Low curvature image simplifiers: global regularity of smooth solutions and Laplacian limiting schemes

Andrea L. Bertozzi* and John B. Greer
Department of Mathematics
Duke University
Durham, NC 27708

June 18, 2003

Abstract

We consider a class of fourth order nonlinear diffusion equations motivated by Tumblin and Turk's (SIGGRAPH 1999) 'low curvature image simplifiers' for image denoising and segmentation. The PDE for the image intensity u is of the form

$$u_t = -\nabla \cdot (g(\Delta u)\nabla \Delta u) + \lambda(f-u),$$

where $g(s) = k^2/\left(k^2+s^2\right)$ is a 'curvature' threshold, and λ denotes a fidelity matching parameter. We derive a priori bounds for Δu that allow us to prove global regularity of smooth solutions in one space dimension, and a geometric constraint for finite-time singularities from smooth initial data in two space dimensions. This is in sharp contrast to the second order Perona-Malik equation (an ill-posed problem), on which the original LCIS method is modelled. The estimates also allow us to design a finite difference scheme that satisfies discrete versions of the estimates, in particular a priori bounds on the smoothness estimator in both one and two space dimensions. We present computational results that show the effectiveness of such algorithms. Our results are connected to recent results for fourth order lubrication-type equations and the design of positivity preserving schemes for such equations. This connection also has relevance for other related fourth order imaging equations (You-Kaveh IEEE Trans. Im. Proc. 2000) and (Osher-Solé-Vese UCLA CAM Rep. 2002).

*address after July 2003: Department of Mathematics, UCLA, Box 951555 Los Angeles, CA 90095-1555

1 Introduction

Nonlinear PDEs are now quite commonly used in image processing for applications ranging from edge detection, denoising, and image inpainting, to texture decomposition. Second order PDEs for image denoising and boundary or edge sharpening date back to the seminal work of Mumford-Shah [24], Rudin-Osher-Fatemi [28], and Perona-Malik [26]. All of these methods have some common features; they are based on a nonlinear version of the heat equation

$$u_t = \nabla \cdot ((g(|\nabla u|)\nabla u) \tag{1}$$

in which the 'thresholding function' g is small in regions of sharp gradients. A number of mathematical issues arise with these equations and their use. For example, Perona-Malik suggest using a smooth g that decays fast enough for large ∇u so that significant diffusion only takes place in regions of small change in the image, i.e. away from edge boundaries. A typical choice might be

$$g(s) = \frac{1}{1 + \left(\frac{s}{k}\right)^2} \tag{2}$$

However, this and similar choices results in a PDE that is linearly ill-posed in regions of high gradients, and the ensuing dynamics results in a characteristic "staircase" instability.

In the past few years, a number of authors have proposed analogous fourth order PDEs for the same applications (i.e. edge detection, image denoising, etc.) with the hope that these methods might perform better than their second order analogues [8, 9, 21, 22, 33, 34, 35]. Indeed there are good reasons to consider fourth order PDEs. Firstly, fourth order linear diffusion damps oscillations at high frequencies (e.g. noise) much faster than second order diffusion. Secondly, there is the possibility of having schemes that include effects of curvature, i.e. the second derivatives of the image, in the dynamics, which opens up a richer set of functional behaviors. On the other hand, the theory of fourth order nonlinear PDEs is much less well-developed than their second order analogues. And such equations often do not possess a maximum principle or comparison principle which could introduce artificial singularities or undesirable behavior in their implementation.

Some examples of fourth order equations include the $L^2 - curvature$ gradient flow method of You and Kaveh [35],

$$u_t = -\Delta(g(\Delta u)\Delta u), \tag{3}$$

and the Perona-Malik analogue by Wei [34],

$$u_t = -\nabla \cdot (g(\nabla u)\nabla \Delta u). \tag{4}$$

Other higher order PDEs in imaging include the texture/cartoon decomposition by Osher, Solé and Vese [25], image inpainting [2, 4, 13], and other examples involving

image restoration and surface reconstruction [8, 9, 31, 21, 22]. This is by no means a comprehensive list and is meant to give some idea of the level activity involving higher order PDE-based methods in imaging.

A concern with all of the above mentioned PDE techniques is that there is little or no theory for these equations and thus little or no understanding of how to implement them effectively to produce the desired effects on images. Indeed images are made of pixels on a grid, so the PDE is meant as a continuum approximation of the true discrete algorithm. Different discretizations of the same PDE could result in different dynamics, as is well-known for problems like convection equations with discontinuities. Recent analysis of a scheme for the second order Perona-Malik equation (1-2) was performed by Esedoglu [12]. Our paper is the first to examine schemes and their connection to the analysis of fourth order PDEs for image processing.

This paper focuses on a class of methods, introduced by Tumblin and Turk, known as 'Low Curvature Image Simplifiers' or LCIS [33]. The general PDE proposed is of the form

$$u_t + \nabla \cdot (g(D_{ij}u)\nabla\Delta u) = 0 \tag{5}$$

where g is a function of the second derivatives of the image intensity function u. Tumblin and Turk consider $g(s)$ as in (2) with the argument of g the square root of the sum of the squares of the entries of the Hessian matrix. However, as recently pointed out by Tumblin [32] this is not rotationally invariant. A better choice, which we use here, is $g(\Delta u)$ i.e. use the Laplacian of u in the argument of g in (2). Thus, we propose to rewrite the LCIS PDE as

$$u_t + \nabla \cdot (g(\Delta u)\nabla\Delta u) = 0. \tag{6}$$

A class of equations including (4-6) was studied in [16] by the authors, who proved global existence of H^1 solutions when the argument of g, in the form of derivatives of the intensity u, is convolved with a standard mollifier kernel. However, as is well known for some second order equations, as in (1), such mollification can turn an ill-posed problem into a well-posed problem [7]. The resulting numerical methods with the mollification appears to smooth out but not remove undesirable artifacts of the method, such as the staircase instability of the Perona-Malik method.

In order to best use any nonlinear PDE proposed for imaging, we need to understand the behavior of the equation without such smoothing kernels. Basic issues that are of interest are well-posedness with smooth initial data (for the continuous problem) and the design of numerical schemes that have certain desired properties on coarse grids and with rough initial data, as is typically the case for image data.

In this paper we study equation (6) in one and two space dimensions. In section 2 we derive the key estimates for this paper, which result in pointwise estimates for Δu in one space dimension. In section 3 we prove well-posedness of (6) for smooth initial data, globally in time in one space dimension and locally in time in 2D, with continuation in 2D unless a specific pointwise blowup of Δu occurs. In section 4,

the key estimates of section 3 lead to the design of new finite difference methods for (6) that have a priori bounds on the discrete Laplacian or 'smoothness estimator' of the image intensity u in both one and two space dimensions. In Section 5, we show some examples that illustrate the performance of these schemes. We discover that in practice the LCIS method performs quite well when used in combination with a fidelity term. That is, we include a forcing function in (6) so that the PDE becomes

$$u_t + \nabla \cdot (g(\Delta u)\nabla \Delta u) = \lambda(f - u), \tag{7}$$

where f is a noisy signal and u is the desired smooth signal. We view the right hand side as a forcing term and solve the equation to steady state. This is in the same spirit as a number of PDE-based variational methods in which the fidelity term has the form of an L^2 functional in the energy. In this case the PDE is not derived in a variational setting so we think of the right hand side as a driving term as in reaction-diffusion equations.

A related method was proposed by Lysaker et. al. [21] for denoising. Their PDE takes the form

$$u_t + (\frac{u_{xx}}{|u_{xx}|})_{xx} + (\frac{u_{yy}}{|u_{yy}|})_{yy} = \lambda(f - u), \tag{8}$$

which is very similar to (7), however it is not rotationally invariant, which is a desired property as pointed out by Tumblin [32]. In one space dimension, the only difference between this method and (7) is that the signum function $y/|y|$ above is replaced by the smoothed function $arctan(y)$ in the LCIS method. We discuss the relationship between these PDEs and a new fourth order equation for texture-cartoon decomposition [25] at the end of this paper.

2 Key estimates

For simplicity, we consider equation (6) with $g = 1/(1+y^2)$. Other forms of g can be considered however the particular structure of the nonlinearity can affect the results. The generalization to (7) with fidelity is discussed at the end of this section. For simplicity of analytical notation, we consider solutions on a periodic unit domain, $\mathbb{T}^N$, $N = 1, 2$. We note that some image processing is performed with Neumann boundary conditions, and this can be mapped onto the periodic problem by reflection symmetry on a domain of size 2^N times the original domain.

We make a key change of variables used by the authors in [17] to prove existence of traveling wave solutions of (6) with an additional advective Burgers term. Note that $g(\Delta u)\nabla \Delta u = \nabla(\arctan(\Delta u))$, and write $\tan w = \Delta u$. The equation satisfied by w is

$$w_t + \cos^2 w \Delta^2 w = 0. \tag{9}$$

This is quite similar to the modified lubrication equation

$$h_t + h^n h_{xxxx} = 0 \tag{10}$$

in one space dimension considered by Bertozzi [5] as a model problem for finite time singularities in thin film equations. Solutions of (10) (on a periodic domain) with positive initial data remain positive whenever $n \geq 5/3$, however numerics show the possibility of finite time singularities for smaller n. A combination of numerical and asymptotic analysis in [5] suggests that the true critical exponent is $n = 3/2$. We note that in one dimension (9) is like (10) with h^n replaced by $\cos^2 h$, i.e. a quadratic degeneracy in $\cos h = 0$. Thus by analogy, we expect (9) to behave like (10) with $n = 2$, near $w = \pm\pi/2$. In particular, in one dimension $\cos^2 w$ is bounded away from zero and thus u_{xx} is bounded from infinity. We make this rigorous below and derive an estimate that also provides bounds in two space dimensions that are useful for designing numerical schemes. Note that the particular form of the nonlinearity $\tan w = \Delta u$ affects the degree of degeneracy in the transformed equation (9) and thus will determine whether finite time singularities are possible in solutions of the PDE. This is quite different from the Perona-Malik equation where the non-monotonicity of $(sg(s))$ is what determines whether the equation is ill-posed and thus if it has singularities.

2.1 A priori H^2 estimate

We note that $w = \arctan \Delta u$ so the equation (9) is only meaningful for $w \in [-\pi/2, \pi/2]$. Thus we have, by assumption, a bound on $w \in L^2[\mathbb{T}^N]$. That w has an a priori H^2 bound follows by integration by parts:

$$\frac{d}{dt}\frac{1}{2}\int (\Delta w)^2\, dx = -\int \cos^2(w)(\Delta^2 w)^2. \tag{11}$$

Combining this with the L^2 estimate implies an a priori bound on $w \in L^\infty([0,T];(H^2[\mathbb{T}^N]))$ in terms of the H^2 norm of the initial data:

$$\frac{1}{2}\int (\Delta w(x,T))^2\, dx + \int_0^T \int \cos^2(w)(\Delta^2 w)^2 dx dt = \int (\Delta w(x,0))^2\, dx. \tag{12}$$

2.2 A priori estimate on $\int w \tan w$ and Δu

The quantity $w \tan w$ is nonnegative on $[-\pi/2, \pi/2]$. Using the above, and again integrating by parts, we note that

$$\frac{d}{dt}\int w \tan w dx = -\int (\Delta w)^2 dx - \int \sin w \cos w\ \Delta^2 w dx.$$

Combining this with (11) gives

$$\frac{d}{dt}[\int w \tan w dx + \frac{1}{2}\int (\Delta w)^2] \leq \int \frac{\sin^2 w}{4} dx - \int (\Delta w)^2 dx$$

which implies

$$\int w(x,T)\tan w(x,T)dx + \frac{1}{2}\int (\Delta w(x,T))^2\,dx + \int_0^T \int (\Delta w(x,t))^2 dxdt \\ \leq \int w(x,0)\tan w(x,0)dx + \frac{1}{2}\int (\Delta w(x,0))^2\,dx + \frac{T}{4}. \quad (13)$$

In the last inequality, we have assumed that the measure of T^N is one, so that

$$\int \sin^2 w \; dx \;\leq\; 1.$$

Combining (12), the Sobolev Lemma, and (13) gives the following result in one space dimension.

Proposition 1. *(Pointwise bound on u_{xx} in 1D) Let $u(x,t)$ be a smooth solution of (6) on the periodic unit interval $\mathbb{T}^1$ over the time interval $[0,T]$, with initial data $u(x,0) = u_0$. Then there exists a constant C depending on T and the H^2 norm of $w_0 = \arctan u_{0xx}$ such that $|u_{xx}(x,t)| < C$ for all $x \in \mathbb{T}^N$ and $t \in [0,T]$.*

Proof. Using the w variables, we note that the a priori H^2 bound implies w is Lipschitz continuous. Let M denote the Lipschitz constant. We show that $|w|$ is bounded away from $\pi/2$. Consider any time $T_0 \in [0,T]$ and Let x_0 be a point in the domain where $|w(\cdot,T_0)| \equiv w_0$ has a maximum. Then by Lipschitz continuity,

$$|w(x)| \geq w_0 - M|x - x_0|.$$

For $w_0 > 1$ we have

$$w_0 \tan w_0 > C(\pi/2 - |w_0|)^{-1},$$

and thus

$$\int w\tan w \geq C\int_\Omega (\delta + M|x-x_0|)^{-1}dx = C_1|\ln\delta| + C_2,$$

where $\delta = \pi/2 - w_0$ and C_i depend only on M and the size of the domain

$$\Omega = \{x \in T^N : w(x) > 1\}.$$

In particular, the a priori bound on $\int w\tan w$ implies an a priori pointwise lower bound on δ and thus on $w(x)$ away from $\pm\pi/2$. Applying the transformation $\tan w = u_{xx}$ finishes the result.

Remarks: In fact the a priori H^2 bound implies $w_x \in C^{1/2}$ in one dimension, which can be used to prove a sharper pointwise bound on δ in terms of the modulus of continuity. In 2D, the Sobolev lemma gives a weaker result and this is not enough to prove a pointwise bound on δ given that the two-dimensional geometry requires integrating the singularity of $w\tan w$ in a disk.

2.3 A priori bounds with fidelity

If we repeat the above estimates including the fidelity term in (7) we see that the $\tan w = \Delta u$ transformation gives the equation

$$(\tan w)_t + \Delta^2 w = \lambda(\Delta f - \tan w). \tag{14}$$

It is an exercise to repeat the above estimates including the terms on the right hand side of (14). The result is an estimate of the form

$$\begin{aligned}\int w(x,T)\tan w(x,T)dx + \frac{1}{2}\int (\Delta w(x,T))^2\,dx \\ + \int_0^T \int [(\Delta w(x,t))^2 + \lambda w(x,t)\tan w(x,t) + \lambda \sin^2 w(x,t)]dxdt \\ \leq \int w(x,0)\tan w(x,0)dx + \frac{1}{2}\int (\Delta w(x,0))^2\,dx + CT, \end{aligned} \tag{15}$$

where the constant C depends on the H^2 norm of the data f and on λ.

Thus we can prove Proposition 1 for the problem with fidelity when the signal f is in H^2. In practice f will be noisy. However the above estimate for smooth f allows us to design a numerical scheme that performs well in practice with noisy signals. This is discussed in sections 4 and 5.

3 Regularity of smooth solutions

In this section we prove regularity of smooth solutions of (6) on $\mathbb{T}^1$ and $\mathbb{T}^2$. We prove the existence of smooth solutions for both cases on some finite time interval. In Section 3.6, we prove that this interval can be extended for all time for solutions on $\mathbb{T}^1$. The analysis is carried out for equation (9) without fidelity; including the fidelity term in (7) is straightforward when the signal f is smooth. In the following two sections 4 and 5, we discuss schemes for the problem with fidelity and show some computational examples with and without fidelity for noisy signals f. The first result we prove is the following general existence theorem. Since we are interested in classical solutions of a fourth order PDE in one and two space dimensions, we take the initial data to be in H^6.

Theorem 1 (Existence). *If $w_0 \in H^6$, then there exists a $T > 0$ such that*

$$\frac{\partial w}{\partial t} = -\cos^2 w \Delta\Delta w \tag{16}$$

has a classical solution $w \in C\left([0,T); C^4(\mathbb{T}^N)\right)$, for $N = 1$ and 2, satisfying

$$w(\cdot, 0) = w_0.$$

Our approach involves energy methods combined with mollifiers, in the same spirit as in Chapter 3 of [23] for the Navier-Stokes equations or in [30] for nonlinear second order parabolic equations. We first prove the existence of smooth solutions to a regularized version of (16):

$$\frac{\partial w^\epsilon}{\partial t} = -J_\epsilon \cos^2 w^\epsilon \Delta\Delta J_\epsilon w^\epsilon \qquad (17)$$
$$w^\epsilon(\cdot, 0) = w_0.$$

J_ϵ is a mollifier on $\mathbb{T}^N$. Then we show that one can pass to the limit in ϵ to obtain a classical solution of the PDE without mollification. In the next subsection we present some basic properties of mollifiers that we use later in this section.

3.1 Notation and properties of mollifiers

We shall use the following notation:

- $\mathbb{T}^N$ is the N-dimensional torus.
- $\| u \|_0$ is the L^2-norm of u on $\mathbb{T}^N$:

$$\| u \|_0 = \left(\int_{\mathbb{T}^N} |u|^2 \right)^{\frac{1}{2}}.$$

- For a multi-index $\alpha = (\alpha_1, \cdots, \alpha_N)$, $\alpha_i \in Z^+ \bigcup \{0\}$,

$$D^\alpha f(x) = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_N}}{\partial x_N^{\alpha_N}} f(x).$$

  We define $|\alpha| = \sum_{i=1}^N \alpha_i$.

- $\| u \|_m$ for $m > 0$ is the H^m-norm of u on $\mathbb{T}^N$:

$$\| u \|_m = \left(\sum_{|\alpha| \le m} \| D^\alpha u \|_0^2 \right)^{\frac{1}{2}}.$$

- $\| u \|_\infty$ is the L^∞-norm of u on $\mathbb{T}^N$:

$$\| u \|_\infty = \operatorname{ess} \sup_{x \in \mathbb{T}^N} |u(x)|.$$

- Given a Banach space X, with norm $\| \cdot \|_X$, $C([0,T]; X)$ is the space of continuous functions mapping $[0,T]$ into X. We give this space the norm

$$\| u \|_{C([0,T];X)} = \sup_{0 \le t \le T} \| u(t) \|_X .$$

Similarly, $C^1((0,T);X)$ is the space of functions in $C([0,T];X)$ with derivatives in $C((0,T);X)$.

- $L^\infty(0,T;X)$ is the space of functions such that $u(t) \in X$ for a.e. $t \in (0,T)$ with finite norm

$$\| u \|_{L^\infty(0,T;X)} = \operatorname{ess\,sup}_{t \in (0,T)} \| u(t) \|_X .$$

- $L^2(0,T;X)$ is the space of functions such that $u(t) \in X$ for a.e. $t \in (0,T)$ with finite norm

$$\| u \|_{L^2(0,T;X)} = \left(\int_0^T \| u(t) \|_X^2 \, dt \right)^{\frac{1}{2}} .$$

Given f defined on $\mathbb{T}^N$, its Fourier coefficients are

$$\hat{f}(k) = \int_{T^N} f(x) e^{-2\pi i k \cdot x} dx$$

for each $k \in Z^N$(see, for example, [29] or [11]). We define the mollification, $J_\epsilon f$, of functions $f \in L^p(\mathbb{T}^N)$, $1 \le p \le \infty$, by

$$J_\epsilon f(x) = \sum_{k \in Z^N} \hat{f}(k) e^{-\epsilon^2 |k|^2 + 2\pi i k \cdot x}. \tag{18}$$

J_ϵ is analogous to mollifiers defined through a convolution operator for functions on R^n [1, 14, 23] . In Lemma 1, we list those properties of mollifiers on $\mathbb{T}^N$ that will be used throughout the paper. These properties use the equivalence of the H^m-norm of f as defined above with the norm

$$\| f \|_m = \left(\sum_{k \in Z^N} |\hat{f}(k)|^2 (1 + |k|^2)^m \right)^{\frac{1}{2}},$$

which is a direct result of the Plancherel identity [11].

Lemma 1. *Let J_ϵ be the mollifier defined in* (18). *Then $J_\epsilon f$ is a C^∞ function with the following properties [15]:*

1. *For all $f \in C^l(\mathbb{T}^N)$, with $l \ge \frac{n}{2}$, $J_\epsilon f \to f$ uniformly and*

$$|J_\epsilon f|_\infty \le |f|_\infty . \tag{19}$$

2. *For all* $f, g \in L^2(\mathbb{T}^N)$,

$$\int_{\mathbb{T}^N} (J_\epsilon f)\, g dx = \int_{\mathbb{T}^N} f\, (J_\epsilon g)\, dx. \tag{20}$$

3. *Mollifiers commute with distribution derivatives,*

$$D^\alpha J_\epsilon f = J_\epsilon D^\alpha f \quad \forall |\alpha| \leq m,\ f \in H^m. \tag{21}$$

4. *For all* $f \in H^s(\mathbb{T}^N)$, $J_\epsilon f$ *converges to* f *in* $H^s(\mathbb{T}^N)$, *and the rate of convergence in the* H^{s-1}*-norm is linear in* ϵ*:*

$$\lim_{\epsilon \to 0} \| J_\epsilon f - f \|_s = 0, \tag{22}$$

$$\| J_\epsilon f - f \|_{s-1} \leq \epsilon \| f \|_s . \tag{23}$$

5. *For all* $f \in H^s(\mathbb{T}^N)$, $\gamma \in Z^+ \cup \{0\}$, $m \in \{0, 1, ..., \gamma\}$, *and* $\epsilon > 0$,

$$\| J_\epsilon f \|_{s+\gamma} \leq \frac{c_{s\gamma}}{\epsilon^\gamma} \| f \|_s, \tag{24}$$

$$|J_\epsilon D^\gamma f|_\infty \leq \frac{c_\gamma}{\epsilon^{\frac{N}{2}+\gamma-m}} \| f \|_m . \tag{25}$$

3.2 A priori bounds in a high norm

Proving existence of solutions to both (16) and (17) relies heavily on the a priori bounds given by the following lemma.

Lemma 2. *Let* w^ϵ *be a solution of* (17) *on* $[0, T) \times \mathbb{T}^N$ *for any positive* ϵ *and* T *satisfying* $\cos^2 w \geq \nu > 0$. *Let* m *be a nonnegative integer and* $t \in [0, T)$. *If* $w_0 \in H^{m+2}$ *then* w^ϵ *satisfies bounds of the form*

$$\| w^\epsilon(t) \|_{m+2} \leq C_1(N, T, \| w_0 \|_{m+2}) F_1(\nu) \tag{26}$$

and

$$\int_0^T \| J_\epsilon w^\epsilon(t) \|_{m+4}^2 \, dt \leq C_2(N, T, \| w_0 \|_{m+2}) F_2(\nu), \tag{27}$$

where F_1 *and* F_2 *are continuously differentiable functions of* ν *away from* $\nu = 0$, *and both* C_1 *and* C_2 *are independent of* ϵ. N *denotes the dimension, which we assume to be one or two.*

Proof.
Case m = 0. Taking the L^2-inner product of (17) with $\Delta\Delta w^\epsilon$ gives

$$\frac{d}{dt}\frac{1}{2} \parallel \Delta w^\epsilon \parallel_0^2 = -\int \cos^2(w^\epsilon)(J_\epsilon \Delta\Delta w^\epsilon)^2 \leq 0, \tag{28}$$

so

$$\parallel \Delta w^\epsilon \parallel_0 \leq \parallel w_0 \parallel_2, \tag{29}$$

and

$$2\int_0^T \int \cos^2(w)(J_\epsilon \Delta\Delta w^\epsilon)^2 = \parallel \Delta w^\epsilon(0) \parallel_0^2 - \parallel \Delta w^\epsilon(T) \parallel_0^2 \leq \parallel \Delta w_0^\epsilon \parallel_0^2 . \tag{30}$$

Note also that w is defined between $\pm\pi/2$ so that there is an implicit a priori L^2 bound. This combined with (29) gives

$$\parallel w(t) \parallel_2 \leq C \parallel w_0 \parallel_2 . \tag{31}$$

Since $\nu > 0$, we also have

$$\int_0^T \int (J_\epsilon \Delta\Delta w^\epsilon)^2 \leq \frac{2}{\nu} \parallel w_0 \parallel_2, \tag{32}$$

and therefore

$$\int_0^T \parallel J_\epsilon w^\epsilon \parallel_4^2 \leq \frac{C(T)}{\nu} \parallel w_0 \parallel_2 . \tag{33}$$

Case m > 0. Assume the lemma holds for $m-1$:

$$\parallel w^\epsilon(t) \parallel_{m+1} \leq C_1(N, T, \parallel w_0 \parallel_{m+1}) F_1(\nu)$$

and

$$\int_0^T \parallel J_\epsilon w^\epsilon(t) \parallel_{m+3}^2 dt \leq C_2(N, T, \parallel w_0 \parallel_{m+1}) F_2(\nu).$$

Pick some multi-index α with $|\alpha| \leq m$. The L^2-inner product of the D^α-derivative of equation (17) with $D^\alpha \Delta\Delta w^\epsilon$ gives

$$\begin{aligned}
\frac{d}{dt}\frac{1}{2} \parallel D^\alpha \Delta w^\epsilon(t) \parallel_0^2 &= -\left(D^\alpha J_\epsilon \cos^2 w^\epsilon \Delta\Delta J_\epsilon w^\epsilon, D^\alpha \Delta\Delta w^\epsilon\right) \\
&= -\left(\cos^2 w^\epsilon D^\alpha \Delta\Delta J_\epsilon w^\epsilon, D^\alpha \Delta\Delta J_\epsilon w^\epsilon\right) \\
&\quad - \sum_{\substack{\mu+\beta=\alpha \\ \beta\neq\alpha}} \left(D^\mu \left(\cos^2 w^\epsilon\right) D^\beta \Delta\Delta J_\epsilon w^\epsilon, D^\alpha \Delta\Delta J_\epsilon w^\epsilon\right) \\
&\leq -\nu \parallel D^\alpha J_\epsilon w^\epsilon \parallel_4^2 + \sum_{\substack{\mu+\beta=\alpha \\ \beta\neq\alpha}} B_{\mu,\beta}.
\end{aligned}$$

As $\cos u$ and all of its derivatives are smooth and bounded, we use the Sobolev Lemma to derive (for $N = 1, 2$)

$$\begin{aligned} \| D^\mu \cos^2 w^\epsilon \|_\infty &\leq C(|\mu|) \left(\| w^\epsilon \|^{|\mu|}_{1+\frac{N}{2}+1} + \| w^\epsilon \|^{|\mu|-1}_{2+\frac{N}{2}+1} + ... + \| w^\epsilon \|_{|\mu|+\frac{N}{2}+1} \right) \\ &\leq C(|\mu|) \left(\| w^\epsilon \|^m_{m+\frac{N}{2}} + \| w^\epsilon \|_{m+\frac{N}{2}+1} \right) \\ &\leq C(m) \left(\| w^\epsilon \|^m_{m+1} + \| w^\epsilon \|_{m+2} \right), \end{aligned}$$

for each multi-index μ. Since $|\beta| \leq |\alpha| - 1$ in each term $B_{\mu,\beta}$,

$$\| D^\beta \Delta \Delta J_\epsilon u^\epsilon \|_0 \leq \| D^\alpha J_\epsilon w^\epsilon \|_3$$

for each β, giving

$$\begin{aligned} \frac{d}{dt} \frac{1}{2} \| D^\alpha \Delta w^\epsilon(t) \|^2_0 &\leq -\nu \| D^\alpha J_\epsilon w^\epsilon \|^2_4 \\ &\quad + C \left(\sum_{0 < |\mu| \leq |\alpha|} \| D^\mu \cos^2 w^\epsilon \|_\infty \right) \| D^\alpha J_\epsilon w^\epsilon \|_3 \| D^\alpha J_\epsilon w^\epsilon \|_4 \\ &\leq -\nu \| D^\alpha J_\epsilon w^\epsilon \|^2_4 \\ &\quad + C(m) \left(\| w^\epsilon \|^m_{m+1} + \| w^\epsilon \|_{m+2} \right) \| D^\alpha J_\epsilon w^\epsilon \|_3 \| D^\alpha J_\epsilon w^\epsilon \|_4 . \end{aligned}$$

Summing over all multi-indices α with $|\alpha| \leq m$, we have

$$\frac{d}{dt} \| \Delta w^\epsilon(t) \|^2_m \leq -2\nu \| J_\epsilon w^\epsilon \|^2_{m+4} + C(m) \left(\| w^\epsilon \|^m_{m+1} + \| w^\epsilon \|_{m+2} \right) \| J_\epsilon w^\epsilon \|_{m+3} \| J_\epsilon w^\epsilon \|_{m+4} .$$

We now combine this with the inductive hypothesis (26) to derive

$$\begin{aligned} \frac{d}{dt} \| \Delta w^\epsilon(t) \|^2_m &\leq -2\nu \| J_\epsilon w^\epsilon \|^2_{m+4} \\ &+ C \left(\| w_0 \|_{m+1} \right) (1 + \| w^\epsilon \|_{m+2}) \| J_\epsilon w^\epsilon \|_{m+3} \| J_\epsilon w^\epsilon \|_{m+4} . \end{aligned} \tag{34}$$

Using the inequality

$$ab \leq \nu a^2 + \left(\frac{1}{4\nu} \right) b^2 \tag{35}$$

with $a = \| J_\epsilon w^\epsilon \|_{m+4}$, we see

$$\frac{d}{dt} \| \Delta w^\epsilon(t) \|^2_m \leq -\nu \| J_\epsilon w^\epsilon \|^2_{m+4} + \frac{C \left(\| w_0 \|_{m+1}, T \right)}{\nu} \left(1 + \| w^\epsilon \|^2_{m+2} \right) \| J_\epsilon w^\epsilon \|^2_{m+3},$$

and therefore

$$\frac{d}{dt} \| \Delta w^\epsilon(t) \|^2_m \leq \frac{C \left(\| w_0 \|_{m+1}, T \right)}{\nu} \left(1 + \| w^\epsilon \|^2_{m+2} \right) \| J_\epsilon w^\epsilon \|^2_{m+3} . \tag{36}$$

Integrating in time, and using the inductive hypothesis (27), we have

$$\| w^\epsilon(t) \|_{m+2} \leq \ (1+ \| w_0 \|_{m+2}) \, e^{\frac{C(\|w_0\|_{m+1},T)}{\nu} \int_0^T \|J_\epsilon w^\epsilon(s)\|^2_{m+3} ds} - 1 \tag{37}$$

for all $0 \leq t \leq T$. Returning to inequality (34) and using

$$a(b-a) = \frac{1}{2}b^2 - \frac{1}{2}a^2 - \frac{1}{2}(a-b)^2 \leq \frac{1}{2}b^2 - \frac{1}{2}a^2,$$

gives

$$\frac{d}{dt} \| \Delta w^\epsilon(t) \|^2_m + \nu \| J_\epsilon w^\epsilon \|^2_{m+4} \leq \ C(\| w_0 \|_{m+1}, T)(1+ \| w^\epsilon \|_{m+2}) \| J_\epsilon w^\epsilon \|^2_{m+3} \, .$$

Integrating in time and using the hypothesis (27) again, results in

$$\int_0^T \| J_\epsilon w^\epsilon \|^2_{m+4} \leq \ C(\| w_0 \|_{m+2}, T) F(\nu) + C' \| w_0 \|^2_{m+2} \, . \tag{38}$$

By induction, bounds of the form (26) and (27) hold for all $m \geq 0$. □

3.3 Existence and uniqueness of solutions to the approximating equations

Theorem 2. *Let $w_0 \in H^m$ for some integer $m \geq 6$. There is a time $T > 0$, independent of ϵ, such that there exists a solution $w^\epsilon \in C^1((0,T), H^m)$ to the regularized equation (17) for every $\epsilon > 0$. Furthermore, there is some positive ν that is independent of ϵ and satisfies $\cos^2 w^\epsilon(t) \geq \nu > 0$ for each ϵ and all $t \in [0,T)$.*

Proof. We use the Picard Theorem on a Banach space, and the continuation property of autonomous ODEs on a Banach space. We state the important theorems here. See [19] for more information.

Theorem (Picard Theorem on a Banach Space). *Let $O \subseteq B$ be an open subset of a Banach space B, and let $F : O \to B$ be a locally Lipschitz continuous mapping. Then for any $X_0 \in O$, there exists a time T such that the ODE*

$$\frac{dX}{dt} = F(X), \quad X|_{t=0} = X_0 \in O, \tag{39}$$

has a unique (local) solution $X \in C^1((-T,T); O)$.

Theorem (Continuation on a Banach Space.). *Let $O \subset B$ be an open subset of a Banach space B, and let $F : O \to B$ be a locally Lipschitz continuous operator. Then the unique solution $X \in C^1([0,T); O)$ to the autonomous ODE*

$$\frac{dX}{dt} = F(X), \quad X|_{t=0} = X_0 \in O, \tag{40}$$

either exists globally in time, or $T < \infty$ and $X(t)$ leaves the open set O as $t \to T$.

Let $M > 0$ be given. We define for $m \geq 6$,

$$O = \left\{ w \in H^m | \parallel w \parallel_m > M, \text{ and } \cos^2(w(x)) > 0 \right\}. \tag{41}$$

Since $C(\mathbb{T}^N)$ is compactly embedded in $H^6(\mathbb{T}^N)$ for $N = 1$ and 2, O is an open set in $H^6(\mathbb{T}^N)$. The Sobolev Lemma also insures that

$$\parallel \cos^2(w_1) - \cos^2(w_2) \parallel_\infty \leq C \parallel w_1 - w_2 \parallel_6$$

for $w_1, w_2 \in O$. Using this estimate we can show that $F^\epsilon(w^\epsilon) = -J_\epsilon \cos^2 w^\epsilon \Delta\Delta J_\epsilon w^\epsilon$ is Lipschitz continuous on H^m. This result is an exercise which uses the properties of mollifiers from section 3.1. Note in particular property 5 which tells us that J_ϵ composed with differential operators is bounded on H^s and bounded from H^m to L^∞. We leave the details of this calculation to the reader; similar arguments can be found in [16] for related PDEs and in [23] for the Navier-Stokes equations and in [30] for second order nonlinear parabolic equations. The existence of smooth solutions to (17) on some finite time interval $[0, T^\epsilon)$ easily follows from the Picard Theorem.

We now find a positive time T that is independent of ϵ, and such that given $w_0 \in H^6$, all of the solutions w^ϵ exist and have $\cos w^\epsilon$ bounded away from zero on the common time interval $[0, T)$. We now use the Sobolev Lemma and Lemma 2 to find $T > 0$ independent of ϵ so that $\cos^2 w^\epsilon \geq \nu$ on some time interval. Denote by σ the minumum value of $\cos^2 w^\epsilon$. Then by the equation we have

$$\begin{aligned} \frac{d\sigma}{dt} &\leq C(\sigma) \parallel J_\epsilon \cos^2 w^\epsilon \Delta\Delta J_\epsilon w^\epsilon \parallel_\infty \\ &\leq C'(\sigma) \parallel w^\epsilon \parallel_6 \\ &\leq C''(\sigma) (\parallel w_0 \parallel_6, T), \end{aligned}$$

where C, C', and C'' are smooth functions of σ away from $\sigma = 0$. This bound is completely independent of ϵ. Gronwall's lemma then ensures the existence of a $\nu > 0$ and a $T > 0$ so that $\cos^2 w^\epsilon > \nu$ for $t < T$.

3.4 Passing to the limit

Proof of Theorem 1. We first show that solutions w^ϵ of the regularized equations form a Cauchy sequence in L^2. We then use the uniform bound on $\parallel w^\epsilon \parallel_6$ and an interpolation lemma to show that w^ϵ forms a Cauchy sequence in H^s for all positive $s < 6$.

Lemma 3. *The family of regularized solutions w^ϵ forms a Cauchy sequence in $C\left([0, T], L^2\left(\mathbb{T}^N\right)\right)$. In other words,*

$$\sup_{0<t<T} \parallel w^\epsilon - w^{\epsilon'} \parallel_0 \leq C(\parallel w_0 \parallel_5, T, \nu) \max(\epsilon, \epsilon'), \tag{42}$$

for all $\epsilon, \epsilon' < 1$.

Proof Taking the L^2-inner-product of $w_t^\epsilon - w_t^{\epsilon'}$ with $w^\epsilon - w^{\epsilon'}$, we have

$$\begin{aligned}
\frac{d}{dt}\frac{1}{2} \parallel w^\epsilon - w^{\epsilon'} \parallel_0^2 &= -\left(J_\epsilon \cos^2 w^\epsilon J_\epsilon \Delta\Delta w^\epsilon - J_{\epsilon'} \cos^2 w^{\epsilon'} J_{\epsilon'} \Delta\Delta w^{\epsilon'}, w^\epsilon - w^{\epsilon'}\right) \\
&\leq \left(\parallel J_\epsilon \cos^2 w^\epsilon J_\epsilon \Delta\Delta w^\epsilon \parallel_0 + \parallel J_{\epsilon'} \cos^2 w^{\epsilon'} J_{\epsilon'} \Delta\Delta w^{\epsilon'} \parallel_0\right) \parallel w^\epsilon - w^{\epsilon'} \parallel_0 \\
&\leq C \max(\epsilon, \epsilon') \left(\parallel w^\epsilon \parallel_5 + \parallel w^{\epsilon'} \parallel_5\right) \parallel w^\epsilon - w^{\epsilon'} \parallel_0 \\
&\leq C \max(\epsilon, \epsilon') \left(\parallel w^\epsilon \parallel_5^2 + \parallel w^{\epsilon'} \parallel_5^2\right) + \parallel w^\epsilon - w^{\epsilon'} \parallel_0^2 \\
&\leq C(\parallel w_0 \parallel_5, T, \nu) \max(\epsilon, \epsilon') + \parallel w^\epsilon - w^{\epsilon'} \parallel_0^2 .
\end{aligned}$$

Using Grönwall's Lemma complete's the proof. □

Lemma 3 implies that w^ϵ converges strongly to some $w \in C\left([0,T), L^2\left(\mathbb{T}^N\right)\right)$; We will use the following lemma (stated and proved in [1]) with the uniform bound on $\parallel w^\epsilon \parallel_6$ given by Lemma 2 to prove convergence in higher Sobolev spaces.

Lemma 4. *Given $s > 0$, there exists a constant C_s so that for all $w \in H^s$ and $0 < s' < s$,*

$$\parallel w \parallel_{s'} \leq C_s \parallel w \parallel_0^{1-s'/s} \parallel w \parallel_s^{s'/s} .$$

Given $s \in (0,6)$, we have

$$\sup_{0<t<T} \parallel w^\epsilon - w \parallel_{s'} \leq C(\parallel w_0 \parallel_6, T, \nu)\epsilon^{1-s'/s}.$$

Picking $s = N/2 + 4$ shows that w^ϵ converges strongly in $C\left([0,T); C^4(\mathbb{T}^N)\right)$, while $\frac{dw^\epsilon}{dt}$ converges strongly in $C\left([0,T); C(\mathbb{T}^N)\right)$. Since $w^\epsilon \to w$, the distribution limit of w_t^ϵ must be w_t, and w is a classical solution of (16).
□

3.5 Uniqueness and regularity

To prove uniqueness of the smooth solution in Theorem 1, we transform back to the u variable. Suppose w_1 and w_2 solve equation (16) with the initial condition $w_1(0) = w_2(0) = w_0 \in H^6$. To prove that $w_1 = w_2$ we consider $u_i = \Delta^{-1}(\tan w_i - \tan w_0)$ where we take the unique inverse with zero mean. Note that since w_i each have the same initial condition and since the mean of $\tan w$ is conserved, showing $u_1 = u_2$ is sufficient to show that $w_1 = w_2$. We accomplish this by proving that $||u_1 - u_2||_0$ decays in time. Note that for the image processing problem, $\int_{\mathbb{T}^2} \tan w_0 = 0$, however we can prove uniqueness for the general problem without the mean zero condition.

By taking the L^2-inner-product of $\frac{d}{dt}(u_1 - u_2)$ with $u_1 - u_2$, we discover

$$\frac{d}{dt}\frac{1}{2} \parallel u_1 - u_2 \parallel^2 =$$
$$-\int (\arctan(\Delta u_1 - \tan w_0) - \arctan(\Delta u_2 - \tan w_0)) ((\Delta u_1 - \tan w_0) - (\Delta u_2 - \tan w_0)) \leq 0,$$

since arctan is monotone increasing. We have just proved

Theorem 3 (Uniqueness). *The smooth solution of Theorem 1 is unique.*

The second result of this section is the fact that the solution is C^∞ after time zero due to the uniform parabolicity of the equation. The theorem below is proved using a standard bootstrap argument.

Theorem 4 (Regularity). *The solution w of PDE is in $C^\infty\left((0,T) \times T^N\right)$.*

Returning to the sequence of regularized problems, we see from Lemma 2 that $J_\epsilon w^\epsilon$ is uniformly bounded in $L^2\left((0,T); H^8\left(\mathbb{T}^N\right)\right)$. It follows that the classical solution w of equation (16) is in $L^2\left((0,T); H^8\left(\mathbb{T}^N\right)\right)$, and therefore $w(t) \in H^8$ for almost every $t \in (0,T)$. Pick t_0 with $w(t_0) \in H^8$. Now consider the new set of regularized equations,

$$w_t^\epsilon = -J_\epsilon \cos^2 w^\epsilon J_\epsilon \Delta\Delta w^\epsilon$$
$$w^\epsilon(t_0) = w(t_0).$$

Proving existence follows a standard bootstrapping argument. Lemma 2 now gives a uniform bound on $\parallel w^\epsilon(t) \parallel_8$ for $t \in (t_0, T)$. Repeating the steps in the above existence proof, and now using Lemma 4 to interpolate between L^2 and H^8, we get strong convergence in $C\left([0,T]; C^6(\mathbb{T}^N)\right)$. Theorem 3 shows that this is the same solution given by Theorem 1. Repeating proves smoothness in space, while differentiating equation (16) in time gives smoothness of all time derivatives of w.

3.6 Global existence

Theorem 5. *Consider PDE* (16) *on* $\mathbb{T}^1$. *The smooth solution w of theorems 1 and 3 can be continued smoothly for all time. The second derivative of u, in particular, is bounded for all time.*

Proof. Suppose $[0,T)$ gives the largest time interval for some finite T. The a priori bound from Section 2 on $\int w \tan w$ implies $\cos^2 w(t) \geq \nu$ for some $\nu > 0$ and all $t \in [0,T)$. Pick ν' with $0 < \nu' < \nu$, and consider the sequence of solutions $\{w^\epsilon\}$ to the regularized PDE (17). Since $w^\epsilon \to w$ strongly in $C\left([0,T); C^4(\mathbb{T}^N)\right)$, we have $\cos^2(w^\epsilon(t)) \geq \nu'$ for small enough ϵ and all $t \in [0,T)$. The argument used in the proof of Theorem 2 can be repeated to extend the solution w to a larger time interval $[0,T')$, contradicting the assumption that $[0,T)$ is the largest such interval.

Remark: Following the same arguments, continuation of smooth solutions on $\mathbb{T}^2$ is achieved provided that $|\cos w|$ is bounded away from zero, which is equivalent to having a pointwise bound on Δu.

Remark on (7) with fidelity: The analysis of this section can be carried out to include the fidelity term $\lambda(f-u)$ in (7) provided that we have smoothness of the signal f. The estimates are the same although the algebra is somewhat messier.

4 Laplacian limiting schemes

Real images are made of pixels and thus in practice we use a discrete version of the PDE to modify the image. The spatial derivatives are approximated using finite differences. The fact that the PDE is highly nonlinear gives us a number of choices for the scheme, even when we restrict ourselves to second order centered differences. The design of schemes for nonlinear fourth order diffusion equations is an active area of research for a different class of fourth order equations. For example, lubrication equations used to model thin film flow have highest order terms of the form

$$u_t = -\nabla \cdot (f(u)\nabla\Delta u),$$

which possess nonlinear 'entropy' estimates in a similar spirit to those derived in section 2. In the last few years numerical schemes possessing analogous discrete estimates have been shown to perform well for the lubrication problems [10, 18, 36] and also for degenerate Cahn-Hilliard equations [3] and PDEs arising in quantum systems [6, 20, 27]. These special discretizations often have, in practice, an advantage over other schemes for computing solutions that have discontinuities in a derivative.

For the original LCIS PDE, Tumblin and Turk considered a straightforward centered difference approximation of (5). Here we propose a method based on the transformation $\tan w = \Delta u$ which results in a scheme that is *guaranteed to limit the smoothness estimator of the image*, i.e. the discrete Laplacian is bounded for all time and the scheme has a global solution. We show in section 5 that the new scheme works very well for denoising of piecewise linear signals.

To numerically integrate (6) we consider a continuous time variable and a discrete spatial variable. We split the equation into a pair of equations that are solved simultaneously:

$$u_t = \Delta w, \quad \tan w = -\Delta u. \tag{43}$$

4.1 A finite difference scheme

Approximate Δ by the standard five point stencil in two dimensions, which is also called the smoothness estimator, i.e.

$$\Delta^h u(i,j) = u(i+1,j) + u(i-1,j) + u(i,j+1) + u(i,j-1) - 4u(i,j). \tag{44}$$

In the imaging literature it is often typical to consider a grid size of one which gives (44). However we note that for consistency with the PDE we might consider a grid size of h, which for (6) is equivalent to using a grid size of one a rescaling the time variable. Plugging (44) into (43) gives the following finite difference equation for our new LCIS method:

$$u_t = \Delta^h w, \quad \tan w = -\Delta^h u. \tag{45}$$

Note then that by differentiating the second equation in time and applying the smoothness estimator to the first equation we have the discrete analogue of (9),

$$(\tan w)_t = -\Delta^h \Delta^h w. \tag{46}$$

By computing the continuous time derivative above and multiplying both sides by $\cos^2 w \Delta^h \Delta^h w$ and performing a standard summation by parts on the left hand side gives the discrete analogue of (12), namely

$$\frac{1}{2}\sum_{ij}(\Delta^h w(i,j,T))^2 + \int_0^T \sum_{i,j} \cos^2(w(i,j))(\Delta^h \Delta^h w(i,j))^2 dt = \sum_{i,j}(\Delta^h w(i,j,0))^2. \tag{47}$$

We can also derive a discrete analogue of (13)

$$\sum_{i,j} w(i,j,T)\tan w(i,j,T)dx + \frac{1}{2}\sum_{i,j}(\Delta^h w(i,j,T))^2 + \int_0^T \sum_{i,j}(\Delta^h w(i,j,t))^2 dt$$
$$\leq \sum_{i,j} w(i,j,0)\tan w(i,j,0) + \frac{1}{2}\sum_{i,j}(\Delta^h w(i,j,0))^2 dx + T/4. \tag{48}$$

Since $w \tan w$ is non-negative we have the a priori bound for each i, j in the sum:

$$w(i,j,T)\tan w(i,j,T) \leq C(T, w_0) \tag{49}$$

where C depends on T and the initial condition as in the right hand side of (48). By definition, a bound on $w \tan w$ bounds the smoothness estimator of u, namely the discrete Laplacian of u, $\Delta^h u$.

Remark: For the continuous problem, we can derive an a priori bound for Δu only in one space dimension, due to the Sobolev lemma and the geometry of space. For the discrete problem we have an a priori bound in all space dimensions. If we put in the dependence on the grid size h, the estimate (49) depends on the size of the grid.

4.2 A scheme with fidelity

The above idea also works for the PDE (7) that includes the fidelity term.

We can rewrite (7) as a pair of equations

$$u_t = \Delta w + \lambda(f - u), \quad \tan w = -\Delta u \tag{50}$$

and discretize the system in the same way as (45)

$$u_t = \Delta^h w + \lambda(f - u), \quad \tan w = -\Delta^h u. \tag{51}$$

To obtain an a priori estimate, the summation by parts is more complicated, however as in the case of (48) it follows the analysis of the continuous case in (15).

$$\sum_{i,j} w(i,j,T)\tan w(i,j,T)dx + \frac{1}{2}\sum_{i,j}(\Delta^h w(i,j,T))^2 + \int_0^T \sum_{i,j}(\Delta^h w(i,j,t))^2 dt$$
$$\leq \sum_{i,j} w(i,j,0)\tan w(i,j,0) + \frac{1}{2}\sum_{i,j}(\Delta^h w(i,j,0))^2 dx + CT/4, \tag{52}$$

where C depends on the discrete H^2 norm of f and on λ. The upshot is that we obtain a similar estimate to (49) for the smoothness estimator in which the a priori bound also depends on the H^2 norm of the signal f. In practice for noisy signals, the H^2 norm could be very large, however the scheme still has w bounded away from $\pm\pi/2$. As we show in simulations below, this is crucial to obtain good performance of the algorithm.

4.3 Discretization in time and timestep control

Many papers in the imaging literature choose forward timestepping methods for solving discrete versions of the PDE. When the diffusion is fourth order, this approach can be prohibitively expensive because high frequency noise on the scale of the grid is damped at a rate that scales like the smallest scale to the fourth power. This type of problem is called "stiff" and requires an implicit time stepping method to avoid the stability constraints imposed by a forward time method.

Here we implement a backward Euler adaptive time step on (51) which means that the solution at each new time level must be computed implicitly.

$$u^{n+1} - u^n = k_n[\Delta^h w^{n+1} + \lambda(f - u^{n+1})], \quad \tan w^{n+1} = -\Delta^h u^{n+1}. \tag{53}$$

Here k_n denotes the nth time step which changes dynamically. Existence of a solution of the continuous time - discrete space problem (51) is guaranteed by the a priori bound on w and classical ODE theory. Thus the discrete time problem (53) will have a bounded solution for sufficiently small time step k_n. The boundedness of w away from $\pm\pi/2$ is a useful flag for the time step control. In the computations presented below, the nonlinear system (53) is solved using Newton's method. For the one dimensional problems discussed in this paper, Newton's method is very fast as

the linear algebra involves a banded matrix which can be efficiently inverted using the LU decomposition.

For the problem with fidelity, we approach a steady state. The adaptive backward Euler timestep allows us to take very large steps as we get close to the steady state.

5 Computational examples

5.1 Dynamic evolution of the PDE without fidelity

We consider a computational example for (6) without fidelity that illustrates two important issues in numerical simulation of such equations. The first issue is one of numerical 'pinning' of the curvature. This results in a solution that very quickly converges to a piecewise linear function on a coarse grid. In this situation, the first derivative appears to have a jump, however w remains bounded away from $\pm\pi/2$ and the PDE actually has a 'near singularity'. The second issue is one of true finite time singularities in the scheme, in which the w variable increases beyond $\pm\pi/2$. For the discrete space, continuous time problem, the above analysis shows that w should remain bounded, thus when taking discrete timesteps, it is important for the timestep to be small enough to prevent such singularities.

To illustrate numerical pinning, we consider the initial data

$$u_0(x) = 1 - \cos(6\pi x) \tag{54}$$

on the interval $[0, 1]$ with periodic boundary conditions. Figure 1 shows the results of the simulation with 500 grid points. Note that the solution very quickly evolves into a sawtooth pattern that decays to zero as time evolves. In this simulation the size of the timestep is controlled so that $w = \arctan \Delta u$ never becomes larger than $\pi/2$ in absolute value. Note that the first derivative appears to form a jump discontinuity. Interestingly the jump is not really a singularity; it is due to the coarsening of the grid. The bottom figure contrasts the same simulation on a fine grid where we see that the corner is smoothly resolved. The continuous PDE does not have singularities, as proved in the previous section.

Figure 2 considers initial data $u_0(x) = 1 - \cos(10\pi x)$ and shows the results when the timestep constraint is not restricted to keep $w = \arctan \Delta u$ bounded away from $\pi/2$ in absolute value. Spurious jumps form at $t = 0.00004$ that result in strange discontinuous behavior of the solution, as shown at later times. This simulation is also computed with 500 grid points. In addition to the timestep control, analogous spurious effects can be seen with a different choice of the spatial discretization.

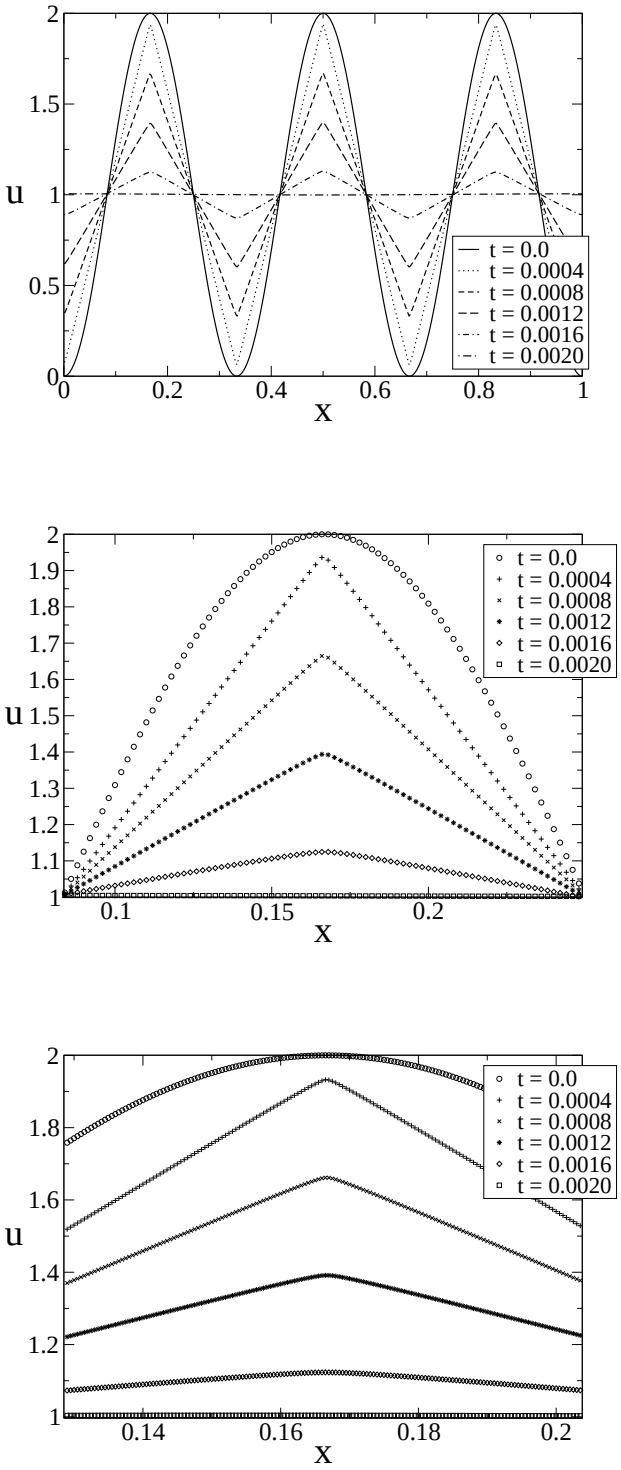

Figure 1: The solution of (6) with initial condition (54) with 500 grid points. The middle figure shows a close up of what appears to be a curvature singularity. The bottom figure shows a closeup of that singularity for a 2000 grid point simulation. The corner is actually smooth at high resolution.

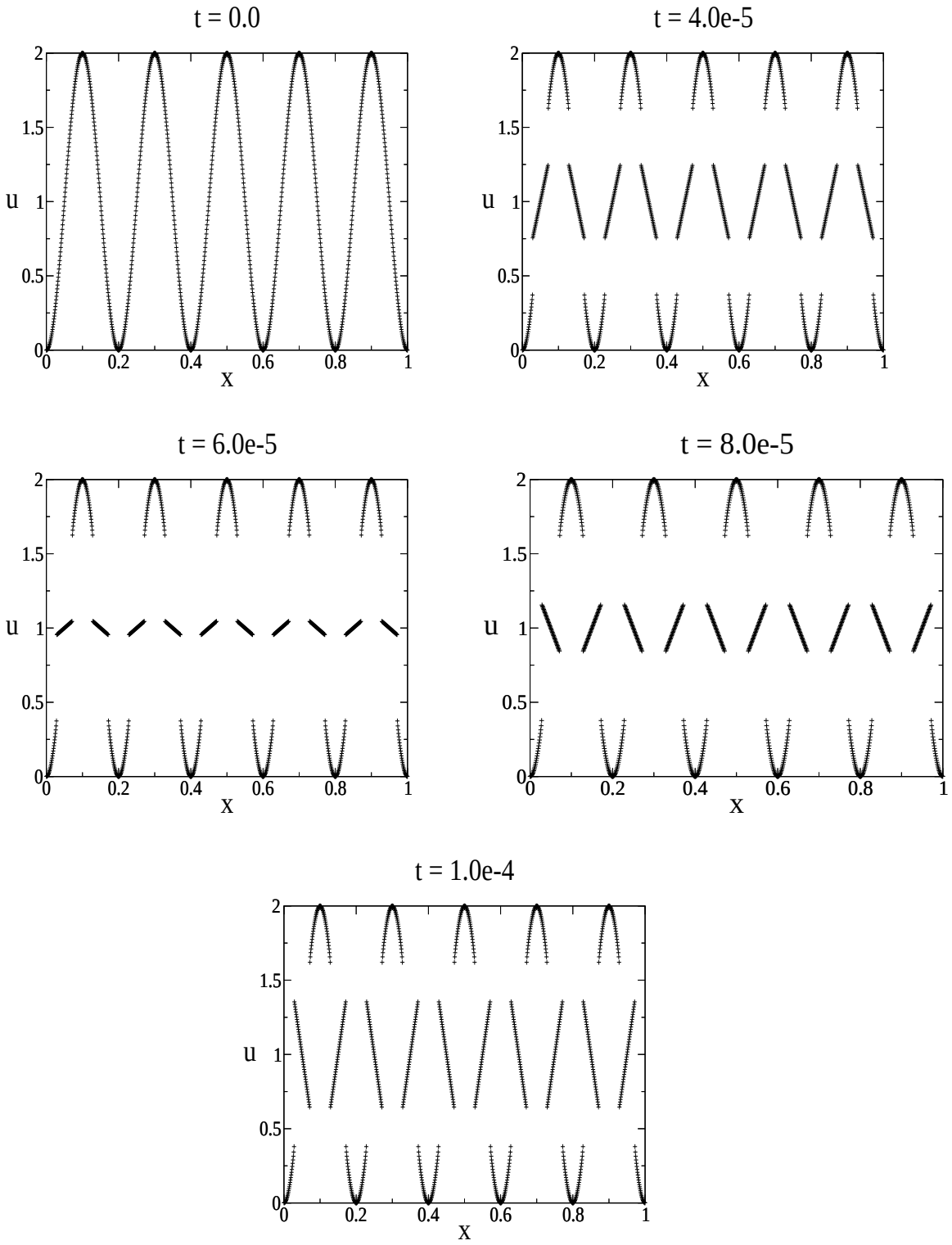

Figure 2: Spurious behavior of the scheme (53) when the timestep is too large, allowing w to cross over $\pm\pi/2$ between $t = 0$ and $t = 4.0 \times 10^{-5}$ (the top two figures).

5.2 Denoising of piecewise linear signals using LCIS with fidelity

The original LCIS scheme was designed for denoising of piecewise linear signals. Here we show computations illustrating that the above numerical method combined with a time-varying fidelity term results in a scheme that is both computationally fast and quite good at denoising of data. We include the fidelity term

$$u_t + \nabla \cdot (g(\Delta u)\nabla \Delta u) = \lambda(f - u). \tag{55}$$

For denoising, f represents the original image or signal with noise, and u is the desired denoised solution. The parameter λ determines how closely we approximate the data. In practice we find that the numerics converges to steady state much faster if we 'ramp up' λ dynamically in the simulation. That is we start with λ very small and adapt it dynamically in correlation with the timestep. We discover empirically that this allows us to take rather large timesteps to converge very quickly to steady state. Using a time varying fidelity parameter λ is not new. The original TV method of Rudin, Osher, and Fatemi [28] takes λ as a functional of u and the variance of the signal noise. The functional is derived from a constrained variational problem.

Our method is not derived from a variational problem, rather the time dependence of λ is more ad-hoc. For the computations shown here, we take $\lambda = 10^{-5}$ initially. The timestep is chosen small enough so that the Newton iteration converges quickly. If the timestep k_n is greater than 1000 times dx^4, λ increases by a factor of 1.005 until it reaches the maximum value that we specify. If the Newton iteration requires decreasing dt, we also decrease λ by 20 percent. It is important not to increase λ too quickly, or u will become singular and both the timestep and λ will need to be decreased temporarily. This can lead to an oscillating effect where u becomes more, then less singular as dt and λ are increased, then decreased. Increasing λ too slowly gives a slower method. At late times, both λ and u reach a steady state. The value of λ_{max} is prespecified, while the steady state u is the result of running the simulation for long enough time.

Figure 3 shows an example in which the signal includes both piecewise linear and piecewise constant components. Noise is added with mean zero and 0.5 standard deviation. We compare the original Perona-Malik method (without fidelity, equations (1-2)) to the LCIS with fidelity. Both simulations use 1000 grid cells and the LCIS method is run with $\lambda_{max} = 0.02$. For Perona-Malik without fidelity we must pick a stopping time. Here we choose t=0.3. The Perona-Malik method does better on the piecewise constant part of the signal, while the LCIS method does better on the piecewise linear part of the signal. Figure 4 shows similar results for a piecewise linear signal, for which the LCIS method is preferred over Perona-Malik.

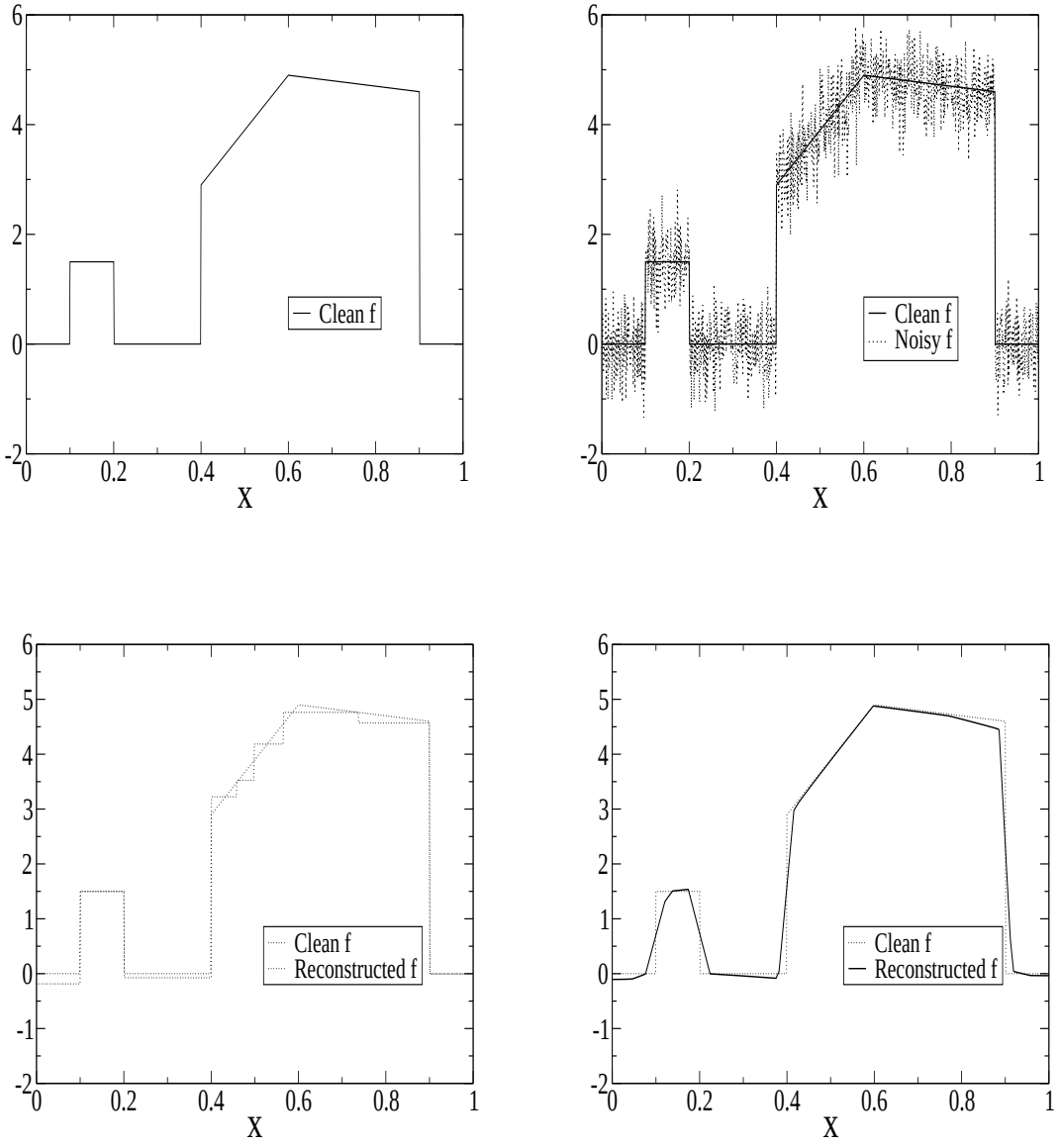

Figure 3: **Top:** The fidelity term, f, before and after adding Gaussian noise with mean 0 and 0.5 standard deviation. **Bottom:** Reconstruction of f. Perona Malik on the left, and LCIS with fidelity term on the right. Both simulations are done with 1000 grid cells. The LCIS reconstruction is given by the steady state with $\lambda = 0.02$.

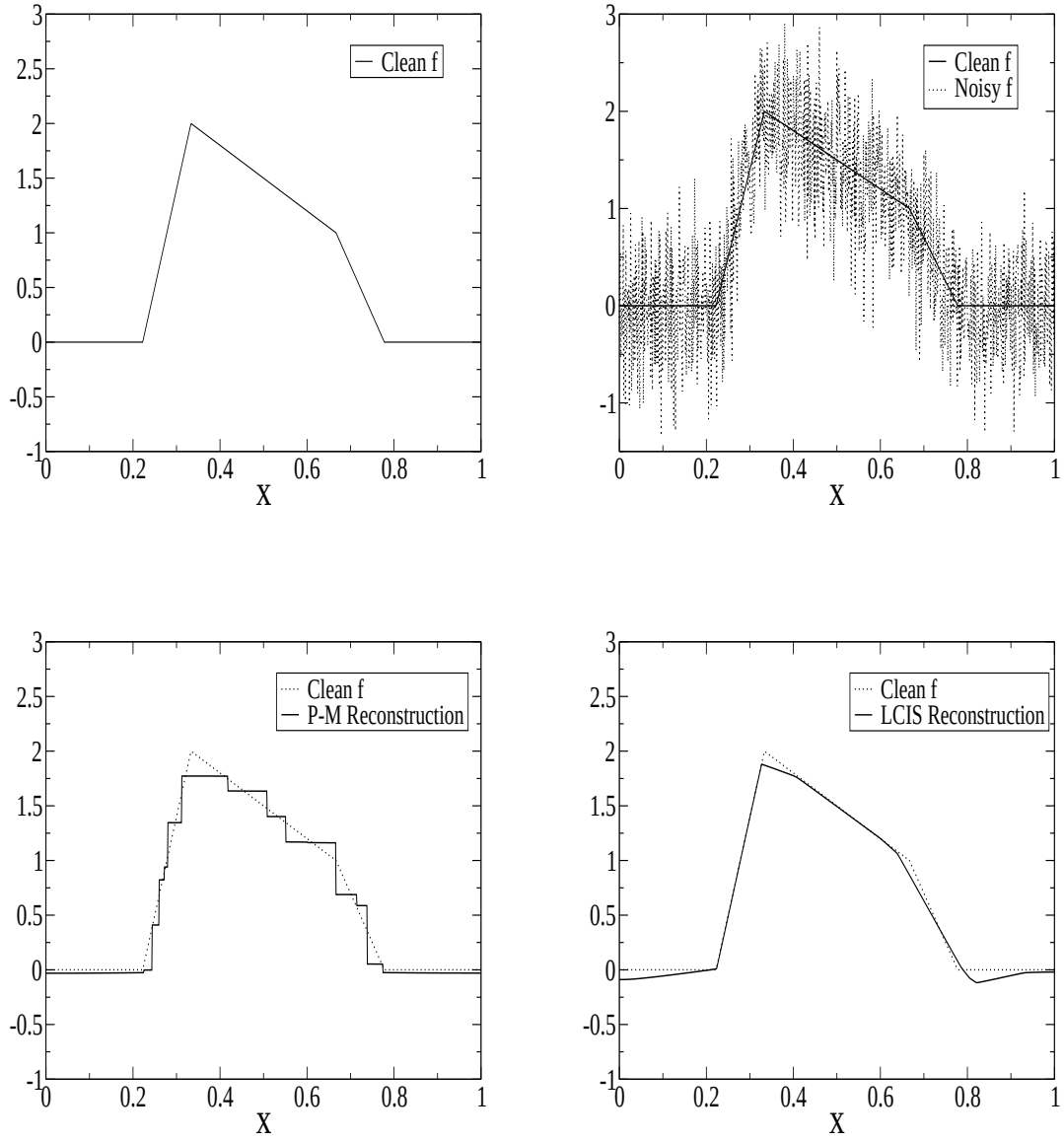

Figure 4: **Top:** The fidelity term, f, before and after adding Gaussian noise with mean 0 and 0.5 standard deviation. **Bottom:** Reconstruction of f. Perona Malik on the left, and LCIS with fidelity term on the right. Both simulations are done with 1000 grid cells. The LCIS reconstruction is given by the steady state with $\lambda = 0.02$. The Perona-Malik reconstruction is given by t = 0.3, while the LCIS steady state is reached at approximately t = 0.003.

6 Discussion

This paper has several results. Firstly we show that the idea of 'Low Curvature Image Simplifiers' originally proposed by Tumblin and Turk [33] can be reformulated in (6) as a fourth order nonlinear PDE in which the nonlinearity depends on the Laplacian (or smoothness estimator) of the image intensity function. Unlike the Perona-Malik equation, we show that this method has globally smooth solutions from smooth initial data in one space dimension, and we prove a priori bounds for solutions in one and two dimensions. Secondly, these bounds illustrate how to design a finite difference scheme that preserves a discrete version of the same estimates. While the analysis is for smooth solutions, we show in computational examples that the resulting discrete scheme performs quite well with 'near singularities' that typically result from piecewise linear data. Finally we consider the performance of this PDE for denoising of piecewise linear and piecewise constant signals. For this purpose we consider the PDE with a time-dependent fidelity term. Our results are compared with the traditional Perona-Malik method and are shown to perform better for piecewise linear data. An interesting point for further study is to better understand the theory for the LCIS equation for noisy signals f and for noisy initial data u_0. Such an approach is the motivation behind the equation discussed in Section 6.2 below.

The analysis of this paper is in part motivated by related analysis of nonlinear fourth order lubrication-type equations. This connection has relevance to two other image processing PDEs which we discuss below.

6.1 A remark on You-Kaveh

We can rewrite the You-Kaveh PDE (3) as

$$w_t = -\Delta^2(g(w)w), \quad w = \Delta u. \tag{56}$$

Consider the function f satisfying $f(g(w)w) = w$ near $w = 0$. Then the above equation above becomes $[f(v)]_t = -\Delta^2 v$ or $v_t = -\Delta^2 v / f'(v)$. By analogy to the behavior of equation (10) we gain insight as to why singularities are observed in one dimensional solutions of (56). For g as in (2), used by You-Kaveh, we note that $g(w)w$ has a derivative that changes sign at $w = 1$, this corresponded to a square-root degeneracy in $1/f'(v)$, i.e. $1/f'(v)$ vanishes like $\sqrt{|v - v_c|}$ at the critical value $v_c = \pm k$. Near the singularity, the behavior should be similar to that of (10) with $n = 1/2$, which was observed in [5] to have finite time singularities in which $h \to 0$. By analogy we should expect finite time singularities in (56). To avoid this, we could choose a different form of g so that the resulting $1/f'(v)$ vanishes faster at the critical value of v. The resulting modification of (3) however results in a PDE for which the second derivative is confined to lie between two fixed finite values. This is not so useful for images in which edges correspond to high derivatives. The advantage of the LCIS method over You-Kaveh is that while the second derivatives are bounded, the

bound is in terms of the initial data, time, and λ. So the method does not enforce an absolute cut off for Δu.

6.2 A remark on Osher-Solé-Vese

A new PDE introduced by Osher, Solé and Vese for cartoon/texture decomposition is [25]

$$u_t = -\Delta\nabla\cdot(\frac{\nabla u}{|\nabla u|}) + \lambda(f-u).$$

where u at steady state represents the cartoon part of the signal f. In one space dimension this equation becomes

$$u_t = -(\frac{u_x}{|u_x|})_{xxx} + \lambda(f-u). \tag{57}$$

Note that if we replace the discontinuous $\frac{u_x}{|u_x|}$ by the regularization $\arctan(u_x/\delta)$ then for $\lambda = 0$ equation (57) maps onto equation (9) through the change of variables $w = arctan(u_x/\delta)$. In particular we have the result that the regularized Osher-Solé-Vese PDE is well-posed for smooth data in one space dimension.

Acknowledgments

We thank Jack Tumblin for useful conversations regarding the original intent of the LCIS method. We thank Selim Esedoglu, Stanley Osher, Guillermo Sapiro, and Demetri Terzopoulos for helpful comments. This work is supported by ONR grants N000140110290 and N000140310073, NSF grant DMS-0074049, and ARO grant DAAD19-02-1-0055.

References

[1] Robert A. Adams. *Sobolev spaces.* Academic Press, New York, 1975.

[2] C. Ballester, M.Bertalmío, V.Caselles, G.Sapiro, and J.Verdera. Filling-in by joint interpolation of vector fields and gray levels. In *IEEE Trans. Image Processing*, volume 10, pages 1200–1211, 2001.

[3] John W. Barrett, James F. Blowey, and Harald Garcke. Finite element approximation of a fourth order degenerate parabolic equation. *Numer. Math.*, 80(4):525–556, October 1998.

[4] Marcelo Bertalmio, Guillermo Sapiro, Vincent Caselles, and Coloma Ballester. Image inpainting. In *SIGGRAPH*, pages 417 – 424, 2000.

[5] Andrea L. Bertozzi. Symmetric singularity formation in lubrication-type equations for interface motion. *SIAM J. Applied Math.*, 56(3):681–714, June 1996.

[6] J. A. Carrillo, A. Jungel, and S. Q. Tang. Positive entropic schemes for a nonlinear fourth-order parabolic equation. *Discrete and Continuous Dynamical Systems B*, 3(1):1–20, 2003.

[7] Francine Catte, Pierre-Louis Lions, Jean-Michel Morel, and Tomeu Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM. J. Num. Anal.*, 29(1):182–193, 1992.

[8] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math.*, 76:167–188, 1997.

[9] T. Chan, A. Marquina, and P. Mulet. High-order total variation-based image restoration. *SIAM J. Sci. Comp.*, 22(2):503–516, 2000.

[10] J. A. Diez and L. Kondic. Computing three-dimensional thin film flows including contact lines. *J. Comp. Phys.*, 183(1):274–306, 2002.

[11] H. Dym and H. P. McKean. *Fourier series and integrals.* Academic Press, New York, 1972. Probability and Mathematical Statistics, No. 14.

[12] S. Esedoglu. An analysis of the Perona-Malik scheme. *Comm. Pure Appl. Math.*, 54:1442–1487, 2001.

[13] S. Esedoglu and J. Shen. Image inpainting by the Mumford-Shah-Euler model, 2002. IMA Preprint 1812, to appear in European J. Appl. Math.

[14] Avner Friedman. *Partial differential equations of parabolic type.* Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.

[15] J. B. Greer. PhD thesis, Duke University, 2003.

[16] J. B. Greer and A. L. Bertozzi. H^1 solutions of a class of fourth order nonlinear equations for image processing. *Discrete and continuous dynamical systems*, 2003. to appear.

[17] J. B. Greer and A. L. Bertozzi. Traveling wave solutions of fourth order PDEs for image processing, 2003.

[18] G. Grün and M. Rumpf. Nonnegativity preserving convergent schemes for the thin film equation. *Num. Math.*, 87(1):113–152, 2000.

[19] P. Hartman. *Ordinary Differential Equations.* Birkhäuser, 1982.

[20] A. Jungel and R. Pinnau. A positivity-preserving numerical scheme for a nonlinear fourth order parabolic system. *SIAM J. Num. Anal.*, 39(2):385–406, 2001.

[21] Marius Lysaker, Arvid Lundervold, and Xue-Cheng Tai. Noise removal using fourth-order partial differential equations with applications to medical magnetic resonance images in space and time, 2002. UCLA CAM preprint 02-44.

[22] Marius Lysaker, Stanley Osher, and Xue-Cheng Tai. Noise removal using smoothed normals and surface fitting, 2003. UCLA CAM preprint 03-03.

[23] A. Majda and A. L. Bertozzi. *Vorticity and Incompressible Flow.* Cambridge University Press, 2002.

[24] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pur. Appl. Math.*, 42:577–685, 1989.

[25] Stanley Osher, Andrés Solé, and Luminita Vese. Image decomposition and restoration using total variation minimization and the H^{-1} norm, 2002. UCLA CAM preprint 02-57, October 2002.

[26] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

[27] R. Pinnau. Numerical approximation of the transient quantum drift diffusion model. *Nonlinear Anal. Th. Meth. Appl.*, 47(9):5849–5860, 2001.

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

[29] Michael E. Taylor. *Partial differential equations. I.* Springer-Verlag, New York, 1996. Basic theory.

[30] Michael E. Taylor. *Partial differential equations. III.* Springer-Verlag, New York, 1997. Nonlinear equations, Corrected reprint of the 1996 original.

[31] D. Terzopoulos. The computation of visible-surface representations. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 10(4), 1988.

[32] Jack Tumblin, 2003. private communication.

[33] Jack Tumblin and Greg Turk. LCIS: A boundary hierarchy for detail-preserving contrast reduction. In *Proceedings of the SIGGRAPH 1999 annual conference on Computer graphics, August 8-13, 1999, Los Angeles, CA USA*, pages 83–90, 1999. http://www.acm.org/pubs/citations/proceedings/graph/311535/p83-tumblin/.

[34] Guo W. Wei. Generalized Perona-Malik equation for image processing. *IEEE Signal Processing Letters*, 6(7):165–167, July 1999.

[35] Yu-Li You and M. Kaveh. Fourth-order partial differential equations for noise removal. *IEEE Trans. Image Process.*, 9(10):1723–1730, 2000.

[36] L. Zhornitskaya and A. L. Bertozzi. Positivity-preserving numerical schemes for lubrication-type equations. *SIAM J. Numer. Anal.*, 37(2):523–555 (electronic), 2000.