National Research Network S92

Industrial Geometry

http://www.industrial-geometry.at

NRN Report No. 103

A Coarea Formula for Anisotropic Total Variation Regularisation

Markus Grasmair

May 2010

Der Wissenschaftsfonds.

A Coarea Formula for Anisotropic Total Variation Regularisation

Markus Grasmair

Computational Science Center
University of Vienna
Nordbergstrasse 15
1090 Wien, Austria

May 10, 2010

Abstract

The coarea formula is an important tool in the analysis of total variation based regularisation methods. It states that the total variation of a function can be computed by integrating the lengths of its level lines. In this note we show that the same formula can be extended to anisotropic total variation, the anisotropy also being dependent on the location in space. As one application, we prove the equivalence of two seemingly different regularisation methods that have been proposed for the removal of multiplicative noise in gray value images.

1 Introduction

Let $\Omega \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be some open and bounded domain, and let $u \in L^1(\Omega)$. The *total variation* of u is defined as

$$|Du|(\Omega) = \sup\left\{ \int_\Omega u(x) \,\operatorname{div} \phi(x)\, dx : \phi \in C^1(\Omega;\mathbb{R}^n),\ |\phi(x)| \le 1 \text{ in } \Omega \right\}.$$

In this expression, $|\phi(x)|$ stands for the Euclidean norm of the vector $\phi(x) \in \mathbb{R}^n$. In case $|Du|(\Omega) < \infty$, there exists a vector valued Radon measure $Du \in \mathcal{M}(\Omega;\mathbb{R}^n)$ satisfying $\int_\Omega u(x) \,\operatorname{div}\phi(x)\, dx = -\int_\Omega \phi(x)\, dDu$ for every function $\phi \in C_c^1(\Omega;\mathbb{R}^n)$. We denote by $BV(\Omega)$ the space of all functions u of bounded variation, that is, $u \in BV(\Omega)$ if and only if $u \in L^1(\Omega)$ and $|Du|(\Omega) < \infty$.

The total variation of u can be computed by summing the perimeters of all level sets of u. More precisely, we define the *perimeter in* Ω of a measurable set $U \subset \Omega$ as

$$\operatorname{Per}(U;\Omega) := |D\chi_U|(\Omega)\,,$$

where χ_U denotes the characteristic function of U, that is, $\chi_U(x) = 1$ for $x \in U$ and $\chi_U(x) = 0$ for $x \notin U$. Then the *coarea formula*

$$|Du|(\Omega) = \int_{-\infty}^{\infty} \operatorname{Per}\bigl(\{u \ge t\};\Omega\bigr)\, dt$$

holds. The significance of the coarea formula lies in the fact that it permits to identify a function $u \in BV(\Omega)$ with the family of all level sets $\{u \geq t\}$, $t \in \mathbb{R}$, while still retaining access to its total variation. This identification leads to various interesting theoretical but also numerical results.

One important application of functions of bounded variation is the field of image processing: Because a function $u \in BV(\Omega)$ may contain discontinuities of dimension $n-1$, it has been argued that the space $BV(\Omega)$ is well suited for modelling natural images, which basically consist of rather uniform (textured) regions well separated by significant edges. For instance, one of the most widely used techniques in image restoration is total variation regularisation, introduced in [15] for the problem of image denoising and generalised in [1] to the stable solution of inverse problems. Given a noisy image, modelled as function $f \in L^2(\Omega)$, one minimises, for some regularisation parameter $\alpha > 0$, the functional

$$\mathcal{T}(u) = \|u - f\|_{L^2}^2 + \alpha |Du|(\Omega) \; . \tag{1}$$

Then the minimiser u_α of $\mathcal{T}$ contains the most important features of the data f while noise, consisting of small oscillations in u, is removed.

The coarea formula can be used to derive regularity properties of the solution of total variation regularisation. For instance, it has been used in [2] (see also [6]) to derive regularity results for the level lines of u_α. Similarly, a consequence of the coarea formula has been used in [7] for showing that total variation regularisation creates no new discontinuities, that is, the jump set of u_α is contained in the jump set of f. Finally, the (discrete) coarea formula is the basis of *graph cut algorithms* (see [8, 9]), a class of interesting methods for the numerical minimisation of $\mathcal{T}$.

Recently, there have been extensions of the standard functional (1), where the isotropic total variation term $|Du|(\Omega)$ is replaced by an anisotropic, space dependent term (see for instance [4, 13, 18]). In this note we will show that the coarea formula can naturally be extended to this situation. As a particular application, we prove the equivalence of two total variation based regularisation methods that have been proposed for the restoration of images corrupted by Poisson noise. This equivalence result is based on the invariance of the sub-differential of the total variation under contrast changes, which in turn is a consequence of the (anisotropic) coarea formula.

2 An Anisotropic Coarea Formula

Let $\Omega \subset \mathbb{R}^n$, $n \geq 2$, be an open and bounded set with Lipschitz boundary. Let $\alpha \colon \Omega \times \mathbb{R}^n \to \mathbb{R}_{>0}$ satisfy the following conditions:

($A1$) The function α is continuous.

($A2$) For every $x \in \Omega$ the function $A \mapsto f(x, A)$ is convex and positively homogeneous.

($A3$) There exists $C > 1$ such that

$$C^{-1}|A| \leq \alpha(x, A) \leq C|A|$$

for every $(x, A) \in \Omega \times \mathbb{R}^n$.

For $u \in L^1(\Omega)$ define

$$\int_\Omega \alpha(Du) := \begin{cases} \int_\Omega \alpha\Big(x, \dfrac{dDu}{d|Du|}(x)\Big)\, d|Du|\,, & \text{if } u \in BV(\Omega)\,, \\ +\infty\,, \text{if } u \notin BV(\Omega)\,. \end{cases}$$

For $U \subset \Omega$ define

$$\operatorname{Per}(U;\alpha;\Omega) := \begin{cases} \int_\Omega \alpha(D\chi_U)\,, & \text{if } \chi_U \in BV(\Omega)\,, \\ +\infty\,, & \text{if } \chi_U \notin BV(\Omega)\,. \end{cases}$$

Here χ_U denotes the characteristic function of U, that is, $\chi_U(x) = 1$ if $x \in U$ and $\chi_U(x) = 0$ if $x \notin U$. For every set U of finite perimeter we have

$$\operatorname{Per}(U;\alpha;\Omega) = \int_{\partial^* U \cap \Omega} \alpha\big(x, \nu_U(x)\big)\, d\mathcal{H}^{n-1}\,,$$

where $\nu_U(x)$ denotes the inner normal to U at $x \in \partial^* U$. Here $\partial^* U$ denotes the reduced boundary of U (cf. [3, Def. 3.54]).

Lemma 1. *Let $u \in BV(\Omega)$. Then*

$$\int_\Omega \alpha(Du) \\ = \sup\Big\{-\int_\Omega u \operatorname{div}\phi\, dx : \phi \in C_c^1(\Omega;\mathbb{R}^n),\ \alpha^0\big(x,\phi(x)\big) \le 1 \text{ for all } x \in \Omega\Big\}\,. \qquad (2)$$

Here, α^0 denotes the polar of the integrand α, which is defined as

$$\alpha^0(x,\xi) := \max\big\{\theta \cdot \xi : \theta \in \mathbb{R}^n,\ \alpha(x,\theta) \le 1\big\}\,.$$

Proof. For every $\phi \in C_c^1(\Omega;\mathbb{R}^n)$ we have

$$-\int_\Omega u(x) \operatorname{div}\phi(x)\, dx = \int_\Omega \phi(x)\, dDu = \int_\Omega \phi(x) \cdot \frac{dDu}{d|Du|}(x)\, d|Du|\,.$$

Moreover,

$$\alpha(x,\theta) = \max\big\{\theta \cdot \xi : \xi \in \mathbb{R}^n,\ \alpha^0(x,\xi) \le 1\big\}$$

for every $x \in \Omega$ and $\theta \in \mathbb{R}^n$. This shows that $\int_\Omega \alpha(Du)$ is larger or equal than the right hand side of (2).

In order to show the converse inequality let $\varepsilon > 0$. There exists a compact set $K \subset \Omega$ such that $|Du|(\Omega \setminus K) \le \varepsilon$ and $dDu/d|Du|$ is continuous on K. Consequently, the function

$$x \mapsto \alpha^0\Big(x, \frac{dDu}{d|Du|}(x)\Big)$$

is continuous on K. Thus, there exists a function $\phi \in C_c^1(\Omega;\mathbb{R}^n)$ satisfying $\alpha^0\big(x,\phi(x)\big) \le 1$ for every $x \in \Omega$ and

$$\phi(x) \cdot \nabla u(x) \ge \alpha\Big(x, \frac{dDu}{d|Du|}(x)\Big) - \varepsilon$$

for every $x \in K$. Consequently,

$$\int_\Omega \alpha(Du) \le \int_K \alpha(Du) + C\varepsilon \le \int_K \phi(x) \cdot \frac{dDu}{d|Du|}(x)\, d|Du| + \varepsilon(C + \mathcal{L}^n(\Omega))$$
$$\le \int_\Omega \phi(x) \cdot \frac{dDu}{d|Du|}(x)\, d|Du| + \varepsilon(2C + \mathcal{L}^n(\Omega))$$
$$= -\int_\Omega u(x)\ \operatorname{div} \phi(x)\, dx + (2C + \mathcal{L}^n(\Omega))\varepsilon\ .$$

Since ε was arbitrary, this proves the assertion. □

Lemma 2. *Let $u \in BV(\Omega)$. Then there exists a sequence $(u_k)_{k\in\mathbb{N}} \subset C^\infty(\Omega) \cap BV(\Omega)$ such that $\|u_k - u\|_1 \to 0$ and $\int_\Omega \alpha(Du_k) \to \int_\Omega \alpha(Du)$.*

Proof. Let $(u_k)_{k\in\mathbb{N}} \subset C^\infty(\Omega) \cap BV(\Omega)$ strictly converge to u, that is, $\|u_k - u\|_1 \to 0$ and $|Du_k|(\Omega) \to |Du|(\Omega)$. Such a sequence exists by [3, Thm. 3.9]. Then the Reshetnyak Continuity Theorem [3, Thm. 2.39] implies that $\int_\Omega \alpha(Du_k) \to \int_\Omega \alpha(Du)$, which proves the assertion. □

Theorem 3. *Let $u \in BV(\Omega)$. Then*

$$\int_\Omega \alpha(Du) = \int_{-\infty}^{+\infty} \operatorname{Per}\big(\{u \ge t\}; \alpha; \Omega\big)\, dt\ . \tag{3}$$

Proof. We follow the presentation of the isotropic case in [3, Thm. 3.40].

Recall the coarea formula

$$\int_\Omega g(x)|\nabla u(x)|\, dx = \int_{\mathbb{R}} \int_{u^{-1}(t)} g(x)\, d\mathcal{H}^{n-1}\, dt\,,$$

for $u\colon \Omega \to \mathbb{R}$ Lipschitz and $g\colon \Omega \to \bar{\mathbb{R}}$ integrable (see [11, Thm. 3.2.12]). Applying this result with $g(x) = \alpha\big(x, \nabla u(x)\big)/|\nabla u(x)|$, one obtains the assertion for every Lipschitz function $u\colon \Omega \to \mathbb{R}$.

Now let $u \in BV(\Omega)$ be arbitrary. Let $\phi \in C_c^1(\Omega; \mathbb{R}^n)$ satisfying $\alpha^0\big(x, \phi(x)\big) \le 1$ for every $x \in \Omega$. Then we obtain, by using Lemma 1, that

$$-\int_\Omega u(x)\ \operatorname{div} \phi(x)\, dx = -\int_\Omega \int_{-\infty}^{+\infty} \chi_{\{u\ge t\}}(x)\ \operatorname{div} \phi(x)\, dt\, dx$$
$$= -\int_{-\infty}^{+\infty} \chi_{\{u\ge t\}}(x)\ \operatorname{div} \phi(x)\, dx\, dt \le \int_{-\infty}^{+\infty} \operatorname{Per}\big(\{u \ge t\}; \alpha; \Omega\big)\, dt\ .$$

Again from Lemma 1 it follows that $\int_\Omega \alpha(Du)$ is smaller or equal than the right hand side of (3).

Now let $(u_k)_{k\in\mathbb{N}} \subset C^\infty(\Omega) \cap BV(\Omega)$ be a sequence satisfying $\|u_k - u\|_1 \to 0$ and $\int_\Omega \alpha(Du_k) \to \int_\Omega \alpha(Du)$. Such a sequence exists by Lemma 2. Possibly passing to a subsequence, it follows that $\chi_{\{u_k\ge t\}}$ converges to $\chi_{\{u\ge t\}}$ with respect to the L^1-norm for almost every $t \in \mathbb{R}$. Therefore

$$\int_{-\infty}^{+\infty} \operatorname{Per}\big(\{u \ge t\}; \alpha; \Omega\big)\, dt \le \int_{-\infty}^{+\infty} \liminf_{k\to\infty} \operatorname{Per}\big(\{u_k \ge t\}; \alpha; \Omega\big)\, dt$$
$$\le \liminf_{k\to\infty} \int_{-\infty}^{+\infty} \operatorname{Per}\big(\{u_k \ge t\}; \alpha; \Omega\big)\, dt = \lim_{k\to\infty} \int_\Omega \alpha(Du_k) = \int_\Omega \alpha(Du)\ .$$

This proves the assertion. □

3 Characterisation of the Subdifferential

In the following we denote by $\mathcal{R}_\alpha\colon L^p(\Omega) \to [0,+\infty]$ the functional

$$\mathcal{R}_\alpha(u) = \int_\Omega \alpha(Du)\ .$$

Lemma 4. *Let $u \in L^p(\Omega) \cap BV(\Omega)$ and $\xi \in L^{p_*}(\Omega)$. Then $\xi \in \partial\mathcal{R}_\alpha(u)$ if and only if*

$$\int_U \xi(x)\,dx \le \operatorname{Per}(U;\alpha;\Omega) \tag{4}$$

for all Borel sets $U \subset \Omega$, and

$$\int_{\{u\ge t\}} \xi(x)\,dx = \operatorname{Per}\big(\{u \ge t\};\alpha;\Omega\big) \tag{5}$$

for almost every $t \in \mathbb{R}$.

Proof. The proof is along the lines of [16, Thm. 4.41], where the analogous result is shown for the isotropic total variation. We still provide the proof for the sake of completeness.

The inclusion $\xi \in \partial\mathcal{R}_\alpha(u)$ is equivalent to

$$\int_\Omega \xi(x)\big(v(x)-u(x)\big)\,dx \le \mathcal{R}_\alpha(v) - \mathcal{R}_\alpha(u) \tag{6}$$

for every $v \in L^p(\Omega) \cap BV(\Omega)$.

We first show that $\xi \in \partial\mathcal{R}_\alpha(u)$ implies (4) and (5). Assume therefore that (6) holds. Let $U \subset \Omega$ be such that $\operatorname{Per}(U;\alpha;\Omega) = \mathcal{R}_\alpha(\chi_U) < \infty$ and $|Du|(\partial^* U) = 0$. Define for $\varepsilon > 0$ the function $v_\varepsilon := u + \varepsilon\chi_U$. Then (6) implies that

$$\varepsilon\int_U \xi(x)\,dx \le \varepsilon\operatorname{Per}(U;\alpha;\Omega)$$

Dividing by ε and letting $\varepsilon \to 0^+$, the inequality (4) follows. For general $U \subset \Omega$ with $\operatorname{Per}(U;\alpha;\Omega) < +\infty$, the claim follows by approximation with sets U_k satisfying $|Du|(\partial^* U_k) = 0$ and $\operatorname{Per}(U_k;\alpha;\Omega) \to \operatorname{Per}(U;\alpha;\Omega)$.

Now define for $t \in \mathbb{R}$ and $\varepsilon > 0$ the function

$$v_{\varepsilon,t}(x) := \begin{cases} u(x)\,, & \text{if } u(x) \le t\,,\\ t\,, & \text{if } t \le u(x) \le t+\varepsilon\,,\\ u(x)-\varepsilon\,, & \text{if } u(x) \ge t+\varepsilon\ .\end{cases}$$

For almost every $t \in \mathbb{R}$ and $\varepsilon > 0$ we have

$$\int_\Omega \xi(x)\big(v_{\varepsilon,t}(x)-u(x)\big)\,dx \ge -\varepsilon\int_{\{u\ge t+\varepsilon\}} \xi(x)\,dx - \varepsilon\int_{\{t<u<t+\varepsilon\}} |\xi(x)|\,dx\ .$$

It follows from Theorem 3 and the definition of $v_{\varepsilon,t}$ that for almost every $t \in \mathbb{R}$

$$\begin{aligned}\mathcal{R}_\alpha(v_{\varepsilon,t}) &= \int_{-\infty}^{\infty} \operatorname{Per}\big(\{v_{\varepsilon,t} \ge s\};\alpha;\Omega\big)\,ds\\ &= \int_{-\infty}^{t} \operatorname{Per}\big(\{u \ge s\};\alpha;\Omega\big)\,ds + \int_{t+\varepsilon}^{\infty} \operatorname{Per}\big(\{u \ge s\};\alpha;\Omega\big)\,ds\\ &= \mathcal{R}_\alpha(u) - \int_t^{t+\varepsilon} \operatorname{Per}\big(\{u \ge s\};\alpha;\Omega\big)\,ds\ .\end{aligned}$$

Thus, (6) implies that

$$-\int_{\{u\geq t+\varepsilon\}}\xi(x)\,dx \leq \int_{\{t<u<t+\varepsilon\}}|\xi(x)|\,dx - \frac{1}{\varepsilon}\int_t^{t+\varepsilon}\operatorname{Per}\bigl(\{u\geq s\};\alpha;\Omega\bigr)\,ds\,.$$

The passage to the limit $\varepsilon \to 0$ therefore implies (5) for almost every $t \in \mathbb{R}$.

Now assume that (4) and (5) hold. Passing to the limit $t \to -\infty$, equation (5) implies that $\int_\Omega \xi(x)\,dx = 0$. Now let $v \in BV(\Omega) \cap L^{p_*}(\Omega)$. In order to show that $\xi \in \partial\mathcal{R}_\alpha(u)$ we have to verify (6).

Since $\int_\Omega \xi(x)\,dx = 0$, it follows from Fubini's Theorem that

$$\int_\Omega \xi(x)\bigl(v(x)-u(x)\bigr)\,dx = \int_{-\infty}^{\infty}\biggl(\int_{\{v\geq t\}}\xi(x)\,dx - \int_{\{u\geq t\}}\xi(x)\,dx\biggr)\,dt\,.$$

This equation together with (4) applied to the sets $\{v \geq t\}$, (5) applied to $\{u \geq t\}$, and Theorem 3 implies that

$$\begin{aligned}&\int_\Omega \xi(x)\bigl(v(x)-u(x)\bigr)\,dx\\ &\qquad\leq \int_{-\infty}^{\infty}\Bigl(\operatorname{Per}\bigl(\{v\geq t\};\alpha;\Omega\bigr)-\operatorname{Per}\bigl(\{u\geq t\};\alpha;\Omega\bigr)\Bigr)\,dt = \mathcal{R}_\alpha(v)-\mathcal{R}_\alpha(u)\,,\end{aligned}$$

which proves the inequality (6). □

As a particular consequence of the last theorem, we obtain that the subdifferential of $\mathcal{R}_\alpha$ is invariant under contrast changes.

Lemma 5. *Let $g\colon \mathbb{R} \to \mathbb{R}$ be a non-decreasing locally absolutely continuous function. If $u \in BV(\Omega)$ is such that $g(u) \in BV(\Omega)$, then $\partial\mathcal{R}_\alpha(u) \subset \partial\mathcal{R}_\alpha\bigl(g(u)\bigr)$.*

Proof. Let $\xi \in \partial\mathcal{R}_\alpha(u)$. Lemma 4 implies that $\int_U \xi \leq \operatorname{Per}(U;\alpha;\Omega)$ for every measurable $U \subset \Omega$. Thus it remains to show that the set

$$E := \Bigl\{t \in \mathbb{R} : \int_{\{g(u)\geq t\}}\xi \neq \operatorname{Per}(\{g(u)\geq t\};\alpha;\Omega)\Bigr\}$$

satisfies $\mathcal{L}^1(E) = 0$.

Define now the function $h\colon \mathbb{R} \to \mathbb{R}\cup\{-\infty\}$ by $h(t) := \inf\bigl\{s : g(s) \geq t\bigr\}$. Then h is a strictly increasing function and

$$\bigl\{x : g\circ u(x) \geq t\bigr\} = \bigl\{x : u(x) \geq h(t)\bigr\} \tag{7}$$

for every $t \in \mathbb{R}$. Moreover, $g \circ h(t) = t$ for every $t \in g(\mathbb{R})$, that is, h is a right inverse of g. Because $\xi \in \partial\mathcal{R}_\alpha(u)$, it follows from Lemma 4 that the set

$$F := \Bigl\{t \in \mathbb{R} : \int_{\{u\geq t\}}\xi \neq \operatorname{Per}(\{u\geq t\};\alpha;\Omega)\Bigr\}$$

satisfies $\mathcal{L}^1(F) = 0$. Because of (7) we have that $h(E) \subset F$. Consequently, the fact that h is a right inverse of g and $E \subset g(\mathbb{R})$ implies that $E = g \circ h(E) \subset g(F)$. Since g is locally absolutely continuous and $\mathcal{L}^1(F) = 0$, it follows that $\mathcal{L}^1(E) \leq \mathcal{L}^1\bigl(g(F)\bigr) = 0$, which proves the assertion. □

4 Rescaling of Minimisation Problems

In [19], the removal of multiplicative noise in images using a total variation regularisation term has been studied. In particular, the authors compare two seemingly different methods that both can be motivated from a *maximum a–posteriori* (MAP) estimation approach for the removal of Poisson noise. Given noisy data $f\colon \Omega \to \mathbb{R}_{>0}$, the application of MAP estimation suggests the similarity term

$$I^{(1)}(u) := \int_\Omega f \log\bigl(f/u\bigr) - f + u\,dx\,,$$

which is the Bregman distance of the Boltzmann–Shannon entropy. Using an isotropic total variation regularisation term, one then obtains, for any regularisation parameter $\alpha > 0$, the function

$$u_\alpha := \arg\min_u \Bigl[I^{(1)}(u) + \alpha|Du|(u)\Bigr] \tag{8}$$

as a guess of the clean underlying image.

The approach in [17] uses the rescaling $w := \log u$ and arrives, discounting constant terms, at the similarity term

$$I^{(2)}(w) := \int_\Omega f\,e^{-w} + w\,dx\,.$$

Again applying isotropic total variation regularisation, one can then define the clean image as

$$u_\alpha := e^{w_\alpha} \qquad \text{with} \qquad w_\alpha \in \arg\min_w \Bigl[I^{(2)}(w) + \alpha|Du|(w)\Bigr]\,. \tag{9}$$

Because the regularisation is applied to u_α in the first functional and to $\log u_\alpha$ in the second one, one might expect, at first glance, that the two methods could give different results. Lemma 5, however, states that the subdifferential of the total variation is invariant under rescalings. Using this invariance, it is then easy to see that, indeed, the equation $u_\alpha = e^{-w_\alpha}$ holds, provided one can write the subdifferentials of the functionals appearing in (8) and (9) as sum of the subdifferentials of the total variation and the respective similarity term (a detailed argument is given below in Lemma 7). Standard prerequisites for this splitting of the subdifferential, however, are not satisfied in this setting, as the domains of both similarity and regularisation term have empty interior. The result below shows that it is nevertheless possible to prove the equivalence of the two regularisation methods. This is achieved by approximating the similarity terms from below by more regular functions for which the addition theorem for the subdifferential applies. A shortened version of the result, only applicable for data f satisfying $0 < \operatorname{ess\,inf} f \le \operatorname{ess\,sup} f < +\infty$, has appeared in [19].

Assume for the following that the set Ω is connected. Let ϕ, $\psi\colon \Omega \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, and assume that there exist possibly unbounded, non-empty, open intervals I_ϕ, I_ψ such that $\operatorname{dom}\phi = \Omega \times I_\phi$ and $\operatorname{dom}\psi = \Omega \times I_\psi$.

Assume, in addition to assumptions $(A1)$–$(A3)$, that the following hold:

$(B1)$ The functions ϕ and ψ are convex Carathéodory integrands.

(*B*2) There exists $\gamma \in L^1(\Omega)$ with $\gamma \geq 0$ almost everywhere, such that

$$\phi(x,s) \geq -\gamma(x) \qquad \text{and} \qquad \psi(x,s) \geq -\gamma(x)$$

for almost every $x \in \Omega$ and every $s \in \mathbb{R}$.

(*B*3) There exist a set $E \subset \Omega$ and a constant $D > 0$ such that $\mathcal{L}^n(E) > 0$ and

$$\phi(x,s) \geq D|s| - \gamma(x)$$

for every $x \in E$ and $s \in \mathbb{R}$.

(*B*4) There exists a locally Lipschitz, non-decreasing mapping $g\colon I_\phi \to I_\psi$ such that

$$\partial\phi(x,s) \subset \partial\psi\big(x, g(s)\big) \tag{10}$$

for almost every $x \in \Omega$ and $s \in I_\phi$.

We define the functionals $\mathcal{T}_\phi$, $\mathcal{T}_\psi\colon L^1(\Omega) \to \mathbb{R} \cup \{+\infty\}$ by

$$\begin{aligned} \mathcal{T}_\phi(u) &:= \mathcal{S}_\phi(u) + \mathcal{R}_\alpha(u) := \int_\Omega \phi\big(x, u(x)\big)\, dx + \int_\Omega \alpha(Du)\,, \\ \mathcal{T}_\psi(u) &:= \mathcal{S}_\psi(u) + \mathcal{R}_\alpha(u) := \int_\Omega \psi\big(x, u(x)\big)\, dx + \int_\Omega \alpha(Du)\,. \end{aligned}$$

Lemma 6. *Let Assumptions* (*B*1), (*B*2), *and* (*B*4) *hold. Then*

$$\partial\mathcal{S}_\phi(u) \subset \partial\mathcal{S}_\psi(g \circ u)$$

for every $u \in L^1(\Omega)$.

Proof. This follows from (10) and the characterisation of the subdifferential of an integral functional (see [14, Cor. 3E]). □

Lemma 7. *Let Assumptions* (*B*1)–(*B*4) *hold. Assume in addition that* $I_\phi = \mathbb{R}$, *g is Lipschitz, and there exist* $D' > 0$ *and* $\gamma' \in L^1(\Omega)$ *with* $\gamma' \geq 0$ *almost everywhere such that*

$$\phi(x,s) \leq D'|s| + \gamma(x) \tag{11}$$

for almost every $x \in \Omega$ *and every* $s \in \mathbb{R}$. *Let* u_ϕ *be a minimiser of* $\mathcal{T}_\phi$. *Then* $g \circ u_\phi$ *minimises* $\mathcal{T}_\psi$.

Proof. Because of (11), the functional $\mathcal{S}_\phi\colon L^1(\Omega) \to [0, +\infty]$ is continuous [12, Cor. 6.51]. Since u_ϕ minimises $\mathcal{T}_\phi$, we therefore have

$$0 \in \partial\mathcal{T}_\phi(u_\phi) = \partial\mathcal{S}_\phi(u_\phi) + \partial\mathcal{R}_\alpha(u_\phi)\,.$$

Lemma 6 implies that $\partial\mathcal{S}_\phi(u_\phi) \subset \partial\mathcal{S}_\psi(g \circ u_\phi)$. Using Lemma 5, we obtain that $\partial\mathcal{R}(u_\phi) \subset \partial\mathcal{R}\big(g \circ u_\phi\big)$. Thus

$$0 \in \partial\mathcal{S}_\phi(u_\phi) + \partial\mathcal{R}_\alpha(u_\phi) \subset \partial\mathcal{S}_\psi(g \circ u_\phi) + \partial\mathcal{R}_\alpha(u_\phi) \subset \partial\mathcal{T}_\psi(g \circ u_\phi)\,,$$

showing that $g \circ u_\phi$ minimises $\mathcal{T}_\psi$. □

Theorem 8. *Let Assumptions* (*B*1)–(*B*4) *hold. Then there exists a minimiser* u_ϕ *of* $\mathcal{T}_\phi$ *such that* $g \circ u_\phi$ *minimises* $\mathcal{T}_\psi$.

Proof. Because ϕ is a Carathéodory integrand, there exists an increasing sequence of compact sets $E_k \subset \Omega$ with $\mathcal{L}^n(\Omega \setminus E_k) < 1/k$ such that the restriction of ϕ to $E_k \times \mathbb{R}$ is continuous (see [12, Thm. 6.35]). Moreover, there exists an increasing sequence $[a_k, b_k] \subset I_\phi$ of compact intervals such that $\bigcup_k [a_k, b_k] = I_\phi$. For $x \in E_k$ define

$$\zeta_k(x) := \min\{\inf \partial\phi(x, a_k), 0\} \qquad \text{and} \qquad \xi_k(x) := \max\{\sup \partial\phi(x, a_k), 0\} .$$

Define the functions

$$\phi^{(k)}(x,s) := \begin{cases} -\gamma(x)\,, & \text{if } x \notin E_k\,,\\ \phi(x,a_k) + (s-a_k)\zeta_k(x)\,, & \text{if } x \in E_k \text{ and } s < a_k\,,\\ \phi(x,s)\,, & \text{if } x \in E_k \text{ and } a_k \le s \le b_k\,,\\ \phi(x,b_k) + (s-b_k)\xi_k(x)\,, & \text{if } x \in E_k \text{ and } b_k < s\,, \end{cases}$$

$$\psi^{(k)}(x,s) := \begin{cases} -\gamma(x)\,, & \text{if } x \notin E_k\,,\\ \psi\big(x,g(a_k)\big) + \big(s-g(a_k)\big)\zeta_k(x)\,, & \text{if } x \in E_k \text{ and } s < g(a_k)\,,\\ \psi(x,s)\,, & \text{if } (x,s) \in E_k \times [g(a_k), g(b_k)]\,,\\ \psi\big(x,g(b_k)\big) + \big(s-g(b_k)\big)\xi_k(x)\,, & \text{if } x \in E_k \text{ and } g(b_k) < s\,, \end{cases}$$

$$g^{(k)}(s) := \begin{cases} g(a_k) + (s-a_k)\,, & \text{if } s < a_k\,,\\ g(s)\,, & \text{if } a_k \le s \le b_k\,,\\ g(b_k) + (s-b_k)\,, & \text{if } b_k < s\,. \end{cases}$$

Then the functions $\phi^{(k)}$ and $\psi^{(k)}$ satisfy the assumptions $(B1)$ and $(B2)$ and $g^{(k)}$ is a Lipschitz function for every k. Moreover, we have

$$\partial\phi^{(k)}(x,s) \subset \partial\psi^{(k)}\big(x, g^{(k)}(s)\big)$$

for almost every $x \in \Omega$ and every $s \in \mathbb{R}$. Finally, the compactness of the set $E_k \times [a_k, b_k]$ implies that

$$D_k := \sup\big\{\max\{|\zeta_k(x)|, |\xi_k(x)|\} : x \in E_k\big\} < +\infty .$$

Consequently, there exists a non-negative function $\gamma_k \in L^1(\Omega)$ such that

$$\phi^{(k)}(x,s) \le D_k|s| + \gamma_k(x)$$

for almost every $x \in \Omega$ and $s \in \mathbb{R}$.

Now let $k' \in \mathbb{N}$ be such that $\mathcal{L}^n(E) > 1/k'$ and define $E' := E \cap E_{k'}$. Then $\mathcal{L}^n(E') > 0$. Since by assumption $\phi(x,s) \ge D|s| - \gamma(x)$ on $E' \times \mathbb{R}$, there exists $k'' \in \mathbb{N}$ such that $\zeta_{k''}(x) \le -D/2$ and $\xi_{k''}(x) \ge D/2$ for every $x \in E'$. Thus the construction of the functions $\phi^{(k)}$ implies the existence of a non-negative function $\gamma' \in L^1(\Omega)$ such that

$$\phi^{(k)}(x,s) \ge D|s|/2 - \gamma'(x) \tag{12}$$

for every $x \in E'$, $s \in \mathbb{R}$, and $k \ge k''$. Now recall that the Poincaré inequality (see [21, Thm. 5.12.7]) implies the existence of a constant $C' > 0$ such that

$$\|u - \mu_{E'}(u)\|_{n/(n-1)} \le C'|Du|(\Omega)$$

for every $u \in L^1(\Omega)$, where

$$\mu_{E'}(u) := \fint_{E'} u(x)\,dx := \frac{1}{\mathcal{L}^n(E')}\int_{E'} u(x)\,dx\;.$$

Moreover, we have by assumption 2 that

$$|Du|(\Omega) \le C\mathcal{R}_\alpha(u)\;.$$

Thus,

$$\begin{aligned}
\|u\|_1 &\le \|u - \mu_{E'}(u)\|_1 + \mathcal{L}^n(\Omega)|\mu_{E'}(u)| \\
&\le \mathcal{L}^n(\Omega)^{1/n} + \|u - \mu_{E'}(u)\|_{n/(n-1)} + \mathcal{L}^n(\Omega)\fint_{E'} |u(x)|\,dx \\
&\le \mathcal{L}^n(\Omega)^{1/n} + C'|Du|(\Omega) + \mathcal{L}^n(\Omega)\fint_{E'} |u(x)|\,dx \\
&\le \mathcal{L}^n(\Omega)^{1/n} + C'C\mathcal{R}_\alpha(u) + \mathcal{L}^n(\Omega)\fint_{E'} |u(x)|\,dx
\end{aligned}$$

for every $u \in L^1(\Omega)$.

Define now

$$\begin{aligned}
\mathcal{T}_\phi^{(k)}(u) &:= \mathcal{S}_\phi^{(k)}(u) + \mathcal{R}_\alpha(u) := \int_\Omega \phi^{(k)}\big(x, u(x)\big)\,dx + \int_\Omega \alpha(Du)\,, \\
\mathcal{T}_\psi^{(k)}(u) &:= \mathcal{S}_\psi^{(k)}(u) + \mathcal{R}_\alpha(u) := \int_\Omega \psi^{(k)}\big(x, u(x)\big)\,dx + \int_\Omega \alpha(Du)\;.
\end{aligned}$$

Now (12) and Assumption $(B2)$ imply that

$$\mathcal{S}_\phi^{(k)}(u) \ge \frac{D}{2}\int_{E'} |u(x)|\,dx - \|\gamma' + \gamma\|_1 = \frac{C\mathcal{L}^n(E')}{2}\fint_{E'} |u(x)|\,dx - \|\gamma' + \gamma\|_1$$

for every $u \in L^1(\Omega)$ and every $k \ge k''$. Consequently we obtain that

$$\|u\|_1 \le \mathcal{L}^n(\Omega)^{1/n} + C'C\mathcal{T}_\phi^{(k)}(u) + \frac{2\mathcal{L}^n(\Omega)}{D\mathcal{L}^n(E')}\mathcal{T}_\phi^{(k)}(u) + \|\gamma' + \gamma\|_1$$

for every $u \in L^1(\Omega)$ and every $k \ge k''$, which proves that, for $k \ge k''$, the functionals $\mathcal{T}_\phi^{(k)}$ are equi-coercive. A similar argumentation proves that the functionals $\mathcal{T}_\psi^{(k)}$ are equi-coercive.

Because the functions $\phi^{(k)}$ are Carathéodory and $\phi^{(k)}(x,s) \ge -\gamma(x)$ almost everywhere, the functionals $\mathcal{S}_\phi^{(k)}$ are lower semi-continuous. This proves the existence of a minimiser $u_\phi^{(k)}$ of $\mathcal{T}_\phi^{(k)}$ for every $k \ge k''$. Applying Lemma 7, we obtain that $g^{(k)} \circ u_\phi^{(k)}$ minimises $\mathcal{T}_\psi^{(k)}$.

By construction, the sequences $\phi^{(k)}$ and $\psi^{(k)}$ are increasing, non-negative, and converge to ϕ and ψ, respectively. Moreover, the functionals $\mathcal{T}_\phi^{(k)}$ and $\mathcal{T}_\psi^{(k)}$ are lower semi-continuous. Standard properties of the Γ-limit (see for instance [5, Rem. 1.40]) and the Monotone Convergence Theorem imply that

$$\mathcal{T}_\phi = \Gamma\text{-}\lim_k \mathcal{T}_\phi^{(k)} \qquad \text{and} \qquad \mathcal{T}_\psi = \Gamma\text{-}\lim_k \mathcal{T}_\psi^{(k)}\;.$$

Because the functionals $\mathcal{T}_\phi$ and $\mathcal{T}_\psi$ are equicoercive, there exists a sequence of minimisers $u_\phi^{(k)}$ of $\mathcal{T}_\phi^{(k)}$ converging to a minimiser u_ϕ of $\mathcal{T}_\phi$, such that the sequence of functions $g^{(k)} \circ u_\phi^{(k)}$, which minimise $\mathcal{T}_\psi^{(k)}$, converges to a minimiser of u_ψ of $\mathcal{T}_\psi$ (see [5, Thm. 1.21]). Since $u_\psi = \lim_k g^{(k)} \circ u_\phi^{(k)} = g \circ u_\phi$, this proves the assertion. □

Example 9. Let $f \colon \mathbb{R} \to \mathbb{R}$ be a measurable function such that $f(x) > 0$ almost everywhere and $\log f \in L^1(\Omega)$. Define

$$\phi(x,s) := s + f(x)\exp(-s) \qquad \text{and} \qquad \psi(x,s) := s + f(x)\log\Bigl(\frac{f(x)}{s}\Bigr) + f(x)\ .$$

Then ϕ and ψ satisfy the conditions $(B1)$–$(B4)$ with $I_\phi = \mathbb{R}$, $I_\psi = (0,+\infty)$, and $g(s) := \exp(s)$. Since the functions ϕ and ψ are strictly convex, and therefore their minimisers unique, Theorem 8 implies that u_ϕ minimises $\mathcal{T}_\phi$, if and only if $\exp(u_\phi)$ minimises $\mathcal{T}_\psi$. In particular, the denoising methods defined in (8) and (9) are equivalent provided $\log f \in L^1(\Omega)$. ■

5 Conclusion

In this note it has been shown that the coarea formula for the total variation can be extended in a natural way to anisotropic total variation like functionals, where the anisotropy varies continuously in space. As a consequence, it follows that the subdifferential of the (anisotropic) total variation functional is invariant under smooth contrast changes. This result is of particular interest for image processing, where similar invariances are sometimes postulated for geometric evolution equations. As an application, we have shown the equivalence of two methods for the removal of multiplicative noise in images. Also, the result applies to more general pairs of functionals that are related by a condition on the subdifferential. The main difficulty in the proof lies in the fact that standard results from convex analysis seem not to be available because of the lack of continuity of the involved functionals.

Acknowledgement

This work has been supported by the Austrian Science Fund (FWF) within the national research network *Industrial Geometry*, project 9203-N12.

References

[1] R. Acar and C. R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Probl.*, 10(6):1217–1229, 1994.

[2] W. K. Allard. Total variation regularization for image denoising. I. Geometric theory. *SIAM J. Math. Anal.*, 39(4):1150–1190, 2007/08.

[3] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.

[4] B. Berkels, M. Burger, M. Droske, O. Nemitz, and M. Rumpf. Cartoon extraction based on anisotropic image classification. In *Vision, Modeling, and Visualization Proceedings*, pages 293–300, 2006.

[5] A. Braides. Γ-*convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[6] V. Caselles, A. Chambolle, S. Moll, and M. Novaga. A characterization of convex calibrable sets in $\mathbb{R}^N$ with respect to anisotropic norms. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 25(4):803–832, 2008.

[7] V. Caselles, A. Chambolle, and M. Novaga. The discontinuity set of solutions of the TV denoising problem and some extensions. *Multiscale Model. Simul.*, 6(3):879–894, 2007.

[8] A. Chambolle. Total variation minimization and a class of binary MRF models. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, volume 3757 of *Lecture Notes in Computer Vision*, pages 136–152. Springer Berlin/Heidelberg, 2005.

[9] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation. I. Fast and exact optimization. *J. Math. Imaging Vision*, 26(3):261–276, 2006.

[10] A. Dold and B. Eckmann, editors. *Nonlinear Operators and the Calculus of Variations, Bruxelles 1975*. Springer Verlag, Berlin, Heidelberg, New York, 1976.

[11] H. Federer. *Geometric Measure Theory*. Die Grundlehren der Mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[12] I. Fonseca and G. Leoni. *Modern methods in the calculus of variations: L^p spaces*. Springer Monographs in Mathematics. Springer, New York, 2007.

[13] M. Grasmair and F. Lenzen. Anisotropic total variation filtering. Reports of FSP S092 - "Industrial Geometry" 84, University of Innsbruck, Austria, 2009. Accepted for publication in *Appl. Math. Optim.*.

[14] R. T. Rockafellar. Integral functionals, normal integrands and measurable selections. In *[10]*, pages 157–207, 1976.

[15] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1–4):259–268, 1992.

[16] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational methods in imaging*, volume 167 of *Applied Mathematical Sciences*. Springer, New York, 2009.

[17] J. Shi and S. Osher. A nonlinear inverse scale space method for a convex multiplicative noise model. *SIAM J. Imaging Sciences*, 1(3):294–321, 2008.

[18] G. Steidl and T. Teuber. Anisotropic smoothing using double orientations. In *[20]*, pages 477–489, 2009.

[19] G. Steidl and T. Teuber. Removing multiplicative noise by Douglas-Rachford splitting methods. *J. Math. Imaging Vision*, 36(2):168–184, 2010.

[20] X. Tai, K. Mørken, M. Lysaker, and K.-A. Lie, editors. *Scale Space and Variational Methods in Computer Vision*, volume 5567 of *Lecture Notes in Computer Science*, Berlin/Heidelberg, 2009. Springer. Second International Conference, SSVM 2009, Voss, Norway, June 1-5, 2009. Proceedings.

[21] W. P. Ziemer. *Weakly Differentiable Functions. Sobolev Spaces and Functions of Bounded Variation*, volume 120 of *Graduate Texts in Mathematics*. Springer Verlag, Berlin etc., 1989.