COMM. MATH. SCI.
Vol. 3, No. 4, pp. 479–492

SLOPE AND G-SET CHARACTERIZATION OF SET-VALUED FUNCTIONS AND APPLICATIONS TO NON-DIFFERENTIABLE OPTIMIZATION PROBLEMS*

O. SCHERZER†, WOTAO YIN‡, AND S. OSHER§

Abstract. In this paper we derive a generalizing concept of G-norms, which we call G-sets, which is used to characterize minimizers of non-differentiable regularization functionals. Moreover, the concept is closely related to the definition of slopes as published in a recent book by Ambrosio, Gigli, Savaré. A paradigm of regularization models fitting in this framework is *robust* bounded variation regularization. Two essential properties of this regularization technique are documented in the literature and it is shown that these properties can also be achieved with metric regularization techniques.

Key words. G-norm, G-sets, bounded variation regularization, slopes, robust regularization.

AMS subject classifications. 65F22, 65J20, 49J40

1. Introduction

In this work we are concerned with characterization of the minimizers of the *robust regularization* functional

$$\mathcal{F}(u) := \int |u-f| + \alpha \|Du\|, \tag{1.1}$$

and the *quantile regularization* functional

$$\mathcal{F}^\beta(u) := \int S^\beta(u) + \alpha \|Du\|,$$

where

$$S^\beta(v) := \begin{cases} (1-\beta)(f-v) & \text{if} \quad f \geq v, \\ \beta(v-f) & \text{if} \quad f \leq v \end{cases}$$

with $0<\beta<1$ and $\|Du\|$ denoting the total variation semi-norm.

The functional $\mathcal{F}(u)$ has been analyzed by Alliney and Nikolova [1, 6, 8, 7]. Recent attempts in characterizing properties of the minimizers of $\mathcal{F}$ have been made by Chan & Esedoglu [3] and in [9]. In the latter work we characterized minimizers of (1.1) using the G-norm introduced by Y. Meyer [5]. The results essentially apply if the zeros of $u_\alpha - f$ are sparse, where u_α denotes a minimizer of the robust regularization functional. This limits the applicability of the results. In this work we derive a general characterization of the minimizing elements. For this purpose we develop the concept of G-sets and G-values, which is a generalization of Y. Meyer's G-norm to *set valued* functions. In general, for the functional (1.1) the characterization of minimizers is no

*Received: May 25, 2005; accepted (in revised version): July 17, 2005. Communicated by Shi Jin.
†Department of Computer Science, University of Innsbruck, Technikerstr. 21a, A-6020 Innsbruck, Austria (otmar.scherzer@uibk.ac.at).
‡Department of Industrial Engineering and Operations Research, Columbia University, New York, NY 10025, USA (wy2002@columbia.edu).
§Department of Mathematics, University of California, Los Angeles, CA 90095-1555, USA (sjo@math.ucla.edu).

longer possible by the G-norm as for instance for the Rudin-Osher-Fatemi model [10] (cf. Meyer [5]).

Moreover, we show a relation between G-values and *slopes* as introduced recently in [2].

The results of this paper allow us to characterize minimizers of $\mathcal{F}$ in a functional analytical framework, and as a byproduct we can generalize the results of Chan & Esedoglu [3]. Moreover, some of the results can easily be generalized to a wider class of metrical regularization techniques.

2. Basic Facts on Minimizers and Notation

It is relatively easy to show that there exists a minimizer u_α of $\mathcal{F}$ in $\mathtt{BV}$, the *space of functions of bounded variation* (cf. Evans & Gariepy [4]),i.e., the space of functions in L^1 with finite total variation.

Note that the minimizing elements do not have to be unique since the functional is *not* strictly convex.

For $v \in \mathtt{BV}$ we let

$$\begin{aligned} \psi_v(x) &= \begin{cases} \operatorname{sgn}(v(x)-f(x)) \text{ if } v(x)-f(x)\neq 0 \\ \qquad\qquad\qquad 0 \text{ if } v(x)-f(x)=0 \end{cases} \in \\ \Psi_v &= \{\zeta \in L^\infty : \\ &\qquad \zeta(x)=\operatorname{sgn}(v(x)-f(x)) \text{ if } v(x)\neq f(x), \zeta\in[-1,1] \text{ else}\}. \end{aligned} \tag{2.1}$$

Moreover, let

$$\begin{aligned} &\eta : \mathbb{R}\times\mathtt{BV}\times\mathtt{BV}\to\mathbb{R}. \\ &\qquad (t,v,h)\to\int(|v+th-f|-|v-f|-t\psi_v h). \end{aligned} \tag{2.2}$$

LEMMA 2.1. *Assume that $v,h\in\mathtt{BV}$, then*

$$\lim_{t\to 0}\frac{\eta(t,v,h)}{|t|}=\int_{\{v=f\}}|h|. \tag{2.3}$$

Proof. The definition of η implies that

$$\left|\frac{\eta(t,v,h)}{|t|}-\int_{\{v=f\}}|h|\right|\leq 2\int_{\{0<|v-f|\leq|th|\}}|h|.$$

The family of functions $g_{|t|}(x):=|h(x)|\chi_{\{0<|v-f|\leq|t||h|\}}(x)$ is monotonically decreasing in $|t|$ and thus by the monotone convergence theorem

$$\begin{aligned} \lim_{|t|\to 0}\int g_{|t|}(x) &= \int|h(x)|\lim_{|t|\to 0}\chi_{0<|v-f|\leq|th|}(x) \\ &= \int|h(x)|\chi_{M_0}(x) \\ &= 0, \end{aligned}$$

where M_0 is a set of measure 0. This gives the assertion. ∎

As a consequence of the above lemma we have that if $\{v=f\}$ has Lebesgue measure 0, then

$$\frac{|\eta(t,v,h)|}{|t|}\to 0. \tag{2.4}$$

The G-norm of a measureable function $v:\mathbb{R}\to\mathbb{R}$ is defined as the minimum of all values $\lambda>0$ satisfying

$$\left|\int vh\right|\leq\lambda\int|\nabla v|\,,\ \text{where } v\in C_0^\infty\,. \tag{2.5}$$

Using (2.4), we can reinterpret the results in [9], which read as follows:

THEOREM 2.2.

1. *Let $\{0=f\}$ have Lebesgue measure 0. Then $\|\psi_0\|_G\leq\alpha$ if and only if $u_\alpha\equiv 0$. Here $\|\cdot\|_G$ denotes the G-norm of ψ_0.*
2. *Let $\{u_\alpha=f\}$ have Lebesgue measure 0. If $\|\psi_0\|_G>\alpha$, then*

$$\|\psi_{u_\alpha}\|_G=\alpha \text{ and } -\int\psi_{u_\alpha}u_\alpha=\alpha\|Du_\alpha\|\,.$$

In the following we generalize the result of Theorem 2.2 and neglect the assumption that $\{u_\alpha=f\}$ has Lebesgue measure zero.

3. Slopes

Let $\phi:\mathcal{B}\to(-\infty,\infty]$ be an extended real functional on a real Banach space $\mathcal{B}$ with proper domain

$$D(\phi):=\{v\in\mathcal{B}:\phi(v)<\infty\}\neq\emptyset.$$

A metric on $\mathcal{B}$ is denoted by $d(\cdot,\cdot)$.

In [2] the following definitions have been given:

1. Local slope:

$$|\partial\phi|(v):=\limsup_{w\to v}\frac{(\phi(v)-\phi(w))^+}{d(v,w)}\,.$$

2. Global slope:

$$\mathcal{I}_\phi(v):=\sup_{w\neq v}\frac{(\phi(v)-\phi(w))^+}{d(v,w)}\,.$$

The following result from [2, Proposition 1.4.4] is used afterward:

THEOREM 3.1. *Let $\phi:\mathcal{B}\to(-\infty,\infty]$ be a convex and lower semi continuous functional. Then*

$$|\partial\phi|(v)=\min\{\|\zeta\|_{\mathcal{B}^*}:\zeta\in\partial\phi(v)\}=\mathcal{I}_\phi(v)\,,$$

where

$$\partial\phi(v)=\{\zeta\in\mathcal{B}^*:\phi(h)-\phi(v)-\langle\zeta,h-v\rangle\geq 0 \text{ for all } h\in\mathcal{B}\}\,,$$

is the sub-gradient (here $\langle\cdot,\cdot\rangle$ denotes the dual pairing) of ϕ at v and $\mathcal{B}^*$ is the dual of $\mathcal{B}$.

The dual of the Sobolev space $\mathcal{B}:=W^{1,1}_{0,\lambda}$, of absolutely integrable functions with absolute integrable derivatives, is denoted by $\mathcal{B}^*$; the intuitive metric on $\mathcal{B}$ is

$$d(v,h):=\int|\nabla v-\nabla h|+\lambda\int|v-h|;.$$

The functional

$$\begin{aligned}\phi : \mathcal{B} &\to [0,\infty]\\ v &\to \int |v|\end{aligned}$$

is convex and lower semi continuous. Note that in order to be able to define the slope via the minimum, $|\cdot|$ has to be lower semi continuous, which is guaranteed if $\lambda > 0$.

Note, we do not notationally distinguish between sub-differential of functions and operators. We also emphasize that a-priori we do not assume that $\partial\phi(v) \neq \emptyset$. We define

$$D(\partial\phi) := \{v \in \mathcal{B} : \partial\phi(v) \neq \emptyset\}.$$

Since by definition C_0^∞ is dense in $W_{0,\lambda}^{1,1}$ with respect to $\|\cdot\|_\lambda$ we therefore have

$$|\partial\phi|(v) = \inf_{\zeta\in\partial\phi(v)} \sup_{\{h\in C_0^\infty : \|h\|_\lambda \leq 1\}} \int \zeta h\,.$$

From Proposition 1.4.4. in [2] it follows that

$$|\partial\phi|(v) = \mathcal{I}_{|\cdot|}(v) := \sup_{v\neq h\in\mathcal{B}} \frac{\left(\int|v| - \int|h|\right)^+}{d(v,h)}\,. \tag{3.1}$$

We have that

$$\begin{aligned}\mathcal{I}_{|\cdot|}(v) &= |\partial\phi|(v)\\ &\geq \inf_{\zeta\in\partial\phi(v)} \sup_{\{h\in C_0^\infty : \|h\|_\lambda = 1\}} \int_{\{v\neq 0\}} \operatorname{sgn}(v)h + \int_{\{v=0\}} \zeta h\\ &\geq \sup_{\{h\in C_0^\infty : \|h\|_\lambda = 1\}} \left(\left|\int_{\{v\neq 0\}} \operatorname{sgn}(v)h\right| - \int_{\{v=0\}} |h|\right)^+\\ &=: G_\lambda(\partial|v|)\,.\end{aligned}$$

For every $h \in C_0^\infty$

$$\begin{aligned}\int (|v| - |h|) &= \int_{\{v\neq 0\}} |v| - \int_{\{v\neq 0\}} |h| - \int_{\{v=0\}} |v-h|\\ &\leq \int_{\{v\neq 0\}} |v| - \int_{\{v\neq 0\}} \operatorname{sgn}(v)h - \int_{\{v=0\}} |v-h|\\ &\leq \left(\int_{\{v\neq 0\}} \operatorname{sgn}(v)(v-h) - \int_{\{v=0\}} |v-h|\right)^+\\ &\leq G_\lambda(\partial|v|)\left(\int |\nabla(v-h)| + \lambda\int|v-h|\right).\end{aligned}$$

This shows that $\mathcal{I}_{|\cdot|}(v) \leq \alpha = G_\lambda(\partial|v|)$. Combination of the two inequalities above shows that

$$\begin{aligned}\mathcal{I}_{|\cdot|}(v) &= \sup_{v\neq h\in\mathcal{B}} \frac{\left(\int|v| - \int|h|\right)^+}{d(v,h)}\\ &= \sup_{\{h\in C_0^\infty : \|h\|_\lambda = 1\}} \left(\left|\int_{\{v\neq 0\}} \operatorname{sgn}(v)h\right| - \int_{\{v=0\}} |h|\right)^+ = G_\lambda(\partial|v|)\,,\end{aligned} \tag{3.2}$$

or in other words the slope of $|\cdot|$ equals the G_λ value of $\partial|\cdot|$.

We apply Theorem 3.1 to the functional

$$\begin{aligned}\tilde{\phi}: L^1 &\to [0,\infty],\\ u &\to \|Du\|\end{aligned}$$

where $\|Du\|$ is the total variation semi-norm of u if $u \in \mathtt{BV}$ and $+\infty$ else. We use the metric induced by the L^1-norm. In this case we have

$$\left|\partial\tilde{\phi}\right|(v) = \min\{\|\zeta\|_{L^\infty} : \zeta \in \partial\tilde{\phi}(v)\}.$$

$\zeta \in \partial\tilde{\phi}(v)$ satisfies

$$\tilde{\phi}(u) - \tilde{\phi}(v) - \langle \zeta, u-v\rangle \geq 0,$$

where $\langle\cdot,\cdot\rangle$ is the dual pairing between $L^\infty = L^{1*}$ and L^1. Formally, the inequality reads as follows

$$\tilde{\phi}(u) - \tilde{\phi}(v) + \int \nabla\cdot\left(\frac{\nabla v}{|\nabla v|}\right)(u-v) \geq 0.$$

Note, that the sub-gradient could be empty, if there does not exist $\zeta \in L^\infty = L^{1*}$ which formally satisfies $\zeta = -\nabla\cdot\left(\frac{\nabla v}{|\nabla v|}\right)$.

Since the functional $\tilde{\phi}$ is weakly lower semi-continuous (cf. Evans & Gariepy [4]), according to Proposition 1.4.4. in [2]

$$\mathcal{I}_{\tilde{\phi}}(v) := \sup \frac{(\|Dv\| - \|Dh\|)^+}{\int |v-h|} = \left|\partial\tilde{\phi}\right|(v). \tag{3.3}$$

In the following we use *directional derivatives* of a function $\phi : \mathcal{B} \to (-\infty,\infty]$ and define

$$|\partial\phi|(v,h) := \lim_{t\to 0^+} \frac{(\phi(v) - \phi(v+th))^+}{t},$$

provided the limit exists.

EXAMPLE 3.2. *From Lemma 2.1 it follows that for* $\phi(\cdot - f) = |\cdot - f|$

$$\lim_{t\to 0^+} \frac{(\phi(v) - \phi(v+th))^+}{t} = \left(-\int \psi_v h - \int_{\{v=f\}} |h|\right)^+ = |\partial\phi|(v,h).$$

For S^β, $0<\beta<1$ *as in the quantile regularization model we define*

$$\psi_v^\beta(x) = \begin{cases} \beta \text{ if } v-f>0,\\ \beta-1 \text{ if } v-f<0,\\ 0 \text{ if } v-f=0. \end{cases} \tag{3.4}$$

We have

$$\psi_v^\beta \in \Psi_v^\beta := \Big\{\zeta \in L^\infty : \zeta(x) = \beta - \chi_{v<f}(x) \text{ if } v(x) \neq f(x) \\ \text{and } \zeta(x) \in [\beta-1,\beta] \text{ if } v(x) = f(x)\Big\}.$$

In a similar manner, we can prove that the directional slope of S^β at v in direction h is

$$\left(-\int \psi_v^\beta h - \int_{\{v=f\}} \beta(h)h\right)^+ ,$$

where

$$\beta(h) = \begin{cases} \beta & \text{if } h>0 \\ (\beta-1) & \text{if } h<0 . \end{cases}$$

4. G-Values

The following generalizing concepts of the G-norm are relevant for our paper:

Definition 4.1. *Let $\Psi : \mathbb{R}^n \to 2^{\mathbb{R}}$ be a set-valued function (here, as usual $2^{\mathbb{R}}$ denotes the power set of $\mathbb{R}$) and let*

$$\Psi := \{\psi : \mathbb{R}^n \to \mathbb{R} \text{ is measurable and } \psi(x) \in \Psi(x) \text{ almost everywhere}\} \neq \emptyset .$$

Note, that notationally we do not distinguish between the set Ψ and the function Ψ. We define the G-value of Ψ as follows:

$$\begin{aligned} G(\Psi) &:= \sup_{\{h \in C_0^\infty : \int |\nabla h| = 1\}} - \sup_{\psi \in \Psi} \int \psi h \\ &= \sup_{\{h \in C_0^\infty : \int |\nabla h| = 1\}} \inf_{\psi \in \Psi} \int (-\psi)(-h) . \end{aligned} \tag{4.1}$$

Note that for the later identity we have used that for $h \in C_0^\infty$ satisfying $\int |\nabla h| = 1$ also $-h$ satisfies these properties.

Note, that if Ψ is single valued and measurable then $G(\Psi)$ is the G-norm of Ψ. The G-norm is the norm of the dual of the space $W_0^{1,1}$, which is the closure of C_0^∞ with respect to the norm $u \to \int |\nabla u|$. The concept can be modified when the closure of C_0^∞ is taken with respect to the norm

$$\|u\|_\lambda := \int (|\nabla u| + \lambda |u|) ,$$

where $\lambda > 0$.

Definition 4.2. *The G_λ-values of Ψ are defined as*

$$G_\lambda(\Psi) := \sup_{\{h \in C_0^\infty : \|h\|_\lambda = 1\}} - \sup_{\psi \in \Psi} \int \psi h . \tag{4.2}$$

We have proven that for $\lambda > 0$ the slope and G_λ values are identical (cf. (3.2)). For $\lambda = 0$ the definition of slopes is not applicable, since $\|u_n - u\|_{L^1}$ is not lower semi-continuous: Meyer [5] has given an example of a function $u \notin L^1$ satisfying $\|Du\| < \infty$.

From Theorem 3.1 it follows that

$$\begin{aligned} |\partial\phi|(v) &= \min_{\zeta \in \partial\phi(v)} \sup_{\{h \in C_0^\infty : \|h\|_\lambda = 1\}} \int \zeta h \\ &= \sup_{\{h \in C_0^\infty : \|h\|_\lambda = 1\}} \inf_{\psi \in \Psi} \int \psi h \\ &= G_\lambda(\partial |v|) . \end{aligned}$$

This essentially shows that in the definition of slopes and G_λ values the sequence of supremum and infimum is interchanged.

For our application the most important example of a set-valued function is

$$\partial |g| := \begin{cases} \mathrm{sgn}(g) \text{ if } & g \neq 0, \\ [-1,1] \text{ if } & g = 0. \end{cases}$$

In the following we derive some G-value properties of $\partial |g|$.

LEMMA 4.3. *For $g \in L^1$, $G(\partial |g|) \leq \alpha$ if and only if*

$$\left(\left| \int_{\{g \neq 0\}} sgn(g) h \right| - \int_{\{g=0\}} |h| \right)^+ \leq \alpha \|Dh\| \text{ for all } h \in \mathtt{BV}. \tag{4.3}$$

Moreover,

$$G(\partial |g|) = \sup_{\{h \in \mathtt{BV}: \|Dh\|=1\}} - \sup_{\psi \in \Psi} \int \psi h .$$

Proof. Since $h \in \mathtt{BV}$ can be approximated by a sequence of functions $h_n \in C_0^\infty$ satisfying $h_n \to h$ in L^1 and $\int |\nabla h_n| \to \|Dh\|$ it follows that

$$\left| \int_{\{g \neq 0\}} \mathrm{sgn}(g) h_n \right| - \int_{\{g=0\}} |h_n| \to \left| \int_{\{g \neq 0\}} \mathrm{sgn}(g) h \right| - \int_{\{g=0\}} |h| .$$

Therefore (4.3) holds for all $h \in \mathtt{BV}$ if it holds for all $h \in C_0^\infty$.

For $h \in C_0^\infty$ let

$$\psi_h := \mathrm{sgn}(h) \chi_{g=0} - \mathrm{sgn}(g) \chi_{g \neq 0} \in \partial |g| .$$

Therefore,

$$\int \psi_h h = - \int_{\{g \neq 0\}} \mathrm{sgn}(g) h + \int_{\{g=0\}} |h| \geq - \int_{\{g \neq 0\}} \mathrm{sgn}(g) h + \int_{\{g=0\}} \psi h$$

for all $\psi \in \partial |g|$. Therefore

$$\begin{aligned} G(\partial |g|) &= \sup_{\{h \in C_0^\infty : \int |\nabla h| = 1\}} \left(\int_{\{g \neq 0\}} \mathrm{sgn}(g) h - \int_{\{g=0\}} |h| \right)^+ \\ &= \sup_{\{h \in C_0^\infty : \int |\nabla h| = 1\}} \max \left(\int_{\{g \neq 0\}} \mathrm{sgn}(g)(\pm h) - \int_{\{g=0\}} |h| \right)^+ \\ &= \sup_{\{h \in C_0^\infty : \int |\nabla h| = 1\}} \left(\left| \int_{\{g \neq 0\}} \mathrm{sgn}(g) h \right| - \int_{\{g=0\}} |h| \right)^+ . \end{aligned}$$

□

The definition of G-values implies also that for every function $h \in C_0^\infty$

$$\inf_{\psi \in \Psi} \int \psi h = - \sup_{\psi \in \Psi} - \int \psi h \leq G(\Psi) \|D(-h)\| = G(\Psi) \|Dh\| . \tag{4.4}$$

We introduce the definition of G-sets, which is most relevant for our work:

DEFINITION 4.4. *Assume $f \in L^1$ and $u \in \mathtt{BV}$. We define the G-set as*

$$\mathcal{G}_u(\partial|u-f|) := \{\alpha \in [0,\infty] : \alpha \text{ satisfies } (4.6)\} . \tag{4.5}$$

Here for every $h \in \mathtt{BV}$

$$-\int_{\{u\neq f\}} sgn(u-f)h - \int_{\{u=f\}} |h| \leq \alpha(\|D(u+h)\| - \|Du\|) . \tag{4.6}$$

Note that for $\alpha \in \mathcal{G}_u(\partial|u-f|)$ it follows that for every $h \in \mathtt{BV}$

$$\left|\int_{\{u\neq f\}} \mathrm{sgn}(u-f)h\right| - \int_{\{u=f\}} |h| \leq \alpha\|Dh\| ,$$

and thus

$$G(\partial(u-f)) \leq \alpha . \tag{4.7}$$

We also note that (4.6) is equivalent to

$$-\int_{\{u\neq f\}} \mathrm{sgn}(u-f)(v-u) - \int_{\{u=f\}} |v-u| \leq \alpha(\|Dv\| - \|Du\|) ,$$

for all $v \in \mathtt{BV}$.

Since any function $v \in \mathtt{BV}$ can be approximated by a sequence of functions $v_n \in C_0^\infty$ it can be approximated in such a way that

$$v_n \to v \text{ in } L^1 \text{ and } \|Dv_n\| \to \|Dv\| ,$$

we have proven the following lemma:

LEMMA 4.5. *Assume $f \in L^1$ and $u \in \mathtt{BV}$. Then $\alpha \in \mathcal{G}_u(\partial|u-f|)$ if and only if for every $v \in C_0^\infty$*

$$-\int_{\{u\neq f\}} sgn(u-f)(v-u) - \int_{\{u=f\}} |v-u| \leq \alpha(\|Dv\| - \|Du\|) . \tag{4.8}$$

5. Properties of Minimizers

In the following we prove a similar result to (3.1).

THEOREM 5.1. *Assume that $f \in L^1$ and $\alpha > 0$. Then $u = u_\alpha$ is a minimizer of $\mathcal{F}$ if and only if $u \in \mathtt{BV}$ and $\alpha \in \mathcal{G}_u(\partial|u-f|)$.*

Proof. Since u_α minimizes $\mathcal{F}$ it follows that for all $h \in \mathtt{BV}$ and $\varepsilon > 0$ that

$$\begin{aligned} &\int |u_\alpha - f| + \alpha\|Du_\alpha\| \\ &\leq \int |u_\alpha + \varepsilon h - f| + \alpha\|D(u_\alpha + \varepsilon h)\| \\ &\leq \int |u_\alpha - f| + \varepsilon \int_{\{u_\alpha \neq f\}} \mathrm{sgn}(u_\alpha - f)h + \eta(\varepsilon, u_\alpha, h) + \alpha\|D(u_\alpha + \varepsilon h)\| . \end{aligned} \tag{5.1}$$

This shows that for every $h \in \mathtt{BV}$

$$
\begin{aligned}
&-\int_{\{u_\alpha \neq f\}} \operatorname{sgn}(u_\alpha - f) h - \int_{\{u_\alpha = f\}} |h| \\
\leq &\alpha \frac{\|D(u_\alpha + \varepsilon h)\| - \|Du_\alpha\|}{\varepsilon} + \left(\eta(\varepsilon, u_\alpha, h) - \int_{\{u_\alpha = f\}} |h| \right).
\end{aligned}
$$

Since $\|Du\|$ is convex the one dimensional function

$$
g(\varepsilon) := \|D(u_\alpha + \varepsilon h)\|
$$

is convex in ε (and thus by Rademacher's theorem differentiable almost everywhere), and thus

$$
\begin{aligned}
\liminf_{\varepsilon \to 0+} \frac{g(\varepsilon) - g(0)}{\varepsilon} &= \liminf_{\varepsilon \to 0+} \frac{\|D(u_\alpha + \varepsilon h)\| - \|Du_\alpha\|}{\varepsilon} \\
&\leq \|D(u_\alpha + h)\| - \|Du_\alpha\| \\
&= g(1) - g(0).
\end{aligned}
$$

The argument can be illustrated with the following drawing cf. Figure 5.1. Since $\eta(\varepsilon, u_\alpha, h) \to \int_{\{u_\alpha = f\}} |h|$ for $\varepsilon \to 0$, we find that

$$
-\int_{\{u_\alpha \neq f\}} \operatorname{sgn}(u_\alpha - f) h - \int_{\{u_\alpha = f\}} |h| \leq \alpha(\|D(u_\alpha + h)\| - \|Du_\alpha\|),
$$

or in other words $\alpha \in \mathcal{G}_{u_\alpha}(\partial |u_\alpha - f|)$.

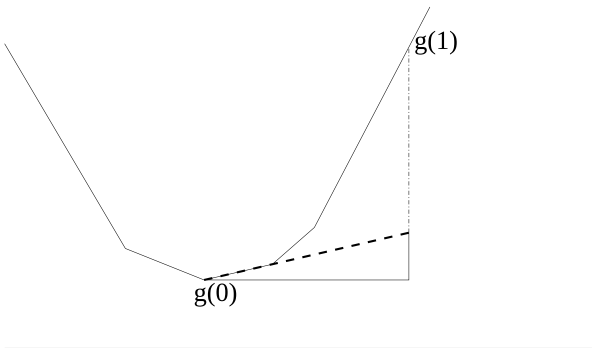

FIG. 5.1. *The directional derivative is below the line connecting $g(1)$ and $g(0)$ in the graph.*

To prove the converse direction we note that from $\alpha \in \mathcal{G}_u(\partial(u - f))$ and the convexity of $\int |u - f|$ it follows that

$$
\begin{aligned}
&\int |u + h - f| + \alpha \|D(u + h)\| \\
\geq &\int |u - f| + \alpha \|Du\| + \int_{\{u \neq f\}} \operatorname{sgn}(u - f) h + \int_{\{u = f\}} |h| \\
&\quad + \alpha(\|D(u + h)\| - \|Du\|) \\
\geq &\int |u - f| + \alpha \|Du\|.
\end{aligned}
$$

Thus u is a global minimizer. □

The following consequences can be derived from Theorem 5.1:

REMARK 5.2.

- *From* $\alpha \in \mathcal{G}_u(\partial|u-f|)$ *it follows by taking in* (4.8) *(with* $v=0$ *and* $v=2u$*) that*

$$\int_{\{u\neq f\}} sgn(u-f)u - \int_{\{u=f\}} |f| \leq -\alpha\|Du\|$$

and

$$-\int_{\{u\neq f\}} sgn(u-f)u - \int_{\{u=f\}} |f| \leq \alpha\|Du\|\,,$$

which shows that

$$\|Du\| \in \left\{ -\int_{\{u\neq f\}} \psi u : \psi \in \partial|u-f| \right\}.$$

- $\alpha \in \mathcal{G}_0(\partial|f|)$ *if and only if for all* $h \in \mathtt{BV}$

$$-\int_{\{f\neq 0\}} sgn(f)h - \int_{\{f=0\}} |h| \leq \alpha\|Dh\|\,.$$

Therefore $\alpha \geq G(\partial|f|)$. *In this case*

$$\inf\{\alpha : \alpha \in \mathcal{G}_0(\partial|f|)\} = G(\partial|f|)\,.$$

- *Together with* (4.7) *it follows that*

$$G(\partial(u_\alpha - f)) \leq \alpha\,.$$

In particular, if $f \in \mathtt{BV}$ *and we take* $h = u_\alpha - f$, *it follows then that*

$$\int |u_\alpha - f| \leq \alpha\|D(u_\alpha - f)\|\,. \tag{5.2}$$

- *Moreover, from Theorem 5.1 it follows that* $u_\alpha = f \in \mathtt{BV}$ *if and only if* $\alpha \in \mathcal{G}_f(\partial|0|)$. *Moreover,* $\alpha \in \mathcal{G}_f(\partial|0|)$ *is equivalent to*

$$-\int |h| \leq \alpha(\|D(f+h)\| - \|Df\|) \text{ for all } h \in \mathtt{BV}\,. \tag{5.3}$$

This in turn is equivalent to $\mathcal{I}_{\tilde{\phi}}(f) = \left|\partial\tilde{\phi}\right|(f) \leq \frac{1}{\alpha}$ *(cf.* (3.3)*).*
From (5.3) *it follows that for all* $c \in \mathbb{R}\backslash\{0\}$

$$-\int |ch| \leq \alpha(\|D(cf+ch)\| - \|D(cf)\|) \text{ for all } h \in \mathtt{BV}\,,$$

or equivalently

$$-\int |h| \leq \alpha(\|D(cf+h)\| - \|D(cf)\|) \text{ for all } h \in \mathtt{BV}\,.$$

This shows that $\alpha \in \mathcal{G}_{cf}(\partial|0|)$.

The G value and the G set are different concepts: If $u_\alpha = f$, what happens if $\alpha \in \mathcal{G}_f(\partial|0|)$, then $G(\partial|0|) = 0$.

A similar result to Theorem 5.1 also applies to the β-quantile regularization:

THEOREM 5.3. *Assume that $f \in L^1$ and $\alpha > 0$. Then $u = u_\alpha$ is a minimizer of $\mathcal{F}^\beta$ if and only if $u \in \mathtt{BV}$ and $\alpha \in \mathcal{G}_u(\partial S^\beta(u))$.*

We note that

$$\mathcal{G}_u(\partial S^\beta(u)) := \{\alpha \in [0,\infty] : \alpha \text{ satisfies (5.5)}\}\,. \tag{5.4}$$

Here for every $h \in \mathtt{BV}$

$$-\beta \int_{\{u>f\}} h - (\beta - 1)\int_{\{u<f\}} h - \int_{\{u=f\}} \beta(h) h \le \alpha\left(\|D(u+h)\| - \|Du\|\right). \tag{5.5}$$

6. Relation to the Literature

Chan & Esedoglu [3] characterized minimizers of the functional (1.1) when $f = \chi_\Omega$ under the assumptions that

$$\|Df\| = \int f \nabla\cdot\vec{\phi} \text{ for some } \vec{\phi} \in C_0^1 \text{ satisfying } \left|\vec{\phi}(x)\right| \le 1 \text{ and } \left|\nabla\cdot\vec{\phi}(x)\right| \le C\,.$$

In this case we have for all $u \in L^1$

$$\frac{\|Df\| - \|Du\|}{\int |u - f|} \le \frac{\int (f-u)\nabla\cdot\vec{\phi}}{\int |u-f|} \le C\,.$$

That is $\left|\partial\tilde{\phi}\right|(f) \le C$, and consequently, if $C \le \frac{1}{\alpha}$, then $u_\alpha = f$.

In particular if $f = \chi_\Omega \in \mathtt{BV}$ we have $u_\alpha \equiv 0$ if and only if for every $h \in \mathtt{BV}$

$$\int_\Omega h - \int_{\mathbb{R}^n\setminus\Omega} |h| \le \alpha\|Dh\|\,.$$

Taking $h = \chi_\Omega$, we find that

$$\frac{\text{meas}(\Omega)}{\text{Per}(\Omega)} \le \alpha\,.$$

Moreover, we have $u_\alpha \equiv f$ if and only if for every $h \in \mathtt{BV}$

$$\alpha\left(\text{Per}(\Omega) - \|D(\chi_\Omega + h)\|\right) \le \int |h|\,. \tag{6.1}$$

Note, that for any function $h \in W_0^{1,1}$ we have $\|Dh\|$ is the norm of the absolute continuous part and $\|D(\chi_\Omega)\|$ is the singular part of the measure $\|D(\chi_\Omega + h)\|$ and therefore

$$\|D(\chi_\Omega + h)\| = \|D(\chi_\Omega)\| + \|Dh\|\,.$$

Therefore, the left hand side of (6.1) is negative and thus (6.1) is satisfied.

If $u_\alpha = f$, then (6.1) provides a restriction on α if $\|Dh\|$ is an appropriate singular measure. Take $h = -c\chi_\Omega$ with $c \in [0,1]$, then from (6.1) it follows that

$$\alpha \le \frac{\text{meas}(\Omega)}{\text{Per}(\Omega)}\,.$$

The technique of the proof of Theorem 5.1 is not limited to the $L^1 - \mathtt{BV}$ regularization technique. Analogous characterization can be proven for the Rudin-Osher-Fatemi model:

THEOREM 6.1. *Assume $f \in L^2$ and $\alpha > 0$. Then $u = u_\alpha$ is a minimizer of the functional*

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

if and only if $u \in L^2$ with finite total variation and for every $h \in L^2$ with finite total variation

$$-\int (u-f)h \leq \alpha\left(\|D(u+h)\| - \|Du\|\right). \tag{6.2}$$

More general, if $\phi(\cdot)$ is a convex, differentiable, coercive (i.e. it satisfies $\phi(\cdot) \geq |\cdot|^p$, $p > 1$) function, then $u = u_\alpha$ is a minimizer of the functional

$$\int \phi(u) + \alpha \|Du\|$$

if and only if $u \in L^p$ with finite total variation and for every $h \in L^q$ with finite total variation

$$-\int \phi'(u)h \leq \alpha\left(\|D(u+h)\| - \|Du\|\right).$$

From (6.2) it follows

1. by taking $h = u$ and $h = -u$ that
$$-\int (u-f)u = \alpha \|Du\|. \tag{6.3}$$
2. From the triangle inequality it follows that for every $h \in L^2$ with finite total variation
$$\left|\int (u-f)h\right| \leq \alpha \|Dh\| \tag{6.4}$$
which in particular guarantees that the G-norm of $u-f$ is less or equal α and together with (6.3) it follows that the G-norm of $u-f$ is α.

If the two items hold, then by taking $h = u + \tilde{h}$ in (6.4) (actually a $W_0^{1,1}$ approximation of h has to be used, to make the the statement rigorous) it follows that

$$\alpha\|Du\| - \int (u-f)\tilde{h} = -\int (u-f)(u+\tilde{h}) \leq \alpha \left\| D(u+\tilde{h}) \right\|. \tag{6.5}$$

This shows that (6.2) holds. Item I and II is the characterization from the book of Meyer [5]. From a continuously differentiable function ϕ the corresponding result can be found in [9].

7. Metrical regularization

A minimizer $u_\alpha = f$ can be guaranteed to be a minimizer of functionals of the form

$$d(u,f) + \alpha\psi(u),$$

where $d(\cdot,\cdot)$ is a metric on a Banach space $\mathcal{B}$ and $\psi(\cdot):\mathcal{B}\to(-\infty,\infty]$ is a convex, lower semi continuous functional. From Proposition 1.4.4 in [2] we know that for $f\in\mathcal{B}$

$$|\partial\psi|(f) = \mathcal{T}_\psi(f) := \sup \frac{(\psi(f)-\psi(u))^+}{d(f,u)}.$$

This shows that

Corollary 7.1. *$u_\alpha = f$ if and only if $|\partial\psi|(f) \le \frac{1}{\alpha}$.*

We have considered already the metric on L^1 and the convex functional $\tilde{\phi}(u) = \|Du\|$, which results in the functional $\mathcal{F}$. Another example of a metric is $d(f,g) = \sqrt{\int |f-g|^2}$. The functional

$$\tilde{\tilde{\phi}} : L^2 \to [0,\infty], \quad u \to \|Du\|$$

is convex and lower semi-continuous. Application of the Corollary 7.1 shows that $u_\alpha = f$ if and only if $\left|\partial\tilde{\tilde{\phi}}\right| \le \frac{1}{\alpha}$. Note that in this case u_α satisfies the Euler equation

$$\frac{u-f}{\sqrt{\int (u-f)^2}} \in \alpha\nabla\cdot\left(\frac{\nabla u}{|\nabla u|}\right).$$

This is variant of the Rudin-Osher-Fatemi functional where the minimizer satisfies similar analytical properties as the minimizers of the functional $\mathcal{F}$. Note however, that the functional is strictly convex and thus the minimizer is unique. For the numerical solution a non-local PDE has to be solved.

Acknowledgment. The work of O. S. has been supported by the Austria Science Foundation (FWF) Projects Y-123INF, FSP 9203-N12 and FSP 9207-N12. The work of W. Yin has been supported by NSF Grant DMS-01-04282, ONR Grant N00014-03-1-0514, and DOE Grant GE-FG01-92ER-25126. The work of S. Osher has been supported by NSF Grants DNS-03-12222 and ACI-03-21917 and ONR Grant N00014-03-0071.

REFERENCES

[1] S. Alliney, *Digital filters as absolute norm regularizers*, IEEE Transactions on Signal Processing, 40, 6, 1548-1562, 1992.

[2] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures* , Birkhäuser, 2005.

[3] T. F. Chan and S. Esedoglu, *Aspects of total variation regularized L^1 function approximation*, CAM Report, 1-29, 2004.

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

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

[6] M. Nikolova, *Minimizers of cost-functions involving nonsmooth data-fidelity terms, application to the processing of outliers*, SIAM J. Numer. Anal., electronic, 40, 965-994, 2002.
[7] M. Nikolova, *A variational approach to remove outliers and impulse noise*, J. Math. Imaging Vision, 20, 99-120, 2004. Special Issue on Mathematics and Image Analysis.
[8] M. Nikolova, *Weakly constrained minimization: application to the estimation of images and signals involving constant regions*, J. Math. Imaging Vision, 21, 155-175, 2004.
[9] S. Osher and O. Scherzer, *G-norm properties of bounded variation regularization*, Comm. Math. Sci, 2, 237-254, 2004.
[10] L. I. Rudin, S. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D, 60, 259-268, 1992.