SpezialForschungsBereich F 32

Karl–Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz

Properties of L^1-TGV^2: The one-dimensional case

K. Bredies K. Kunisch T. Valkonen

SFB-Report No. 2011-006 March 2011

A–8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)

SFB sponsors:

- **Austrian Science Fund** (FWF)
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**

The final publication is available at www.sciencedirect.com
DOI 10.1016/j.jmaa.2012.08.053

Properties of L^1-TGV^2: The one-dimensional case

Kristian Bredies, Karl Kunisch, Tuomo Valkonen

Institute for Mathematics and Scientific Computing, Karl-Franzens University of Graz, Austria

Abstract

We study properties of solutions to second order Total Generalized Variation (TGV^2) regularized L^1-fitting problems in dimension one. Special attention is paid to the analysis of the structure of the solutions, their regularity and the effect of the regularization parameters.

31 March 2011 / Updated: 14 September 2012

Keywords: Total generalized variation, robust data fitting, regularization techniques

1. The L^1-TGV^2 functional

In this work we study the variational problem

$$\min_u \|u-f\|_1 + \mathrm{TGV}^2_{\vec{\alpha}}(u), \tag{1.1}$$

where f is the input data, u denotes a possible solution, and $\vec{\alpha} = (\beta, \alpha) > 0$ componentwise stands for a vector-valued regularization parameter. The precise definition of $\mathrm{TGV}^2_{\vec{\alpha}}$ will be given below. For the moment it suffices to know that it is a flexible regularization functional which adapts to first and second order smoothness of the data. The $\mathrm{TGV}^2_{\vec{\alpha}}$-functional is a special case of the $\mathrm{TGV}^k_{\vec{\alpha}}$-functional, where $k \geq 2$, which was introduced in [1]. In [1] basic analytical properties of $\mathrm{TGV}^k_{\vec{\alpha}}$ and numerical results with an L^2 data-fitting term for the cases $k=2$ and $k=3$ are provided. One way of interpreting $\mathrm{TGV}^k_{\vec{\alpha}}$, consists in realizing that it regularizes independently on different regularity scales of the function that it is applied to. Compared to the TV-functional [2] we recall that constant functions are in the kernel of TV, while polynomials of degree strictly less than k constitute the kernel of $\mathrm{TGV}^k_{\vec{\alpha}}$, see [1].

Since its introduction, $\mathrm{TGV}^k_{\vec{\alpha}}$ functionals have proved to be effective in diverse mathematical imaging problems, including denoising [1], reconstruction of magnetic resonance images from highly undersampled data [3], deconvolution [4] and fusion of stereographical data [5]. This success motivates, besides inherent mathematical interest, to further investigate analytical properties of these functionals. Within the present paper we focus on the case $k=2$ in spatial dimension one and aim at getting detailed insight on structural properties of the solution of (1.1). These include regularity properties, higher degree staircasing effects, and features of the solution that are inherited from the data f. We expect that generalizations to $k \geq 3$ are possible.

The advantages and differences of the L^1 data-fitting term over the L^2-performance criterion are well reported in the literature. From the point of view of robust statistics L^1 should be preferred over an L^2 fidelity term, since the latter magnifies errors introduced by outliers. Geometric features and scale separation properties of the L^1 criterion are reported in e.g. [6, 7, 8, 9]. All these papers address the L^1-TV, as opposed to the TGV case, which is in the focus of the present paper. The numerical realization of the L^1-TV problem is typically considered in

Email addresses: kristian.bredies@uni-graz.at (Kristian Bredies), karl.kunisch@uni-graz.at (Karl Kunisch), tuomo.valkonen@iki.fi (Tuomo Valkonen)

the discrete formulation with L^1 replaced by ℓ^1. Among the techniques that were analyzed we mention linear programming, generalized reweighted least-squares, and splitting techniques, see e.g. [10, 11], and semi-smooth Newton methods [12].

We point out two other papers whose focus is, at least in part, the investigation of the effect of higher-order regularization in one dimension. In [13], using Fenchel duality, explicit solutions for TV-based higher-order penalization are derived. In [14], the authors provide a detailed analysis of a non-convex second-order functional and, in particular, it is proved that staircasing does not occur. The L^1-TGV^2 model considered in the current paper also provides solutions without zero-order staircasing.

The subsequent sections are structured as follows. Section 2 contains notation that will be used throughout the paper as well as a summary of useful facts on functions of bounded variation with special attention paid to the one-dimensional case. The precise problem formulation involving the TGV^2-functional is contained in Section 3. Introducing a set-valued generalization of the sign operation that is applicable to Radon measures allows an elegant description of necessary and sufficient optimality conditions. In Section 4 monotonicity and staircasing properties of the solution to (1.1), as well as its jump set, are analyzed. It is shown that zero degree staircasing, well-known for the solutions of BV-regularized problems, cannot occur. Instead, it is replaced by staircasing of degree one for solutions to (1.1) which is less detectable by the human eye. The optimality conditions allow to argue that certain regularity properties of the data f, like absolute and Lipschitz continuity, as well as piecewise affinity are inherited by the solution to (1.1). This is treated in Section 5. Section 6 focuses on the effect of the regularization parameter on the solution. The asymptotic behavior of the solution and monotonicity properties of the performance and complexity summand in the cost functional are proved. Further threshold bounds on the solution in terms of the regularization parameters are obtained. The paper concludes with examples illustrating these bounds.

2. Notation and preliminaries

2.1. Measures and functions

For a function $u : \Omega \to \mathbb{R}$, we denote by $|u|$ the pointwise absolute value: $|u|(x) := |u(x)|$.

A function $u : (a,b) \to \mathbb{R}$ is said to be *piecewise affine*, if there are finitely many disjoint (open) intervals $I_1, \ldots, I_N$ such that $(a,b) = \bigcup_{i=1}^N \bar{I}_i$, and u is affine on each I_i. Here $\bar{I}_i$ denotes the relative closure of I_i in (a,b).

Let $\Omega \subset \mathbb{R}$ be a Borel set and let $\mathscr{M}(\Omega, \mathbb{R}^n)$ denote the space of (vector-valued) Radon measures on Ω. The total variation measure of $\mu \in \mathscr{M}(\Omega)$ is denoted $|\mu|$, and we define the norm $\|\mu\|_{\mathscr{M}(\Omega)} := |\mu|(\Omega)$, see e.g. [15].

For each $\mu \in \mathscr{M}(\Omega, \mathbb{R}^n)$ there exists a polar decomposition $\mu = \mathrm{sgn}(\mu)|\mu|$ with $\mathrm{sgn}(\mu) \in L^\infty(\Omega, |\mu|)$ and $\|\mathrm{sgn}(\mu)\|_\infty \leq 1$. The notation $\mu \ll \nu$ denotes the fact that the measure μ is absolutely continuous with respect to the measure ν.

We denote by $\mathscr{L}^m$ the Lebesgue measure on $\mathbb{R}^m$, while $\mathscr{H}^k$ denotes the k-dimensional Hausdorff measure on a suitable ambient space. The Dirac measure concentrated at x is denoted δ_x. The restriction of a Radon measure μ to a Borel set A is denoted $\mu \llcorner A$, where $(\mu \llcorner A)(B) := \mu(A \cap B)$.

Finally, we like to recall the definition of the Radon norm for distributions. A distribution u on Ω is a Radon measure (in the sense that there is a $\mu \in \mathscr{M}(\Omega)$ such that $\int_\Omega v \, \mathrm{d}\mu = \langle u, v \rangle$ for all $v \in \mathscr{C}_c^\infty(\Omega)$) if and only if

$$\|u\|_{\mathscr{M}} = \sup \left\{ \langle u, v \rangle \;\middle|\; v \in \mathscr{C}_c^\infty(\Omega),\ \|v\|_\infty \leq 1 \right\} \tag{2.1}$$

is finite. In particular, if finite, the supremum coincides with the norm in $\mathscr{M}(\Omega)$.

Therefore, we have, for distributions u and w, the identities

$$\|Dw\|_{\mathscr{M}} = \sup \left\{ \langle w, v' \rangle \;\middle|\; v \in \mathscr{C}_c^\infty(\Omega),\ \|v\|_\infty \leq 1 \right\} \tag{2.2}$$

and

$$\|Du - w\|_{\mathscr{M}} = \sup \left\{ \langle w, \omega \rangle + \langle u, \omega' \rangle \;\middle|\; \omega \in \mathscr{C}_c^\infty(\Omega),\ \|\omega\|_\infty \leq 1 \right\} \tag{2.3}$$

where the value ∞ is possibly attained.

If u can be identified with an element in the dual space $\mathscr{C}_0^k(\Omega)^*$, then by density the set of test functions $\mathscr{C}_c^\infty(\Omega)$ can be replaced by $\mathscr{C}_0^k(\Omega)$.

2.2. Functions of bounded variation

Following, e.g., [16], a function $u \in L^1(\Omega)$ on a non-empty open set $\Omega \subset \mathbb{R}$ is said to be of *bounded variation*, denoted $u \in \mathrm{BV}(\Omega)$, if the distributional derivative Du is a (vector-valued) Radon measure. In other words

$$\int_\Omega u'\phi \, \mathrm{d}x = -\int_\Omega \phi \, \mathrm{d}Du, \quad \text{for all} \quad \phi \in \mathscr{C}_c^\infty(\Omega).$$

In $\mathrm{BV}(\Omega)$ we define the norm $\|u\|_1 + \|Du\|_{\mathscr{M}}$ and the BV-seminorm by $\mathrm{TV}(u) = \|Du\|_{\mathscr{M}}$. A sequence $\{u^i\}_{i=0}^\infty$ in $\mathrm{BV}(\Omega)$ converges strongly to $u \in \mathrm{BV}(\Omega)$ if both $\|u^i - u\|_{L^1(\Omega)} \to 0$ and $\|Du^i - Du\|_{\mathscr{M}(\Omega)} \to 0$. Weak convergence is defined as $u^i \to u$ strongly in $L^1(\Omega)$ and $Du^i \to Du$ weakly* in $\mathscr{M}(\Omega, \mathbb{R}^m)$.

In the following, let $\Omega = (a, b)$ and $u \in \mathrm{BV}(\Omega)$. Recall that $x \in \Omega$ is called a *Lebesgue point* if there exists a $\tilde{u}(x)$ such that

$$\lim_{\rho \searrow 0} \frac{1}{2\rho} \int_{x-\rho}^{x+\rho} |\tilde{u}(x) - u(y)| \, \mathrm{d}y = 0.$$

The set of points where this limit does not exist is called the *approximate discontinuity set*, denoted by S_u. In the one-dimensional case, the *approximate left and right limits*, $u^-(x)$ and $u^+(x)$ exist for every $x \in \Omega$ and are defined by satisfying

$$\lim_{\rho \searrow 0} \frac{1}{\rho} \int_x^{x+\rho} |u^+(x) - u(y)| \, \mathrm{d}y = 0 \text{ and } \lim_{\rho \searrow 0} \frac{1}{\rho} \int_{x-\rho}^{x} |u^-(x) - u(y)| \, \mathrm{d}y = 0,$$

respectively. The set of points x where $u^-(x) \neq u^+(x)$, which is called the *jump set* J_u of u, is known to be at most countable and to coincide with S_u.

We can decompose the distributional derivative of a $u \in \mathrm{BV}(\Omega)$ as $Du = D^a u + D^j u + D^c u$, where $D^a u = u' \mathscr{L}^1$ is the *absolutely continuous part*, with u' the approximate differential, $D^j u$ represents the *jump part* which can be represented as

$$D^j u = (u^+ - u^-) \mathscr{H}^0 \llcorner J_u,$$

and $D^c u$ is the *Cantor part* which vanishes on any Borel set σ-finite with respect to $\mathscr{H}^0$, i.e., at most countable sets. The singular parts of D are denoted by $D^s = D^j + D^c$. For $u \in \mathscr{C}^1(\bar{\Omega})$ the approximate differential coincides with the common notion of a derivative.

For $u \in \mathrm{BV}(\Omega)$ we will be mostly working with *good representatives* as defined in [16, Theorem 3.28]. These are functions $\tilde{u} : \Omega \to \mathbb{R}$ which are continuous outside J_u and satisfy for some unique $c_u \in \mathbb{R}$ that

$$c_u + Du((a, t)) \le \tilde{u}(t) \le c_u + Du((a, t]) \quad \text{for all} \quad t \in (a, b).$$

They are uniquely determined on $\Omega \setminus J_u$ and indeed representatives of the equivalence class u. In particular, u^- and u^+ are good representatives which are moreover left and right continuous, respectively.

3. Problem formulation and optimality conditions

Assumption 3.1. Throughout this paper, unless otherwise stated, we assume that $\Omega = (a, b) \subset \mathbb{R}$.

We write problem (1.1) as

$$\min_{u \in \mathrm{BV}(\Omega)} F(u), \quad \text{where} \quad F(u) := \|f - u\|_{L^1(\Omega)} + \mathrm{TGV}^2_{\vec{\alpha}}(u) \tag{P}$$

for $\vec{\alpha} = (\beta, \alpha) > 0$ componentwise and

$$\mathrm{TGV}^2_{\vec{\alpha}}(u) := \sup \left\{ \int_\Omega u v'' \, \mathrm{d}x \;\middle|\; v \in \mathscr{C}_c^2(\Omega), \ \|v\|_\infty \le \beta, \ \|v'\|_\infty \le \alpha \right\}, \tag{TGV$^{\text{sup}}$}$$

also called the *predual* or *supremum definition* of $\mathrm{TGV}^2_{\vec{\alpha}}(u)$.

Here we prefer to work with the *minimum characterization* of $\mathrm{TGV}^2_{\vec{\alpha}}(u)$, expressed as

$$\mathrm{TGV}^{2,\min}_{(\beta,\alpha)}(u) := \min_{w\in \mathrm{BV}(\Omega)} \big(\alpha\|Du-w\|_{\mathscr{M}(\Omega)} + \beta\|Dw\|_{\mathscr{M}(\Omega)}\big). \tag{TGV$^{\min}$}$$

Observe that the minimization problem in (TGV$^{\min}$) is just L^1-TV for u', as the singular part $D^s u$ cannot be approximated by w keeping $\|Dw\|_{\mathscr{M}}$ bounded. In [1] it was shown that for $u \in \mathscr{C}^\infty(\Omega)$, we have

$$\mathrm{TGV}^{2,\min}_{\vec{\alpha}}(u) = \mathrm{TGV}^2_{\vec{\alpha}}(u).$$

In the following, we will prove this equivalence for general $u \in L^1(\Omega)$ along with showing the equivalence of $\|\cdot\|_{\mathrm{BV}(\Omega)}$ to the norm

$$\|\cdot\|_{\mathrm{BGV}^2_{\vec{\alpha}}} := \|\cdot\|_{L^1(\Omega)} + \mathrm{TGV}^2_{\vec{\alpha}}.$$

This allows us to reformulate (P) as a minimization problem in $\mathrm{BV}(\Omega)^2$:

$$\min_{(u,w)\in \mathrm{BV}(\Omega)^2} \|f-u\|_{L^1(\Omega)} + \alpha\|Du-w\|_{\mathscr{M}} + \beta\|Dw\|_{\mathscr{M}}. \tag{P^2}$$

The existence of solutions to (P^2) and, consequently, to (P) follows from the fact that a minimizing sequence $\{(u^n, w^n)\}$ has to be bounded in $\mathrm{BV}(\Omega)^2$ and converges, by compactness, subsequentially to some element (u^*, w^*) in $L^1(\Omega)^2$ which is a minimizer by lower semi-continuity of the objective functional.

Proposition 3.2. *For $u \in L^1(\Omega)$ the supremum definition* (TGV$^{\sup}$) *and the minimum characterization* (TGV$^{\min}$) *coincide, that is*

$$\min_{w\in \mathrm{BV}(\Omega)} \alpha\|Du-w\|_{\mathscr{M}} + \beta\|Dw\|_{\mathscr{M}} = \mathrm{TGV}^2_{\vec{\alpha}}(u) = \sup\left\{\int_\Omega u v''\,\mathrm{d}x \;\middle|\; v \in \mathscr{C}^2_c(\Omega),\ \|v\|_\infty \le \beta,\ \|v'\|_\infty \le \alpha\right\}.$$

Proof. The verification of this claim will be based on Fenchel-Rockafellar duality. As (TGV$^{\min}$) possesses a structure which is similar to an infimal convolution, it comes as no surprise that an equivalent characterization involves multiple constraints associated with the closed predual balls with respect to $\|\cdot\|_{\mathscr{M}(\Omega)}$. In order to employ duality, first observe that the supremum in (TGV$^{\sup}$) can also be written as the negative infimum

$$\mathrm{TGV}^2_{\vec{\alpha}}(u) = -\inf\left\{-\int_\Omega u v''\,\mathrm{d}x \;\middle|\; v \in \mathscr{C}^2_c(\Omega),\ \|v\|_\infty \le \beta,\ \|v'\|_\infty \le \alpha\right\}. \tag{3.1}$$

Moreover, by density of $\mathscr{C}^2_c(\Omega)$ in $\mathscr{C}^2_0(\Omega)$ with respect to the $\mathscr{C}^2$-norm, $\mathscr{C}^2_c(\Omega)$ in (3.1) can be replaced by $\mathscr{C}^2_0(\Omega)$. We therefore introduce $X = \mathscr{C}^2_0(\Omega)$, $Y = \mathscr{C}^1_0(\Omega)$ and the operator $\Lambda : v \mapsto v'$, for which $\Lambda \in \mathscr{L}(X, Y)$. Defining furthermore

$$F_1 : X \to (-\infty,\infty], \quad F_1(v) = I_{\{\|\cdot\|_\infty \le \beta\}}(v),$$
$$F_2 : Y \to (-\infty,\infty], \quad F_2(\omega) = I_{\{\|\cdot\|_\infty \le \alpha\}}(\omega) - \int_\Omega u\omega'\,\mathrm{d}x,$$

the infimum in (3.1) can be expressed as

$$\inf_{v\in X} F_1(v) + F_2(\Lambda v).$$

According to [17], the Fenchel-Rockafellar duality formula is valid if

$$Y = \bigcup_{\lambda\ge 0} \lambda\big(\mathrm{dom}(F_2) - \Lambda\,\mathrm{dom}(F_1)\big),$$

where $\mathrm{dom}(F_1)$ and $\mathrm{dom}(F_2)$ denote the effective domains of F_1 and F_2, respectively, i.e., the set where F_1 (resp. F_2) admits finite values.

Since each $\omega \in Y$ can be written as $\omega = \lambda(\lambda^{-1}\omega)$ with $\lambda > 0$ such that $\|\lambda^{-1}\omega\|_\infty \le \alpha$ and $0 \in \mathrm{dom}(F_1)$, this is immediately clear. Consequently, we know that

$$\min_{w\in Y^*} F_1^*(-\Lambda^* w) + F_2^*(w) = -\inf_{v\in X} F_1(v) + F_2(\Lambda v) = \mathrm{TGV}_{\vec{\alpha}}^2(u).$$

In particular, the infimum on the left is attained. Computing $F_1^*(-\Lambda^* w)$ gives

$$F_1^*(-\Lambda^* w) = \sup\left\{\langle w, -v'\rangle \;\middle|\; v \in \mathscr{C}_0^2(\Omega),\ \|v\|_\infty \le \beta\right\} = \beta\|Dw\|_{\mathscr{M}}$$

according to (2.2) and noting that $-\Lambda^* w$ can be interpreted as an element of $\mathscr{C}_0^2(\Omega)^*$. Likewise, (2.3) gives

$$F_2^*(w) = \sup\left\{\langle w, \omega\rangle + \langle u, \omega'\rangle \;\middle|\; \omega \in \mathscr{C}_0^1(\Omega),\ \|\omega\|_\infty \le \alpha\right\} = \alpha\|Du - w\|_{\mathscr{M}}.$$

These considerations yield the desired identity. □

Lemma 3.3. *There exist constants $0 < c < C < \infty$ such that for $u \in L^1(\Omega)$, we have*

$$c\big(\|u\|_{L^1(\Omega)} + \mathrm{TV}(u)\big) \le \|u\|_{L^1(\Omega)} + \mathrm{TGV}_{\vec{\alpha}}^2(u) \le C\big(\|u\|_{L^1(\Omega)} + \mathrm{TV}(u)\big).$$

Proof. The inequality

$$\|u\|_{L^1(\Omega)} + \mathrm{TGV}_{\alpha}^2(u) \le \max(1,\alpha)\big(\|u\|_{L^1(\Omega)} + \mathrm{TV}(u)\big)$$

is trivial: By Proposition 3.2, we can employ the minimum characterization ($\mathrm{TGV}^{\min}$) and take $w = 0$.

In order to complete the proof we have to show

$$c\big(\|u\|_{L^1(\Omega)} + \mathrm{TV}(u)\big) \le \|u\|_{L^1(\Omega)} + \mathrm{TGV}_{\vec{\alpha}}^2(u) \tag{3.2}$$

for some $c > 0$. We may assume that $\|Du\|_{\mathscr{M}(\Omega)} < \infty$, since otherwise the claim is trivial, both sides of the inequality being infinite. We begin by showing that, for some constant $C_1 = C_1(\Omega) > 0$,

$$\|Du\|_{\mathscr{M}(\Omega)} \le C_1\big(\|Du - \bar{w}\|_{\mathscr{M}(\Omega)} + \|u\|_{L^1(\Omega)}\big), \quad \text{for all} \quad \bar{w} \in \mathbb{R}. \tag{3.3}$$

Indeed, let us take $v(x) := \bar{w}x + h$ for some $h \in \mathbb{R}$ such that $\int_\Omega v = \int_\Omega u$. By the continuity of the differential operator $D : v \mapsto v'$ on affine functions, there exists a constant $C_2 = C_2(\Omega)$ such that $\|Dv\|_{L^1(\Omega)} \le C_2\|v\|_{L^1(\Omega)}$. It follows that

$$\begin{aligned}\|Du\|_{\mathscr{M}(\Omega)} &\le \|D(u-v)\|_{\mathscr{M}(\Omega)} + \|Dv\|_{L^1(\Omega)}\\ &\le \|D(u-v)\|_{\mathscr{M}(\Omega)} + C_2\|v\|_{L^1(\Omega)}\\ &\le \|D(u-v)\|_{\mathscr{M}(\Omega)} + C_2\|u-v\|_{L^1(\Omega)} + C_2\|u\|_{L^1(\Omega)}.\end{aligned}$$

Applying the Poincaré inequality [16, p. 152] to the middle term in the last expression, where we observe that $\int(u-v) = 0$ by construction of v, we obtain for a constant C_1 independent of u

$$\|Du\|_{\mathscr{M}(\Omega)} \le C_1\big(\|D(u-v)\|_{\mathscr{M}(\Omega)} + \|u\|_{L^1(\Omega)}\big).$$

Since $Dv = \bar{w}$, we may deduce (3.3).

Next we take $w \in \mathrm{BV}(\Omega)$ and let $\bar{w} := (b-a)^{-1}\int_\Omega w(x)\,\mathrm{d}x$. Then another application of the Poincaré inequality shows that there is a constant $C_3 = C_3(\Omega, \vec{\alpha})$ such that

$$\begin{aligned}\|Du - \bar{w}\|_{\mathscr{M}(\Omega)} &\le \|Du - w\|_{\mathscr{M}(\Omega)} + \|w - \bar{w}\|_{L^1(\Omega)}\\ &\le C_3\big(\alpha\|Du - w\|_{\mathscr{M}(\Omega)} + \beta\|Dw\|_{\mathscr{M}(\Omega)}\big).\end{aligned} \tag{3.4}$$

Combining (3.3) and (3.4) and taking the infimum over $w \in \mathrm{BV}(\Omega)$ now yields (3.2) by Proposition 3.2, concluding the proof. □

For stating optimality conditions based on subdifferential calculus, let us study the subdifferential of the L^1-norm and the norm in $\mathscr{M}(\Omega)$. For this purpose we need the following generalization of the sign function.

Definition 3.4. Let $\mu \in \mathscr{M}(\Omega)$. Then, $\operatorname{sgn}(\mu)$ denotes the unique element in $L^\infty(\Omega, |\mu|)$ for which $\mu = \operatorname{sgn}(\mu)|\mu|$. Moreover, the *set-valued sign* is defined as

$$\operatorname{Sgn}(\mu) = \{v \in L^\infty(\Omega) \cap L^\infty(\Omega, |\mu|) \,\big|\, \|v\|_\infty \le 1,\ \|v\|_{\infty,|\mu|} \le 1, v = \operatorname{sgn}(\mu),\ |\mu|\text{-almost everywhere}\},$$

with $\|v\|_{\infty,|\mu|}$ denoting the $|\mu|$-essential supremum of $|v|$.

For $u \in L^1(\Omega)$, we moreover define $\operatorname{Sgn}(u) = \operatorname{Sgn}(u\mathscr{L}^1)$.

It is obvious that if $u \in L^1(\Omega)$, then $v \in L^\infty(\Omega)$ belongs to $\operatorname{Sgn}(u)$ if and only if $v(t) = u(t)/|u(t)|$ almost everywhere in $\{u \neq 0\}$ and $v(t) \in [-1,1]$ almost everywhere in $\{u = 0\}$. Hence, the set-valued sign of $\mu \in \mathscr{M}(\Omega)$ can be regarded as the generalization of the sign to Radon measures.

Having this notion, the subgradient of the norm in $L^1(\Omega)$ and $\mathscr{M}(\Omega)$ can be characterized, for the latter at least for predual elements.

Lemma 3.5. *The following identities hold:*

(i) If $u \in L^1(\Omega)$, then $\partial\|\cdot\|_1(u) = \operatorname{Sgn}(u)$.

(ii) If $\mu \in \mathscr{M}(\Omega)$, then $\partial\|\cdot\|_{\mathscr{M}}(\mu) \cap \mathscr{C}_0(\Omega) = \operatorname{Sgn}(\mu) \cap \mathscr{C}_0(\Omega)$.

Proof. For the first part, note that from subdifferential calculus, $\omega \in L^\infty(\Omega)$ is in $\partial\|\cdot\|_1(u)$ if and only if

$$\|\omega\|_\infty \le 1 \text{ and } \int_\Omega \omega u \,\mathrm{d}x = \int_\Omega |u| \,\mathrm{d}x.$$

The latter expression is equivalent to $\int_{\{u\neq 0\}} \Big(\frac{u}{|u|} - \omega\Big)|u| \,\mathrm{d}x = 0$. Consequently $\omega = \frac{u}{|u|}$ almost everywhere in $\{u \neq 0\}$, and hence the equivalence holds as stated.

For the second part, recall that for a given $\mu \in \mathscr{M}(\Omega)$, $v \in \mathscr{C}_0(\Omega)$ implies $v \in L^\infty(\Omega) \cap L^\infty(\Omega, |\mu|)$ with $\|v\|_{\infty,|\mu|} \le \|v\|_\infty$. Now, $v \in \mathscr{C}_0(\Omega)$ satisfies

$$v \in \partial\|\cdot\|_{\mathscr{M}}(\mu) \cap \mathscr{C}_0(\Omega) \quad \text{if and only if} \quad \|v\|_\infty \le 1 \quad \text{and} \quad \langle \mu, v\rangle = \|\mu\|_{\mathscr{M}}.$$

By the decomposition $\mu = \operatorname{sgn}(\mu)|\mu|$ and $\|\mu\|_{\mathscr{M}} = \int_\Omega 1 \,\mathrm{d}|\mu|$, the latter is equivalent to

$$\|v\|_\infty \le 1 \qquad \text{and} \qquad \int_\Omega (\operatorname{sgn}(\mu) - v) \,\mathrm{d}|\mu| = 0$$

and this, in turn, to $v = \operatorname{sgn}(\mu)$, $|\mu|$-almost everywhere. Therefore, the characterization holds as stated. □

Proposition 3.6. *An element $u \in \mathrm{BV}(\Omega)$ is a minimizer for* (P) *and a $w \in \mathrm{BV}(\Omega)$ minimizes* ($\mathrm{TGV}^{\min}$) *for this u if and only if there exists a $v \in H_0^2(\Omega)$ such that*

$$v'' \in \operatorname{Sgn}(f-u), \tag{O_f}$$
$$-v' \in \alpha \operatorname{Sgn}(Du - w), \tag{O_α}$$
$$v \in \beta \operatorname{Sgn}(Dw). \tag{O_β}$$

Proof. We will show that the maximization problem

$$\max\left\{\int_\Omega f v'' \,\mathrm{d}x \,\middle|\, v \in H_0^2(\Omega),\ \|v\|_\infty \le \beta,\ \|v'\|_\infty \le \alpha,\ \|v''\|_\infty \le 1\right\} \tag{P'}$$

can be regarded as the predual problem for (P^2) and derive the optimality conditions from Fenchel-Rockafellar duality. First, note that (P′) has a solution $v^* \in H_0^2(\Omega)$ since the functional to maximize is weakly continuous and the constraints correspond to a non-empty, convex, closed and bounded subset of $H_0^2(\Omega)$. Hence, writing the maximum in (P′) is justified.

For the purpose of establishing Fenchel-Rockafellar duality, we introduce

$$X = H_0^2(\Omega) \times H_0^1(\Omega), \qquad Y = H_0^1(\Omega) \times L^2(\Omega),$$

and the linear and continuous mapping $\Lambda : X \to Y$ according to $\Lambda(v,\omega) = (\omega + v', \omega')$. Furthermore, let

$$F_1 : X \to (-\infty, \infty], \quad F_1(v,\omega) = I_{\{\|\cdot\|_\infty \le \beta\}}(v) + I_{\{\|\cdot\|_\infty \le \alpha\}}(\omega),$$

$$F_2 : Y \to (-\infty, \infty], \quad F_2(\phi,\psi) = I_{\{0\}}(\phi) + \int_\Omega f\psi \,\mathrm{d}x + I_{\{\|\cdot\|_\infty \le 1\}}(\psi).$$

It is easy to see that (P′) is equivalent to

$$\max\ (\mathrm{P}') = -\inf_{(v,\omega)\in X} F_1\big((v,\omega)\big) + F_2\big(\Lambda(v,\omega)\big).$$

To employ Fenchel-Rockafellar duality in this situation, we again establish the sufficient condition

$$Y = \bigcup_{\lambda \ge 0} \lambda\big(\mathrm{dom}(F_2) - \Lambda\,\mathrm{dom}(F_1)\big). \tag{3.5}$$

Let $(\phi,\psi) \in Y$ be given. In order to obtain the desired representation of this part, we have to "split off" a suitable affine part from ψ. Therefore, we choose $\psi_0 = h_0 + h_1 x$ with $h_0, h_1 \in \mathbb{R}$ such that

$$\int_\Omega \psi_0(x)\,\mathrm{d}x = \int_\Omega \psi(x)\,\mathrm{d}x, \qquad \int_\Omega x\psi_0(x)\,\mathrm{d}x = \int_\Omega x\psi(x) + \phi(x)\,\mathrm{d}x$$

is satisfied (this linear system of equations for (h_0, h_1) can easily seen to be uniquely solvable). Furthermore, we construct

$$\omega(x) = \int_a^x (\psi_0 - \psi)(y)\,\mathrm{d}y, \qquad v(x) = -\int_a^x (\phi + \omega)(y)\,\mathrm{d}y.$$

Note that $\omega \in H_0^1(\Omega)$: Indeed, $-\omega' = \psi - \psi_0 \in L^2(\Omega)$, $\omega(a) = 0$ by construction and

$$\omega(b) = \int_a^b (\psi_0 - \psi)(x)\,\mathrm{d}x = 0.$$

Likewise we find $v \in H_0^2(\Omega)$. In fact, $-v' = \omega + \phi \in H_0^1(\Omega)$, $v(a) = 0$, and by Fubini's theorem it follows that

$$\begin{aligned} v(b) &= -\int_a^b \omega(x) + \phi(x)\,\mathrm{d}x = \int_a^b \int_a^x (\psi - \psi_0)(y)\,\mathrm{d}y - \phi(x)\,\mathrm{d}x \\ &= \int_a^b \int_y^b 1\,\mathrm{d}x\,(\psi - \psi_0)(y)\,\mathrm{d}y - \int_a^b \phi(x)\,\mathrm{d}x \\ &= \int_a^b (b-x)(\psi - \psi_0)(x) - \phi(x)\,\mathrm{d}x \\ &= \int_a^b x\psi_0(x)\,\mathrm{d}x - \int_a^b x\psi(x) + \phi(x)\,\mathrm{d}x = 0. \end{aligned}$$

Therefore, $(v,\omega) \in X$ with

$$(\phi,\psi) = (0,\psi_0) - (\omega + v', \omega') = (0,\psi_0) - \Lambda(v,\omega).$$

By choosing $\lambda > 0$ appropriately, we can now achieve that

$$\|\lambda^{-1}\psi_0\|_\infty \le 1, \qquad \|\lambda^{-1}\omega\|_\infty \le \alpha, \qquad \|\lambda^{-1}v\|_\infty \le \beta,$$

and since $\lambda^{-1}\Lambda(v,\omega) = \Lambda(\lambda^{-1}v, \lambda^{-1}\omega)$, the representation

$$(\phi,\psi) = \lambda\big(\underbrace{(0,\lambda^{-1}\psi_0)}_{\in \mathrm{dom}(F_2)} - \Lambda\underbrace{(\lambda^{-1}v,\lambda^{-1}\omega)}_{\in \mathrm{dom}(F_1)}\big).$$

Since $(\phi,\psi) \in Y$ was arbitrary, (3.5) is established.

Therefore, we have

$$\Big(\min_{(v,\omega)\in X} F_1(v,\omega) + F_2\big(\Lambda(v,\omega)\big)\Big) + \Big(\min_{(w,u)\in Y^*} F_1^*\big(-\Lambda^*(w,u)\big) + F_2^*(w,u)\Big) = 0,$$

in particular the minimum is attained at some $(w^*,u^*) \in Y^*$. Interpreting $(\phi,\psi) \in H_0^2(\Omega)^* \times H_0^1(\Omega)^* = X^*$ as distributions of order 1 and 0, respectively, the functional dual to F_1 can be expressed as

$$F_1^*\big((\phi,\psi)\big) = \sup_{\substack{(v,\omega)\in X,\\ \|v\|_\infty \le \beta,\ \|\omega\|_\infty\le\alpha}} \langle \phi,\, v\rangle + \langle \psi,\, \omega\rangle = \alpha\|\psi\|_{\mathcal{M}} + \beta\|\phi\|_{\mathcal{M}}$$

by virtue of (2.1). Noting that $-\Lambda^*(w,u) = (Dw, Du - w)$ in the distributional sense, it follows

$$F_1^*\big(-\Lambda^*(w,u)\big) = \alpha\|Du - w\|_{\mathcal{M}} + \beta\|Dw\|_{\mathcal{M}}.$$

Likewise, we deduce

$$F_2^*\big((w,u)\big) = \sup_{\substack{(\phi,\psi)\in Y,\\ \phi=0,\ \|\psi\|_\infty\le 1}} \langle \phi,\, w\rangle + \int_\Omega (u-f)\psi \,\mathrm{d}x = \|f-u\|_1$$

leading to max (P′) = min (P) as claimed. Moreover, the optimality conditions can be expressed in terms of subgradients: A primal-dual pair $\big((w,u),(v,\omega)\big) \in Y^* \times X$ is optimal if and only if

$$(v,\omega) \in \partial F_1^*\big(-\Lambda^*(w,u)\big) \qquad \text{and} \qquad \Lambda(v,\omega) \in \partial F_2^*\big((w,u)\big).$$

Using that $\partial 0 = \{0\}$, the results of Lemma 3.5 as well as the subdifferentiation rule $\partial\|f - \cdot\|_1(u) = -\partial\|\cdot\|_1(f-u)$, this means

$$\begin{cases} v \in \beta\,\mathrm{Sgn}(Dw), \\ \omega \in \alpha\,\mathrm{Sgn}(Du - w), \end{cases} \qquad \begin{cases} \omega + v' = 0 \\ \quad \omega' \in -\mathrm{Sgn}(f-u). \end{cases}$$

Using $\omega = -v'$ and, consequently $\omega' = -v''$, the characterization (O_f)–(O_β) follows. □

4. The structure of the solutions

4.1. First-degree "staircasing" and monotonicity

In the L^1-TV case, i.e., for the problem

$$\min_{u\in \mathrm{BV}(\Omega)} \|u-f\|_{L^1(\Omega)} + \alpha\|Du\|_{\mathcal{M}(\Omega)},$$

the conditions (O_f)–(O_β) are replaced by the simpler conditions

$$v' \in \mathrm{Sgn}(f-u), \tag{4.1}$$

$$-v \in \alpha\,\mathrm{Sgn}(Du). \tag{4.2}$$

These conditions imply the well-known "staircasing of degree zero" phenomenon: u is piecewise constant when it does not equal f. In fact, arguing formally, if $u(x) < f(x)$ then $u < f$ in a neighborhood I of x and by (4.1) we have that $v' = 1$ and hence v is affine on I. Therefore by (4.2), $u' = 0$ and hence u is constant on I. For the case of an L^2 data fitting term the staircasing effect associated to the TV-regularization term was already analyzed in detail for dimension one in [18].

For L^1-TGV^2 we get a similar staircasing phenomenon "of the first degree", meaning that $u'' = 0$ in a suitable sense when u does not equal f.

Definition 4.1. Let $u \in \mathrm{BV}(\Omega)$ for $\Omega \subset \mathbb{R}$. For $x \in \Omega$, we define

$$\overline{u}(x) = \max\{u^+(x), u^-(x)\}, \quad \text{and} \quad \underline{u}(x) = \min\{u^+(x), u^-(x)\}$$

(equating $u^+ = u^- = \tilde{u}$ on $\Omega \setminus J_u$).

Observe that $\overline{u}$ and $\underline{u}$ are "good representatives" of u which have the property to be upper and lower semi-continuous, respectively. Indeed, if we have $\overline{u}(x) > t$ for some $x \in \Omega$ and $t \in \mathbb{R}$, then by right continuity, also $u^+ > \frac{1}{2}(\overline{u}(x) + t)$ in some right neighborhood $I^+ = (x, x+\delta)$. If $u^-(y) \le t$ for a $y \in I^+$, then by left continuity $u^- \le \frac{1}{2}(\overline{u}(x) + t)$ on some open subset of I^+ which contradicts that $u^- = u^+$ a.e. in Ω. Hence $\overline{u} > t$ on I^+. Analogously one gets $\overline{u} > t$ in a left neighborhood I^- of x. This shows the upper semi-continuity. With the same arguments, one obtains that $\underline{u}$ is lower semi-continuous.

In particular, these functions are continuous on $\Omega \setminus J_u$.

Proposition 4.2. *Let $f \in \mathrm{BV}(\Omega)$, and suppose that $u \in \mathrm{BV}(\Omega)$ solves (P) with the minimum in (TGV$^{\min}$) achieved by $w \in \mathrm{BV}(\Omega)$. Suppose $\overline{u} < \underline{f}$ on an open interval $I \subset \Omega$. Then we have*

(i) $(Du - w)\llcorner I = 0$, i.e., $u' = w$ on I and $|D^s u|(I) = 0$.

(ii) $w' = 0$ on I and $0 \le -Dw\llcorner I \ll \delta_x$ for some $x \in I$.

(iii) The function $w = u'$ is non-increasing on I.

If, on the other hand, $\underline{u} > \overline{f}$ on I, then in addition to (i), we have

(ii') $w' = 0$ on I and $0 \le Dw\llcorner I \ll \delta_x$ for some $x \in I$.

(iii') The function $w = u'$ is non-decreasing on I.

Proof. We consider the case $\overline{u} < \underline{f}$, as the case $\underline{u} > \overline{f}$ can be shown with identical arguments or, for instance, by observing that $-u$ is a solution of (P) with data $-f$ and a minimum in (TGV$^{\min}$) achieved by $-w$.

From (O_f) it first of all follows that $v'' = 1$ a.e. on I. In particular, v' is strictly monotone. Next, it follows from (O_α) that

$$-v' \in \alpha\, \mathrm{Sgn}(Du - w).$$

Since v' is strictly monotone and I is open, we must have $v' \in (-\alpha, \alpha)$ on I. This forces $Du - w\llcorner I = 0$. Hence $u' = w$ on I and $|D^s u|\llcorner I = 0$. This concludes the proof of (i).

On the other hand, (O_β) gives

$$v \in \beta\, \mathrm{Sgn}(Dw)$$

The fact that $v'' = 1$ implies that v is a quadratic function that reaches its minimum on I in at most one point $x \in I$. Elsewhere on the open set I we must have $v \in (-\beta, \beta)$. This forces $-Dw\llcorner I \ll \delta_x$ on I (with $x \in I$ arbitrary if v does not admit a minimum on I) as well as $0 \ge D^s w$. Therefore also $w' = 0$ on I. This concludes the proof of (ii). Property (iii) is an immediate consequence of (ii). □

Remark 4.3. Using the arguments of the proof of Proposition 4.2 it is now simple to argue rigorously staircasing of degree zero for the L^1-TV case. Iteration of this reasoning implies "staircasing of degree $k-1$" for TGV^k.

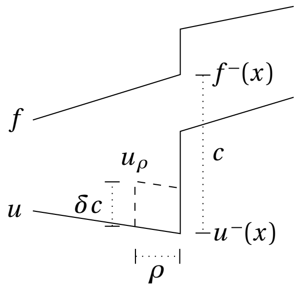

Figure 1. The construction in the proof of Proposition 4.6.

Corollary 4.4. *Let $f \in \mathrm{BV}(\Omega)$, and suppose $u \in \mathrm{BV}(\Omega)$ solves* (P) *with the minimum in* ($\mathrm{TGV}^{\min}$) *achieved by $w \in \mathrm{BV}(\Omega)$. Let*

$$A_{u,f} := \tilde{A}_{u,f} \cup \tilde{A}_{f,u}, \quad \text{where} \quad \begin{cases} \tilde{A}_{u,f} := \{x \in \Omega \mid \overline{u}(x) < \underline{f}(x)\}, \\ \tilde{A}_{f,u} := \{x \in \Omega \mid \overline{f}(x) < \underline{u}(x)\}. \end{cases}$$

Then $w = u'$ and $w' = 0$ on $A_{u,f}$. Moreover, $|D^s u|(A_{u,f}) = 0$, and the set $A_{u,f}$ as well as $\tilde{A}_{u,f}$ and $\tilde{A}_{f,u}$ are open.

Proof. *This is an immediate consequence of the semi-continuity of $\overline{u}$, $\underline{u}$, $\overline{f}$ and $\underline{f}$ as well as Proposition 4.2.* □

4.2. Structure of the jump set

Proposition 4.2 already tells us that $J_u \cap A_{u,f} = \emptyset$, that is, u has no jumps on any open set where it does not equal f in a suitable sense. Our next proposition strengthens this result, in particular showing the behavior on $\partial A_{u,f}$. It shows that the jumps of u are contained in the jumps of f in the sense of graphs.

Definition 4.5. For $f \in \mathrm{BV}(\Omega)$, let us define the *jump graph* as

$$G_f := \{(x,t) \in \Omega \times \mathbb{R} \mid x \in J_f,\ t \in [\underline{f}(x), \overline{f}(x)]\}.$$

Proposition 4.6. *Let $f \in \mathrm{BV}(\Omega)$, and suppose $u \in \mathrm{BV}(\Omega)$ solves* (P)*. Then $G_u \subset G_f$, and, in particular, $J_u \subset J_f$.*

Proof. Again we use the particular properties of BV-functions in the one-dimensional case as outlined in Section 2.2, in particular the left and right limits $f^\pm(x)$ always exist and give left and right continuous functions, respectively. Moreover, $S_f = J_f$. We choose $x \in J_u$, and consider only the case $u^-(x) < u^+(x)$, the opposite case being similar. To show that $G_u \subset G_f$, we have to show that $\underline{f}(x) \le u^-(x)$ and $u^+(x) \le \overline{f}(x)$. Since the proofs of these two properties are analogous, we study only the first one.

To reach a contradiction, we assume that $\underline{f}(x) > u^-(x)$, which implies that $f^-(x) > u^-(x)$. We denote the difference by $c := f^-(x) - u^-(x) > 0$ and choose $\gamma \in (0, 1/2]$ such that

$$c\gamma \le u^+(x) - u^-(x). \tag{4.3}$$

We consider the functions $u_\rho := u + c\gamma\chi_{B_\rho}$ for $B_\rho := [x-\rho, x]$; see Figure 1 for a sketch of the construction. Then

$$\int_\Omega |f(y) - u_\rho(y)|\,\mathrm{d}y = \int_{\Omega\setminus B_\rho} |f(y) - u(y)|\,\mathrm{d}y + \int_{B_\rho} |f(y) - u(y) - c\gamma|\,\mathrm{d}y. \tag{4.4}$$

Due to the left continuity of f^- and u^- we can find a $\rho > 0$ such that $f^-(y) - u^-(y) \ge c\gamma$ for $y \in B_\rho$. Consequently, as u^- and f^- are in the equivalence classes u and f,

$$\begin{aligned} \int_{B_\rho} |f(y) - u(y) - c\gamma|\,\mathrm{d}y &= \int_{B_\rho} f^-(y) - u^-(y) - c\gamma\,\mathrm{d}y \\ &< \int_{B_\rho} f^-(y) - u^-(y)\,\mathrm{d}y = \int_{B_\rho} |f(y) - u(y)|\,\mathrm{d}y. \end{aligned} \tag{4.5}$$

Observe, finally, that by the definition of u_ρ, we have

$$Du_\rho = Du + c\gamma(\delta_{x-\rho} - \delta_x).$$

Therefore, by the choice (4.3), a part of the jump of u at x of mass $c\gamma \leq u^+(x) - u^-(x)$ is shifted to $x-\rho$. It follows that

$$\begin{aligned}
\|Du_\rho - w\|_{\mathscr{M}(\Omega)} &= \|D^a u - w\|_{L^1(\Omega)} + \|D^s u + c\gamma(\delta_{x-\rho} - \delta_x)\|_{\mathscr{M}(\Omega)} \\
&\leq \|D^a u - w\|_{L^1(\Omega)} + \left\|D^s u - \left(u^+(x) - u^-(x)\right)\delta_x\right\|_{\mathscr{M}(\Omega)} + c\gamma + \left(u^+(x) - u^-(x) - c\gamma\right) \\
&= \|D^a u - w\|_{L^1(\Omega)} + \left\|D^s u - \left(u^+(x) - u^-(x)\right)\delta_x\right\|_{\mathscr{M}(\Omega)} + \left\|\left(u^+(x) - u^-(x)\right)\delta_x\right\|_{\mathscr{M}(\Omega)} \\
&= \|D^a u - w\|_{L^1(\Omega)} + \|D^s u\|_{\mathscr{M}(\Omega)} = \|Du - w\|_{\mathscr{M}(\Omega)}.
\end{aligned}$$

Consequently $\|Du - w\|_{\mathscr{M}(\Omega)}$ in (P), where $w \in \mathrm{BV}(\Omega)$, does not grow when replacing u by u_ρ. Minding (4.4) and (4.5), this shows that $F(u_\rho) < F(u)$, so u cannot be optimal. Hence we have found the desired contradiction and can conclude the proof. □

Remark 4.7. The same argument works in $\mathbb{R}^1$ for general L^1-TGV^k, $(k \geq 1)$, so, in particular L^1-TV. For the L^2-TV problem

$$\min_{u \in \mathrm{BV}(\Omega)} \|f - u\|^2_{L^2(\Omega)} + \alpha \|Du\|_{\mathscr{M}(\Omega)},$$

with $f \in \mathrm{BV}(\Omega) \cap L^\infty(\Omega)$ and $\Omega \subset \mathbb{R}^n$, $n \geq 1$, the property $J_u \subset J_f$, up to a set of $\mathscr{H}^{n-1}$ measure zero, has already been shown by a different technique in [19].

4.3. Summary

We summarize the findings of this section in the following theorem.

Theorem 4.8. *Suppose $u \in \mathrm{BV}(\Omega)$ solves (P) for $f \in \mathrm{BV}(\Omega)$. Then there exists an open set $A_{u,f}$, which is the union of at most countably many disjoint open intervals $I_i = (a_i, b_i)$, $(i = 0, 1, 2, \ldots)$, such that*

(i) $u = f$ on $\Omega \setminus \overline{A}_{u,f}$.

Moreover, for each $i = 0, 1, 2, \ldots$, there exist points $x_i \in I_i$ such that the following hold.

(ii) $\underline{f}(a_i) \leq u^+(a_i) \leq \overline{f}(a_i)$, and $\underline{f}(b_i) \leq u^-(b_i) \leq \overline{f}(b_i)$, if $a_i \neq a$ and $b_i \neq b$.

(iii) Both $u|(a_i, x_i)$ and $u|(x_i, b_i)$ are affine. Moreover, $u^-(x_i) = u^+(x_i)$, that is, u is continuous on I_i.

(iv) Either $u < \underline{f}$ or $u > \overline{f}$ on I_i. In the former case, $(u')^-(x_i) \geq (u')^+(x_i)$. In the latter case, $(u')^-(x_i) \leq (u')^+(x_i)$.

Proof. This is an immediate consequence of Propositions 4.2 & 4.6. □

5. Preserved properties

5.1. Continuity

Proposition 5.1. *Suppose that $f \in \mathrm{BV}(\Omega)$ is (absolutely) continuous and that $u \in \mathrm{BV}(\Omega)$ solves (P). Then u is (absolutely) continuous.*

Proof. In the one-dimensional case under consideration, u has a continuous representative on $\Omega \setminus J_u$. The preservation of continuity therefore follows from the fact that $J_u \subset J_f = \emptyset$ which was established in Proposition 4.6.

Next we show the preservation of absolute continuity. We write $A_{u,f} = \bigcup_{i=1}^\infty I_i$, where the intervals I_i are open and disjoint. We then write

$$f(t) = c + \int_a^t f'(s)\,\mathrm{d}s, \quad (t \in \Omega).$$

Such a representation holds thanks to the absolute continuity of f. Minding that, by Proposition 4.2, u is also absolutely continuous on $\bigcup_{i=1}^{j} I_i \subset A_{u,f}$, we define

$$g_j(t) := \begin{cases} f'(t), & t \in \Omega \setminus \bigcup_{i=1}^{j} I_i, \\ u'(t), & t \in \bigcup_{i=1}^{j} I_i, \end{cases}$$

and

$$u_j(t) := c_j + \int_a^t g_j(s)\,\mathrm{d}s$$

where $c_j = c$ if $a \notin \bigcup_{i=1}^{j} \overline{I}_i$ and $c_j = u^+(a)$ otherwise. Clearly $u_j' = g_j$. The idea is that u_j is formed from f by replacing it by u on each of the intervals $I_i = (c_i, d_i)$, $(i = 1, \ldots, j)$, where $u(c_i) = f(c_i)$ and $u(d_i) = f(d_i)$. Thus $u_j = f$ on $\Omega \setminus \bigcup_{i=1}^{j} I_i$.

We finally let

$$g(t) := \begin{cases} f'(t), & t \in \Omega \setminus A_{u,f}, \\ u'(t), & t \in A_{u,f}, \end{cases}$$

If we then show that $u_j \to u$ in $L^1(\Omega)$ and $u_j' = g_j \to g$ in $L^1(\Omega)$, it follows that $u_j \to u$ in $W^{1,1}(\Omega)$ and u is absolutely continuous with $u' = g$.

First, we indeed observe that

$$\lim_{j\to\infty} \|g_j - g\|_{L^1(\Omega)} = \lim_{j\to\infty} \sum_{i=j+1}^{\infty} \|u' - f'\|_{L^1(I_i)} = 0,$$

thanks to $\|u' - f'\|_{L^1(A_{u,f})} < \infty$ and $\mathscr{L}^1(\bigcup_{i=j+1}^{\infty} I_i) \to 0$ as $j \to \infty$.

Second, we observe analogously that

$$\lim_{j\to\infty} \|u_j - u\|_{L^1(\Omega)} = \lim_{j\to\infty} \sum_{i=j+1}^{\infty} \|u - f\|_{L^1(I_i)} = 0.$$

This concludes the proof. □

Theorem 5.2. *Let $f : \Omega \to \mathbb{R}$ be Lipschitz continuous with Lipschitz constant L, and suppose that $u \in \mathrm{BV}(\Omega)$ solves* (P). *Then u is Lipschitz continuous with Lipschitz constant at most L.*

Proof. By Proposition 5.1 the function u is absolutely continuous. Hence $A_{u,f}$ is open and $|u'| \le L$ pointwise a.e. on $\Omega \setminus A_{u,f}$, since $u = f$ on this set. The theorem will be verified if we show that $|u'| \le L$ a.e. in $A_{u,f} = \tilde{A}_{u,f} \cup \tilde{A}_{f,u}$.

Let $(c_1, c_2) \subset \tilde{A}_{u,f}$ be a maximal interval. By Proposition 4.2 we have $0 \le -Dw \llcorner (c_1, c_2) \ll \delta_x$ for some $x \in (c_1, c_2)$ and hence

$$u' = w_1 \text{ in } (c_1, x) \quad \text{and} \quad u' = w_2 \text{ in } (x, c_2)$$

with $w_1 \ge w_2$. Suppose that $w_1 > L$. Then $c_1 = a$, since otherwise $u(c_1) = f(c_1)$ which is a contradiction to

$$u(c_1) = u(x) + w_1(c_1 - x) < f(x) - L(x - c_1) \le f(x) + f(c_1) - f(x) = f(c_1).$$

By (O_f) therefore $v'' = 1$ in (a, c_2) and, using $v \in H_0^2(\Omega)$, we have $v > 0$ in (a, c_2).

Let $c_3 \in [x, b]$ be chosen as the maximal real number such that $w = w_1$ in (a, c_3). If $c_3 < b$ then $v(c_3) \in \{-\beta, \beta\}$: If this is not the case, then $|v| < \beta$ in a neighborhood of c_3 implying by (O_β) that $Dw = 0$ and hence, $w = w_1$ in a neighborhood of c_3 which contradicts the maximality of c_3. In either of the cases $v(c_3) \in \{-\beta, \beta\}$, v admits a local maximum at some $x^* \in (a, c_3]$. If $c_3 = b$ this is also true for some $x^* \in (a, b)$, since $v \in H_0^2(\Omega)$. The following argument will imply that $c_3 = b$.

Since $v'(x^*) = 0$ and $v' \in \mathscr{C}([a,b])$ there exists $\epsilon > 0$ such that $|v'| < \alpha$ in $(x^* - \epsilon, x^*) =: \Omega' \subset \Omega$. If $\{f = u\}$ on a non-null subset $\mathscr{S}$ of Ω', then $f' = u'$ on $\mathscr{S}$. Using that $u' = w_1$ in Ω' by (O_α) we have $w_1 = f' \leq L$ on $\mathscr{S}$, which is a contradiction. Therefore $f \neq u$ a.e. in Ω'. By (O_f) this leads to $|v''| = 1$ a.e. in Ω'. If $v'' = 1$ a.e. in Ω', then $v(y) > v(x^*)$ for $y \in \Omega'$, where we use that $v'(x^*) = 0$. Hence $v(x^*)$ is not a local maximum of v, which gives a contradiction. Consequently $v'' = -1$ on some non-null subset $\Omega'' \subset \Omega'$. In particular $f < u$ on Ω'' by (O_f).

By the last argument there exists a nontrivial open interval (c_4, c_5), with $c_4 < x^*$ and $c_5 \leq c_3 \leq b$ in which $f < u$ and $u' = w_1$, and c_5 is chosen to be maximal. The case $c_5 < b$ cannot occur, since for any $x \in I$ we have

$$u(c_5) = u(x) + w_1(c_5 - x) > f(x) + L(c_5 - x) \geq f(x) + f(c_5) - f(x) = f(c_5),$$

which contradicts $u(c_5) = f(c_5)$. Hence $c_3 = b$, $u > f$ on (c_4, b), and $v'' = -1$ on (c_4, b). Consequently v is quadratic on (c_4, b). Since $x^* \in (c_4, b)$ and $v'(x^*) = 0$ this contradicts $v \in H_0^2(\Omega)$. Therefore $w_2 \leq w_1 \leq L$ and $u' \leq L$ on (c_1, c_2). Since (c_1, c_2) was an arbitrary maximal interval in $\tilde{A}_{u,f}$ we have $u' \leq L$ in $\tilde{A}_{u,f}$.

Likewise one can prove that $-L > w_2$ leads to a contradiction by changing in the above arguments the left and right sides of Ω. The case $(c_1, c_2) \subset \tilde{A}_{f,u}$, with $u > f$ in (c_1, c_2) and $w_1 \leq w_2$ can be treated analogously by swapping signs and considering minima instead of maxima. This concludes the proof. □

5.2. *Piecewise affinity*

Theorem 5.3. *Let $f : \Omega \to \mathbb{R}$ be piecewise affine, and suppose that $u \in \mathrm{BV}(\Omega)$ solves (P). Then u is piecewise affine.*

Proof. By Proposition 4.2, u is piecewise affine on any open interval in $A_{u,f}$. Clearly u is also piecewise affine on any open interval in $\Omega \setminus A_{u,f}$, since it is equal to f there. Therefore it suffices to show that $A_{u,f}$ is the union of at most finitely many open intervals.

Let f be affine on the segments (a_i, a_{i+1}) with discontinuities at a_i and choose $I = (c, d)$ as a maximal interval of $A_{u,f}$ that lies strictly inside Ω. Then by Proposition 4.2, u is piecewise affine on I with at most one point of discontinuity of $w = u'$. Moreover $u(c) = f(c)$ and $u(d) = f(d)$, and without loss of generality we may assume that $u < f$ in I. Let j be chosen such that $c \in [a_j, a_{j+1})$. Since $u(c) = f(c)$ and $u < f$ on I yield $(u')^+(c) < (f')^+(c)$, and since by Proposition 4.2, w is non-increasing, so that $u' < f' = (f')^+(c)$ on (a_j, a_{j+1}), it is impossible that $d \in [a_j, a_{j+1}]$. Hence the extremities of I must lie in different segments of (a_i, a_{i+1}), and it follows that there is only a finite number of such intervals. □

Remark 5.4. The proof also shows that u does not oscillate away from f in the middle of an interval I on which f is affine.

6. The effect of the regularization parameters

6.1. *Convergence*

In this subsection, we consider problem (P) with the regularization term weighted for simplicity with a single parameter $\lambda > 0$, that is we consider

$$\min_{u \in \mathrm{BV}(\Omega)} F_\lambda(u), \quad F_\lambda(u) := \|f - u\|_{L^1(\Omega)} + \lambda \,\mathrm{TGV}^2_{\vec{\alpha}}(u). \tag{P_λ}$$

Proposition 6.1. *Assume that $f \in \mathrm{BV}(\Omega)$. For $\alpha, \beta > 0$ fixed, let u_λ be a solution of (P_λ) with $\lambda > 0$. Then*

(i) $u_\lambda \to f$ strongly in $L^1(\Omega)$ as $\lambda \searrow 0$.

(ii) Every sequence $\lambda_i \nearrow \infty$ has a subsequence $\{\lambda_{i_j}\}_{j=0}^\infty$, such that $u_{\lambda_{i_j}} \to f^*$ weakly in $\mathrm{BV}(\Omega)$ as $j \to \infty$, where f^* is a solution to the L^1-regression problem

$$\min_{u \text{ affine}} \|f - u\|_{L^1(\Omega)}. \tag{6.1}$$

(iii) The function $\lambda \mapsto \|f - u_\lambda\|_{L^1(\Omega)}$ is non-decreasing, while the function $\lambda \mapsto \mathrm{TGV}^2_{\vec{\alpha}}(u_\lambda)$ is non-increasing.

Proof. The proof of (i) is elementary: Suppose that $u_\lambda \not\to f$ in $L^1(\Omega)$. Then there exist $\delta > 0$ and a sequence $\lambda_i \searrow 0$, $(i = 0, 1, 2, \dots)$, such that

$$\delta \le \|u_{\lambda_i} - f\|_{L^1(\Omega)} \le F_{\lambda_i}(u_{\lambda_i}) \le F_{\lambda_i}(f).$$

But $F_{\lambda_i}(f) \to 0$ as $i \to \infty$, which gives a contradiction to the above inequality.

The proof of (ii) is somewhat more involved. First of all, we observe that $\mathrm{TGV}^2(u) = 0$ for affine functions u. Since u_λ solves (P_λ), we find that

$$\min_{v \text{ affine}} \|(f - u_\lambda) - v\|_{L^1(\Omega)} = \|f - u_\lambda\|_{L^1(\Omega)},$$

and consequently

$$u_\lambda \in X := \{u \in L^1(\Omega) \mid 0 \in \mathscr{R}(f - u)\}, \tag{6.2}$$

where $\mathscr{R}(f)$ is the solution set of (6.1). Note that X is closed with respect to strong convergence in $L^1(\Omega)$. In fact, let $\{u_i\}$ denote a sequence in X with limit u. For arbitrary $\epsilon > 0$ we have $\|u - u^i\|_{L^1(\Omega)} < \epsilon/2$ for i large enough. Then for such i

$$\begin{aligned}
\|f - u\|_{L^1(\Omega)} &\le \|f - u^i\|_{L^1(\Omega)} + \|u^i - u\|_{L^1(\Omega)} \\
&= \min_{v \text{ affine}} \|(f - u^i) - v\|_{L^1(\Omega)} + \|u^i - u\|_{L^1(\Omega)} \\
&\le \min_{v \text{ affine}} \|(f - u) - v\|_{L^1(\Omega)} + 2\|u^i - u\|_{L^1(\Omega)} \\
&\le \min_{v \text{ affine}} \|(f - u) - v\|_{L^1(\Omega)} + \epsilon.
\end{aligned}$$

Let $w = w_\lambda$ be such that the minimum in ($\mathrm{TGV}^{\min}$) is achieved for $u = u_\lambda$. Further, denote the mean

$$\bar{u}_\lambda := [\mathscr{L}^1(\Omega)]^{-1} \int_\Omega u_\lambda \, d\mathscr{L}^1,$$

and similarly let $\bar{w}_\lambda$ be the mean of w_λ on Ω. We define $u_\lambda^a(t) := t\bar{w}_\lambda + c_\lambda$, where $c_\lambda \in \mathbb{R}$ is chosen such that $\bar{u}_\lambda^a = \bar{u}_\lambda$. The Poincaré inequality [16, Theorem 3.44], applied twice, then gives for a constant C dependent on Ω alone, and a constant C' dependent on $\vec{\alpha}$ and C, that

$$\begin{aligned}
\|u_\lambda - u_\lambda^a\|_{L^1(\Omega)} &\le C\|Du_\lambda - Du_\lambda^a\|_{\mathscr{M}(\Omega)} \\
&= C\|Du_\lambda - \bar{w}_\lambda\|_{\mathscr{M}(\Omega)} \\
&\le C\|Du_\lambda - w_\lambda\|_{\mathscr{M}(\Omega)} + C\|w_\lambda - \bar{w}_\lambda\|_{L^1(\Omega)} \\
&\le C\|Du_\lambda - w_\lambda\|_{\mathscr{M}(\Omega)} + C^2\|Dw_\lambda\|_{\mathscr{M}(\Omega)} \\
&\le C' \,\mathrm{TGV}^2_{\vec{\alpha}}(u_\lambda).
\end{aligned} \tag{6.3}$$

Observe then that $\{F_\lambda(u_\lambda)\}_{\lambda>0}$ is bounded, because

$$F_\lambda(u_\lambda) \le F_\lambda(f^*) = \|f - f^*\|_{L^1(\Omega)} < \infty.$$

Thus $\mathrm{TGV}^2_{\vec{\alpha}}(u_\lambda) \to 0$, for $\lambda \nearrow \infty$, and hence by (6.3)

$$\|u_\lambda - u_\lambda^a\|_{L^1(\Omega)} \to 0, \qquad \text{for } \lambda \nearrow \infty. \tag{6.4}$$

Observe now that $\{u_{\lambda_i}\}_{i=0}^\infty$ is bounded in $L^1(\Omega)$ since $\{F_\lambda(u_\lambda)\}_{\lambda>0}$ being bounded. Hence $\{u_{\lambda_i}^a\}_{i=0}^\infty$ is bounded by (6.4). Since the functions $u_{\lambda_i}^a$, $(i = 0, 1, 2, \dots)$, are affine, we may therefore find an unrelabeled subsequence $\lambda_i \nearrow \infty$, such that $u_{\lambda_i}^a \to u^a$ strongly in $L^1(\Omega)$ for some affine function u^a. Consequently also $u_{\lambda_i} \to u^a$ strongly in $L^1(\Omega)$. Since $u_{\lambda_i} \in X$ and since X is closed it follows that $0 \in \mathscr{R}(f - u^a)$, which by u^a being affine implies that u^a solves (6.1). This establishes that u_{λ_i} converges strongly in $L^1(\Omega)$ to a solution of (6.1).

We still need to bound $\{\|Du_{\lambda_i}\|_{\mathcal{M}(\Omega)}\}_{i=0}^\infty$ to get weak convergence in $\mathrm{BV}(\Omega)$. Towards this end, we observe from (6.3) and the discussion following it that $\|Du_{\lambda_i} - Du^a_{\lambda_i}\|_{\mathcal{M}(\Omega)} \to 0$. But $\{\|Du^a_{\lambda_i}\|_{\mathcal{M}(\Omega)}\}_{i=0}^\infty$ is bounded since $\{u^a_{\lambda_i}\}_{i=0}^\infty$ is bounded in $L^1(\Omega)$ and the functions $u^a_{\lambda_i}$ are affine. Therefore $\{\|Du_{\lambda_i}\|_{\mathcal{M}(\Omega)}\}_{i=0}^\infty$ is also bounded. This completes the proof of claim (ii).

Claim (iii) follows by a generic argument. Let $\mu > \lambda$. We then have

$$\begin{aligned}
&\|f - u_\lambda\|_{L^1(\Omega)} + \lambda \mathrm{TGV}^2_{\vec{\alpha}}(u_\lambda) \le \|f - u_\mu\|_{L^1(\Omega)} + \lambda \mathrm{TGV}^2_{\vec{\alpha}}(u_\mu), \quad \text{and}\\
&\|f - u_\mu\|_{L^1(\Omega)} + \mu \mathrm{TGV}^2_{\vec{\alpha}}(u_\mu) \le \|f - u_\lambda\|_{L^1(\Omega)} + \mu \mathrm{TGV}^2_{\vec{\alpha}}(u_\lambda).
\end{aligned}$$

Therefore, summing, we find that

$$(\mu - \lambda)\mathrm{TGV}^2_{\vec{\alpha}}(u_\mu) \le (\mu - \lambda)\mathrm{TGV}^2_{\vec{\alpha}}(u_\lambda),$$

so that $\mathrm{TGV}^2_{\vec{\alpha}}(u_\mu) \le \mathrm{TGV}^2_{\vec{\alpha}}(u_\lambda)$, if $\mu > \lambda$. This shows that $\lambda \mapsto \mathrm{TGV}^2_{\vec{\alpha}}(u_\lambda)$ is non-increasing. Next, we deduce that

$$\begin{aligned}
\|f - u_\lambda\|_{L^1(\Omega)} + \lambda \mathrm{TGV}^2_{\vec{\alpha}}(u_\lambda) &\le \|f - u_\mu\|_{L^1(\Omega)} + \lambda \mathrm{TGV}^2_{\vec{\alpha}}(u_\mu)\\
&\le \|f - u_\mu\|_{L^1(\Omega)} + \lambda \mathrm{TGV}^2_{\vec{\alpha}}(u_\lambda),
\end{aligned}$$

which shows that

$$\|f - u_\lambda\|_{L^1(\Omega)} \le \|f - u_\mu\|_{L^1(\Omega)},$$

concluding the proof of the claim and the lemma. □

Remark 6.2. With reference to (ii) above, note that as $\mathrm{TGV}^2_{\vec{\alpha}}(u) = 0$ forces $Dw = 0$ and thus u' to be a constant, we find that f^* is a solution of the constrained problem

$$\min_{u \in \mathrm{BV}(\Omega)} \|f - u\|_{L^1(\Omega)} \text{ subject to } \mathrm{TGV}^2_{\vec{\alpha}}(u) = 0.$$

Remark 6.3. In the following we will see that, actually, $u_{\lambda^*} = f^*$ for sufficiently large λ^*. The convergence proof above remains valid with minor modifications also for $\lambda \to \infty$ replaced by $\lambda \nearrow \lambda^*$. In fact, from this case utilizing Proposition 6.1 (iii) we have $\lim_{\lambda \nearrow \lambda^*} \mathrm{TGV}^2_{\vec{\alpha}}(u_\lambda) = 0$. Hence we obtain (6.4) with $\lambda \to \infty$ replaced by $\lambda \nearrow \lambda^*$ using (6.3).

6.2. *Thresholding*

We next derive bounds on $\vec{\alpha}$ ensuring that either $u = f^*$ or $u = f$ solve (P). We begin with the L^1-regression case.

Proposition 6.4. *There exists $\alpha^*, \beta^* \in (0, \infty)$, such that whenever $f \in \mathrm{BV}(\Omega)$, $\alpha \ge \alpha^*$, and $\beta \ge \beta^*$, then (P) is solved by an L^1-regression f^* of f.*

Proof. The proof is based on the Poincaré inequality argument found in the proof of Proposition 6.1. Let $u \in \mathrm{BV}(\Omega)$ be arbitrary. Then for any $w \in \mathrm{BV}(\Omega)$ let $u^a(t) := t\bar{w} + c$ where c is chosen such that $\bar{u}^a = \bar{u}$. Then

$$\begin{aligned}
\|f - f^*\|_{L^1(\Omega)} &= \min_{v \text{ affine}} \|f - v\|_{L^1(\Omega)}\\
&\le \min_{v \text{ affine}} \left(\|f - u\|_{L^1(\Omega)} + \|u - v\|_{L^1(\Omega)}\right)\\
&\le \|f - u\|_{L^1(\Omega)} + \|u - u^a\|_{L^1(\Omega)}.
\end{aligned}$$

According to (6.3) we have

$$\|u - u^a\|_{L^1(\Omega)} \le C\|Du - w\|_{\mathcal{M}(\Omega)} + C^2\|Dw\|_{\mathcal{M}(\Omega)},$$

where C is the constant for the Poincaré inequality in Ω. Now, choosing w such that it achieves the minimum in ($\mathrm{TGV}^{\min}$) for the chosen u it follows that

$$\|f - f^*\|_{L^1(\Omega)} \le \|f - u\|_{L^1(\Omega)} + \mathrm{TGV}^2_{\vec{\alpha}}(u) \quad \text{for all } u \in \mathrm{BV}(\Omega),$$

provided that $\alpha \ge C$ and $\beta \ge C^2$. Thus $\alpha^* = C$ and $\beta^* = C^2$ satisfy the claims of the proposition independently of f. □

We next derive bounds on $\vec{\alpha}$ ensuring that $u = f$ for the solution of (P), at least for reasonably simple f. Similar results for L^1-TV can be found in [6, 7].

Notation. Let $f : \Omega \to \mathbb{R}$ be piecewise affine with $I_1, \ldots, I_{N_f}$ the maximal disjoint ordered (open) intervals on each of which f is affine. We denote

$$\delta_f := \min_{i=1,\ldots,N_f} \mathcal{L}^1(I_i).$$

Proposition 6.5. *Let $f : \Omega \to \mathbb{R}$ be piecewise affine with $J_f = \emptyset$ and*

$$\delta_f \geq \begin{cases} 2\beta/\alpha + \alpha, & \text{if } \alpha \leq \sqrt{2\beta}, \\ 2\sqrt{2\beta}, & \text{if } \alpha \geq \sqrt{2\beta}. \end{cases} \tag{6.5}$$

Then $u = f$ whenever u is a solution of (P).

Proof. We study when the optimality conditions (O_f)–(O_β) hold with $u = f$ and $w = f'$. For this purpose we need to find $v \in H_0^2(\Omega)$, satisfying

$$\begin{aligned} v'' &\in \operatorname{Sgn}(0), \\ -v' &\in \alpha \operatorname{Sgn}(0), \quad \text{and} \\ v &\in \beta \operatorname{Sgn}(D^j f'). \end{aligned}$$

Let $I_1, \ldots, I_{N_f}$ be the intervals of affinity of f, with $I_i = (a_i, b_i)$, with $a_1 = a$, $b_{N_f} = b$, $a_{i+1} = b_i$, $i = 2, \ldots, N_f - 1$. Also let $d_i \in \{-1, +1\}$ denote the direction of the jump of f' at a_i, ($i = 2, \ldots, N_f$). Then the optimality conditions reduce into

$$v''(t) \in [-1, 1], \quad (t \in \Omega), \tag{6.6}$$

$$v'(t) \in [-\alpha, \alpha], \quad (t \in \Omega), \tag{6.7}$$

$$v(t) \in [-\beta, \beta], \quad (t \in \Omega), \quad \text{and} \tag{6.8}$$

$$v(a_1) = 0,\ v(b_{N_f}) = 0,\ v(a_i) = \beta d_i. \tag{6.9}$$

Let us set $\delta_* := 2\beta/\alpha + \alpha$ and suppose $\delta_* \geq 2\alpha$. Then $(\alpha, \delta_* - \alpha)$ is a nonempty open interval and we can set

$$r(t) := \begin{cases} t^2/2, & t \in (0, \alpha], \\ -\alpha^2/2 + \alpha t, & t \in (\alpha, \delta_* - \alpha], \\ -\alpha^2 + \alpha\delta_* - (\delta_* - t)^2/2, & t \in (\delta_* - \alpha, \delta_*], \\ 2\beta, & t \in (\delta_*, \infty). \end{cases}$$

We can check that $r \in \mathscr{C}^1([0,\infty)) \cap H^2_{\mathrm{loc}}((0,\infty))$. Continuity at δ_* requires that $r(\delta_*) = 2\beta$: the condition for the latter is just $-\alpha^2 + \alpha\delta_* = 2\beta$, which suggested the definition

$$\delta_* = 2\beta/\alpha + \alpha.$$

Moreover we note that $r(0) = 0$. If $\alpha \leq \sqrt{2\beta}$, which corresponds to the first case of (6.5), this implies the requirement that $\delta_* \geq 2\alpha$. The derivatives of r satisfy $r' \in [-\alpha, \alpha]$, $r'' \in [-1, 1]$ almost everywhere in $(0, \infty)$. We now define the dual variable by assigning its values on each of the interval I_i, $i = 1, \ldots, N_f$, according to

$$v(t) = \beta d_i + c_i r(t - a_i), \quad \text{for } t \in I_i,$$

where $d_1 = 0$ and

$$c_i = \begin{cases} d_2/2 & \text{for } i = 1, \\ (d_{i+1} - d_i)/2 & \text{for } i = 2, \ldots, N_f - 1, \\ -d_{N_f - 1}/2 & \text{for } i = N_f \end{cases}$$

with the jump directions d_i at the jump points defined in (6.9). Note that $c_i \in \{-1,0,1\}$ for $i = 2,\dots,N_f - 1$ and $c_1, c_{N_f} \in \{-1/2, 1/2\}$. Since, by assumption, $\delta_f \ge \delta_*$, we have $r(b_i - a_i) = 2\beta$, and therefore $v \in \mathscr{C}(\bar\Omega)$. Moreover $r'(0) = r'(b_i - a_i) = 0$ and this implies that $v \in \mathscr{C}^1(\bar\Omega)$. Finally we chose $v(a_1) = v(b_{N_f}) = 0$ and hence $v \in H_0^2(\Omega)$. Since $c_i \in \{-1,0,1\}$ it follows that $v' \in [-\alpha,\alpha]$, $v'' \in [-1,1]$. By construction we can also see that $v(t) \in [-\beta,\beta]$. Thus we find that v satisfies (6.6)–(6.9).

To cover the second case of (6.5), suppose that $\alpha \ge \sqrt{2\beta}$. Setting $\tilde\delta := 2\sqrt{2\beta}$, observe that $\tilde\delta \le 2\alpha$ and $\tilde\delta \le \delta_*$ (with equality at $\alpha = \sqrt{2\beta}$). We now define

$$\tilde r(t) := \begin{cases} t^2/2, & t \in (0, \tilde\delta/2), \\ \tilde\delta^2/4 - (\tilde\delta - t)^2/2, & t \in (\tilde\delta/2, \tilde\delta), \\ 2\beta, & t \in (\tilde\delta, \infty). \end{cases}$$

Then $\tilde r(0) = 0$ and $r(\tilde\delta) = 2\beta$ by the choice of $\tilde\delta$. Clearly again $\tilde r \in \mathscr{C}^1([0,\infty)) \cap H^2_{\mathrm{loc}}((0,\infty))$ with $\tilde r'(0) = 0$ and $\tilde r'(T) = 0$ for any $T \ge \tilde\delta$, as well as $r''(t) \in [-1,1]$ for a.e. $t \in (0,\infty)$ and $r'(t) \in [-\alpha,\alpha]$ for all $t \in [0,\infty)$. Defining v as above with $\tilde r$ in place of r, similar reasoning shows that (6.6)–(6.9) hold. □

Remark 6.6. Observe that as the intervals on which f is affine get smaller, β also has to become smaller to guarantee "locking" $u = f$ by Proposition 6.5. An example illustrating this point is provided by Example 6.12 below.

In the following proposition we consider the case of piecewise affine functions allowing for jumps in the function values as well as in the derivative.

Proposition 6.7. *Let $f : \Omega \to \mathbb{R}$ be piecewise affine with $J_f \cap J_{f'} = \emptyset$ and*

$$\alpha \le \sqrt{2\beta} \quad \textit{and} \quad 2\alpha + 4\beta/\alpha \le \delta_f. \tag{6.10}$$

Then $u = f$ whenever u is a solution of (P).

Proof. We shall adapt the proof of Proposition 6.5. With (6.10) holding, the first case of (6.5) holds as well and we can use the function r of the proof of Proposition 6.5.

The optimality conditions with $u = f$ and $w = D^a f$ are satisfied if we find $v \in H_0^2(\Omega)$ satisfying

$$\begin{aligned} v'' &\in \mathrm{Sgn}(0), \\ -v' &\in \alpha\,\mathrm{Sgn}(D^j f), \quad \text{and} \\ v &\in \beta\,\mathrm{Sgn}(D^j D^a f), \end{aligned}$$

or equivalently if (6.6) – (6.8) hold and (6.9) is replaced by

$$\begin{aligned} &v(a_1) = 0,\ v(b_{N_f}) = 0,\ v(a_i) = \beta d_i, \text{ if } a_i \in J_{f'}, \\ &v'(a_i) = \alpha d_i, \text{if } a_i \in J_f, \text{ for } i = 2,\dots,N_f, \end{aligned} \tag{6.11}$$

where, as above, $d_i \in \{-1,1\}$ if $a_i \in J_{f'} \cup J_f$, with the sign depending on whether the jump of f' and f in a_i, respectively, is negative or positive.

The function r needs to be modified to guarantee that the last requirement in (6.11) holds. Note at first that $r'(\delta_*/2) = \alpha$. This follows from the fact that $\alpha \le \frac{\delta_*}{2} \le \delta_* - \alpha$, which is implied by $\alpha^2 \le 2\beta$. The idea is now to refine the partition $I_1,\dots,I_{N_f}$ such that the construction of Proposition 6.5 can be applied. For $I_1,\dots,I_{N_f}$, with $I_i = (a_i, b_i)$ and $b_i - a_i \ge \delta_f$, denoting the intervals on which f is affine, we consider intervals

$$\tilde I_i := I_i \setminus \bigcup_{x \in J_f} I^x, \quad (i = 1,\dots,N_f),$$

and

$$I^x := (x - \delta_*/2, x + \delta_*/2), \quad (x \in J_f).$$

Recall here that $J_f \cup J_{f'} = \{a_2, \ldots, a_{N_f}\}$. The condition $\delta_f \geq \frac{4\beta}{\alpha} + 2\alpha$ guarantees that $\delta_* = \frac{2\beta}{\alpha} + \alpha \leq \frac{\delta_f}{2}$ and hence $\mathscr{L}^1(\tilde{I}_i) \geq \delta_*$ and $\mathscr{L}^1(I^x) = \delta_*$, for each $x \in J_f$. The intervals $\tilde{I}_i$ and $I^x, x \in J_f$, form a new partition $\tilde{I}_i = (\tilde{a}_i, \tilde{a}_{i+1})$ of (a, b), with $i = 1, \ldots, \tilde{M}$, for some $\tilde{M}$, and $\tilde{a}_1 = a$, $\tilde{a}_{\tilde{M}+1} = b$. Each jump point of f is the midpoint of some interval I^x, each jump point of f' is a boundary point of some $\tilde{I}_i$. If $\tilde{a}_i$ coincides with some $a_j \in J_{f'}$, then $v(\tilde{a}_i) = v(a_j)$ is already defined there. Otherwise, is $\tilde{a}_i$ is an endpoint of some I^x, say the left endpoint. For $\tilde{a}_i \in J_{f'}$ we set $\tilde{d}_i$ according to the direction of the jump of f', while for $I^x = (\tilde{a}_i, \tilde{a}_{i+1})$, we set $\tilde{d}_i = 1$, $\tilde{d}_{i+1} = -1$, if the jump of f in x is positive, and $\tilde{d}_i = -1$, $\tilde{d}_{i+1} = 1$, if it is negative. Now v on (a, b) can be defined as in the proof of Proposition 6.5, with a_i replaced by $\tilde{a}_i$, d_i replaced by $\tilde{d}_i$, and $N_f = \tilde{M}$. For $\tilde{a}_i = a_j$, with $a_j \in J_{f'}$, we have $v(\tilde{a}_i) = v(a_j) = \beta d_j$, and for $\tilde{a}_i = a_j$, with $a_j \in J_f$ we have that a_j is the midpoint of the interval $\tilde{I}_i$ and hence $v'(\tilde{a}_i) = v'(a_j) = \alpha d_j$. Thus this v is our desired dual variable. □

Remark 6.8. The "locking" of u to the data f, as studied in Proposition 6.7, does not necessarily hold for any values of α and β when $J_f \cap J_{f'} \neq \emptyset$. This point will be demonstrated in Example 6.11 below. Moreover, Example 6.12 below demonstrates that locking may not be achieved for functions that are not (finitely) piecewise affine, even when the function is continuous.

Generally, we have the following "partial locking" result.

Proposition 6.9. *Suppose u is a solution of* (P) *with the minimum in* (TGV$^{\min}$) *achieved by $w \in \mathrm{BV}(\Omega)$. Then we have the following.*

(i) If $\overline{u} < \underline{f}$ or $\underline{u} > \overline{f}$ an open interval I, then $\mathscr{L}^1(I) \leq 2\alpha$.

(ii) If $u' \in \mathrm{BV}(\Omega)$ and $\overline{w} < \underline{u'}$ or $\underline{w} > \overline{u'}$ a.e. on an open interval I, such that $f' \in \mathrm{BV}(I)$, then $\mathscr{L}^1(I) \leq 2\beta/\alpha$.

Proof. We first show point (i), considering only the case $\overline{u} < \underline{f}$, as the case $\underline{u} > \overline{f}$ is analogous. We choose arbitrary $x, x' \in I$ with $x < x'$. By the necessary optimality condition (O_α), we have $v'(x), v'(x') \in [-\alpha, \alpha]$. while, since $v'' = 1$ on I by (O_f), we get

$$v'(x') - v'(x) = \int \chi_{[x,x']} v'' \, \mathrm{d}\mathscr{L}^1 = \pm(x' - x).$$

We therefore deduce that $|x' - x| \leq 2\alpha$ and $\mathscr{L}^1(I) \leq 2\alpha$.

To show point (ii), we simply employ in the above proof, the condition (O_β) in place of (O_α), to get $v(x), v(x') \in [-\beta, \beta]$. Then we use the condition (O_α) in place of (O_f) to get $v(x') - v(x) = \int \chi_{[x,x']} v' \, \mathrm{d}\mathscr{L}^1 = \mp\alpha(x' - x)$. Thus we deduce $\alpha|x' - x| \leq 2\beta$. □

Propositions 6.5 and 6.9 imply the following corollary.

Corollary 6.10. *Let $f : \Omega \to \mathbb{R}$ be piecewise affine with $J_f = \emptyset$ and suppose that $\delta_f > 2\beta/\alpha$ and* (6.5) *hold. Then the optimal solution satisfies $u = f$ and $w = f'$.*

Proof. By Proposition 6.5 condition (6.5) implies that $u = f$. Since $\delta_f > 2\beta/\alpha$, Proposition 6.9 shows that $\mathscr{L}^1(I) < \delta_f$ for any open interval I such that $\overline{w} < \underline{f'}$ or $\underline{w} > \overline{f'}$. It follows that $w(x_i) = f'(x_i)$ at some $x_i \in I_i$ for each $i = 1, \ldots, N_f$, and, consequently $\|Dw\|_{\mathscr{M}} \geq \|Df'\|_{\mathscr{M}}$. It is hence optimal to pick $w = f'$. □

6.3. Examples

We next study some counter-examples regarding the thresholds on $\vec{\alpha}$ which guarantee that $u = f$ solves (P). The next example demonstrates that for general piecewise affine f, (P) may not be solved by $u = f$ for arbitrary choice of $\alpha, \beta > 0$.

Example 6.11. Let us take the domain $\Omega := (-1, 1)$ and consider on Ω the function

$$f(t) := \begin{cases} 0, & \text{for } t \leq 0, \\ 1 - t, & \text{for } t > 0. \end{cases}$$

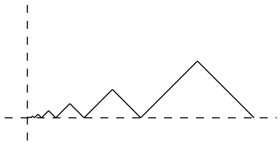

Figure 2. Function f of Example 6.12.

We study again the optimality conditions (O_f)–(O_β). The conditions (O_α),(O_β) state that for some $v \in H_0^2(\Omega)$

$$\begin{aligned} -v' &\in \alpha\,\mathrm{Sgn}(Du - w), \quad \text{and} \\ v &\in \beta\,\mathrm{Sgn}(Dw). \end{aligned} \tag{6.12}$$

These conditions are compatible with v and v' having zero traces on $\partial\Omega$, i.e. $v \in H_0^2(\Omega)$, only if there exists $\delta > 0$ such that $w = u'$ and $w' = 0$ on $(-1, -1+\delta) \cup (1-\delta, 1)$.

Suppose then that $u = f$ solves (P). Consider $w \in \mathrm{BV}(\Omega)$ that minimizes

$$\alpha\|w - Df\|_{\mathcal{M}(\Omega)} + \beta\|Dw\|_{\mathcal{M}(\Omega)} = \alpha + \alpha\|w + \chi_{(0,1)}\|_{L^1(\Omega)} + \beta\|Dw\|_{\mathcal{M}(\Omega)}.$$

From the reasoning above, we have that for some $\delta > 0$, $w = 0$ on $(-1, -1+\delta)$, and $w = -1$ on $(1-\delta, 1)$. But then necessarily $\|Dw\|_{\mathcal{M}(\Omega)} \geq 1$, which gives $w = -\chi_{(0,1)}$ as the optimal choice.

We next show that $u = f$ and $w = -\chi_{(0,1)}$ cannot solve (P). The optimality conditions (O_α),(O_β) for this choice would state that

$$\begin{aligned} -v' &\in \alpha\,\mathrm{Sgn}(\delta_0), \quad \text{and} \\ v &\in \beta\,\mathrm{Sgn}(-\delta_0), \end{aligned}$$

so that $v'(0) = -\alpha$, and $v(0) = -\beta$. But, minding that $v \in H_0^2(\Omega)$, the function v is differentiable with v' (Lipschitz) continuous. Since $v \geq -\beta$ by (6.12), clearly we cannot then have $v(0) = -\beta$ with $v'(0) = -\alpha < 0$. Thus $u = f$ cannot solve (P).

Our final example concerns functions with countably many affine parts, but no jumps.

Example 6.12. Let us consider $\Omega = (0,1)$ and the sawtooth function

$$f(t) := \int_0^t \sum_{i=2}^{\infty} \big(\chi_{(2\cdot 2^{-i}, 3\cdot 2^{-i})}(s) - \chi_{(3\cdot 2^{-i}, 4\cdot 2^{-i})}(s)\big)\, \mathrm{d}s$$

depicted in Figure 2. The function f is absolutely continuous with countably many affine parts, but $Df' = 2\sum_{i=2}^{\infty}(\delta_{2\cdot 2^{-i}} - \delta_{3\cdot 2^{-i}})$, so that f' has infinite variation on $(0,\delta)$ for any $\delta > 0$.

Suppose u and w solve (P) for f. As in Example 6.11 above, there must exist $\delta > 0$ such that $w = u'$ and $w' = 0$ on $(0,\delta)$. If $u = f$, this would imply that w has infinite variation on $(0,\delta)$, and so clearly cannot minimize $\alpha\|Df - w\|_{\mathcal{M}(\Omega)} + \beta\|Dw\|_{\mathcal{M}(\Omega)}$. We conclude that $u = f$ cannot solve (P).

Acknowledgements

The authors would like to thank the anonymous referee for many helpful comments and suggestions. This work was supported by the Austrian Science Fund (FWF) under grant SFB F32 (SFB "Mathematical Optimization and Applications in Biomedical Sciences").

References

[1] K. Bredies, K. Kunisch, T. Pock, Total generalized variation, SIAM J. Imaging Sciences 3 (2011) 492–526.

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

[3] F. Knoll, K. Bredies, T. Pock, R. Stollberger, Second order total generalized variation (TGV) for MRI, Magnetic Resonance in Medicine 65 (2) (2011) 480–491.

[4] K. Bredies, T. Valkonen, Inverse problems with second-order total generalized variation constraints, in: Proceedings of SampTA 2011, 9th International Conference on Sampling Theory and Applications, 2011.

[5] T. Pock, L. Zebedin, H. Bischof, TGV-fusion, in: C. Calude, G. Rozenberg, A. Salomaa (Eds.), Rainbow of Computer Science, Vol. 6570 of Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2011, pp. 245–258.

[6] T. F. Chan, S. Esedoglu, Aspects of total variation regularized L^1 function approximation, SIAM Journal on Applied Mathematics 65 (2005) 1817–1837.

[7] V. Duval, J. F. Aujol, Y. Gousseau, The TVL1 model: A geometric point of view, SIAM J. Multiscale Modeling & Simulation 8 (2009) 154–189.

[8] M. Nikolova, Minimizers of cost functions involving nonsmooth data-fideltiy terms. Application of processing of outliers, SIAM J. on Numerical Analysis 40 (2002) 965–994.

[9] W. Yin, D. Goldfarb, S. Osher, The total variation regularized L^1 model for multiscale decomposition, SIAM J. Multiscale Modeling & Simulation 6 (2007) 190–211.

[10] C. Clason, B. Jin, K. Kunisch, A duality-based splitting method for ℓ^1-TV image restoration with automatic choice of regularization parameter, SIAM J. Sci. Comp. 32 (2010) 1484–1505.

[11] J. Yang, Y. Zhang, W. Yin, An efficient TVL1 algorithm for deblurring multichannel images corrupted by impulsive noise, SIAM J. Sci. Comp. 31 (2009) 2842–2865.

[12] Y. Dong, M. Hintermüller, M. Neri, An efficient primal-dual method for the $L^1 - TV$ image restoration, SIAM J. Sci. Comp 2 (2009) 1168–1189.

[13] C. Pöschl, O. Scherzer, Characterization of minimizers of convex regularization functionals, in: Frames and operator theory in analysis and signal processing, Vol. 451 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2008, pp. 219–248.

[14] G. Dal Maso, I. Fonseca, G. Leoni, M. Morini, A higher order model for image restoration: the one-dimensional case, SIAM J. Math. Anal. 40 (6) (2009) 2351–2391.

[15] W. Rudin, Real and Complex Analysis, McGraw-Hill Book Company, 1966.

[16] L. Ambrosio, N. Fusco, D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, Oxford University Press, 2000.

[17] H. Attouch, H. Brezis, Duality for the sum of convex functions in general Banach spaces, in: J. A. Barroso (Ed.), Aspects of Mathematics and its Applications, Vol. 34 of North-Holland Mathematical Library, Elsevier, 1986, pp. 125–133.

[18] W. Ring, Structural properties of solutions to total variation regularization problems, ESAIM: Mathematical Modelling and Numerical Analysis 34 (2000) 799–810.

[19] V. Caselles, A. Chambolle, M. Novaga, The discontinuity set of solutions of the TV denoising problem and some extensions, Multiscale Modeling and Simulation 6 (3) (2008) 879–894.