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

Kristian Bredies * Karl Kunisch* Tuomo Valkonen*

May 5, 2011

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.

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

1 The L^1-TGV^2 functional

In this work we study the variational problem

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

where f is the input data, u denotes a solution, and $\vec{\alpha} = (\beta, \alpha) > 0$ stands for a vector-valued regularization parameter. The precise definition of TGV^2 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 $\text{TGV}^2_{\vec{\alpha}}$-functional is a special case of the $\text{TGV}^k_{\vec{\alpha}}$-functional, where $k \geq 2$, which was introduced in [3]. In [3] basic analytical properties of $\text{TGV}^k_{\vec{\alpha}}$ and numerical results with an L^2 data-fitting term for the cases $k = 2$ and $k = 3$ are provided. One way of interpreting $\text{TGV}^k_{\vec{\alpha}}$, consists in realizing that it regularizes independently on different regularity scales of the function that it is applied to. Compared to the BV-functional [10] we recall that constant functions are in the kernel of BV, while polynomials of degree constitute the kernel of $\text{TGV}^k_{\vec{\alpha}}$, see [3].

The reason for focussing on the case $k = 2$ in spatial dimension one is given by the fact that we aim at getting detailed insight into the effect $\text{TGV}^2_{\vec{\alpha}}$

*Institute for Mathematics and Scientific Computing, Karl-Franzens University of Graz, Austria. E-mail: `kristian.bredies@uni-graz.at`, `karl.kunisch@uni-graz.at`, `tuomo.valkonen@iki.fi`.

on the structure of the solution to (1.1). 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. [5, 8, 9, 13]. All of these papers address the L^1-BV, as opposed to the TGV case, which is in the focus of the present paper. The numerical realization of the L^1-BV problem is typically considered in 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. [6, 12], and semi-smooth Newton methods [7].

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 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, is replaced by staircasing of degree one for solutions to (1.1). 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 $I_1, \dots, I_N$ 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 $\Omega \subset \mathbb{R}^m$ be a Borel set and let $\mathcal{M}(\Omega, \mathbb{R}^n)$ denote the space of (vector-valued) Radon measures on Ω. The total variation measure of $\mu \in \mathcal{M}(\Omega)$ is denoted $|\mu|$, and we define the norm $\|\mu\|_{\mathcal{M}(\Omega)} := |\mu|(\Omega)$, see e.g. [11].

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

We denote by $\mathcal{L}^m$ the Lebesgue measure on $\mathbb{R}^m$, while $\mathcal{H}^k$ 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 what the Radon norm means for distributions. A distribution u on Ω is a Radon measure (in the sense that there is a $\mu \in \mathcal{M}(\Omega)$ such that $\int_\Omega v \,\mathrm{d}\mu = \langle u, v\rangle$ for all $v \in \mathcal{C}_\mathrm{c}^\infty(\Omega)$) if and only if

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

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

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

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

and

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

where the value ∞ is possibly attained.

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

2.2 Functions of bounded variation

Following, e.g., [1], a function $u \in L^1(\Omega)$ on a non-empty open set $\Omega \subset \mathbb{R}^m$ 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

$$\int_\Omega u \operatorname{div} \phi \,\mathrm{d}x = -\int_\Omega \phi \,\mathrm{d}Du, \quad \text{for all} \quad \phi \in \mathcal{C}_\mathrm{c}^\infty(\Omega, \mathbb{R}^m).$$

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

In the following, let $m = 1$, $\Omega = (a, b)$ 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

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

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 $x \in B$ and are defined by satisfying

$$\lim_{\rho \searrow 0} \frac{1}{\rho} \int_0^\rho |u^+(x) - u(y)| \, \mathrm{d}y = 0 \quad \text{and} \quad \lim_{\rho \searrow 0} \frac{1}{\rho} \int_{-\rho}^0 |u^-(x) - u(y)| \, \mathrm{d}y = 0,$$

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 $D^a u = u' \mathcal{L}^1$ is the *absolutely continuous part*, with u' the approximate differential, $D^j u$ represents the *jump part* which can be represented as

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

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

For $u \in \mathrm{BV}(\Omega)$ we will be mostly working with *good representatives* as defined in [1, 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

$$\tilde{u}(t) \in c_u + Du((a,t)) + [0,1] Du(\{t\}) \quad \text{for all} \quad t \in (a,b).$$

In this sense, u^- and u^+ are good representatives of u.

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

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

also called the *predual* or *supremum definition* of $\mathrm{TGV}^2_{\vec{\alpha}}(u)$. The existence of solutions to (P) follows from the lower semi-continuity of the TGV seminorm shown in [3].

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

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

Observe that the minimization problem in ($\mathrm{TGV}^{\min}$) is just L^1-TV for u', as the singular part $D^s u$ cannot be approximated by w. In [3] it was shown that for $u \in \mathcal{C}^\infty(\Omega)$, 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

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

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

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

Proof. First observe that the supremum in ($\mathrm{TGV}^{\sup}$) can also be written as the negative infimum

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

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

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

the infimum in (3.1) can be expressed as

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

We employ the Fenchel-Rockafellar duality formula for this setting. According to [2] this is justified 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

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

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

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

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

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

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

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

Applying the Poincaré inequality [1, 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

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

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

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

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 $\mathcal{M}(\Omega)$. For this purpose we need the following generalization of the sign function.

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

$$\begin{aligned} \mathrm{Sgn}(\mu) = \{v \in L^\infty(\Omega) \cap L^\infty(\Omega, |\mu|) \bigm| \|v\|_\infty \leq 1,\ \|v\|_{\infty,|\mu|} \leq 1, \\ v = \mathrm{sgn}(\mu),\ |\mu|\text{-almost everywhere}\}, \end{aligned}$$

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

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

It is obvious that if $u \in L^1(\Omega)$, then $v \in L^\infty(\Omega)$ belongs to $\mathrm{Sgn}(u)$ 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 $\mu \in \mathcal{M}(\Omega)$ 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 $\mathcal{M}(\Omega)$ 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 $\partial\|\cdot\|_1(u) = \mathrm{Sgn}(u)$.

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

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

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

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 $\mu \in \mathcal{M}(\Omega)$, $v \in \mathcal{C}_0(\Omega)$ implies $v \in L^\infty(\Omega) \cap L^\infty(\Omega, |\mu|)$ with $\|v\|_{\infty,|\mu|} \leq \|v\|_\infty$. Now, $v \in \mathcal{C}_0(\Omega)$ satisfies

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

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

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

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

Proposition 3.6. *A pair $(u, w) \in \mathrm{BV}(\Omega)^2$ is a minimizer for (P) if and only if there exists a $v \in H_0^2(\Omega)$ such that*

$$v'' \in \operatorname{Sgn}(f - u), \tag{O$_f$}$$

$$-v' \in \alpha \operatorname{Sgn}(Du - w), \tag{O$_\alpha$}$$

$$v \in \beta \operatorname{Sgn}(Dw). \tag{O$_\beta$}$$

Proof. We will show that the maximization problem

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

can be regarded the predual problem for (P) 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

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

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

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

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

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

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

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

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

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

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

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

$$v' \in \operatorname{Sgn}(f-u), \tag{4.1}$$
$$-v \in \alpha \operatorname{Sgn}(Du). \tag{4.2}$$

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. By (4.2) therefore $u' = 0$ and hence u is constant on I.

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 then set

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

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

Observe that $\overline{u}$ and $\underline{u}$ are "good representatives" of u. In particular, they are continuous on $\Omega \setminus J_u$.

Lemma 4.2. *Let $f, u \in \mathrm{BV}(\Omega)$. Then the set of $x \in \Omega$ with $\overline{u}(x) < \underline{f}(x)$ (resp. $\underline{u}(x) > \overline{f}(x)$) is open.*

Proof. Suppose that $\overline{u}(x) < \underline{f}(x)$. We may further assume $u = \overline{u}$ and $f = \underline{f}$. We let $d := f(x) - u(x) > 0$. We may then find $\delta > 0$ such that

$$|Df|((x, x+\delta)) + |Du|((x, x+\delta)) < d/3$$

and

$$|Df|((x-\delta, x)) + |Du|((x-\delta, x)) < d/3.$$

Here we use that $0 = \lim_{i\to\infty} |Df|((x, x+\frac{1}{i})) = |Df|(\cap_{i\in\mathbb{N}}(x, x+\frac{1}{i}))$ and analogously for u. The characterization of "good representatives" in Subsection 2.2, shows that for a constant c_u we have

$$u(t) \in c_u + Du((a,t)) + [0,1]Du(\{t\}),$$

so that, for $\epsilon \in (0, \delta)$, we have

$$u(x+\epsilon) \le u(x) + |Du|((x, \epsilon]).$$

Likewise we find for $\epsilon \in (0, \delta)$ that

$$f(x+\epsilon) \ge f(x) - |Df|((x, \epsilon]).$$

Consequently, we obtain $u - f \le -d/3$ on $(x, x+\delta)$. A similar calculation can be performed on $(x-\delta, x)$. We find that $u < f$ on $I := (x-\delta, x+\delta)$. $\square$

Proposition 4.3. *Let $f \in \mathrm{BV}(\Omega)$, and suppose that $u \in \mathrm{BV}(\Omega)$ solves* (P) *with the minimum in* ($\mathrm{TGV}^{\min}$) *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 analogous arguments.

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

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

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

$$v \in \beta \operatorname{Sgn}(Dw)$$

The fact that, $v'' = 1$ implies that v is a quadratic function that reaches its minimum on I at exactly one point $x \in \overline{I}$. Elsewhere on the open set I we must have $v \in (-\beta, \beta)$. This forces $-Dw\llcorner I \ll \delta_x$ 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 4.4. Using the arguments of the proof of Proposition 4.3 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.

Corollary 4.5. *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*

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

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 Lemma 4.2 and Proposition 4.3. □

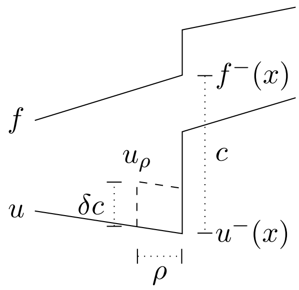


Figure 1: The construction in the proof of Proposition 4.7.

4.2 Structure of the jump set

Proposition 4.3 already tells as 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 4.6. For $f \in \mathrm{BV}(\Omega)$, let us define the *jump graph* as

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

Proposition 4.7. *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 $S_f = J_f$. 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 $\underline{f}(x) \leq u^-(x)$ and $u^+(x) \leq \overline{f}(x)$. 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

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

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

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

We claim that

$$\int_{B_\rho} |f(y) - u(y) - c\gamma|\, \mathrm{d}y < \int_{B_\rho} |f(y) - u(y)|\, \mathrm{d}y \tag{4.5}$$

for some small $\rho > 0$.

In B_ρ, we have pointwise almost everywhere

$$|f - u - c\gamma| \le (1-\gamma)|f-u| + \gamma|f-u-c|,$$

so it suffices to show that

$$\int_{B_\rho} |f(y) - u(y) - c| \, \mathrm{d}y < \int_{B_\rho} |f(y) - u(y)| \, \mathrm{d}y \tag{4.6}$$

By the definition of approximate limits, $(f-u)^- = f^- - u^- = c$, and we get

$$\begin{aligned} 0 &\le \lim_{\rho \searrow 0} \frac{1}{\rho} \int_{B_\rho} |(f-u)(y) - c| \, \mathrm{d}y \\ &\le \lim_{\rho \searrow 0} \frac{1}{\rho} \int_{B_\rho} |f - f^-| \, \mathrm{d}y + \lim_{\rho \searrow 0} \frac{1}{\rho} \int_{B_\rho} |u - u^-| \, \mathrm{d}y = 0. \end{aligned}$$

This implies that

$$\lim_{\rho \searrow 0} \frac{1}{\rho} \int_{B_\rho} |(f-u)(y)| \, \mathrm{d}y \ge \lim_{\rho \searrow 0} \left(c - \frac{1}{\rho} \int_{B_\rho} |(f-u)(y) - c| \, \mathrm{d}y \right) = c.$$

and establishes the existence of some small $\rho > 0$ such that (4.6) and consequently (4.5) hold. Moreover, since J_u is at most countable ρ can be chosen such that $x - \rho \notin J_u$. Recalling (4.4), this implies

$$\|f - u_\rho\|_{L^1(\Omega)} < \|f - u\|_{L^1(\Omega)}. \tag{4.7}$$

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 $c\gamma \le |D^j u(\{x\})|$ is shifted to $x - \rho \notin J_u$. It follows that

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

Consequently $\|Du - w\|_{\mathcal{M}(\Omega)}$ in (P), where $w \in \mathrm{BV}(\Omega)$, is not affected by replacing u by u_ρ. Minding (4.7), 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 4.8. The same argument works in $\mathbb{R}^1$ for general L^1-TGV^k, $(k \ge 1)$, so, in particular L^1-TV. For the L^2-TV problem

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

with $f \in \mathrm{BV}(\Omega) \cap L^\infty(\Omega)$ and $\Omega \subset \mathbb{R}^n$, $n \ge 1$, the property $J_u \subset J_f$, up to a set of $\mathcal{H}^{n-1}$ measure zero, has already been shown by a different technique in [4].

4.3 Summary

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

Theorem 4.9. *Suppose* $u \in \mathrm{BV}(\Omega)$ *solves* (P) *for* $f \in \mathrm{BV}(\Omega)$. *Then there exists an open set* $A_{u,f}$, *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* $\underline{f}(b_i) \leq u^-(b_i) \leq \overline{f}(b_i)$.

(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 4.3 & 4.7. □

5 Preserved properties

5.1 Continuity

Proposition 5.1. *Suppose that* $f : \Omega \to \mathbb{R}$ *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 4.7.

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

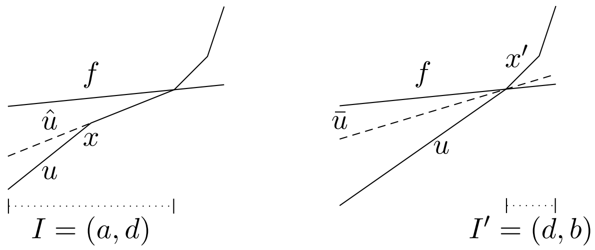
$$f(t) = c + \int_a^t f'(s) \, \mathrm{d}s, \quad (t \in \Omega).$$

Such a representation holds thanks to the absolute continuity of f. Minding that, by Proposition 4.3, 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

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



(a) Illustration of $\hat{u}$

(b) Illustration of $\bar{u}$

Figure 2: Constructions in the proof of Theorem 5.2

Clearly $u_j' = g_j$. If $a \notin \bigcup_{i=1}^j \overline{I}_i$, then $c_j = c$. Otherwise $c_j = u^+(a)$. The idea is that u_j is formed from f by replacing it by u on each of the intervals $I_i = (c_i, d_i)$, $(i = 1, \dots, j)$, 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 $u' = g$.

Firstly, 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 $\mathcal{L}^1(\bigcup_{i=j+1}^\infty I_i) \to 0$.

Secondly, 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. Observe that the set $A_{u,f}$ is open, and $|u'| \le L$ pointwise a.e. on $\Omega \setminus \overline{A}_{u,f}$, as $u = f$ on this set. We want to show that $|u'| \le L$ pointwise a.e. on all of Ω.

Let $w \in \mathrm{BV}(\Omega)$ be a function achieving the minimum in ($\mathrm{TGV}^{\min}$). By Proposition 4.3, we have $u' = w$ and $w' = 0$ on $A_{u,f}$. Thus $|u'| \le L$ a.e. on

$A_{u,f}$ will follow if we show $|w| \leq L$ on $A_{u,f}$. Since $\partial A_{u,f}$ is $\mathcal{L}^1$-negligible and u is absolutely continuous by Proposition 5.1, a referral to the fundamental theorem of calculus then shows that u has Lipschitz factor at most L, as claimed.

We may study w separately on the sets $\tilde{A}_{u,f}$ and $\tilde{A}_{f,u}$, whose union constitutes $A_{u,f}$. Since the proof is analogous (with some sign changes) in both cases, we concentrate on $\tilde{A}_{u,f}$, i.e., on the case $u < f$, and show that $|w| \leq L$ on $\tilde{A}_{u,f}$. Let $I \subset \tilde{A}_{u,f}$ be a maximal open interval, that is, there exists no open interval $I' \neq I$ with $I \subset I' \subset \tilde{A}_{u,f}$. (Recall that, $\tilde{A}_{u,f} \subset \mathbb{R}$ being open, it is the union of countably many *disjoint* open intervals.) By Proposition 4.3, $0 \leq -Dw \llcorner I \ll \delta_x$ for some $x \in I$, so that for some $w_1, w_2 \in \mathbb{R}$, $w_1 \geq w_2$, we have $w = w_1$ on (c, x) and $w = w_2$ on (x, d). Moreover, since $u < f$ on I, there exists $\epsilon > 0$ such that $u(x) + \epsilon < f(x)$.

To reach a contradiction suppose that $w_1 > L$. Then

$$u' = w_1 > L \geq f' \quad \text{on } (c, x). \tag{5.1}$$

It follows that $u(y) + \epsilon \leq f(y)$ for $y \in (c, x)$. By the maximality of I, we deduce that $c = a$. Thus, on $(a, x]$, we have $u + \epsilon < f$ and u is affine with slope w_1. Suppose $w_2 < w_1$, and set $\hat{L} := \max\{L, w_2\}$ as well as (see Figure 2(a))

$$\hat{u}(y) := \begin{cases} u(y), & y \in [x, b), \\ u(x) + \hat{L}(y - x), & y \in (a, x), \end{cases} \qquad \hat{w}(y) := \begin{cases} w(y), & y \in [x, b), \\ \hat{L}, & y \in (a, x). \end{cases}$$

Using (5.1) it follows that $u < \hat{u} \leq f$ on (a, x). Moreover $\hat{u}' - \hat{w} = u' - w$ on $[x, b)$, $\hat{u}' - \hat{w} = 0$ on (a, x), and

$$|\hat{w}^+(x) - \hat{w}^-(x)| < |w^+(x) - w^-(x)|,$$

where we use that $\hat{L} \leq w_2$. It is thus easily seen that $F(\hat{u}) < F(u)$, contradicting the optimality of u for (P). Hence, if $w_1 > L$, then $w_2 = w_1$, so that u is affine on $I = (a, d)$ with slope $w_1 = w_2$.

Let us now assume that $d < b$. Setting $x = d$ and $\hat{L} = \max\{L, w^+(d)\}$ an argument by contradiction using the above construction shows that $w^+(d) \geq w_1$. By the maximality of I we have $u(d) = f(d)$. Let $h \in \overline{(A_{u,f} \cap (d, b))}$ be the point closest to d in this set. Then $u = f$ on $I_0 := [d, h]$. Suppose $h > d$. The function w then minimises $\|f' - w\|_1$ on (d, h) subject to the boundary values $w^+(d)$ on d and $w^-(h)$ on h. Since $f' \leq L$ and $w^+(d) \geq w_1$, it follows that $w \geq w_1$ a.e. on (d, h). Indeed, if we had $\underline{w}(y') < w_1$, then (see the proof of Lemma 4.2) there would exist $y \in (d, y']$ and $\tilde{L} \in [\underline{w}(y), \overline{w}(y)] \cap [L, w^+(d))$ such that $f' \leq \tilde{w} < w$ on (d, y) for

$$\tilde{w}(x) = \begin{cases} \tilde{L}, & x \in (d, y), \\ w(x), & \text{otherwise.} \end{cases}$$

Then also $|D\tilde{w}|(\Omega) \leq |Dw|(\Omega)$, which would contradict the optimality of w. Since now $w \geq w_1 > L$ a.e. on I_0, we may actually assume (choosing $\tilde{L} = w_1$ and $y = h$) that $w = w_1$ a.e. on I_0. In fact, if $w^+(h) < w_1$ (when $h < b$), we could take $y = h$ and $\tilde{L} = \max\{L, w^+(h)\}$, contradicting that $w^+(d) \geq w_1$. Thus $w^+(h) \geq w_1$ (when $h < b$) and $w \geq w_1$ a.e. on $I \cup I_0 = (a, h]$. The same conclusion holds already by earlier reasoning when $h = d$.

Suppose $h < b$. It now follows from $u(h) = f(h)$ that $u > f$ on an interval $I' = (h, e) \subset \Omega$. Indeed, since $h \in \partial A_{u,f} \cap (d, b)$ and $w^+(h) \geq w_1 > L$, we have the existence of $\epsilon > 0$ such that $B := (h, h+\epsilon) \cap A_{u,f}$ satisfies $\mathcal{L}^1(B) > 0$ and $u' = w > L$ a.e. on B (see again the proof of Lemma 4.2), while $u = f$ a.e. on $(h, h+\epsilon) \setminus B$. From here it follows by integration that $B = (h, h+\epsilon)$ and $u > f$ on $I' = B$. We may again take I' maximal, in the sense defined above. Using (ii') of Proposition 4.3 we obtain the existence of $x' \in (h, e)$ such that for some $w_1', w_2' \in \mathbb{R}$, with $w_1' \leq w_2'$, we have $w = w_1' = w^+(d) \geq w_1 > L$ on (h, x') and $w = w_2'$ on (x', e). But then $w > L$ on I', which implies that $u^-(e) > f^-(e)$. By the maximality of I', we thus necessarily have $e = b$. This implies that $\Omega = I \cup I_0 \cup I'$.

When $h = b$, we take $I' = \emptyset$.

When $d = b$, we take $h = b$ and $I_0 = I' = \emptyset$.

Let us now define (see Figure 2(b))

$$\bar{u}(y) := \begin{cases} u(d) + L(y-d), & y \in I, \\ u(h) + L(y-h), & y \in I', \\ u(y), & y \in I_0, \end{cases} \qquad \bar{w}(y) := L, \qquad (y \in (a,b)).$$

Then $u < \bar{u} \leq f$ on (a, d), and $f \leq \bar{u} < u$ on (h, b). It follows that $\|\bar{u} - f\|_{L^1(\Omega)} < \|u - f\|_{L^1(\Omega)}$. Trivially also $D\bar{u} - \bar{w} = 0$, and $D\bar{w} = 0$, so that clearly $F(\bar{u}) < F(u)$. This provides the desired contradiction to the assumption $w_1 > L$. Since $w_2 \leq w_1$, it follows that $w \leq L$ on I. A proof by contradiction completely analogous to the one above further shows that $w_2 \geq -L$, so that $|w| \leq L$ on I. 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 4.3, u is piecewise affine on any open interval $I \Subset A_{u,f}$. Trivially this results extends to any open interval $I \subset A_{u,f}$. Clearly u is also piecewise affine on any open interval $I \subset \Omega \setminus A_{u,f}$, since it is equal to f there. The only problem therefore lies in showing that $A_{u,f}$ is the union of at most finitely many open intervals. From there the same result follows for $\Omega \setminus \overline{A}_{u,f}$.

Let $I \in \{I_1, \ldots, I_N\}$ be one of the finitely many open intervals on which f is affine. It then suffices to show that $A_{u,f} \cap I$ consists of finitely many open intervals.

Recall that by Proposition 4.7 and continuity of f on I we may assume that u continuous on I. Let $I' = (c, d) \subset A_{u,f} \cap I$ be maximal. In view of the continuity of u on I this means that for $x \in \{c, d\}$, either $u(x) = f(x)$ or $x \in \partial I$. If no such interval exists, then due to Lemma 4.2, we have $u = f$ on I and there is nothing to prove.

Next we note that there are at most two sub-intervals $I' \subset A_{u,f} \cap I$ sharing a boundary point of I. It therefore suffices to study the number of subintervals $I' = (c, d)$ with $c, d \notin \partial I$. For such intervals, both $u(c) = f(c)$ and $u(d) = f(d)$, while $u \neq f$ on I'. Let us assume that $u < f$ on I'. The opposite case can be treated analogously. By Proposition 4.3, u is piecewise affine on I, with at most one point of discontinuity for $w = u'$. Consequently, the assumptions $u(c) = f(c)$ and $u < f$ on I yield $w^+(c) < (f')^+(c)$. By Proposition 4.3, moreover, w is non-increasing, so that $w < f' = (f')^+(c)$ on I'. Consequently $u(d) = f(d)$ is impossible. This contradiction shows that $A_{u,f} \cap I$ consists of at most two intervals, specifically those with at least one boundary point meeting a boundary point of I. □

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 regularisation 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

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

Proposition 6.1. *For $\alpha, \beta > 0$ fixed, let u_λ be a solution of (P$_\lambda$) 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 $u_{\lambda_{i_j}} \rightharpoonup f^*$ 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 $\lambda_i \searrow 0$, $(i = 0, 1, 2, \ldots)$, such that

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

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

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

Note that X is a 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

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

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

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

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 [1, 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

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

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

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

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

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

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 $\{u^a_{\lambda_i}\}_{i=0}^\infty$ is bounded by (6.4). Since the functions $u^a_{\lambda_i}$, $(i = 0, 1, 2, \ldots)$, are affine, we may therefore find an unrelabeled subsequence $\lambda_i \nearrow \infty$, such that $u^a_{\lambda_i} \to u^a$ 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 $(f - u^a)^* = 0$, which by u^a being affine implies that $f^* = u^a$ solves (6.1). This establishes that $u_{\lambda_i} \to f^*$ strongly in $L^1(\Omega)$.

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 also when $\lambda \nearrow \lambda^*$ where $u_{\lambda^*} = f^*$ at λ^*.

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 $f \in \mathrm{BV}(\Omega)$, $\alpha \geq \alpha^*$, and $\beta \geq \beta^*$, then (P) is solved by the 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 $u \in BV(\Omega)$ 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

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

According to (6.3) we have

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

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

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

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

We next derive bounds on $\vec{\alpha}$ that ensure that $u = f$ for the solution of (P), at least for reasonably simple f. Similar results for L^1-TV can be found in [5, 8].

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

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

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

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

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 $d_{a_i} \in \{-1, +1\}$ denote the direction of the jump of f' at a_i, ($i = 2, \ldots, N_f$). Then the optimality conditions reduce into

$$v''(t) \in [-1, 1], \quad (t \in \Omega), \tag{6.6}$$
$$v'(t) \in [-\alpha, \alpha], \quad (t \in \Omega), \tag{6.7}$$
$$v(t) \in [-\beta, \beta], \quad (t \in \Omega), \quad \text{and} \tag{6.8}$$
$$v(a_1) = 0,\ v(b_{N_f}) = 0,\ v(a_i) = \beta d_{a_i}. \tag{6.9}$$

Let us set $\delta_* := 2\beta/\alpha + \alpha$ and suppose $\delta_* \geq 2\alpha$. Then $(\alpha, \delta_* - \alpha)$ is a welldefined 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 $r \in C^1([0, \infty)) \cap H^2_{loc}((0, \infty))$. 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 Ω. We now define the dual variable by assigning its values on each of the interval I_i, $i = 1, \ldots, N_f$, according to

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

with

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

with jump directions d_i at the jump points defined in (6.9). Note that $c_i \in \{-1, 0, 1\}$. Since, by assumption, $\delta_f \geq \delta_*$, we have $r(b_i - a_i) = 2\beta$, and therefore $v \in C(\bar{\Omega})$. Moreover $r'(0) = r'(b_i - a_i) = 0$ and this implies that $v \in C^1(\bar{\Omega})$. Finally we chose v such that $v(a_1) = v(b_{N_f}) = 0$ 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 it follows 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 $\alpha \geq \sqrt{2\beta}$. Setting $\tilde{\delta} := 2\sqrt{2\beta}$, observe that $\tilde{\delta} \leq 2\alpha$ and $\tilde{\delta} \leq \delta_*$ (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 $\tilde{r} \in H^2_{\mathrm{loc}}((0, \infty))$ with $\tilde{r}'(0) = 0$ and $\tilde{r}'(T) = 0$ for any $T \geq \tilde{\delta}$, 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*

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

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

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

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

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

where, as above, $d_{a_i} \in \{-1, 1\}$ if $a_i \in J_{f'} \cup J_f$, with the sign depending on whether the jump 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 $\alpha \leq \frac{\delta_*}{2} \leq \delta_* - \alpha$, which is implied by $\alpha^2 \leq 2\beta$. The idea now is to add "extra points" to $J_{f'}$ around points of J_f. For $I_1, \dots, I_{N_f}$, with $I_i = (a_i, b_i)$ and $b_i - a_i \geq \delta_f$, 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 $J_f \cup J_{f'} = \{a_2, \dots, a_{N_f}\}$. 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 $\mathcal{L}^1(\tilde{I}_i) \geq \delta_*$ and $\mathcal{L}^1(I^*) = \delta_*$. 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 $i = 1, \dots, \tilde{M}$, for some $\tilde{M}$, and $\tilde{a}_1 = a$, $\tilde{a}_{\tilde{M}+1} = b$. Each jumppoint of f is the midpoint of some interval $\tilde{I}_i$, each jumppoint 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. Then we define the values of v as $v(\tilde{a}_i) = -\beta$ and for the right endpoint $v(\tilde{a}_{i+1}) = +\beta$ if the jump of f is positive and s $v(\tilde{a}_i) = \beta$ and for the right endpoint $v(\tilde{a}_{i+1}) = -\beta$ if the jump is negative. Now v on (a, b) can be defined as in the proof of Proposition 6.5, with a_i replaced by $\tilde{a}_i$, and $N_f = \tilde{M}$. For $\tilde{a}_i = a_j$, with $a_j \in J_{f'}$, we have $v(\tilde{a}_i) = v(a_j) = \beta d_{a_j}$ 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 $v'(\tilde{a}_i) = v'(a_j) = \alpha d_{a_j}$. 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 achived 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 $\mathcal{L}^1(I) \leq 2\alpha$.

(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 $\mathcal{L}^1(I) \leq 2\beta/\alpha$.

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

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

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

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 $v(x') - v(x) = \int \chi_{[x,x']} v' \, \mathrm{d}\mathcal{L}^1 = \mp\alpha(x' - x)$. 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 $\mathcal{L}^1(I) < \delta_f$ 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 $i = 1, \ldots, N_f$. But then it is 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

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

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 $(-1, -1 + \delta) \cup (1 - \delta, 1)$.

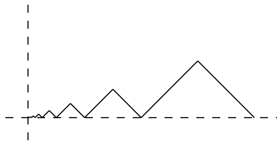


Figure 3: Function f of Example 6.12.

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

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

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

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

so that $v'(0) = -\alpha$, and $v(0) = -\beta$. But, minding that $v \in H_0^2(\Omega)$, the function v is differentiable in the distributional sense 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

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

depicted in Figure 3. The function f is absolutely continuous with countably many affine parts, but $Df' = \sum_{i=2}^{\infty}(\delta_{2\cdot 2^{-i}} - \delta_{3\cdot 2^{-i}})$, 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 $\|Df - w\|_{\mathcal{M}(\Omega)} + \|Dw\|_{\mathcal{M}(\Omega)}$. We conclude that $u = f$ cannot solve (P).

References

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

[2] H. Attouch and H. Brezis. Duality for the sum of convex functions in general Banach spaces. In Jorge Alberto Barroso, editor, *Aspects of Mathematics and its Applications*, volume 34 of *North-Holland Mathematical Library*, pages 125–133. Elsevier, 1986.

[3] K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM J. Imaging Sciences*, 3:492–526, 2011.

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

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

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

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

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

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

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

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

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

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