\textbf{H1 SOLUTIONS OF A CLASS OF FOURTH ORDER NONLINEAR EQUATIONS FOR IMAGE PROCESSING}

JOHN B. GREER AND ANDREA L. BERTOZZI

\textbf{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 $J_\sigma$ is a standard mollifier. Using ODE methods on Sobolev spaces, we prove existence and uniqueness of solutions of this problem for $H^2$ initial data.

1. \textbf{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

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

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 clamps...
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:

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

An example typically used in applications is

\[
g(s) = \frac{1}{1 + (Ks)^2},
\]

where \( K \) is a subjective parameter. The Perona-Malik equation replaces the dif-
fusion coefficient in the heat equation with this edge indicator:

\[
\begin{align*}
ux + \nabla \cdot (g(|\nabla u|)\nabla u) = 0 \\
u(0, x) = u_0(x).
\end{align*}
\]

(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 \( \nabla \cdot \mathbf{z} \). 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

\[
\begin{align*}
ux - \nabla \cdot [g(|\nabla \ast_{\sigma} u|)\nabla u] = 0 \\
u(x, 0) = u_0(x),
\end{align*}
\]

(3)

where \( \ast_{\sigma} \) 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, \( \ast_{\sigma} \) 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,
\( \sigma \) is often made to depend on time, with \( \sigma \) 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 "bloody" [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

$$u_t + \nabla \cdot [g(m) \nabla u] = 0,$$

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:

$$u_t + \Delta [g(m) \Delta u] = 0,$$

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]:

$$u_t + \nabla \cdot [g(D^k J_q u) \nabla u] = 0$$

$$u(x, 0) = u_0(x).$$

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

$$u^e_t + J_e \nabla \cdot [g(D^k J_q u^e) J_e \nabla u^e] = 0$$

$$u^e(x, 0) = u_0(x),$$

where $J_e$ is again a mollifier. After proving existence of solutions $u^e$ 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 $\epsilon$ 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 $\alpha = (\alpha_1, \ldots, \alpha_N)$, $\alpha_i \in \mathbb{Z} \cup \{0\}$,
  \[ D^\alpha f(x) = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_N}}{\partial x_N^{\alpha_N}} f(x). \]

  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 $L^\infty$-norm of $u$ on $T^N$:
  \[ \|u\|_{\infty} = \sup_{x \in T^N} |u(x)|. \]

- 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
  \[ \|u\|_{L^\infty(0,T;X)} = \sup_{t \in (0,T)} \|u(t)\|_X. \]

- $L^2(0,T;X)$ 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
\[ \hat{f}(k) = \int_{T^N} f(x)e^{-2\pi ik \cdot x} \, dx \]
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
\[ J_{\epsilon} f(x) = \sum_{k \in Z^N} \hat{f}(k)e^{-\epsilon^2 \|k\|^2 + 2\pi ik \cdot x}. \]  
(8)

$J_{\epsilon}$ is analogous to mollifiers defined through a convolution operator for functions on $R^n$ [1, 19, 21]. 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 $J_{\epsilon}$ be the mollifier defined in (8). Then $J_{\epsilon} f$ is a $C^\infty$ function with the following properties:

1. For all $f \in C^l(T^N)$, with $l \geq \frac{N}{2}$, $J_{\epsilon} f \to f$ uniformly and
   \[ |J_{\epsilon} f|_{L^\infty} \leq |f|_{L^\infty}. \]  
   (9)

2. For all $f, g \in L^2(T^N)$,
   \[ \int_{T^N} (J_{\epsilon} f)(x) g(x) \, dx = \int_{T^N} f(x) (J_{\epsilon} g)(x) \, dx. \]  
   (10)

3. Mollifiers commute with distribution derivatives,
   \[ D^\alpha J_{\epsilon} f = J_{\epsilon} D^\alpha f \quad \forall |\alpha| \leq m, \ f \in H^m. \]  
   (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 $\epsilon$:
   \[ \lim_{\epsilon \to 0} ||J_{\epsilon} f - f||_s = 0. \]  
   (12)

   \[ ||J_{\epsilon} f - f||_{s-1} \leq \epsilon ||f||_s \]  
   (13)

5. For all $f \in H^s(T^N)$, $\gamma \in Z^+ \cup \{0\}$, $m \in \{0, 1, \ldots, \gamma\}$, and $\epsilon > 0$,
   \[ ||J_{\epsilon} f||_{s+\gamma} \leq \frac{C_\gamma}{\epsilon^{\gamma}} ||f||_s \]  
   (14)

   \[ |J_{\epsilon} D^\gamma f|_{L^\infty} \leq \frac{C_\gamma}{\epsilon^{\gamma+m}} \| f \|_m \]  
   (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

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

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

Theorem (Continuation 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 operator. Then the unique solution $X \in C^1([0,T];O)$ to the autonomous ODE

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

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

Let $g'_\epsilon = g([D^k J_\epsilon u_\epsilon^n])$, and $F_\epsilon(u_\epsilon^n) = J_\epsilon \nabla \cdot (g'_\epsilon J_\epsilon \Delta u_\epsilon^n)$. To use the Picard Theorem, we need to show that $F_\epsilon$ 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 $\alpha$, there exist constants $C_1$ and $C_0$, both depending on $k, \sigma, N$ and $\alpha$, such that

$$\| D^\alpha (g([D^k J_\sigma u_1]) - g([D^k J_\sigma u_2])) \|_\infty \leq C_1 \| u_1 - u_2 \|_1, \quad \text{and}$$

$$\| D^\alpha (g([D^k J_\sigma u_1]) - g([D^k J_\sigma u_2])) \|_\infty \leq C_0 \| u_1 - u_2 \|_0. \quad \text{ (18)}$$

Proof

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

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

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

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

Similarly,

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

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. \hfill $\square$

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. $\epsilon$ or $\sigma$), we write $C$ as $C$(dependencies) (e.g. $C(\epsilon, \sigma)$).
Lemma 3. For each multi-index $\alpha$ there exists a constant $C$ (depending on $k$, $N$, $\alpha$ and $\sigma$) such that on a bounded set $\{u : \|u\|_1 \leq M\}$,
\[ \| D^{\alpha}(g(D^k J_\sigma u_1) - g(D^k J_\sigma u_2)) \|_0 \leq C \| u_1 - u_2 \|_1. \] (20)

Proof
First let $|\alpha| = 0$. As shown in the proof of Lemma 2, $g$ is Lipschitz on bounded sets in $\mathcal{H}^1$, so
\[ \| g(D^k J_\sigma u_1) - g(D^k J_\sigma u_2) \|_0 = \left( \int \| g(D^k J_\sigma u_1) - g(D^k J_\sigma u_2) \|^2 \right)^{\frac{1}{2}} \leq \left( \int C^2 \| J_\sigma D^k(u_1 - u_2) \|^2 \right)^{\frac{1}{2}} = C \| J_\sigma D^k(u_1 - u_2) \|_0 \leq \frac{C}{\sigma^{k-1}} \| u_1 - u_2 \|_1. \]

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

Proof of Theorem 1
Now we show that $F_\varepsilon$ is locally Lipschitz continuous in $\mathcal{H}^m$ for each $m \geq 1$.
\[ \| F_\varepsilon(u_1^\varepsilon) - F_\varepsilon(u_2^\varepsilon) \|_m = \| J_\varepsilon \nabla \cdot \left( (g_1 J_\varepsilon \nabla (u_1^\varepsilon - u_2^\varepsilon) - J_\varepsilon \nabla \Delta u_1^\varepsilon) + (g_2^\varepsilon - g_2^\varepsilon) J_\varepsilon \nabla \Delta u_2^\varepsilon \right) \|_m \]
\[ \leq \| J_\varepsilon \nabla \cdot (g_1 J_\varepsilon \nabla (u_1^\varepsilon - u_2^\varepsilon)) \|_m + \| J_\varepsilon \nabla \cdot ((g_1^\varepsilon - g_2^\varepsilon) J_\varepsilon \nabla \Delta u_2^\varepsilon) \|_m = A + B. \]

Using the mollifier inequalities and above lemmas, we have
\[ A \leq \frac{C}{\varepsilon} \| g_1 J_\varepsilon \nabla \Delta (u_1^\varepsilon - u_2^\varepsilon) \|_m \]
\[ \leq \frac{C}{\varepsilon} \left( \| g_1 \|_\infty + \| D^m J_\varepsilon \nabla \Delta (u_1^\varepsilon - u_2^\varepsilon) \|_0 + \| D^m g_1 \|_0 \| D^m J_\varepsilon \nabla \Delta (u_1^\varepsilon - u_2^\varepsilon) \|_\infty \right) \]
\[ \leq \frac{C}{\varepsilon} \left( \| g_1 \|_\infty + \frac{1}{\varepsilon^3} \| u_1^\varepsilon - u_2^\varepsilon \|_m + \| D^m g_1 \|_0 \frac{1}{\varepsilon^{i+2}} \| u_1^\varepsilon - u_2^\varepsilon \|_1 \right) \]
\[ \leq \frac{C}{\varepsilon^{\frac{i}{2} + 1}} \left( \| g_1 \|_\infty + \| D^m g_1 \|_0 \right) \| u_1^\varepsilon - u_2^\varepsilon \|_m \]
\[ \leq C(\varepsilon, m, \sigma) \| u_1^\varepsilon \|_1 \| u_1^\varepsilon - u_2^\varepsilon \|_m. \]

Also using Sobolev inequalities and the above lemmas give
\[ B \leq \frac{C}{\varepsilon} \| (g_1^\varepsilon - g_2^\varepsilon) J_\varepsilon \nabla \Delta u_2^\varepsilon \|_m \]
\[ \leq \frac{C}{\varepsilon} \left( \| D^m J_\varepsilon \nabla \Delta u_2^\varepsilon \|_0 \| u_1^\varepsilon - u_2^\varepsilon \|_1 + \| D^m (g_1^\varepsilon - g_2^\varepsilon) \|_0 \| J_\varepsilon \nabla \Delta u_2^\varepsilon \|_\infty \right) \]
\[ \leq \frac{C(\sigma)}{\varepsilon} \left( \| u_2^\varepsilon \|_1 \| u_1^\varepsilon - u_2^\varepsilon \|_1 + \| u_2^\varepsilon \|_1 \| u_1^\varepsilon - u_2^\varepsilon \|_1 \right) \]
\[ \leq \frac{C(\sigma)}{\varepsilon^{\frac{i}{2} + m + 3}} \| u_2^\varepsilon \|_1 \| u_1^\varepsilon - u_2^\varepsilon \|_m. \]

The final result is that
\[ \| F_\varepsilon(u_1^\varepsilon) - F_\varepsilon(u_2^\varepsilon) \|_m \leq C \left( \| u_1^\varepsilon \|_1, \varepsilon, \sigma, N, m \right) \| u_1^\varepsilon - u_2^\varepsilon \|_m, \] (21)
proving that $F_\varepsilon$ is locally Lipschitz continuous on any open set

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

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

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

for some $T > 0$.

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

$$\sup_{0 \leq t \leq T} \| u^\varepsilon(t) \|_1 \leq C \| u_0 \|_1 . \quad (22)$$

Take the $L^2$-inner product of equation (7) with $\Delta u^\varepsilon$ and integrate by parts in the term with the time derivative to obtain

$$\frac{d}{dt} \frac{1}{2} \| \nabla u^\varepsilon \|_0^2 - \int J_\varepsilon \nabla \cdot \left[ g(|D^k J_\varepsilon u^\varepsilon|) J_\varepsilon \nabla \Delta u^\varepsilon \right] \Delta u^\varepsilon = 0.$$

Integrating by parts and using properties of mollifiers we have

$$\frac{d}{dt} \frac{1}{2} \| \nabla u^\varepsilon \|_0^2 + \int g(|D^k J_\varepsilon u^\varepsilon|) J_\varepsilon \nabla \Delta u^\varepsilon \nabla u^\varepsilon = 0 \quad (23)$$

So

$$\frac{d}{dt} \frac{1}{2} \| \nabla u^\varepsilon \|_0^2 - \int g(|D^k J_\varepsilon u^\varepsilon|) J_\varepsilon \nabla \Delta u^\varepsilon \nabla u^\varepsilon \leq 0,$$

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

$$\| \nabla u^\varepsilon(t) \|_0 \leq \| \nabla u_0 \|_0 \quad (24)$$

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

$$\| u^\varepsilon(t) \|_1 \leq C \| u_0 \|_1 . \quad (25)$$

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

$$\frac{1}{2} (\| \nabla u^\varepsilon(T) \|_0^2 - \| \nabla u_0 \|_0^2) + \int_0^T \int g(|D^k J_\varepsilon u^\varepsilon|) J_\varepsilon \nabla \Delta u^\varepsilon \nabla u^\varepsilon = 0,$$

so

$$\int_0^T \int g(|D^k J_\varepsilon u^\varepsilon|) J_\varepsilon \nabla \Delta u^\varepsilon \nabla u^\varepsilon \leq \frac{1}{2} \| \nabla u_0 \|_0^2 . \quad (26)$$

We will now find a bound on $\| u^\varepsilon \|_m \cdot$ Letting $u_1 = 0$ in (21), we have

$$\frac{d}{dt} \| u^\varepsilon \|_m \leq C(\| u^\varepsilon \|_1, \varepsilon, \sigma, N, n) \| u^\varepsilon \|_m . \quad (27)$$

(25) and (27) give

$$\frac{d}{dt} \| u^\varepsilon \|_m \leq C(\| u_0 \|_1, \varepsilon, \sigma, N, n) \| u^\varepsilon \|_m \cdot$$

Gr"onwall’s inequality implies

$$\| u^\varepsilon(*, T) \|_m \leq A \varepsilon^C T, \quad (28)$$
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. \qed

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

$$ g([D^k J_\sigma u]) \geq \nu > 0, \quad (29) $$

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

$$ \int_{\Omega} \frac{\partial u}{\partial t}(t) v - \int_{\Omega} g([D^k J_\sigma u(t)]) \nabla u(t) \nabla v = 0 \quad (30) $$

$$ u(x, 0) = u_0(x) \quad (31) $$

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 $t > 0$.

We prove the following theorem.

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

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

satisfying (31).

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

4.1. A Bound on $\| u^\epsilon \|_m$ that is independent of $\epsilon$. In order to prove the existence of solutions of equation (6), we pass to a limit in $\epsilon$ in the approximating equations (7). To do this, we must find bounds on solutions of (7) that are independent of $\epsilon$. 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 $\int_0^T \| J_\epsilon u^\epsilon(t) \|_6^2 \, dt$, 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

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

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

\[
\frac{\partial u_\epsilon}{\partial t} = -J_\epsilon g(J^k J_\sigma u^\epsilon | \Delta \Delta J_\epsilon u^\epsilon - J_\epsilon \nabla (g(J_\sigma D^k u)) \cdot J_\epsilon \nabla \Delta u^\epsilon
\]

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

\[
\frac{d}{dt} \| D^\alpha \nabla u^\epsilon(t) \|_0^2 = 2(D^\alpha J_\epsilon L_\sigma J_\epsilon u^\epsilon, D^\alpha \Delta u^\epsilon) + 2(D^\alpha B_\epsilon, J_\epsilon D^\alpha \Delta u^\epsilon)
\]

\[
= 2(A + B).
\]

\[
A = (g(J_\sigma D^k u^\epsilon | D^\alpha \Delta J_\epsilon u^\epsilon, D^\alpha \Delta J_\epsilon u^\epsilon)
+ \sum_{\mu + |\beta| = \alpha} (D^\mu (g(J_\sigma D^k u^\epsilon)) D^\beta \Delta J_\epsilon u^\epsilon, D^\alpha \Delta J_\epsilon u^\epsilon)
\]

\[
= A_1 + A_2,
\]

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

\[
A_1 = \sum_{i,j} (\partial_i (g(J_\sigma D^k u^\epsilon)) \partial_i \partial_j^2 v, \partial_j^2 v) - (\partial_i (g(J_\sigma D^k u^\epsilon)) \partial_i \partial_j^2 v, \partial_j^2 v)
\]

\[
= \sum_{i,j} - (g(J_\sigma D^k u^\epsilon)) \partial_i \partial_j^2 v, \partial_i \partial_j^2 v) - (\partial_i (g(J_\sigma D^k u^\epsilon)) \partial_i \partial_j^2 v, \partial_j^2 v)
\]

\[
\leq -\nu \| v \|_0^2 - \sum_{i,j} (\partial_i (g(J_\sigma D^k u^\epsilon)) \partial_i \partial_j^2 v, \partial_j^2 v),
\]

with \( \nu \) given by (29).

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

\[
\| D^\mu g(J^k J_\sigma u) \|_{L^\infty} \leq C(\mu, \sigma, k, N, \| u_0 \|_1)
\]

for each multi-index \( \mu \). We will now use this bound to estimate \( A_2 \) and the remaining term of \( A_1 \). Since \( |\beta| = |\mu| - 1 \) in each term of \( A_2 \),

\[
\| D^\beta \Delta J_\epsilon u^\epsilon \| \leq \| D^\alpha J_\epsilon u^\epsilon \|_3
\]

for each \( \beta \), giving
\[ A \leq -\nu \| D^\alpha J u^\varepsilon \|^3_3 + C(\sigma, k, \| u_0 \|_1) \| D^\alpha J u^\varepsilon \|_3 \| D^\alpha J \varepsilon u^\varepsilon \|_2 + C \left( \sum_{0 < |\alpha| \leq |\alpha|} \| D^\mu g[J_{\sigma\varepsilon} D^k u^\varepsilon] \|_{\infty} \right) \| D^\alpha J u^\varepsilon \|_3 \| D^\alpha J \varepsilon u^\varepsilon \|_2 \]

\[ \leq -\nu \| D^\alpha J u^\varepsilon \|^3_3 + C(\sigma, k, N, \| u_0 \|_1) \| D^\alpha J u^\varepsilon \|_3 \| D^\alpha J \varepsilon u^\varepsilon \|_2 \]  

(38)

Next consider \( B_c = \nabla g[D^k J_{\sigma\varepsilon} u^\varepsilon] \cdot \nabla \Delta J u^\varepsilon \). Using estimates of \( \| D^\alpha g[D^k J_{\sigma\varepsilon} u^\varepsilon] \|_{L^\infty} \) as above, we have

\[ \| D^\alpha B_c \|_0 \leq C(\sigma, \alpha, k, \| u_0 \|_1) \| D^\alpha J \varepsilon u^\varepsilon \|_3 \]

so

\[ (D^\alpha B_c, D^\alpha J \Delta u^\varepsilon) \leq C(\sigma, \alpha, k, \| u_0 \|_1) \| D^\alpha J u^\varepsilon \|_3 \| D^\alpha J \varepsilon u^\varepsilon \|_2 \]  

(39)

Using our bounds (38) and (39) on \( A \) and \( B \), we have an upper bound for

\[ \frac{d}{dt} \| D^\alpha \nabla u^\varepsilon(t) \|_{m-1}^2 \leq -2\nu \| J_{\varepsilon u^\varepsilon} \|_{m+2}^2 + C(\sigma, m, k, \| u_0 \|_1) \| J_{\varepsilon u^\varepsilon} \|_{m+2} \| J_{\varepsilon u^\varepsilon} \|_{m+1} \]  

(40)

Using the inequality

\[ ab \leq \nu a^2 + \left( \frac{1}{4\nu} \right) b^2 \]  

with \( a = \| J_{\varepsilon u^\varepsilon} \|_{m+2} \), we see

\[ \frac{d}{dt} \| \nabla u^\varepsilon(t) \|_{m-1}^2 \leq -\nu \| J_{\varepsilon u^\varepsilon} \|_{m+2}^2 + C(\sigma, m, k, \nu, \| u_0 \|_1) \| J_{\varepsilon u^\varepsilon} \|_{m+1}^2, \]

and therefore

\[ \frac{d}{dt} \| \nabla u^\varepsilon(t) \|_{m-1}^2 \leq C(\sigma, m, k, \nu, N_\varepsilon, \| u_0 \|_1) \| J_{\varepsilon u^\varepsilon} \|_{m+1}^2. \]  

(42)

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

\[ \| u^\varepsilon(t) \|_m \leq C(\sigma, m, k, \nu, N_\varepsilon, \| u_0 \|_{m, T}) \]  

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

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

we have

\[ \frac{d}{dt} \| \nabla u^\varepsilon(t) \|_{m-1}^2 + \nu \| J_{\varepsilon u^\varepsilon} \|_{m+2}^2 \leq \frac{1}{2}C(\sigma, m, k, \nu, N_\varepsilon, \| u_0 \|_1) \| J_{\varepsilon u^\varepsilon} \|_{m+1}^2, \]

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

\[ \int_0^T \| J_{\varepsilon u^\varepsilon} \|_{m+2}^2 \leq C(\sigma, m, k, \nu, N_\varepsilon, \| u_0 \|_{m}) + C_2 \| u_0 \|_m^2 \]  

(44)

By induction, bounds of the form (33) and (34) hold for all \( m \geq 1 \).
4.2. Proof of Theorem 2. Theorem 1 showed that there exists a smooth solution \( u^\epsilon \) to
\[
\begin{align*}
  u^\epsilon_t + J^\epsilon \nabla \cdot \left[ g(|D^k J^\sigma u^\epsilon|) J^\epsilon \nabla \Delta u^\epsilon \right] &= 0 \\
  u^\epsilon(x,0) &= J^\epsilon u_0(x),
\end{align*}
\]
for all \( \epsilon > 0 \). Lemma 4 gives us the following bounds, all of which are uniform in \( \epsilon \):
\[
\begin{align*}
  \| J^\epsilon u^\epsilon \|_{L^2(0,T;H^3)} &\leq C_1 (\sigma, k, T; \|u_0\|_{H^1}) \\
  \| u^\epsilon \|_{L^\infty(0,T;H^1)} &\leq \| u_0 \|_{H^1}.
\end{align*}
\]
Also from the approximating equations (45), for all \( v \in H^1 \) we have
\[
\int_T \frac{d}{dt}(J^\epsilon u^\epsilon)(t)v = \int_T g(|D^k J^\sigma u^\epsilon|) J^\epsilon \nabla \Delta u^\epsilon \nabla J^\epsilon v 
\leq \| J^\epsilon u^\epsilon(t) \|_3 \| v \|_1,
\]
so by (46),
\[
\| \frac{d}{dt} u^\epsilon \|_{L^2(0,T;H^{-1})} \leq C_2 (\sigma, k, T; \|u_0\|_{H^1}).
\]
The uniform bounds (46) and (48) imply the existence of a subsequence \( \{ u^\epsilon \} \) such that
\[
J^\epsilon u^\epsilon \rightharpoonup u \text{ in } L^2(0,T;H^3) \quad \text{(49)}
\]
\[
\frac{d}{dt} u^\epsilon \rightharpoonup \frac{d}{dt} u \text{ in } L^2(0,T;H^{-1}) \quad \text{(50)}
\]
where \( \rightharpoonup \) is used to denote weak convergence. Also by (46),
\[
\frac{\partial^3 J^\epsilon u^\epsilon}{\partial x_i \partial x_j \partial x_k} \text{ is bounded in } L^2(0,T;L^2),
\]
and passing to subsequences again we have
\[
\frac{\partial^3 J^\epsilon u^\epsilon}{\partial x_i \partial x_j \partial x_k} \rightarrow \frac{\partial^3 u}{\partial x_i \partial x_j \partial x_k} \text{ in } L^2(0,T;L^2) \quad \text{(51)}
\]
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 embedding \( E_0 \rightarrow E \) is compact and the embedding \( E \rightarrow E_1 \) is continuous. If \( \{ u_k \} \) is a bounded sequence in \( L^2(0,T;E_0) \) and \( \{ \frac{d}{dt} u_k \} \) 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
\[
\begin{align*}
  u^\epsilon \rightarrow u \text{ strongly in } L^2(0,T;L^2) \text{ and } C([0,T];H^{-1}).
\end{align*}
\]
This along with the convergence properties of mollifiers show that
\[ u^\epsilon(0) \rightarrow u_0 \text{ in } H^{-1}. \]

Let \( v \in C([0,T];H^1) \). Integrate (30) over \((0,T)\) for each \( u^\epsilon \):

\[
\int_0^T \int_{\mathcal{T}^N} \frac{\partial u^\epsilon}{\partial t}(t)v(t) - \int_0^T \int_{\mathcal{T}^N} g(L^k J_\sigma u^\epsilon) \nabla \Delta J_k u^\epsilon(t) \nabla J_k v(t) = 0. \quad (53)
\]

We show that the limit of this equation is

\[
\int_0^T \int_{\mathcal{T}^N} \left( \frac{\partial u}{\partial t}(t) - \frac{\partial u^\epsilon}{\partial t}(t) \right) v(t) \]

\[
- \int_0^T \int_{\mathcal{T}^N} g(L^k J_\sigma u^\epsilon) \nabla \Delta J_k u^\epsilon(t) \nabla J_k v(t) - g(L^k J_\sigma u) \nabla \Delta u(t) \nabla v(t) = 0. \quad (54)
\]

As \( \epsilon \rightarrow 0 \). To do this, we subtract (54) from (53) and show

\[
\lim_{\epsilon \to 0} \int_0^T \int_{\mathcal{T}^N} \left( \frac{\partial u}{\partial t}(t) - \frac{\partial u^\epsilon}{\partial t}(t) \right) v(t)
- \int_0^T \int_{\mathcal{T}^N} g(L^k J_\sigma u^\epsilon) \nabla \Delta J_k u^\epsilon(t) \nabla J_k v(t) - g(L^k J_\sigma u) \nabla \Delta u(t) \nabla v(t) = 0.
\]

Consider the second integral.

\[
\left| \int_0^T \int_{\mathcal{T}^N} g(L^k J_\sigma u^\epsilon) \nabla \Delta J_k u^\epsilon(t) \nabla J_k v(t) - g(L^k J_\sigma u) \nabla \Delta u(t) \nabla v(t) \right|
\]

\[
\leq \sum_{k,j=1}^N \left| \int_0^T \int_{\mathcal{T}^N} \left( g(L^k J_\sigma u^\epsilon) - g(L^k J_\sigma u) \right) \partial_i \partial_j^2 J_k u^\epsilon \partial_i J_k v \right|
+ \left| \int_0^T \int_{\mathcal{T}^N} g(L^k J_\sigma u) \left( \partial_i \partial_j^2 J_k u^\epsilon - \partial_i \partial_j^2 u \right) \partial_i J_k v \right|
+ \left| \int_0^T \int_{\mathcal{T}^N} g(L^k J_\sigma u) \partial_i \partial_j^2 u \left( \partial_i J_k v - \partial_i v \right) \right|
= \sum_{k,j=1}^N A_{ij} + B_{ij} + C_{ij}.
\]

Using Lemma 2,

\[
A_{ij} \leq C(\sigma) \int_0^T \| u^\epsilon(t) - u(t) \|_0 \int_{\mathcal{T}^N} \| \partial_i \partial_j^2 J_k u^\epsilon \partial_i J_k v \|
\leq C(\sigma, v) \int_0^T \| u^\epsilon(t) - u(t) \|_0 \| \partial_i \partial_j^2 J_k u^\epsilon \|_0
\leq C(\sigma, v) \| J_k u^\epsilon \|_{L^2(0,T;H^2)} \| u^\epsilon - u \|_{L^2(0,T;L^2)},
\]

which tends to 0 by (46) and the strong convergence of \( u^\epsilon \) to \( u \) in \( L^2(0,T;L^2) \).

Since \( g \) is bounded above by a constant, each \( B_{ij} \rightarrow 0 \) by the weak convergence of \( \partial_i \partial_j^2 J_k u^\epsilon \) in \( L^2(0,T;L^2) \). Also since \( \frac{\partial u^\epsilon}{\partial t} \|_{L^2(0,T;H^2)} \leq 0 \), \( C_{ij} \rightarrow 0 \) by the strong convergence of \( \partial_i J_k v \) to \( \partial i v \) in \( L^2 \).
Finally, from (48) we have

$$\int_0^T \int_{T^N} \frac{\partial u}{\partial \ell}(t)v \rightarrow \int_0^T \int_{T^N} \frac{\partial u}{\partial \ell}(t)v$$

as $\epsilon \rightarrow 0$. □

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), $u_1, u_2 \in L^2(0,T; H^2)$. So for almost every $t \in (0,T]$, $\Delta (u_1 - u_2) \in H^1$ and

$$\int_{T^N} \frac{d u_1}{d t} \Delta (u_1 - u_2) - \int_{T^N} g(D^k J_{\sigma} u_1) \nabla \Delta (u_1 - u_2) = 0 \quad (55)$$

$$\int_{T^N} \frac{d u_2}{d t} \Delta (u_1 - u_2) - \int_{T^N} g(D^k J_{\sigma} u_2) \nabla \Delta (u_1 - u_2) = 0 \quad (56)$$

$$u_1(x,0) = u_2(x,0) = u_0(x). \quad (57)$$

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

$$\frac{1}{2} \frac{d}{dt} \| \nabla (u_1 - u_2) \|^2_{0} + \nu \| \nabla \Delta (u_1 - u_2) \|^2_{0} \leq$$

$$\| g(D^k J_{\sigma} u_1) - g(D^k J_{\sigma} u_2) \|_{\infty} \| \nabla u_2 \|_{0} \| \nabla \Delta (u_1 - u_2) \|_{0},$$

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

$$\frac{d}{dt} \| \nabla (u_1 - u_2) \|^2_{0} \leq \frac{C(k_2 \sigma)}{4\nu} \| \nabla (u_1 - u_2) \|^2_{0} \| \nabla \Delta u_2 \|^2_{0}. \quad (58)$$

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

$$\| \nabla (u_1 - u_2)(t) \|^2_{0} \leq \frac{C(k_2 \sigma)}{4\nu} \| u_2 \|_{L^2(0,T; H^2)} \int_0^t \| \nabla (u_1 - u_2)(s) \|^2_{0} ds. \quad (59)$$

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, $u \in C^\infty ((0,\infty) \times T^N)$. 
**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,

$$
\ddot{u} + \nabla \cdot [g(|D^k J_\sigma u|) \nabla \Delta u] = 0
$$

$$
\dot{u}(x, t_0) = \dot{u}_0(x) = u(t_0) \in H^3.
$$

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):

$$
u^\ell + J_\varepsilon \nabla \cdot [g(|D^k J_\sigma u^\ell|) J_\varepsilon \nabla \Delta u^\ell] = 0
$$

$$
u^\ell(x, t_0) = J_\varepsilon u_0(x).
$$

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

$$
\| J_\varepsilon u^\ell \|_{L^2(0,T;H^5)} \leq C_1(\sigma, k, \varepsilon, \| u_0 \|_{H^3})
$$

$$
\| u^\ell \|_{L^\infty(0,T;H^3)} \leq C_2(\sigma, k, \varepsilon, \| u_0 \|_{H^3}).
$$

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

$$
J_\varepsilon u^\ell \rightharpoonup u \text{ in } L^2(0,T;H^5)
$$

$$
\frac{d u^\ell}{dt} \to \frac{d u}{dt} \text{ in } L^2(0,T;H^{-1})
$$

$$
u^\ell \to u \text{ strongly in } L^2(0,T;H^1) \text{ and } C([0,T];H^{-1})
$$

$$u^\ell(t_0) \to u_0 \in H^{-1}.
$$

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

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

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

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

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

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

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

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

$u_t$ 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 $u \in C^\infty((0,\infty) \times T^N). \square$
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 $\sigma$. 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. $C^{3/2}$ 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

$$u_t = \nabla^2 u \cdot \nabla \Delta u + \nu \nabla \cdot \left[ g(|D^2 J_{\sigma}u|) \nabla u \right], \tag{71}$$

with small $\nu > 0$. If one replaces the second-order diffusion with fourth-order diffusion of the type discussed in this paper, the $H^4$-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

FOURTH ORDER EQUATIONS FOR IMAGE PROCESSING


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