DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **9**, Number **4**, Xxxx **2003**

Website: http://AIMsciences.org

pp. **000–000**

H^1 SOLUTIONS OF A CLASS OF FOURTH ORDER NONLINEAR EQUATIONS FOR IMAGE PROCESSING

JOHN B. GREER AND ANDREA L. BERTOZZI

Abstract. Recently fourth order equations of the form $u_t = -\nabla\cdot((\mathcal{G}(J_\sigma u))\nabla\Delta u)$ have been proposed for noise reduction and simplification of two dimensional images. The operator $\mathcal{G}$ is a nonlinear functional involving the gradient or Hessian of its argument, with decay in the far field. The operator J_σ is a standard mollifier. Using ODE methods on Sobolev spaces, we prove existence and uniqueness of solutions of this problem for H^1 initial data.

1. **Introduction.** Image restoration and smoothing is important in problems ranging from medical diagnostic tests to defense applications such as target recognition. The review article [11] provides a historical description of the use of PDEs in image processing. Consider a gray-scale image, represented through pixels on a grid, by a scalar function u. The image is evolved according to a particular differential equation

$$\frac{\partial u}{\partial t} = N(u(x,y,t)) \tag{1}$$

where N is in general a nonlinear partial differential operator. The time variable is artificial in that it measures the degree of processing (e.g. smoothing or inpainting) of the image as opposed to a real time. The true problem in image processing is discrete in space, with data represented on a Cartesian grid of pixels. Thus, the PDE plays the role of a continuum approximation of the true discrete problem. This is in contrast to a numerical approximation of a continuum model from physics, where the discrete approximation is necessary in order to solve the problem on a computer with a finite memory and number of operations. Despite the role reversal, much intuition can be derived from physical models and this has led to some novel new PDEs proposed for performing specialized tasks on images. Many image processing problems involve some sort of energy minimization, and thus the PDE is related to a variational problem. The methods that we consider here are not derived from a variational setting and thus we consider dynamic methods for studying regularity properties of these equations.

Before the development of nonlinear PDE-based methods, the problem of noise reduction in images was treated through linear filtering, in which the image intensity function is convolved with a Gaussian. The method of linear filtering was introduced by Marr and Hildreth [26], and then further developed by Witkin [36], Koenderink [23], and Canny [10]. It is equivalent to solving the heat equation with initial data given by the noisy image intensity function. Although this technique quickly damps

Key words and phrases. Fourth order partial differential equations, image processing, anisotropic diffusion, mollifiers, ODEs on a Banach space, higher order nonlinear PDEs.

The authors thank Stanley Osher, Guillermo Sapiro, Michael Taylor, and the referees for useful comments. This work was supported by ONR grant N00014011029 and NSF grant DMS-0074049.

out any noise in the image, it also badly blurs edges, often leaving objects in the image unrecognizable. We consider the problem of reducing overall noise while preserving or sharpening edges. A related problem is to detect edges while ignoring fine scale scatter.

A well-known method, involving nonlinear diffusion, was proposed by Perona and Malik [28]. They define an edge indicator, g to be a smooth non-increasing function satisfying the following:

- $g(0) = 1$
- $g(s) > 0$
- $\lim_{s\to\infty} \frac{d^n}{ds^n} g(s) = 0$ for each integer $n \geq 0$.

An example typically used in applications is

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

where K is a subjective parameter. The Perona-Malik equation replaces the diffusion coefficient in the heat equation with this edge indicator:

$$\begin{aligned} u_t &= \nabla \cdot (g(|\nabla u|)\nabla u) \\ u(0,x) &= u_0(x). \end{aligned} \tag{2}$$

A related method for image denoising was introduced by Rudin and Osher [27, 29, 30] using $g(\nabla u) = 1/|\nabla u|$, which arises from a variational formulation of the problem using TV norms [2, 3, 4, 5].

Equation (2) is a continuum approximation of a discrete algorithm, also proposed by Perona and Malik [18, 28]. Although this method is effective, the continuum equation (2) is ill-posed. Specifically in regions of high gradients, backward diffusion is introduced in the direction of the gradient. This can be seen by computing the linearization of (2) about a constant gradient solution $\vec{c} \cdot \vec{x}$. In practice, the ill-posedness results in a mild instability in the discrete problem. Regions of high gradients develop a "staircase" instability that involves dynamic coarsening of the steps as time evolves [18, 22]. To make the equation well posed, a simple adjustment with practical applications, is to include a short range mollifier in the nonlinear diffusion [12]. The new, well-posed equation is given by

$$\begin{aligned} u_t - \nabla \cdot [g(|\nabla J_\sigma u|)\nabla u] &= 0 \\ u(x,0) &= u_0(x), \end{aligned} \tag{3}$$

where J_σ is a mollifier, as described in Section 2. Existence and uniqueness of solutions to this modified Perona-Malik equation has been proved for initial data $u_0 \in L^2(\Omega)$ [12].

In terms of the image processing, J_σ causes g to measure edges of u with a small amount of linear filtering. This was originally recommended by Perona and Malik to prevent their algorithm from making edges where they are unwanted, such as in the noise. This small amount of linear filtering allows g to measure edges of u in a more "global" sense, so that it is not easily affected by local discretization. In practice, σ is often made to depend on time, with σ larger in earlier iterations, when noise is most prevalent [35]. While use of the mollifier may seem to be counterproductive, since the original intention was to avoid the blurring caused by linear filtering, the results can be quite impressive, and are in fact a great improvement over linear filtering.

Although use of the Perona-Malik equation effectively reduces noise while maintaining edges, the resulting images have large jumps in intensity, which make them appear "blocky" [35]. To make the images more pleasing to the eye, it would be useful to reduce this effect. One way of doing so is to use a higher-order version of the Perona-Malik equation, examples of which are given in [33], [34], and [37].

In separate recent studies, Tumblin and Turk [33] and Wei [34] propose equations of the form

$$u_t + \nabla \cdot [g(m)\nabla \Delta u] = 0, \tag{4}$$

where g is defined as above, and m is some measurement of u. In [34], m is merely the Euclidean norm of the gradient of u. In [33], m is a measure of the curvature of u, rather than its gradient; in two dimensions they take

$$m^2 = \frac{1}{2}\left(\left(\frac{\partial^2}{\partial x_1^2}u\right)^2 + \left(\frac{\partial^2}{\partial x_2^2}u\right)^2\right) + \left(\frac{\partial^2}{\partial x_1 \partial x_2}u\right)^2.$$

You and Kaveh [37] introduce a different form of fourth-order diffusion,

$$u_t + \Delta\left[g(m)\Delta u\right] = 0, \tag{5}$$

where $m = |\Delta u|$. (5) is derived from a variational formulation, and it would be interesting to compare this problem with the second-order method introduced by Rudin and Osher. Unlike the equations considered in [33] and [34], (5) is self-adjoint. (5) was derived in a variational setting, and the natural 'energy' for that problem is the L^2 energy $\int |u|^2$, while it is preferable to consider the H^1 energy with (4).

Both [33] and [37] discover that computations with their methods result in piecewise linear solutions as $t \to \infty$. While the second-order Perona-Malik equation produces image intensity functions with sharp jumps in intensity, the fourth-order methods result in image intensity functions with jumps in gradient. This is far less obvious to the human eye, and has the potential to give cleaner images.

Fourth order equations of the type described above are very new and no analytical results have been derived for these problems. This paper proves existence and uniqueness of H^1 solutions of the higher-order version of the Perona-Malik equation discussed in [34] and [33]:

$$\begin{aligned} u_t + \nabla \cdot \left[g(|D^k J_\sigma u|)\nabla \Delta u\right] &= 0 \\ u(x,0) &= u_0(x). \end{aligned} \tag{6}$$

Our technique uses the Picard Theorem on a Banach space combined with energy methods. We consider approximate equations

$$\begin{aligned} u_t^\epsilon + J_\epsilon \nabla \cdot \left[g(|D^k J_\sigma u^\epsilon|) J_\epsilon \nabla \Delta u^\epsilon\right] &= 0 \\ u^\epsilon(x,0) &= u_0(x), \end{aligned} \tag{7}$$

where J_ϵ is again a mollifier. After proving existence of solutions u^ϵ to the approximating equations, we find a subsequence converging to a solution of (6). This constructive method is used in [32] for nonlinear second order parabolic equations and in [25] for both the Euler and Navier-Stokes equations for incompressible fluids. In this and other problems, the argument hinges on a basic energy estimate that is independent of the regularization parameter ϵ in the above. For Euler and Navier-Stokes, the energy is the physical kinetic energy of the fluid. For the problem considered here, it is the H^1 energy $\int |\nabla u|^2$. The general method is not restricted to parabolic equations as the Euler example of [25] shows.

For simplicity in this paper, we consider solutions on the N-dimensional torus, T^N. This is reasonable for problems in image processing, where computations are often performed on a square with Neumann boundary conditions. Mathematically, the periodic case is a generalization of the Neumann case, in the sense that the Neumann problem is a projection onto a subspace satisfying reflection symmetry of the full periodic problem on a larger cube.

Section 2 introduces notation and basic properties of mollifiers. In Section 3 we prove the existence of solutions to the regularized equations (7) with $u_0 \in H^m(T^N)$, $m \geq 1$. Then in Section 4, we assume u_0 is in $H^1(T^N)$ and prove the existence of a solution u to a weak form of (6). We show that this solution is unique in Section 5, and in Section 6 we prove that it is actually a strong solution of (6) that is infinitely differentiable away from $t = 0$.

2. **Notation and Properties of Mollifiers.** We shall use the following notation:

- T^N is the N-dimensional torus.
- $\| u \|_0$ is the L^2-norm of u on T^N :
$$\| u \|_0 = \left(\int_{T^N} |u|^2 \right)^{\frac{1}{2}}.$$
- 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 T^N :
$$\| u \|_m = \left(\sum_{|\alpha| \leq m} \| D^\alpha u \|_0^2 \right)^{\frac{1}{2}}.$$
- $\| u \|_\infty$ is the L^∞-norm of u on T^N :
$$\| u \|_\infty = \operatorname{ess} \sup_{x \in T^N} |u(x)|.$$
- 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
$$\| u \|_{C([0,T];X)} = \sup_{0 \leq t \leq T} \| u(t) \|_X .$$
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 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, [31] or [16]). We define the mollification, $J_\epsilon f$, of functions $f \in L^p(T^N)$, $1 \leq p \leq \infty$, by

$$J_\epsilon f(x) = \sum_{k \in Z^N} \hat{f}(k) e^{-\epsilon^2 |k|^2 + 2\pi i k \cdot x}. \tag{8}$$

J_ϵ is analogous to mollifiers defined through a convolution operator for functions on R^n [1, 19, 25] . In Lemma 1, we list those properties of mollifiers on 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 [16].

Lemma 1. *Let J_ϵ be the mollifier defined in* (8). *Then $J_\epsilon f$ is a C^∞ function with the following properties:*

1. *For all $f \in C^l\left(T^N\right)$, with $l \geq \frac{n}{2}$, $J_\epsilon f \to f$ uniformly and*
$$|J_\epsilon f|_{L^\infty} \leq |f|_{L^\infty}. \tag{9}$$
2. *For all $f, g \in L^2(T^N)$,*
$$\int_{T^N} (J_\epsilon f)\, g dx = \int_{T^N} f\, (J_\epsilon g)\, dx. \tag{10}$$
3. *Mollifiers commute with distribution derivatives,*
$$D^\alpha J_\epsilon f = J_\epsilon D^\alpha f \quad \forall |\alpha| \leq m, \ f \in H^m. \tag{11}$$
4. *For all $f \in H^s(T^N)$, $J_\epsilon f$ converges to f in $H^s(T^N)$, and the rate of convergence in the H^{s-1}-norm is linear in ϵ:*
$$\lim_{\epsilon \to 0} \| J_\epsilon f - f \|_s = 0, \tag{12}$$
$$\| J_\epsilon f - f \|_{s-1} \leq \epsilon \| f \|_s \tag{13}$$
5. *For all $f \in H^s(T^N)$, $\gamma \in Z^+ \cup \{0\}$, $m \in \{0, 1, ..., \gamma\}$, and $\epsilon > 0$,*
$$\| J_\epsilon f \|_{s+\gamma} \leq \frac{c_{s\gamma}}{\epsilon^\gamma} \| f \|_s \tag{14}$$
$$|J_\epsilon D^\gamma f|_{L^\infty} \leq \frac{c_\gamma}{\epsilon^{\frac{N}{2}+\gamma-m}} \| f \|_m \tag{15}$$

3. Existence and Uniqueness of Solutions to the Approximating Equations.

We first prove the following theorem.

Theorem 1. *Let $u_0 \in H^m$ for some integer $m \geq 1$. For every $\epsilon > 0$, there exists a solution $u^\epsilon \in C^1((0,\infty), H^m)$ to the regularized equation* (7)*, with $u^\epsilon(x, 0) = u_0(x)$.*

Our proof uses 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 [21] 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*

$$\frac{dX}{dt} = F(X), \quad X|_{t=0} = X_0 \in O, \tag{16}$$

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*

$$\frac{dX}{dt} = F(X), \quad X|_{t=0} = X_0 \in O, \tag{17}$$

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

Let $g_i^\epsilon = g(|D^k J_\sigma u_i^\epsilon|)$, and $F_\epsilon(u_i^\epsilon) = J_\epsilon \nabla \cdot (g_i^\epsilon J_\epsilon \nabla \Delta u_i^\epsilon)$. To use the Picard Theorem, we need to show that F_ϵ is locally Lipschitz continuous. The smoothness requirements on g will allow us to prove two lemmas toward that end.

Lemma 2. *For u_1, u_2 contained in a bounded set in H^1, and for every multi-index α, there exist constants C_1 and C_0, both depending on k, σ, N and α, such that*

$$\| D^\alpha (g(|D^k J_\sigma u_1|) - g(|D^k J_\sigma u_2|)) \|_\infty \le C_1 \| u_1 - u_2 \|_1, \tag{18}$$

and

$$\| D^\alpha (g(|D^k J_\sigma u_1|) - g(|D^k J_\sigma u_2|)) \|_\infty \le C_0 \| u_1 - u_2 \|_0 . \tag{19}$$

Proof

Suppose we are considering the set $O = \{u : \| u \|_1 \le M\}$. First let $|\alpha| = 0$.

$$\| D^k J_\sigma u \|_\infty \le \frac{C}{\sigma^{\frac{N}{2}+k-1}} \| u \|_1 \le \frac{C}{\sigma^{\frac{N}{2}+k-1}} M,$$

showing that $\left|D^k J_\sigma u\right|$ is bounded for u in O. Since g is Lipschitz continuous, there exists a C' such that

$$\| g(|D^k J_\sigma u_1|) - g(|D^k J_\sigma u_2|) \|_\infty \le C' \| D^k J_\sigma (u_1 - u_2) \|_\infty \le \frac{C''}{\sigma^{\frac{N}{2}+k-1}} \| u_1 - u_2 \|_1 .$$

Similarly,

$$\| g(|D^k J_\sigma u_1|) - g(|D^k J_\sigma u_2|) \|_\infty \le \frac{C}{\sigma^{\frac{N}{2}+k}} \| u_1 - u_2 \|_0 .$$

For $|\alpha| > 0$, we use the chain rule and the fact that g is smooth (and therefore $g^{(m)}$ is locally Lipschitz for each positive integer m) to give a proof as above. $\square$

Remark: Although the proof of this lemma uses C' and C'' to denote new constants, for simplicity we refrain from doing so in the remainder of this paper. Instead C is used to denote a general constant. When C depends on a parameter that we feel is important to mention (e.g. ϵ or σ), we write C as C(dependencies) (e.g. $C(\epsilon, \sigma)$).

Lemma 3. *For each multi-index α there exists a constant C (depending on k, N, α and σ) such that on a bounded set $\{u : \| u \|_1 \leq M\}$,*

$$\| D^\alpha (g(|D^k J_\sigma u_1|) - g(|D^k J_\sigma u_2|)) \|_0 \leq C \| u_1 - u_2 \|_1 . \tag{20}$$

Proof
First let $|\alpha| = 0$. As shown in the proof of Lemma 2, g is Lipschitz on bounded sets in H^1, so

$$\begin{aligned} \| g(|D^k J_\sigma u_1|) - g(|D^k J_\sigma u_2|) \|_0 &= \left(\int \left| g(\left|D^k J_\sigma u_1\right|) - g(\left|D^k J_\sigma u_2\right|) \right|^2 \right)^{\frac{1}{2}} \\ &\leq \left(\int C^2 \left| J_\sigma D^k (u_1 - u_2) \right|^2 \right)^{\frac{1}{2}} = C \| J_\sigma D^k (u_1 - u_2) \|_0 \\ &\leq C \| J_\sigma (u_1 - u_2) \|_k \leq \frac{C}{\sigma^{k-1}} \| u_1 - u_2 \|_1 . \end{aligned}$$

The chain rule and the smoothness of g give similar results for $|\alpha| > 0$. □

Proof of Theorem 1
Now we show that F_ϵ is locally Lipschitz continuous in H^m for each $m \geq 1$.

$$\begin{aligned} &\| F_\epsilon(u_1^\epsilon) - F_\epsilon(u_2^\epsilon) \|_m = \\ &\qquad \| J_\epsilon \nabla \cdot [(g_1^\epsilon (J_\epsilon \nabla \Delta u_1^\epsilon - J_\epsilon \nabla \Delta u_2^\epsilon)) + (g_1^\epsilon - g_2^\epsilon) J_\epsilon \nabla \Delta u_2^\epsilon] \|_m \\ &\leq \quad \| J_\epsilon \nabla \cdot (g_1^\epsilon (J_\epsilon \nabla \Delta (u_1^\epsilon - u_2^\epsilon))) \|_m + \| J_\epsilon \nabla \cdot ((g_1^\epsilon - g_2^\epsilon) J_\epsilon \nabla \Delta u_2^\epsilon) \|_m \\ &= \quad A + B. \end{aligned}$$

Using the mollifier inequalities and above lemmas, we have

$$\begin{aligned} A &\leq \frac{C}{\epsilon} \| g_1^\epsilon J_\epsilon \nabla \Delta (u_1^\epsilon - u_2^\epsilon) \|_m \\ &\leq \frac{C}{\epsilon} \{ \| g_1^\epsilon \|_\infty \| D^m J_\epsilon \nabla \Delta (u_1^\epsilon - u_2^\epsilon) \|_0 + \| D^m g_1^\epsilon \|_0 \| J_\epsilon \nabla \Delta (u_1^\epsilon - u_2^\epsilon) \|_\infty \} \\ &\leq \frac{C}{\epsilon} \left\{ \| g_1^\epsilon \|_\infty \frac{1}{\epsilon^3} \| u_1^\epsilon - u_2^\epsilon \|_m + \| D^m g_1^\epsilon \|_0 \frac{1}{\epsilon^{\frac{N}{2}+2}} \| u_1^\epsilon - u_2^\epsilon \|_1 \right\} \\ &\leq \frac{C}{\epsilon^{\frac{N}{2}+4}} \{ \| g_1^\epsilon \|_\infty + \| D^m g_1^\epsilon \|_0 \} \| u_1^\epsilon - u_2^\epsilon \|_m \\ &\leq C(\epsilon, m, \sigma) \| u_1 \|_1 \| u_1^\epsilon - u_2^\epsilon \|_m . \end{aligned}$$

Also using Sobolev inequalities and the above lemmas give

$$\begin{aligned} B &\leq \frac{C}{\epsilon} \| (g_1^\epsilon - g_2^\epsilon) J_\epsilon \nabla \Delta u_2^\epsilon \|_m \\ &\leq \frac{C}{\epsilon} \{ \| D^m J_\epsilon \nabla \Delta u_2^\epsilon \|_0 \| u_1^\epsilon - u_2^\epsilon \|_1 + \| D^m (g_1^\epsilon - g_2^\epsilon) \|_0 \| J_\epsilon \nabla \Delta u_2^\epsilon \|_\infty \} \\ &\leq \frac{C(\sigma)}{\epsilon} \left\{ \frac{1}{\epsilon^{m+2}} \| u_2^\epsilon \|_1 \| u_1^\epsilon - u_2^\epsilon \|_1 + \frac{1}{\epsilon^{\frac{N}{2}+2}} \| u_2^\epsilon \|_1 \| u_1^\epsilon - u_2^\epsilon \|_1 \right\} \\ &\leq \frac{C(\sigma)}{\epsilon^{\frac{N}{2}+m+3}} \| u_2^\epsilon \|_1 \| u_1^\epsilon - u_2^\epsilon \|_m . \end{aligned}$$

The final result is that

$$\| F_\epsilon(u_1^\epsilon) - F_\epsilon(u_2^\epsilon) \|_m \quad \leq \quad C \left(\| u_j^\epsilon \|_1, \epsilon, \sigma, N, m \right) \| u_1^\epsilon - u_2^\epsilon \|_m, \tag{21}$$

proving that F_ϵ is locally Lipschitz continuous on any open set

$$O^M = \{u \in H^m | \parallel u \parallel_m < M, m \geq 1\}.$$

The Picard Theorem tells us that equation (7) has a unique solution

$$u^\epsilon \in C^1([0,T), H^m)$$

for some $T > 0$.

Now we check how long solutions to (7) can be continued. To do this, we first show that

$$\sup_{0 \leq t \leq T} \parallel u^\epsilon(t) \parallel_1 \leq C \parallel u_0 \parallel_1 . \tag{22}$$

Take the L^2-inner product of equation (7) with Δu^ϵ and integrate by parts in the term with the time derivative to obtain

$$\frac{d}{dt}\frac{1}{2} \parallel \nabla u^\epsilon \parallel_0^2 - \int J_\epsilon \nabla \cdot \left[g(|D^k J_\sigma u^\epsilon|) J_\epsilon \nabla \Delta u^\epsilon \right] \Delta u^\epsilon = 0.$$

Integrating by parts and using properties of mollifiers we have

$$\frac{d}{dt}\frac{1}{2} \parallel \nabla u^\epsilon \parallel_0^2 + \int g(|D^k J_\sigma u^\epsilon|) |J_\epsilon \nabla \Delta u|^2 = 0. \tag{23}$$

So

$$\frac{d}{dt}\frac{1}{2} \parallel \nabla u^\epsilon \parallel_0^2 = - \int g(|D^k J_\sigma u^\epsilon|) |J_\epsilon \nabla \Delta u|^2 \leq 0,$$

since $g(s) \geq 0$. We now have the L^2-bound

$$\parallel \nabla u^\epsilon(t) \parallel_0 \leq \parallel \nabla u_0 \parallel_0 \tag{24}$$

for solutions of (7). Since $\int u^\epsilon$ is a conserved quantity, (24) implies

$$\parallel u^\epsilon(t) \parallel_1 \leq C \parallel u_0 \parallel_1 . \tag{25}$$

Remark: By integrating equation (23) over time, we get another useful result:

$$\frac{1}{2} \left(\parallel \nabla u^\epsilon(T) \parallel_0^2 - \parallel \nabla u_0 \parallel_0^2 \right) + \int_0^T \int g(|D^k J_\sigma u^\epsilon|) |J_\epsilon \nabla \Delta u|^2 = 0,$$

so

$$\int_0^T \int g(|D^k J_\sigma u^\epsilon|) |J_\epsilon \nabla \Delta u|^2 \leq \frac{1}{2} \parallel \nabla u_0 \parallel_0^2 . \tag{26}$$

We will now find a bound on $\parallel u^\epsilon \parallel_m$. Letting $u_2^\epsilon = 0$ in (21), we have

$$\frac{d}{dt} \parallel u^\epsilon \parallel_m \leq C(\parallel u^\epsilon \parallel_1, \epsilon, \sigma, N, n) \parallel u^\epsilon \parallel_m . \tag{27}$$

(25) and (27) give

$$\frac{d}{dt} \parallel u^\epsilon \parallel_m \leq C\left(\parallel u_0 \parallel_1, \epsilon, \sigma, N, n\right) \parallel u^\epsilon \parallel_m .$$

Grönwall's inequality implies

$$\parallel u^\epsilon(\cdot, T) \parallel_m \leq Ae^{CT}, \tag{28}$$

where A is a constant depending on $\| u_0 \|_m$. Invoking the continuation property of autonomous ODEs on a Banach space, we can extend the solution of (7) to any time interval $[0, T)$. The theorem follows. $\square$

Remark: The proof of Lemma 2 showed that for bounded sets of u in H^1, $\left|D^k J_\sigma u(t)\right|$ is bounded. Given (25), there exists some $\nu > 0$ such that

$$g(|D^k J_\sigma u|) \geq \nu > 0. \tag{29}$$

This is important for later estimates.

4. **Weak solution with initial data in H^1.** We call u a weak solution of equation (6) if it satisfies

$$\int_{T^N} \frac{\partial u}{\partial t}(t)\, v - \int_{T^N} g(|D^k J_\sigma u(t)|) \nabla \Delta u(t) \nabla v = 0 \tag{30}$$

$$u(x, 0) = u_0(x) \tag{31}$$

for any $v \in H^1$ a.e. in $[0, T]$. It is easy to check that (30) is equivalent to (6) for smooth enough u. Here we have fixed a time interval $[0, T]$, but this choice is arbitrary, due to the existence of solutions to the approximating equations for all $t \geq 0$.

We prove the following theorem.

Theorem 2. *If $u_0 \in H^1$, then equation (30) has a solution*

$$u \in L^2\left(0, T; H^3\right) \cap C\left([0, T]; H^{-1}\right) \cap L^\infty([0, T], H^1) \tag{32}$$

satisfying (31).

The proof requires passing to a limit in ϵ in the approximating equations (7). To prepare for this, in Subsection 4.1 we establish bounds on solutions of (7) that are independent of ϵ. We prove Theorem 2 in Subsection 4.2.

4.1. **A Bound on $\| u^\epsilon \|_m$ that is independent of ϵ.** In order to prove the existence of solutions of equation (6), we pass to a limit in ϵ in the approximating equations (7). To do this, we must find bounds on solutions of (7) that are independent of ϵ. From (25) and (26), we have

$$\| u^\epsilon(t) \|_1 \leq C_1 \| u_0 \|_1,$$

for $t > 0$, and

$$\int_0^T \| J_\epsilon u^\epsilon(t) \|_3^2 \, dt \leq C_2 \| u_0 \|_1^2 .$$

We use these bounds to estimate $\| u^\epsilon(t) \|_3$ and $\int_0^T \| J_\epsilon u^\epsilon(t) \|_5^2 \, dt$, then repeat this process to get bounds on even higher Sobolev norms.

Lemma 4. *Let u be a solution of (7) for any $\epsilon > 0$. For all integers $m \geq 1$ and $0 < t < T$, for any $T > 0$, if $u_0 \in H^m$, we have bounds of the form*

$$\| u^\epsilon(t) \|_m \leq C(N, \sigma, k, \nu, T, \| u_0 \|_m) \tag{33}$$

and

$$\int_0^T \| J_\epsilon u^\epsilon(t) \|_{m+2}^2 \, dt \ \leq \ C(N,\sigma,k,\nu,T,\| u_0 \|_m). \tag{34}$$

Remark: Our proof of Theorem 2 uses (33) and (34) in the case $m=1$. We prove that solutions of the type given by Theorem 2 are infinitely differentiable in space time with the use of (33) and (34) for each positive integer m.

Proof of Lemma 4

Since equations (33) and (34) hold for $m=1$, assume they hold for $m-1$. We show that bounds of a similar type are true for m. Rewrite (7) as

$$\begin{aligned}\frac{\partial u_\epsilon}{\partial t} &= -J_\epsilon g(\left|D^k J_\sigma u^\epsilon\right|)\Delta\Delta J_\epsilon u^\epsilon - J_\epsilon \nabla\left(g(\left|J_\sigma D^k u\right|)\right)\cdot J_\epsilon \nabla\Delta u^\epsilon \\ &= -J_\epsilon L_\sigma J_\epsilon u^\epsilon - J_\epsilon B_\epsilon,\end{aligned}$$

where $L_\sigma = g(\left|D^k J_\sigma u^\epsilon\right|)\Delta\Delta$, and $B_\epsilon = \nabla\left(g(\left|J_\sigma D^k u^\epsilon\right|)\right)\cdot J_\epsilon \nabla\Delta u^\epsilon$. Take the $\alpha-$derivative (for some multi-index α) of the above equation, and then take the L^2-inner product of each term with $D^\alpha \Delta u^\epsilon$ to get

$$\begin{aligned}\frac{d}{dt}\| D^\alpha \nabla u^\epsilon(t) \|_0^2 &= 2\left(D^\alpha J_\epsilon L_\sigma J_\epsilon u^\epsilon, D^\alpha \Delta u^\epsilon\right) + 2\left(D^\alpha B_\epsilon, J_\epsilon D^\alpha \Delta u^\epsilon\right) \\ &= 2(A+B).\end{aligned} \tag{35}$$

$$\begin{aligned} A &= \left(g(\left|J_\sigma D^k u^\epsilon\right|)D^\alpha \Delta\Delta J_\epsilon u^\epsilon, D^\alpha \Delta J_\epsilon u^\epsilon\right) \\ &\quad + \sum_{\substack{\mu+\beta=\alpha \\ \beta\neq\alpha}} \left(D^\mu\left(g(\left|J_\sigma D^k u^\epsilon\right|)\right) D^\beta \Delta\Delta J_\epsilon u^\epsilon, D^\alpha \Delta J_\epsilon u^\epsilon\right) \\ &= A_1 + A_2.\end{aligned} \tag{36}$$

Let $v = D^\alpha J_\epsilon u^\epsilon$. Then the Leibniz rule gives

$$\begin{aligned} A_1 &= \sum_{i,j}\left(\partial_i\left(g(\left|J_\sigma D^k u^\epsilon\right|)\partial_i\partial_j^2 v\right), \partial_j^2 v\right) - \left(\partial_i\left(g(\left|J_\sigma D^k u^\epsilon\right|)\right)\partial_i\partial_j^2 v, \partial_j^2 v\right) \\ &= \sum_{i,j} -\left(g(\left|J_\sigma D^k u^\epsilon\right|)\partial_i\partial_j^2 v, \partial_i\partial_j^2 v\right) - \left(\partial_i\left(g(\left|J_\sigma D^k u^\epsilon\right|)\right)\partial_i\partial_j^2 v, \partial_j^2 v\right) \\ &\leq -\nu \| v \|_3^2 - \sum_{i,j}\left(\partial_i\left(g(\left|J_\sigma D^k u^\epsilon\right|)\right)\partial_i\partial_j^2 v, \partial_j^2 v\right),\end{aligned}$$

with ν given by (29).

The smoothness properties of g, along with the chain rule, mollifier inequalities, and estimate (22) give

$$\| D^\mu g(\left|D^k J_\sigma u\right|) \|_{L^\infty} \ \leq \ C\left(\mu,\sigma,k,N,\| u_0 \|_1\right) \tag{37}$$

for each multi-index μ. We will now use this bound to estimate A_2 and the remaining term of A_1. Since $|\beta| \leq |\alpha| - 1$ in each term of A_2,

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

for each β, giving

$$
\begin{aligned}
A \;&\le\; -\nu \parallel D^\alpha J_\epsilon u^\epsilon \parallel_3^2 + C\,(\sigma,k,\parallel u_0 \parallel_1) \parallel D^\alpha J_\epsilon u^\epsilon \parallel_3 \parallel D^\alpha J_\epsilon u^\epsilon \parallel_2 \\
&+ C\left(\sum_{0<|\mu|\;\le\;|\alpha|} \parallel D^\mu g(\left|J_\sigma D^k u^\epsilon\right|) \parallel_\infty\right) \parallel D^\alpha J_\epsilon u^\epsilon \parallel_3 \parallel D^\alpha J_\epsilon u^\epsilon \parallel_2 \\
&\le\; -\nu \parallel D^\alpha J_\epsilon u^\epsilon \parallel_3^2 + C\,(\sigma,k,N,\parallel u_0 \parallel_1) \parallel D^\alpha J_\epsilon u^\epsilon \parallel_3 \parallel D^\alpha J_\epsilon u^\epsilon \parallel_2 \qquad (38)
\end{aligned}
$$

Next consider $B_\epsilon = \nabla g(\left|D^k J_\sigma u^\epsilon\right|) \cdot \nabla \Delta J_\epsilon u^\epsilon$. Using estimates of $\parallel D^\alpha g(\left|D^k J_\sigma u\right|) \parallel_{L^\infty}$ as above, we have

$$\parallel D^\alpha B_\epsilon \parallel_0 \;\le\; C_1\,(\sigma,\alpha,k,\parallel u_0 \parallel_1) \parallel D^\alpha J_\epsilon u^\epsilon \parallel_3,$$

so

$$(D^\alpha B_\epsilon, D^\alpha J_\epsilon \Delta u^\epsilon) \;\le\; C_1\,(\sigma,\alpha,k,\parallel u_0 \parallel_1) \parallel D^\alpha J_\epsilon u^\epsilon \parallel_3 \parallel D^\alpha J_\epsilon u^\epsilon \parallel_2 \qquad (39)$$

Using our bounds (38) and (39) on A and B, we have an upper bound for $\frac{d}{dt} \parallel D^\alpha \nabla J_\epsilon u \parallel_0^2$. Summing over all multi-indices α with $|\alpha| \;\le\; m-1$, we have

$$
\begin{aligned}
&\frac{d}{dt} \parallel \nabla u^\epsilon(t) \parallel_{m-1}^2 \le \\
&\qquad -2\nu \parallel J_\epsilon u^\epsilon \parallel_{m+2}^2 + C\,(\sigma,m,k,\parallel u_0 \parallel_1) \parallel J_\epsilon u^\epsilon \parallel_{m+2} \parallel J_\epsilon u^\epsilon \parallel_{m+1} \qquad (40)
\end{aligned}
$$

Using the inequality

$$ab \;\le\; \nu a^2 + \left(\frac{1}{4\nu}\right) b^2 \qquad (41)$$

with $a = \parallel J_\epsilon u^\epsilon \parallel_{m+2}$, we see

$$\frac{d}{dt} \parallel \nabla u^\epsilon(t) \parallel_{m-1}^2 \le\; -\nu \parallel J_\epsilon u^\epsilon \parallel_{m+2}^2 + C\,(\sigma,m,k,\nu,\parallel u_0 \parallel_1) \parallel J_\epsilon u^\epsilon \parallel_{m+1}^2,$$

and therefore

$$\frac{d}{dt} \parallel \nabla u^\epsilon(t) \parallel_{m-1}^2 \le\; C\,(\sigma,m,k,\nu,N,\parallel u_0 \parallel_1) \parallel J_\epsilon u^\epsilon(t) \parallel_{m+1}^2 . \qquad (42)$$

Integrating in time, and using the inductive hypothesis (34), we have

$$\parallel u^\epsilon(t) \parallel_m \;\le\; C\,(\sigma,m,k,\nu,N,\parallel u_0 \parallel_m, T) \qquad (43)$$

for all $0 \le t \le T$. Returning to inequality (40) and using

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

we have

$$\frac{d}{dt} \parallel \nabla u^\epsilon(t) \parallel_{m-1}^2 + \nu \parallel J_\epsilon u^\epsilon \parallel_{m+2}^2 \;\le\; \frac{1}{2} C\,(\sigma,m,k,\nu,N,\parallel u_0 \parallel_1) \parallel J_\epsilon u^\epsilon \parallel_{m+1}^2 .$$

Once again integrating in time and using the hypothesis (34), we see

$$\int_0^T \parallel J_\epsilon u^\epsilon \parallel_{m+2}^2 \;\le\; C\,(\sigma,m,k,\nu,N,\parallel u_0 \parallel_m) + C_2 \parallel u_0 \parallel_m^2 \qquad (44)$$

By induction, bounds of the form (33) and (34) hold for all $m \ge 1$. $\square$

4.2. **Proof of Theorem 2.** Theorem 1 showed that there exists a smooth solution u^ϵ to

$$\begin{aligned} u_t^\epsilon + J_\epsilon \nabla \cdot \left[g(|D^k J_\sigma u^\epsilon|) J_\epsilon \nabla \Delta u^\epsilon \right] = 0 \\ u^\epsilon(x,0) = J_\epsilon u_0(x), \end{aligned} \tag{45}$$

for all $\epsilon > 0$. Lemma 4 gives us the following bounds, all of which are uniform in ϵ :

$$\| J_\epsilon u^\epsilon \|_{L^2(0,T;H^3)} \leq C_1 \left(\sigma, k, T, \| u_0 \|_{H^1}\right) \tag{46}$$

$$\| u^\epsilon \|_{L^\infty(0,T;H^1)} \leq \| u_0 \|_{H^1} . \tag{47}$$

Also from the approximating equations (45), for all $v \in H^1$ we have

$$\begin{aligned} \int_{T^N} \frac{du^\epsilon}{dt}(t) v &= \int_{T^N} g(|D^k J_\sigma u^\epsilon|) J_\epsilon \nabla \Delta u^\epsilon \nabla J_\epsilon v \\ &\leq \| J_\epsilon u^\epsilon(t) \|_3 \| v \|_1, \end{aligned}$$

so by (46),

$$\| \frac{du^\epsilon}{dt} \|_{L^2(0,T;H^{-1})} \leq C_2 \left(\sigma, k, T, \| u_0 \|_{H^1}\right). \tag{48}$$

The uniform bounds (46) and (48) imply the existence of a subsequence $\{u^\epsilon\}$ such that

$$J_\epsilon u^\epsilon \rightharpoonup u \text{ in } L^2\left(0,T;H^3\right) \tag{49}$$

$$\frac{du^\epsilon}{dt} \rightharpoonup \frac{du}{dt} \text{ in } L^2\left(0,T;H^{-1}\right), \tag{50}$$

where $\rightharpoonup$ is used to denote weak convergence. Also by (46),

$$\frac{\partial^3 J_\epsilon u^\epsilon}{\partial x_i \partial x_j \partial x_k} \text{ is bounded in } L^2\left(0,T;L^2\right),$$

and passing to subsequences again we have

$$\frac{\partial^3 J_\epsilon u^\epsilon}{\partial x_i \partial x_j \partial x_k} \rightharpoonup \frac{\partial^3 u}{\partial x_i \partial x_j \partial x_k} \text{ in } L^2\left(0,T;L^2\right). \tag{51}$$

We now use the Lions-Aubin Lemma. We give a short statement of a form of the Lemma which we need, but the more general version with proof can be found in [24].

Lemma 5 (Lions-Aubin Lemma)**.** *Let E_0, E, and E_1 be reflexive Banach spaces such that $E_0 \subset E \subset E_1$. Suppose that the imbedding $E_0 \to E$ is compact and the imbedding $E \to E_1$ is continuous. If $\{u_k\}$ is a bounded sequence in $L^2(0,T;E_0)$ and $\{\frac{du_k}{dt}\}$ is a bounded sequence in $L^2(0,T;E_1)$, then there exists a subsequence of $\{u_k\}$ which converges strongly both in $L^2(0,T;E)$ and in $C([0,T];E_1)$.*

Using the Lions-Aubin Lemma with the bounds (47)and (48), a subsequence can be chosen so that

$$u^\epsilon \to u \text{ strongly in } L^2\left(0,T;L^2\right) \text{ and } C\left([0,T];H^{-1}\right). \tag{52}$$

This along with the convergence properties of mollifiers show that

$$u^{\epsilon}(0) \to u_0 \text{ in } H^{-1}.$$

Let $v \in C\left([0,T]; H^1\right)$. Integrate (30) over $(0,T)$ for each u^{ϵ} :

$$\int_0^T \int_{T^N} \frac{\partial u^{\epsilon}}{\partial t}(t)\, v(t) - \int_0^T \int_{T^N} g(|D^k J_\sigma u^{\epsilon}|) \nabla \Delta J_\epsilon u^{\epsilon}(t) \nabla J_\epsilon v(t) = 0. \tag{53}$$

We show that the limit of this equation is

$$\int_0^T \int_{T^N} \frac{\partial u}{\partial t}(t)\, v(t) - \int_0^T \int_{T^N} g(|D^k J_\sigma u|) \nabla \Delta u(t) \nabla v(t) = 0 \tag{54}$$

as $\epsilon \to 0$. To do this, we subtract (54) from (53) and show

$$\begin{aligned}
&\lim_{\epsilon \to 0} \int_0^T \int_{T^N} \left(\frac{\partial u^{\epsilon}}{\partial t}(t) - \frac{\partial u}{\partial t}(t) \right) v(t) \\
&- \int_0^T \int_{T^N} g(|D^k J_\sigma u^{\epsilon}|) \nabla \Delta J_\epsilon u^{\epsilon}(t) \nabla J_\epsilon v(t) - g(|D^k J_\sigma u|) \nabla \Delta u(t) \nabla v(t) \\
&= 0.
\end{aligned}$$

Consider the second integral.

$$\begin{aligned}
&\left| \int_0^T \int_{T^N} g(|D^k J_\sigma u^{\epsilon}|) \nabla \Delta J_\epsilon u^{\epsilon}(t) \nabla J_\epsilon v(t) - g(|D^k J_\sigma u|) \nabla \Delta u(t) \nabla v(t) \right| \\
&\leq \sum_{i,j=1}^{N} \left| \int_0^T \int_{T^N} \left(g(|D^k J_\sigma u^{\epsilon}|) - g(|D^k J_\sigma u|) \right) \partial_i \partial_j^2 J_\epsilon u^{\epsilon} \partial_i J_\epsilon v \right| \\
&+ \left| \int_0^T \int_{T^N} g(|D^k J_\sigma u|) \left(\partial_i \partial_j^2 J_\epsilon u^{\epsilon} - \partial_i \partial_j^2 u \right) \partial_i J_\epsilon v \right| \\
&+ \left| \int_0^T \int_{T^N} g(|D^k J_\sigma u|) \partial_i \partial_j^2 u \left(\partial_i J_\epsilon v - \partial_i v \right) \right| \\
&= \sum_{i,j=1}^{N} A_{ij} + B_{ij} + C_{ij}.
\end{aligned}$$

Using Lemma 2,

$$\begin{aligned}
A_{ij} &\leq C(\sigma) \int_0^T \| u^{\epsilon}(t) - u(t) \|_0 \int_{T^N} \left| \partial_i \partial_j^2 J_\epsilon u^{\epsilon} \partial_i J_\epsilon v \right| \\
&\leq C(\sigma, v) \int_0^T \| u^{\epsilon}(t) - u(t) \|_0 \| \partial_i \partial_j^2 J_\epsilon u^{\epsilon} \|_0 \\
&\leq C(\sigma, v) \| J_\epsilon u^{\epsilon} \|_{L^2(0,T;H^3)} \| u^{\epsilon} - u \|_{L^2(0,T;L^2)},
\end{aligned}$$

which tends to 0 by (46) and the strong convergence of u^{ϵ} to u in $L^2\left(0,T;L^2\right)$.

Since g is bounded above by a constant, each $B_{ij} \to 0$ by the weak convergence of $\partial_i \partial_j^2 J_\epsilon u^{\epsilon}$ in $L^2\left(0,T;L^2\right)$. Also since $\frac{\partial^3 u}{\partial x_i \partial x_j \partial x_k} \in L^2\left(0,T;L^2\right)$, $C_{ij} \to 0$ by the strong convergence of $\partial_i J_\epsilon v$ to $\partial_i v$ in L^2.

Finally, from (48) we have

$$\int_0^T \int_{T^N} \frac{\partial u^\epsilon}{\partial t}(t)\, v \to \int_0^T \int_{T^N} \frac{\partial u}{\partial t}(t)\, v$$

as $\epsilon \to 0$. $\square$

5. **Uniqueness of Weak Solutions.**

Theorem 3. *Let $u_0 \in H^1$. The solution u of* (30) *satisfying* (32) *is unique.*

Proof
Let u_1 and u_2 be two solutions of (30) with $u_1(x,0) = u_2(x,0) = u_0$. Assuming they satisfy (32), $u_1, u_2 \in L^2(0,T;H^3)$. So for almost every $t \in (0,T]$, $\Delta(u_1 - u_2) \in H^1$ and

$$\int_{T^N} \frac{du_1}{dt}\Delta(u_1-u_2) - \int_{T^N} g(|D^k J_\sigma u_1|)\nabla\Delta u_1 \nabla\Delta(u_1-u_2) = 0 \tag{55}$$

$$\int_{T^N} \frac{du_2}{dt}\Delta(u_1-u_2) - \int_{T^N} g(|D^k J_\sigma u_2|)\nabla\Delta u_2 \nabla\Delta(u_1-u_2) = 0 \tag{56}$$

$$u_1(x,0) = u_2(x,0) = u_0(x). \tag{57}$$

Subtracting (55) from (56), and using energy estimates such as those done in Subsection 4.1, we have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt} \| \nabla(u_1-u_2) \|_0^2 + \nu \| \nabla\Delta(u_1-u_2) \|_0^2 \leq \\ \| g(|D^k J_\sigma u_1|) - g(|D^k J_\sigma u_2|) \|_\infty \| \nabla\Delta u_2 \|_0 \| \nabla\Delta(u_1-u_2) \|_0,\end{aligned}$$

where ν is given by (29). Using Lemma 2 and (41) as in Section 3, we arrive at

$$\frac{d}{dt} \| \nabla(u_1-u_2) \|_0^2 \leq \frac{C(k,\sigma)}{4\nu} \| \nabla(u_1-u_2) \|_0^2 \| \nabla\Delta u_2 \|_0^2 . \tag{58}$$

Integrating on some time interval $(0,t)$, $t < T$ gives us

$$\| \nabla(u_1-u_2)(t) \|_0^2 \leq \frac{C(k,\sigma)}{4\nu} \| u_2 \|_{L^2(0,T;H^3)} \int_0^t \| \nabla(u_1-u_2)(s) \|_0^2 \, ds. \tag{59}$$

Using the integral form of Grönwall's inequality, we have

$$\| \nabla u_1(t) - \nabla u_2(t) \|_0 = 0 \text{ for all } t \in [0,T].$$

Since the mean $\int u$ is a conserved quantity, the above result is sufficient to prove uniqueness.

6. **Regularity of Solutions.**

Theorem 4. *Let $u_0 \in H^1$. The solution u of* (30) *that satisfies* (32) *is actually a strong solution of* (6)*. Furthermore, $u \in C^\infty\left((0,\infty) \times T^N\right)$.*

Proof

Our proof is a standard bootstrapping argument. Since solutions u of (30) are in $L^2(0,T;H^3)$, we can pick a $t_0 \in (0,T)$ arbitrarily close to 0, with $u(t_0) \in H^3$. Consider the new initial value problem,

$$\begin{aligned} \tilde{u}_t + \nabla \cdot \left[g(|D^k J_\sigma \tilde{u}|) \nabla \Delta \tilde{u} \right] = 0 \\ \tilde{u}(x,t_0) = \tilde{u}_0(x) = u(t_0) \in H^3. \end{aligned} \tag{60}$$

We define a weak solution as in (30). By our uniqueness results, $\tilde{u}(t) = u(t)$ for $t \geq t_0$. Dropping the tilde, we consider approximating equations like (45):

$$\begin{aligned} u_t^\epsilon + J_\epsilon \nabla \cdot \left[g(|D^k J_\sigma u^\epsilon|) J_\epsilon \nabla \Delta u^\epsilon \right] = 0 \\ u^\epsilon(x,t_0) = J_\epsilon u_0(x). \end{aligned} \tag{61}$$

From the estimates (33) and (34) given in Lemma 4, we see

$$\| J_\epsilon u^\epsilon \|_{L^2(0,T;H^5)} \leq C_1 \left(\sigma, k, T, \| u_0 \|_{H^3}\right) \tag{62}$$
$$\| u^\epsilon \|_{L^\infty(0,T;H^3)} \leq C_2 \left(\sigma, k, T, \| u_0 \|_{H^3}\right). \tag{63}$$
$$\tag{64}$$

Our earlier estimate (48) still holds as well. Proceeding as in the proof of Theorem 2, we find a subsequence $\{u^\epsilon\}$ such that

$$J_\epsilon u^\epsilon \rightharpoonup u \text{ in } L^2\left(0,T;H^5\right) \tag{65}$$
$$\frac{du^\epsilon}{dt} \rightharpoonup \frac{du}{dt} \text{ in } L^2\left(0,T;H^{-1}\right) \tag{66}$$
$$u^\epsilon \to u \text{ strongly in } L^2\left(0,T;H^1\right) \text{ and } C\left([0,T];H^{-1}\right) \tag{67}$$
$$u^\epsilon(t_0) \to u_0 \in H^{-1}. \tag{68}$$

It can then be shown, as in the proof of Theorem 2, that u solves the weak form of (60), and

$$u \in L^2\left(t_0,T;H^5\right) \cap C\left([t_0,T];H^{-1}\right) \cap L^\infty([t_0,T],H^3). \tag{69}$$

Furthermore, since the solution of (30) was in H^3 for all $t \in (0,T]$ except on a set of measure 0, $t_0 > 0$ could be chosen arbitrarily small and we have

$$u \in L^2\left(0,T;H^5\right) \cap C\left([0,T];H^{-1}\right) \cap L^\infty((0,T],H^3). \tag{70}$$

Now we pick a new time $t_0 > 0$ such that $u(t_0) \in H^5$ and repeat the above to show

$$u \in L^2\left(0,T;H^7\right) \cap C\left([0,T];H^{-1}\right) \cap L^\infty((0,T],H^5).$$

This argument can be continued to show that $u(t)$ is in as high a Sobolev space as we would like for $t > 0$. It follows that u is actually a strong solution of (6). Since

$$u_t = -\nabla \cdot \left[g(|D^k J_\sigma u|) \nabla \Delta u \right],$$

u_t is in H^m for all m. Taking time derivatives of (6), we see that u has smooth time derivatives of all orders. We therefore have $u \in C^\infty\left((0,\infty) \times T^N\right)$. $\square$

7. Remarks.

The family of equations considered here all have the basic H^1 energy $\int |\nabla u|^2$ with an a priori bound independent of the mollifying parameter σ. Thus, on mathematical grounds, it makes sense to consider H^1 initial data for this problem. The computations in [33] and [34] suggest that the method performs well on problems in which the image or signal has jumps in the gradient. These computations are performed with $\sigma = 0$. Tumblin and Turk describe jumps in gradient that appear in their numerical solutions, however the solutions remain continuous [33]. An interesting open problem is well-posedness of these nonlinear diffusion equations for H^1 data, which can include discontinuities and singularities in the gradient. For images with jump discontinuities in the image intensity function, as is often the case near sharp edges, the H^1 constraint is more severe. We can think of sharp edges as two dimensional objects in which the discontinuity is one-dimensional in nature. Thus the H^1 constraint requires that sharp edges be represented by singularities that are at worst square-root cusps (i.e. $C^{1/2}$ functions in the direction of the gradient). In practice, this smoothing could be done over a very small number of pixels, and it would be interesting to know more about the performance and regularity properties of these equations for initial data that is L^2 or only slightly smoothed from L^2.

We also briefly mention that higher-order nonlinear partial differential equations have been proposed for "image inpainting" [8, 13, 15, 14, 17]. Image inpainting is the process of restoring regions of missing information in photographs (e.g. holes and scratches) by using the information surrounding these regions. In [8] the authors propose inpainting with the equation

$$u_t = \nabla^{\perp} u \cdot \nabla \Delta u + \nu \nabla \cdot \left[g(|D^k J_\sigma u|) \nabla u \right], \tag{71}$$

with small $\nu > 0$. If one replaces the second-order diffusion with fourth-order diffusion of the type discussed in this paper, the H^1-norm of the image intensity function dissipates. It would be interesting to compare the effects of inpainting with these two types of diffusion and also for the closely related method of Navier-Stokes based inpainting [7].

A final comment is that there has been a burst of activity during the last decade in analysis of fourth order nonlinear parabolic equations arising in thin films [9]. These problems have nonlinear diffusion that is typically degenerate and depends on the solution itself, where u represents the thickness of the film. Several recent papers [6, 20, 38] address numerical schemes for weak solutions, where the solution vanishes on a set of positive measure. It would be interesting to explore the connection between ideas from this field and the image processing problem.

REFERENCES

[1] Robert A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[2] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. The dirichlet problem for the total variation flow. *J. Functional Analysis*, 180:347–403, 2001.

[3] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Minimizing total variational flow. *Differential and integral equations*, 14(3):321–360, 2001.

[4] F. Andreu, V. Caselles, and J. M. Mazón. Existence and uniqueness of solution for a parabolic quasilinear problem for linear growth functionals wtih l^1 data, 2001. Mathematisches Annalen, to appear, see http://www.iua.upf.es/ vcaselles/ for the preprint.

[5] F. Andreu, V. Caselles, and J. M. Mazón. A parabolic quasilinear problem for linear growth functionals, 2001. to appear in Revista Matemática Iberoamericana, see http://www.iua.upf.es/ vcaselles/ for the preprint.

[6] 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.
[7] Marcelo Bertalmio, Andrea Bertozzi, and Guillermo Sapiro. Navier-Stokes, fluid-dynamics and image and video inpainting. In *IEEE CVPR 2001, Hawaii*, volume I, December 2001.
[8] Marcelo Bertalmio, Guillermo Sapiro, Vincent Caselles, and Coloma Ballester. Image inpainting. In *SIGGRAPH*, pages 417 – 424, 2000.
[9] Andrea L. Bertozzi. The mathematics of moving contact lines in thin liquid films. *Notices Am. Math. Soc.*, 45(6):689–697, 1998.
[10] J Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8:679–698, 1986.
[11] Vicent Caselles, Jean-Michel Morel, Guillermo Sapiro, and Allen Tannenbaum. Introduction to the special issue on partial differential equations and geometry-driven diffusion in image processing and analysis. *IEEE Transactions on image processing*, 7(3):269–273, March 1998.
[12] 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.
[13] C.Ballester, M.Bertalmo, V.Caselles, G.Sapiro, and J.Verdera. Filling-in by joint interpolation of vector fields and gray levels. In *IEEE Trans. Image Processing*, August 2001.
[14] T. F. Chan, S. H. Kang, and J. Shen. Euler's elastica and curvature based inpaintings, 2002.
[15] T. F. Chan and J. Shen. Non-texture inpainting by curvature-driven diffusions (CDD), 2002.
[16] H. Dym and H. P. McKean. *Fourier series and integrals*. Academic Press, New York, 1972. Probability and Mathematical Statistics, No. 14.
[17] S. Esedoglu and J. Shen. Digital inpaintings based on the Mumford-Shah-Euler image model, 2002.
[18] Selim Esedoglu. *An analysis of the Perona-Malik Scheme*. PhD thesis, Courant Institute of Mathematical Sciences, NYU, July 2000.
[19] Avner Friedman. *Partial differential equations of parabolic type*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.
[20] G. Grün and M. Rumpf. Nonnegativity-preserving schemes for lubrication type equations. *Numer. Math.*, 87:113–152, 2000.
[21] Philip Hartman. *Ordinary differential equations*. John Wiley & Sons Inc., New York, 1964.
[22] Satyanad Kichenassamy. The Perona-Malik paradox. *SIAM J. Appl. Math.*, 57(5):1328–1342, 1997.
[23] Jan J. Koenderink. The structure of images. *Biol. Cybernet.*, 50(5):363–370, 1984.
[24] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod, 1969.
[25] A. Majda and A. L. Bertozzi. *Vorticity and Incompressible Flow*. Cambridge University Press, 2002.
[26] D. Marr and E. Hildreth. Theory of edge detection. *Proc. Roy. Soc. London Ser. B*, pages 187–217, 1980.
[27] S. Osher and L. Rudin. Feature oriented image enhancement using shock filters. *SIAM J. Num. Anal.*, 27:919–940, 1990.
[28] 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.
[29] L. Rudin and S. Osher. Total variation based image restoration with free local constraints. In *Proc IEEE ICIP, Austin*, volume 1, pages 31–35, November 1994.
[30] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
[31] Michael E. Taylor. *Partial differential equations. I*. Springer-Verlag, New York, 1996. Basic theory.
[32] Michael E. Taylor. *Partial differential equations. III*. Springer-Verlag, New York, 1997. Nonlinear equations, Corrected reprint of the 1996 original.
[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] R. Whitaker and S. Pizer. A multi-scale approach to nonuniform diffusion. *CVGIP:Image Understanding*, 57(1):99–110, 1993.

[36] A. Witkin. Scale-space filtering. In *Proceedings of IJCAI, Karlsruhe*, pages 1019–1021, 1983.
[37] Yu-Li You and M. Kaveh. Fourth-order partial differential equations for noise removal. *IEEE Trans. Image Process.*, 9(10):1723–1730, 2000.
[38] L. Zhornitskaya and A. L. Bertozzi. Positivity preserving numerical schemes for lubrication-type equations. *S.I.A.M. J. Num Anal.*, 37(2):523–555, 2000.

E-mail address: `jbg33@math.duke.edu, bertozzi@math.duke.edu`