UC Berkeley
UC Berkeley Previously Published Works

Title
Well-posedness and dispersive decay of small data solutions for the Benjamin-Ono equation

Permalink

Journal

ISSN

Authors
Mihaela IFRIM
Daniel TATARU

Publication Date

DOI

WELL-POSEDNESS AND DISPERSIVE DECAY OF SMALL DATA SOLUTIONS FOR THE BENJAMIN-ONO EQUATION

MIHAELA IFRIM AND DANIEL TATARU

ABSTRACT. This article represents a first step toward understanding the long time dynamics of solutions for the Benjamin-Ono equation. While this problem is known to be both completely integrable and globally well-posed in L^2, much less seems to be known concerning its long time dynamics. Here, we prove that for small localized data the solutions have (nearly) dispersive dynamics almost globally in time. An additional objective is to revisit the L^2 theory for the Benjamin-Ono equation and provide a simpler, self-contained approach.

RÉSUMÉ. Cet article représente une première étape vers la compréhension du comportement en temps long pour l'équation de Benjamin-Ono. Tandis que ce problème est à la fois complètement intégrable et globalement bien posé en L^2, beaucoup moins semble être connu en ce qui concerne son comportement en temps long. Nous montrons ici que pour de données petites et localisée, les solutions ont une dynamique dispersive presque globalement en temps. Un autre objectif est de revoir la thórie L^2 pour Benjamin-Ono et de fournir une approche plus simple et autonome.

CONTENTS

1. Introduction

In this article we consider the Benjamin-Ono equation

$$(1.1) \qquad (\partial_t + H\partial_x^2)\phi = \frac{1}{2}\partial_x(\phi^2), \qquad \phi(0) = \phi_0,$$

where ϕ is a real valued function $\phi : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$. H denotes the Hilbert transform on the real line; we use the convention that its symbol is

$$H(\xi) = -i\,\mathrm{sgn}\,\xi$$

as in Tao [36] and opposite to Kenig-Martel [26]. Thus, dispersive waves travel to the right and solitons to the left.

The Benjamin-Ono equation is a model for the propagation of one dimensional internal waves (see [4]). Among others, it describes the physical phenomena of wave propagation at the interface of layers of fluids with different densities (see Benjamin [4] and Ono [31]). It also belongs to a larger class of equation modeling this type of phenomena, some of which are certainly more physically relevant than others.

Equation (1.1) is known to be completely integrable. In particular it has an associated Lax pair, an inverse scattering transform and an infinite hierarchy of conservation laws. For further information in this direction we refer the reader to [24] and references therein. We list only some of these conserved energies, which hold for smooth solutions (for example $H_x^3(\mathbb{R})$). Integrating by parts, one sees that this problem has conserved mass,

$$E_0 = \int \phi^2\,dx,$$

momentum

$$E_1 = \int \phi H\phi_x - \frac{1}{3}\phi^3\,dx,$$

as well as energy

$$E_2 = \int \phi_x^2 - \frac{3}{4}\phi^2 H\phi_x + \frac{1}{8}\phi^4\,dx.$$

More generally, at each nonnegative integer k we similarly have a conserved energy E_k corresponding at leading order to the $\dot{H}^{\frac{k}{2}}$ norm of ϕ.

This is closely related to the Hamiltonian structure of the equation, which uses the symplectic form

$$\omega(\psi_1, \psi_2) = \int \psi_1 \partial_x \psi_2\,dx$$

with associated map $J = \partial_x$. Then the Benjamin-Ono equation is generated by the Hamiltonian E_1 and symplectic form ω. E_0 generates the group of translations. All higher order conserved energies can be viewed in turn as Hamiltonians for a sequence of commuting flows, which are known as the Benjamin-Ono hierarchy of equations.

The Benjamin-Ono equation is a dispersive equation, i.e. the group velocity of waves depends on the frequency. Precisely, the dispersion relation for the linear part is given by

$$\omega(\xi) = -\xi|\xi|,$$

and the group velocity for waves of frequency ξ is $v = 2|\xi|$. Here we are considering real solutions, so the positive and negative frequencies are matched. However, if one were to restrict the linear Benjamin-Ono waves to either positive or negative frequencies then we

obtain a linear Schrödinger equation with a choice of signs. Thus one expects that many features arising in the study of nonlinear Schrödinger equations will also appear in the study of Benjamin-Ono.

Last but not least, when working with Benjamin-Ono one has to take into account its quasilinear character. A cursory examination of the equation might lead one to the conclusion that it is in effect semilinear. It is only a deeper analysis see [30] [28] which reveals the fact that the derivative in the nonlinearity is strong enough to insure that the nonlinearity is non-perturbative, and that only countinuous dependence on the initial data may hold, even at high regularity.

Considering local and global well-posedness results in Sobolev spaces H^s, a natural threshold is given by the fact that the Benjamin-Ono equation has a scale invariance,

$$(1.2) \qquad\qquad \phi(t, x) \to \lambda \phi(\lambda^2 t, \lambda x),$$

and the scale invariant Sobolev space associated to this scaling is $\dot{H}^{-\frac{1}{2}}$.

There have been many developments in the well-posedness theory for the Benjamin-Ono equations, see: $[5, 22, 23, 25, 28, 30, 32, 33, 36]$. Well-posedness in weighted Sobolev spaces was considered in [8] and [9], while soliton stability was studied in [10, 26]. These is also closely related work on an extended class of equations, called generalized Benjamin-Ono equations, for which we refer the reader to [13], [14] and references therein. More extensive discussion of Benjamin-Ono and related fluid models can be found in the survey papers [1] and [27].

Presently, for the Cauchy problem at low regularity, the existence and uniqueness result at the level of $H^s(\mathbf{R})$ data is now known for the Sobolev index $s \geq 0$. Well-posedness in the range $-\frac{1}{2} \leq s < 0$ appears to be an open question. We now review some of the key thresholds in this analysis.

The H^3 well posedness result was obtained by Saut in [33], using energy estimates. For convenience we use his result as a starting point for our work, which is why we recall it here:

Theorem 1. *The Benjamin-Ono equation is globally well-posed in H^3.*

The H^1 threshold is another important one, and it was reached by Tao [36]; his article is highly relevant to the present work, and it is where the idea of renormalization is first used in the study of Benjamin-Ono equation.

The L^2 threshold was first reached by Ionescu and Kenig [23], essentially by implementing Tao's renormalization argument in the context of a much more involved and more delicate functional setting, inspired in part from the work of the second author [29] and of Tao [36] on wave maps. This is imposed by the fact that the derivative in the nonlinearity is borderline from the perspective of bilinear estimates, i.e. there is no room for high frequency losses. An attempt to simplify the L^2 theory was later made by Molinet-Pilod [29]; however, their approach still involves a rather complicated functional structure, involving not only $X^{s,b}$ spaces but additional weighted mixed norms in frequency.

Our first goal here is to revisit the L^2 theory for the Benjamin-Ono equation, and (re)prove the following theorem:

Theorem 2. *The Benjamin-Ono equation is globally well-posed in L^2.*

Since the L^2 norm of the solutions is conserved, this is in effect a local in time result, trivially propagated in time by the conservation of mass. In particular it says little about the long time properties of the flow, which will be our primary target here.

Given the quasilinear nature of the Benjamin-Ono equation, here it is important to specify the meaning of well-posedness. This is summarized in the following properties:

(i) Existence of regular solutions: For each initial data $\phi_0 \in H^3$ there exists a unique global solution $\phi \in C(\mathbb{R}; H^3)$.

(ii) Existence and uniqueness of rough solutions: For each initial data $\phi_0 \in L^2$ there exists a solution $\phi \in C(\mathbb{R}; L^2)$, which is the unique limit of regular solutions.

(iii) Continuous dependence : The data to solution map $\phi_0 \to \phi$ is continuous from L^2 into $C(L^2)$, locally in time.

(iv) Higher regularity: The data to solution map $\phi_0 \to \phi$ is continuous from H^s into $C(H^s)$, locally in time, for each $s > 0$.

(v) Weak Lipschitz dependence: The flow map for L^2 solutions is locally Lipschitz in the $H^{-\frac{1}{2}}$ topology.

The weak Lipschitz dependence part appears to be a new result, even though certain estimates for differences of solutions are part of the prior proofs in [23] and [29].

Our approach to this result is based on the idea of normal forms, introduced by Shatah [34] [14]in the dispersive realm in the context of studying the long time behavior of dispersive pde's. Here we turn it around and consider it in the context of studying local well-posedness. In doing this, the chief difficulty we face is that the standard normal form method does not readily apply for quasilinear equations.

One very robust adaptation of the normal form method to quasilinear equations, called "quasilinear modified energy method" was introduced earlier by the authors and collaborators in [15], and then further developed in the water wave context first in [17] and later in [11, 18–20]. There the idea is to modify the energies, rather than apply a normal form transform to the equations; this method is then successfully used in the study of long time behavior of solutions. Alazard and Delort [2,3] have also developed another way of constructing the same type of almost conserved energies by using a partial normal form transformation to symmetrize the equation, effectively diagonalizing the leading part of the energy.

The present paper provides a different quasilinear adaptation of the normal form method. Here we do transform the equation, but not with a direct quadratic normal form correction (which would not work). Instead we split the quadratic nonlinearity in two parts, a milder part and a paradifferential part[1]. Then we construct our normal form correction in two steps: first a direct quadratic correction for the milder part, and then a renormalization type correction for the paradifferential part. For the second step we use a paradifferential version of Tao's renormalization argument [36].

Compared with the prior proofs of L^2 well-posedness in [23] and [29], our functional setting is extremely simple, using only Strichartz norms and bilinear L^2 bounds. Furthermore, the bilinear L^2 estimates are proved in full strength but used only in a very mild way, in order to remove certain logarithmic divergences which would otherwise arise. The (minor) price to pay is that the argument is now phrased as a bootstrap argument, same as in [36]. However this is quite natural in a quasilinear context.

One additional natural goal in this problem is the enhanced uniqueness question, namely to provide relaxed conditions which must be imposed on an arbitrary L^2 solution in order to

---

1. This splitting is of course not a new idea, and it has been used for some time in the study of quasilinear problems

compel it to agree with the L^2 solution provided in the theorem. This problem has received substantial attention in the literature but is beyond the scope of the present paper. Instead we refer the reader to the most up to date results in [29].

We now arrive at the primary goal of this paper. The question we consider concerns the long time behavior of Benjamin-Ono solutions with small localized data. Precisely, we are asking what is the optimal time-scale up to which the solutions have linear dispersive decay. Our main result is likely optimal, and asserts that this holds almost globally in time:

Theorem 3. *Assume that the initial data ϕ_0 for* (1.1) *satisfies*

$$(1.3) \qquad \|\phi_0\|_{L^2} + \|x\phi_0\|_{L^2} \le \epsilon \ll 1.$$

Then the solution ϕ satisfies the dispersive decay bounds

$$(1.4) \qquad |\phi(t,x)| + |H\phi(t,x)| \lesssim \epsilon |t|^{-\frac{1}{2}} \langle x_- t^{-\frac{1}{2}} \rangle^{-\frac{1}{2}}$$

up to time

$$|t| \lesssim T_\epsilon := e^{\frac{c}{\epsilon}}, \qquad c \ll 1.$$

The novelty in our result is that the solution exhibits dispersive decay. We also remark that better decay holds in the region $x < 0$. This is because of the dispersion relation, which sends all the propagating waves to the right.

A key ingredient of the proof of our result is a seemingly new conservation law for the Benjamin Ono equation, which is akin to a normal form associated to a corresponding linear conservation law.

This result closely resembles the authors' recent work in [21] (see also further references therein) on the cubic nonlinear Schrödinger problem (NLS)

$$(1.5) \qquad iu_t - u_{xx} = \pm u^3, \qquad u(0) = u_0,$$

with the same assumptions on the initial data. However, our result here is only almost global, unlike the global NLS result in [21].

To understand why the cubic NLS problem serves as a good comparison, we first note that both the Benjamin-Ono equation and the cubic NLS problem have $\dot{H}^{-\frac{1}{2}}$ scaling. Further, for a restricted frequency range of nonlinear interactions in the Benjamin-Ono equation, away from zero frequency, a normal form transformation turns the quadratic BO nonlinearity into a cubic NLS type problem for which the methods of [21] apply. Thus, one might naively expect a similar global result. However, it appears that the Benjamin-Ono equation exhibits more complicated long range dynamics near frequency zero, which have yet to be completely understood.

One way to heuristically explain these differences is provided by the the inverse scattering point of view. While the small data cubic focusing NLS problem has no solitons, on the other hand in the Benjamin-Ono case the problem could have solitons for arbitrarily small localized data. As our result can only hold in a non-soliton regime, the interesting question then becomes what is the lowest time-scale where solitons can emerge from small localized data. A direct computation [2] shows that this is indeed the almost global time scale, thus justifying our result.

---

2. This is based on the inverse scattering theory for the Benjamin-Ono equation, and will be described in subsequent work.

We further observe that our result opens the way for the next natural step, which is to understand the global in time behavior of solutions, where in the small data case one expects a dichotomy between dispersive solutions and dispersive solutions plus one soliton:

Conjecture 4 (Soliton resolution). *Any global Benjamin-Ono solution which has small data as in (1.3) must either be dispersive, or it must resolve into a soliton and a dispersive part.*

2. DEFINITIONS AND REVIEW OF NOTATIONS

The big O notation: We use the notation $A \lesssim B$ or $A = O(B)$ to denote the estimate that $|A| \leq CB$, where C is a universal constant which will not depend on ϵ. If X is a Banach space, we use $O_X(B)$ to denote any element in X with norm $O(B)$; explicitly we say $u = O_X(B)$ if $\|u\|_X \leq CB$. We use $\langle x \rangle$ to denote the quantity $\langle x \rangle := (1 + |x|^2)^{1/2}$.

Littlewood-Paley decomposition: One important tool in dealing with dispersive equations is the Littlewood-Paley decomposition. We recall its definition and also its usefulness in the next paragraph. We begin with the Riesz decomposition

$$1 = P_- + P_+,$$

where $P_\pm$ are the Fourier projections to $\pm[0, \infty)$; from

$$\widehat{Hf}(\xi) = -i\,\mathrm{sgn}(\xi)\,\hat{f}(\xi),$$

we observe that

(2.1) $$iH = P_+ - P_-.$$

Let ψ be a bump function adapted to $[-2, 2]$ and equal to 1 on $[-1, 1]$. We define the Littlewood-Paley operators P_k and $P_{\leq k} = P_{<k+1}$ for $k \geq 0$ by defining

$$\widehat{P_{\leq k}f}(\xi) := \psi(\xi/2^k)\hat{f}(\xi)$$

for all $k \geq 0$, and $P_k := P_{\leq k} - P_{\leq k-1}$ (with the convention $P_{\leq -1} = 0$). Note that all the operators P_k, $P_{\leq k}$ are bounded on all translation-invariant Banach spaces, thanks to Minkowski's inequality. We define $P_{>k} := P_{\geq k-1} := 1 - P_{\leq k}$.

For simplicity, and because $P_\pm$ commutes with the Littlewood-Paley projections P_k, $P_{<k}$, we will introduce the following notation $P_k^\pm := P_k P_\pm$, respectively $P_{<k}^\pm := P_\pm P_{<k}$. In the same spirit, we introduce the notations $\phi_k^+ := P_k^+ \phi$, and $\phi_k^- := P_k^- \phi$, respectively.

Given the projectors P_k, we also introduce additional projectors $\tilde{P}_k$ with slightly enlarged support (say by 2^{k-4}) and symbol equal to 1 in the support of P_k.

From Plancherel's theorem we have the bound

(2.2) $$\|f\|_{H_x^s} \approx \left(\sum_{k=0}^{\infty} \|P_k f\|_{H_x^s}^2\right)^{1/2} \approx \left(\sum_{k=0}^{\infty} 2^{ks}\|P_k f\|_{L_x^2}^2\right)^{1/2}$$

for any $s \in \mathbb{R}$.

Multi-linear expressions. We shall now make use of a convenient notation for describing multi-linear expressions of product type, as in [37]. By $L(\phi_1, \cdots, \phi_n)$ we denote a translation invariant expression of the form

$$L(\phi_1, \cdots, \phi_n)(x) = \int K(y)\phi_1(x + y_1) \cdots \phi_n(x + y_n)\, dy,$$

where $K \in L^1$. More generally, one can replace $K dy$ by any bounded measure. By L_k we denote such multilinear expressions whose output is localized at frequency 2^k.

This L notation is extremely handy for expressions such as the ones we encounter here; for example we can re-express the normal form (4.12) in a simpler way as shown in Section 4.2. It also behaves well with respect to reiteration, e.g.

$$L(L(u, v), w) = L(u, v, w).$$

Multilinear L type expressions can easily be estimated in terms of linear bounds for their entries. For instance we have

$$\|L(u_1, u_2)\|_{L^r} \lesssim \|u_1\|_{L^{p_1}} \|u_1\|_{L^{p_2}}, \qquad \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{r}.$$

A slightly more involved situation arises in this article when we seek to use bilinear bounds in estimates for an L form. There we need to account for the effect of uncorrelated translations, which are allowed given the integral bound on the kernel of L. To account for that we use the translation group $\{T_y\}_{y \in \mathbb{R}}$,

$$(T_y u)(x) = u(x + y),$$

and estimate, say, a trilinear form as follows:

$$\|L(u_1, u_2, u_3)\|_{L^r} \lesssim \|u_1\|_{L^{p_1}} \sup_{y \in \mathbb{R}} \|u_2 T_y u_3\|_{L^{p_2}}, \qquad \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{r}.$$

On occasion we will write this in a shorter form

$$\|L(u_1, u_2, u_3)\|_{L^r} \lesssim \|u_1\|_{L^{p_1}} \|L(u_2, u_3)\|_{L^{p_2}}.$$

To prove the boundedness in L^2 of the normal form transformation, we will use the following proposition from Tao [37]; for completeness we recall it below:

Lemma 2.1 (Leibnitz rule for P_k). *We have the commutator identity*

(2.3) $$[P_k, f]\, g = L(\partial_x f, 2^{-k} g).$$

When classifying cubic terms (and not only) obtained after implementing a normal form transformation, we observe that having a commutator structure is a desired feature. In particular Lemma 2.1 tells us that when one of the entry (call it g) has frequency $\sim 2^k$ and the other entry (call it f) has frequency $\lesssim 2^k$, then $P_k(fg) - f P_k g$ effectively shifts a derivative from the high-frequency function g to the low-frequency function f. This shift will generally ensure that all such commutator terms will be easily estimated.

Frequency envelopes. Before stating one of the main theorems of this paper, we revisit the *frequency envelope* notion; it will turn out to be very useful, and also an elegant tool later in the proof of the local well-posedness result, both in the proof of the a-priori bounds for solutions for the Cauchy problem (1.1) with data in L^2, which we state in Section 4.2, and in the proof of the bounds for the linearized equation, in the following section.

Following Tao's paper [36] we say that a sequence $c_k \in l^2$ is an L^2 frequency envelope for $\phi \in L^2$ if

    i) $\sum_{k=0}^{\infty} c_k^2 \lesssim 1$;

    ii) it is slowly varying, $c_j/c_k \leq 2^{\delta|j-k|}$, with δ a very small universal constant;

    iii) it bounds the dyadic norms of ϕ, namely $\|P_k\phi\|_{L^2} \leq c_k$.

Given a frequency envelope c_k we define

$$c_{\leq k} = \Big(\sum_{j\leq k} c_j^2\Big)^{\frac{1}{2}}, \qquad c_{\geq k} = \Big(\sum_{j\geq k} c_j^2\Big)^{\frac{1}{2}}.$$

Remark 2.2. *To avoid dealing with certain issues arising at low frequencies, we can harmlessly make the extra assumption that $c_0 \approx 1$.*

Remark 2.3. *Another useful variation is to weaken the slowly varying assumption to*

$$2^{-\delta|j-k|} \leq c_j/c_k \leq 2^{C|j-k|}, \qquad j < k,$$

where C is a fixed but possibly large constant. All the results in this paper are compatible with this choice. This offers the extra flexibility of providing higher regularity results by the same argument.

3. THE LINEAR FLOW

Here we consider the linear Benjamin-Ono flow,

$$(3.1) \qquad\qquad (\partial_t + H\partial^2)\psi = 0, \qquad \psi(0) = \psi_0.$$

Its solution $\phi(t) = e^{-tH\partial^2}\psi_0$ has conserved L^2 norm, and satisfies standard dispersive bounds:

Proposition 3.1. *The linear Benjamin-Ono flow satisfies the dispersive bound*

$$(3.2) \qquad\qquad \|e^{-tH\partial^2}\|_{L^1\to L^\infty} \lesssim t^{-\frac{1}{2}}.$$

This is a well known result. For convenience we outline the classical proof, and then provide a second, energy estimates based proof.

First proof of Proposition 3.1. Applying the spatial Fourier transform and solving the corresponding differential equation we obtain the following solution of the linear Benjamin-Ono equation

$$(3.3) \qquad\qquad \psi(t,x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i|\xi|\xi t + i\xi(x-y)}\psi_0(y)\, dyd\xi.$$

We change coordinates $\xi \to t^{-\frac{1}{2}}\eta$ and rewrite (3.3) as

$$\psi(t,x) = t^{-\frac{1}{2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i|\eta|\eta + i\eta t^{-\frac{1}{2}}(x-y)}\psi_0(y)\, dyd\eta,$$

8

which can be further seen as a convolution

$$\psi(t, x) = t^{-\frac{1}{2}} A(t^{-\frac{1}{2}} x) * \psi_0(x),$$

where $A(x)$ is an oscillatory integral

$$A(x) := \int_{-\infty}^{\infty} e^{-i|\eta|\eta + i\eta x} \, d\eta.$$

It remains to show that A is bounded, which follows by a standard stationary phase argument, with a minor complication arising from the fact that the phase is not C^2 at $\eta = 0$. $\square$

The second proof will also give us a good starting point in our study of the dispersive properties for the nonlinear equation. This is based on using the operator

$$L = x - 2tH\partial_x,$$

which is the push forward of x along the linear flow,

$$L(t) = e^{-tH\partial^2} x e^{tH\partial^2},$$

and thus commutes with the linear operator,

$$[L, \partial_t + H\partial^2] = 0.$$

In particular this shows that for solutions ψ to the homogeneous equation, the quantity $\|L\psi\|_{L^2}^2$ is also a conserved quantity.

Second proof of Proposition 3.1. We rewrite the dispersive estimate in the form

$$\|e^{-tH\partial^2} \delta_0\|_{L^\infty} \lesssim t^{-\frac{1}{2}}.$$

We approximate δ_0 with standard bump functions $\alpha_\epsilon(x) = \epsilon^{-1}\alpha(x/\epsilon)$, where α is a C_0^∞ function with integral one. It suffices to show the uniform bound

(3.4) $$\|e^{-tH\partial^2} \alpha_\epsilon\|_{L^\infty} \lesssim t^{-\frac{1}{2}}.$$

The functions α_ϵ satisfy the L^2 bound

$$\|\alpha_\epsilon\|_{L^2} \lesssim \epsilon^{-\frac{1}{2}}, \qquad \|x\alpha_\epsilon\|_{L^2} \lesssim \epsilon^{\frac{1}{2}}.$$

By energy estimates, this implies that

$$\|e^{-tH\partial^2} \alpha_\epsilon\|_{L^2} \lesssim \epsilon^{-\frac{1}{2}}, \qquad \|Le^{-tH\partial^2} \alpha_\epsilon\|_{L^2} \lesssim \epsilon^{\frac{1}{2}}.$$

Then the bound (3.4) is a consequence of the following

Lemma 3.2. *The following pointwise bound holds:*

(3.5) $$\|\psi\|_{L^\infty} + \|H\psi\|_{L^\infty} \lesssim t^{-\frac{1}{2}} \|\psi\|_{L^2}^{\frac{1}{2}} \|L\psi\|_{L^2}^{\frac{1}{2}}.$$

We remark that the operator L is elliptic in the region $x < 0$, therefore a better pointwise bound is expected there. Indeed, we have the estimate

(3.6) $$|\psi(t, x)| + |H\psi(t, x)| \leq t^{-\frac{1}{2}} (1 + |x_-|t^{-\frac{1}{2}})^{-\frac{1}{4}} \|\psi\|_{L^2}^{\frac{1}{2}} \|L\psi\|_{L^2}^{\frac{1}{2}},$$

where x_- stands for the negative part of x. To avoid repetition we do not prove this here, but it does follow from the analysis in the last section of the paper.

9

Proof. Denote

$$c = \int_{\mathbb{R}} \psi \, dx.$$

We first observe that we have

(3.7) $$c^2 \lesssim \|\psi\|_{L^2} \|L\psi\|_{L^2}.$$

All three quantities are constant along the linear Benjamin-Ono flow, so it suffices to verify this at $t = 0$. But there this inequality becomes

$$c^2 \lesssim \|\psi\|_{L^2} \|x\psi\|_{L^2},$$

which is straightforward using Hölder's inequality on each dyadic spatial region.

Next we establish the uniform $t^{-\frac{1}{2}}$ pointwise bound. We rescale to $t = 1$. Denote $u = P^+ \psi$, so that $\psi = 2\Re u$ and $H\psi = 2\Im u$. Hence it suffices to obtain the pointwise bound for u.

We begin with the relation

$$(x + 2i\partial)u = P^+ L\psi + c,$$

where the c term arises from the commutator of P^+ and x. We rewrite this as

$$\partial_x(ue^{\frac{ix^2}{4}}) = \frac{1}{2i} e^{\frac{ix^2}{4}} (P^+ L\psi + c).$$

Let F be a bounded antiderivative for $\frac{1}{2i} e^{\frac{ix^2}{4}}$. Then we introduce the auxiliary function

$$v = u e^{\frac{ix^2}{4t}} - cF,$$

which satisfies

$$\partial_x v = \frac{1}{2i} e^{\frac{ix^2}{4}} (P^+ L\psi).$$

In view of the previous bound (3.7) for c, it remains to show that

(3.8) $$\|v\|_{L^\infty}^2 \lesssim c^2 + \|v_x\|_{L^2} \|v + cF\|_{L^2}.$$

On each interval I of length R we have by Hölder's inequality

$$\|v\|_{L^\infty(I)} \lesssim R^{\frac{1}{2}} \|v_x\|_{L^2(I)} + R^{-\frac{1}{2}} \|v\|_{L^2(I)}.$$

Thus we obtain

$$\|v\|_{L^\infty}^2 \lesssim R\|v_x\|_{L^2}^2 + R^{-1}(\|v + cF\|_{L^2}^2 + c^2 R) = c^2 + R\|v_x\|_{L^2}^2 + R^{-1}\|v + cF\|_{L^2}^2,$$

and (3.8) follows by optimizing the value for R.

□

□

One standard consequence of the dispersive estimates is the Strichartz inequality, which applies to solutions to the inhomogeneous linear Benjamin-Ono equation.

(3.9) $$(\partial_t + H\partial^2)\psi = f, \qquad \psi(0) = \psi_0.$$

We define the Strichartz space S associated to the L^2 flow by

$$S = L_t^\infty L_x^2 \cap L_t^4 L_x^\infty,$$

as well as its dual

$$S' = L_t^1 L_x^2 + L_t^{\frac{4}{3}} L_x^1.$$

We will also use the notation
$$S^s = \langle D \rangle^{-s} S$$
to denote the similar spaces associated to the flow in H^s.

The Strichartz estimates in the L^2 setting are summarized in the following

Lemma 3.3. *Assume that ψ solves (3.9) in $[0,T] \times \mathbb{R}$. Then the following estimate holds.*

$$\|\psi\|_S \lesssim \|\psi_0\|_{L^2} + \|f\|_{S'}. \tag{3.10}$$

We remark that these Strichartz estimates can also be viewed as a consequence[3] of the similar estimates for the linear Schrödinger equation. This is because the two flows agree when restricted to functions with frequency localization in $\mathbb{R}^+$.

We also remark that we have the following Besov version of the estimates,

$$\|\psi\|_{\ell^2 S} \lesssim \|\psi_0\|_{L^2} + \|f\|_{\ell^2 S'}, \tag{3.11}$$

where

$$\|\psi\|_{\ell^2 S}^2 = \sum_k \|\psi_k\|_S^2, \qquad \|\psi\|_{\ell^2 S'}^2 = \sum_k \|\psi_k\|_{S'}^2.$$

The last property of the linear Benjamin-Ono equation we will use here is the bilinear L^2 estimate, which is as follows:

Lemma 3.4. *Let ψ^1, ψ^2 be two solutions to the inhomogeneous Schrödinger equation with data ψ_0^1, ψ_0^2 and inhomogeneous terms f^1 and f^2. Assume that the sets*

$$E_i = \{|\xi|, \xi \in supp \ \hat{\psi}^i\}$$

are disjoint. Then we have

$$\|\psi^1 \psi^2\|_{L^2} \lesssim \frac{1}{dist(E_1, E_2)} (\|\psi_0^1\|_{L^2} + \|f^1\|_{S'})(\|\psi_0^2\|_{L^2} + \|f^2\|_{S'}). \tag{3.12}$$

These bounds also follow from the similar bounds for the Schrödinger equation, where only the separation of the supports of the Fourier transforms is required. They can be obtained in a standard manner from the similar bound for products of solutions to the homogenous equation, for which we reffer the reader to [35].

One corollary of this applies in the case when we look at the product of two solutions which are supported in different dyadic regions:

Corollary 3.5. *Assume that ψ^1 and ψ^2 as above are supported in dyadic regions $|\xi| \approx 2^j$ and $|\xi| \approx 2^k$, $|j - k| > 2$, then*

$$\|\psi^1 \psi^2\|_{L^2} \lesssim 2^{-\frac{\max\{j,k\}}{2}} (\|\psi_0^1\|_{L^2} + \|f^1\|_{S'})(\|\psi_0^2\|_{L^2} + \|f^2\|_{S'}). \tag{3.13}$$

Another useful case is when we look at the product of two solutions which are supported in the same dyadic region, but with frequency separation:

Corollary 3.6. *Assume that ψ^1 and ψ^2 as above are supported in the dyadic region $|\xi| \approx 2^k$, but have $O(2^k)$ frequency separation between their supports. Then*

$$\|\psi^1 \psi^2\|_{L^2} \lesssim 2^{-\frac{k}{2}} (\|\psi_0^1\|_{L^2} + \|f^1\|_{S'})(\|\psi_0^2\|_{L^2} + \|f^2\|_{S'}). \tag{3.14}$$

---

3. Exept for the $L_t^4 L_x^\infty$ bound, as the Hilbert transform is not bounded in L^∞.

4. Normal form analysis and a-priori bounds

In this section we establish apriori L^2 bounds for regular (H_x^3) solutions for the Cauchy problem (1.1). First, we observe from the scale invariance (1.2) of the equation (1.1) that it suffices to work with solutions for which the L^2 norm is small, in which case it is natural to consider these solutions on the time interval $[-1, 1]$ (i.e., we set $T := 1$).

Precisely we may assume that the initial satisfies

$$(4.1) \qquad \|\phi(0)\|_{L_x^2} \leq \epsilon.$$

Then our main apriori estimate is as follows:

Theorem 5. *Let ϕ be an H_x^3 solution to* (1.1) *with small initial data as in* (4.1). *Let $\{c_k\}_{k=0}^{\infty} \in l^2$ so that ϵc_k is a frequency envelope for the initial $\phi(0)$ in L^2. Then we have the Strichartz bounds*

$$(4.2) \qquad \|\phi_k\|_{S^0([-1,1] \times \mathbf{R})} \lesssim \epsilon c_k,$$

as well as the bilinear bounds

$$(4.3) \qquad \|\phi_j \cdot \phi_k\|_{L^2} \lesssim 2^{-\frac{\max\{j,k\}}{2}} \epsilon^2 c_k \, c_j, \qquad j \neq k.$$

Here, the implicit constants do not depend on the H_x^3 norm of the initial data $\phi(0)$, but they will depend on $\|\phi(0)\|_{L^2}$. A standard iteration method will not work, because the linear part of the Benjamin-Ono equation does not have enough smoothing to compensate for the derivative in the nonlinearity. To resolve this difficulty we use ideas related to the normal form method, first introduced by Shatah in [34] in the context of dispersive PDEs. The main principle in the normal form method is to apply a quadratic correction to the unknown in order to replace a nonresonant quadratic nonlinearity by a milder cubic nonlinearity. Unfortunately this method does not apply directly here, because some terms in the quadratic correction are unbounded, and so are some of the cubic terms generated by the correction. To bypass this issue, here we develop a more favorable implementation of normal form analysis. This is carried out in two steps:

— a partial normal form transformation which is bounded and removes some of the quadratic nonlinearity

— a conjugation via a suitable exponential (also called gauge transform, [36]) which removes in a bounded way the remaining part of the quadratic nonlinearity.

This will transform the Benjamin-Ono equation (1.1) into an equation where the the quadratic terms have been removed and replaced by cubic perturbative terms.

4.1. The quadratic normal form analysis.
In this subsection we formally derive the normal form transformation for the Benjamin-Ono equation, (1.1). Even though we will not make use of it directly we will still use portions of it to remove certain ranges of frequency interactions from the quadratic nonlinearity.

Before going further, we emphasizes that by a *normal form* we refer to any type of transformation which will remove nonresonant quadratic terms; all such transformations are uniquely determined up to quadratic terms.

The normal form idea goes back to Birkhoff which used it in the context of ordinary differential equations. Later, Shatah [34] was the first to implement it in the context of partial differential equations. In general, the fact that one can compute such a normal form

for a partial differential equation with quadratic nonresonant interactions is not sufficient, unless the transformation is invertible, and, as seen in other works, in addition, good energy estimates are required. In the context of quasilinear equations one almost never expects the normal form transformation to be bounded, and new ideas are needed. In the Benjamin-Ono setting such ideas were first introduced by Tao [36] whose renormalization is a partial normal form transformation in disguise. More recently, other ideas have been introduced in the quasilinear context by Wu [39], Hunter-Ifrim [16], Hunter-Ifrim-Tataru [15], Alazard-Delort [2,3] and Hunter-Ifrim-Tataru [17].

In particular, for the Benjamin-Ono equation we seek a quadratic transformation

$$\tilde{\phi} = \phi + B(\phi, \phi),$$

so that the new variable $\tilde{\phi}$ solves an equation with a cubic nonlinearity,

$$(\partial_t + H\partial_x^2)\tilde{\phi} = Q(\phi, \phi, \phi),$$

where B and Q are translation invariant bilinear, respectively trilinear forms.

A direct computation yields an explicit formal spatial expression of the normal form transformation:

(4.4)
$$\tilde{\phi} = \phi - \frac{1}{4}H\phi \cdot \partial_x^{-1}\phi - \frac{1}{4}H\left(\phi \cdot \partial_x^{-1}\phi\right).$$

Note that at low frequencies (4.4) is not invertible, which tends to be a problem if one wants to apply the normal form transformation directly.

4.2. A modified normal form analysis.

We begin by writing the Benjamin-Ono equation (1.1) in a paradifferential form, i.e., we localize ourselves at a frequency 2^k, and then project the equation either onto negative or positive frequencies:

$$(\partial_t \mp i\partial_x^2)\phi_k^{\pm} = P_k^{\pm}(\phi \cdot \phi_x).$$

Since ϕ is real, ϕ^- is the complex conjugate of ϕ^+ so it suffices to work with the latter.

Thus, the Benjamin-Ono equation for the positive frequency Littlewood-Paley components ϕ_k^+ is

(4.5)
$$\left(i\partial_t + \partial_x^2\right)\phi_k^+ = iP_k^+(\phi \cdot \phi_x).$$

Heuristically, the worst term in $P_k^+(\phi \cdot \phi_x)$ occurs when ϕ_x is at high frequency and ϕ is at low frequency. We can approximate $P_k^+(\phi \cdot \phi_x)$, by its leading paradifferential component $\phi_{<k} \cdot \partial_x\phi_k^+$; the remaining part of the nonlinearity will be harmless. More explicitly we can eliminate it by means of a bounded normal form transformation.

We will extract out the main term $i\phi_{<k} \cdot \partial_x\phi_k^+$ from the right hand side nonlinearity and move it to the left, obtaining

(4.6)
$$\left(i\partial_t + \partial_x^2 - i\phi_{<k} \cdot \partial_x\right)\phi_k^+ = iP_k^+\left(\phi_{\geq k} \cdot \phi_x\right) + i\left[P_k^+, \phi_{<k}\right]\phi_x.$$

For reasons which will become apparent later on when we do the exponential conjugation, it is convenient to add an additional lower order term on the left hand side (and thus also on the right). Denoting by $A_{BO}^{k,+}$ the operator

(4.7)
$$A_{BO}^{k,+} := i\partial_t + \partial_x^2 - i\phi_{<k} \cdot \partial_x + \frac{1}{2}\left(H + i\right)\partial_x\phi_{<k}$$

13

we rewrite the equation (4.6) in the form

$$(4.8) \qquad A_{BO}^{k,+} \, \phi_k^+ = i P_k^+ \left(\phi_{\geq k} \cdot \phi_x \right) + i \left[P_k^+ , \, \phi_{<k} \right] \phi_x + \frac{1}{2} \left(H + i \right) \partial_x \phi_{<k} \cdot \phi_k^+ .$$

Note the key property that the operator $A_{BO}^{k,+}$ is symmetric, which in particular tells us that the L^2 norm is conserved in the corresponding linear evolution.

The case $k = 0$ is mildly different in this discussion. There we need no paradifferential component, and also we want to avoid the operator P_0^+ which does not have a smooth symbol. Thus we will work with the equation

$$(4.9) \qquad (\partial_t + H \partial_x^2) \phi_0 = P_0(\phi_0 \phi_x) + P_0(\phi_{>0} \phi_x),$$

where the first term on the right is purely a low frequency term and will play only a perturbative role.

The next step is to eliminate the terms on the right hand side of (4.8) using a normal form transformation

$$(4.10) \qquad \tilde{\phi}_k^+ := \phi_k^+ + B_k(\phi, \phi).$$

Such a transformation is easily computed and formally is given by the expression

$$(4.11) \qquad B_k(\phi, \phi) = \frac{1}{2} H P_k^+ \phi \cdot \partial_x^{-1} P_{<k} \phi - \frac{1}{4} P_k^+ \left(H \phi \cdot \partial_x^{-1} \phi \right) - \frac{1}{4} P_k^+ H \left(\phi \cdot \partial_x^{-1} \phi \right) .$$

One can view this as a subset of the normal form transformation computed for the full equation, see (4.4). Unfortunately, as written, the terms in this expression are not well defined because $\partial_x^{-1} \phi$ is only defined modulo constants. To avoid this problem we separate the low-high interactions which yield a well defined commutator, and we rewrite $B_k(\phi, \phi)$ in a better fashion as

$$(4.12) \qquad B_k(\phi, \phi) = -\frac{1}{2} \left[P_k^+ H , \, \partial_x^{-1} \phi_{<k} \right] \phi - \frac{1}{4} P_k^+ \left(H \phi \cdot \partial_x^{-1} \phi_{\geq k} \right) - \frac{1}{4} P_k^+ H \left(\phi \cdot \partial_x^{-1} \phi_{\geq k} \right) .$$

In the case $k = 0$ we will keep the first term on the right and apply a quadratic correction to remove the second. This yields

$$(4.13) \qquad B_0(\phi, \phi) = -\frac{1}{4} P_0^+ \left[H \phi \cdot \partial_x^{-1} \phi_{\geq 1} \right] - \frac{1}{4} P_0^+ H \left[\phi \cdot \partial_x^{-1} \phi_{\geq 1} \right] .$$

Remark 4.1. *The normal form transformation associated to (4.5) is the normal form derived in (4.4), but with the additional P_k^+ applied to it. Thus, the second and the third term in (4.11) are the projection P_k^+ of (4.4), which , in particular, implies that the linear Schrödinger operator $i\partial_t + \partial_x^2$ applied to these two terms will eliminate entirely the nonlinearity $P_k^+ (\phi \cdot \phi_x)$. The first term in (4.11) introduces the paradifferential corrections moved to the left of (4.8), and also has the property that it removes the unbounded part in the second and third term.*

Replying ϕ_k^+ with $\tilde{\phi}_k^+$ removes all the quadratic terms on the right and leaves us with an equation of the form

$$(4.14) \qquad A_{BO}^{k,+} \, \tilde{\phi}_k^+ = Q_k^3(\phi, \phi, \phi),$$

where $Q_k^3(\phi, \phi, \phi)$ contains only cubic terms in ϕ. We will examine $Q_k^3(\phi, \phi, \phi)$ in greater detail later in Lemma 4.2, where its full expression is given.

The case $k = 0$ is again special. Here the first normal form transformation does not eliminate the low-low frequency interactions, and our intermediate equation has the form

$$(4.15) \qquad (i\partial_t + \partial_x^2)\, \tilde{\phi}_0^+ = Q_0^2(\phi, \phi) + Q_0^3(\phi, \phi, \phi),$$

where Q_0^2 contains all the low-low frequency interactions

$$Q_0^2(\phi, \phi) := P_0^+ \left(\phi_0 \cdot \phi_x \right).$$

The second stage in our normal form analysis is to perform a second bounded normal form transformation that will remove the paradifferential terms in the left hand side of (4.14); this will be a renormalization, following the idea introduced by Tao ([36]). To achieve this we introduce and initialize the spatial primitive $\Phi(t, x)$ of $\phi(t, x)$, exactly as in Tao [36]. It turns out that $\Phi(t, x)$ is necessarily a real valued function that solves the equation

$$(4.16) \qquad \Phi_t + H\Phi_{xx} = \Phi_x^2,$$

which holds globally in time and space. Here, the initial condition imposed is $\Phi(0,0) = 0$. Thus,

$$(4.17) \qquad \Phi_x(t, x) = \frac{1}{2}\phi(t, x).$$

The idea in [36] was that in order to get bounds on ϕ it suffices to obtain appropriate bounds on $\Phi(t, x)$ which are one higher degree of reqularity as (4.17) suggests. Here we instead use Φ merely in an auxiliary role, in order to define the second normal form transformation. This is

$$(4.18) \qquad \psi_k^+ := \tilde{\phi}_k^+ \cdot e^{-i\Phi_{<k}}.$$

The transformation (4.18) is akin to a Cole-Hopf transformation, and expanding it up to quadratic terms, one observes that the expression obtained works as a normal form transformation, i.e., it removes the paradifferential quadratic terms. The difference is that the exponential will be a bounded transformation, whereas the corresponding quadratic normal form is not. One also sees the difference reflected at the level of cubic or higher order terms obtained after implementing these transformation (obviously they will differ).

By applying this *Cole-Hopf type* transformation, we rewrite the equation (4.14) as a a nonlinear Schrödinger equation for our final normal form variable ψ_k, with only cubic and quartic nonlinear terms:

$$(4.19) \qquad (i\partial_t + \partial_x^2)\, \psi_k^+ = [\tilde{Q}_k^3(\phi, \phi, \phi) + \tilde{Q}_k^4(\phi, \phi, \phi, \phi)]e^{-i\Phi_{<k}},$$

where $\tilde{Q}_k^3$ and $\tilde{Q}_k^4$ contain only cubic, respectively quartic terms; these are also computed in Lemma 4.2.

The case $k = 0$ is special here as well, in that no renormalization is needed. There we simply set $\psi_0 = \tilde{\phi}_0$, and use the equation (4.15).

This concludes the algebraic part of the analysis. Our next goal is study the analytic properties of our multilinear forms:

Lemma 4.2. *The quadratic form B_k can be expressed as*

$$(4.20) \qquad B_k(\phi, \phi) = 2^{-k}L_k(\phi_{<k}, \phi_k) + \sum_{j \geq k} 2^{-j}L_k(\phi_j, \phi_j) = 2^{-k}L_k(\phi, \phi).$$

The cubic and quartic expressions Q_k^3, $\tilde{Q}_k^3$ and $\tilde{Q}_k^4$ are translation invariant multilinear forms of the type

$$Q_k^3(\phi, \phi, \phi) = L_k(\phi, \phi, \phi) + L_k(H\phi, \phi, \phi),$$

(4.21)
$$\tilde{Q}_k^3(\phi, \phi, \phi) = L_k(\phi, \phi, \phi) + L_k(H\phi, \phi, \phi),$$

$$\tilde{Q}_k^4(\phi, \phi, \phi, \phi) = L_k(\phi, \phi, \phi, \phi) + L_k(H\phi, \phi, \phi, \phi),$$

all with output at frequency 2^k.

Proof. We recall that B_k is given in (4.12). For the first term we use Lemma 2.1. For the two remaining terms we split the unlocalized ϕ factor into $\phi_{<k} + \phi_{\geq k}$. The contribution of $\phi_{<k}$ is as before, while in the remaining bilinear term in $\phi_{\geq k}$ the frequencies of the two inputs must be balanced at some frequency 2^j where j ranges in the region $j \geq k$. For the last expression of B_k we simply observe that

(4.22)
$$\partial_x^{-1}\phi_{\geq k} = 2^{-k}L(\phi).$$

Next we consider Q_k^3 which is obtained by a direct computation
(4.23)
$$Q_k^3(\phi, \phi, \phi) = -\frac{1}{2}i \left[P_k^+ H, P_{<k}(\phi^2)\right]\phi - \frac{1}{2}i \left[P_k^+ H, \partial_x^{-1}\phi_{<k}\right]\partial_x(\phi^2) - \frac{1}{4}iP_k^+\left(H\partial_x(\phi^2) \cdot \partial_x^{-1}\phi_{\geq k}\right)$$

$$- \frac{1}{4}iP_k^+\left(H\phi \cdot P_{\geq k}(\phi^2)\right) - \frac{1}{4}iP_k^+ H\left(\partial_x(\phi^2) \cdot \partial_x^{-1}\phi_{\geq k}\right) - \frac{1}{4}iP_k^+ H\left(\phi \cdot P_{\geq k}(\phi^2)\right)$$

$$- iP_{<k}\phi \cdot \left\{-\frac{1}{2}\left[P_k^+ H, \phi_{<k}\right]\phi - \frac{1}{2}\left[P_k^+ H, \partial_x^{-1}\phi_{<k}\right]\phi_x - \frac{1}{4}P_k^+\left(H\phi_x \cdot \partial_x^{-1}\phi_{\geq k}\right)\right.$$

$$\left. - \frac{1}{4}P_k^+\left(H\phi \cdot \phi_{\geq k}\right) - \frac{1}{4}P_K^+ H\left(\phi_x \cdot \partial_x^{-1}\phi_{\geq k}\right) - \frac{1}{4}P_k^+ H\left(\phi \cdot \phi_{\geq k}\right)\right\}$$

$$- \frac{1}{2}\partial_x(H + i)\phi_{<k} \cdot B_k(\phi, \phi).$$

We consider each term separately. For the commutator terms we use Lemma 2.1 to eliminate all the inverse derivatives. This yields a factor of 2^{-k} which in turn is used to cancel the remaining derivative in the expressions. For instance consider the second term

$$\left[P_k^+ H, \partial_x^{-1}\phi_{<k}\right]\partial_x(\phi^2) = \left[P_k^+ H, \partial_x^{-1}\phi_{<k}\right]\tilde{P}_k\partial_x(\phi^2)$$

$$= L(\phi_{<k}, 2^{-k}\tilde{P}_k\partial_x(\phi^2))$$

$$= L(\phi_{<k}, \phi^2)$$

$$= L(\phi_{<k}, \phi, \phi).$$

The remaining terms are all similar. We consider for example the third term

$$P_k^+\left(H\partial_x(\phi^2) \cdot \partial_x^{-1}\phi_{\geq k}\right) = P_k^+\partial_x\left(H(\phi^2) \cdot \partial_x^{-1}\phi_{\geq k}\right) - P_k^+\left(H(\phi^2) \cdot \phi_{\geq k}\right).$$

The derivative in the first term yields a 2^k factor, and we can use (4.22), and the second term is straightforward.

For $\tilde{Q}_k^3$ an easy computation yields

$$\tilde{Q}_k^3(\phi, \phi, \phi) = Q_k^3(\phi, \phi, \phi) + \frac{1}{2}\phi_k^+ \cdot P_{<k}(\phi^2) - \frac{1}{4}\phi_k^+ \cdot (P_{<k}\phi)^2,$$

and both extra terms are straightforward.

16

Finally, $\tilde{Q}_k^4(\phi, \phi, \phi, \phi)$ is given by

$$\tilde{Q}_k^4(\phi, \phi, \phi, \phi) = \frac{1}{4} B_k(\phi, \phi) \cdot \left\{ 2P_{<k}(\phi^2) - (P_{<k}\phi)^2 \right\},$$

and the result follows from the one for the $B_k(\phi, \phi)$.

$\square$

4.3. The bootstrap argument.
We now finalize the proof of Theorem 5 using a standard continuity argument based on the H_x^3 global well-posedness theory. Given $0 < t_0 \leq 1$ we denote by

$$M(t_0) := \sup_j c_j^{-2} \|P_k\phi\|_{S^0[0,t_0]}^2 + \sup_{j \neq k \in \mathbf{N}} \sup_{y \in \mathbb{R}} c_j^{-1} \cdot c_k^{-1} \cdot \|\phi_j \cdot T_y\phi_k\|_{L^2[0,t_0]}.$$

Here, in the second term, the role of the condition $j \neq k$ is to insure that ϕ_j and ϕ_k have $O(2^{\max\{j,k\}})$ separated frequency localizations. However, by a slight abuse of notation, we also allow bilinear expressions of the form $P_k^1\phi \cdot P_k^2\phi$, where P_k^1 and P_k^2 are both projectors at frequency 2^k but with at least 2^{k-4} separation between the **absolute values** of the frequencies in their support.

We also remark here on the role played by the translation operator T_y. This is needed in order for us to be able to use thee bilinear bounds in estimating multilinear L type expressions.

We seek to show that

$$M(1) \lesssim \epsilon^2.$$

As ϕ is an H^3 solution, it is easy to see that $M(t)$ is continuous as a function of t, and

$$\lim_{t \searrow 0} M(t) \lesssim \epsilon^2.$$

This is because the only nonzero component of the S norm in the limit $t \to 0$ is the energy norm, which converges to the energy norm of the data.

Thus, by a continuity argument it suffices to make the bootstrap assumption

$$M(t_0) \leq C^2 \epsilon^2$$

and then show that

$$M(t_0) \lesssim \epsilon^2 + C^6 \epsilon^6.$$

This suffices provided that C is large enough (independent of ϵ) and ϵ is sufficiently small (depending on C). From here on $t_0 \in (0, 1]$ is fixed and not needed in the argument, so we drop it from the notations.

Given our bootstrap assumption, we have the starting estimates

(4.24)
$$\|\phi_k\|_{S^0} \lesssim C\epsilon c_k,$$

and

(4.25)
$$\|\phi_j \cdot T_y\phi_k\|_{L^2} \lesssim 2^{-\frac{\max\{j,k\}}{2}} C^2 \epsilon^2 c_j c_k, \qquad j \neq k, \qquad y \in \mathbb{R}.$$

where in the bilinear case, as discussed above, we also allow $j = k$ provided the two localization multipliers are at least 2^{k-4} separated. This separation threshold is fixed once and for all. On the other hand, when we prove that the bilinear estimates hold, no such sharp threshold is needed.

Our strategy will be to establish these bounds for the normal form variables ψ_k, and then to transfer them to the original solution ϕ by inverting the normal form transformations and estimating errors.

We obtain bounds for the normal form variables ψ_k^+. For this we estimate the initial data for ψ_k in L^2, and then the right hand side in the Schrodinger equation (4.19) for ψ_k^+ in $L^1 L^2$. For the initial data we have

Lemma 4.3. *Assume* (4.1). *Then we have*

(4.26)
$$\|\psi_k^+(0)\|_{L^2} \lesssim c_k \epsilon.$$

Proof. We begin by recalling the definition of $\psi(t, x)$:

$$\psi(t, x) = \tilde{\phi}_k^+ e^{-i\Phi_{<k}}.$$

The L_x^2 norms of ψ_k and $\tilde{\phi}_k^+$ are equivalent since the conjugation with the exponential is harmless. Thus, we need to prove that L^2 norm of $\tilde{\phi}_k^+$ is comparable with the L^2 norm of ϕ_k^+. The two variables are related via the relation (4.10). Thus, we reduce our problem to the study of the L^2 bound for the bilinear form $B_k(\phi, \phi)$. From Lemma (4.2) we know that

$$B_k(\phi, \phi) = 2^{-k} L_k(\phi_{<k}, \phi_k) + \sum_{j \geq k} 2^{-j} L_k(\phi_j, \phi_j),$$

so we estimate each term separately. For the first term we use the the smallness of the initial data in the L^2 norm, together with Bernstein's inequality, which we use for the low frequency term

$$\|2^{-k} L_k(\phi_{<k}, \phi_k)\|_{L^2} \lesssim 2^{-\frac{k}{2}} \cdot \epsilon \cdot \|\phi(0)\|_{L^2} = 2^{-\frac{k}{2}} \cdot \epsilon^2 \cdot c_k.$$

For the second component of $B_k(\phi, \phi)$, we again use Bernstein's inequality

$$\|\sum_{j \geq k} 2^{-j} L_k(\phi_j, \phi_j)\|_{L_x^2} \lesssim \sum_{j \geq k} 2^{-\frac{j}{2}} \cdot \epsilon \cdot \|\phi_j(0)\|_{L^2} \lesssim \sum_{j \geq k} 2^{-\frac{j}{2}} \cdot \epsilon^2 \cdot c_j \lesssim 2^{-\frac{k}{2}} \cdot c_k \cdot \epsilon^2.$$

This concludes the proof.

$\square$

Next we consider the right hand side in the ψ_k equation:

Lemma 4.4. *Assume* (4.24) *and* (4.25). *Then we have*

(4.27)
$$\|\tilde{Q}_k^3\|_{L^1 L^2} + \|\tilde{Q}_k^4\|_{L^1 L^2} \lesssim C^3 \epsilon^3 c_k.$$

A similar estimate holds for the quadratic term Q_0^2 which appears in the case $k = 0$, but that is quite straightforward.

Proof. We start by estimating the first term in (4.27). For completeness we recall the expression of $\tilde{Q}_k^3$ from Lemma 4.2:

$$\tilde{Q}_k^3(\phi, \phi, \phi) = L_k(\phi, \phi, \phi) + L_k(H\phi, \phi, \phi).$$

Here H plays no role so it suffices to discuss the first term. To estimate the trilinear expression $L_k(\phi, \phi, \phi)$ we do a frequency analysis. We begin by assuming that the first entry of L_k is localized at frequency 2^{k}_1, the second at frequency 2^{k_2}, and finally the third one is at frequency 2^{k_3}. As the output is at frequency 2^k, there are three possible cases:

18

— If $2^k < 2^{k_1} < 2^{k_2} = 2^{k_3}$, then we can use the bilinear Strichartz estimate for the imbalanced frequencies, and the Strichartz inequality for the remaining term to arrive at

$$\|L_k(\phi_{k_1}, \phi_{k_2}, \phi_{k_3})\|_{L_t^{\frac{4}{3}} L_x^2} \lesssim \|L(\phi_{k_1}, \phi_{k_3})\|_{L_{t,x}^2} \cdot \|\phi_{k_2}\|_{L_t^4 L_x^\infty}$$

$$\lesssim 2^{-\frac{k_3}{2}} \cdot C^2 \epsilon^2 \cdot c_{k_1} \cdot c_{k_3} \cdot \|\phi_{k_2}\|_{L_t^4 L_x^\infty}$$

$$\lesssim 2^{-\frac{k_3}{2}} \cdot C^3 \cdot \epsilon^3 \cdot c_{k_1} c_{k_2} c_{k_3} \lesssim 2^{-\frac{k}{2}} \cdot C^3 \cdot \epsilon^3 \cdot c_k^3.$$

— If $2^{k_1} = 2^{k_2} = 2^{k_3} \approx 2^k$, then we use directly the Strichartz estimates

$$\|L_k(\phi_{k_1}, \phi_{k_2}, \phi_{k_3})\|_{L_t^2 L_x^2} \lesssim \|\phi_{k_1}\|_{L_t^\infty L_x^2} \cdot \|\phi_{k_2}\|_{L_t^4 L_x^\infty} \cdot \|\phi_{k_3}\|_{L_t^4 L_x^\infty} \lesssim C^3 \epsilon^3 c_k^3.$$

— If $2^{k_1} = 2^{k_2} = 2^{k_3} \gg 2^k$ then the frequencies of the three entries must add to $O(2^k)$. Then the absolute values of at least two of the three frequencies must have at least a 2^{k_3-4} separation. Thus, the bilinear Strichartz estimate applies, and the same estimate as in the first case follows in the same manner.

This concludes the bound for $\tilde{Q}_k^3$.

Finally, the $L_t^1 L_x^2$ bound for

$$\tilde{Q}_k^4(\phi, \phi, \phi, \phi) = \frac{1}{4} B_k(\phi, \phi) \cdot \left\{ 2P_{<k}(\phi^2) - (P_{<k}\phi)^2 \right\},$$

follows from the L^2 bound for $B_k(\phi, \phi)$ obtained in Lemma 4.3 together with the $L_t^4 L_x^\infty$ bounds for the remaining factors. To bound these terms we do similar estimates as the ones in Lemma 4.3. $\qquad\square$

Given the bounds in the two above lemmas we have the Strichartz estimates for ψ_k:

$$\|\psi_k\|_{S^0} \lesssim \|\psi_k(0)\|_{L_x^2} + \|\tilde{Q}_k^3(\phi, \phi, \phi) + \tilde{Q}_k^4(\phi, \phi, \phi, \phi)\|_{L_t^1 L_x^2} \lesssim c_k \left(\epsilon + \epsilon^3 C^3 \right).$$

This implies the same estimate for $\tilde{\phi}_k^+$. Further we claim that the same holds for ϕ_k^+. For this we need to estimate $B_k(\phi, \phi)$ in S^0. We recall that

$$B_k(\phi, \phi) = 2^{-k} L_k(P_{<k}\phi, P_k\phi) + \sum_{j \geq k} 2^{-j} L_k(\phi_j, \phi_j).$$

We now estimate

$$\|B_k(\phi, \phi)\|_{S^0} \lesssim 2^{-k} \|\phi_k\|_{S^0} \|\phi_{<k}\|_{L^\infty} + \sum_{j \geq k} 2^{-j} \|\phi_j\|_{S^0} \|\phi_j\|_{L^\infty}$$

$$\lesssim C\epsilon^2 c_k 2^{-\frac{k}{2}} + \sum_{j \geq k} C\epsilon^2 c_j 2^{-\frac{j}{2}}$$

$$\lesssim C\epsilon^2 c_k 2^{-\frac{k}{2}}.$$

Here we have used Bernstein's inequality to estimate the L^∞ norm in term of the mass, and the slowly varying property of the c_k's for the last series summation. This concludes the Strichartz component of the bootstrap argument.

For later use, we observe that the same argument as above but with without using Bernstein's inequality yields the bound

(4.28) $$\|\psi_k - e^{-i\Phi_{<k}}\phi_k^+\|_{L^2 L^\infty \cap L^4 L^2} \lesssim 2^{-k} \epsilon^2 C^2 c_k$$

as a consequence of a similar bound for B_k.

We now consider the bilinear estimates in our bootstrap argument. We drop the translations from the notations, as they play no role in the argument. Also to fix the notations, in what follows we assume that $j < k$. The case when $j = k$ but we have frequency separation is completely similar.

We would like to start from the bilinear bounds for ψ_k, which solve suitable inhomogeneous linear Schrödinger equations. However, the difficulty we face is that, unlike $\tilde{\phi}_k^+$, ψ_k are no longer properly localized in frequency, therefore for $j \neq k$, ψ_j and ψ_k are no longer frequency separated. To remedy this we introduce additional truncation operators $\tilde{P}_j$ and $\tilde{P}_k$ which still have $2^{\max\{j,k\}}$ separated supports but whose symbols are identically 1 in the support of P_j, respectively P_k. Then the bilinear L^2 bound in Lemma 3.4 yields

$$\|\tilde{P}_j\psi_j \cdot \tilde{P}_k\psi_k\|_{L^2} \lesssim \epsilon^2 c_j c_k 2^{-\frac{\max\{j,k\}}{2}}(\epsilon^2 + C^6\epsilon^6).$$

It remains to transfer this bound to $\phi_j^+ \phi_k^+$. We expand

$$\tilde{P}_j\psi_j\tilde{P}_k\psi_k - \phi_j^+ e^{-i\Phi_{<j}}\phi_k^+ e^{-i\Phi_{<k}} = \tilde{P}_j\psi_j(\tilde{P}_k\psi_k - \phi_k^+ e^{-i\Phi_{<k}}) + (\tilde{P}_j\psi_j - \phi_j^+ e^{-i\Phi_{<j}})\phi_k^+ e^{-i\Phi_{<k}}.$$

For the first term we use the bound (4.28) for the second factor combined with the Strichartz bound for the second,

$$\|\tilde{P}_j\psi_j(\tilde{P}_k\psi_k - \phi_k^+ e^{-i\Phi_{<k}})\|_{L^2} \lesssim \|\psi_j\|_{L^\infty L^2}\|\psi_k - \phi_k^+ e^{-i\Phi_{<k}}\|_{L^2 L^\infty} \lesssim \epsilon^3 C^2 c_j c_k 2^{-k},$$

which is better than we need. It remains to consider the second term, where we freely drop the exponential. There the above argument no longer suffices, as it will only yield a 2^{-k} low frequency gain.

We use the commutator Lemma 2.1 to express the difference in the second term as

$$\begin{aligned}
\tilde{P}_j\psi_j - \phi_j^+ e^{-i\Phi_{<j}} &= (\tilde{P}_j - 1)(\tilde{\phi}_j^+ e^{-i\Phi_{<j}}) + B_j(\phi,\phi)e^{-i\Phi_{<j}} \\
&= [\tilde{P}_j - 1, e^{-i\Phi_{<j}}]\phi_j^+ e^{-i\Phi_{<j}}) + (\tilde{P}_j - 1)(B_j(\phi,\phi)e^{-i\Phi_{<j}} + +B_j(\phi,\phi)e^{-i\Phi_{<j}} \\
&= 2^{-j}L(\partial_x e^{-i\Phi_{<j}}, \phi_j^+) + L(B_j(\phi,\phi), e^{-i\Phi_{<j}}) \\
&= 2^{-j}L(\phi_{<j}, \phi_j, e^{-i\Phi_{<j}}) + \sum_{l>j} 2^{-l}L(\phi_l, \phi_l, e^{-i\Phi_{<j}}).
\end{aligned}$$

Now we multiply this by ϕ_k^+, and estimate in L^2 using our bootstrap hypothesis. For $l \neq k$ we can use a bilinear L^2 estimate combined with an L^∞ bound obtained via Bernstein's inequality. For $l = k$ we use three Strichartz bounds. The exponential is harmlessly discarded in all cases. We obtain

$$\|(\tilde{P}_j\psi_j - \phi_j^+ e^{-i\Phi_{<j}})\phi_k^+\|_{L^2} \lesssim \epsilon^3 C^2(c_j c_k 2^{-\frac{j}{2}}2^{-\frac{k}{2}} + \sum_{l>j} c_l c_k 2^{-\frac{l}{2}}2^{-\frac{k}{2}}) = \epsilon^3 C^2 c_j c_k 2^{-\frac{j}{2}}2^{-\frac{k}{2}}$$

which suffices.

5. Bounds for the linearized equation

In this section we consider the linearized Benjamin-Ono equation equation,

$$(5.1) \qquad (\partial_t + H\partial_x^2)v = \partial_x(\phi v).$$

Understanding the properties of the linearized flow is critical for any local well-posedness result.

Unfortunately, studying the linearized problem in L^2 presents considerable difficulty. One way to think about this is that L^2 well-posedness for the linearized equation would yield Lipschitz dependence in L^2 for the solution to data map, which is known to be false.

Another way is to observe that by duality, L^2 well-posedness implies $\dot{H}^{-1}$ well-posedness, and then, by interpolation, $\dot{H}^s$ well-posedness for $s \in [0,1]$. This last consideration shows that the weakest (and most robust) local well-posedness result we could prove for the linearized equation is in $\dot{H}^{-\frac{1}{2}}$.

Since we are concerned with local well-posedness here, we will harmlessly replace the homogeneous space $\dot{H}^{-\frac{1}{2}}$ with $H^{-\frac{1}{2}}$. Then we will prove the following:

Theorem 6. *Let ϕ be an H^3 solution to the Benjamin-Ono equation in $[0,1]$ with small mass, as in (4.1). Then the linearized equation (5.1) is well-posed in $H^{-\frac{1}{2}}$ with a uniform bound*

$$(5.2) \qquad \|v\|_{C(0,1;H^{-\frac{1}{2}})} \lesssim \|v_0\|_{H^{-\frac{1}{2}}}$$

with a universal implicit constant (i.e., not depending on the H^3 norm of ϕ).

We remark that as part of the proof we also show that the solutions to the linearized equation satisfy appropriate Strichartz and bilinear L^2 bounds expressed in terms of the frequency envelope of the initial data.

The rest of the section is devoted to the proof of the theorem. We begin by considering more regular solutions:

Lemma 5.1. *Assume that ϕ is an H^3 solution to the Benjamin-Ono equation. Then the linearized equation (5.1) is well-posed in H^1, with uniform bounds*

$$(5.3) \qquad \|v\|_{C(0,1;H^1)} \lesssim \|v_0\|_{H^1}.$$

Compared with the main theorem, here the implicit constant is allowed to depend on the H^3 norm of ϕ.

Proof. The lemma is proved using energy estimates. We begin with the easier L^2 well-posedness. On one hand, for solutions for (5.1) we have the bound

$$\frac{d}{dt}\|v\|_{L^2}^2 = \int_{\mathbb{R}} v\partial_x(\phi v)\,dx = \frac{1}{2}\int_{\mathbb{R}} v^2\partial_x\phi\,dx \lesssim \|\phi_x\|_{L^\infty}\|v\|_{L^2}^2,$$

which by Gronwall's inequality shows that

$$\|v\|_{L_t^\infty L_x^2} \lesssim \|v_0\|_{L_x^2},$$

thereby proving uniqueness. On the other hand, for the (backward) adjoint problem

$$(5.4) \qquad (\partial_t + H\partial_x^2)w = \phi\partial_x w, \qquad w(1) = w_1$$

we similarly have

$$\|w\|_{L_t^\infty L_x^2} \lesssim \|w_1\|_{L_x^2},$$

which proves existence for the direct problem.

To establish H^1 well-posedness in a similar manner we rewrite our evolution as a system for $(v, v_1 := \partial_x v)$,

$$\begin{cases} (\partial_t + H\partial_x^2)v = \partial_x(\phi v), \\ (\partial_t + H\partial_x^2)v_1 = \partial_x(\phi v_1) + \phi_x v_1 + \phi_{xx} v. \end{cases}$$

21

An argument similar to the above one shows that this system is also L^2 well-posed. Further, if initially we have $v_1 = v_x$ then this condition is easily propagated in time. This concludes the proof of the lemma. □

Given the last Lemma 5.1, in order to prove Theorem 6, it suffices to show that the H^1 solutions v given by the Lemma 5.1 satisfy the bound (5.2). It is convenient in effect to prove stronger bounds. To state them we assume that $\|v(0)\|_{H^{-\frac{1}{2}}} \leq 1$, and consider a frequency envelope d_k for $v(0)$ in $H^{-\frac{1}{2}}$. Without any restriction in generality we may assume that $c_k \leq d_k$, where c_k represents an L^2 frequency envelope for $\phi(0)$ as in the previous section. With these notations, we aim to prove that the dyadic pieces v_k of v satisfy the Strichartz estimates

$$\|v_k\|_{S^0} \lesssim 2^{\frac{k}{2}} d_k,$$

as well as the bilinear L^2 estimates

$$\|L(v_j, \phi_k)\|_{L^2} \lesssim \epsilon d_j c_k 2^{\frac{j}{2}} \cdot 2^{-\frac{\min\{j,k\}}{2}}.$$

Again, here we allow for $j = k$ under a 2^{k-4} frequency separation condition. Since v is already in H^1 and ϕ is in H^3, a continuity argument shows that it suffices to make the bootstrap assumptions

(5.5) $$\|v_k\|_{S^0} \leq C 2^{\frac{k}{2}} d_k,$$

(5.6) $$\sup_{y \in \mathbb{R}} \|v_j T_y \phi_k\|_{L^2} \lesssim C \epsilon d_j c_k 2^{\frac{j}{2}} 2^{-\frac{\min\{j,k\}}{2}}, \quad j \neq k,$$

and prove that

(5.7) $$\|v_k\|_{S^0} \lesssim (1 + \epsilon C) 2^{\frac{k}{2}} d_k,$$

respectively

(5.8) $$\sup_{y \in \mathbb{R}} \|v_j T_y \phi_k\|_{L^2} \lesssim \epsilon(1 + \epsilon C) d_j c_k 2^{\frac{j}{2}} 2^{-\frac{\min\{j,k\}}{2}}, \quad j \neq k.$$

We proceed in the same manner as for the nonlinear equation, rewriting the linearized equation in paradifferential form as

(5.9) $$A_{BO}^{k,+} v_k^+ = i P_k^+ \partial_x(\phi \cdot v) - i \phi_{<k} \partial_x v_k^+ + \frac{1}{2} \partial_x (H + i) \phi_{<k} \cdot v_k^+.$$

Here, in a similar manner as before, we isolate the case $k = 0$, where no paradifferential terms are kept on the left.

The next step is to use a normal form transformation to eliminate quadratic terms on the right, and replace them by cubic terms. The difference with respect to the prior computation is that here we leave certain quadratic terms on the right, because their corresponding normal form correction would be too singular. To understand why this is so we begin with a formal computation which is based on our prior analysis for the main problem. Precisely, the normal form which eliminates the full quadratic nonlinearity in the linearized equation (i.e. the first term on the right in (5.9)) is obtained by linearizing the normal for for the full equation, and is given by

(5.10) $$-\frac{1}{4} P_k^+ \left[H v \cdot \partial_x^{-1} \phi \right] - \frac{1}{4} P_k^+ H \left[v \cdot \partial_x^{-1} \phi \right] - \frac{1}{4} P_k^+ \left[H \phi \cdot \partial_x^{-1} v \right] - \frac{1}{4} P_k^+ H \left[\phi \cdot \partial_x^{-1} v \right].$$

On the other hand, the correction which eliminates the paradifferential component (i.e., last two terms in (5.9)) is given by

$$(5.11) \qquad \frac{1}{2} H P_k^+ v \cdot \partial_x^{-1} P_{<k} \phi,$$

which corresponds to an asymmetric version of the first term in B_k in (4.10). Thus, the full normal form correction for the right hand side of the equation (5.9) is (5.10) + (5.11). The term in (5.11) together with the last two entries in (5.10) yield a commutator structure as in B_k in the previous section. To obtain a similar commutator structure for the first two terms in (5.10) we would need an additional correction

$$(5.12) \qquad \frac{1}{2} H P_k^+ \phi \cdot \partial_x^{-1} P_{<k} v.$$

Precisely, if we add the three expressions above we obtain the linearization of B_k,

$$(5.10) + (5.11) + (5.12) = 2 B_k(v, \phi),$$

where B_k stands for the symmetric bilinear form associated to the quadratic form B_k defined in (4.12). Hence, our desired normal form correction is

$$(5.10) + (5.11) = 2 B_k(v, \phi) - (5.12).$$

Unfortunately the expression (5.12) contains $\partial_x^{-1} v$ which is ill defined at low frequencies. Unlike in the analysis of the main equation in the previous section, here we also have no commutator structure to compensate. To avoid this problem we exclude the frequencies < 1 in v from the (5.12) part of the normal form correction. Thus, our quadratic normal form correction will be

$$(5.13) \qquad B_k^{lin}(\phi, v) = 2 B_k(v, \phi) - \frac{1}{2} H P_k^+ \phi \cdot \partial_x^{-1} v_{(0,k)}.$$

This serves as a quadratic correction for the full quadratic terms in the right hand side of (5.9), except for the term which corresponds to the frequencies of size $O(1)$ in w, namely the expression

$$Q_k^{2,lin}(\phi, v) = i v_0 \partial_x \phi_k^+ - \frac{1}{2} \partial_x (H + i) v_0 \cdot \phi_k^+.$$

Following the same procedure as in the normal form transformation for the full equation we denote the first normal form correction in the linearized equation by

$$(5.14) \qquad \tilde{v}_k^+ := v_k^+ + 2 B_k^{lin}(\phi, v).$$

The equation for $\tilde{v}_k^+$ has the form

$$(5.15) \qquad A_{BO}^{k,+} \tilde{v}_k^+ = Q_k^{3,lin}(\phi, \phi, v) + Q_k^{2,lin}(v_0, \phi_k).$$

Here $Q_k^{2,lin}$ is as above, whereas $Q_k^{3,lin}$ contains the linearization of Q_k^3 plus the extra contribution arising from the second term in B_k^{lin}, namely

$$(5.16) \qquad Q_k^{3,lin}(\phi, \phi, v) = 3 Q_k^3(\phi, \phi, v) + \frac{i}{2} \phi_k^+ P_{(0,k)}(v\phi) + \frac{i}{2} P_k^+ \partial_x(\phi^2) \partial_x^{-1} v_{(0,k)}.$$

Again there is a straightforward adjustment in this analysis for the case $k = 0$, following the model in the previous section. This adds a trivial low frequency quadratic term on the right.

Finally, for $k > 0$ we renormalize $\tilde{v}_k^+$ to

$$w_k := e^{-i\Phi_{<k}} \tilde{v}_k^+,$$

which in turn solves the inhomogeneous Schrödinger equation

(5.17) $\left(i\partial_t + \partial_x^2\right)w_k = [Q_k^{2,lin}(\phi_0, v_k^+) + 3\tilde{Q}_k^{3,lin}(\phi, \phi, v) + \tilde{Q}_k^{4,lin}(v, \phi, \phi, \phi)]e^{-i\Phi_{<k}},$

where

$$\tilde{Q}_k^{3,lin}(v, \phi, \phi) = Q_k^{3,lin}(v, \phi, \phi) + \frac{1}{4}v_k^+\left(2 \cdot P_{<k}(\phi^2) - (P_{<k}\phi)^2\right),$$

and

$$\tilde{Q}_k^{4,lin}(v, \phi, \phi) = Q_k^{3,lin}(v, \phi, \phi) + \frac{1}{4}B_k^{lin}(v, \phi)\left(2 \cdot P_{<k}(\phi^2) - (P_{<k}\phi)^2\right).$$

Our goal is now to estimate the initial data for w_k in L^2, and the inhomogeneous term in $L_t^1 L_x^2$. We begin with the initial data, for which we have

Lemma 5.2. *The initial data for w_k satisfies*

(5.18) $$\|w_k(0)\|_{L^2} \lesssim 2^{\frac{k}{2}}d_k.$$

Proof. It suffices to prove the similar estimate for $\tilde{v}_k$, which in turn reduces to estimating $B_k^{lin}(\phi, v)$. The same argument as in the proof of Lemma 4.3 yields

$$\|B_k^{lin}(\phi, v)\|_{L^2} \lesssim k\epsilon d_k,$$

which is stronger than we need. $\qquad\square$

Next we consider the inhomogeneous term:

Lemma 5.3. *The inhomogeneous terms in the w_k equation satisfy*

(5.19) $$\|Q_k^{2,lin}\|_{L^1 L^2} + \|\tilde{Q}_k^{3,lin}\|_{L^1 L^2} + \|\tilde{Q}_k^{4,lin}\|_{L^1 L^2} \lesssim 2^{\frac{k}{2}}C\epsilon d_k.$$

Proof. We begin with $Q_k^{2,lin}$, which is easily estimated in L^2 using the bilinear Strichartz estimates (5.6) in our bootstrap assumption.

All terms in the cubic part $\tilde{Q}_k^{3,lin}$ have the form $L_k(\phi, \phi, v)$ possibly with an added harmless Hilbert transform, except for the expression $P_k^+ \partial_x(\phi^2)\partial_x^{-1}v_{(0,k)}$. For this we have the bound

$$\|L_k(\phi, \phi, v)\|_{L^1 L^2} \lesssim 2^{\frac{k}{2}}C^2\epsilon^2 d_k.$$

The proof is identical to the similar argument for the similar bound in Lemma 4.27; we remark that the only difference occurs in the case when v has the highest frequency, which is larger than 2^k.

We now consider the remaining expression $P_k^+\partial_x(\phi^2)\partial_x^{-1}v_{(0,k)}$, which admits the expansion

$$P_k^+\partial_x(\phi^2)\partial_x^{-1}v_{(0,k)} = \sum_{j\in(0,k)}2^{-j}2^k L_k(\phi_k, \phi_{<k}, v_j) + \sum_{j\in(0,k)}\sum_{l\geq k}2^{-j}2^k L_k(\phi_l, \phi_l, v_j).$$

Here we necessarily have two unbalanced frequencies, therefore this expression is estimated by a direct application of the bilinear L^2 bound plus a Strichartz estimate.

The bound for the quartic term is identical to the one in Lemma 4.27. $\qquad\square$

Now we proceed to recover the Strichartz and bilinear L^2 bounds. In view of the last two Lemmas we do have the Strichartz bounds for w_k, and thus for $\tilde{v}_k$. On the other hand for the quadratic correction $B_k^{lin}(\phi, v)$ we have

$$B_k^{lin}(\phi, v) = 2^{-k}L(\phi_{<k}, v_k) + \sum_{j\in(0,k)}2^{-j}L(v_j, \phi_k) + \sum_{j\geq k}2^{-j}L(\phi_j, v_j).$$

24

Therefore, applying one Strichartz and one Bernstein inequality, we obtain

$$\|B_k^{lin}(\phi, v)\|_S \lesssim C\epsilon d_k,$$

which suffices in order to transfer the Strichartz bounds to v_k.

To recover the bilinear L^2 bounds we again follow the argument in the proof of Theorem 5. Our starting point is the bilinear L^2 bound

$$\|\tilde{P}_j v_j \cdot \tilde{P}_k \psi_k\|_{L^2} \lesssim C\epsilon d_j c_k 2^{\frac{j}{2}} 2^{-\frac{\max\{j,k\}}{2}}$$

which is a consequence of Lemma 3.4. To fix the notations we assume that $j < k$; the opposite case is similar. To transfer this bound to $v_j^+ \phi_k^+$ we write

$$\tilde{P}_j v_j \tilde{P}_k \psi_k - \phi_j^+ e^{-i\Phi_{<j}} \phi_k^+ e^{-i\Phi_{<k}} = \tilde{P}_j v_j (\tilde{P}_k \psi_k - \phi_k^+ e^{-i\Phi_{<k}}) + (\tilde{P}_j v_j - v_j^+ e^{-i\Phi_{<j}}) \phi_k^+ e^{-i\Phi_{<k}}.$$

For the first term we use the bound (4.28) for the second factor combined with the Strichartz bound for the first factor. It remains to consider the second term. We freely drop the exponential, and use the commutator result in Lemma 2.1 to express the difference in the second term as

$$\begin{aligned}
\tilde{P}_j w_j - v_j^+ e^{-i\Phi_{<j}} &= (\tilde{P}_j - 1)(\tilde{v}^+ e^{-i\Psi_{<j}}) + B_j^{lin}(\phi, v) e^{-i\Phi_{<j}} \\
&= [\tilde{P}_j - 1, e^{-i\Phi_{<j}}]v_j^+ + (\tilde{P}_j - 1)(B_j(\phi, v) e^{-i\Psi_{<j}}) + B_j(\phi, v) e^{-i\Phi_{<j}} \\
&= 2^{-j} L(\partial_x e^{-i\Phi_{<j}}, \phi_j^+) + L(B_j(\phi, v), e^{-i\Phi_{<j}}) \\
&= 2^{-j} L(\phi_{<j}, v_j, e^{-i\Phi_{<j}}) + 2^{-j} L(v_{<j}, \phi_j, e^{-i\Phi_{<j}}) + L(\partial^{-1} v_{(0,j)}, \phi_j, e^{-i\Phi_{<j}}) \\
&\quad + \sum_{l>j} 2^{-l} L(v_l, \phi_l, e^{-i\Phi_{<j}}).
\end{aligned}$$

Now we multiply this by ϕ_k^+, and estimate in L^2 using our bootstrap hypothesis. For $l \neq k$ we can use a bilinear L^2 estimate combined with an L^∞ bound obtained via Bernstein's inequality. For $l = k$ we use three Strichartz bounds. The exponential is harmlessly discarded in all cases. We obtain

$$\|(\tilde{P}_j w_j - \phi_j^+ e^{-i\Phi_{<j}}) \phi_k^+\|_{L^2} \lesssim C\epsilon^2 2^{-\frac{k}{2}} d_j d_k$$

which suffices. The same argument applies when the roles of j and k are interchanged.

6. L^2 WELL-POSEDNESS FOR BENJAMIN-ONO

Here we prove our main result in Theorem 2. By scaling we can assume that our initial data satisfies

$$\|\phi_0\|_{L^2} \leq \epsilon \ll 1, \tag{6.1}$$

and prove well-posedness up to time $T = 1$. We know that if in addition $\phi_0 \in H^3$ then solutions exist, are unique and satisfy the bounds in Theorem 5. For H^3 data we can also use the bounds for the linearized equation in Theorem 6 to compare two solutions,

$$\|\phi^{(1)} - \phi^{(2)}\|_{S^{-\frac{1}{2}}} \lesssim \|\phi^{(1)}(0) - \phi^{(2)}(0)\|_{C(0,1;H^{-\frac{1}{2}})}. \tag{6.2}$$

We call this property *weak Lipschitz dependence on the initial data*.

We next use the above Lipschitz property to construct solutions for L^2 data. Given any initial data $\phi_0 \in L^2$ satisfying (6.1), we consider the corresponding regularized data

$$\phi^{(n)}(0) = P_{<n}\phi_0.$$

These satisfy uniformly the bound (6.1), and further they admit a uniform frequency envelope ϵc_k in L^2,

$$\|P_k \phi^{(n)}(0)\|_{L^2} \leq \epsilon c_k.$$

By virtue of Theorem 5, the corresponding solutions $\phi^{(k)}$ exist in $[0,1]$, and satisfy the uniform bounds

(6.3) $$\|P_k \phi^{(n)}\|_S \lesssim \epsilon c_k.$$

On the other hand, the differences satisfy

$$\|\phi^{(n)} - \phi^{(m)}\|_{S^{-\frac{1}{2}}} \lesssim \|\phi^{(1)}(0) - \phi^{(2)}(0)\|_{H^{-\frac{1}{2})}} \lesssim (2^{-n} + 2^{-m})\epsilon.$$

Thus the sequence $\phi^{(n)}$ converges to some function ϕ in $S^{-\frac{1}{2}}$,

$$\|\phi^{(n)} - \phi\|_{S^{-\frac{1}{2}}} \lesssim 2^{-n}\epsilon.$$

In particular we have convergence in S for each dyadic component, therefore the function ϕ inherits the dyadic bounds in (6.3),

(6.4) $$\|P_k \phi\|_S \lesssim \epsilon c_k.$$

This further allows us to prove convergence in $\ell^2 S$. For fixed k we write

$$\limsup \|\phi^{(n)} - \phi\|_{\ell^2 S} \leq \limsup \|P_{<k}(\phi^{(n)} - \phi)\|_{\ell^2 S} + \|P_{\geq k}\phi\|_{\ell^2 S} + \limsup \|P_{\geq k}\phi^{(n)}\|_{\ell^2 S} \leq c_{\geq k}.$$

Letting $k \to \infty$ we obtain

$$\lim \|\phi^{(n)} - \phi\|_{\ell^2 S} = 0.$$

Finally, this property also implies uniform convergence in $C(0,1;L^2)$; this in turn allows us to pass to the limit in the Benjamin-Ono equation, and prove that the limit ϕ solves the Benjamin-Ono equation in the sense of distributions.

Thus, for each initial data $\phi_0 \in L^2$ we have obtained a weak solution $\phi \in \ell^2 S$, as the limit of the solutions with regularized data. Further, this solution satisfies the frequency envelope bound (6.4).

Now we consider the dependence of these weak solutions on the initial data. First of all, the $\ell^2 S$ convergence allows us to pass to the limit in (6.2), therefore (6.2) extends to these weak solutions. Finally, we show that these weak solutions depend continuously on the initial data in L^2. To see that, we consider a sequence of data $\phi^{(n)}(0)$ satisfying (6.1) uniformly, so that

$$\phi^{(n)}(0) \to \phi_0 \qquad \text{in } L^2.$$

Then by the weak Lipschitz dependence we have

$$\phi^{(n)} \to \phi \text{ in } S^{-\frac{1}{2}}.$$

Hence for the corresponding solutions we estimate

$$\phi^{(n)} - \phi = P_{<k}(\phi^{(n)} - \phi) + P_{\geq k}\phi^{(n)} - P_{\geq k}\phi.$$

Here the first term on the right converges to zero in $\ell^2 S$ as $n \to \infty$ by the weak Lipschitz dependence (6.1), and the last term converges to zero as $k \to \infty$ by the frequency envelope bound (6.4). Hence letting in order first $n \to \infty$ then $k \to \infty$ we have

$$\limsup_{n\to\infty} \|\phi^{(n)} - \phi\|_{\ell^2 S} \leq \|P_{\geq k}\phi\|_{\ell^2 S} + \limsup_{n\to\infty} \|P_{\geq k}\phi^{(n)}\|_{\ell^2 S}$$

26

and then
$$\limsup_{n\to\infty} \|\phi^{(n)} - \phi\|_{\ell^2 S} \leq \lim_{k\to\infty} \limsup_{n\to\infty} \|P_{\geq k}\phi^{(n)}\|_{\ell^2 S}.$$
It remains to show that this last right hand side vanishes. For this we use the frequency envelope bound (6.4) applied to $\phi^{(n)}$ as follows.

Given $\delta > 0$, we have
$$\|\phi^{(n)}(0) - \phi_0\|_{L^2} \leq \delta, \qquad n \geq n_\delta.$$
Suppose ϵc_k is an L^2 frequency envelope for ϕ_0, and δd_k is an L^2 frequency envelope for $\phi^{(n)}(0) - \phi_0$. Here d_k is a normalized frequency envelope, which however may depend on n. Then $\epsilon c_k + \delta d_k$ is an L^2 frequency envelope for $\phi^{(n)}(0)$. Hence by (6.4) we obtain for $n \geq n_\delta$
$$\|P_{\geq k}\phi^{(n)}\|_{\ell^2 S} \lesssim \epsilon c_{\leq k} + \delta d_{\leq k} \lesssim \epsilon c_{\leq k} + \delta.$$
Thus
$$\limsup_{n\to\infty} \|P_{\geq k}\phi^{(n)}\|_{\ell^2 S} \lesssim \epsilon c_{\leq k} + \delta,$$
and letting $k \to \infty$ we have
$$\lim_{k\to\infty} \limsup_{n\to\infty} \|P_{\geq k}\phi^{(n)}\|_{\ell^2 S} \lesssim \delta.$$
But $\delta > 0$ was arbitrary. Hence
$$\lim_{k\to\infty} \limsup_{n\to\infty} \|P_{\geq k}\phi^{(n)}\|_{\ell^2 S} = 0,$$
and the proof of the theorem is concluded.

7. THE SCALING CONSERVATION LAW

As discussed in the previous section, for the linear equation (3.1) with localized data we can measure the initial data localization with an x weight, and then propagate this information along the flow using the following relation:
$$\|x\psi(0)\|_{L^2} = \|L\psi(t)\|_{L^2} = \|(x - 2tH\partial_x)\psi(t)\|_{L^2}^2.$$
The question we ask here is whether there is a nonlinear counterpart to that. To understand this issue we expand
$$\|(x - 2tH\partial_x)\phi(t)\|_{L^2}^2 = \int x^2\phi^2 - 4xt\phi H\phi_x + 4t^2\phi_x^2 \, dx \, dt,$$
where we recognize the linear mass, momentum and energy densities.

To define the nonlinear counterpart of this we introduce the nonlinear mass, momentum and energy densities as
$$m = \phi^2,$$
$$p = \phi H\phi_x - \frac{1}{3}\phi^3,$$
$$e = \phi_x^2 - \frac{3}{4}\phi^2 H\phi_x + \frac{1}{8}\phi^4.$$
Then we set
$$G(\phi) = \int x^2 m - 4xtp + 4t^2 e \, dx.$$
For this we claim that the following holds:

27

Proposition 7.1. *Let ϕ be a solution to the Benjamin-Ono equation for which the initial data satisfies $\phi_0 \in H^2$, $x\phi_0 \in L^2$. Then*
a) $L\phi \in C_{loc}(\mathbb{R}; L^2(\mathbb{R}))$.
b) The expression $G(\phi)$ is conserved along the flow.
c) We have the representation

$$(7.1) \qquad\qquad G(\phi) = \|L^{NL}\phi\|_{L^2}^2$$

where

$$(7.2) \qquad\qquad L^{NL}\phi = x\phi - 2t(H\phi_x - \frac{1}{8}(3\phi^2 - (H\phi)^2)).$$

Here one can view the expression $L^{NL}\phi$ as a normal form correction to $L\phi$. While such a correction is perhaps expected to exist, what is remarkable is that it is both nonsingular and exactly conserved.

Proof. a) We first show that the solution ϕ satisfies

$$(7.3) \qquad\qquad \|x\phi(t)\|_{L^2} \lesssim_{\phi_0} \langle t \rangle.$$

For this we truncate the weight to x_R, which is chosen to be a smooth function which equals x for $|x| < R/2$ and R for $|x| > R$. Then we establish the uniform bound

$$(7.4) \qquad\qquad \frac{d}{dt}\|x_R\phi\|_{L^2}^2 \lesssim_{\phi_0} 1 + \|x_R\phi\|_{L^2}.$$

Indeed, we have

$$
\begin{aligned}
\frac{d}{dt}\|x_R\phi\|_{L^2}^2 &= \int_{\mathbb{R}} x_R^2 \phi(-H\partial_x^2\phi + \phi\phi_x)\,dx \\
&= \int_{\mathbb{R}} x_R^2 \phi_x H\phi_x\,dx + \int_{\mathbb{R}} 2x_R x_R'(\phi H\phi_x - \frac{1}{3}\phi^3)\,dx \\
&= \int_{\mathbb{R}} x_R \phi_x[x_R, H]\phi_x\,dx + \int_{\mathbb{R}} 2x_R x_R'(\phi H\phi_x - \frac{1}{3}\phi^3)\,dx \\
&= \int_{\mathbb{R}} -x_R'\phi[x_R, H]\phi_x - x_R\phi\partial_x[x_R, H]\phi_x\,dx + \int_{\mathbb{R}} 2x_R x_R'(\phi H\phi_x - \frac{1}{3}\phi^3)\,dx.
\end{aligned}
$$

Then it suffices to establish the commutator bounds

$$\|[x_R, H]\partial_x\|_{L^2 \to L^2} \lesssim 1, \qquad \|\partial_x[x_R, H]\|_{L^2 \to L^2} \lesssim 1$$

But these are both standard Coifman-Meyer estimates, which require only $x_R' \in BMO$.
Combining (7.3) with the uniform H^1 bound, we obtain

$$\|L\phi\|_{L^2} \lesssim_{\phi_0} \langle t \rangle.$$

To establish the continuity in time of $L\phi$, we write the evolution equation

$$(\partial_t + H\partial_x^2)L\phi = L\phi\phi_x + H\phi_x\phi_x,$$

and observe that this equation is strongly well-posed in L^2.
b) Integrating by parts we write

$$\frac{d}{dt}G(\phi) = \int_{\mathbb{R}} x^2(m_t + 2p_x) - 4xt(p_t + 2e_x)\,dx.$$

28

It remains to show that the two terms above vanish. For the first we compute

$$m_t + 2p_x = -2\phi H\phi_{xx} + 2\phi^2\phi_x + 2(\phi H\phi_x)_x - 2\phi^2\phi_x = 2\phi_x H\phi_x.$$

Integrating, we can commute in the x to get

$$\int x^2(m_t + 2p_x)\,dx = 2\int x^2\phi_x H\phi_x\,dx = \int x\phi_x H(x\phi_x)\,dx = 0$$

using the antisymmetry of H.

For the second term we write

$$p_t + 2e_x = -H\phi_{xx}H\phi_x + \phi\phi_{xxx} + \phi\phi_x H\phi_x + \phi H(\phi\phi_x)_x + \phi^2 H\phi_{xx} - \phi^3\phi_x$$

$$+ 4\phi_x\phi_{xx} - 3\phi\phi_x H\phi_x - \frac{3}{2}\phi^2 H\phi_{xx} + \phi^3\phi_x$$

$$= \partial_x\left(-\frac{1}{2}(H\phi_x)^2 + \frac{3}{2}\phi_x^2 + \phi\phi_{xx}\right) + \partial_x\left(\phi H(\phi\phi_x) - \frac{1}{2}\phi^2 H(\phi_x)\right)$$

$$- \phi_x H(\phi\phi_x) - \phi\phi_x H\phi_x.$$

Integrating by parts we have

$$\int x(p_t + 2e_x)\,dx = -\int -\frac{1}{2}(H\phi_x)^2 + \frac{3}{2}\phi_x^2 + \phi\phi_{xx} + \phi H(\phi\phi_x) - \frac{1}{2}\phi^2 H(\phi_x)\,dx$$

$$- \int x(\phi_x H(\phi\phi_x) + \phi\phi_x H\phi_x)\,dx.$$

To get zero in the first integral we integrate by parts and use the antisymmetry of H together with $H^2 = -I$. In the second integral we can freely commute x under one H and then use the antisymmetry of H.

c) We compute the expression

$$Err(\phi) = G(\phi) - \int_{\mathbb{R}} (x\phi - 2t(H\phi_x - \frac{1}{8}(3\phi^2 - (H\phi)^2)))^2\,dx.$$

The quadratic terms easily cancel, so we are first left with an xt term,

$$Err_1(\phi) = \int -4xt\left(-\frac{1}{3}\phi^3 + \frac{1}{8}\phi(3\phi^2 - (H\phi)^2)\right)dx.$$

For this to cancel we need

$$\int x\phi^3\,dx = 3\int x\phi(H\phi)^2\,dx.$$

Splitting into positive and negative frequencies

$$\phi = \phi^+ + \phi^-. \qquad H\phi = \frac{1}{i}(\phi^+ - \phi^-),$$

the cross terms cancel and we are left with having to prove that

$$\int x(\phi^+)^3\,dx = \int x(\phi^-)^3\,dx = 0.$$

where $\phi^- = \overline{\phi^+}$. By density it suffices to establish this for Schwartz functions ϕ. Then the Fourier transform of ϕ^+ is supported in $\mathbb{R}^+$, and is smooth except for a jump at frequency 0. It follows that the Fourier transform of $(\phi^+)^3$ is also supported in $\mathbb{R}^+$ but of class $C^{1,1}$ at

29

zero, i.e. with a second derivative jump. Hence the Fourier transform of $(\phi^+)^3$ vanishes at zero and the conclusion follows.

Secondly, we are left with a t^2 term, namely

$$Err_2(\phi) = \int 4t^2(-\frac{3}{4}\phi^2 H\phi_x + \frac{1}{4}(3\phi^2 - (H\phi)^2)H\phi_x) + 4t^2(\frac{1}{8}\phi^4 - \frac{1}{64}(3\phi^2 - (H\phi)^2)^2)\,dx.$$

The first term cancels since we can integrate out the triple $H\phi$ term. For the second we compute

$$8\phi^4 - (3\phi^2 - (H\phi)^2)^2 = -\phi^4 + 6\phi^2(H\phi)^2 - (H\phi)^4 = -2(\phi^-)^4 - 2(\phi^+)^4,$$

which again suffices, by the same argument as in the first case. □

We further show that this bound naturally extends to L^2 solutions:

Proposition 7.2. *Let ϕ be a solution to the Benjamin-Ono equation whose initial data satisfies $\phi_0 \in L^2$, $x\phi_0 \in L^2$. Then ϕ satisfies the bounds*

$$(7.5) \qquad\qquad\qquad\qquad \|L\phi\|_{L^2} \lesssim_{\phi_0} \langle t\rangle,$$

$$(7.6) \qquad\qquad\qquad\qquad \|\phi\|_{L^\infty} \lesssim_{\phi_0} t^{-\frac{1}{2}}\langle t^{\frac{1}{2}}\rangle.$$

Furthermore $L^{NL}\phi \in C(\mathbb{R}; L^2)$ and has conserved L^2 norm.

We remark that both bounds (7.5) and (7.6) are sharp, as they must apply to solitons.

Proof. Since the solution to data map is continuous in L^2, it suffices to prove (7.5) and (7.6) for H^2 solutions. Then we a-priori know that $L\phi \in L^2$ and $\phi \in L^\infty$, and we can take advantage of the $\|L^{NL}\phi\|_{L^2}$ conservation law. Hence we can use (3.5) to estimate

$$\|L\phi\|_{L^2} \lesssim \|L^{NL}\phi\|_{L^2} + t\|\phi\|_{L^\infty}\|\phi\|_{L^2} \lesssim \|L^{NL}\phi\|_{L^2} + t^{\frac{1}{2}}\|L\phi\|_{L^2}^{\frac{1}{2}}\|\phi\|_{L^2}^{\frac{3}{2}},$$

which by Cauchy-Schwarz inequality yields

$$\|L\phi\|_{L^2} \lesssim \|L^{NL}\phi\|_{L^2} + t\|\phi\|_{L^2}^3.$$

Now the pointwise bound bound for ϕ follows by reapplying (3.5).

For the last part, we first approximate the initial data ϕ_0 with H^2 data ϕ_0^n so that

$$\|\phi_0^n - \phi_0\|_{L^2} \to 0, \qquad \|x(\phi_0^n - \phi_0)\|_{L^2} \to 0.$$

Then we have $\|L^{NL}\phi^n\|_{L^2} \to \|L^{NL}\phi(0)\|_{L^2}$. Since $\phi^n \to \phi_0$ in L^2_{loc}, taking weak limits, we obtain

$$\|L^{NL}\phi\|_{L^\infty L^2} = \|L^{NL}\phi(0)\|_{L^2}.$$

Repeating the argument but with initialization at a different time t we similarly obtain

$$\|L^{NL}\phi\|_{L^\infty_t L^2_x} = \|L^{NL}\phi(t)\|_{L^2_x}.$$

Hence $\|L^{NL}\phi\|_{L^2}$ is constant in time. Then, the L^2 continuity follows from the corresponding weak continuity, which in turn follows from the strong L^2 continuity of ϕ. □

8. The uniform pointwise decay bound

In this section we establish our main pointwise decay bound for ϕ, namely

$$(8.1) \qquad \|\phi(t)\|_{L^\infty} + \|H\phi(t)\|_{L^\infty} \le C\epsilon\langle t\rangle^{-\frac{1}{2}}, \qquad |t| \le e^{\frac{c}{\epsilon}}$$

with a large universal constant C and a small universal constant c, to be chosen later.

Since the Benjamin-Ono equation is well-posed in L^2, with continuous dependence on the initial data, by density it suffices to prove our assertion under the additional assumption that $\phi_0 \in H^2$. This guarantees that the norm $\|u(t)\|_{L^\infty}$ is continuous as a function of time. Then it suffices to establish the desired conclusion (8.1) in any time interval $[0, T]$ under the additional bootstrap assumption

$$(8.2) \qquad \|\phi(t)\|_{L^\infty} + \|H\phi(t)\|_{L^\infty} \le 2C\epsilon\langle t\rangle^{-\frac{1}{2}}, \qquad |t| \le T \le e^{\frac{c}{\epsilon}}.$$

We will combine the above bootstrap assumption with the bounds arising from the following conservation laws:

$$(8.3) \qquad \|\phi(t)\|_{L^2} \le \epsilon,$$

$$(8.4) \qquad \|L^{NL}\phi(t)\|_{L^2} \le \epsilon,$$

$$(8.5) \qquad \int_{-\infty}^{\infty} \phi\, dx = c, \qquad |c| \le \epsilon.$$

We recall that L^{NL} is given by

$$L^{NL}\phi = x\phi - 2t\left[H\phi_x - \frac{1}{8}(3\phi^2 - (H\phi)^2)\right].$$

One difficulty here is that the quadratic term in $L^{NL}\phi$ cannot be treated perturbatively. However, as it turns out, we can take advantage of its structure in a simple fashion.

As a preliminary step, we establish a bound on the function

$$\partial^{-1}\phi(x) := \int_{-\infty}^{x} \phi(y)\, dy$$

as follows:

$$(8.6) \qquad |\partial^{-1}\phi(x)| \lesssim C\epsilon + C^2\epsilon^2 \log\langle t/x\rangle.$$

Assume first that $x \le -\sqrt{t}$. Then we write

$$\phi = \frac{1}{x}L^{NL}(\phi) + \frac{2t}{x}H\phi_x - \frac{t}{4x}(3\phi^2 - (H\phi)^2)).$$

Integrating by parts, we have

$$\partial^{-1}\phi(x) = \frac{2t}{x}H\phi(x) + \int_{-\infty}^{x} \frac{2t}{y^2}H\phi(y) + \frac{1}{x}L^{NL}(\phi) - \frac{t}{4y}(3\phi^2 - (H\phi)^2)\, dy.$$

For the first two terms we have a straightforward $\dfrac{C\epsilon\sqrt{t}}{|x|}$ bound due to (8.2). For the third term we use (8.4) and the Cauchy-Schwarz inequality. For the last integral term we use the L^2 bound (8.3) for $x < -t$ and the L^∞ bound (8.2) for $-t \le x \le -\sqrt{t}$ to get a bound of $C^2\epsilon^2 \log\langle t/x\rangle$.

31

This gives the desired bound in the region $x \leq -\sqrt{t}$. A similar argument yields the bound for $x \geq \sqrt{t}$, where in addition we use the conservation law (8.5) for $\int \phi \, dy$ to connect $\pm \infty$. Finally, for the inner region $|x| \leq \sqrt{t}$ we use directly the pointwise bound (8.2) on ϕ. This concludes the proof of (8.6).

Now we return to the pointwise bounds on ϕ and $H\phi$. Without using any bound for t, we will establish the estimate

(8.7) $$\|\phi(t)\|^2_{L^\infty} + \|H\phi(t)\|^2_{L^\infty} \lesssim \epsilon^2 t^{-1} (1 + C + C^3 \epsilon \log t + C^4 \epsilon^2 \log^2 t).$$

In order to retrieve the desired bound (8.1) we first choose $C \gg 1$ in order to account for the first two terms, and then restrict t to the range $C\epsilon \log t \ll 1$ for the last two terms. This determines the small constant c in (8.1).

To establish (8.7) we first use the expression for $L^{NL}(\phi)$ to compute

$$\frac{d}{dx}(|\phi|^2 + |H\phi|^2) = \frac{1}{t}F_1 + \frac{1}{t}F_2 + \frac{1}{4}F_3,$$

where

$$F_1 = \phi H L^{NL}(\phi) + H\phi L^{NL}(\phi), \qquad F_2 = x\phi H\phi - \phi H(x\phi),$$
$$F_3 = -\phi H(3\phi^2 - (H\phi)^2) + H\phi(3\phi^2 - (H\phi)^2).$$

We will estimate separately the contributions of F_1, F_2 and F_3. For F_1 we combine (8.3) and (8.4) to obtain

$$\|F_1\|_{L^1} \lesssim \epsilon^2,$$

which suffices. For F_2 we commute x with H to rewrite it as

$$F_2(x) = \phi(x) \int_{-\infty}^{\infty} \phi(y) \, dy,$$

which we can integrate using (8.6).

Finally, for F_3 we use the identity

$$H(\phi^2 - (H\phi)^2) = 2\phi H\phi$$

to rewrite it as

$$F_3 = -\phi H(\phi^2 + (H\phi)^2) - H\phi(\phi^2 + (H\phi)^2).$$

This now has a commutator structure, which allows us to write

$$\int_{-\infty}^{x_0} F_3(x) \, dx = -\int_{-\infty}^{x_0} \int_{x_0}^{\infty} \phi(x) \frac{1}{x-y}(\phi^2 + (H\phi)^2)(y) \, dy \, dx.$$

Here the key feature is that x and y are separated. We now estimate the last integral. We consider several cases:

a) If $|x - y| \lesssim \sqrt{t}$ then direct integration using (8.2) yields a bound of $C^3 \epsilon^3 t^{-1}$.

b) If $|x - y| > t$ then we use (8.3) to bound $\phi^2 + (H\phi)^2$ in L^1. Denoting $x_1 = \min\{x_0, y-t\}$, we are left with an integral of the form

$$\int_{-\infty}^{x_1} \frac{1}{x-y}\phi(x) \, dx = \frac{1}{x_1 - y}\partial^{-1}\phi(x_1) - \int_{-\infty}^{x_1} \frac{1}{(x-y)^2}\partial^{-1}\phi(x) \, dx.$$

As $|x_1 - y| > t$ from (8.6) we obtain a bound of
$$t^{-1}(C\epsilon^3 + C^2\epsilon^4 \log t).$$

c) $x - y \approx r \in [\sqrt{t}, t]$. Then we use (8.2) to bound $\phi^2 + (H\phi)^2$ in L^∞ and argue as in case (b) to obtain a bound of
$$t^{-1}(C^3\epsilon^3 + C^4\epsilon^4 \log t).$$
Then the dyadic r summation adds another $\log t$ factor.

9. THE ELLIPTIC REGION

Here we improve the pointwise bound on ϕ in the elliptic region $x < -\sqrt{t}$. Precisely, we will show that for $t < e^{\frac{c}{\epsilon}}$ we have

(9.1) $$|\phi(x)| + |H\phi(x)| \lesssim \epsilon t^{-\frac{1}{4}} x^{-\frac{1}{2}}, \qquad x \geq \sqrt{t}.$$

To prove this we take advantage of the ellipticity of the linear part $x - 2tH\partial_x$ of the operator L^{NL} in the region $x \geq \sqrt{t}$. For this linear part we claim the bound

(9.2) $$\|x\chi\phi\|_{L^2}^2 + \|t\chi\phi_x\|_{L^2}^2 \lesssim \|(x - 2tH\partial_x)\phi\|_{L^2}^2 + t^{\frac{3}{2}}\|\phi\|_{L^\infty}^2 + t^{\frac{1}{2}}\|\partial^{-1}\phi\|_{L^\infty}^2,$$

where χ is a smooth cutoff function which selects the region $\{x > \sqrt{t}\}$.

Assuming we have this, using also (8.1), (8.4) and (8.6) we obtain
$$\|x\chi\phi\|_{L^2}^2 + \|t\chi\phi_x\|_{L^2}^2 \lesssim \epsilon t^{\frac{1}{2}} + t^2\|\chi(\phi^2 + (H\phi)^2)\|_{L^2}^2.$$

We claim that we can dispense with the second term on the right. Indeed, we can easily use (8.1) to bound the ϕ^2 contribution by
$$\|\chi\phi^2\|_{L^2} \lesssim \|\phi\|_{L^\infty}\|\chi\phi\|_{L^2} \lesssim \epsilon t^{-1}\|x\chi\phi\|_{L^2}.$$

The $(H\phi)^2$ contribution is estimated in the same manner, but in addition we also need to bound the commutator

(9.3) $$\|[H,\chi]\phi\|_{L^2} \lesssim \|\phi\|_{L^\infty} + t^{-\frac{1}{2}}\|\partial^{-1}\phi\|_{L^\infty}.$$

Assuming we also have this commutator bound, it follows that

(9.4) $$\|x\chi\phi\|_{L^2}^2 + \|t(\chi\phi)_x\|_{L^2}^2 \lesssim \epsilon t^{\frac{1}{2}}.$$

This directly yields the desired pointwise bound (9.1) for ϕ.

Now we prove the $H\phi$ part of (9.1). For $x \approx r > t^{\frac{1}{2}}$ we decompose
$$\phi = \chi_r\phi + (1 - \chi_r)\phi,$$

where χ_r is a smooth bump function selecting this dyadic region.

For the contribution of the first term we use interpolation to write
$$\|H(\chi_r\phi)\|_{L^\infty} \lesssim \|\chi_r\phi\|_{L^2}^{\frac{1}{2}}\|\partial_x(\chi_r\phi)\|_{L^2}^{\frac{1}{2}} \lesssim \epsilon(t^{\frac{1}{4}}r^{-1})^{\frac{1}{2}}(t^{-\frac{3}{4}})^{\frac{1}{2}} = \epsilon t^{-\frac{1}{4}}r^{-\frac{1}{2}}.$$

For the second term we use the kernel for the Hilbert transform,
$$H[(1 - \chi_r)\phi](x) = \int \frac{1}{x - y}[(1 - \chi_r)\phi](y)\, dy.$$

For the contribution of the region $y > t^{\frac{1}{2}}$ we use the pointwise bound (9.1) on ϕ and directly integrate. For the contribution of the region $y < t^{\frac{1}{2}}$ we integrate by parts and use the bound (8.6) on $\partial^{-1}\phi$. This concludes the proof of the $H\phi$ bound in (9.1).

33

It remains to prove the bounds (9.2) and (9.3). Both are scale invariant in time, so without any restriction in generality we can assume that $t = 1$.

Proof of (9.3). The kernel $K(x, y)$ of $[\chi, H]$ is given by

$$K(x, y) = \frac{\chi(x) - \chi(y)}{x - y},$$

and thus satisfies

$$(1 + |x| + |y|)|K(x, y)| + (1 + |x| + |y|)^2 |\nabla_{x,y} K(x, y)| \lesssim 1$$

Then we write

$$\int_{\mathbb{R}} K(x, y)\phi(y)dy = -\int_{\mathbb{R}} K_y(x, y)\partial^{-1}\phi(y)dy$$

and then take absolute values and estimate.

Proof of (9.2). We multiply $(x - 2H\partial_x)\phi$ by $\chi := \chi_{\geq 1}(x)$, square and integrate. have

$$\|\chi(x - 2H\partial_x)\phi\|_{L^2}^2 - \|\chi x\phi\|_{L^2}^2 - 2\|\chi|x|^{\frac{1}{2}}|D|^{\frac{1}{2}}\phi\|_{L^2}^2 - \|\chi\phi_x\|_{L^2}^2 = \langle(T_1 + T_2)\phi, \phi\rangle$$

where

$$T_1 = |D|\chi^2|D| + \partial_x\chi^2\partial_x, \qquad T_2 = \chi^2 x|D| + |D|\chi^2 x - 2|D|^{\frac{1}{2}}\chi^2 x|D|^{\frac{1}{2}}.$$

Then it suffices to show that

(9.5) $$|\langle T_{1,2}\phi, \phi\rangle| \lesssim \|\phi\|_{L^\infty}^2 + \|\partial^{-1}\phi\|_{L^\infty}^2.$$

To achieve this we estimate the kernels $K_{1,2}$ of $T_{1,2}$. In order to compute the kernels K_1 and K_2 we observe that both T_1 and T_2 have a commutator structure

(9.6) $$T_1 = \partial_x \left[\left[\chi^2, H\right], H\right]\partial_x, \quad T_2 = \left[\left[|D|^{\frac{1}{2}}, \chi^2\right], |D|^{\frac{1}{2}}\right].$$

We first consider T_1 for which we claim that its kernel K_1 satisfies the bound

(9.7) $$|K_1(x, y)| \lesssim \frac{1}{(1 + |x|)(1 + |y|)(1 + |x| + |y|)}.$$

This suffices for the estimate (9.5).

To prove (9.7) we observe that instead of analyzing the kernel $K_1(x, y)$, we can analyze the kernel $\tilde{K}_1$:

$$K_1(x, y) = \partial_x \partial_y \tilde{K}_1(x, y),$$

where $\tilde{K}_1$ is the corresponding kernel of the commutator $[[\chi^2, H], H]$, and is given by

$$\tilde{K}_1(x, y) = \int \frac{\chi^2(x) - \chi^2(y)}{x - z} \cdot \frac{1}{z - y} - \frac{\chi^2(z) - \chi^2(y)}{z - z} \cdot \frac{1}{x - z} dz.$$

We can rewrite $\tilde{K}_1$ using the symmetry $z \to x + y - z$

$$\tilde{K}_1(x, y) = \int \frac{\chi^2(x) + \chi^2(y) - \chi^2(z) - \chi^2(x + y - z)}{(x - z)(y - z)} dz.$$

Secondly, in a similar fashion, we compute the kernel K_2 of T_2,

(9.8) $$K_2(x, y) = \int \frac{\chi^2(x) + \chi^2(y) - \chi^2(x + y - z) - \chi^2(z)}{|x - z|^{\frac{3}{2}}|y - z|^{\frac{3}{2}}} dz,$$

34

where again the numerator vanishes of order one at $x = z$ and $y = z$. For this kernel we distinguish two regions:

— $|x| + |y| \lesssim 1$; in this region a direct computation shows that the kernel K_2 has a mild logarithmic singularity on the diagonal $x = y$,

$$|K_2(x, y)| \leq 1 + |\log |x - y||.$$

— $|x| + |y| \gg 1$; in this region the kernel K_2 is smooth and can be shown to satisfy the bound

$$|K_2^{low}(x, y)| \lesssim \frac{(1 + \min\{|x|, |y|\})^{\frac{1}{2}}}{(1 + |x| + |y|)^{\frac{3}{2}}}.$$

This does not suffice for the bound (9.5). However after differentiation it improves to

$$|\partial_x \partial_y K_2^{low}(x, y)| \lesssim \frac{1}{(1 + \min\{|x|, |y|\})^{\frac{1}{2}}(1 + |x| + |y|)^{\frac{5}{2}}},$$

and that is enough to obtain (9.5).

References

[1] L. Abdelouhab, J.L. Bona, M. Felland, and J.-C. Saut, Nonlocal models for nonlinear dispersive waves, *physica D*, 40, 360-392, 1989

[2] T. Alazard, J.M. Delort, Global solutions and asymptotic behavior for two dimensional gravity water waves, *Annales Scientifiques de l cole Normale Suprieure*, 48(5), 2013

[3] T. Alazard, J.M. Delort, Sobolev estimates for two dimensional gravity water waves, *Astrisque*, No. 374 (2015), viii+241 pp.

[4] T.B. Benjamin, Internal waves of permanent form in fluids of great depth, *J. Fluid Mech.*, 29 (1967), 559-592.

[5] N. Burq and F. Planchon, On well-posedness for the Benjamin-Ono equation, *Math. Ann.*, (2008), Volume 340, Issue 3, pp 497-542.

[6] R. R. Coifman and Y. Meyer, Le double commutateur, *Analyse Harmonique d'Orsay*, 1976

[7] A. S. Fokas, P.M. Santini, Bi-Hamiltonian formulation of the Kadomtsev-Petviashvili and Benjamin-Ono equations. *J. Math. Phys.* 29 (1988), no. 3, 604-617

[8] G. Fonseca, G. Ponce. The IVP for the Benjamin-Ono equation in weighted Sobolev spaces, *Journal of Functional Analysis*, Vol. 260, Issue 2, 15 January 2011, 436-459

[9] G. Fonseca, G. Ponce, F. Linares, The IVP for the Benjamin-Ono equation in weighted Sobolev spaces II, *Journal of Functional Analysis*, Vol. 262, Issue 5, 1 March 2012, 2031-2049

[10] Gustafson, Stephen, Takaoka, Hideo; Tsai, Tai-Peng Stability in H1/2 of the sum of K solitons for the Benjamin-Ono equation. *J. Math. Phys.*, 50 (2009), no. 1

[11] B. Harrop-Griffiths, M. Ifrim, and D. Tataru, Finite depth gravity water waves in holomorphic coordinates, to appear in *Annals of PDE*, http://arxiv.org/abs/1607.02409, 2016

[12] B. Harrop-Griffiths, M. Ifrim, and D. Tataru, The lifespan of small data solutions to the KP-I, to appear *International Mathematics Research Notices*, http://arxiv.org/pdf/1409.4487.pdf, 2014

[13] S. Herr, Well-posedness for equations of Benjamin-Ono type, *Illinois J. Math.*, Vol. 51, Number 3 (2007), 951-976.

[14] S. Herr, A. Ionescu, D. Alexandru, C. E. Kenig and H. Koch. A para-differential renormalization technique for nonlinear dispersive equations, *Comm. Partial Differential Equations* 35, no. 10, 1827-1875, 2010

[15] J. K. Hunter, M. Ifrim, D. Tataru, and T. K. Wong, Long time solutions for a Burgers-Hilbert equation via a modified energy method, *Proceedings of the AMS*, 2013

[16] J. K. Hunter, M. Ifrim, Enhanced lifespan of smooth solutions of a Burgers-Hilbert equation, *SIAM Journal on Mathematical Analysis*, Volume 44, Issue 3, 2012, 1279-2235.

[17] J. K. Hunter, M. Ifrim, and D. Tataru, Two dimensional water waves in holomorphic coordinates, *Comm. Math. Phys.*, Vol.346(2), pp. 483-552, 2016

[18] M. Ifrim and D. Tataru, Two dimensional gravity water waves with constant vorticity: I. Cubic lifespan, 2015 *e-print* available at http://arxiv.org/abs/1510.07732, 2015

[19] M. Ifrim and D. Tataru, The lifespan of small data solutions in two dimensional capillary water waves, 2014 *e-print* available at http://arxiv.org/pdf/1406.5471v2.pdf, to appear in *Archive for Rational Mechanics and Analysis*

[20] M. Ifrim and D. Tataru, Two dimensional water waves in holomorphic coordinates II: global solutions, *Bull. Soc. Math. France*, Vol. 144(2), pp. 369-394, 2016

[21] M. Ifrim and D. Tataru, Global bounds for the cubic nonlinear Schrödinger equation (NLS) in one space dimension, *Nonlinearity*, Volume 28, Number 8, 2015.

[22]] R. J. Iorio, On the Cauchy problem for the Benjamin-Ono equation, *Comm. Partial Differential Equations*, 11(1986), 10311081.

[23] A. D. Ionescu and C. E. Kenig, Global well-posedness of the Benjamin-Ono equation in low regularity spaces, *Journal of the American Mathematical Society*, 20 (2007), 753-798.

[24] D. J. Kaup, T. I. Lakoba, and Y. Matsuno, Complete integrability of the Benjamin-Ono equation by means of action-angle variables, *Phys. Lett., A* 238, No.2-3, 123-133 (1998).

[25] C. E. Kenig and K. D. Koenig, On the local well-posedness of the Benjamin-Ono and modified Benjamin-Ono equations, *Math. Res. Lett.*, 10 (2003), 879-895.

[26] C. E. Kenig and Y. Martel , Asymptotic stability of solitons for the Benjamin-Ono equation, *Rev. Mat. Iberoamericana*, Volume 25, Number 3 (2009), 909-970.

[27] C. Klein, J.-C. Saut, IST Versus PDE: A Comparative Study, *Fields Institute Communications*, Springer, Vol. 75, 383-449, 2015

[28] H. Koch and N. Tzvetkov, On the local well-posedness of the Benjamin-Ono equation in $H^s(\mathbf{R})$, *IMRN.*, (2003), 1449-1464.

[29] L. Molinet and D.Pilot, The Cauchy problem for the Benjamin-Ono equation in L^2 revisited, *Analysis and PDE* Vol. 5, No. 2, 2012

[30] L. Molinet, J.C. Saut, and N. Tzvetkov, Ill-posedness issues for the Benjamin-Ono and relatedequations, *SIAM J. Math. Anal.*, 33 (2001), 982-988.

[31] H. Ono. Algebraic solitary waves in stratified fluids. *J. Phys. Soc. Japan*, 39:1082-1091, 1975.

[32] G. Ponce, On the global well-posedness of the Benjamin-Ono equation, *Diff. Integral Eq.*, 4 (1991), 527-542.

[33] J.C. Saut, Sur quelques généralisations de l´equation de Korteweg -de Vries, *J. Math. Pures Appl.*, 58 (1979) 21-61.

[34] J. Shatah, Normal forms and quadratic nonlinear Klein-Gordon equations, *Comm. Pure Appl. Math.*, Vol. 38, (1985), 685-696.

[35] Tao, Terence, Multilinear weighted convolution of L2-functions, and applications to nonlinear dispersive equations. *Amer. J. Math.* 123 (2001), no. 5, 839908.

[36] T. Tao, Global well-posedness of the Benjamin-Ono equation in $^1(\mathbb{R})$, *J. Hyperbolic Differ. Equ.*, 1 (2004), no. 1, 2749.

[37] T. Tao, Global regularity of wave maps II. Small energy in two dimensions, *Commun. Math. Phys.*, 224, 443 544 (2001)

[38] D. Tataru, On global existence and scattering for the wave maps equation, *Amer. J. Math.*, 123 (2001), no. 3, 385–423

[39] S. Wu, Almost global wellposedness of the 2-D full water wave problem, *Invent. Math.*, 177, 2009, 45-135.

Department of Mathematics, University of California at Berkeley
E-mail address: ifrim@math.berkeley.edu

Department of Mathematics, University of California at Berkeley
E-mail address: tataru@math.berkeley.edu