REGULARITY THEORY
FOR PARABOLIC NONLINEAR INTEGRAL OPERATORS

LUIS CAFFARELLI, CHI HIN CHAN, AND ALEXIS VASSEUR

1. Introduction

The purpose of this work is to develop a regularity theory for nonlocal evolution equations of variational type with “measurable” kernels. More precisely, we consider solutions of the evolution equations of the type

\[ w_t(t,x) = \int [w(t,y) - w(t,x)] K(t,x,y) \, dy, \]

where all that is required of the kernel \( K \) is that there exists \( 0 < s < 2 \) and \( 0 < \Lambda \), such that

\[
\text{symmetry in } x,y : \quad K(t,x,y) = K(t,y,x) \text{ for any } x \neq y,
\]

\[
1_{\{|x-y| \leq 3\}} \frac{1}{\Lambda} |x-y|^{-(N+s)} \leq K(t,x,y) \leq \Lambda |x-y|^{-\left(N+s\right)}.
\]

The symmetry of the kernel \( K \) makes of the operator

\[ \int [w(y) - w(x)] K(x,y) \, dy \]

the Euler-Lagrange equation of the energy integral

\[ E(w) = \int \int [w(x) - w(y)]^2 K(x,y) \, dx \, dy. \]

It suggests a mathematical treatment based on the De Giorgi-Nash-Moser ideas [12, 20] from the calculus of variations. In fact, one of the immediate applications of our result is to nonlinear variational integrals

\[ E_\phi(w) = \int \int \phi(w(x) - w(y)) K(x-y) \, dx \, dy, \]

for \( \phi \) a \( C^2 \) strictly convex functional. Indeed, the fact that \( K(x,y) \) has the special form \( K(x-y) \) makes the equation translation invariant, and as in the second-order case, this implies that first derivatives of \( w \) satisfy an equation of the type (1.1). Our
results are basically that solutions with initial data in \( L^2 \) become instantaneously bounded and Hölder continuous.

Kassmann showed previously a similar result for the stationary case in a very interesting paper where he developed the corresponding Moser scheme \([15]\). In a later paper \([1]\), Barlow, Bass, Chen, and Kassmann considered a similar time-dependent equation with less restrictive assumptions on the kernel. Under their assumptions, they constructed an example where Hölder continuity does not hold, due to the fact that their equations are of variable order and their assumptions are not scale invariant. Note, however, that they do prove a Harnack inequality under their general assumptions. For the time-dependent case, a version of the result was proved by Komatsu \([17]\), using the corresponding Nash scheme (see also Chen \([10]\)).

Other results have been obtained for the nondivergence case. Along these lines, see Bass and Kassmann \([3, 2]\) (see also \([4]\)). There are also recent works of Silvestre (see \([21]\), \([7]\), and the references therein).

We were motivated by our work on Navier-Stokes \([23]\) and the quasi-geostrophic equations \([8]\). In this work, the full regularity of the solutions to the surface quasi-geostrophic equation is shown in the critical case. It was followed by several works on the same subject in the super-critical case (see for instance \([11]\)). Note also that the result was obtained, using completely different techniques by Kiselev, Nazarov and Volberg \([16]\). Our approach led to some progress in the supercritical case (see \([22, 9]\)). It follows pretty much the lines of De Giorgi’s work \([12]\) with a different localization scheme. Nonlinear equations of this form appear extensively in the phase transition literature (see Giacomin, Lebowitz, and Presutti \([13]\)) and more recently on issues of image processing (see Gilboa and Osher \([14]\)).

\section{Presentation of the Results}

Consider the variational integral

\[
V(\theta) = \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \phi(\theta(y) - \theta(x)) K(y - x) dy dx,
\]

for \( \phi : \mathbb{R} \to [0, \infty) \) an even convex function of class \( C^2(\mathbb{R}) \) satisfying the conditions

\begin{align}
\phi(0) &= 0, \\
\Lambda^{-1/2} \leq \phi''(x) \leq \Lambda^{1/2}, \quad x \in \mathbb{R},
\end{align}

for a given constant \( 1 < \Lambda < \infty \).

The kernel \( K : \mathbb{R}^N - \{0\} \to (0, \infty) \) is supposed to satisfy the following conditions for \( 0 < s < 2 \):

\begin{align}
K(-x) &= K(x), \quad \text{for any } x \in \mathbb{R}^N - \{0\}, \\
\frac{\Lambda^{-1/2}}{|x|^{N+s}} \leq K(x) \leq \frac{\Lambda^{1/2}}{|x|^{N+s}}, \quad \text{for any } x \in \mathbb{R}^N - \{0\}.
\end{align}

With the above setting, the corresponding Euler-Lagrange equation for the variational integral \( \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \phi(\theta(y) - \theta(x)) K(y - x) dy dx \) is given by

\[
- \int_{\mathbb{R}^N} \phi'(\theta(y) - \theta(x)) K(y - x) dy = 0.
\]

We are considering in this paper the associated time-dependent problem:

\begin{align}
\partial_t \theta(t, x) - \int_{\mathbb{R}^N} \phi'(\theta(t, y) - \theta(t, x)) K(y - x) dy &= 0.
\end{align}
The main goal of this paper is to address the regularity problem for solutions to the above parabolic-type equation and establish the following main theorem.

**Theorem 2.1.** Consider an even convex function \( \phi \) verifying Hypothesis (2.1) and a kernel \( K \) verifying Hypothesis (2.2) for \( 0 < s < 2 \). Then, for any initial datum \( \theta_0 \in H^1(\mathbb{R}^N) \), there exists a global classical solution to equation (2.3) with \( \theta(0,\cdot) = \theta_0 \) in the sense that \( \|\theta(t,\cdot) - \theta_0\|_{L^2(\mathbb{R}^N)} \to 0 \), as \( t \to 0 \). Moreover \( \nabla_x \theta \in C^\alpha((t_0,\infty) \times \mathbb{R}^N) \) for any \( t_0 > 0 \).

The existence of weak solutions with nonincreasing energy can be constructed following [5]. To address the regularity problem for solutions to equation (2.3), we follow the classical idea of De Giorgi and look at the first derivative \( D\theta \) of a solution \( \theta \) to equation (2.3). First, we use the change of variable \( y = x + z \) to rewrite equation (2.3) as follows:

\[
\partial_t \theta - \int_{\mathbb{R}^N} \phi'(\theta(x + z) - \theta(x))K(z)dz = 0.
\]

Now, we consider \( w = D_e \theta \), the derivative in the direction \( e \) of \( \theta \). Derivating (formally) equation (2.3) in the direction \( e \) we find

\[
\partial_t w - \int_{\mathbb{R}^N} \phi''(\theta(x + z) - \theta(x))\{w(x + z) - w(x)\}K(z)dz = 0.
\]

We then perform a change of variable back to \( y = x + z \) to rewrite the above equation in the following way:

\[
\partial_t w - \int_{\mathbb{R}^N} \phi''(\theta(y) - \theta(x))\{w(y) - w(x)\}K(y - x)dz = 0.
\]

Consider the new kernel \( K(t, x, y) = \phi''(\theta(t, y) - \theta(t, x))K(y - x) \) (with an obvious slight abuse of notation). Since \( \phi \) is an even function, \( \phi'' \) is also an even function, and hence the new kernel \( K(t, x, y) \) is symmetric in \( x \) and \( y \). Moreover, Hypotheses (2.2) and (2.1) imply that \( K(t, x, y) \) satisfies the condition

\[
1_{\{|x-y| \leq \lambda\}} \frac{A^{-1}}{|x-y|^{N+s}} \leq K(t, x, y) \leq \frac{A}{|x-y|^{N+s}}.
\]

As a result, the function \( w = D_e \theta \) satisfies equation (1.1) with the kernel \( K(t, x, y) \) verifying Hypothesis (1.2). Our goal is then to show that the solutions to equation (1.1) are in \( C^\alpha \).

To make the argument rigorous, we will consider the difference quotient \( D^b_e \theta(\cdot) = \frac{1}{h} \{\theta(\cdot + he) - \theta(\cdot)\} \). We use again the version (2.3) of equation (2.3). For any given \( \eta \in C^\infty_c(\mathbb{R}^N) \), we use the difference quotient \( D^b_e \eta \) to test against it, and we get

\[
\int_{\mathbb{R}^N} \partial_t \theta(t, x)D^b_e \eta(x)dx - \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \phi'(\theta(t, x + z) - \theta(t, x))D^b_e \eta(x)dxK(z)dz = 0.
\]

Using discrete integration by parts, \( \int_{\mathbb{R}^N} f(x)D^b_e g(x)dx = -\int_{\mathbb{R}^N} D^b_e f(x)g(x)dx \), we find

\[
\int_{\mathbb{R}^N} \partial_t D^b_e \theta(t, x) \cdot \eta(x)dx = \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} D^b_e \phi'(\theta(\cdot + z) - \theta(\cdot)) \eta(x)dzK(z)dz = 0.
\]

The change of variable \( y = x + z \) leads to

\[
\int_{\mathbb{R}^N} \partial_t D^b_e \theta(t, x) \cdot \eta(x)dx = \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} D^b_e \phi'(\theta(\cdot + y - x) - \theta(\cdot)) \eta(x)K(y - x)dxdy = 0.
\]
Setting \( Y \) if the function \( \phi \) is an even function, so \( \phi' \) is an odd function and consequently
\[
D_h^c[\phi'(\cdot + y - x) - \phi'(\cdot)](x) = -D_h^c[\phi'(\cdot + x - y) - \phi'(\cdot)](y).
\]
Using also the symmetry of \( K \), we can symmetrize the operator to get
\[
(2.5)
\int_{\mathbb{R}^N} \partial_t D_h^c \theta(t, x) \cdot \eta(x) dx
- \frac{1}{2} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} D_h^c[\phi'(\cdot + y - x) - \phi'(\cdot)](x) \cdot [\eta(x) - \eta(y)] K(y - x) dx dy = 0.
\]
Setting \( Y = \theta(y + he) - \theta(x + he) \) and \( X = \theta(y) - \theta(x) \), we get
\[
D_h^c[\phi'(\cdot + y - x) - \phi'(\cdot)](x) = \frac{1}{h} \{ \phi'(\theta(y + he_i) - \theta(x + he_i)) - \phi'(\theta(y) - \theta(x)) \}
= \frac{Y - X}{h} \int_0^1 \phi''(X + s(Y - X)) ds
= [D_h^c \theta(y) - D_h^c \theta(x)] \int_0^1 \phi''(1 - s)[\theta(t, y) - \theta(t, x)]
+ s[\theta(t, y + he) - \theta(t, x + he)] ds.
\]
Hence, \( w = D_h^c \theta \) solves the following equation:
\[
\int_{\mathbb{R}^N} \partial_t w(t, x) \eta(x) dx + \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} K_h(t, x, y)[\eta(x) - \eta(y)] [w(t, x) - w(t, y)] dy dx = 0,
\]
where
\[
K_h(t, x, y) = K(y - x) \int_0^1 \phi''((1 - s)[\theta(t, y) - \theta(t, x)] + s[\theta(t, y + he) - \theta(t, x + he)]) ds.
\]
Note that this new kernel verifies independently on \( h \) the properties (1.2) with the same \( \Lambda \).

Theorem 2.2 is then a consequence of the following theorem.

**Theorem 2.2.** Let \( w \) be a weak solution of (1.1) with a kernel satisfying the properties (1.2). Then for every \( t_0 > 0, w \in C^{\alpha}(t_0, \infty) \times \mathbb{R}^N \). The constant \( \alpha \) and the norm of \( w \) depend only on \( t_0, N, \|w^0\|_{L^2}, \) and \( \Lambda \).

Passing into the limit \( h \to 0 \) gives the result of Theorem 2.1. The rest of the paper is dedicated to the proof of Theorem 2.2.

**Remark.** 1. With the Hypothesis of Theorem 2.1 if in addition \( \theta_0 \in H^{1+s}(\mathbb{R}^N) \), then \( \partial_t \theta(0, \cdot) \in L^2(\mathbb{R}^N) \). Then, applying again Theorem 2.2 on \( w = \partial_t \theta \) shows that \( \partial_t \theta \in C^{\alpha} \) and so \( \theta \in C^{1,\alpha} \) on \( (t_0, \infty) \times \mathbb{R}^N \) for any \( t_0 > 0 \). As an alternative to the incremental quotient, an approximation can be used to justify this formal argument. We provide such an approximation scheme in the appendix.

2. If \( s \leq 1 \), the theorem shows that the solutions are classical. For \( 1 < s < 2 \), if the function \( \phi \in C^{3,\alpha} \), and \( K(z) = |z|^{-(N+s)} \), then \( \nabla \theta \) is \( C^{1,\alpha} \) in time, and \( C^{3,\alpha} \) in \( x \). In this case, the solution is also classical. We provide the proof of this fact in the appendix.
3. The first De Giorgi’s lemma

In this section and the next section, we focus on the differential equation stated in the sense of weak formulation in (1.1). We rewrite it in the following way:

\[
\int_{\mathbb{R}^N} \partial_t w(t, x) \cdot \eta(x) dx + B[w(t, \cdot), \eta] = 0, \forall \eta \in C_0^\infty(\mathbb{R}^N),
\]

(3.1)

\[
B[u, v] = \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} K(t, x, y) [u(x) - u(y)] \cdot [v(x) - v(y)] dxdy,
\]

where the kernel \( K(t, x, y) \) is assumed to satisfy the Hypothesis (1.2). We first introduce the following function \( \psi \):

\[
\psi(x) = (|x|^{\frac{2}{s}} - 1)_+.
\]

For any \( L \geq 0 \), we define

\[
\psi_L(x) = L + \psi(x).
\]

(3.2)

With the above setting, the first De Giorgi’s lemma is as follows.

**Lemma 3.1.** Let \( \Lambda \) be the given constant in condition (1.2). Then, there exists a constant \( \epsilon_0 \in (0, 1) \), depending only on \( N, s, \) and \( \Lambda \), such that for any solution \( w: [-2, 0] \times \mathbb{R}^N \to \mathbb{R} \) to (3.1), the following implication for \( w \) holds true.

If it is verified that

\[
\int_{-2}^0 \int_{\mathbb{R}^N} [w(t, x) - \psi(x)]_+^2 dx dt \leq \epsilon_0,
\]

then we have

\[
w(t, x) \leq \frac{1}{2} + \psi(x)
\]

for \( (t, x) \in [-1, 0] \times \mathbb{R}^N \). (Hence, we have in particular that \( w \leq 1/2 \) on \([-1, 0] \times B(1) \).)

The main difficulty in our approach is due to the nonlocal operator. In [8], a localization of the problem was performed at the cost of adding one more variable to the problem. This was based on the “Dirichlet to Neuman” map. This approach still works for any fractional Laplacian (see Caffarelli and Silvestre [6]). However it breaks down for general kernels as (1.2). Instead, we keep track of the far away behavior of the solution via the function \( \psi \).

**Remark.** All the computations on weak solutions in the proof can be justified by replacing the variable kernel in a neighborhood of the origin by the fractional Laplacian through a smooth cutoff, and smoothing out the kernel outside of this neighborhood (in the same spirit as the approximation scheme provided in the appendix). Then the equation becomes a fractional heat equation with a smooth right-hand side, thus \( C^2 \) in space. This makes the integrals involved uniformly convergent. Once the a priori Hölder continuity is proven, we pass to the limit. Actually, such an approximation can be performed for the original nonlinear problem of Theorem 2.1. We provide a more detailed proof of this in the appendix.

**Proof.** We split the proof into several steps.
First step: Energy estimates. Let \( w : [-2, 0] \times \mathbb{R}^N \to \mathbb{R} \) be a solution to equation (3.1). For \( 0 < L \leq 1 \), we consider the truncated function \([w - \psi_L]_+\), where \( \psi_L \) is defined by (3.3). Then, we take the test function \( \eta \) to be \([w - \psi_L]_+\) in the weak formulation of equation (3.1), which gives

\[
0 = \frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}^N} [w - \psi_L]_+^2 dx + B[w, (w - \psi_L)_+] + \frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}^N} [w - \psi_L]_+^2 dx + B[(w - \psi_L)_+, (w - \psi_L)_+] + B[(w - \psi_L)_-, (w - \psi_L)_+] + B[\psi_L, (w - \psi_L)_+].
\]

By symmetry we end up with

\[
B[(w - \psi_L)_-, (w - \psi_L)_+] = 2 \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} K(t, x, y) (w - \psi_L)_+(x)(w - \psi_L)_{\text{neg}}(y) dx dy,
\]

where we denote \((w - \psi_L)_{\text{neg}} = -(w - \psi_L)_- \geq 0\). In particular,

\[
B[(w - \psi_L)_-, (w - \psi_L)_+] \geq 0.
\]

This “good term” is not fully exploited in this section. It will be used in a crucial way in the next section. The remainder can be written as:

\[
B[\psi_L, (w - \psi_L)_+]
\]

\[
= \frac{1}{2} \int \int_{|x - y| \geq 1} K(t, x, y) [\psi_L(x) - \psi_L(y)] \cdot [(w - \psi_L)_+(x) - (w - \psi_L)_+(y)] dx dy
\]

\[
+ \frac{1}{2} \int \int_{|x - y| < 1} K(t, x, y) \{\psi_L(x) - \psi_L(y)\} \cdot [(w - \psi_L)_+(x) - (w - \psi_L)_+(y)] dx dy.
\]

Using the inequality \(|\psi(x) - \psi(y)| \leq 2 |y - x|^2\), for any \( x \) and \( y \) with \(|y - x| \geq 1\), we get the following estimation of the “far-away” contribution:

\[
\left| \int \int_{|x - y| \geq 1} K(t, x, y) [\psi_L(x) - \psi_L(y)] \cdot [w - \psi_L]_+(x) dx dy \right|
\]

\[
= \left| \int \int_{|x - y| \geq 1} K(t, x, y) \psi(x) - \psi(y) \cdot [w - \psi_L]_+(x) dx dy \right|
\]

\[
\leq \int_{\mathbb{R}^N} \int_{|y - x| \geq 1} \frac{2A}{|x - y|^{N+s}} 2 |y - x|^2 dx \cdot (w - \psi_L)_+(x) dx
\]

\[
= 4A \left| \int_{\mathbb{R}^N} (w - \psi_L)_+(x) dx \right| \leq C \int_{\mathbb{R}^N} (w - \psi_L)_+(x) dx.
\]

By symmetry we end up with

\[
(3.6)
\]

\[
\left| \int \int_{|x - y| \geq 1} K(t, x, y) [\psi_L(x) - \psi_L(y)] \cdot [(w - \psi_L)_+(x) - (w - \psi_L)_+(y)] dx dy \right|
\]

\[
\leq C \int_{\mathbb{R}^N} (w - \psi_L)_+(x) dx.
\]
The other part of the remainder can be controlled in the following way:

\[(3.7) \left| \int \int_{|x-y|<1} K(t, x, y)[\psi_L(x) - \psi_L(y)] \cdot \{(w - \psi_L)_+(x) - (w - \psi_L)_+(y)\}dxdy \right| \leq 2 \int \int_{|x-y|<1} K(t, x, y)\chi_{\{|w-\psi_L|(x)>0\}}|\psi_L(x) - \psi_L(y)| \cdot |(w - \psi_L)_+(x) - (w - \psi_L)_+(y)|dxdy \]

where, in the above inequality, we have used the fact that

\[|(w - \psi_L)_+(x) - (w - \psi_L)_+(y)| \leq \{\chi_{\{|w-\psi_L|(x)>0\}} + \chi_{\{|w-\psi_L|(y)>0\}}\}|(w - \psi_L)_+(x) - (w - \psi_L)_+(y)|,\]

and the symmetry in \(x\) and \(y\).

Now, by Hölder’s inequality, and using the elementary inequality \(|\psi(y) - \psi(x)| < |y - x|\), for any \(x, y\) in \(\mathbb{R}^N\), we can have the following estimation:

\[(3.8) \leq a \cdot \int \int_{|x-y|<1} K(t, x, y)\{(w - \psi_L)_+(x) - (w - \psi_L)_+(y)\}^2dydx + \frac{1}{a} \cdot \int \int_{|x-y|<1} K(t, x, y)|\psi(x) - \psi(y)|^2 \cdot \chi_{\{|w-\psi_L|(x)>0\}}dydx,\]

in which the arbitrary \(a > 0\) will be chosen later. Finally

\[\int \int_{|x-y|<1} K(t, x, y)|\psi(x) - \psi(y)|^2dy \cdot \chi_{\{|w-\psi_L|(x)>0\}}dx \leq \frac{2A}{\int_{|x-y|<1} |x-y|^N+\varepsilon^2} \cdot \chi_{\{|w-\psi_L|(x)>0\}}dx = C_s \int_{\mathbb{R}^N} \chi_{\{|w-\psi_L|(x)>0\}}dx.\]

Pulling this inequality in (3.7) with \(a = 1/2\), and gathering it together with (3.5), (3.6), (3.7), we can rewrite the energy inequality as

\[(3.9) \frac{d}{dt} \int_{\mathbb{R}^N} [w - \psi_L]^2 dx + \frac{1}{2} B[(w - \psi_L)_+, (w - \psi_L)_+] \leq C_{N, A, s} \{ \int_{\mathbb{R}^N} (w - \psi_L)_+(x) dx + \int_{\mathbb{R}^N} \chi_{\{|w-\psi_L|(x)>0\}}dx \},\]

where \(C_{N, A, s}\) is some universal constant depending only on \(N\) and \(\Lambda\) and \(s\). Next, in order to employ the Sobolev embedding theorem, we need to compare
\[ B[(w - \psi_L)_{+}, (w - \psi_L)_{+}] \text{ with } \|(w - \psi_L)_{+}\|_{H^{\frac{s}{2}}(\mathbb{R}^N)}^2 \text{ as follows:} \]
\[
\|(w - \psi_L)_{+}\|_{H^{\frac{s}{2}}(\mathbb{R}^N)}^2 = \int \int_{|x-y| \leq 2} \frac{\{(w - \psi_L)_{+}(x) - (w - \psi_L)_{+}(y)\}^2}{|x-y|^{N+s}} dx dy \\
+ \int \int_{|x-y| > 2} \frac{\{(w - \psi_L)_{+}(x) - (w - \psi_L)_{+}(y)\}^2}{|x-y|^{N+s}} dx dy \\
\leq \Lambda \cdot B[(w - \psi_L)_{+}, (w - \psi_L)_{+}] \\
+ 2 \int \int_{|x-y| > 2} \frac{1}{|x-y|^{N+s}} \{(w - \psi_L)_{+}^2(x) + (w - \psi_L)_{+}^2(y)\} dx dy \\
\leq \Lambda \cdot B[(w - \psi_L)_{+}, (w - \psi_L)_{+}] + C \int_{\mathbb{R}^N} (w - \psi_L)_{+}^2 dx. \]

Hence,
\[
\frac{d}{dt} \int_{\mathbb{R}^N} [w - \psi_L]_{+}^2 dx + \frac{1}{\Lambda} \|(w - \psi_L)_{+}\|_{H^{\frac{s}{2}}(\mathbb{R}^N)}^2 \\
\leq C_{N,A,s} \left\{ \int_{\mathbb{R}^N} (w - \psi_L)_{+} dx + \int_{\mathbb{R}^N} \chi_{\{w - \psi_L > 0\}} dx + \int_{\mathbb{R}^N} (w - \psi_L)_{+}^2 dx \right\}. \tag{3.10} \]

**Second step: Nonlinear recurrence.** From this energy inequality, we establish a nonlinear recurrence relation to the following sequence of truncated energy:

\[
U_k = \sup_{t \in [T_k, 0]} \int_{\mathbb{R}^N} (w - \psi_L)_k^2(t, x) dx + \int_{T_k}^0 \|(w - \psi_L)_k^+(t, \cdot)\|_{H^{\frac{s}{2}}(\mathbb{R}^N)}^2 dt, \]

where, in the above expression, \(T_k = -1 - \frac{1}{2^k}\) and \(L_k = \frac{1}{2} (1 - \frac{1}{2^k})\). Moreover, we will use the abbreviation \(Q_k = [T_k, 0] \times \mathbb{R}^N\).

Now, let us consider two variables \(\sigma, t\) that satisfy \(T_{k-1} \leq \sigma \leq T_k \leq t \leq 0\). By taking the time integral over \([\sigma, t]\) in inequality \((3.10)\), we obtain
\[
\int_{\mathbb{R}^N} [w - \psi_L]_{+}^2(t, x) dx + \int_{\sigma}^{t} \|(w - \psi_L)_k^+(t, \cdot)\|_{H^{\frac{s}{2}}(\mathbb{R}^N)}^2 ds \\
\leq \int_{\mathbb{R}^N} [w - \psi_L]_{+}^2(\sigma, x) dx \\
+ C_{N,A,s} \left\{ \int_{\sigma}^{t} \int_{\mathbb{R}^N} (w - \psi_L)_{+} + \chi_{\{w - \psi_L > 0\}} + (w - \psi_L)_{+}^2 dx ds \right\}. \tag{3.11} \]

Next, by first taking the average over \(\sigma \in [T_{k-1}, T_k]\), and then taking the sup over \(t \in [T_k, 0]\) in the above inequality, we deduce from the above inequality that
\[
(3.11) \quad U_k \leq 2^k (1 + C_{N,s,\Lambda}) \left\{ \int_{Q_{k-1}} (w - \psi_L)_k + \chi_{\{w - \psi_L > 0\}} + (w - \psi_L)_k^2 dx ds \right\}. \]

We now use the very classical Sobolev embedding for fractional spaces (see for instance [15]). The Sobolev embedding theorem \(H^{\frac{s}{2}}(\mathbb{R}^N) \subset L^{2\frac{N}{N-2}}(\mathbb{R}^N)\) and interpolation give
\[
\|(w - \psi_L)_{+}\|_{L^{2\left(1 + \frac{s}{N}\right)}(Q_k)} \leq C_N U_{k}^{\frac{s}{2}}. \]
Using the Tchebychev inequality we get
\[
\int_{Q_{k-1}} (w - \psi_{L_k})^+ \leq \int_{Q_{k-1}} (w - \psi_{L})^+ \chi_{\{w - \psi_{L_{k-1}} > \frac{1}{2s+\pi}\}} \\
\leq (2^{k+1})^{1+\frac{2}{N}} \int_{Q_{k-1}} (w - \psi_{L_{k-1}})^{(1+\frac{2}{N})} \\
\leq (2^{k+1})^{2(1+\frac{2}{N})} C_N^{2(1+\frac{2}{N})} U_{k-1}^{1+\frac{2}{N}} ;
\]
\[
\int_{Q_{k-1}} \chi_{\{w - \psi_{L_k} > 0\}} \leq (2^{k+1})^{2(1+\frac{2}{N})} \int_{Q_{k-1}} (w - \psi_{L_{k-1}})^{(1+\frac{2}{N})} \\
\leq (2^{k+1})^{2(1+\frac{2}{N})} C_N^{2(1+\frac{2}{N})} U_{k-1}^{1+\frac{2}{N}} ;
\]
\[
\int_{Q_{k-1}} (w - \psi_{L_k})^2_+ \leq \int_{Q_{k-1}} (w - \psi_{L})^2_+ \chi_{\{w - \psi_{L_{k-1}} > \frac{1}{2s+\pi}\}} \\
\leq (2^{k+1})^{\frac{2}{N}} \int_{Q_{k-1}} (w - \psi_{L_{k-1}})^{(1+\frac{2}{N})} \\
\leq (2^{k+1})^{\frac{2}{N}} C_N^{2(1+\frac{2}{N})} U_{k-1}^{1+\frac{2}{N}} .
\]

The above three inequalities, together with inequality (3.11), give
\[(3.12) \quad U_k \leq \{C_{N,A,s}\}^k U_{k-1}^{1+\frac{2}{N}} , \quad \forall k \geq 0 ,
\]
for some universal constant \(C_{N,A,s}\) depending only on \(N, s,\) and \(\Lambda\). Due to the nonlinear recurrence relation (3.12) for \(U_k\), we know there exists some sufficiently small universal constant \(\epsilon_0 = \epsilon_0(C_{N,A,s})\), depending only on \(C_{N,A,s}\), such that the following implication is valid.

If \(U_1 \leq \epsilon_0\), then it follows that \(\lim_{k \to \infty} U_k = 0\).

Equation (3.11) with the Tchebychev inequality gives that
\[
U_1 \leq C \int_{-2}^0 \int_{\mathbb{R}^N} |w - \psi|^2 \, dx \, dt,
\]
and \(U_k\) converges to 0 implies that
\[
w \leq \psi + \frac{1}{2}, \quad t \in [-1,0] \times \mathbb{R}^N .
\]

We have the following corollary of Lemma 3.1. It shows that any solutions are indeed bounded for \(t > 0\).

**Corollary 3.2.** Any solution to (1.1) with initial value in \(L^2(\mathbb{R}^N)\) is uniformly bounded on \((t_0, \infty) \times \mathbb{R}^N\) for any \(0 < t_0 < 2\). Indeed:

\[
\sup_{t > t_0, x \in \mathbb{R}^N} |w(t, x)| \leq \frac{\|w^0\|_{L^2}}{2\sqrt{\epsilon_0(t_0/2)^{(N/s+1)/2}}}. 
\]

**Proof.** Fixing \(0 < t_0 < 2\) and \(x_0 \in \mathbb{R}^N\), for any \(t > -2, x \in \mathbb{R}^N\), we consider
\[
\tilde{w}(t, x) = \frac{(t_0/2)^{(N/s+1)/2} \sqrt{\epsilon_0}}{\|w^0\|_{L^2}} w(t_0 + t(t_0/2), x_0 + x(t_0/2)^{1/s}).
\]

The function \(\tilde{w}\) still satisfies equation (3.5) with another kernel verifying Hypothesis (1.2) with the same constant \(\Lambda\). From the decreasing of energy, \(\tilde{w}\) satisfies the assumptions of Lemma 3.1. Hence \(\tilde{w}(0,0) \leq 1/2\). Working with \(-\tilde{w}\) gives that \(-\tilde{w}(0,0) \leq 1/2\) too. \(\square\)
We define \( \tilde{\psi}(x) = (|x|^{s/4} - 1)_+ \). We can rewrite the main lemma of this section in the following way. It will be useful for the next section.

**Corollary 3.3.** Let \( \Lambda \) be the given constant in condition (1.2). Then, there exists a constant \( \delta \in (0, 1) \), depending only on \( N, s, \) and \( \Lambda \), such that for any solution \( w : [-2, 0] \times \mathbb{R}^N \to \mathbb{R} \) to (3.1) satisfying

\[
  w(t, x) \leq 1 + \tilde{\psi}(x) \quad \text{on } [-2, 0] \times \mathbb{R}^N
\]

and

\[
  |\{w > 0\} \cap (-2, 0) \times B_2| \leq \delta,
\]

we have

\[
  w(t, x) \leq \frac{1}{2}, \quad (t, x) \in [-1, 0] \times B_1.
\]

**Proof.** Consider \( R \geq 2^{\frac{s}{2}} \) such that \( 1 + \tilde{\psi}(|y| + 1) \leq \tilde{\psi}(|y|) \), for any \( |y| \geq R \). Note that \( R \) depends only on \( s \). For any \( (t_0, x_0) \in [-1, 0] \times B_1 \) we introduce \( w_R \) defined on \( (-2, 0) \times \mathbb{R}^N \) by

\[
  w_R(s, y) = w(t_0 + \frac{s}{R^s}, x_0 + \frac{y}{R}).
\]

Note that \( w_R \) satisfies equation (3.1) with another kernel \( K_R(t, x, y) \) given by

\[
  K_R(t, x, y) = \frac{1}{R^{N+s}} K(t_0 + \frac{t}{R^s}, x_0 + \frac{x}{R}, x_0 + \frac{y}{R}).
\]

The point here is that the kernel \( K_R(t, x, y) \) satisfies the following constraint:

\[
  \frac{1}{\Lambda} \mathbf{1}_{\{|x-y| \leq 3R\}} \frac{1}{|x-y|^{N+s}} \leq K_R(t, x, y) \leq \frac{1}{\Lambda} \mathbf{1}_{|x-y|^{N+s}}.
\]

which is **stronger** than the one in hypothesis (1.2) (since \( R \geq 2^{\frac{s}{2}} > 1 \)). Hence, the conclusion of Lemma 3.1 can be applied to \( w_R \).

Next, we show that \( \tilde{\psi}(x_0 + \frac{y}{R}) \leq \tilde{\psi}(x_0 + \cdot) \). Since \( |x_0| \leq 1 \), we have that \( \{y \in \mathbb{R}^N : \tilde{\psi}(x_0 + y) = 0\} = \{y \in \mathbb{R}^N : |y + x_0| \leq 1\} \) contains the origin \( O \) of \( \mathbb{R}^N \). This in turn ensures that \( \{y \in \mathbb{R}^N : \tilde{\psi}(x_0 + \frac{y}{R}) = 0\} = \{y \in \mathbb{R}^N : |y + Rx_0| \leq R\} \), which is the dilation of the disc \( \{y \in \mathbb{R}^N : |y + x_0| \leq 1\} \) about the origin \( O \) with a factor of \( R \), must contain the disc \( \{y \in \mathbb{R}^N : |y + x_0| \leq 1\} \) itself. That is, we have (3.13)

\[
  \{y \in \mathbb{R}^N : \tilde{\psi}(x_0 + y) = 0\} = \{y \in \mathbb{R}^N : |y + x_0| \leq 1\} \subset \{y \in \mathbb{R}^N : |y + Rx_0| \leq R\}.
\]

Then, we observe that the graph of \( \tilde{\psi}(x_0 + \cdot) \) is obtained by dilating (flattering) the graph of \( \tilde{\psi}(x_0 + \cdot) \) spatially by a factor of \( R \geq 2^{\frac{s}{2}} \). This indicates that the pointwise slope of \( \tilde{\psi}(x_0 + \frac{y}{R}) \) (outside of the disc \( \{y \in \mathbb{R}^N : |y + Rx_0| \leq R\} \)) should be less than that of \( \tilde{\psi}(x_0 + \cdot) \), which together with the relation in (3.13) implies that the graph of \( \tilde{\psi}(x_0 + \frac{y}{R}) \) must be strictly below the graph of \( \tilde{\psi}(x_0 + \cdot) \). That is, we must have the desired relation \( \tilde{\psi}(x_0 + \frac{y}{R}) \leq \tilde{\psi}(x_0 + \cdot) \).

Since \( \tilde{\psi} \) increases with respect to \( |x| \), for \( |x| > 1 \) we have (3.14)

\[
  w_R(s, y) = w(t_0 + \frac{s}{R^s}, x_0 + \frac{y}{R}) \leq 1 + \tilde{\psi}(x_0 + \frac{y}{R}) \leq 1 + \tilde{\psi}(x_0 + y) \leq 1 + \tilde{\psi}(|y| + 1).
\]
So, from the definition of $R$, for $|y| \geq R$ we have $w_R(s, y) \leq \psi(y)$ (recall that $1 + \bar{\psi}(|y| + 1) \leq \psi(|y|)$, for any $|y| \geq R$). Hence, from the hypothesis we have

$$
\int_{-2}^{0} \int_{\mathbb{R}^N} [w_R(s, y) - \psi(y)]^2_+ = \int_{-2}^{0} \int_{|y| \leq R} [w_R(s, y) - \psi(y)]^2_+ dy \, ds \\
\leq \int_{-2}^{0} \int_{|y| \leq R} [w(t_0 + \frac{s}{R^2}, x_0 + \frac{y}{R})]^2_+ dy \, ds \\
\leq R^{N+s} \int_{t_0}^{t_0} \int_{|x_0| + B(1)} [w(s, y)]^2_+ dy \, ds \\
\leq R^{N+s} \int_{-2}^{0} \int_{B(2)} [w(s, y)]^2_+ dy \, ds \\
\leq R^{N+s} (1 + \bar{\psi}(2))^2 \delta.
$$

(3.15)

Notice that in the above estimation, we have used the facts that $\{x_0\} + B(1) \subset B(2)$ (since $|x_0| \leq 1$) and that $[t_0 - \frac{2}{R^2}, t_0] \subset [-2, 0]$ (since $-1 \leq t_0 \leq 0$ and $R \geq 2^{\frac{1}{2}}$).

So, choosing $\delta = R^{-N-s}(1 + \bar{\psi}(2))^{-2} \varepsilon_0$ gives that $w(t_0, x_0) \leq 1/2$ for $(t_0, x_0) \in (-1, 0) \times B_1$. \qed

4. THE SECOND DE GIORGI’S LEMMA

This section is dedicated to a lemma of local decrease of the oscillation of a solution to equation (3.1). We define the following function:

$$F(x) = \sup(-1, \inf(0, |x|^2 - 9)).$$

(4.1)

Note that $F$ is Lipschitz, compactly supported in $B_3$, and equal to $-1$ in $B_2$. For $\lambda < 1/3$, we define

$$\psi(\lambda)(x) = \begin{cases} 
0 & \text{if } |x| \leq \frac{1}{\lambda^{1/3}} \\
((|x| - 1/\lambda^{1/3})^{4/3} - 1) & \text{if } |x| \geq \frac{1}{\lambda^{1/3}}.
\end{cases}$$

The normalized lemma will involve three consecutive cutoffs:

$$\varphi_0 = 1 + \psi + F,$$
$$\varphi_1 = 1 + \psi + \lambda F,$$
$$\varphi_2 = 1 + \psi + \lambda^2 F.$$

We prove the following lemma:

**Lemma 4.1.** Let $\Lambda$ be the given constant in condition (1.2) and $\delta$ the constant defined in Corollary 3.3. Then, there exists $\mu > 0$, $\gamma > 0$, and $\lambda \in (0, 1)$, depending only on $N$, $\Lambda$, and $s$, such that for any solution $w : [-3, 0] \times \mathbb{R}^N \to \mathbb{R}$ to (3.1) satisfying

$$w(t, x) \leq 1 + \psi(\lambda)(x) \quad \text{on } [-3, 0] \times \mathbb{R}^N,$$

$$|\{w < \varphi_0\} \cap ((-3, -2) \times B_1)| \geq \mu,$$

then we have either

$$|\{w > \varphi_2\} \cap ((-2, 0) \times \mathbb{R}^N)| \leq \delta.$$
or

\[ |\{ \varphi_0 < w < \varphi_2 \} \cap (-3,0) \times \mathbb{R}^N| \geq \gamma. \]

The lemma says that in going from the \( \varphi_0 \) cutoff to the \( \varphi_2 \) cutoff, i.e., from the set \( \{ w > \varphi_0 \} \) to \( \{ w > \varphi_2 \} \) “some mass” is lost; i.e., if \( |\{ w > \varphi_2 \}| \) is not yet subcritical (i.e., \( \leq \delta \)), then

\[ |\{ w > \varphi_2 \}| \leq |\{ w > \varphi_0 \}| - \gamma. \]

**Proof.** In all the proof, we denote by \( C \) constants which depend only on \( s, N \) and \( \Lambda \), but which can change from one line to another. We may fix any \( 0 < \mu < \frac{1}{8} \).

We will fix \( \delta \) smaller than the one in Corollary 3.3 and such that the term \( C\delta \) in (4.8) is smaller than \( \frac{1}{4} \). The task consists now in showing that for \( 0 < \lambda < \frac{1}{3} \) small enough, there exists a \( \gamma > 0 \) for which the lemma holds. The constraints on \( \lambda \) are (4.3), (4.5), (4.7), and (4.9). We split the proof into several steps.

**First step: The energy inequality.** We start again with the energy inequality (3.4), but use better the “good” term

\[ B((w - \varphi)_+,(w - \varphi)_-) = \int_{\mathbb{R}^N} (w - \varphi)_+(x)K(t,x,y)(w - \varphi)_{\text{neg}}(y)) \, dx \, dy \]

that we just neglected before.

We have, for \( \varphi_1 \) the intermediate cutoff (see (3.4)):

\[ \int_{T_1}^{T_2} ((w - \varphi_1)_+)^2 \, dx \bigg|_T + \int_{T_1}^{T_2} B((w - \varphi_1)_+,(w - \varphi_1)_+) \, dt \]

\[ = -\int_{T_1}^{T_2} B((w - \varphi_1)_+,(w - \varphi_1)_+) \, dt - \int_{T_1}^{T_2} B((w - \varphi_1)_+,(w - \varphi_1)_-) \, dt. \]

The remainder term can be controlled in the following way:

\[ B((w - \varphi_1)_+,(w - \varphi_1)_+) \leq \frac{1}{2} B((w - \varphi_1)_+,(w - \varphi_1)_+) \]

\[ + 2 \int \int [\varphi_1(x) - \varphi_1(y)]K(x,y)[\varphi_1(x) - \varphi_1(y)][\chi_{B_3}(x)]. \]

The first term \( \frac{1}{2} B((w - \varphi_1)_+,(w - \varphi_1)_+) \) is absorbed on the left. The second term is smaller than

\[ 4\lambda^2 \int \int [F(x) - F(y)]K(x,y)[F(x) - F(y)] \]

\[ + 4 \int \int [\psi_\lambda(x) - \psi_\lambda(y)]K(x,y)[\psi_\lambda(x) - \psi_\lambda(y)][\chi_{B_3}(x)], \]
which is smaller than $C \lambda^2$. This is obvious for the first term since $F$ is Lipschitz and compactly supported. Since $\psi_3(x) = 0$ for $|x| < 3$, the second term is equal to

$$4 \int \int \psi_3(y)^2 [\chi_{B_3}(x)] K(t, x, y) \, dx \, dy$$

$$\leq 4A|B_3| \int \frac{\left( \frac{|y| - 1/\lambda^4/s}{|y| - 3N/s} \right)^2}{(|y| - 3N/s)^{N/s}} \, dy$$

$$\leq 4A|B_3| \lambda^2 \int \frac{\left( \frac{|z| - 1/\lambda^4/s}{|z| - 3\lambda^4/s} \right)^2}{(|z| - 3N/s)^{N/s}} \, dz$$

$$\leq 4A|B_3| \lambda^2 \int \frac{\left( \frac{|z| - 1/\lambda^4/s}{|z| - 1/3} \right)^2}{(|z| - 1/3)^{N/s}} \, dz$$

$$\leq C \lambda^2,$$

since $\lambda < 1/3$.

This leaves us with the inequality

$$\int (w - \varphi_1)^2 dx \bigg|_{T_1}^{T_2} + \frac{1}{2} \int_{T_1}^{T_2} B((w - \varphi_1)_+, (w - \varphi_1)_+) \, dt$$

$$+ \int_{T_1}^{T_2} \int_{\mathbb{R}^N} (w - \varphi_1)_+ (x) K(x, y) (w - \varphi_1)_{\text{neg}}(y) \, dx \, dy \, dt \leq C \lambda^2 (T_2 - T_1).$$

In particular, since the second and third terms are positive, we get that for $-3 < T_1 < T_2 < 0$:

$$H(t) = \int_{\mathbb{R}^N} (w - \varphi_1)_+^2 (t, x) \, dx$$

satisfies

$$H'(t) \leq C \lambda^2.$$

Next, observe that, since $[w(t, \cdot) - \varphi^+]_+ \leq \lambda \mathbf{1}_{B(3)}$, we have $H(t) = \int_{B(3)} |w(t, \cdot) - \varphi^+]_+^2 \leq \lambda^2 |B(3)|$. Hence the following estimation is valid for any $-3 < T_1 < T_2 < 0$:

$$\int_{T_1}^{T_2} \int_{\mathbb{R}^N} (w - \varphi_1)_+ (x) K(x, y) (w - \varphi_1)_{\text{neg}}(y) \, dx \, dy \, dt$$

$$\leq C \lambda^2 |T_2 - T_1| + |H(T_2) - H(T_1)|$$

$$\leq C \lambda^2.$$

Note that, up to now, those estimates hold for any $0 < \lambda < 1/3$.

**Second step:** An estimate on those time slices where the “good” extra term helps. Remember that $\mu < 1/8$ is fixed from the beginning of the proof.

From our hypothesis

$$|\{w < \varphi_0\} \cap ((-3, -2) \times B_1)| \geq \mu,$$

the set of times $\Sigma$ in $(-3, -2)$ for which $|\{w(\cdot, T) < \varphi_0\} \cap B_1| \geq \mu/4$ has at least measure $\mu/(2|B_1|)$.

We estimate now that except for a few of those time slices, $\int_{\mathbb{R}^N} (w - \varphi_1)_+^2 \, dx$ is very tiny.
\[
C \lambda^2 \geq \int_{-3}^{-2} B((w - \varphi_1)_+, (w - \varphi_1)_-) dt \geq C \Lambda^{-1} \frac{\mu}{8} \int_{\Sigma} \int_{\mathbb{R}^N} (w - \varphi_1)_+ \, dx \, dt
\]

since \((w - \varphi_1)_+ \leq \lambda\).

In other words,
\[
\int_{\Sigma} \int_{\mathbb{R}^N} [(w - \varphi_1)_+](t, x) \, dx \, dt \leq \mathcal{C} \frac{\lambda^3}{\mu} \leq \lambda^{3-1/8}
\]

if \(\lambda\) is small enough such that
\[
\lambda \leq \left(\frac{\mu}{\mathcal{C}}\right)^{8/9}.
\]

In particular, from Tchebychev’s inequality:
\[
\int (w - \varphi_1)_+^2(t, x) \, dx \leq \lambda^{3-1/4}
\]

for all \(t \in \Sigma\), except for a very small subset \(F\) of \(t\)'s of measure smaller than \(\lambda^{1/8}\).

We need it still much smaller than \(\mu \sim |\Sigma|\). Indeed, if \(\lambda\) is small enough such that
\[
\lambda \leq \left(\frac{\mu}{4|B_1|}\right)^{8/9},
\]

then (4.3) holds on a set of \(t\)'s in \([-3, -2]\) of measure greater than \(\mu/(4|B_1|)\).

**Third step: In search of an intermediate set, where \(w\) is between \(\varphi_0\) and \(\varphi_2\).** Let us go now to \((w - \varphi_2)_+\).

Assume that for at least one time \(T_0 > -2\),
\[
| \{(x) \mid (w - \varphi_2)_+(T_0, x) > 0\} | > \delta/2,
\]

i.e., goes over critical for the first lemma, and let’s go backwards in time until we reach a slice of time \(T_1 \in \Sigma\), where
\[
\int_{\mathbb{R}^N} (w - \varphi_1)_+^2(T_1, x) \, dx \leq \lambda^{3-1/4}.
\]

At \(T_0\), for the intermediate cutoff, \(\varphi_1\), we have
\[
\int_{\mathbb{R}^N} (w - \varphi_1)_+^2(T_0, x) \, dx \geq \int (\varphi_1 - \varphi_2)^2 \chi_{\{(w - \varphi_2)_+ > 0\}}
\]

(4.6)
\[
\geq \int (\lambda - \lambda^2)^2 F^2(x) \chi_{\{(w - \varphi_2)_+ > 0\}} \geq C_F \lambda^2 \delta^3,
\]

where the constant \(C_F\) depends only on the fixed function \(F\). Indeed we have \(\lambda < 1/2\), and \(F\) is increasing with respect to \(|x|\) and smaller than \(-C(3 - |x|)\) for \(|x| < 3\) close to 3. Hence, the integral is minimum when all the mass \(\{(w - \varphi_2)_+ > 0\}\) is concentrated on \(3 - C\delta < |x| < 3\).

Now, at \(T_1\),
\[
\int_{\mathbb{R}^N} (w - \varphi_1)_+^2(T_1, x) \, dx \leq \lambda^{3-1/4}.
\]
Thus, for \( \lambda \) small enough such that
\[
\lambda^{1-1/4} \leq C_F \frac{\delta^3}{64},
\]
in going from \( T_0 \) backwards to \( T_1 \), \( H(t) = \int_{\mathbb{R}^N} (w - \varphi_1)^2(t, x) \, dx \) has crossed a range from \( C_F \lambda^2 \frac{\delta^3}{4} \) to \( C_F \lambda^2 \frac{\delta^3}{16} \). Since \( H'(t) \leq C \lambda^2 \), in order to do so the energy \( H(t) \) of the truncation \( [w - \varphi_1]_+(t, \cdot) \) has to pass through a range of times \( D \), of at least length \( \sim \delta^3 \), where
\[
D = \{ t \in (T_1, T_0) : C_F \lambda^2 \frac{\delta^3}{16} < H(t) < C_F \lambda^2 \frac{\delta^3}{4} \}.
\]

We want to show that in this range, we pick up an intermediate set, of nontrivial measure, where \( (w - \varphi_0)_+ > 0 \) and \( (w - \varphi_2)_+ = 0 \), implying that the measure
\[
\mathcal{A}_2 = |\{(w - \varphi_2)_+ > 0 \} \cap \{ t \in (-3, 0) \}| \quad \text{effectively decreases some fixed amount from}
\]
\[
\mathcal{A}_0 = |\{(w - \varphi_0)_+ > 0 \} \cap \{ t \in (-3, 0) \}|.
\]

In these ranges of times \( t \in D \), given the gap between \( \varphi_1 \) and \( \varphi_2 \), we should have
\[
|\{(w - \varphi_2)_+(t, \cdot) > 0 \}| \leq \frac{\delta}{2}.
\]
Indeed, (4.8) can be shown as follows. Assume towards a contradiction that we have \( |\{(w - \varphi_2)_+ (\tau, \cdot) > 0 \}| > \frac{\delta}{2} \) for some \( \tau \in D \). Then, at the time slice \( \tau \), we can go through the same estimation in (4.6) to deduce that we would have \( H(\tau) \geq C_F \frac{\lambda^2}{4} \delta^3 \), which is in contradiction with \( \tau \in D \) by the definition of the set \( D \).

As said at the beginning of the proof, we may consider a \( \delta \) such that \( C\delta < 1/4 \). Moreover, those times of \( D \) for which
\[
|\{(w - \varphi_0)_+ \leq 0 \} \cap B_2| \geq \mu
\]
are in an exceptional subset \( \mathcal{F} \) of very small size. Indeed,
\[
C\lambda^2 \geq \int_{-3}^{0} \int_{\mathbb{R}^N} (w - \varphi_1)_+ K(t, x, y)(w - \varphi_1)_{\text{neg}} \, dx \, dt \\
\geq C\mu \int_{\mathcal{F}} \int_{B_2} (w - \varphi_1)_+ \, dx \, dt \\
\geq C\mu |\mathcal{F}| \frac{(C_F \lambda^2 \delta^3/16)}{\lambda},
\]
where, in the last line of the above estimation, we have used \( H(t) = \int_{\mathbb{R}^N} [w - \varphi_1](t, \cdot) \geq C_F \frac{\lambda^2}{16} \delta^3 \) (for \( t \in \mathcal{F} \in D \)) which is ensured by the definition of the timeset \( D \). Hence
\[
|\mathcal{F}| \leq C \frac{\mu}{\mu \delta^3}.
\]
Therefore, for \( \lambda \) small enough such that
\[
\lambda \leq \mu \delta^3 |D|/(2C),
\]
we have
\[
|\mathcal{F}| \leq \frac{|D|}{2}.
\]
Note that the constraint (4.9) can be expressed depending only on $s$, $N$, $\Lambda$, $\delta$, and, $\mu$, since $|D| < C\delta^3$. For these times in $D$ that are not in $F$, we have:

$$A(t) = |\{\varphi_0 \leq w(t, \cdot) \leq \varphi_2\}| \geq 1/2.$$  

That is,

$$|\{\varphi_0 < w < \varphi_2\} \cap ((-3, 0) \times \mathbb{R}^N)| \geq \int_{-3}^{0} A(t) \, dt \geq \int_{D \setminus F} A(t) \, dt \geq \frac{|D|}{4} \geq C\delta^3. \quad \square$$

5. PROOF OF THE $C^\alpha$ REGULARITY

We are now ready to show the following oscillation lemma. First, for $\lambda$ as in the previous section, we define for any $\varepsilon > 0$,

$$\psi_{\varepsilon, \lambda}(x) = \begin{cases} 0 & \text{if } |x| \leq \frac{1}{\lambda^{4/s}} \\ \frac{((|x| - 1/\lambda^{4/s})^+ - 1)}{\lambda^{2k_0}} & \text{if } |x| \geq \frac{1}{\lambda^{4/s}}. \end{cases}$$

Lemma 5.1. There exists $\varepsilon > 0$ and $\lambda^*$ such that for any solution to (3.1) in $[-3, 0] \times \mathbb{R}^N$ such that

$$-1 - \psi_{\varepsilon, \lambda} \leq w \leq 1 + \psi_{\varepsilon, \lambda},$$

we have

$$\sup_{[-1,0] \times B_1} w - \inf_{[-1,0] \times B_1} w \leq 2 - \lambda^*.$$  

Proof. We may assume that

$$|\{w < \varphi_0\} \cap ((-3, -2) \times B_1)| > \mu.$$  

Otherwise this is verified by $-w$, and we may work on this function.

Consider $k_0 = |(-3, 0) \times B_3|/\gamma$. Then we fix $\varepsilon$ small enough such that

$$\frac{(|x|^+ - 1)}{\lambda^{2k_0}} \leq (|x|^{s/4} - 1)^+$$

for all $x$. We may take $\varepsilon = (s/4)\lambda^{2k_0}$ for instance. For $k \leq k_0$, we consider the sequence

$$w_{k+1} = \frac{1}{\lambda^2}(w_k - (1 - \lambda^2)), \quad w_0 = w.$$  

By induction, we have that

$$(w_k)(t, x) \leq 1 + \frac{1}{\lambda^{2k}} \psi_{\varepsilon, \lambda}(x), \quad t \in (-3, 0), x \in \mathbb{R}^N.$$  

So, for $k \leq k_0$ we have $w_k \leq 1 + \psi_{\varepsilon, \lambda}$.

By construction $|\{w_k < \varphi_0\} \cap (-3, -2) \times B_1|$ is increasing, thus greater than $\mu$ for any $k$. Hence, we can apply Lemma 4.1 on $w_k$. As long as $|\{w_k > \varphi_2\} \cap ((-2, 0) \times \mathbb{R}^N)| \geq \delta$, we have

$$|\{w_{k+1} > \varphi_0\}| = |\{w_{k+1} > \varphi_2\}| + |\{\varphi_0 < w_{k+1} < \varphi_2\}|$$

and

$$|\{w_{k+1} > \varphi_2\}| \leq |\{w_{k+1} > \varphi_0\}| - \gamma \leq |\{w_k > \varphi_2\}| - \gamma \leq |(-3, 0) \times B_3| - k\gamma.$$
This cannot be true up to \( k_0 \). So there exists \( k \leq k_0 \) such that
\[
|\{w_k > \varphi_2\} \cap ((-2, 0) \times \mathbb{R}^N)| \leq \delta.
\]
We can then apply the first De Giorgi lemma on \( w_{k+1} \). Indeed,
\[
w_{k+1} \leq 1 + \psi_\lambda \leq 1 + \psi_1 \quad \text{on } (-3, 0) \times \mathbb{R}^N,
\]
and
\[
|\{w_{k+1} > 0\} \cap ((-2, 0) \times B_2)| \leq |\{w_k > \varphi_2\} \cap ((-2, 0) \times B_2)|
\]
\[
\leq |\{w_k > \varphi_2\} \cap ((-2, 0) \times \mathbb{R}^N)| \leq \delta.
\]
Hence, from Corollary 3.3 we have
\[
w_{k+1} \leq 1/2 \quad \text{on } (-1, 0) \times B_1.
\]
This gives the result with
\[
\lambda^* = \frac{\lambda^{2k_0}}{2}.
\]

The \( C^\alpha \) regularity follows in a standard way. More precisely, by using Lemma 5.1 we can now complete the proof of Theorem 2.2 as follows.

**Proof.** For any \((t_0, x_0) \in (0, \infty) \times \mathbb{R}^N\), consider first \( K_0 = \inf(1, t_0/4)^{1/s} \) and
\[
w_0(t, x) = w(t_0 + K_0 t, x_0 + K_0 x).
\]
This function still satisfies an equation of the type of (3.1) in \((-4, 0) \times \mathbb{R}^N\) with a kernel \( K \) satisfying (1.2). From Corollary 3.2 it is bounded on \((-3, 0) \times \mathbb{R}^N\).

Consider \( K < 1 \) such that
\[
\frac{1}{1 - (\lambda^*/2)} \psi_{\lambda, \varepsilon}(Kx) \leq \psi_{\lambda, \varepsilon}(x), \quad \text{for } |x| \geq 1/K.
\]
The coefficient \( K \) depends only on \( \lambda, \lambda^* \) and \( \varepsilon \). Then we define by induction:
\[
w_1(t, x) = \frac{w_0(t, x)}{\|w_0\|_{L^\infty}}, \quad (t, x) \in (-3, 0) \times \mathbb{R}^N,
\]
\[
w_{k+1}(t, x) = \frac{1}{1 - \lambda^*/4} (w_k(K^* t, Kx) - \bar{w}_k), \quad (t, x) \in (-3, 0) \times \mathbb{R}^N,
\]
where
\[
\bar{w}_k = \frac{1}{|B_1|} \int_{-1}^0 \int_{B_1} w_k(t, x) \, dx \, dt.
\]
By construction, \( w_k \) satisfies the hypothesis of Lemma 5.1 for any \( k \). Hence:
\[
\sup_{(t_0 + (-K^* t_0), 0) \times (x_0 + B_{K^*} k)} w - \inf_{(t_0 + (-K^* t_0), 0) \times (x_0 + B_{K^*} k)} w \leq C(1 - \lambda^*/4)^k.
\]
So, \( w \) is \( C^\alpha \) with
\[
\alpha = \frac{\ln(1 - \lambda^*/4)}{\ln(K^*)}.
\]
□
APPENDIX A. CLASSICAL SOLUTIONS IF $\phi$ IS SMOOTH ENOUGH

We consider the case $1 < s < 2$. We show in this appendix that if $\phi \in C^{3,s-1+\varepsilon}$ and $K(z) = |z|^{-(N+s)}$, then the solution to (2.3) is, indeed, a classical solution.

We linearize the equation of the first derivatives $w = \partial_x \theta$ which is $C^{\alpha}$. Since $\phi$ is even, $\phi''(0) = 0$, and

$$\partial_t w - \phi''(0) \Delta^{1/2} w = \int_{\mathbb{R}^N} R(t, x, y) K(x-y)(w(y) - w(x)) \, dy = F(t, x).$$

Since $w = \nabla \theta$ is $C^{\alpha}$, the right-hand side term can be shown to be also $C^{\alpha/2}$ in space and time. To show that it is $C^{\alpha/2}$ in $x$, we write

$$F(x + z) - F(x)$$

$$= \int \frac{R(x, y + x)}{|y|^{N+s}} (w(x + z + y) - w(x + z) - w(x + y) + w(x)) \, dy$$

$$+ \int |R(x + z, y + x + z) - R(x, y + x)| \frac{w(x + z + y) - w(x + z)}{|y|^{N+s}} \, dy.$$

From (A.1), the first term can be controlled by $C|z|^\alpha$. For the second term, note that $\phi''$ is $C^1$ and $\theta$ is $C^1$ in $x$, so, using (A.1) again, we find

$$[R(x + z, y + x + z) - R(x, y + x)]$$

$$\leq C |R(x + z, y + x + z) - R(x, y + x)|^{\alpha/2}$$

$$\times (|R(x + z, y + x + z)|^{1-\alpha/2} + |R(x, y + x)|^{1-\alpha/2})$$

$$\leq C |z|^\alpha \inf(|y|^{-(1-\alpha/2)(s+\varepsilon)}, 1).$$

This second term can be controlled by $C|z|^{\alpha/2}$.

Then, from (A.2), the $C^{\alpha/2}$ regularity of $F$, and using Madych-Rivi`ere [19], $\partial_t w \in C^{\alpha/2}$ and $w$ is $C^{s,\alpha/2}$ in $x$. Thus all the terms in the equation are classical.

APPENDIX B. THE APPROXIMATION SCHEME FOR THE NONLINEAR PROBLEM

We can approach (a priori) weak solutions to the nonlinear problem of Theorem 2.1 by classical solutions of approximated equations using the following approximation scheme. We consider $K(z) = |z|^{-(N+s)}$. For $\varepsilon > 0$, we replace the function $\phi$ by an approximated function $\phi_{\varepsilon}$ such that

$$\phi_{\varepsilon}(X) = |X|^2, \quad \text{for } |X| < \varepsilon,$$

and we consider $\theta_{\varepsilon}$, the solution to

$$\partial_t \theta_{\varepsilon}(t, x) - \int_{\mathbb{R}^N} \phi_{\varepsilon}'(\theta_{\varepsilon}(t, y) - \theta_{\varepsilon}(t, x)) K(y-x) \, dy = 0.$$

Note, first, that the weak solutions $\theta_{\varepsilon}$ are uniformly Lipschitz. This estimate does not need approximation (it is the boundedness of $w_{\varepsilon} = \partial_x \theta_{\varepsilon}$). It can be done using the incremental quotients (as in the second-order case). So the result is also true for $\theta$. The corresponding operator for $\nabla_x \theta_{\varepsilon}$ is

$$\partial_x w_{\varepsilon} = \int_{\mathbb{R}^N} \phi_{\varepsilon}''(\theta_{\varepsilon}(x) - \theta_{\varepsilon}(y))(w_{\varepsilon}(y) - w_{\varepsilon}(x)) K(x-y) \, dy.$$
But \( \phi''_\varepsilon = 1 \) for \(|\theta_\varepsilon(x) - \theta_\varepsilon(y)| < \varepsilon \). Since \( \theta_\varepsilon \) is Lipschitz, \( \phi''_\varepsilon = 1 \) for \(|x - y| \leq C\varepsilon \). Thus, the quantity

\[
\int_{\mathbb{R}^N} [\phi''_\varepsilon(\theta_\varepsilon(x) - \theta_\varepsilon(y)) - 1] [w_\varepsilon(x) - w_\varepsilon(y)] K(x - y) \, dy
\]

is bounded. It is equal to

\[
\partial_t w_\varepsilon - \Delta^{s/2} w_\varepsilon.
\]

By classical theory, this implies that \( w_\varepsilon \) is \( C^\alpha \) in space and time. In turn, it implies that \( \text{(B.1)} \) is \( C^\alpha \), and so \( \text{(B.2)} \) is \( C^\beta \) (see Appendix A). This implies that \( w_\varepsilon \) has enough regularity to make the equation classical.

We can now show the convergence of \( \theta_\varepsilon \) to \( \theta \). We can rewrite the equation on \( \theta_\varepsilon \) as

\[
\partial_t \theta_\varepsilon = \int_{\mathbb{R}^N} \phi'(\theta_\varepsilon(x) - \theta_\varepsilon(y)) K(x - y) \, dy
\]

\[
+ \int_{\mathbb{R}^N} [\phi'_\varepsilon(\theta_\varepsilon(x) - \theta_\varepsilon(y)) - \phi'(\theta_\varepsilon(x) - \theta_\varepsilon(y))] K(x - y) \, dy.
\]

Hence, \( \theta - \theta_\varepsilon \) verifies the linearized equation

\[
\partial_t (\theta - \theta_\varepsilon) = \int_{\mathbb{R}^N} K(t, x, y) [(\theta - \theta_\varepsilon)(y) - (\theta - \theta_\varepsilon)(x)] \, dy
\]

\[
+ \int_{\mathbb{R}^N} [\phi'_\varepsilon(\theta_\varepsilon(x) - \theta_\varepsilon(y)) - \phi'(\theta_\varepsilon(x) - \theta_\varepsilon(y))] K(x - y) \, dy,
\]

where

\[
K(t, x, y) = \int_0^1 \phi''((1 - s)(\theta(x) - \theta(y)) + s(\theta_\varepsilon(x) - \theta_\varepsilon(y))) \, ds.
\]

The quantity \( [\phi'_\varepsilon(\theta_\varepsilon(x) - \theta_\varepsilon(y)) - \phi'(\theta_\varepsilon(x) - \theta_\varepsilon(y))] \) is bounded by \( C \min(\varepsilon, |\theta_\varepsilon(x) - \theta_\varepsilon(y)|) \). We multiply by \( (\theta - \theta_\varepsilon) \) and integrate (note that \( \theta - \theta_\varepsilon = 0 \) at \( t = 0 \)) to obtain

\[
\sup_T \int_{\mathbb{R}^N} (\theta - \theta_\varepsilon)^2(t) \, dx + \frac{1}{2} \int_0^T \int_{\mathbb{R}^{2N}} [(\theta - \theta_\varepsilon)(x) - (\theta - \theta_\varepsilon)(y)]^2 K(t, x, y) \, dx \, dy \, dt
\]

\[
\leq \int_0^T \int_{\mathbb{R}^N} [(\theta - \theta_\varepsilon)(x) - (\theta - \theta_\varepsilon)(y)][\phi'_\varepsilon(\theta_\varepsilon(x) - \theta_\varepsilon(y))]
\]

\[
- \phi'(\theta_\varepsilon(x) - \theta_\varepsilon(y))] K(x - y) \, dy \, dt.
\]

The quantity \( \phi'' \) is bounded from below, so \( K(x - y) \leq CK(t, x, y) \). By Cauchy we get

\[
\sup_T \int_{\mathbb{R}^N} (\theta - \theta_\varepsilon)^2(t) \, dx + \frac{1}{2} \int_0^T \int_{\mathbb{R}^{2N}} [(\theta - \theta_\varepsilon)(x) - (\theta - \theta_\varepsilon)(y)]^2 K(t, x, y) \, dx \, dy \, dt
\]

\[
\leq C \int_{\mathbb{R}^N} [\min(\varepsilon, \theta_\varepsilon(x) - \theta_\varepsilon(y))]^2 K(x - y) \, dy
\]

\[
\leq C\varepsilon^{2-s-\delta},
\]

for any \( \delta > 0 \). Hence \( \theta_\varepsilon \) converges to \( \theta \) in \( L^\infty(0, T; L^2(\mathbb{R}^N)) \cap L^2(0, T; \dot{H}^{s/2}(\mathbb{R}^N)) \).
References


