radon measures on Ω, i.e., the dual of $\mathcal{C}_0(\Omega)$ (see Th. 6.19, p. 130 in [16]). $BV(\Omega)$ is a Banach space if we define the norm on $BV(\Omega)$ by

$$\|f\|_{BV} = \|f\|_{L^1} + \sum_{i=1}^{n} \left|\frac{\partial f}{\partial x_i}\right|(\Omega) \tag{2.2}$$

where $\left|\frac{\partial f}{\partial x_i}\right|(\Omega)$ denotes the total variation of the measure $\frac{\partial f}{\partial x_i}$. We shall write (2.2) in the short form

$$\|f\|_{BV} = \|f\|_{L^1} + |\nabla f|(\Omega). \tag{2.3}$$

The term $|\nabla f|(\Omega)$ can alternatively be written as

$$|\nabla f|(\Omega) = \sup\left\{\int_\Omega f(\mathbf{x})\operatorname{div}\mathbf{v}(\mathbf{x})\,d\mathbf{x} : \mathbf{v} \in \mathcal{C}_0^\infty(\Omega)^n, |\mathbf{v}(\mathbf{x})|_\infty \leq 1 \text{ for all } x \in \Omega\right\}. \tag{2.4}$$

Remark 1. If we choose a different vector norm for $\mathbf{v}$ in (2.4), we get a different (but equivalent) norm on $BV(\Omega)$. We choose the $|\cdot|_\infty$-norm because then (2.2) and (2.4) are equivalent definitions. For any other norm for $\mathbf{v}$, the term corresponding to $|\nabla f|(\Omega)$ cannot be expressed as a simple sum of variations, or any other simple function of the terms $\left|\frac{\partial f}{\partial x_i}\right|(\Omega)$. Using our definition, we get a norm which is easy to analyze, but from the practical point of view the norm has the drawback of being anisotropic.

Remark 2. For $u \in BV(\Omega)$ with $\frac{\partial u}{\partial x_i} \in L^1(\Omega) \subset \mathcal{M}(\Omega)$ we have (see Th. 6.13, p. 125 in [16])

$$\left| \frac{\partial u}{\partial x_i} \right| (\Omega) = \int_\Omega \left| \frac{\partial u}{\partial x_i}(\mathbf{x}) \right| d\mathbf{x}.$$

Thus, the Sobolev space

$$W^{1,1}(\Omega) = \{ f \in L^1(\Omega) \ : \ \frac{\partial f}{\partial x_i} \in L^1(\Omega), \ i = 1, \dots, n \}$$

is isometrically embedded in $BV(\Omega)$. Although $W^{1,1}(\Omega)$ is not dense in $BV(\Omega)$ in the norm topology, we can approximate functions in $BV(\Omega)$ even by $\mathcal{C}^\infty$-functions if we use a slightly weaker topology (see [8]: p. 172 ff).

We have the following properties.

Proposition 1.

1. Lower Semicontinuity:
   If $\{u_j\}_{j=1}^\infty \subset BV(\Omega)$ and $u_j \to u$ in $L^1(\Omega)$ then

$$\left| \frac{\partial u}{\partial x_i} \right| (\Omega) \leq \liminf_{j \to \infty} \left| \frac{\partial u_j}{\partial x_i} \right| (\Omega). \tag{2.5}$$

2. Compactness:
   For every bounded sequence $\{u_j\}_{j=1}^\infty \subset BV(\Omega)$ and for every $p \in [1, \frac{n}{n-1})$ there exists a subsequence $\{u_{j_k}\}_{k=1}^\infty$ and a function $u \in BV(\Omega)$ such that $u_{j_k} \to u$ in $L^p(\Omega)$. If $n = 1$ then "∞" is substituted for "$\frac{n}{n-1}$".

3. Sobolev Inequality:
   There exists a constant $C = C(n, \Omega)$ such that

$$\| u - \tilde{u} \|_{L^{\frac{n}{n-1}}(\Omega)} \leq C \, |\nabla u|(\Omega), \text{ for all } u \in BV(\Omega), \tag{2.6}$$

where the constant function $\tilde{u}$ is defined by $\tilde{u} = \frac{1}{\text{meas}(\Omega)} \int_\Omega u \, d\mathbf{x}$. If $n = 1$ the $L^{\frac{n}{n-1}}$-norm is understood to be the L^∞-norm.

For proofs of the above described properties we refer to [4, 10, 18].
We consider the space $\mathcal{M}_n(\Omega) = \overset{n}{\underset{i=1}{\otimes}} \mathcal{M}(\Omega)$. A norm on $\mathcal{M}_n(\Omega)$ is given by

$$\| \boldsymbol{\mu} \|_{\mathcal{M}_n} = \sum_{i=1}^n |\mu_i|(\Omega) \tag{2.7}$$

where $|\mu_i|(\Omega)$ denotes the total variation of the measure μ_i and

$$\boldsymbol{\mu} = (\mu_1, \dots, \mu_n).$$

Remark 3. We can equivalently define the norm on $\mathcal{M}_n(\Omega)$ by

$$|\boldsymbol{\mu}| = \sup \left\{ \sum_{i=1}^n \int_\Omega v_i(\mathbf{x}) \, d\mu_i(\mathbf{x}) \ : \ \mathbf{v} \in \mathcal{C}_0^\infty(\Omega)^n, \ |\mathbf{v}(\mathbf{x})|_\infty \leq 1 \text{ for all } x \in \Omega \right\}. \tag{2.8}$$

For $u \in BV(\Omega)$ we have $\nabla u \in \mathcal{M}_n(\Omega)$ and definition (2.8) coincides with (2.4).

Let $\mu \in \mathcal{M}(\Omega)$. Then, by Lebesgue's decomposition theorem (see Th. 6.10, p. 121 in [16]) it is possible to write μ in a unique way as

$$\mu = \mu_a + \mu_s \tag{2.9}$$

with μ_a being absolutely continuous with respect to the Lebesgue measure, and μ_s being concentrated on a set of Lebesgue measure 0. For $\boldsymbol{\mu} \in \mathcal{M}_n(\Omega)$, we have the Lebesgue decomposition

$$\boldsymbol{\mu} = \boldsymbol{\mu}_a + \boldsymbol{\mu}_s \text{ with } \boldsymbol{\mu}_a = (\mu_{1,a}, \dots, \mu_{n,a}), \quad \boldsymbol{\mu}_s = (\mu_{1,s}, \dots, \mu_{n,s}). \tag{2.10}$$

If $\mu_{i,s}$ is concentrated on the zero set $S_i \subset \Omega$, then $\boldsymbol{\mu}_s$ is concentrated on $S = \cup_{i=1}^n S_i$ which is also of Lebesgue measure 0. Thus $\boldsymbol{\mu}_s$ is a vector-valued singular measure. Moreover, $\boldsymbol{\mu}_s$ and $\boldsymbol{\mu}_a$ provide the unique decomposition of $\boldsymbol{\mu}$ with this property.

Lemma 1. *We have*

$$|\boldsymbol{\mu}|(\Omega) = |\boldsymbol{\mu}_a|(\Omega) + |\boldsymbol{\mu}_s|(\Omega) \tag{2.11}$$

where $\boldsymbol{\mu} = \boldsymbol{\mu}_a + \boldsymbol{\mu}_s$ is the Lebesgue decomposition of $\boldsymbol{\mu} \in \mathcal{M}_n(\Omega)$ into its absolutely continuous and singular parts, as given in (2.10).

Proof. We first prove the result for dimension $n = 1$. Obviously we have $|\mu|(\Omega) \leq |\mu_a|(\Omega) + |\mu_s|(\Omega)$ by the triangle inequality on $\mathcal{M}(\Omega)$.

To prove the converse inequality, suppose we are given two partitions $\{E_i\}_{i=1}^\infty$ and $\{F_n\}_{n=1}^\infty$ of Ω, *i.e.* two countable collections of measurable sets with $E_i \cap E_j = \emptyset$ if $i \neq j$; $F_n \cap F_m = \emptyset$ if $n \neq m$, and $\cup_i E_i = \cup_n F_n = \Omega$. Since μ_a and μ_s are mutually singular, there exist two measurable sets A and B with $A \cap B = \emptyset$ and $A \cup B = \Omega$ such that $\mu_a(E) = \mu(A \cap E)$ and $\mu_s(E) = \mu(B \cap E)$ for every measurable set E. We have

$$\sum_{i=1}^\infty |\mu_a(E_i)| + \sum_{n=1}^\infty |\mu_s(F_n)| = \sum_{i=1}^\infty |\mu(E_i \cap A)| + \sum_{n=1}^\infty |\mu(F_n \cap B)| \leq |\mu|(\Omega)$$

since $\{E_i \cap A\}_{i=1}^\infty \cup \{F_n \cap B\}_{n=1}^\infty$ is a partition of Ω. If we take the supremum over all partitions $\{E_i\}_{i=1}^\infty$ and $\{F_n\}_{n=1}^\infty$ we get

$$|\mu_a|(\Omega) + |\mu_s|(\Omega) \leq |\mu|(\Omega).$$

The result for dimension $n > 1$ follows immediately from the definition of $|\boldsymbol{\mu}|(\Omega)$ and from the one-dimensional case. $\square$

Remark 4. We can isometrically identify absolutely continuous measures with L^1-functions. Hence we can write (2.9) in the form $\mu = f + \mu_s$ with some $f \in L^1(\Omega)$. This yields the following decomposition of $\mathcal{M}(\Omega)$ into a direct sum

$$\mathcal{M}(\Omega) = L^1(\Omega) \oplus \mathcal{P} \tag{2.12}$$

where
$$\mathcal{P} = \{\mu_s \in \mathcal{M}(\Omega) : \mu_s \text{ is singular with respect to the Lebesgue measure on } \Omega\}.$$
The decomposition for the vector-valued case reads as

$$\mathcal{M}_n(\Omega) = L_n^1(\Omega) \oplus \mathcal{P}^n \tag{2.13}$$

with $L_n^1(\Omega) = \overset{n}{\underset{i=1}{\otimes}} L^1(\Omega)$ and $\mathcal{P}^n = \overset{n}{\underset{i=1}{\otimes}} \mathcal{P}$.

We consider the operator

$$\nabla : BV(\Omega) \to \mathcal{M}_n(\Omega).$$

Obviously, ∇ is a bounded linear operator on $BV(\Omega)$. We define the dual operator:

$$-\mathfrak{Div} : \mathcal{M}_n(\Omega)^* \to BV(\Omega)^*, \quad -\mathfrak{Div} = \nabla^*. \tag{2.14}$$

Here and for the rest of the paper "$*$" denotes the dual for both operators and spaces.

3. PROBLEM FORMULATION, SOLVABILITY AND OPTIMALITY SYSTEMS

Let $\Omega \subset \mathbb{R}^n$ be as in Section 2, $p \in [1, \frac{n}{n-1})$ if $n \geq 2$, and $p \in [1, \infty)$ if $n = 1$. Suppose Y is a reflexive Banach space and $K : L^p(\Omega) \to Y$ is a bounded linear operator. We consider the following unconstrained optimization problem

$$\text{minimize } \frac{1}{2}\|Ku - z\|_Y^2 + \alpha|\nabla u|(\Omega) \text{ over } u \in BV(\Omega) \tag{P}$$

with given $z \in Y$. With $p < \frac{n}{n-1}$, and Ω bounded, we have $BV(\Omega)$ continuously embedded in $L^p(\Omega)$ due to Proposition 1.3. Hence, the minimization over $BV(\Omega)$ in (P) is well defined.

Solvability of problem (P) has been considered in a slightly less general framework in [3]. We present the following solvability result for (P).

Proposition 2. *Suppose K and Y are given as above with $\ker K$ containing no constant, non-zero function. Then, for every $z \in Y$ there exists a solution $\overline{u} \in BV(\Omega)$ to problem (P). If K is injective and the norm on Y is strictly convex then the solution is unique.*

Proof. Suppose $\{u_j\}_{j=1}^{\infty}$ is a minimizing sequence for problem (P). We prove that $\{u_j\}$ is bounded in $BV(\Omega)$. From the Sobolev Inequality (2.6) we conclude that $\{v_j\} = \{u_j - \tilde{u}_j\}$ is bounded in $L^{\frac{n}{n-1}}(\Omega)$ and hence also in $L^p(\Omega)$, with $\tilde{u}_j = \frac{1}{\text{meas}(\Omega)} \int_\Omega u_j \, \mathrm{d}x$. We claim that $\{\tilde{u}_j\}$ (as a sequence of constant functions in $L^1(\Omega)$) is also bounded. Suppose otherwise that we have $\tilde{u}_j = r_j \chi_\Omega$ with $|r_j| \to \infty$ as $j \to \infty$. Then $\|K\tilde{u}_j\|_Y = |r_j| \|K\chi_\Omega\|_Y \to \infty$ as $j \to \infty$ because $\ker K$ does not contain any constant, non-zero function. Since $\|Ku_j\|_Y \geq \|K\tilde{u}_j\|_Y - \|Kv_j\|_y$, we find that $\{Ku_j\}$ must be unbounded in Y, contradicting the assumption that $\{Ku_j\}$ is a minimizing sequence for (P). Therefore, $\{u_j\} = \{v_j + \tilde{u}_j\}$ is bounded in $L^1(\Omega)$. Since $\{|\nabla u_j|(\Omega)\}$ is bounded, we conclude that $\{u_j\}$ is bounded in $BV(\Omega)$.

Using Proposition 1.2, there exists a subsequence, again denoted by $\{u_j\}$, satisfying $u_j \to \overline{u}$ in $L^p(\Omega)$ with $\overline{u} \in BV(\Omega)$. From the continuity of K and Proposition 1.1 we conclude

$$\frac{1}{2}\|K\overline{u} - z\|_Y^2 + \alpha|\nabla\overline{u}|(\Omega) \leq \liminf_{j \to \infty} \left(\frac{1}{2}\|Ku_j - z\|_Y^2 + \alpha|\nabla u_j|(\Omega) \right)$$

$$= \inf_{u \in BV(\Omega)} \left(\frac{1}{2}\|Ku - z\|_Y^2 + \alpha|\nabla u|(\Omega) \right)$$

which proves that $\overline{u}$ is a solution to (P).

To prove uniqueness of the minimizer, it is sufficient to show that the cost functional in (P) is strictly convex. This, however, is an immediate consequence of the injectivity of K and the strict convexity of $\| \cdot \|_Y^2$. $\square$

We shall now give a characterization of the solution $\overline{u}$ to Problem (P) by an appropriate optimality system. In this context we shall use the following notations. Let $(\nabla\overline{u})_a = \{(\frac{\partial\overline{u}}{\partial x_i})_a\}_{i=1}^n$ and $(\nabla\overline{u})_s = \{(\frac{\partial\overline{u}}{\partial x_i})_s\}_{i=1}^n$ denote

the absolute continuous and singular parts of $\nabla \overline{u} \in \mathcal{M}_n(\Omega)$. Moreover let $\mathrm{sign}(x)$ denote the set-valued mapping

$$\mathrm{sign}(x) = \begin{cases} -1 & \text{if } x < 0, \\ [-1,1] & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

The duality mapping $J : Y \to Y^*$ is defined as the subdifferential of the convex function $y \mapsto \frac{1}{2}\|y\|_Y^2$ on Y (see Barbu [2, p. 60] for details). Finally let $L^q(\Omega)$ denote the dual space to $L^p(\Omega)$, *i.e.* let $q \in (1, \infty]$ be given such that $\frac{1}{p} + \frac{1}{q} = 1$. With these definitions we find the following optimality system.

Proposition 3. *Let $\overline{u} \in BV(\Omega)$ be a solution to (P) and let Y^* be a strictly convex Banach space. Then there exists $\hat{\lambda} \in L_n^\infty(\Omega)$, with $\mathrm{div}\,\hat{\lambda} \in L^q(\Omega)$ such that:*

$$K^* J(K\overline{u} - z) = \alpha \, \mathrm{div}\, \hat{\lambda} \in L^q(\Omega) \tag{O_1}$$

$$\hat{\lambda}_i(\mathbf{x}) \in \mathrm{sign}\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a (\mathbf{x}) \tag{O_2}$$

almost everywhere on Ω for $i = 1, \ldots, n$.

Proof. The proof proceeds in two steps. We first derive optimality conditions using techniques from convex analysis. In the second step we consider the absolute continuous and the singular parts of $\nabla \overline{u}$ separately and we show that the Lagrange multiplier $\overline{\lambda}$ found in the first step can be replaced by a smoother function $\hat{\lambda}$.

We consider the cost functional

$$F(u) = F_1(u) + \alpha F_2(u)$$

with $F_1(u) = \frac{1}{2}\|Ku - z\|_Y^2$ and $F_2(u) = |\nabla u|(\Omega)$ on $BV(\Omega)$. Obviously F_1, F_2, and F are continuous convex functions on $BV(\Omega)$. Therefore, $\overline{u} \in BV(\Omega)$ is a solution to (P) if and only if

$$0 \in \partial F(\overline{u}) = \partial F_1(\overline{u}) + \alpha \partial F_2(\overline{u}), \tag{3.1}$$

where the ∂-operator denotes the subdifferential of a convex function (see Prop. 1, p. 124 and Cor. 1, p. 121 in [7] for details). F_2 is in fact the composition of the affine mapping $u \mapsto Ku - z$ and the convex function $y \mapsto \frac{1}{2}\|y\|_Y^2$. Therefore, using the arguments in Prop. 12, p. 119 of [7], we find that

$$\partial F_1(\overline{u}) = K^* J(K\overline{u} - z).$$

The duality mapping is single valued due to the strict convexity of Y^* (see Ex. 1, p. 60 and Th. 1.2, p. 2 in [2]), hence $\partial F_1(\overline{u})$ is also single valued. Since $K^* : Y^* \to L^q(\Omega)$, we obtain $\partial F_1(\overline{u}) \in L^q(\Omega) \subset (BV(\Omega))^*$.

The mapping F_2 is the composition of the bounded linear operator $\nabla : BV(\Omega) \to \mathcal{M}_n(\Omega)$ and the convex function $\tilde{F}_2 : \mu \mapsto |\mu|(\Omega)$ on $\mathcal{M}_n(\Omega)$. Using again [7, Prop. 12, p. 119] we obtain that

$$\psi \in \partial F_2(\overline{u}) \subset (BV(\Omega))^*$$

if and only if

$$\psi = -\mathfrak{Div}\,\overline{\lambda} \quad \text{and} \quad \overline{\lambda} \in \partial \tilde{F}_2(\nabla \overline{u}). \tag{3.2}$$

(Recall the definition of $-\mathfrak{Div}$ in (2.14).) From equation (3.2) it follows from the definition of ∂ that

$$\langle \overline{\lambda}, \mu - \nabla \overline{u}\rangle_{\mathcal{M}_n^*, \mathcal{M}_n} + |\nabla \overline{u}|(\Omega) \leq |\mu|(\Omega) \tag{3.3}$$

for all $\mu \in \mathcal{M}_n(\Omega)$. Here and in the following we denote the duality pairing of a space X and its dual X^* by angle-brackets $\langle \cdot, \cdot \rangle_{X, X^*}$. The necessary optimality condition (3.1) therefore reads as: $\overline{u} \in BV(\Omega)$ is a solution to (P) if and only if there exists $\overline{\lambda} \in \mathcal{M}_n^*(\Omega)$ satisfying (3.3) such that

$$\alpha \, \mathfrak{Div} \, \overline{\lambda} = K^* J(K\overline{u} - z). \tag{3.4}$$

We now reformulate the optimality system (3.3) and (3.4), to get a more convenient characterization of the solution $\overline{u}$ from which we can derive structural properties of $\overline{u}$. We first consider (3.3). Setting $\mu = 0$ and $\mu = 2\nabla \overline{u}$ in (3.3) we obtain

$$-\langle \overline{\lambda}, \nabla \overline{u} \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} \leq -|\nabla \overline{u}|(\Omega)$$

and

$$\langle \overline{\lambda}, \nabla \overline{u} \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} \leq |\nabla \overline{u}|(\Omega)$$

hence

$$\langle \overline{\lambda}, \nabla \overline{u} \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} = |\nabla \overline{u}|(\Omega).$$

Inserting this back into (3.3) we obtain

$$\langle \overline{\lambda}, \mu \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} \leq |\mu|(\Omega)| \tag{3.5}$$

for all $\mu \in \mathcal{M}_n(\Omega)$.

We claim that equality holds in (3.5) not only for $\mu = \nabla \overline{u}$, but also for $\mu = (\nabla \overline{u})_a$ and $\mu = (\nabla \overline{u})_s$. Suppose otherwise that strict inequality holds in at least one relation. Then using Lemma 1 we have:

$$|\nabla \overline{u}|(\Omega) = (|\nabla \overline{u}|)_a(\Omega) + (|\nabla \overline{u}|)_s(\Omega) > \langle \overline{\lambda}, (\nabla \overline{u})_a \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} + \langle \overline{\lambda}, (\nabla \overline{u})_s \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} = \langle \overline{\lambda}, \nabla \overline{u} \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} = |\nabla \overline{u}|(\Omega).$$

Thus, we cannot have a strict inequality in either case. We obtain:

$$\langle \overline{\lambda}, (\nabla \overline{u})_a \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} = |(\nabla \overline{u})_a|(\Omega) \tag{3.6}$$

and

$$\langle \overline{\lambda}, (\nabla \overline{u})_s \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} = |(\nabla \overline{u})_s|(\Omega). \tag{3.7}$$

We now show that we can replace the multiplier $\overline{\lambda} \in \mathcal{M}_n(\Omega)^*$ by a smoother function $\hat{\lambda}$. Here we closely follow Th. 2.4 of [4]. We consider the restriction of the functional $\mu \mapsto \langle \overline{\lambda}, \mu \rangle_{\mathcal{M}_n^*, \mathcal{M}_n}$ to $L_n^1(\Omega) \subset \mathcal{M}_n(\Omega)$

$$\hat{\lambda} = \overline{\lambda}|_{L_n^1(\Omega)}. \tag{3.8}$$

Since $\hat{\lambda}$ is a bounded linear functional on $L_n^1(\Omega)$, we have

$$\hat{\lambda} \in L_n^\infty(\Omega) \tag{3.9}$$

and, by (3.6) and (3.5):

$$\langle \hat{\lambda}, (\nabla \overline{u})_a \rangle_{L_n^\infty, L_n^1} = \sum_{i=1}^n \int_\Omega \hat{\lambda}_i(\mathbf{x}) \left(\frac{\partial \overline{u}}{\partial x_i}\right)_a (\mathbf{x}) \, d\mathbf{x} = \sum_{i=1}^n \int_\Omega \left|\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a (\mathbf{x})\right| d\mathbf{x} \tag{3.10}$$

$$\langle \hat{\lambda}, \mathbf{f} \rangle_{L_n^\infty, L_n^1} = \sum_{i=1}^n \int_\Omega \hat{\lambda}_i(\mathbf{x}) \, f_i(\mathbf{x}) \, d\mathbf{x} \le \sum_{i=1}^n \int_\Omega |f_i(\mathbf{x})| \, d\mathbf{x} \text{ for all } \mathbf{f} \in L_n^1(\Omega). \tag{3.11}$$

(Recall that we use the anisotropic norm (2.2) on $BV(\Omega)$.) From (3.11) we conclude that $|\hat{\lambda}_i(\mathbf{x})| \le 1$ almost everywhere on Ω. Consequently (3.10) implies that

$$\hat{\lambda}_i(\mathbf{x}) \in \text{sign}\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a (\mathbf{x}) \quad \text{almost everywhere on } \Omega \text{ for } i = 1, \ldots, n.$$

We now come back to the optimality condition (3.4). Since the right hand side of (3.4) is in $L^q(\Omega)$ we have $\mathfrak{Div}\,\overline{\lambda} \in L^q(\Omega)$. We shall now calculate the distributional divergence div $\hat{\lambda}$ of $\hat{\lambda}$ and we shall show that it coincides with $\mathfrak{Div}\,\overline{\lambda}$. Let $\mathcal{D}(\Omega)$ denote the space of all infinitely differentiable function with compact support in Ω and let $\mathcal{D}'(\Omega)$ denote its dual. Let $\mathcal{D}_n(\Omega)$ and $\mathcal{D}'_n(\Omega)$ denote the corresponding spaces for functions with values in $\mathbb{R}^n$. We have $\hat{\lambda} \in L_n^\infty(\Omega) \subset \mathcal{D}'_n(\Omega)$. Suppose $\varphi \in \mathcal{D}(\Omega)$. Then we have

$$\langle \text{div}\,\hat{\lambda}, \varphi \rangle_{\mathcal{D}', \mathcal{D}} = -\langle \hat{\lambda}, \nabla\varphi \rangle_{L_n^\infty, L_n^1} = -\langle \overline{\lambda}, \nabla\varphi \rangle_{\mathcal{M}_n^*, \mathcal{M}_n} = \langle \mathfrak{Div}\,\overline{\lambda}, \varphi \rangle_{BV^*, BV} = \langle \mathfrak{Div}\,\overline{\lambda}, \varphi \rangle_{L^q, L^p}. \tag{3.12}$$

The last equality holds because $\mathfrak{Div}\,\overline{\lambda} \in L^q(\Omega)$. Since φ was arbitrary in $\mathcal{D}(\Omega)$ and since $\mathcal{D}(\Omega)$ is dense in $L^p(\Omega)$ we find

$$\text{div}\,\hat{\lambda} = \mathfrak{Div}\,\overline{\lambda} \in L^q(\Omega).$$

Inserting this into (3.4) completes the proof. □

4. The one-dimensional case

Theorem 1. *Let $\Omega = (a, b)$ be a bounded open interval in $\mathbb{R}$, let $p \in (1, \infty)$, and suppose that the assumptions of Propositions 2 and 3 hold. Suppose moreover that the data $z \in Y$ are given such that, for every measurable set $E \subset (a, b)$ with $\text{meas}(E) > 0$, the equation*

$$K^* J(Ku - z) = 0 \text{ on } E \tag{4.1}$$

does not have a solution $u \in BV(a, b)$. Let $\overline{u}$ denote the solution to (P). Then we have

$$(\overline{u}')_a = 0.$$

Hence, $\overline{u}' = (\overline{u}')_s \in \mathcal{P}$, and $\overline{u}' = 0$ almost everywhere on (a, b).

Proof. We use the optimality condition (O_1). In the one-dimensional case, we have $\hat{\lambda} \in W^{1,q}(a, b)$ and

$$\alpha\,\hat{\lambda}' = K^* J(K\overline{u} - z) \in L^q(a, b). \tag{4.2}$$

Suppose that the absolutely continuous part $(\overline{u}')_a$ in the Lebesgue decomposition $\overline{u}' = (\overline{u}')_a + (\overline{u}')_s$ of $\overline{u}'$ does not vanish. Let us assume without loss of generality that $(\overline{u}')_a > 0$ on some measurable set E with $\text{meas}(E) > 0$. Then, by (O_2) we have $\hat{\lambda}(x) = 1$ on E and hence $\hat{\lambda}'(x) = 0$ almost everywhere on E (see Lem. 7.7, p. 145 in [9]). If we insert this into (4.2), we see that $\overline{u}$ solves (4.1) on some set of positive Lebesgue measure in contradiction to our assumption. Hence $(\overline{u}')_a = 0$ and thus $\overline{u}' = (\overline{u}')_s \in \mathcal{P}$. □

Remark 5. Theorem 1 states that under certain conditions on the data z, the measure $\overline{u}'$ is concentrated on a subset of (a, b) which is of Lebesgue measure 0. In this sense, we can say that $\overline{u}' = 0$ almost everywhere on (a, b), and thus "$\overline{u}$ is constant almost everywhere".

Remark 6. It is a common situation for inverse problems that the data z are given by some noisy measurement of "ideal" data $z_0 \in K(BV(a, b))$. Typically $K(BV(a, b))$ is dense in Y but does not coincide with Y and the perturbed data z are outside this range. Therefore the optimization problem

$$\text{minimize}\|Ku - z\|^2 \text{ over } u \in BV(a, b) \tag{4.3}$$

does not have a solution, and consequently the optimality condition

$$K^*J(Ku - z) = 0 \tag{4.4}$$

does not have a solution $u \in BV(a, b)$. Condition (4.1) can therefore be seen as a generalization of $z \notin K(BV(a, b))$.

We now investigate more closely two special situations which are important in practical applications. In the first case $K = \text{id}$, which corresponds to the problem of noise removal from a noise-corrupted signal. In the second case K is some smoothing integral operator, which corresponds to the problem of deblurring a blurred signal.

Proposition 4 (Denoising). *Suppose $Y = L^2(a, b)$, $p = q = 2$, and $K = \text{id}$. Assume moreover that $z \in L^2(a, b)$ is given such that for any $E \subset (a, b)$ with positive Lebesgue measure, z does not coincide with some function $u \in BV(a, b)$ on E. Let $\overline{u}$ denote the solution to (P). Then we have $(\overline{u}')_a = 0$, and hence $\overline{u}' = (\overline{u}')_s \in \mathcal{P}$.*

Proof. In this special situation, we have $K = K^* = \text{id}$, and also $J = \text{id}$ (the duality map of the Hilbert space $L^2(a, b)$). Condition (4.1) then reads as: for any $E \subset (a, b)$ of positive measure, $u - z = 0$ on E does not have a solution $u \in BV(a, b)$. This is the condition formulated in the Proposition. $\square$

In the deblurring case, we find that $\overline{u}'$ is a singular measure, under even weaker assumptions on the data z.

Proposition 5 (Deblurring). *Suppose $Y = L^2(a, b)$, $p = q = 2$, and*

$$Ku = \int_a^b k(x - \cdot)\, u(x)\, \mathrm{d}x,$$

with k analytic, and $k(x) = k(-x)$. Suppose also that K is injective. Assume moreover that

$$z \notin K(BV(a, b)). \tag{4.5}$$

Let $\overline{u}$ denote the solution to (P). Then we have $(\overline{u}')_a = 0$ and hence $\overline{u}' = (\overline{u}')_s \in \mathcal{P}$.

Proof. Under the above assumptions it is easy to see that K is selfadjoint on $L^2(a, b)$, and as in Proposition 4 $J = \text{id}$. Condition (4.1) then reads as: for any $E \subset (a, b)$ of positive measure, the problem

$$K(Ku - z) = 0 \text{ on } E$$

does not have a solution $u \in BV(a, b)$. Suppose this condition was not satisfied, *i.e.*, there exists $u \in BV(a, b)$ such that $K(Ku - z) = 0$ on E for some set E of positive measure. Since k is analytic, we find that $K(Ku - z)$ is an analytic function which is 0 on some set of positive measure. This, however, means $K(Ku - z) \equiv 0$ on (a, b), and by the injectivity of K we obtain $Ku - z \equiv 0$ on (a, b), *i.e.*, $z \in K(BV(a, b))$ in contradiction to our assumption. $\square$

5. THE TWO-DIMENSIONAL CASE

In dimension one, an important point in the proof of Theorem 1 is the fact that we have $\hat{\lambda}' \in L^q$, hence $\hat{\lambda} \in W^{1,q}$, and we can use Lemma 7.7 in [9] to conclude that the distributional derivative of $\hat{\lambda}$ is zero almost everywhere on the set where $\hat{\lambda}$ is constant. In dimension two or higher, we only have $\operatorname{div} \hat{\lambda} \in L^q(\Omega)$ by (O_1), which is weaker than $\hat{\lambda} \in W_2^{1,q}(\Omega)$. We therefore cannot use the same argument, and a closer inspection shows that the reasoning in [9] cannot easily be adapted to the present situation. In the two-dimensional situation we therefore get only a weaker result, claiming that the gradient $\nabla \overline{u}$ of the solution cannot have two non-zero components on any *open* subset of Ω. This corresponds to the observation in the introductory section that, at each point, the gradient of the solution is parallel to one of the coordinate directions.

Theorem 2. *Suppose $\Omega \subset \mathbb{R}^2$ is an open domain, $p \in (1, \infty)$, and suppose that the assumptions of Propositions 2 and 3 hold. Suppose moreover that the data $z \in Y$ are given such that for every open set $U \subset \Omega$ the equation*

$$K^* J(Ku - z) = 0 \text{ on } U \tag{5.1}$$

does not have a solution $u \in BV(\Omega)$. Let $\overline{u}$ denote the solution to (P). *Then, there is no open subset Ω' of Ω on which both components $\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a$, $i = 1, 2$, have constant, non-zero sign.*

Proof. We use the optimality condition (O_1). Suppose, we can find an open $\Omega' \subset \Omega$, where $\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a$ has constant sign for $i = 1, 2$. Suppose, without loss of generality, that both components are positive. Then we have, by (O_2), that $\hat{\lambda}_i = 1$ on Ω' and therefore $\operatorname{div} \hat{\lambda} = 0$ on Ω'. Thus, we have $K^* J(Ku - z) = 0$ on Ω', contradicting our assumption. $\qquad \square$

Remark 7. The following conjecture is the analogue to Lemma 7.7 in [9] for vectorfields with divergence in $L^q(\Omega)$. If the conjecture holds we are able to prove a stronger result.

Conjecture 1. *Let $\lambda \in L_2^\infty(\Omega)$, with $\operatorname{div} \lambda \in L^q(\Omega)$ for some $1 < q < \infty$. Suppose that λ is constant on some measurable set $E \subset \Omega$. Then $\operatorname{div} \lambda = 0$ almost everywhere on E.*

We consider again the solution $\overline{u}$ to problem (P). If we suppose that the data $z \in Y$ are given such that, for every measurable set $E \subset \Omega$ with $\operatorname{meas}(E) > 0$, the equation

$$K^* J(Ku - z) = 0 \text{ on } E \tag{5.2}$$

does not have a solution $u \in BV(\Omega)$, and under the hypothesis that Conjecture 1 holds we get the following result: Almost everywhere on Ω we have

$$\left(\frac{\partial \overline{u}}{\partial x_1}\right)_a = 0 \text{ or } \left(\frac{\partial \overline{u}}{\partial x_2}\right)_a = 0.$$

Remark 8. The result in Theorem 2 depends on the fact that $\hat{\lambda}_i \in \operatorname{sign}\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a$ which is only true for the anisotropic norm on $BV(\Omega)$ chosen in (2.4). For any other choice of the norm, this structural result does not hold in the above form. This is different from the situation in dimension one, where we have a unique vector norm.

As in dimension one, we consider the denoising and deblurring problem in more detail. The proofs are completely analogous to the proofs of Propositions 4 and 5 and are therefore omitted.

Proposition 6 (Denoising). *Suppose $Y = L^2(\Omega)$, $p = q = 2$, and $K = \operatorname{id}$. Assume moreover that $z \in L^2(\Omega)$ is given such that for any open subset $U \subset \Omega$, z does not coincide with some function $u \in BV(\Omega)$ on U. Let $\overline{u}$*

denote the solution to problem (P). *Then there is* no *open subset* Ω' *of* Ω *on which both components* $\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a$, $i = 1, 2$, *have constant, non-zero sign.*

Proposition 7 (Deblurring). *Suppose* $Y = L^2(a, b)$, $p = q = 2$, *and*

$$Ku = \int_\Omega k(x - \cdot) \, u(x) \, \mathrm{d}x,$$

with k *analytic, and* $k(x) = k(-x)$. *Suppose also that* K *is injective. Assume moreover that*

$$z \notin K(BV(\Omega)). \tag{5.3}$$

Let $\overline{u}$ *denote the solution to problem* (P). *Then, there is* no *open subset* Ω' *of* Ω *on which both components* $\left(\frac{\partial \overline{u}}{\partial x_i}\right)_a$, $i = 1, 2$, *have constant, non-zero sign.*

Remark 9. It is obvious how the two-dimensional results can be generalized to $n > 2$ dimensions. As for $n = 2$ we cannot have that all components of the gradient are not zero on any open subset of Ω.

REFERENCES

[1] R. Acar and C.R. Vogel, Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems* **10** (1994) 1217–1229.

[2] V. Barbu, Analysis and Control of Nonlinear Infinite Dimensional Systems. *Math. Sci. Engrg.* **190** (1993).

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

[4] G. Chavent and K. Kunisch, Regularization of linear least squares problems by total bounded variation. *ESAIM Control Optim. Calc. Var.* **2** (1997) 359–376.

[5] D. Dobson and O. Scherzer, Analysis of regularized total variation penalty methods for denoising. *Inverse Problems* **12** (1996) 601–617.

[6] D.C. Dobson and F. Santosa, Recovery of blocky images from noisy and blurred data. *SIAM J. Appl. Math.* **56** (1996) 1181–1192.

[7] I. Ekeland and T. Turnbull, Infinite-Dimensional Optimization and Convexity. *Chicago Lectures in Math.*, The University of Chicago Press, Chicago and London (1983).

[8] L. Evans and R. Gariepy, *Measure Theory and Fine Properties of Functions.* CRC Press, Boca Raton (1992).

[9] D. Gilbarg and N.S. Trudinger, Elliptic Partial Differential Equations of Second Order. *Grundlehren Math. Wiss.* **224** (1977).

[10] E. Giusti, Minimal Surfaces and Functions of Bounded Variation. *Monogr. Math.* **80** (1984).

[11] K. Ito and K. Kunisch, An active set strategy based on the augmented lagrantian formulation for image restauration. *RAIRO Modél. Math. Anal. Numér.* **33** (1999) 1–21.

[12] K. Ito and K. Kunisch, BV-type regularization methods for convoluted objects with edge-flat-grey scales. *Inverse Problems* **16** (2000) 909–928.

[13] M.Z. Nashed and O. Scherzer, Least squares and bounded variation regularization with nondifferentiable functionals. *Numer. Funct. Anal. Optim.* **19** (1998) 873–901.

[14] M. Nikolova, Local strong homogeneity of a regularized estimator. *SIAM J. Appl. Math.* (to appear).

[15] L. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithm. *Physica D* **60** (1992) 259–268.

[16] W. Rudin, *Real and Complex Analysis*, 3rd edn. McGraw-Hill, New York–St Louis–San Francisco (1987).

[17] C. Vogel and M. Oman, Iterative methods for total variation denoising. *SIAM J. Sci. Comp.* **17** (1996) 227–238.

[18] W.P. Ziemer, Weakly Differentiable Functions. *Grad. Texts in Math.* **120** (1989).