ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 40, Number 2, 2010

ON THE KANTOROVICH THEOREM AND THE REGULARIZATION OF TOTAL VARIATION DENOISING PROBLEMS

L.A. MELARA AND A.J. KEARSLEY

ABSTRACT. Total variation methods are an optimization-based approach for solving image restoration problems. The mathematical formulation results in an equality constrained optimization problem, a solution which can be obtained using Newton's method. This note is motivated by the numerical results of an augmented Lagrangian homotopy method for the regularization of total variation problems. The numerical technique uses the regularization parameter as a homotopy parameter which is reduced. As a result, a sequence of equality constrained optimization problems is solved using Newton's method. In this report, the convergence of an augmented Lagrangian homotopy method for total variation minimization is addressed. We present a relationship between the homotopy parameter and the radius of the Kantorovich ball.

1. Introduction. An image comes from a continuous setting, [**2**]. Various devices can be used to capture the image, for example, a digital camera. To construct a digital image, a mesh is superimposed on the photograph and each box is assigned a number representing the average intensity of the brightness in the box. Each box is called a pixel. In image processing, restoring an image is a fundamental task. Roughly speaking, restoration methods can be separated into three categories: statistical methods, transform-based methods and optimization-based methods, [**11**]. In this report, we focus on optimization-based methods, in particular, total variation methods, first introduced in [**15, 16**].

We begin with a brief mathematical formulation of the equality constrained optimization problem. Let $\mathbf{x} = (x, y) \in \Omega$, where Ω is a convex, polygonal bounded region of $\mathbf{R}^2$, and consider $u : \Omega \to \mathbf{R}$ where values of the function $u = u(\mathbf{x})$ represent the intensities of a given image. A common assumption is that the corruption of an image results from the following operations

$$u_0 = Ku^* + \eta,$$

Received by the editors on April 12, 2007, and in revised form on December 21, 2007.

DOI:10.1216/RMJ-2010-40-2-595

where the action of the linear operator K on the true image u^* (a blur, for instance) and the perturbation by random noise η result in the observed image u_0. The operator $K : L^p(\Omega) \to L^2(\Omega)$ is linear and continuous, [**3**–**5, 7, 17**].

In total variation methods, the true image u^* is a solution to the equality constrained optimization problem

$$\min_u \int_\Omega |\nabla u| \, d\mathbf{x} \text{ such that } 1/2 \left(\|Ku - u_0\|^2_{L^2(\Omega)} - \sigma^2 \right) = 0, \tag{1}$$

where

$$\int_\Omega |\nabla u| \, d\mathbf{x} = \int_\Omega \sqrt{(\partial u/\partial x)^2 + (\partial u/\partial y)^2} \, d\mathbf{x},$$

and the standard deviation σ is given by

$$\sigma = \left(\int_\Omega |u^* - u_0|^2 \, d\mathbf{x} \right)^{1/2}, \tag{2}$$

and $\sigma > 0$. The objective function and the equality constraint contribute differently to the image restoration process. The total variation term is the objective function, and this term is important since it preserves the sharp edges of an image. The sharp edges are due to drastic changes in color in an image, for example from black to white. The equality constraint is included to ensure that the solution to (1) does match the true image since σ contains information about the noise added to u^*. In [**1, 8, 18**], the authors addressed the case when σ is unknown.

The work in [**5**] established a link between the constrained optimization problem (1) and the unconstrained problem:

$$\text{Minimize/find a critical point of } \int_\Omega |\nabla u| \, d\mathbf{x} + \frac{\lambda}{2} \|Ku - u_0\|^2_{L^2(\Omega)},$$

where the Lagrange multiplier λ is assumed to be nonnegative. The authors also described a relaxation method for computing a solution and provided a convergence proof.

Solving the constrained optimization problem is difficult due to the evaluation of the unbounded differential operator R where $Ru = |\nabla u|$, [**3**–**5, 7, 17**]. Much work has been done in [**3**–**5, 7, 8, 17**] addressing the appropriate space of functions containing the solution u^*. This space is the bounded variation space $\mathrm{BV}(\Omega)$, see [**1, 3–5, 7, 8, 17, 18**] for more details.

In [**17**], the author proved the existence and convergence of a solution for the denoising-deblurring variation problem using lower semi-continuity results for convex functionals of measures. In her work, she presented numerical results obtained using a finite difference implementation. The report in [**3**] considered the reconstruction model proposed by Geman and Geman. In their work, the Fenchel-Legendre transform is used to reduce the energy model to a sequence of quadratic minimization problems. The authors of [**1**] considered the optimization problem

$$\min_{u\in L^2(\Omega)} \|Ku-u_0\|^2_{L^2(\Omega)} + \beta\int_\Omega \sqrt{|\nabla u|^2+\varepsilon^2}\,d\mathbf{x}, \tag{3}$$

where the term $\int_\Omega \sqrt{|\nabla u|^2+\varepsilon^2}\,d\mathbf{x}$ is a regularization term and β is a penalty parameter. This approach is known in the inverse problems community as Tikhonov regularization, [**1**]. Their work focused on an analysis of bounded variation methods for the operators $Ku=u_0$; for example, they analyzed convexity, semi-continuity and compactness of the regularization term as well as the well-posedness of unconstrained minimization problems, [**1**]. The work assumes that $u^*\in\mathrm{BV}(\Omega)$. Vogel and Oman in [**18**] considered the problem

$$\min_{u\in L^2(\Omega)} \frac{1}{2}\|u-u_0\|^2_{L^2(\Omega)} + \beta\int_\Omega \sqrt{|\nabla u|^2+\varepsilon^2}\,d\mathbf{x}, \tag{4}$$

where $K=I$ and I is the identity. The authors presented a fixed point algorithm which is a variant of a cell-centered finite difference multi-grid method of Ewing and Shen, [**18**]. Whereas in [**7**], Dobson and Scherzer also addressed the choice of appropriate space for a solution u^* to (4) for an unbounded, densely defined linear 'differentiation' operator:

$$R: D(L)\subseteq H^1(\Omega)\subseteq L^2(\Omega)\to (L^2(\Omega))^d, \quad u\to\nabla u.$$

Since $Ru=|\nabla u|$ is unbounded for $u^*\notin\mathrm{BV}(\Omega)$, the approach in [**7**] 'stabilized' the differentiation operator by approximating Ru with

$$R(I+\gamma R^*R)^{-1}u,$$

where γ is the regularization parameter and R^* is the adjoint of R. This stabilization permits the existence of a solution u^* in the Hilbert space $L^2(\Omega)$, [**7**]. The term Ru can also be stabilized by approximating with finite difference quotients, finite element methods, etc., see [**7**]. We assume $K = I$. Although the solutions of the problems (5) and (6) live in the BV (Ω) space, we work with $H^1(\Omega)$ solutions. We use finite element methods to approximate the operator Ru, thus 'stabilizing' the evaluation of Ru, [**7**]. The computational technique presented in [**13**] is associated with this work and P_1 finite element methods were implemented to approximate Ru. In our formulation we seek an optimal solution $u^* \in H^1(\Omega)$ for the optimization problem (1), [**7, 13**].

In [**13**], the authors consider the regularized equality constrained optimization problem

$$\min_{u \in H^1(\Omega)} \int_\Omega \sqrt{|\nabla u|^2 + \varepsilon^2}\, d\mathbf{x} \qquad \text{such that} \quad 1/2 \left(\|u - u_0\|^2_{L^2(\Omega)} - \sigma^2 \right) = 0, \tag{5}$$

where $\varepsilon > 0$ is the regularization parameter. The Lagrangian functional corresponding to the regularized optimization problem (5) is

$$\ell_\varepsilon(u, \lambda) = \int_\Omega \sqrt{|\nabla u|^2 + \varepsilon^2}\, d\mathbf{x} + \lambda/2 \left(\|u - u_0\|^2_{L^2(\Omega)} - \sigma^2 \right), \tag{6}$$

where $\lambda \in \mathbf{R}$ is the Lagrange multiplier associated to the equality constraint. The directional derivative of (6) in direction $q \in H^1(\Omega)$ is

$$\ell'_\varepsilon(u, \lambda)\, q = \int_\Omega \frac{\nabla u \cdot \nabla q}{\sqrt{|\nabla u|^2 + \varepsilon^2}} + \lambda(u - u_0) q \, d\mathbf{x}. \tag{7}$$

The objective function in (1) is regularized in (5) to desingularize the gradient of the Lagrangian with respect to u. From (7) we obtain the L^2-representation of the gradient, given by:

$$\mathcal{M}_\varepsilon(u, \lambda) = -\nabla \cdot \left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon^2}} \right) + \lambda(u - u_0), \quad \mathbf{x} \in \Omega, \tag{8}$$

where

$$\frac{\partial u}{\partial n} = 0, \quad \mathbf{x} \in \Gamma, \text{ with } \Gamma = \partial\Omega,$$

is required to ensure a convenient computation of the gradient. Enforcing the homogeneous Neumann boundary condition is a commonly employed practice. The homogeneous Neumann boundary condition comes from the implicit boundary condition,

$$\frac{\nabla u \cdot n}{\sqrt{\varepsilon^2 + |\nabla u|^2}} = 0, \quad \mathbf{x} \in \Gamma.$$

Note the first term of $\mathcal{M}_\varepsilon(u, \lambda)$ in (8) is well-defined for a constant function u.

Newton's method is used to compute roots $u^* \in H^1(\Omega)$ of (8) for fixed λ, [**13**]. This requires the Hessian of the Lagrangian (6):

$$\begin{aligned}
&\frac{d}{dt} \langle \ell_\epsilon'(u + tq, \lambda),\, q\rangle \,|_{t=0} \\
&= \int_\Omega \left[-\nabla \cdot \left(\frac{\nabla q}{\sqrt{|\nabla u|^2 + \varepsilon^2}} \right) + \nabla \cdot \left(\frac{(\nabla u \cdot \nabla q)\nabla u}{\left(\sqrt{|\nabla u|^2 + \varepsilon^2}\right)^3} \right) + \lambda\, q \right] q \, d\mathbf{x}, \\
&= \langle Aq, q\rangle_{L^2(\Omega)}.
\end{aligned}$$

The use of the Kantorovich Theorem in conjunction with Newton's method for nonlinear operator equations has been studied by Tapia in [**9, 10, 12**]. Therefore, using Newton's method in [**13**] led us to the Kantorovich theorem. The numerical technique implemented in [**13**] indicates that, by constructing a homotopy on the regularization parameter ε, Newton's method generates more desirable solutions to (1). The Kantorovich theorem establishes criteria for an initial iterate of Newton's method so the iterative procedure attains convergence. We present a relationship between the regularization parameter ε and the radius of the Kantorovich ball. This paper seeks to shed light on the results obtained in [**13**] by addressing the role of the Kantorovich theorem in the homotopy technique. The paper is organized as follows. In the next section we show the Kantorovich constants depend on the homotopy parameter ε. The last section states our conclusions.

2. Convergence results. We denote the directional derivative of $\ell_\varepsilon(u, \lambda)$ in direction $q \in H^1(\Omega)$ by

$$L_\varepsilon(u) = \ell'_\varepsilon(u,\, \lambda)\, q,$$

where $L_\varepsilon : H^1(\Omega) \to \mathbf{R}$.

Similarly, the second directional derivative of $\ell_\varepsilon(u, \lambda)$ in the direction $q \in H^1(\Omega)$ is denoted by

$$J_\varepsilon(u) = \langle Aq, q\rangle_{L^2(\Omega)} = \langle q, Aq\rangle_{L^2(\Omega)},$$

where A is symmetric and $J_\varepsilon : H^1(\Omega) \to \mathbf{R}$. Both operators L_ε and J_ε are clearly bounded operators with operator norms,

$$|||L_\varepsilon||| = \sup_{q\neq \mathbf{0}} \frac{|L_\varepsilon(u)|}{\|q\|_{H^1(\Omega)}} \text{ and } |||J_\varepsilon||| = \sup_{q\neq \mathbf{0}} \frac{|J_\varepsilon(u)|}{\|q\|^2_{H^1(\Omega)}}.$$

The application of Newton's method to our discrete equations can be written, for $n = 1, 2, \dots,$

$$u^{(n+1)} = u^{(n)} - J_\varepsilon(u^{(n)})^{-1}\, L_\varepsilon(u^{(n)}),$$

where $J_\varepsilon(u)^{-1}$ is the inverse operator of $J_\varepsilon(u)$. We present the following propositions which establish properties of the Jacobian $J_\varepsilon(u)$. For the numerical implementation of this iterative procedure, see [**13**]. These properties prove the required assumptions for the Kantorovich theorem, presented at the end of this section.

Proposition 1. *Given $m, p \in H^1(\Omega)$, the following inequality holds*

$$|||J_\varepsilon(m) - J_\varepsilon(p)||| \le c_1(\varepsilon)||m - p||_{H^1(\Omega)}, \tag{9}$$

with

$$c_1(\varepsilon) = 6 \max_{\mathbf{x}\in\Omega} \left\{ \frac{1}{|\nabla m|^2 + \varepsilon^2}, \frac{1}{[(|\nabla p|^2 + \varepsilon^2)(|\nabla m|^2 + \varepsilon^2)]^{1/2}}, \frac{1}{|\nabla p|^2 + \varepsilon^2} \right\}. \tag{10}$$

Proof. Consider the difference of the bilinear forms for any $q \in H^1(\Omega)$,

$$J_\varepsilon(m) - J_\varepsilon(p) = \int_\Omega (\nabla q \cdot \nabla q) \left(\frac{1}{\sqrt{|\nabla m|^2 + \varepsilon^2}} - \frac{1}{\sqrt{|\nabla p|^2 + \varepsilon^2}} \right) d\mathbf{x}$$
$$+ \int_\Omega \left(\frac{(\nabla p \cdot \nabla q)(\nabla p \cdot \nabla q)}{(\sqrt{|\nabla p|^2 + \varepsilon^2})^3} \right) - \left(\frac{(\nabla m \cdot \nabla q)(\nabla m \cdot \nabla q)}{(\sqrt{|\nabla m|^2 + \varepsilon^2})^3} \right) d\mathbf{x}.$$

We take the absolute value of both sides and apply the Cauchy-Schwarz and triangle inequalities to obtain

$$\begin{aligned}
&|J_\varepsilon(m) - J_\varepsilon(p)| \\
&\le \int_\Omega |\nabla q|^2 \left| \left(\frac{1}{(|\nabla m|^2+\varepsilon^2)^{1/2}} - \frac{1}{(|\nabla p|^2+\varepsilon^2)^{1/2}} \right) \right| d\mathbf{x} \\
&\quad + \int_\Omega |\nabla q|^2 \left(\frac{|\nabla p|}{(|\nabla p|^2+\varepsilon^2)^{3/2}} + \frac{|\nabla m|}{(|\nabla m|^2+\varepsilon^2)^{3/2}} \right) |\nabla m - \nabla p| \, d\mathbf{x} \\
&\le \int_\Omega |\nabla q|^2 \left(\frac{|\nabla m - \nabla p|}{(|\nabla m|^2+\varepsilon^2)^{1/2}(|\nabla p|^2+\varepsilon^2)^{1/2}} \right) d\mathbf{x} \\
&\quad + \int_\Omega |\nabla q|^2 \left(\frac{|\nabla p|}{(|\nabla p|^2+\varepsilon^2)^{3/2}} + \frac{|\nabla m|}{(|\nabla m|^2+\varepsilon^2)^{3/2}} \right) |\nabla m - \nabla p| \, d\mathbf{x} \\
&\quad + \int_\Omega |\nabla q|^2 |\nabla m||\nabla p| \left| \frac{1}{(|\nabla p|+\varepsilon^2)^{3/2}} - \frac{1}{(|\nabla m|+\varepsilon^2)^{3/2}} \right| d\mathbf{x} \\
&\le \int_\Omega |\nabla q|^2 |\nabla m - \nabla p| \\
&\quad \times \left(\frac{1}{|\nabla p|^2+\varepsilon^2} + \frac{1}{(|\nabla p|^2+\varepsilon^2)^{1/2}(|\nabla m|^2+\varepsilon^2)^{1/2}} \right. \\
&\qquad \left. + \frac{1}{|\nabla m|^2+\varepsilon^2} \right) d\mathbf{x} \\
&\quad + \int_\Omega |\nabla q|^2 |\nabla m||\nabla p| \\
&\qquad \times \left| \frac{1}{(|\nabla p|+\varepsilon^2)^{1/2}} - \frac{1}{(|\nabla m|+\varepsilon^2)^{1/2}} \right| \\
&\qquad \times \left(\frac{1}{|\nabla p|^2+\varepsilon^2} \right. \\
&\qquad\quad \left. + \frac{1}{(|\nabla p|^2+\varepsilon^2)^{1/2}(|\nabla m|^2+\varepsilon^2)^{1/2}} + \frac{1}{|\nabla m|^2+\varepsilon^2} \right) d\mathbf{x} \\
&\le \int_\Omega |\nabla q|^2 |\nabla m - \nabla p| \times \\
&\quad 2\left(\frac{1}{|\nabla p|^2+\varepsilon^2} + \frac{1}{(|\nabla p|^2+\varepsilon^2)^{1/2}(|\nabla m|^2+\varepsilon^2)^{1/2}} + \frac{1}{|\nabla m|^2+\varepsilon^2} \right) d\mathbf{x} \\
&\le c_1(\varepsilon) \|\nabla q\|^2_{L^4(\Omega)} \|\nabla m - \nabla p\|_{L^2(\Omega)} \\
&\le c_1(\varepsilon) \|\nabla q\|^2_{L^2(\Omega)} \|\nabla m - \nabla p\|_{L^2(\Omega)}
\end{aligned}$$

$$\leq c_1(\varepsilon)\|q\|^2_{H^1(\Omega)}\|m-p\|_{H^1(\Omega)},$$

where $c_1(\varepsilon)$ is as in (10). This establishes inequality (9). □

Proposition 2. *Let $u \in H^1(\Omega)$. If the Jacobian $J_\varepsilon(u)$ is symmetric positive definite, then it has a bounded inverse with bound*

$$|||J_\varepsilon(u)^{-1}||| \leq c_2(\varepsilon) = (\min\{\overline{v}, \lambda\})^{-1}, \tag{11}$$

where

$$\overline{v} = \min_{\mathbf{x}\in\Omega}\left\{\frac{\varepsilon^2}{(\sqrt{|\nabla u|^2+\varepsilon^2})^3}\right\}.$$

Proof. For any $m \in H^1(\Omega)$, we have that

$$J_\varepsilon(u) = \int_\Omega \left(-\frac{(\nabla m \cdot \nabla u)^2}{(\sqrt{|\nabla u|^2+\varepsilon^2})^3} + \frac{(\nabla m \cdot \nabla m)}{\sqrt{|\nabla u|^2+\varepsilon^2}} + \lambda\, m^2 \right) d\mathbf{x}.$$

Taking absolute values and applying the Cauchy-Schwarz inequality, we have

$$\begin{aligned}|J_\varepsilon(u)| &\geq \left| \int_\Omega \frac{\varepsilon^2|\nabla m|^2}{(\sqrt{|\nabla u|^2+\varepsilon^2})^3} + \lambda m^2\, d\mathbf{x} \right| \\ &\geq \min_{\mathbf{x}\in\Omega}\left\{\frac{\varepsilon^2}{(\sqrt{|\nabla u|^2+\varepsilon^2})^3}\right\}\int_\Omega |\nabla m|^2\, d\mathbf{x} + \lambda \int_\Omega m^2\, d\mathbf{x}.\end{aligned}$$

Letting

$$\overline{v} = \min_{\mathbf{x}\in\Omega}\left\{\frac{\varepsilon^2}{(\sqrt{|\nabla u|^2+\varepsilon^2})^3}\right\},$$

we have

$$|J_\varepsilon(u)| \geq \|m\|^2_{H^1(\Omega)} \min\{\overline{v}, \lambda\}.$$

Therefore, we obtain that

$$|J_\varepsilon(u)^{-1}| \leq \|m\|^2_{H^1(\Omega)} (\min\{\overline{v}, \lambda\})^{-1},$$

for all $m \in H^1(\Omega)$. This establishes (11) with constant $c_2(\varepsilon)$. □

Let $u^{(0)}$ be an initial iterate for Newton's method. Denote the p-ball of radius r centered at the point $u^{(0)}$ by $B_p(u^{(0)}, r) = \{u : \|u - u^{(0)}\|_p < r\}$. Recall the following theorem, see for example Dennis and Schnabel in [**6**].

Theorem 3 (Kantorovich). *Consider a function, $L_\varepsilon : \mathbf{R}^n \to \mathbf{R}^n$ that is defined on a convex set $C \subseteq \mathbf{R}^n$. Let the Jacobian operator of L_ε be J_ε and further assume that J_ε is a Lipschitz function with Lipschitz constant α_L. Assume that $u^{(0)}$ is some starting point selected from C and that the following are all satisfied for some $u^{(0)} \in C$,*

1. $|||J_\varepsilon(u) - J_\varepsilon(v)||| \le \alpha_L \|u - v\|$, *for all* $u, v \in C$,
2. $|||J_\varepsilon(u^{(0)})^{-1}||| \le \alpha_1$,
3. $|||J_\varepsilon(u^{(0)})^{-1} L_\varepsilon(u^{(0)})||| \le \alpha_2$.

If

$$\delta = \alpha_L \alpha_1 \alpha_2 \le 1/2,$$

and if $B_p(u^{(0)}, r) \subseteq C$ with

$$r = \alpha_2(1 - \sqrt{1 - 2\delta})/\delta, \tag{12}$$

then the Newtonian sequence $\{u^{(n)}\}$ given by

$$u^{(n+1)} = u^{(n)} - J_\varepsilon(u^{(n)})^{-1} L_\varepsilon(u^{(n)}), \tag{13}$$

is well defined, remains in the ball $B_p(u^{(0)}, r)$, and converges to the unique solution of $L_\varepsilon(u^*) = 0$ inside $B_p(u^{(0)}, r)$.

The Kantorovich theorem establishes criteria on the initial iterate $u^{(0)}$ of Newton's method for the iterative method to attain convergence. The statement of the Kantorovich theorem does not assume the existence of $u^* \in H^1(\Omega)$ nor the nonsingularity of $J_\varepsilon(u^*)$, [**6**]. However, since Ru is approximated using finite element methods, we can assume that a solution $u^* \in H^1(\Omega)$ of (5) exists.

Therefore, condition 3 holds because $|||L_\varepsilon|||$ and $|||J_\varepsilon^{-1}|||$ are bounded. Then, the sequence $\{u^{(n)}\}$ generated by Newton's method in (13), is well defined, converges to u^* and obeys

$$\|u^{(n+1)} - u^*\|_{H^1(\Omega)} \le C(\alpha_L, \alpha_1)\|u^{(n)} - u^*\|^2_{H^1(\Omega)},$$

where $C(\alpha_L, \alpha_1) > 0$, α_L is the Lipschitz constant presented in Theorem 3 and $|||J_\varepsilon(u^*)^{-1}||| \le \alpha_1/2$, [**6**]. The latter inequality can be obtained via induction in the proof of local convergence for Newton's method, see [**6**]. From Propositions 1 and 2, we see that the constants α_L and α_1 are dependent upon ε. Thus, the radius r of the Kantorovich ball given in (12) depends upon the regularization parameter ε, hence we have $r = r(\varepsilon)$. We note that $\alpha_L = c_1(\varepsilon)$ and $\alpha_1 = c_2(\varepsilon)$. By Theorem 3, we have $\delta = \alpha_L\alpha_1\alpha_2 \le 1/2$; hence,

$$(14)\qquad 0 < r(\varepsilon) = \frac{\alpha_2(1-\sqrt{1-2\delta})}{\delta} \le \frac{1}{\alpha_L\alpha_1} = \frac{1}{c_1(\varepsilon)c_2(\varepsilon)}.$$

From (10) and (11) we have

$$c_1(\varepsilon) = 6\max_{\mathbf{x}\in\Omega}\left\{\frac{1}{|\nabla m|^2+\varepsilon^2}, \frac{1}{[(|\nabla p|^2+\varepsilon^2)(|\nabla m|^2+\varepsilon^2)]^{1/2}}, \frac{1}{|\nabla p|^2+\varepsilon^2}\right\},$$

and

$$c_2(\varepsilon) = (\min\{\overline{v}, \lambda\})^{-1},$$

respectively. The term $\overline{v}$ was given in Proposition 1. We take the limit as $\varepsilon \to 0$ in (14) and obtain

$$\lim_{\varepsilon\longrightarrow 0} r(\varepsilon) = 0,$$

since $[c_1(\varepsilon)]^{-1} > 0$ has a constant limit as $\varepsilon \to 0$. Furthermore, from condition 2, we assume that $|\nabla u^{(0)}| \ne 0$ since $u^{(0)}$ is the observed noisy image. Therefore, for $[c_2(\varepsilon)]^{-1} > 0$, we take the minimum between $\overline{v}$ and λ for each value of ε. For sufficiently small values of ε, the minimum between $\overline{v}$ and λ will be $\overline{v}$. So, as $\varepsilon \to 0 \Rightarrow [c_2(\varepsilon)]^{-1} \to 0$.

The Kantorovich theorem is presented in a discrete setting. The results presented in Propositions 1 and 2 do hold in a discrete setting as well. Let

$$\Omega = \bigcup_{k=1}^{N(Q)} Q_k,$$

where $N(Q)$ is the total number of partitions in Ω with the partitions having the form

$$Q_k = ((i-1)h, ih) \times ((j-1)h, jh),$$

for $k = 1, 2, \dots, N(Q)$ with appropriate values of i and j; and $h = 1/N$ is the mesh size for N partitions along both coordinate axes. Then, the approximation for the partial derivatives of ∇u is

$$\partial u/\partial x = (u(x_{i+1}, y_j) - u(x_{i-1}, y_j))/(2h) \text{ and}$$
$$\partial u/\partial y = (u(x_i, y_{j+1}) - u(x_i, y_{j-1}))/(2h),$$

where $u(x_i, y_j)$ are function values at the nodes (x_i, y_j) with $x_i = ih$ and $y_j = jh$. Using this approximation we define the following equivalence between the Hilbert space norm $\|\cdot\|_{L^2(Q_k)}$ of ∇u and the vector norm $\|\cdot\|_2$ of ∇u

$$\|\nabla u\|_{L^2(Q_k)} = h\|\nabla u\|_2.$$

Similarly, using a midpoint quadrature rule yields

$$\|u\|_{L^2(Q_k)} = h\|u\|_2.$$

Combining these approximations for functions u and v and the Cauchy-Schwarz inequality yields

$$\begin{aligned}\int_\Omega \nabla u \cdot \nabla v \, d\mathbf{x} &= \sum_{k=1}^{N(Q)} \int_{Q_k} \nabla u \cdot \nabla v \, d\mathbf{x} \\ &= h^2 \sum_{k=1}^{N(Q)} [\nabla u \cdot \nabla v]_{Q_k} \\ &\le h^2 \sum_{k=1}^{N(Q)} [\|\nabla u\|_2 \|\nabla v\|_2]_{Q_k}.\end{aligned}$$

These techniques can be used to modify the proofs of Propositions 1 and 2 for a discrete setting.

3. Conclusion. In [**13**], the augmented Lagrangian was constructed to numerically approximate a minimizer of the optimization problem (5). Throughout this presentation, we have focused on the Lagrangian associated to the optimization problem (5). The minimizer of the augmented Lagrangian is also the solution to (5) for the well-chosen penalty parameter, [**14**]. In [**14**] it is shown that the Hessian of the objective function is positive definite in the nullspace of the gradient of the equality constraint. This requirement is often referred to as the sufficiency condition. Therefore, the numerical solutions computed in [**13**] are also saddle points of the Lagrangian presented here. Taken together this suggests that the radius of the Kantorovich ball increases as ε increases.

REFERENCES

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

2. G. Aubert and P. Kornprobst, *Mathematical problems in image processing*, Springer, New York, 2000.

3. G. Aubert and L. Vese, *A variational method in image recovery*, SIAM J. Numer. Anal. **34** (1997), 1948–1979.

4. M. Bertalmio, V. Caselles, B. Rouge and A. Sole, *TV based image restoration with local constraints*, J. Scien. Computing **19** (2003), 95–122.

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

6. J.E. Dennis and R.B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

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

8. D.C. Dobson and C.R. Vogel, *Convergence of an iterative method for total variation denoising*, SIAM J. Numer. Anal. **34** (1997), 1779–1791.

9. A.A. Goldstein and R.A. Tapia, *Kantorovich estimates for some convex programs*, SIAM J. Control **11** (1973), 375–377.

10. W.B. Gragg and R.A. Tapia, , *Optimal error bounds for the Newton-Kantorovich theorem*, SIAM J. Numer. Anal. **11** (1974), 10–13.

11. K. Majava, *Optimization-based techniques for image restoration*, Ph.D. Thesis, Department of Mathematical Information Technology, University of Jyväskylä, 2001.

12. M.A. McCarthy and R.A. Tapia, *Computable a posteriori l_∞ error bounds for the approximate solution of two-point boundary value problems*, SIAM J. Numer. Anal. **12** (1975), 919–937.

13. L.A. Melara, A.J. Kearsley and R.A. Tapia, *A homotopy method in the regularization of total variation denoising problems*, J. Optim. Theory Appl. **133** (2007), **15–25**.

14. J. Nocedal and S.J. Wright, *Numerical optimization*, Springer Series in Operations Research, Springer-Verlag, New York, NY, 1999.

15. L. Rudin and S. Osher, *Total variation based image restoration with free local constraints*, Proceedings of the International Conference on Image Processing, Volume I, 31–35, IEEE, 1994.

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

17. L. Vese, *A Study in the BV space of a denoising-deblurring variational problem*, Appl. Math. Optim. 44 (2001), 131–161.

18. C.R. Vogel and M.E. Oman, *Iterative methods for total variation denoising*, SIAM J. Sci. Comput. **17** (1996), 227–238.

Department of Mathematics, Shippensburg University, Shippensburg, PA 17257
Email address: lamelara@ship.edu

National Institute of Standards and Technology, Gaithersburg, MD 20899-8910
Email address: ajk@nist.gov