

## DISPERSIVE DECAY FOR THE 1D KLEIN-GORDON EQUATION WITH VARIABLE COEFFICIENT NONLINEARITIES

JACOB STERBENZ

ABSTRACT. We study the 1D Klein-Gordon equation with variable coefficient nonlinearity. This problem exhibits an interesting resonant interaction between the spatial frequencies of the nonlinear coefficients and the temporal oscillations of the solutions. In the case when only the cubic coefficients are variable we prove  $L^\infty$  scattering and smoothness of the solution in weighted spaces with the help of both quadratic and cubic normal forms transformations. In the case of cubic interactions these normal forms appear to be novel.

### CONTENTS

1. Introduction	2081
2. Equations and coordinates	2083
3. Norms	2086
4. Main constructions and estimates	2087
5. Fourier analytic setup	2090
6. The quadratic correction	2103
7. The cubic correction	2105
Acknowledgements	2113
References	2113

### 1. INTRODUCTION

In this paper we initiate the study of 1D Klein-Gordon equations with variable coefficient nonlinearities. Specifically we investigate small compactly supported solutions to the problem:

$$(1) \quad (\square + 1)\phi = \alpha_0\phi^2 + \beta(x)\phi^3, \quad \square = \partial_t^2 - \partial_x^2,$$

where  $\beta(x) = \beta_0 + \beta_1(x)$ , with  $\beta_1$  a real valued Schwartz function, and  $\alpha_0, \beta_0 \in \mathbb{R}$ . To motivate this recall the two well known one dimensional nonlinear Klein-Gordon equations:

$$\square\phi - 2\phi + 2\phi^3 = 0 \quad (\phi^4 \text{ model}),$$

$$\square u + \sin(u) = 0 \quad (\text{Sin-Gordon equation}).$$

Both of these equations have static “kink” type solutions which are respectively (see [M-S]):

$$\phi_0 = \tanh(x), \quad u_0 = 4 \arctan(e^x).$$

---

Received by the editors July 31, 2013 and, in revised form, September 1, 2013, April 28, 2014, and May 22, 2014.

2010 *Mathematics Subject Classification*. Primary 35L70.

The author was supported in part by NSF grant DMS-1001675.

Linearization of their respective equations around these solutions leads to nonlinear Klein-Gordon equations of the form (1), albeit with potentials and nonlinear quadratic coefficients as well. The asymptotic stability problem of small solutions to such generalizations of (1) appears to be quite difficult, so in the present work we focus attention on the more modest goal of understanding the simplified model equation (1). Our main result here is the following:

**Theorem 1.1.** *Let  $\phi(t, x)$  be the global solution to the equation (1) with sufficiently small and smooth compactly supported initial data  $\phi(0) = \phi_0$  and  $\phi_t(0) = \phi_1$ . Then for  $t \geq 0$  the solution  $\phi(t, x)$  obeys the  $L^\infty$  estimate:*

$$(2) \quad |\phi(t, x)| \lesssim \frac{C(\phi_0, \phi_1)}{(1 + |\rho|)^{\frac{1}{2}}}, \quad \rho = |t^2 - x^2|^{1/2}.$$

Notice that solutions to the scalar equation (1) enjoy the conserved energy:

$$E(t) = \frac{1}{2} \int_{\mathbb{R}} (\phi_t^2 + \phi_x^2 + \phi^2 - \frac{2}{3} \alpha_0 \phi^3 - \frac{1}{2} \beta \phi^4)(t) dx,$$

so global existence from small energy class initial data is not an issue. On the other hand the dispersion rate for solutions to the 1D linear homogeneous Klein-Gordon equation  $(\square + 1)\phi = 0$  is only  $t^{-\frac{1}{2}}$ , placing the problem (1) well out of reach from the point of view of  $L^1 - L^\infty$  type dispersive (or Strichartz) estimates. Moreover, even using the strongest decay estimates one can muster based on a combination of the vector-field method of Klainerman [K1] and normal forms method of Shatah [Sh] the problem (1) is still out of reach unless one takes into account the long range behavior of the cubic nonlinearity. In fact a deep result of Delort [D1] shows that the global solutions to (1) with constant coefficients have the asymptotic behavior:

$$(3) \quad \phi(t, x) \sim \rho^{-1/2} e^{i\psi(\rho, x/\rho)} a(x/\rho) + \rho^{-1/2} e^{-i\psi(\rho, x/\rho)} \overline{a(x/\rho)},$$

where the phase is not linear but instead given by the expression:

$$\psi(\rho, x/\rho) = \rho - \left( \frac{5}{3} \alpha_0^2 + \frac{3}{2} \beta_0 \right) \sqrt{1 - (x/\rho)^2} |a(x/\rho)|^2 \ln \rho.$$

In particular this shows that any direct attempt to control solutions of (1) globally in time via Duhamel's principle for the linear Klein-Gordon equation must fail.

In the present paper we push beyond the analysis of [D1] to include the case of variable coefficient cubic nonlinearities of the form explained above. The main difficulty in this analysis stems from the fact that the vector-field technique, needed to handle the (nonlocalized) constant coefficient nonlinearities involving  $\alpha_0$  and  $\beta_0$ , is to some extent incompatible with coefficients depending on the spatial variable. To see this note that the linear Klein-Gordon equation  $\square + 1$  is not scale invariant and hence there is no vector-field that preserves free solutions based on scalings. The only useful weighted derivative that commutes with the linear flow appears to be the Lorentz boost  $\partial_y = t\partial_x + x\partial_t$ , and differentiation by it of coefficients depending on  $x$  leads to very badly diverging sources terms.

At first it might seem as if the lack of good vector-fields places (1) well out of reach of the classical methods of Klainerman and Shatah. However, if one pushes the method of Shatah in a new direction to include not only quadratic but variable coefficient cubic normal forms, then (1) can be rewritten in such a way that it is amenable to estimates involving the Lorentz boost. However, in order to make this analysis work we must develop a Littlewood-Paley type calculus as well as

bilinear  $\Psi$ DO operators and cubic paraproducts in terms of  $\partial_y$ , instead of the usual translation derivative  $\partial_x$ . This circle of ideas is interesting in its own right and we believe it has applications problems which are more general than what we consider here.

The method of the present paper owes much to previous work of Lindblad-Soffer [L-S1] and [L-S2] which gave a simplified proof of Delort’s global existence theorem from [D1] in the case when  $\alpha_0 = \beta_1 \equiv 0$ . This simplified method bypasses the phase corrections (3) and provides sharp  $L^\infty$  estimates directly in terms of an asymptotic form of the equation (1). In the context of our work this method is further clarified by writing the asymptotic equation in a paradifferential form with respect to  $\partial_y$  based frequency cutoffs, in which case the  $L^\infty$  estimates (and even the phase corrections if one wishes) become particularly transparent.

The remainder of our paper is organized as follows: In the next section we set up hyperboloid coordinates so that  $\partial_y$  becomes a coordinate derivative. Essentially all of the analysis in the remainder of the paper is carried out in these coordinates. At this point it is also natural to give a heuristic overview of the proof before getting into technical details.

In Section 3 we introduce the main function spaces we’ll use in our argument, and a few of their basic properties.

In Section 4 we state the quadratic and cubic normal forms transformations we use in an abstract form, as well as some nonlinear energy and  $L^\infty$  estimates, and then use these to derive Theorem 1.1.

In Section 5 we set up all the harmonic analysis used later in the normal forms constructions. This involves a number of lemmas whose proofs are for the most part independent of the nonlinear problem.

In Section 6 we construct and estimate our quadratic normal forms transformations, which are defined according to the classical method of Shatah [Sh].

In Section 7 we construct and estimate our variable coefficient cubic normal forms transformation.

**1.1. Basic notation.** We’ll use the standard notation  $A \lesssim B$ ,  $A \approx B$ ,  $A \gtrsim B$  to denote  $A \leq CB$ ,  $C^{-1}B \leq B \leq CB$ , and  $CA \geq B$  from some implicit  $C > 0$ . The notation  $\langle q \rangle = (1 + q^2)^{\frac{1}{2}}$  will be used for scalars and operators. We also set  $(a)_+ = \max\{a, 0\}$ .

If  $N$  is a Banach space with norm  $\| \cdot \|_N$  and  $L$  is an invertible linear operator on  $N$  we denote by  $LN$  the Banach space with norm  $\| u \|_{LN} = \| L^{-1}u \|_N$ . More generally if  $L : N_1 \rightarrow N_2$  is continuous then we’ll often denote this by the inclusion  $LN_1 \subseteq N_2$ , which is consistent with the previous convention if  $L$  is invertible.

In this work all analysis is done with respect to one spatial variable, thus unless otherwise stated  $L^p$ ,  $H^1$ , etc., denote  $L^p(\mathbb{R})$ ,  $H^1(\mathbb{R})$ , etc. We denote by  $\dot{H}^1$  the space with norm  $\| u \|_{\dot{H}^1} = \| \partial_x u \|_{L^2}$ , and by  $\mathcal{S} = \mathcal{S}(\mathbb{R})$  the one dimensional Schwartz space. We use the convention that for  $\varphi \in \mathcal{S}$  a quantity  $C(\| \varphi \|_{\mathcal{S}})$  denotes a positive constant depending on finitely many Schwartz seminorms of  $\varphi$ .

## 2. EQUATIONS AND COORDINATES

Inside the forward light cone  $t > |x|$  we set new coordinates:

$$(4) \quad x = \rho \sinh(y), \quad t = \rho \cosh(y), \quad \partial_y = t\partial_x + x\partial_t, \quad \partial_\rho = \rho^{-1}(t\partial_t + x\partial_x).$$

In the coordinates (4) the Minkowski metric takes the simple form  $-dt^2 + dx^2 = -d\rho^2 + \rho^2 dy^2$  and therefore, the linear Klein-Gordon operator may be written as:

$$(5) \quad \square + 1 = \partial_t^2 - \partial_x^2 + 1 = \partial_\rho^2 + \frac{1}{\rho} \partial_\rho - \frac{1}{\rho^2} \partial_y^2 + 1.$$

Conjugating the RHS above by  $\rho^{\frac{1}{2}}$  we have:

$$(6) \quad \rho^{\frac{1}{2}}(\square + 1) = (\square_{\mathcal{H}} + 1)\rho^{\frac{1}{2}}, \quad \square_{\mathcal{H}} = \partial_\rho^2 - \rho^{-2} \partial_y^2 + \frac{1}{4} \rho^{-2}.$$

In particular we may write the equation (1) as follows:

$$(7) \quad (\square_{\mathcal{H}} + 1)u = \rho^{-\frac{1}{2}} \alpha_0 u^2 + \rho^{-1} \beta (\rho \sinh(y)) u^3, \quad \text{where } u = \rho^{\frac{1}{2}} \phi.$$

Therefore, after translating the initial data problem forward in time by a bounded amount, to prove Theorem 1.1 it suffices to show:

**Theorem 2.1.** *Let  $u(\rho, y)$  be a global solution to the equation (7) with sufficiently small and smooth compactly supported initial data  $u(1) = u_0$  and  $u_\rho(1) = u_1$ . Then for  $\rho \geq 1$  one has the uniform  $L^\infty$  estimate  $|u| \lesssim C(u_0, u_1)$ .*

**2.1. Heuristic overview of the proof.** A naive approach to proving Theorem 2.1 would be to attempt uniform control in time of the Sobolev norm  $\|u(\rho)\|_{H_y^1}$ . We see from equation (7) this is almost possible if  $\alpha_0 = 0$  and  $\beta$  is constant. By Duhamel’s principle the contribution of the nonlinear terms to an  $L^\infty(H_y^1)$  bound for the solution would be  $\int_1^T \rho^{-1} \|u(\rho)\|_{H_y^1}^3 d\rho$ , which just fails. A closer inspection shows this failure is really only a matter of relatively low frequencies because one only needs  $H_y^{\frac{1}{2}+\epsilon}$  to control  $u$  in  $L_y^\infty$ . Thus, the only significant contribution to the Duhamel integral is  $\int_1^T \rho^{-1} \|P_{<c \ln(\rho)} u(\rho)\|_{H_y^1}^3 d\rho$  where  $P_{<c \ln(\rho)}$  denotes projection onto spatial frequencies  $|\xi| \lesssim \rho^c$ . To handle this low frequency logarithmic divergence one needs to directly manipulate the equation. By frequency projecting the evolution one finds:

$$(8) \quad (\square_{\mathcal{H}} + 1)P_{<c \ln(\rho)} u = \rho^{-1} \beta (P_{<c \ln(\rho)} u)^3 + \{\text{better terms}\},$$

where the “better terms” on the RHS are  $O(\rho^{-1-\frac{1}{10}c})$  and hence integrable. For small  $c$  the equation on the LHS above is essentially an ODE in time (thanks to  $\rho^{-2} \partial_y^2$ ), so one may obtain  $L^\infty$  estimates directly for  $P_{<c \ln(\rho)} u$  by simply using the almost Hamiltonian structure of  $\ddot{u} + u = \rho^{-1} \beta u^3$ . One just has to deal with small (integrable) errors when bounding  $\frac{1}{2}(\dot{u}^2 + u^2)(\rho)$ . Combining this  $L_y^\infty$  control for low frequencies with a slow growth estimate of the form  $\|u(\rho)\|_{H_y^1} \lesssim \epsilon \rho^\delta$  guarantees uniform  $L_y^\infty$  bounds for all of  $u$ , and this suffices to stabilize slow  $H_y^1$  growth via Duhamel’s principle and a bootstrapping argument.

Now up the ante to considering (7) with  $\alpha_0 \neq 0$  but  $\beta$  still constant. The only difference here is that we need to employ a suitable version of Shatah normal forms to remove the quadratic term. In doing so one produces additional cubic terms of the type  $\rho^{-1} \alpha_0^2 u \dot{u}^2$ , but these are harmless in the nonlinear frequency localized  $L_y^\infty$  estimates of the previous paragraph because one still retains the needed Hamiltonian structure. The implementation of Shatah normal forms in hyperbolic coordinates is more or less straightforward, and moreover has a nice semi-classical flavor due to the rescaling of the  $\partial_y$  derivatives in time.

Finally we come to the full problem considered here, which is to allow a local non-constant perturbation of  $\beta$  with respect to the spatial variable. At first it would appear that the method sketched in the previous paragraphs breaks down completely because one has  $x \approx \rho y$  in hyperbolic coordinates, so differentiation of  $\beta(x)$  by  $\partial_y$  leads to very badly diverging terms. However, after some frequency localizations one can show that the only troublesome contribution stems from the interactions  $\beta_{high}u_{low}^3$ . Asymptotically one can replace the low frequency factors with “plane waves”  $u_{low} \approx a_+e^{i\rho} + a_-e^{-i\rho}$ . This leads to an ansatz for the worst contribution from the nonlinearity which is roughly

$$\mathcal{N}_{cubic} = (\square_{\mathcal{H}} + 1)^{-1}(c_{\pm}\rho^{-1}\beta_{high}e^{\pm i\rho} + d_{\pm}\rho^{-1}\beta_{high}e^{\pm 3i\rho})$$

for various coefficient  $c_{\pm}, d_{\pm}$ . The nice feature of this equation is that it can be solved more or less explicitly and estimated via stationary phase calculations. The resulting expression  $\mathcal{N}_{cubic}$  can be thought of as a “cubic normal forms” correction and is given in terms of trilinear paraproducts applied to  $u$ . See formula (77) below. This expression carries all the relevant information about the resonant interaction between the spacial frequencies of  $\beta$  and the time oscillations of the solution  $u$ . With a little bit of work one can show that the remaining quantity  $w = u - \mathcal{N}_{cubic}$  is amenable to the direct  $H_y^1$  energy and nonlinear  $L^\infty$  estimates mentioned in the previous paragraphs. This allows one to close the proof.

In order to implement the strategy sketched above, we introduce a number of function spaces in the next section. The  $N$  space is used to bound source terms, and is simply a Duhamel type norm that allows one to conclude  $L_\rho^\infty(H_y^1)$  estimates. We also introduce two solution spaces  $S$  and  $\dot{S}$ . Since our  $L_\rho^\infty(H_y^1)$  estimates necessarily grow in time (again, due to low frequencies), one must add a sharp (nongrowing)  $L_y^\infty$  component to the solution norms. This latter part cannot be recovered directly in terms of the  $N$  norm (at least for low frequencies), but is instead bounded (for low frequencies) by directly integrating the equation as explained above. The  $\dot{S}$  norm is slightly weaker than the  $S$  norm and is used to control  $\dot{u}$ . Since this quantity occurs so many times when computing normal forms expressions, it deserves to have its own function space which helps to streamline many of the product estimates which occur in the sequel.

**2.2. A preliminary coefficient estimate.** We end this section with an elementary lemma that lets us write the nonlinearity of (7) in a more useful form:

**Lemma 2.2** (Coefficient approximation). *Let  $\varphi \in \mathcal{S}$ . Then if we set  $F(\rho, y) = \varphi(\rho y) - \varphi(\rho \sinh(y))$ , one has the estimate:*

$$(9) \quad \|\partial_\rho^n(\rho^{-1}\partial_y)^m F\|_{L_y^p} \leq C_{n,m}(\|\varphi\|_S)\rho^{-2-\frac{1}{p}}.$$

*Proof of the lemma.* The proof is by induction on the number of derivatives. When  $n = m = 0$  we have the bound  $|F(\rho, y)| \lesssim \rho^{-2}\langle \rho y \rangle^{-2}$  which is easy when  $|y| \geq \epsilon$ , and for  $|y| < \epsilon$  follows from the mean value theorem which gives:

$$F(\rho, y) = \varphi'(\xi(\rho, y))O(\rho y^3), \quad \text{where } |\xi(\rho, y)| \geq \rho|y|.$$

Taking first derivatives we have:

$$\partial_\rho F(\rho, y) = \rho^{-1}\tilde{F}(\rho, y), \quad \tilde{F}(\rho, y) = \tilde{\varphi}(\rho y) - \tilde{\varphi}(\rho \sinh(y)),$$

where  $\tilde{\varphi}(x) = x\varphi'(x)$ . And:

$$\begin{aligned} \rho^{-1}\partial_y F(\rho, y) &= (\varphi'(\rho y) - \varphi'(\rho \sinh(y))) - G(\rho, y), \\ G(\rho, y) &= \rho^{-2}(1 + \cosh(y))^{-1}\tilde{\tilde{\varphi}}(\rho \sinh(y)), \end{aligned}$$

where  $\tilde{\tilde{\varphi}}(x) = x^2\varphi'(x)$ . Then the proof concludes by induction and the easy estimate:

$$\|\partial_\rho^n(\rho^{-1}\partial_y)^m G\|_{L^p} \leq C_{n,m}(\|\tilde{\tilde{\varphi}}\|_S)\rho^{-2-\frac{1}{p}},$$

for  $G$  of the form above. □

### 3. NORMS

The main norms we work with in the paper are as follows.

**Definition 3.1** (Function spaces). For sufficiently smooth and well localized functions we define the solution and source spaces, where we use the shorthand  $\dot{u} = \partial_\rho u$ :

$$\|u\|_{S[1,T]} = \sup_{1 \leq \rho \leq T} \|u(\rho)\|_{S(\rho)},$$

where  $\|u(\rho)\|_{S(\rho)} = \rho^{-\delta} \|(u, \dot{u}, \rho^{-1}\partial_y u)(\rho)\|_{H_y^1} + \|(u, \dot{u})(\rho)\|_{L_y^\infty \cap L_y^2}$ ,

$$\|F\|_{N[1,T]} = \sup_{1 \leq \rho \leq T} \|F(\rho)\|_{N(\rho)},$$

where  $\|F(\rho)\|_{N(\rho)} = \rho^{1-\delta} \|F(\rho)\|_{H_y^1} + \rho \|F(\rho)\|_{L_y^\infty \cap L_y^2}$ .

We also introduce an auxiliary space to hold time derivatives of solutions to the Klein-Gordon equation:

$$\begin{aligned} \|u(\rho)\|_{\dot{S}(\rho)} &= \rho^{-\delta} \|u\|_{H_y^1} + \|u\|_{L_y^\infty \cap L_y^2} \\ &+ \inf_{v \in S(\rho)} (\rho^{-\delta} \|\dot{u} - \rho^{-2}\partial_y^2 v\|_{H_y^1} + \|\dot{u} - \rho^{-2}\partial_y^2 v\|_{L_y^\infty \cap L_y^2} + \|v\|_{S(\rho)}). \end{aligned}$$

Then define  $\|u\|_{\dot{S}[1,T]} = \sup_{1 \leq \rho \leq T} \|u(\rho)\|_{\dot{S}(\rho)}$  as usual.

These norms have the obvious uniform inclusions:

$$(10) \quad S(\rho) \subseteq \dot{S}(\rho), \quad \rho^{-1}\dot{S}(\rho) \subseteq N(\rho), \quad \rho^{-1}\partial_\rho S(\rho) \subseteq N(\rho),$$

which also directly implies the corresponding embeddings for  $S[1, T]$ ,  $\dot{S}[1, T]$ , and  $N[1, T]$ . There are a number of other basic estimates for norms that will come up later, so we collect them here.

**Lemma 3.2** (Basic properties of the norms). *The norms defined above have the following algebraic properties:*

$$(11) \quad \|FG\|_{N(\rho)} \lesssim \rho^{-1} \|F\|_{N(\rho)} \|G\|_{N(\rho)}, \quad \|(u, \dot{u})F\|_{N(\rho)} \lesssim \|u\|_{S(\rho)} \|F\|_{N(\rho)},$$

$$(12) \quad \|uv\|_{S(\rho)} \lesssim \|u\|_{S(\rho)} \|v\|_{S(\rho)}, \quad \|uv\|_{\dot{S}(\rho)} \lesssim \|u\|_{S(\rho)} \|v\|_{\dot{S}(\rho)}.$$

In addition, the  $S(\rho)$  and  $\dot{S}(\rho)$  norms are related as follows:

$$(13) \quad \begin{aligned} \|(1 - \rho^{-2}\partial_y^2)^{-\frac{1}{2}} u\|_{S(\rho)} &\lesssim \|u\|_{\dot{S}(\rho)}, \\ \|\dot{u}\|_{\dot{S}(\rho)} &\lesssim \|u\|_{S(\rho)} + \rho^{-\delta} \|(\square_{\mathcal{H}} + 1)u\|_{H_y^1} + \|(\square_{\mathcal{H}} + 1)u\|_{L_y^2 \cap L_y^\infty}. \end{aligned}$$

Finally, one has the Duhamel type estimate:

$$(14) \quad \sup_{1 \leq \rho \leq T} \rho^{-\delta} \|(u, \dot{u}, \rho^{-1}\partial_y u)(\rho)\|_{H_y^1} \lesssim \|(u_0, u_1)\|_{H_y^2 \times H_y^1} + \delta^{-1} \|F\|_{N[1,T]},$$

for solutions to the problem  $(\square_{\mathcal{H}} + 1)u = F$  with  $u(1) = u_0$  and  $\partial_\rho u(1) = u_1$ .

The proofs of the above estimates are for the most part immediate, with the exception of the second bound on line (12) and the first bound on line (13). These will be shown in Section 5.

*Remark 3.3.* Note that the embedding  $(\square_{\mathcal{H}} + 1)^{-1}N[1, T] \hookrightarrow S[1, T]$  is not uniform and grows like  $T^\delta$ . To overcome this obstacle we will need to use nonlinear estimates to recover the  $L_y^\infty \cap L_y^2$  norm of  $u$  at low frequencies.

Finally, we list here a nonlinear notation for norms. We let  $\mathcal{Q}_k(a, b, c, \dots)$  denote any polynomial which has a term of at least degree  $k$  in the first entry as a factor in each monomial. A typical example would be the function  $\mathcal{Q}_2(a, b) = (a^2 + a^4)(1 + b^2)$  which would then be used as:

$$(15) \quad \mathcal{Q}_2(\|u\|_{S[1,T]}, \|F\|_{S[1,T]}) = (\|u\|_{S[1,T]}^2 + \|u\|_{S[1,T]}^4)(1 + \|F\|_{S[1,T]}^2),$$

where  $F$  is some auxiliary quantity appearing in an estimate. The exact form of  $\mathcal{Q}$  may change from line to line. In one place in the sequel we'll also use  $\mathcal{Q}_{\frac{3}{2}}(\|u\|_{S[1,T]})$  with the obvious interpretation.

#### 4. MAIN CONSTRUCTIONS AND ESTIMATES

The main construction of this paper is the following:

**Theorem 4.1** (Normal forms corrections). *Let  $\alpha_0, \beta_0 \in \mathbb{R}$  and  $\beta_1 = \beta_1(\rho y)$  for some  $\beta_1(x) \in \mathcal{S}$ , and set  $\beta = \beta_0 + \beta_1$ . Let  $u$  and  $w$  solve the equations:*

$$(16) \quad (\square_{\mathcal{H}} + 1)u = \rho^{-\frac{1}{2}}\alpha_0 u^2 + \rho^{-1}\beta u^3 + F, \quad (\square_{\mathcal{H}} + 1)w = \rho^{-1}\beta_1 w^3 + G,$$

on the time interval  $[1, T]$ . Then there exist nonlinear expressions  $\mathcal{N}_{quad} = \mathcal{N}_{quad}(u, \dot{u})$  and  $\mathcal{R}_{quad} = \mathcal{R}_{quad}(u, \dot{u}, F, \dot{F})$  and  $\tilde{\mathcal{R}}_{quad} = \tilde{\mathcal{R}}_{quad}(u, \dot{u}, F, \dot{F})$ ,  $\mathcal{N}_{cubic} = \mathcal{N}_{cubic}(w, \dot{w})$  and  $\mathcal{R}_{cubic} = \mathcal{R}_{cubic}(w, \dot{w}, G, \dot{G})$ , such that one has the algebraic identities:

$$(17) \quad (\square_{\mathcal{H}} + 1)(u - \mathcal{N}_{quad}) = \rho^{-1}\beta u^3 + F + \mathcal{R}_{quad} \\ = \rho^{-1}(\beta + 2\alpha_0^2)u^3 - \frac{8}{3}\rho^{-1}\alpha_0^2\dot{u}^2u + F + \tilde{\mathcal{R}}_{quad},$$

$$(18) \quad (\square_{\mathcal{H}} + 1)(w - \mathcal{N}_{cubic}) = G + \mathcal{R}_{cubic}.$$

For the quadratic corrections one has the estimates:

$$(19) \quad \|\rho^{\frac{1}{2}}\mathcal{N}_{quad}\|_{S[1,T]} + \|(\mathcal{R}_{quad}, \dot{\mathcal{R}}_{quad})\|_{N[1,T]} + \|\tilde{\mathcal{R}}_{quad}\|_{L_\rho^1(L_y^\infty)[1,T]} \\ \lesssim \|\rho^{\frac{1}{2}}F\|_{S[1,T]}^2 + \|\rho^{\frac{1}{2}}\dot{F}\|_{S[1,T]}^2 + \|\rho(F, \dot{F})\|_{L_{\rho,y}^\infty[1,T]}^2 \\ + \mathcal{Q}_2(\|u\|_{S[1,T]}, \|\rho^{\frac{1}{2}}F\|_{S[1,T]}, \|\rho^{\frac{1}{2}}\dot{F}\|_{S[1,T]}, \|\rho(F, \dot{F})\|_{L_{\rho,y}^\infty[1,T]}).$$

In the case of the cubic corrections one has:

$$(20) \quad \|\mathcal{N}_{cubic}\|_{S[1,T]} + \|\mathcal{R}_{cubic}\|_{N[1,T]} \\ \lesssim \mathcal{Q}_2(\|w\|_{S[1,T]}, \|G\|_{N[1,T]}, \|\rho\dot{G}\|_{L_\rho^\infty(L_y^2)[1,T]}) + \mathcal{Q}_3(\|G\|_{N[1,T]}).$$

To this information we add the following nonlinear  $N \rightarrow S$  space bound:

**Proposition 4.2** (Nonlinear estimates). *Let  $a = a_0 + a_1(\rho y)$  and  $b = b_0 + b_1(\rho y)$  for real functions  $a, b$  with  $a_1, b_1 \in \mathcal{S}$  and  $a_0, b_0 \in \mathbb{R}$ . Let  $u, w$  solve the inhomogeneous nonlinear equations:*

$$(21) \quad (\square_{\mathcal{H}} + 1 + \rho^{-\frac{1}{2}}au + \rho^{-1}bu^2)u = F,$$

$$(22) \quad (\square_{\mathcal{H}} + 1 + \rho^{-1}aw^2 + \rho^{-1}bw^2)w = G,$$

on the interval  $[1, T]$ , with initial data  $u(1) = u_0$  and  $\partial_\rho u(1) = u_1$  and similarly for  $w$ . Then one has the nonlinear estimates:

$$(23) \quad \|(u, \dot{u})\|_{L^\infty(L^2_y)[1, T]} \lesssim \|(u_0, u_1)\|_{H^1_y \times L^2_y} + \mathcal{Q}_{\frac{3}{2}}(\|u\|_{S[1, T]}) + \|F\|_{L^1_\rho(L^2_y)[1, T]}.$$

$$(24) \quad \|(w, \dot{w})\|_{L^\infty_{\rho, y}[1, T]} \lesssim \|(w_0, w_1)\|_{H^2_y \times H^1_y} + \sup_{1 \leq \rho \leq T} \rho^{-c} \|w(\rho)\|_{S(\rho)} \\ + \mathcal{Q}_2(\|w\|_{S[1, T]}) + \|G\|_{L^1_\rho(L^\infty_y)[1, T]}.$$

for some sufficiently small  $c > 0$  in the second estimate.

Combining Theorem 4.1 and Proposition 4.2 we have the following a-priori bound which implies Theorem 2.1 through the usual local existence and continuity argument.

**Theorem 4.3** (Main a-priori estimate). *Let  $\beta = \beta_0 + \beta_1$  with  $\beta_1 \in \mathcal{S}$ , and let  $u$  solve the equation:*

$$(\square_{\mathcal{H}} + 1)u = \rho^{-\frac{1}{2}}\alpha_0 u^2 + \rho^{-1}\beta(\rho \sinh(y))u^3,$$

on  $[1, T]$  with initial data  $u(1) = u_0$  and  $\partial_\rho u(1) = u_1$ . Then one has the following a-priori estimate for some  $C \geq 2$ :

$$(25) \quad \|u\|_{S[1, T]} \lesssim \|(u_0, u_1)\|_{H^2_y \times H^1_y} + \|u\|_{S[1, T]}^{\frac{3}{2}} + \|u\|_{S[1, T]}^C.$$

*Proof that Theorem 4.1 and Proposition 4.2 imply Theorem 4.3.* The proof is in a series of steps.

*Step 1* (Coefficient reduction and the  $L^2_y$  estimate). First we write:

$$(\square_{\mathcal{H}} + 1)u = \rho^{-\frac{1}{2}}\alpha_0 u^2 + \rho^{-1}\beta(\rho y)u^3 + F,$$

where  $F = \rho^{-1}(\beta_1(\rho \sinh(y)) - \beta_1(\rho y))u^3$ . For the functions  $F(\rho, y, u)$  one has from Lemma 2.2 and the various product estimates and inclusions in Lemma 3.2 the collection of bounds:

$$(26) \quad \|F\|_{L^1_\rho(L^2_y \cap L^\infty_y)[1, T]} + \|\rho^{\frac{1}{2}}F\|_{S[1, T]} \\ + \|\rho^{\frac{1}{2}}\dot{F}\|_{\dot{S}[1, T]} + \|\rho(F, \dot{F})\|_{L^\infty_{\rho, y}[1, T]} + \|(F, \dot{F})\|_{N[1, T]} \\ \lesssim \mathcal{Q}_3(\|u\|_{S[1, T]}).$$

By combining the  $L^1L^2$  estimate for  $F$  above with the nonlinear energy estimate (23) we have:

$$(27) \quad \|(u, \dot{u})\|_{L^\infty(L^2_y)[0, T]} \lesssim \|(u_0, u_1)\|_{H^1_y \times L^2_y} + \mathcal{Q}_{\frac{3}{2}}(\|u\|_{S[1, T]}).$$

*Step 2* (Quadratic correction and the  $L_y^\infty$  estimate). Next, introduce  $w = u - \mathcal{N}_{quad}$  which solves the equation:

$$(\square_{\mathcal{H}} + 1)w = \rho^{-1}(\beta + 2\alpha_0^2)w^3 - \frac{8}{3}\rho^{-1}\alpha_0^2\dot{w}^2w + \tilde{F},$$

where

$$\tilde{F} = \rho^{-1}(\beta + 2\alpha_0^2)(u^3 - w^3) - \frac{8}{3}\rho^{-1}\alpha_0^2(\dot{u}^2u - \dot{w}^2w) + F + \tilde{\mathcal{R}}_{quad},$$

and where  $\tilde{\mathcal{R}}_{quad}$  is defined on the far RHS of line (17) and satisfies the relevant portion of estimate (19). Combining with the improved  $S$  norm bounds for  $\mathcal{N}_{quad} = u - w$  from line (19), and using the  $L^1L^\infty$  part of (26) as well, we have  $\|\tilde{F}\|_{L_\rho^1(L_y^\infty)[1,T]} \lesssim \mathcal{Q}_2(\|u\|_{S[1,T]})$ . Therefore by estimate (24) we control  $\|(w, \dot{w})\|_{L_{\rho,y}^\infty[1,T]}$ , and by a further application of the  $S$  estimate (19) for  $\mathcal{N}_{quad}$  (which also gives an estimate for the data  $\|(w_0, w_1)\|_{H_y^2 \times H_y^1}$  thanks to  $S(1) \subseteq H_y^2 \times H_y^1$ ) we have:

$$(28) \quad \|(u, \dot{u})\|_{L_{\rho,y}^\infty[1,T]} \lesssim \sup_{1 \leq \rho \leq T} \rho^{-c} \|u(\rho)\|_{S(\rho)} + \|(u_0, u_1)\|_{H_y^2 \times H_y^1} + \mathcal{Q}_2(\|u\|_{S[1,T]}).$$

*Step 3* (Cubic correction and the  $\rho^\delta H_y^1$  estimate). Finally, we rewrite the equation for  $w$  above as:

$$(\square_{\mathcal{H}} + 1)w = \rho^{-1}\beta_1w^3 + G,$$

where

$$G = \rho^{-1}\beta_1(u^3 - w^3) + \rho^{-1}\beta_0u^3 + F + \mathcal{R}_{quad}.$$

In order to invoke the cubic part of Theorem 4.1 we need to show the estimate:

$$(29) \quad \|G\|_{N[1,T]} + \|\rho\partial_\rho G\|_{L_\rho^\infty(L_y^2)[1,T]} \lesssim \mathcal{Q}_2(\|u\|_{S[1,T]}).$$

For the quantities  $F$  and  $\mathcal{R}_{quad}$  we use (26) and (19). For the  $N[1, T]$  norm of the remaining cubic expressions in  $u$  and  $w$  we use the inclusions on line (10) and the second algebra estimate from line (11), as well as estimate (19) for  $\mathcal{N}_{quad} = u - w$  combined with  $\rho^{-\frac{1}{2}}\beta_1 \in S(\rho)$ . The  $L^\infty L^2$  bound for  $\rho\partial_\rho$  applied to the cubic terms in  $G$  is immediate from the  $S$  norm bounds for  $u$  and  $w$ .

Now form the quantity  $v = w - \mathcal{N}_{cubic}$  which satisfies equation (18) and estimates (20). Then by applying (14) to the equation for  $v$  and using estimates (20) for  $\mathcal{R}_{cubic}$  and (29) for  $G$  in the resulting RHS, we control  $\|\rho^{-\delta}(v, \dot{v}, \rho^{-1}\partial_y v)(\rho)\|_{L_\rho^\infty(H_y^1)}$ . Finally, we can switch back to the original  $u$  via the estimates for  $\mathcal{N}_{cubic}$  on line (20) and the estimates for  $\mathcal{N}_{quad}$  on line (19). All together this gives:

$$(30) \quad \|\rho^{-\delta}(u, \dot{u}, \rho^{-1}\partial_y u)(\rho)\|_{L_\rho^\infty(H_y^1)[1,T]} \lesssim \|(u_0, u_1)\|_{H_y^2 \times H_y^1} + \mathcal{Q}_2(\|u\|_{S[1,T]}).$$

By adding this last bound to (27) and (28) above we have (25) for some  $C \geq 2$ .

*Step 4* (Combination). By adding this last bound to (27) and (28) above, we have for  $T_0 > 1$  sufficiently large the following estimate uniform in  $T \geq 1$ :

$$(31) \quad \|u\|_{S[1,T]} \lesssim \|u\|_{S[1,T_0]} + \|(u_0, u_1)\|_{H_y^2 \times H_y^1} + \mathcal{Q}_{\frac{3}{2}}(\|u\|_{S[1,T]}).$$

On the other hand from (30) and Sobolev embedding we easily have:

$$\|u\|_{S[1,T_0]} \lesssim T_0^\delta (\|(u_0, u_1)\|_{H_y^2 \times H_y^1} + \mathcal{Q}_2(\|u\|_{S[1,T]})).$$

Substituting this for the first term on RHS (31) yields (25) for some sufficiently large  $C \geq 2$ . □

5. FOURIER ANALYTIC SETUP

We assume the reader is familiar with basic Littlewood-Paley theory, which we only employ in the spatial variable. We let  $P_k$  be a smooth cutoff on frequency  $\xi \sim 2^k$ , and set  $P_{<k} = \sum_{j < k} P_j$ . We use  $u_k = P_k u$  and  $u_{<k} = P_{<k} u$  etc. Recall that  $P_k$  and  $P_{<k}$  are uniformly bounded on all  $L^p$  spaces. As a consequence it is easy to check that  $P_k$  and  $P_{<k}$  are bounded with respect to  $S$ ,  $\dot{S}$ , and  $N$  norms.

In the  $(\rho, y)$  variables there is no action at low frequencies for the evolution of (5), so we use the convention that  $P_0 = P_{<0}$  whenever we are in these variables. When in the  $(\rho, y)$  variables we'll use subscripts such as  $L^p_y, H^1_y$  to remind the reader of this. However, in some places we rescale back to  $x$  variables via  $x = \rho y$ , in which case we let  $P_k$  denote cutoff on  $\xi \sim 2^k$  for all  $k \in \mathbb{Z}$ .

Several basic things that will be used freely are:

$$\|P_k u\|_{L^p(\mathbb{R})} \lesssim 2^{\frac{1}{q} - \frac{1}{p}} \|P_k u\|_{L^q(\mathbb{R})} \text{ and } \|\partial_x P_k u\|_{L^p(\mathbb{R})} \approx 2^k \|P_k u\|_{L^p(\mathbb{R})},$$

for all  $1 \leq q \leq p \leq \infty$ . A direct consequence of these is the following bound that will be used many times in the sequel:

$$(32) \quad 2^{(k - \ln_2(\rho))^+} \|P_k u\|_{L^\infty_y} + \|P_k \dot{u}\|_{L^\infty_y} \lesssim \rho^\delta 2^{-\frac{1}{2}k} \|u\|_{S(\rho)}, \quad \|P_k u\|_{L^\infty_y} \lesssim \rho^\delta 2^{-\frac{1}{2}k} \|u\|_{\dot{S}(\rho)}.$$

In particular notice that this gives  $\|\rho^{-1} \partial_y u\|_{L^\infty_y} \lesssim \|u\|_{S(\rho)}$ , which will be used freely in the sequel.

Finally, we make the special notational convention that  $D_y = \frac{1}{i\rho} \partial_y$ , which can be thought of as a semi-classical derivative where  $h = \rho^{-1}$ . However, note this assignment is time dependent so one has the identity  $[\partial_\rho, D_y] = -\rho^{-1} D_y$ . We set  $D_x = \frac{1}{i} \partial_x, D_\rho = \frac{1}{i} \partial_\rho$ , etc., as usual.

**5.1. Multilinear estimates Part I: Paraproducts.** We list here a number of elementary product estimates for frequency localized functions. Our purpose is to set these up in a notational way that will be convenient for later use. The details are standard.

**Lemma 5.1** (Basic high-low product estimate). *Let  $k \in \mathbb{Z}$ . For test functions  $f_k, u_{<k+O(1)}^{(i)}$  and  $s_i \in \mathbb{N}$  one has:*

$$(33) \quad \|\partial_y^{s_0} (f_k \prod_{i=1}^n \partial_y^{s_i} u_{<k+O(1)}^{(i)})\|_{L^r} \lesssim \|(\partial_y)^{\sum s_i} f_k\|_{L^{p_0}} \prod_{i=1}^n \|u_{<k+O(1)}^{(i)}\|_{L^{p_i}},$$

where  $\sum \frac{1}{p_i} = \frac{1}{r}$ .

From this we derive a number of special cases which will be useful later.

**Lemma 5.2** (High-low  $S(\rho)$  and  $N(\rho)$  product estimates). *Let  $k > 0$ . For test functions  $f_k, u_{<k}^{(i)}, H_{<k+C}$  and  $s_i \in \mathbb{N}$  one has:*

$$(34) \quad \|\rho^{-1} f_k \prod_{i=1}^n u_{<k}^{(i)}\|_{S(\rho)} \lesssim \sum_{j=1}^2 \|(\rho^{-1} f_k, \rho^{-\delta} D_y^j f_k, \rho^{-\delta} D_y^{j-1} \dot{f}_k)\|_{L^\infty_y} \prod_{i=1}^n \|u_{<k}^{(i)}\|_{\dot{S}(\rho)},$$

$$(35) \quad \|\rho^{-1} f_k \prod_{i=1}^n D_y^{s_i} u_{<k}^{(i)}\|_{N(\rho)} \lesssim \|(D_y)^{\sum s_i} f_k\|_{L^\infty_y} \prod_{i=1}^n \|u_{<k}^{(i)}\|_{\tilde{S}(\rho)}, \text{ if } \sum s_i \geq 1,$$

(36)

$$\begin{aligned} & \|\rho^{-1} f_k H_{<k+C} \prod_{i=1}^n D_y^{s_i} u_{<k}^{(i)}\|_{N(\rho)} \\ & \lesssim \rho^{-\delta} \|(\rho^{2\delta-1} (D_y)^{\sum s_i} f_k, (D_y)^{1+\sum s_i} f_k)\|_{L_y^\infty} \prod_{i=1}^n \|u_{<k}^{(i)}\|_{\tilde{S}(\rho)} \|\rho H_{<k+C}\|_{L_y^2}. \end{aligned}$$

In the last two estimates we have set  $\tilde{S}(\rho) = \rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty$  at fixed time.

*Proof of Lemma 5.1.* By the Leibniz rule and a possible redefinition of the  $s_i$  it suffices to consider the product  $\partial_y^{s_0} f_k \cdot \prod_{i=1}^n \partial_y^{s_i} u_{<k+O(1)}^{(i)}$  instead. By Hölder’s inequality we may assume each of the  $s_i \geq 1$ , for  $i = 1, \dots, n$ . Breaking the  $u_{<k+O(1)}^{(i)}$  into dyadic frequencies and using Hölder’s inequality we have:

$$\begin{aligned} & \|\partial_y^{s_0} f_k \cdot \prod_{i=1}^n \partial_y^{s_i} u_{<k+O(1)}^{(i)}\|_{L^r} \\ & \lesssim \sum_{k_i < k+O(1)} 2^{\sum_{i=1}^n s_i k_i} \|\partial_y^{s_0} f_k\|_{L^{p_0}} \prod_{i=1}^n \|P_{k_i} u_{<k+O(1)}^{(i)}\|_{L^{p_i}}, \\ & \lesssim 2^{(\sum_{i=1}^n s_i)k} \|\partial_y^{s_0} f_k\|_{L^{p_0}} \prod_{i=1}^n \|u_{<k+O(1)}^{(i)}\|_{L^{p_i}}. \end{aligned}$$

The proof concludes by trading dyadic weights for derivatives. □

*Proof of estimate (34).* First consider the portion of the  $S(\rho)$  norm which does not contain  $\partial_\rho$  derivatives. Then the  $L_y^2 \cap L_y^\infty$  and  $\rho^\delta H_y^1$  estimates are all an immediate consequence of (33) by always putting  $f_k$  and its derivatives in  $L_y^\infty$ , one of the  $u_{<k}^{(i)}$  in either  $L_y^\infty$  or  $L_y^2$ , and the rest of the  $u_{<k}^{(i)}$  in  $L_y^\infty$ .

Now consider the portion of  $S(\rho)$  containing  $\partial_\rho$  derivatives. If  $\partial_\rho$  lands on  $f_k$  then the proof is similar to what was just discussed. If the  $\partial_\rho$  lands on a  $u_{<k}^{(i)}$  assume WLOG it is  $u_{<k}^{(1)}$ . To use the structure of the  $\dot{S}$  space we write  $\dot{u}_{<k}^{(1)} = P_{<k+C}(\dot{u}_{<k}^{(1)} + D_y^2 v) - D_y^2 v_{<k+C} = w_{<k+C} - D_y^2 v_{<k+C}$  where  $\|w_{<k+C}\|_{\rho^\delta H_y^1 \cap L_y^\infty \cap L_y^2} + \|v_{<k+C}\|_{S(\rho)} \lesssim \|u_{<k}^{(1)}\|_{\dot{S}(\rho)}$ . Therefore, modulo bounds which are again of the form just discussed we need to prove:

$$\begin{aligned} & \|\rho^{-1} f_k D_y^2 v_{<k+C} \prod_{i=2}^n u_{<k}^{(i)}\|_{\rho^\delta H_y^1 \cap L_y^\infty \cap L_y^2} \\ & \lesssim \rho^{-\delta} \|D_y f_k\|_{L_y^\infty} \|v_{<k+C}\|_{S(\rho)} \prod_{i=2}^n \|u_{<k}^{(i)}\|_{\dot{S}(\rho)}. \end{aligned}$$

In all cases the desired estimate again follows from (33) by putting  $D_y f_k$  in  $L_y^\infty$ ,  $D_y v_{<k+C}$  in  $L_y^\infty$  or  $L_y^2$ ,  $D_y^2 v_{<k+C}$  in  $L_y^2$ , and the rest of the  $u_{<k}^{(i)}$  in  $L_y^\infty$ . □

*Proof of estimate (35).* The estimate for the  $L_y^2 \cap L_y^\infty$  part of the  $N(\rho)$  norm follows at once from (33). In the case of the  $\dot{H}_y^1$  portion of the norm we need to use the

condition that  $\sum s_i \geq 1$ . WLOG assume that  $s_1 \geq 1$ . Then from (33) we have:

$$\rho^{-\delta} \|\partial_y(f_k \prod_{i=1}^n D_y^{s_i} u_{<k}^{(i)})\|_{L_y^2} \lesssim \|(D_y)^{\sum s_i} f_k\|_{L_y^\infty} \rho^{-\delta} \|\partial_y u_{<k}^{(1)}\|_{L_y^2} \prod_{i=2}^n \|u_{<k}^{(i)}\|_{L_y^\infty},$$

which is sufficient to achieve RHS (35). □

*Proof of estimate (36).* For the  $L_y^2$  and  $\rho^\delta H_y^1$  portions of the  $N(\rho)$  norms this again boils down to simple manipulations involving (33) and the details are left to the reader.

To prove (36) for the  $L^\infty$  portion of  $N(\rho)$  we combine (33) with a Sobolev embedding which gives:

$$\begin{aligned} & \|f_k H_{<k+C} \prod_{i=1}^n D_y^{s_i} u_{<k}^{(i)}\|_{L_y^\infty} \\ & \lesssim 2^{\frac{1}{2}k} \rho^{-1} \|(D_y)^{\sum s_i} f_k\|_{L_y^\infty} \prod_{i=1}^n \|u_{<k}^{(i)}\|_{L_y^\infty} \|\rho H_{<k+C}\|_{L_y^2}. \end{aligned}$$

Then RHS above is bounded in terms of RHS (36) by splitting into cases based on  $2^{\frac{1}{2}k} \leq \rho^\delta$  and vice versa. In the latter case we use  $2^{\frac{1}{2}k} \rho^{-1} P_k \lesssim \rho^{-\delta} D_y P_k$ . □

**5.2. Proof of the basic  $S, \dot{S}$  space properties.** The purpose of this subsection is to prove Lemma 3.2. The proofs of both estimates on line (11), the first estimate on line (12), the second estimate on line (13), and the estimate on line (14) are all elementary and left to the reader. Here we'll handle the second estimate on line (12) and the first estimate on line (13).

*Proof of the estimate  $\|\langle D_y \rangle^{-1} u\|_{S(\rho)} \lesssim \|u\|_{\dot{S}(\rho)}$ .* The main step is to show the operator  $\langle D_y \rangle^{-1}$  is bounded on all  $L_y^p$  spaces. After rescaling this is the same as showing that  $\langle D_x \rangle^{-1}$  is bounded on all  $L^p$  spaces which easily follows by showing the associated convolution kernel is in  $L^1$  (it has only a logarithmic singularity at the origin and is rapidly decaying). This shows that

$$\|\langle D_y \rangle^{-1} u\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty} \lesssim \|u\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty}.$$

The estimate for  $D_y \langle D_y \rangle^{-1} u$  in  $\rho^\delta H_y^1$  follows directly from Plancherel.

It remains to bound  $\partial_\rho \langle D_y \rangle^{-1} u$  in  $\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty$ . Modulo a commutator which produces terms like what we have already bounded in the previous paragraph, it suffices to consider  $\langle D_y \rangle^{-1} \dot{u}$ . Choose a  $v$  with  $\|\dot{u} + D_y^2 v\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty} + \|v\|_{S(\rho)} \lesssim \|u\|_{\dot{S}(\rho)}$ . Then we have:

$$\begin{aligned} \|\langle D_y \rangle^{-1} \dot{u}\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty} & \lesssim \|\langle D_y \rangle^{-1} (\dot{u} + D_y^2 v)\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty} \\ & \quad + \|\langle D_y \rangle^{-1} D_y^2 v\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty}. \end{aligned}$$

The first term on the RHS above is of the form of the previous paragraph. To handle the second we split  $v = v_{<\ln_2(\rho)} + v_{\geq \ln_2(\rho)}$ . For  $v_{<\ln_2(\rho)}$  use  $D_y^2 P_{<\ln_2(\rho)}$  is bounded on any  $L_y^p$  space. For  $v_{\geq \ln_2(\rho)}$  we use the embedding:

$$(37) \quad P_{\geq \ln_2(\rho)} : \rho^\delta H_y^1 \longrightarrow L_y^2 \cap L_y^\infty,$$

(which follows from (32)), and then the fact that  $D_y \langle D_y \rangle^{-1}$  is bounded on  $\rho^\delta H_y^1$ . □

*Proof of the estimate*  $\|uv\|_{\dot{S}(\rho)} \lesssim \|u\|_{S(\rho)}\|v\|_{\dot{S}(\rho)}$ . The bound for the portion of  $\dot{S}(\rho)$  which does not contain  $\dot{v}$  follows directly from the definition of the norms, the Leibniz rule, and Hölder’s inequality. Here we focus attention on showing there exists a  $w \in S(\rho)$  such that:

$$(38) \quad \|u\dot{v} + D_y^2 w\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty} + \|w\|_{S(\rho)} \lesssim \|u\|_{S(\rho)}\|v\|_{\dot{S}(\rho)}.$$

Let  $\tilde{w} \in S(\rho)$  such that  $\|\dot{v} + D_y^2 \tilde{w}\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty} + \|\tilde{w}\|_{S(\rho)} \lesssim \|v\|_{\dot{S}(\rho)}$ , then set  $w = \sum_{k \geq \ln_2(\rho)} u_{<k-C} \tilde{w}_k$ . Then  $\|w\|_{S(\rho)} \lesssim \|u\|_{S(\rho)}\|\tilde{w}\|_{S(\rho)}$  follows easily by using the product estimate (33) and orthogonality to gain an  $\rho^\delta H_y^1$  bound and then using (37). Now decompose the product  $u\dot{v}$  into a number of frequency interactions and estimate each term separately.

*Case 1* (The product  $u\dot{v}_{<\ln_2(\rho)}$ ). Here we compute:

$$\begin{aligned} \|u\dot{v}_{<\ln_2(\rho)}\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty} &\lesssim \|u(\dot{v}_{<\ln_2(\rho)} + D_y^2 \tilde{w}_{<\ln_2(\rho)})\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty} \\ &\quad + \|uD_y^2 \tilde{w}_{<\ln_2(\rho)}\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty}, \end{aligned}$$

and use the fact that  $\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty$  is an algebra followed by the boundedness of  $D_y^2 P_{<\ln_2(\rho)}$  on  $S(\rho)$  to achieve RHS (38).

*Case 2* (The product  $\sum_{k \geq \ln_2(\rho)} u_{<k-C} \dot{v}_k$ ). To bound this use orthogonality of the sum and the embedding (37), which reduces matters to bounding the RHS of:

$$\begin{aligned} \|u_{<k-C} \dot{v}_k + D_y^2(u_{<k-C} \tilde{w}_k)\|_{H_y^1} &\lesssim \|u_{<k-C}(\dot{v}_k + D_y^2 \tilde{w}_k)\|_{H_y^1} \\ &\quad + \|D_y^2(u_{<k-C} \tilde{w}_k) - u_{<k-C} D_y^2 \tilde{w}_k\|_{H_y^1}. \end{aligned}$$

The first term on the RHS is estimated by the Leibniz rule, while the second uses the fact that one derivative transfers to the low frequency term. Note that  $\|D_y P_{<k-C} u\|_{L^\infty} \lesssim \|u\|_{S(\rho)}$  follows easily from (32).

*Case 3* (The product  $\sum_{k \geq \ln_2(\rho)} u_{\geq k-C} \dot{v}_k$ ). For this term we first use the inequality:

$$(39) \quad \begin{aligned} \sum_{k \geq \ln_2(\rho)} \|u_{\geq k-C} \dot{v}_k\|_{H_y^1} &\lesssim \sum_{k \geq \ln_2(\rho)} \|u_{\geq k-C}(\dot{v}_k + D_y^2 \tilde{w}_k)\|_{H_y^1} \\ &\quad + \sum_{k \geq \ln_2(\rho)} \|u_{\geq k-C} D_y^2 \tilde{w}_k\|_{H_y^1}. \end{aligned}$$

Using paraproduct estimates similar to (33) and orthogonality we have the fixed frequency bounds:

$$\begin{aligned} \|u_{\geq k-C}(\dot{v}_k + D_y^2 \tilde{w}_k)\|_{H_y^1} &\lesssim \|u_{\geq k-C}\|_{H_y^1} \|\dot{v}_k + D_y^2 \tilde{w}_k\|_{L_y^\infty}, \\ \|u_{\geq k-C} D_y^2 \tilde{w}_k\|_{H_y^1} &\lesssim \|D_y u_{\geq k-C}\|_{H_y^1} \|D_y \tilde{w}_k\|_{L_y^\infty}. \end{aligned}$$

Using bounds similar to (32) we gain a factor of  $\rho^\delta 2^{-\frac{1}{2}k}$  from each RHS  $L_y^\infty$  term. Therefore:

$$LHS(39) \lesssim \sum_{k \geq \ln_2(\rho)} \rho^{2\delta} 2^{-\frac{1}{2}k} \|u\|_{S(\rho)}\|v\|_{\dot{S}(\rho)},$$

which suffices to give a uniform  $H_y^1$  bound in this case, and in particular one also gains control of  $L_y^2 \cap L_y^\infty$ . □

**5.3. Some estimates for coefficients.** We record here a number of asymptotic calculations for nonlinear coefficients such as  $\beta_1(\rho y)$ .

**Lemma 5.3** (An estimate for the cubic coefficient). *For functions of the form  $\varphi(\rho y)$  with  $\varphi \in \mathcal{S}$ , and for  $(n, m) \in \mathbb{N} \times \mathbb{Z}$ , we have for  $k > 0$  the uniform bound:*

$$(40) \quad \|\partial_\rho^n D_y^m P_k \varphi\|_{L_y^p} \leq C_{n,m,N} (\|\varphi\|_{\mathcal{S}}) \rho^{-n-m-1} 2^{(m+1-\frac{1}{p})k} 2^{-N(k-\ln_2(\rho))_+}.$$

*Proof of lemma.* We have  $D_y^m P_k \varphi(\rho y) = (2^k/\rho)^m \tilde{P}_k \varphi(\rho y)$  where the symbol of  $\tilde{P}_k$  has similar support and smoothness properties as  $P_k$ . Differentiation by  $\partial_\rho$  produces  $\partial_\rho D_y^m P_k \varphi(\rho y) = 2^{mk} \rho^{-m-1} \tilde{P}_k \tilde{\varphi}$  where  $\tilde{\varphi}(x) = -m\varphi(x) + x\varphi'(x)$ . Thus, by induction one has  $\partial_\rho^n D_y^m \varphi(\rho y) = \rho^{-n} (2^k/\rho)^m \tilde{P}_k \psi(\rho y)$  for some  $\psi \in \mathcal{S}$  and  $\tilde{P}_k$  with similar properties as  $P_k$ . Next, we have from change of variable and the  $L_x^1 \rightarrow L_x^p$  Bernstein's inequality:

$$\begin{aligned} \|\tilde{P}_k \psi(\rho y)\|_{L_y^p} &= \rho^{-\frac{1}{p}} \|\tilde{P}_{k-\ln_2(\rho)} \psi\|_{L_x^p} \lesssim \rho^{-\frac{1}{p}} 2^{-N(k-\ln_2(\rho))_+} \|\tilde{P}_{k-\ln_2(\rho)} \langle D_x \rangle^N \psi\|_{L_x^p}, \\ &\leq C_N (\|\psi\|_{\mathcal{S}}) \rho^{-1} 2^{(1-\frac{1}{p})k} 2^{-N(k-\ln_2(\rho))_+}. \end{aligned}$$

□

**Lemma 5.4** (A stationary phase calculation). *Let  $\lambda > 1$  and let  $h_\lambda(t)$  be a smooth function supported where  $t \approx \lambda$  with  $|\partial_t^n h_\lambda| \lesssim \lambda^{-n}$ . Let  $\varphi \in C_0^\infty$  and set:*

$$I_\pm(t, x) = \int_1^t \int e^{i(\pm(t-s)\langle \xi \rangle + 3s+x\xi)} h_\lambda(s) \varphi(\xi) d\xi ds.$$

*Then for fixed  $\epsilon > 0$  sufficiently small one has the asymptotic estimate:*

$$(41) \quad |D_t^n D_x^m I_\pm(t, x)| \leq C_{n,m,N}(\varphi) (\chi_{[-1+\epsilon, 1-\epsilon]}(x/t) + \langle (t, x) \rangle^{-N}).$$

*Here the estimates are uniform over any collection of functions  $h_\lambda$  which satisfy the above derivative bounds with a uniform constant.*

*Proof of lemma.* Note first that we may reduce to the case  $n = m = 0$ , because differentiation by  $D_x$  produces an integral of the same form, while differentiation by  $D_t$  produces an integral of the same form or a boundary term that is  $e^{i3t} h_\lambda(t) \hat{\varphi}(-x)$  which obviously satisfies estimate (41).

Next, we dispense with two easy special cases:

*Case 1* (The region  $|x/t| > 1 - \epsilon$  for  $0 < \epsilon \ll 1$ ). If  $\epsilon \ll 1$  is sufficiently small then for either  $I_\pm$  in this region one has  $|\partial_\xi \psi| \approx |x|$  where  $\psi$  denotes the phase. Thus  $|I_\pm| = O(\langle x \rangle^{-N})$  follows at once from integration by parts.

*Case 2* (The integral  $I_-$  where  $|x/t| < 1$ ). In this case the phase is nonstationary so one gets a uniform bound via integration by parts on time with respect to  $s$ .

Now we move to the main case:

*Case 3* (The integral  $I_+$  where  $|x/t| < 1$ ). Writing the phase as  $\psi(s, \xi; t, x)$  we have:

$$\psi_s = -\langle \xi \rangle + 3, \quad \psi_\xi = \langle \xi \rangle^{-1} (\xi(t-s) + \langle \xi \rangle x).$$

So the stationary points are  $\xi_\pm = \pm\sqrt{8}$  and  $s_\pm = t \pm \frac{3}{\sqrt{8}}x$ . A rough form of the Taylor expansion of  $\psi$  close to its stationary points is:

$$(42) \quad \psi(s, \xi; t, x) = \psi(s_\pm, \xi_\pm; t, x) + (s - s_\pm)(\xi - \xi_\pm) q_1(\xi) + x(\xi - \xi_\pm)^2 q_2(\xi),$$

where  $|q_2| \approx |q_1| \approx 1$  are smooth functions of  $\xi$ . To take advantage of this we break the integral up into two pieces  $I_+ = I_+^{sp} + I_+^{rem}$  where  $I_+^{sp}$  denotes the  $I_+$  integral with integrand multiplied by a cutoff  $\chi^{sp}(s, \xi; t, x)$  which is supported on the region:

$$\mathcal{A}^{sp} = \{|\xi - \xi_{\pm}| \ll 1\} \cap \{|s - s_{\pm}| \ll \langle x \rangle\} \cap \{|s - 1| \gtrsim \langle x \rangle^{\frac{1}{2}}\} \cap \{|t - s| \gtrsim \langle x \rangle^{\frac{1}{2}}\},$$

and chosen in such a way that  $|\partial_{\xi}^a \partial_s^b \chi^{sp}| \lesssim \langle x \rangle^{-\frac{1}{2}b}$ . The analysis proceeds via further sub-cases:

*Case 3a.1* (The estimate for  $I_+^{rem}$  when  $|\xi - \xi_{\pm}| \gtrsim 1$ ). This is similar to the estimate for  $I_-$  above, just integrate by parts one time with respect to  $\partial_s$ .

*Case 3a.2* (The estimate for  $I_+^{rem}$  when  $|\xi - \xi_{\pm}| \ll 1$  and  $|s - s_{\pm}| \gtrsim \langle x \rangle$ ). Here the rough Taylor expansion (42) shows that the differentiated phase  $\psi_{\xi}$  is  $O(s - s_{\pm})$ . Therefore, integrating by parts twice with respect to  $\xi$  yields an integrable weight of the form  $\langle s - s_{\pm} \rangle^{-2}$ .

*Case 3a.3.* The estimate for  $I_+^{rem}$  when  $|\xi - \xi_{\pm}| \ll 1$  and  $|s - s_{\pm}| \ll \langle x \rangle$ , and either  $|s - 1| \ll \langle x \rangle^{\frac{1}{2}}$  or  $|t - s| \ll \langle x \rangle^{\frac{1}{2}}$ . A standard stationary phase calculation in one variable shows:

$$\left| \int e^{i(\pm(t-s)\langle \xi \rangle + 3s + x\xi)} (1 - \chi^{sf}(s, \xi; t, x)) h_{\lambda}(s) \varphi(\xi) d\xi \right| \lesssim \langle x \rangle^{-\frac{1}{2}},$$

which is sufficient to integrate over this region.

*Case 3b* (The estimate for  $I_+^{sp}$ ). The form of the phase on line (42) shows that if we make the change of variables  $s' = \langle x \rangle^{-\frac{1}{2}}(s - s_{\pm})$  and  $\xi' = \langle x \rangle^{\frac{1}{2}}(\xi - \xi_{\pm})$  then in the new variables we have that the hessian  $\psi''$  is  $O(1)$  and nondegenerate at  $s' = \xi' = 0$ , and the phase may be written in the form:

$$\psi = \psi_0 + as'\xi' + b\xi'^2 + r(s', \xi'), \quad \text{where } a \neq 0,$$

and the remainder satisfies  $|\nabla^{\alpha} r(s', \xi')| \ll (|s'| + |\xi'|)^{2-|\alpha|}$  on the region  $\mathcal{A}^{sp}$  defined above. Moreover on this region in the  $(s', \xi')$  variables one has that the amplitude  $\Phi = \chi^{sp} h_{\lambda} \varphi$  obeys derivative bounds of the form  $|\nabla_{(s', \xi')}^{\alpha} \Phi| \lesssim 1$ . Thus, after a further uniformly smooth change of variables we may write:

$$I_+^{sp} = e^{i\psi_0} \iint e^{i(z_1^2 - z_2^2)} \Phi(z) dz, \quad \text{where } \Phi \in C_0^{\infty} \text{ and } |\nabla_z^{\alpha} \Phi| \lesssim 1,$$

and the uniform bound  $I_+^{sp} = O(1)$  follows from integration by parts away from the origin. □

**5.4. Proof of the nonlinear  $L^2$  and  $L^{\infty}$  estimates.** Here we prove Proposition 4.2. First note that (23) is a standard energy estimate and follows via multiplication of the equation (21) by  $\dot{u}$  and then integrating by parts. Details are left to the reader.

We obtain the  $L^{\infty}$  bound (24) in essentially the same way after a frequency localization. Notice that by estimate (32) and the form of RHS (24) we only need to establish this for frequencies  $k < c \ln_2(\rho)$ , as long as we choose  $c$  so that  $\delta \ll c \ll 1$ . Our goal then is to show:

$$(43) \quad \begin{aligned} \|(\dot{w}_{<c \ln_2(\rho)}, \dot{w}_{<c \ln_2(\rho)})\|_{L_{\rho, y}^{\infty}[1, T]} &\lesssim \|(\dot{w}_0, \dot{w}_1)\|_{H_y^2 \times H_y^1} + \sup_{1 \leq \rho \leq T} \rho^{-\tilde{c}} \|\dot{w}(\rho)\|_{S(\rho)} \\ &+ \mathcal{Q}_2(\|\dot{w}\|_{S[1, T]}) + \|G\|_{L_{\rho}^1(L_y^{\infty})[1, T]}, \end{aligned}$$

where  $\tilde{c} = \frac{1}{8}c$  and where our convention is  $\dot{w}_{<k} = P_{<k}\partial_\rho w$ . To this end we write the  $P_{<c \ln_2(\rho)}$  truncated version of (22) as:

$$(44) \quad (\square_{\mathcal{H}} + 1)w_{<c \ln_2(\rho)} + \rho^{-1}a_{<c \ln_2(\rho)+C}w^3_{<c \ln_2(\rho)} + \rho^{-1}b_{<c \ln_2(\rho)+C}\dot{w}^2_{<c \ln_2(\rho)}w_{<c \ln_2(\rho)} = \mathcal{R}_c(w, \dot{w}) + G_{<c \ln_2(\rho)}.$$

The proof of (43) is in a series of steps.

*Step 1* (Estimate for the remainder). Here we establish the preliminary estimate:

$$(45) \quad \|\mathcal{R}_c(w, \dot{w})\|_{L_y^\infty} \lesssim \rho^{-1-\frac{1}{2}c+\delta} \mathcal{Q}_1(\|w\|_{S(\rho)}).$$

This is shown for the various parts of  $\mathcal{R}_c$  separately.

*Case 1* (The term  $\partial_\rho^2, P_{<c \ln_2(\rho)}w$ ). We have

$$[\partial_\rho^2, P_{<c \ln_2(\rho)}]w = \rho^{-1}P'_{k=c \ln_2(\rho)+O(1)}\dot{w} + \rho^{-2}P''_{k=c \ln_2(\rho)+O(1)}w$$

for cutoffs  $P, P''$  with standard  $L^1$  kernels. Then (45) for this term follows from estimate (32).

*Case 2* (The term  $T = \rho^{-1}P_{<c \ln_2(\rho)}(aw^3) - \rho^{-1}a_{<c \ln_2(\rho)+C}w^3_{<c \ln_2(\rho)}$ ). We decompose this further into two subpieces by writing  $T = T_1 - T_2$  where:

$$T_1 = \rho^{-1}P_{<c \ln_2(\rho)}[a(w^3 - w^3_{<c \ln_2(\rho)})], \quad T_2 = \rho^{-1}P_{\geq c \ln_2(\rho)}(a_{<c \ln_2(\rho)+C}w^3_{<c \ln_2(\rho)}).$$

*Case 2a* (The bound for  $T_1$ ). Here we write  $w^3 - (w_{<k})^3 = (w_{\geq k})^3 + 3(w_{\geq k})^2w_{<k} + 3w_{\geq k}w^2_{<k}$ , then (45) follows from summing over the Sobolev estimate (32) for one of the  $w_{\geq c \ln_2(\rho)}$ .

*Case 2b* (The bound for  $T_2$ ). We may write  $a_{<c \ln_2(\rho)} = P_{k=c \ln_2(\rho)+O(C)}a_1 + P_{<c \ln_2(\rho)-C}a$ . Since we are assuming  $a_1 \in \mathcal{S}$  the coefficient bound (40) gives (45) for the first term on the RHS of this formula. To handle the second term we write:

$$P_{k \geq c \ln_2(\rho)}(a_{<c \ln_2(\rho)-C}w^3_{<c \ln_2(\rho)}) = P_{k \geq c \ln_2(\rho)}(a_{<c \ln_2(\rho)-C}(w^3_{<c \ln_2(\rho)} - w^3_{<c \ln_2(\rho)-C})),$$

and then we are again in the situation of Case 2.a above.

*Case 3* (The term  $\rho^{-1}P_{<c \ln_2(\rho)}(b\dot{w}^2w) - \rho^{-1}b_{<c \ln_2(\rho)+C}\dot{w}^2_{<c \ln_2(\rho)}w_{<c \ln_2(\rho)}$ ). This is similar to the previous case.

*Step 2* (The integral estimate). We multiply the equation (44) by  $\partial_\rho w_{<c \ln_2(\rho)}$  and then first integrate by parts in time followed by taking the sup in  $y$ . This yields the estimate:

$$(46) \quad \|(\partial_\rho w_{<c \ln_2(\rho)}, w_{<c \ln_2(\rho)})\|_{L_y^\infty}^2 \Big|_1^T \lesssim I_1 + I_2 + N_1 + N_2,$$

where:

$$I_1 = \sup_y \left| \int_1^T \rho^{-1}a_{<c \ln_2(\rho)+C}w^3_{<c \ln_2(\rho)}\partial_\rho w_{<c \ln_2(\rho)}d\rho \right|,$$

$$I_2 = \sup_y \left| \int_1^T \rho^{-1}b_{<c \ln_2(\rho)+C}\dot{w}^2_{<c \ln_2(\rho)}w_{<c \ln_2(\rho)}\partial_\rho w_{<c \ln_2(\rho)}d\rho \right|,$$

and:

$$N_1 = \|\rho^{-2}(w_{<c \ln_2(\rho)}, \partial_y^2 w_{<c \ln_2(\rho)}) \partial_\rho w_{<c \ln_2(\rho)}\|_{L^1_\rho(L^\infty_y)[1, T]},$$

$$N_2 = \|(\mathcal{R}_c, G_{<c \ln_2(\rho)}) \partial_\rho w_{<c \ln_2(\rho)}\|_{L^1_\rho(L^\infty_y)[1, T]}.$$

Our goal for the remainder of the proof is to show:

$$(47) \quad I_1 + I_2 + N_1 + N_2 \lesssim \epsilon \|\partial_\rho w_{<c \ln_2(\rho)}\|_{L^2_{\rho, y}[0, T]}^2 + \sup_{0 \leq \rho \leq T} \rho^{-2\tilde{c}} \|w(\rho)\|_{S(\rho)}^2 + \mathcal{Q}_4(\|w\|_{S[1, T]}) + \epsilon^{-1} \|G\|_{L^1_\rho(L^\infty_y)}^2,$$

which will be done for each term separately. With (47) in hand, estimate (43) follows from supping (46) over times in the interval  $[0, T]$ , taking  $\epsilon$  sufficiently small to absorb the first term on RHS (47), and then using commutator formula:

$$(48) \quad \dot{w}_{<c \ln_2(\rho)} = \partial_\rho w_{<c \ln_2(\rho)} - \rho^{-1} P'_{k=c \ln_2(\rho)+O(1)} w,$$

to return to the original  $\dot{w}_{<c \ln_2(\rho)}$  variable.

*Case 1* (The integral  $I_1$ ). Here we show  $I_1 \lesssim \mathcal{Q}_4(\|w\|_{S[1, T]})$ . To prove it we integrate by parts one time with respect to  $\rho$ , noting that the weight  $|\partial_\rho(\rho^{-1} a_{<c \ln_2(\rho)+C})| \lesssim \rho^{-2}$  is integrable.

*Case 2* (The integral  $I_2$ ). This term is a little more involved. First use the identity (for time dependent frequency cutoff and  $b_{<c \ln_2(\rho)+C}$  replaced by  $b$  to save notation):

$$\begin{aligned} \rho^{-1} b \dot{w}_{<k}^2 w_{<k} \partial_\rho w_{<k} &= \frac{1}{2} \partial_\rho (\rho^{-1} b \dot{w}_{<k}^2 w_{<k}^2) - \frac{1}{2} \partial_\rho (\rho^{-1} b) \dot{w}_{<k}^2 w_{<k}^2 \\ &\quad - \rho^{-1} b \dot{w}_{<k} \partial_\rho \dot{w}_{<k} w_{<k}^2. \end{aligned}$$

The integral of the first and second terms on the RHS above are directly bounded in terms of RHS (47). For the last term on RHS above we use the identity:

$$(49) \quad \begin{aligned} \partial_\rho \dot{w}_{<c \ln_2(\rho)} &= -(1 + 4^{-1} \rho^{-2}) w_{<c \ln_2(\rho)} + (\square_{\mathcal{H}} + 1) w_{<c \ln_2(\rho)} \\ &\quad - \partial_\rho [\partial_\rho, P_{<c \ln_2(\rho)}] w + \rho^{-2} \partial_y^2 w_{<c \ln_2(\rho)}. \end{aligned}$$

All of the terms on the RHS (49) yield integrable contributions when multiplied by  $\rho^{-1} b \dot{w}_{<c \ln_2(\rho)} w_{<c \ln_2(\rho)}^2$ , possibly after an addition integration by parts. We explain each case separately.

For the first term on RHS (49) multiplied by  $\rho^{-1} b \dot{w}_{<c \ln_2(\rho)} w_{<c \ln_2(\rho)}^2$  we either have an integrable contribution thanks to the weight  $\rho^{-2}$ , or we are back in the situation of Case 1 above after a use of identity (48) to write things as an absolutely integrable contribution plus a time derivative.

For the second term on RHS (49) multiplied by  $\rho^{-1} b \dot{w}_{<c \ln_2(\rho)} w_{<c \ln_2(\rho)}^2$  we have an integrable contribution by inspection of the terms in equation (44) and using estimate (45).

For the third term on RHS (49) multiplied by  $\rho^{-1} b \dot{w}_{<c \ln_2(\rho)} w_{<c \ln_2(\rho)}^2$  we expand:

$$\partial_\rho [\partial_\rho, P_{<c \ln_2(\rho)}] = \rho^{-1} P'_{k=c \ln_2(\rho)+O(1)} (\partial_\rho - \rho^{-1}) + \rho^{-2} P''_{k=c \ln_2(\rho)+O(1)},$$

which directly leads to absolutely integrable contributions.

Finally, for the last term on RHS (49) multiplied by  $\rho^{-1} b \dot{w}_{<c \ln_2(\rho)} w_{<c \ln_2(\rho)}^2$  we first sum over (32) to get:

$$(50) \quad \|\rho^{-2} \partial_y^2 P_{<c \ln_2(\rho)} w\|_{L^\infty} \lesssim \rho^{-2+\frac{3}{2}c+\delta} \|w\|_{S(\rho)}.$$

Again this leads to an absolutely integrable contribution.

Note that up to this point all of our estimates have been in terms of the last two expressions on RHS (47).

*Case 3* (The term  $N_1$ ). For this we use identity (48) and estimate (50) which directly gives the bound  $N_1 \lesssim \sup_{0 \leq \rho \leq T} \rho^{-c} \|w(\rho)\|_{S(\rho)}^2$ .

*Case 4* (The term  $N_2$ ). For this we directly use estimate (45) and Young’s inequality which gives:

$$N_2 \lesssim \epsilon \| \partial_\rho w_{<c \ln_2(\rho)} \|_{L_{\rho,y}^\infty[0,T]}^2 + \sup_{0 \leq \rho \leq T} \rho^{-2\tilde{c}} \|w(\rho)\|_{S(\rho)}^2 + \mathcal{Q}_4(\|w\|_{S[1,T]}) + \epsilon^{-1} \|G\|_{L_b^1(L_y^\infty)}^2.$$

**5.5. Multilinear estimates Part II: Semi-classical bilinear operators.** To set this up we define the  $h = \rho^{-1}$  semi-classical Fourier transform as:

$$\widehat{u}(\xi) = \mathcal{F}_\rho(v)(\xi) = \rho \int_{\mathbb{R}} e^{-i\rho\xi y} u(y) dy,$$

with inversion formula:

$$u(y) = \mathcal{F}_\rho^{-1}(\widehat{u})(y) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{i\rho y \xi} \widehat{u}(\xi) d\xi.$$

In particular the symbol of  $D_y$  is  $\xi$  in this picture. One must keep in mind that the operation  $u \rightarrow \widehat{u}$  is time dependent; in particular one has the identity:

$$(51) \quad [\partial_\rho, \mathcal{F}_\rho] = \rho^{-1} \partial_\xi \xi \mathcal{F}_\rho.$$

Next, given a symbol  $k(\xi, \eta)$  we use it to define a semi-classical bilinear operator as follows:

$$(52) \quad K({}^1D_y, {}^2D_y)[u, v] = \frac{1}{4\pi^2} \iint k(\xi, \eta) e^{i\rho y(\xi+\eta)} \widehat{u}(\xi) \widehat{v}(\eta) d\xi d\eta.$$

The class of symbols we will work with are the following.

**Definition 5.5** ( $S^{a,b;c}$  symbols). We say that a symbol  $k(\xi, \eta)$  is in  $S^{a,b;c}$  if  $k = p/q$  where  $q \approx (\langle \xi \rangle + \langle \eta \rangle)^c$  and:

$$(53) \quad |(\langle \xi \rangle \partial_\xi)^l (\langle \eta \rangle \partial_\eta)^m p| \lesssim C_{l,m} \langle \xi \rangle^a \langle \eta \rangle^b, \quad |(\langle \xi \rangle \partial_\xi)^l (\langle \eta \rangle \partial_\eta)^m q| \lesssim C_{l,m} (\langle \xi \rangle + \langle \eta \rangle)^c.$$

We say that  $k \in S_0^{a,b;c}$  if in addition to the above bounds one has  $k(0, 0) = 0$ .

Our main algebraic result for the operators  $K({}^1D_y, {}^2D_y)$  is the following.

**Lemma 5.6** (Bilinear  $\Psi$ DO Calculus). *Let  $k \in S^{a,b;c}$  and let  $K({}^1D_y, {}^2D_y)$  be the corresponding operator as defined on line (52). Then for all  $u, v \in \mathcal{S}$  one has the following identities:*

$$(54) \quad D_y K({}^1D_y, {}^2D_y)[u, v] = K({}^1D_y, {}^2D_y)[D_y u, v] + K({}^1D_y, {}^2D_y)[u, D_y v].$$

and

$$(55) \quad \partial_\rho K({}^1D_y, {}^2D_y)[u, v] = K({}^1D_y, {}^2D_y)[\partial_\rho u, v] + K({}^1D_y, {}^2D_y)[u, \partial_\rho v] - \rho^{-1} \partial_1 K({}^1D_y, {}^2D_y)[D_y u, v] - \rho^{-1} \partial_2 K({}^1D_y, {}^2D_y)[u, D_y v].$$

Here  $\partial_1 K$  and  $\partial_2 K$  have symbols  $\partial_\xi k$  and  $\partial_\eta k$  respectively.

*Proof.* The first identity follows at once by differentiation of the formula (52). The second formula follows as well by combining (52) with (51), and then integrating by parts with respect to  $\xi$ .  $\square$

Our main analytical result is:

**Proposition 5.7** (Estimates for Bilinear  $\Psi$ DO). *Let  $a, b, c \geq 0$  and let  $k \in S^{a,b;c}$  with  $K = K({}^1D_y, {}^2D_y)$  the corresponding operator defined by  $\mathcal{F}_\rho$ . Then one has the following mappings:*

$$(56) \quad K : S(\rho) \times S(\rho) \longrightarrow S(\rho), \quad \text{if} \quad a + b \leq 1 + c, \quad a \leq c, \quad b \leq c,$$

$$(57) \quad K : S(\rho) \times \dot{S}(\rho) \longrightarrow S(\rho), \quad \text{if} \quad a + b \leq c, \quad a \leq c, \quad b \leq c - 1,$$

$$(58) \quad K : \dot{S}(\rho) \times \dot{S}(\rho) \longrightarrow S(\rho), \quad \text{if} \quad a + b \leq c - 1, \quad a \leq c - 1, \quad b \leq c - 1.$$

Finally, if  $k \in S_0^{a,b;c}$  then under the same restrictions as above one has the improved  $L^\infty$  bound:

$$(59) \quad K : A \times B \longrightarrow \rho^{-\frac{1}{2}+\delta} L_y^\infty, \quad A = S(\rho), \dot{S}(\rho), \quad B = S(\rho), \dot{S}(\rho).$$

The last estimate of this proposition has a slightly more useful form that we will use in the sequel:

**Proposition 5.8.** *Let  $k \in S^{a,b;c}$  under the same restrictions as on lines (56)–(58). Then for  $A, B$  any combination of  $A = S(\rho), \dot{S}(\rho)$  and  $B = S(\rho), \dot{S}(\rho)$  we have:*

$$(60) \quad \| K({}^1D_y, {}^2D_y)[u, v] - k(0, 0)uv \|_{L_y^\infty} \lesssim \rho^{-\frac{1}{2}+\delta} \| u \|_A \| v \|_B.$$

To prove Proposition 5.7 we will rely on two lemmas.

**Lemma 5.9** (Dyadic Kernel Bounds). *Let  $k \in S^{a,b;c}$  with  $a, b, c \geq 0$ , and let  $K$  be the corresponding operator in the case  $\rho = 1$ . Then for all  $1 \leq p, q, r \leq \infty$  with  $p^{-1} + q^{-1} = r^{-1}$ , and integers  $0 \leq k \leq k'$ , one has the uniform family of bounds:*

$$(61) \quad K : P_k L^p \times P_{k'} L^q \longrightarrow 2^{ak} 2^{(b-c)k'} L^r.$$

Here we set  $P_k = P_{\leq 0}$  in the case  $k = 0$  and similarly for  $k'$ .

**Lemma 5.10** (Model Estimates). *Let  $k \in S^{a,b;c}$  with  $a, b, c \geq 0$  and such that  $a + b \leq c + 1$ ,  $a \leq c$ , and  $b \leq c$ . Let  $K$  be the corresponding operator in the case  $\rho = 1$ . Then one has the three embeddings:*

$$(62) \quad K : H^1 \times \dot{H}^1 \cap L^\infty \longrightarrow L^2,$$

$$(63) \quad K : L^2 \times \dot{H}^1 \cap \dot{H}^2 \cap L^\infty \longrightarrow L^2,$$

$$(64) \quad K : \dot{H}^1 \cap L^\infty \times \dot{H}^1 \cap \dot{H}^2 \cap L^\infty \longrightarrow L^\infty.$$

*Proof of Lemma 5.9.* The symbol of  $K({}^1D, {}^2D)[P_k u, P_{k'} v]$  is

$$l(\xi, \eta) = k(\xi, \eta) p_k(\xi) p_{k'}(\eta),$$

and one directly verifies that  $2^{-ak+(c-b)k'} l \in S^{0,0;0}$ , so it suffices to prove estimate (61) in this case.

Assuming now that  $k \in S^{0,0;0}$ , we have that the convolution kernel of the localized operator  $L({}^1D, {}^2D)$  is given by:

$$L({}^1D, {}^2D)(x, y) = \frac{1}{4\pi^2} \iint e^{i(x\xi+y\eta)} k(\xi, \eta) p_k(\xi) p_{k'}(\eta) d\xi d\eta,$$

so it suffices to show  $k \in S^{0,0;0}$  implies  $\|L\|_{L^1(dx dy)} \lesssim 1$  with a uniform bound depending on only finitely many of the seminorms (53). Integrating by parts two times when either  $|x| \geq 2^{-k}$  or  $|y| \geq 2^{-k'}$ , and doing nothing otherwise, and evaluating the resulting integral gives:

$$|L({}^1D, {}^2D)(x, y)| \lesssim 2^{k+k'} (1 + 2^k|x|)^{-2} (1 + 2^{k'}|y|)^{-2}.$$

The desired  $L^1$  bound follows. □

*Proof of Lemma 5.10.* The proof contains a number of steps and subcases. Here we use the convention that  $P_0 = P_{\leq 0}$  in all Littlewood-Paley decompositions.

*Step 1 (Low frequency factor).* We first dispense with the special case of the contribution of  $P_{\leq 0}v$ . Breaking  $u$  up into dyadic frequencies it suffices to show:

$$\begin{aligned} \left\| \sum_{k \geq 0} K({}^1D, {}^2D)[P_k u, P_{\leq 0}v] \right\|_{L^2} &\lesssim \|u\|_{L^2} \|v\|_{L^\infty}, \\ \left\| \sum_{k \geq 0} K({}^1D, {}^2D)[P_k u, P_{\leq 0}v] \right\|_{L^\infty} &\lesssim \|u\|_{L^\infty \cap \dot{H}^1} \|v\|_{L^\infty}. \end{aligned}$$

The first estimate follows from (61) and the fact that the LHS sum is essentially orthogonal. The second bound follows by breaking the sum into  $k < C$  and  $k > C$ , and using (61) in the  $L^\infty \cdot L^\infty$  form in both cases followed by:

$$\sum_{k > C} \|P_k u\|_{L^\infty} \lesssim \sum_{k > C} 2^{-\frac{1}{2}k} \|P_k u\|_{\dot{H}^1} \lesssim \|u\|_{\dot{H}^1},$$

in the case of  $P_{>C}u$ .

It now suffices to prove (62) and (63) when the second factor is  $P_{>0}v$ , and (64) in the case where the factors are  $P_{>0}u$  and  $P_{>0}v$  (by symmetry).

*Step 2 (Estimates (62) and (63) with  $P_{>0}v$ ).* Because of the frequency restriction on  $v$  we may assume its norm is  $H^1$  in (62), and  $H^2$  in (63). Now break the product up into a sum over all frequencies and group terms:

$$\begin{aligned} T_1 &= \sum_{\substack{k, k' \geq 0 \\ k > k' + C}} K({}^1D, {}^2D)[P_k u, P_{k'} v], & T_2 &= \sum_{\substack{k, k' \geq 0 \\ k' > k + C}} K({}^1D, {}^2D)[P_k u, P_{k'} v], \\ T_3 &= \sum_{\substack{k, k' \geq 0 \\ |k - k'| \leq C}} K({}^1D, {}^2D)[P_k u, P_{k'} v]. \end{aligned}$$

Now there are two subcases:

*Case 1 (High-low and low-high products).* Here we consider the sums  $T_1$  and  $T_2$ . Since the sum is essentially orthogonal in the high frequency factor we can reduce to a fixed high frequency. Putting the low frequency factor in  $L^\infty$  via Sobolev and using (61) we have:

$$(65) \quad \|K({}^1D, {}^2D)[P_k u, P_{k'} v]\|_{L^2} \lesssim 2^{\frac{3}{2} \min\{k, k'\}} \|P_k u\|_{L^2} \|P_{k'} v\|_{L^2}.$$

In the case of estimates (62) and (63) we bound the RHS above as (resp):

$$\begin{aligned} \text{RHS(65)} &\lesssim 2^{-\frac{1}{2} \min\{k,k'\}} 2^{-|k-k'|} \|P_k u\|_{H^1} \|P_{k'} v\|_{H^1}, \\ \text{RHS(65)} &\lesssim 2^{-\frac{1}{2} \min\{k,k'\}} \|P_k u\|_{L^2} \|P_{k'} v\|_{H^2}. \end{aligned}$$

In either case one may sum the low frequency term.

*Case 2 (High-high products).* In this case we bring the sum outside the  $L^2$  norm and then use (65) which directly gives  $\sum_k 2^{-\frac{1}{2}k}$  (*RHS norms*) for both (62) and (63).

*Step 3 (Estimate (64) with  $P_{>0}u$  and  $P_{>0}v$ ).* This is similar to the argument above. First break into all possibly frequency combinations and bring the resulting sum outside the norm. For fixed frequency we get from (61) and Sobolev embedding the estimate:

$$\begin{aligned} \|K({}^1D, {}^2D)[P_k u, P_{k'} v]\|_{L^\infty} &\lesssim 2^{\min\{k,k'\}} 2^{-\frac{1}{2}k} 2^{-\frac{3}{2}k'} \|P_k u\|_{H^1} \|P_{k'} v\|_{H^2} \\ &\lesssim 2^{-\frac{1}{2}|k-k'|} \|P_k u\|_{H^1} \|P_{k'} v\|_{H^2}. \end{aligned}$$

One may sum the RHS over all  $k, k' > 0$  using the  $\ell^2$  Young’s inequality and orthogonality. □

*Proof of Proposition 5.7.* The proof is in a series of steps.

*Step 1 (Reduction and rescaling).* First we note that by the first estimate on line (13) we only have to consider the  $S(\rho) \times S(\rho)$  case of estimates (56)–(59).

For this estimate we consider  $\dot{u}$  and  $\dot{v}$  as separate variables, which is permissible because (55) shows that differentiation by  $\partial_\rho$  preserves the symbol class of  $k(\xi, \eta)$ . Next, we rescale everything to  $\rho = 1$ . Setting new variables  $x = \rho y$  we have that  $D_y = D_x$  where  $D_x = \frac{1}{i}\partial_x$ , and the bilinear operators  $K({}^1D, {}^2D)$  are still defined by (52) but now with  $\rho = 1$ . The norm  $S(\rho)$  changes under rescaling to a new norm  $\tilde{S}(\rho)$ :

$$\|u\|_{\tilde{S}(\rho)} = \rho^{-\frac{1}{2}} \|(u, \dot{u})\|_{L_x^2} + \rho^{\frac{1}{2}-\delta} \|(u, \dot{u}, D_x u)\|_{\dot{H}_x^1} + \|(u, \dot{u})\|_{L_x^\infty}.$$

In the rescaled picture we are trying to show the two embeddings:

$$(66) \quad K : \tilde{S}(\rho) \times \tilde{S}(\rho) \longrightarrow \tilde{S}(\rho), \quad K : \tilde{S}(\rho) \times \tilde{S}(\rho) \longrightarrow \rho^{-\frac{1}{2}+\delta} L^\infty,$$

in the case  $k \in S^{a,b;c}$  with  $a + b \leq c + 1$ ,  $a \leq c$ , and  $b \leq c$  where in addition we assume  $k \in S_0^{a,b;c}$  for the second estimate. In the case of the first bound above we must also consider the case where one of the factors is  $\dot{u}$  or  $\dot{v}$ . We prove these bounds separately for each constituent of the  $\tilde{S}(\rho)$  norm.

*Case 1 (Contribution of the  $L^2$  part).* To estimate the  $L^2$  portion of the  $\tilde{S}(\rho)$  norm in the first bound on line (66) we use the estimates:

$$\rho^{-\frac{1}{2}} \|(u, \dot{u})\|_{H^1} \lesssim \|u\|_{\tilde{S}(\rho)}, \quad \|v\|_{L^\infty \cap \dot{H}^1} \lesssim \|v\|_{\tilde{S}(\rho)},$$

where the first bound is used for the time differentiated factor, and we conclude via estimate (62).

*Case 2 (Contribution of the  $\dot{H}^1$  part).* For the  $\dot{H}^1$  portion of the  $\tilde{S}(\rho)$  norm in the first bound on line (66) we need to differentiate with respect to each of  $D_x, D_x^2$ . In the case of a time differentiated factor we only need to apply  $D_x$ . There are two

cases depending on whether the derivatives split under Leibniz’s rule, or all go to one factor. In the first case we use:

$$\| (u, \dot{u}, D_x u) \|_{L^\infty \cap \dot{H}^1} \lesssim \| u \|_{\tilde{S}(\rho)}, \quad \rho^{-\delta + \frac{1}{2}} \| D_x u \|_{H^1} \lesssim \| u \|_{\tilde{S}(\rho)},$$

and then conclude via (62). In the second case we use:

$$\| u \|_{L^\infty \cap \dot{H}^1 \cap \dot{H}^2} \lesssim \| u \|_{\tilde{S}(\rho)}, \quad \rho^{-\delta + \frac{1}{2}} \| (D_x u, D_x^2 u, D_x \dot{u}) \|_{L^2} \lesssim \| u \|_{\tilde{S}(\rho)},$$

and then conclude via (63).

*Case 3* (Contribution of the  $L^\infty$  part). To estimate the  $L^\infty$  portion of the  $\tilde{S}(\rho)$  norm in the first bound on line (66) we use:

$$\| u \|_{L^\infty \cap \dot{H}^1 \cap \dot{H}^2} \lesssim \| u \|_{\tilde{S}(\rho)}, \quad \| \dot{u} \|_{L^\infty \cap \dot{H}^1} \lesssim \| u \|_{\tilde{S}(\rho)},$$

and then conclude via (64).

*Step 2* (Proof of the improved  $L^\infty$  norm). To prove the second bound on line (66) it suffices to consider instead  $P_{\leq 0} K({}^1D, {}^2D)[u, v]$  because we already have the  $\tilde{S}(\rho)$  estimate and can use  $P_{> 0} \dot{H}^1 \subseteq L^\infty$  for the complement. Breaking up the expression into frequencies we have:

$$(67) \quad \begin{aligned} P_{\leq 0} K({}^1D, {}^2D)[u, v] &= P_{\leq 0} K({}^1D, {}^2D)[P_{< C} u, P_{< C} v] \\ &+ \sum_{\substack{\max\{k, k'\} \geq C \\ k = k' + O(1)}} P_{\leq 0} K({}^1D, {}^2D)[P_k u, P_{k'} v]. \end{aligned}$$

For the second term on RHS we use embedding (64) and the fact that for high frequencies one has:

$$\rho^{-\delta + \frac{1}{2}} \| P_k u \|_{L^\infty \cap \dot{H}^1 \cap \dot{H}^2} \lesssim \| u \|_{\tilde{S}(\rho)}, \quad \text{for } k \geq C.$$

The frequency sum is then bounded by  $\ell^2$  Cauchy-Schwartz and orthogonality.

For the low frequency part on RHS (67) we need to use the condition  $k \in S_0^{a,b;c}$ . Assume WLOG that the vanishing condition is due to the first factor, in which case the symbol of the total operator is  $p_0(\xi + \eta)k(\xi, \eta)p_{< C}(\xi)p_{< C}(\eta) = \tilde{k}(\xi, \eta)p_{< C}(\xi)p_{< C}(\eta)$  for some smooth compactly supported  $\tilde{k}$ . Thus:

$$\| P_{\leq 0} K({}^1D, {}^2D)[P_{< C} u, P_{< C} v] \|_{L^\infty} \lesssim \| D_x P_{< C} u \|_{L^\infty} \| v \|_{L^\infty},$$

and we conclude with  $D_x P_{< C} \dot{H}^1 \subseteq L^\infty$  and  $\rho^{\frac{1}{2} - \delta} \tilde{S}(\rho) \subseteq \dot{H}^1$ . □

Finally, we prove Proposition 5.8.

*Proof.* For this proof we switch back to the  $y$  variable. Given any symbol  $k \in S^{a,b;c}$  the new symbol  $\tilde{k}(\xi, \eta) = k(\xi, \eta) - k(0, 0)p_{< 0}(\xi)p_{< 0}(\eta)$  is such that  $\tilde{k} \in S_0^{a,b;c}$  with the same indices. Therefore, by estimate (59) it suffices to show that:

$$\| uv - P_{< \ln_2(\rho)} u P_{< \ln_2(\rho)} v \|_{L^\infty} \lesssim \rho^{-\frac{1}{2} + \delta} \| u \|_A \| v \|_B,$$

for any combination of the cases  $A = S(\rho), \dot{S}(\rho)$  and  $B = S(\rho), \dot{S}(\rho)$ . Expanding the difference it suffices to show (by symmetry):

$$\| P_{\geq \ln_2(\rho)} u \|_{L^\infty} \lesssim \rho^{-\frac{1}{2} + \delta} \| u \|_A, \quad A = S(\rho), \dot{S}(\rho).$$

This follows at once from (32). □

6. THE QUADRATIC CORRECTION

The main result of this section is the following:

**Theorem 6.1** (Quadratic normal forms). *Let  $u$  be a sufficiently smooth and well localized solution to:*

$$(68) \quad (\square_{\mathcal{H}} + 1)u = \rho^{-\frac{1}{2}}\alpha_0 u^2 + \rho^{-1}\beta u^3 + F,$$

on the time interval  $[1, T]$ . Then there exist three nonlinear quantities  $\mathcal{N}_{quad} = \mathcal{N}_{quad}(u, \dot{u})$ ,  $\mathcal{R}_{quad} = \mathcal{R}_{quad}(u, \dot{u}, F, \dot{F})$ , and  $\tilde{\mathcal{R}}_{quad} = \tilde{\mathcal{R}}_{quad}(u, \dot{u}, F, \dot{F})$  such that one has the algebraic identity:

$$(69) \quad (\square_{\mathcal{H}} + 1)\mathcal{N}_{quad} = \rho^{-\frac{1}{2}}\alpha_0 u^2 + \mathcal{R}_{quad} = \rho^{-\frac{1}{2}}\alpha_0 u^2 - 2\rho^{-1}\alpha_0^2 u^3 + \frac{8}{3}\rho^{-1}\alpha_0^2 \dot{u}^2 u + \tilde{\mathcal{R}}_{quad}.$$

In addition one has the estimates:

$$(70) \quad \begin{aligned} & \|\rho^{\frac{1}{2}}\mathcal{N}_{quad}\|_{S[1, T]} + \|(\mathcal{R}_{quad}, \dot{\mathcal{R}}_{quad})\|_{N[1, T]} + \|\tilde{\mathcal{R}}_{quad}\|_{L^1_{\rho}(L^{\infty}_{\dot{y}})[1, T]} \\ & \lesssim \|\rho^{\frac{1}{2}}F\|_{S[1, T]}^2 + \|\rho^{\frac{1}{2}}\dot{F}\|_{\dot{S}[1, T]}^2 + \|\rho(F, \dot{F})\|_{L^{\infty}_{\rho, y}[1, T]}^2 \\ & \quad + \mathcal{Q}_2(\|u\|_{S[1, T]}, \|\rho^{\frac{1}{2}}F\|_{S[1, T]}, \|\rho^{\frac{1}{2}}\dot{F}\|_{\dot{S}[1, T]}, \|\rho(F, \dot{F})\|_{L^{\infty}_{\rho, y}[1, T]}). \end{aligned}$$

We construct  $\mathcal{N}_{quad}$  according to the classical method of Shatah [Sh]:

$$(71) \quad \mathcal{N}_{quad} = \rho^{-\frac{1}{2}}K_0[u, u] + \rho^{-\frac{1}{2}}K_2[\dot{u}, \dot{u}],$$

where  $K_0, K_2$  are bilinear operators as defined on line (52), which we also assume to be symmetric. To compute them we use the notation  $[\partial_{\rho}^2, \rho^{-\frac{1}{2}}K]$  to denote the operator:

$$[\partial_{\rho}^2, \rho^{-\frac{1}{2}}K][u, v] = \partial_{\rho}^2(\rho^{-\frac{1}{2}}K[u, v]) - \rho^{-\frac{1}{2}}K[\ddot{u}, v] - \rho^{-\frac{1}{2}}K[u, \ddot{v}] - 2\rho^{-\frac{1}{2}}K[\dot{u}, \dot{v}].$$

Then a short calculation shows that for  $\mathcal{N}_{quad}$  as defined above:

$$(\square_{\mathcal{H}} + 1)\mathcal{N}_{quad} = T_1 + T_2 + T_3 - \frac{1}{4}\rho^{-2}\mathcal{N}_{quad},$$

where if we set  $G = (\square_{\mathcal{H}} + 1)u$  then:

$$\begin{aligned} T_1 &= \rho^{-\frac{1}{2}}\left[2K_2[(D_y^2 + 1)u, (D_y^2 + 1)u] + 2K_0[D_y u, D_y u] + 2K_2[D_y \dot{u}, D_y \dot{u}] \right. \\ & \quad \left. - K_0[u, u] + 2K_0[\dot{u}, \dot{u}] - K_2[\dot{u}, \dot{u}]\right], \\ T_2 &= \rho^{-\frac{1}{2}}\left[2K_0[G, u] + 2K_2[G, G] + 2K_2[\dot{G}, \dot{u}] - 4K_2[(D_y^2 + 1)u, G] - \rho^{-2}K_2[G, u]\right], \\ T_3 &= [\partial_{\rho}^2, \rho^{-\frac{1}{2}}K_0][u, u] + [\partial_{\rho}^2, \rho^{-\frac{1}{2}}K_2][\dot{u}, \dot{u}] + 4\rho^{-\frac{3}{2}}K_2[D_y^2 u, \dot{u}] + \rho^{-\frac{7}{2}}K_2[u, \dot{u}] \\ & \quad + \rho^{-\frac{5}{2}}K_2[(D_y^2 + 1)u, u] + \frac{1}{8}\rho^{-\frac{9}{2}}K_2[u, u]. \end{aligned}$$

The term  $T_1$  is no better than the original quadratic nonlinearity, so one chooses the  $K_i$  specifically to achieve  $T_1 = \rho^{-\frac{1}{2}}\alpha_0 u^2$ . A standard calculation (see Section 7.8 of [H1]) shows that to do this one must define the symbols of  $K_0$  and  $K_2$  as follows:

$$k_0 = \alpha_0(1 - 2\xi\eta)q^{-1}, \quad k_2 = 2\alpha_0 q^{-1}, \quad \text{where } q = (4\xi^2 + 4\eta^2 + 4\xi\eta + 3).$$

Notice that these symbols are such that  $k_0 \in S^{1,1;2}$  and  $k_2 \in S^{0,0;2}$ . To estimate  $\mathcal{N}_{quad}$  and the error terms  $T_2$  and  $T_3$  listed above we will use:

**Proposition 6.2** (Quadratic NF estimates). *Let  $(\square_{\mathcal{H}}+1)u = \rho^{-\frac{1}{2}}\alpha_0u^2 + \rho^{-1}\beta u^3 + F := G$ . Then one has the estimate:*

$$(72) \quad \begin{aligned} &\rho^{\frac{1}{2}}\|G\|_{S(\rho)} + \rho^{\frac{1}{2}}\|\dot{G}\|_{\dot{S}(\rho)} + \|\dot{u}\|_{\dot{S}(\rho)} + \|\ddot{u} + D_y^2u\|_{S(\rho)} \\ &\lesssim \mathcal{Q}_1(\|u\|_{S(\rho)}, \|F\|_{S(\rho)}) + \|\rho^{\frac{1}{2}}F\|_{S(\rho)} + \|\rho^{\frac{1}{2}}\dot{F}\|_{\dot{S}(\rho)}. \end{aligned}$$

In addition if we set  $G = \rho^{-\frac{1}{2}}\alpha_0u^2 + \rho^{-1}\beta u^3 + F$  in  $T_2$  and  $T_3$  above, then one has the fixed time estimates:

$$(73) \quad \rho\|T_2\|_{S(\rho)} + \rho^{\frac{3}{2}}\|T_3\|_{S(\rho)} \lesssim (RHS)(70),$$

$$(74) \quad \|T_2 + \rho^{-1}(2\alpha_0^2u^3 - \frac{8}{3}\alpha_0^2\dot{u}^2u)\|_{L_y^\infty} \lesssim \rho^{-\frac{3}{2}+\delta}(RHS)(70).$$

First we shall use this Proposition to demonstrate Theorem 6.1.

*Proof that Proposition 6.2 implies Theorem 6.1.* With the choice (71) the estimate for  $\rho^{\frac{1}{2}}\mathcal{N}_{quad}$  on line (70) is an immediate consequence of (56) and (58) and the estimate for  $\dot{u}$  on line (72).

The identity (69) is satisfied if we define:

$$\begin{aligned} \mathcal{R}_{quad} &= T_2 + T_3 - \frac{1}{4}\rho^{-2}\mathcal{N}_{quad}, \\ \tilde{\mathcal{R}}_{quad} &= (2\rho^{-1}\alpha_0^2u^3 - \frac{8}{3}\rho^{-1}\alpha_0^2\dot{u}^2u + T_2) + T_3 - \frac{1}{4}\rho^{-2}\mathcal{N}_{quad}. \end{aligned}$$

Then the estimate for  $(\mathcal{R}_{quad}, \tilde{\mathcal{R}}_{quad})$  on line (70) follows immediately from the second and third embeddings on line (10), estimate (73), and the estimate for  $\mathcal{N}_{quad}$  already shown. The estimate for  $\tilde{\mathcal{R}}_{quad}$  on line (70) is immediate from the estimate for  $T_3$  on line (73), estimate (74), and the estimate for  $\mathcal{N}_{quad}$ .  $\square$

Now we prove Proposition 6.2 which is done separately for the various estimates involved. First the preliminary bounds.

*Proof of estimate (72).* We consider each term on the LHS of (72) separately.

*Case 1* (The  $S(\rho)$  estimate for  $G$ ). For the part of  $G$  which contains  $F$  this is immediate. For the quadratic terms in  $G$  the bound is also immediate from the first algebra estimate on line (12). For the cubic part of  $G$  we can again use (12) once we know  $\rho^{-\frac{1}{2}}\beta_1(\rho y) \in S(\rho)$  uniformly. The main thing to check is the  $H_y^1$  portion of the  $S(\rho)$  norm which has the worst behavior. For this we compute:

$$\|\beta_1\|_{H_y^1} \lesssim \rho\|(\beta_1, \beta'_1)\|_{L_y^2} \lesssim \rho^{\frac{1}{2}}, \quad \text{where } \beta'_1 = (\partial_x\beta_1)(\rho y),$$

which suffices. Similar  $H_y^1$  bounds for  $\partial_\rho(\rho^{-\frac{1}{2}}\beta_1)$  and  $D_y(\rho^{-\frac{1}{2}}\beta_1)$  are immediate.

*Case 2* (The  $\dot{S}(\rho)$  estimate for  $\dot{u}$ ). Using the second bound on line (13), this boils down to estimating  $\|G\|_{\rho^\delta H_y^1 \cap L_y^2 \cap L_y^\infty}$  which is already contained in the  $S(\rho)$  bound for  $G$  above.

*Case 3* (The  $S(\rho)$  estimate for  $\ddot{u} + D_y^2u$ ). Since  $\ddot{u} + D_y^2u = G - (1 + \frac{1}{4}\rho^{-2})u$  the bound follows from the  $S(\rho)$  estimates already proved for  $G$  and  $u$ .

Case 4 (The  $\dot{S}(\rho)$  estimate for  $\dot{G}$ ). Again it suffices to focus attention on the nonlinear terms. Differentiating the quadratic and cubic terms in  $G$  with respect to  $\partial_\rho$  and using the second algebra estimate on line (12), the desired bound follows at once from the  $\dot{S}(\rho)$  estimate for  $\dot{u}$  and the  $S(\rho)$  estimate for  $\rho^{-\frac{1}{2}}\beta_1$  proved previously.  $\square$

Finally, we prove the error bounds for  $T_2$  and  $T_3$  in Proposition 6.2.

*Proof of estimates (73) and (74).* There are a number of cases.

Case 1 (The  $S(\rho)$  estimate for  $T_2$  on line (73)). This follows immediately from the  $S(\rho)$  and  $\dot{S}(\rho)$  bounds for  $G, \dot{G}$ , and  $\dot{u}$  on line (72) and the bilinear estimates (56)–(58).

Case 2 (The improved  $L_y^\infty$  estimate for  $T_2$  on line (74)). Here we use estimate (60) which shows that in light of (72) we already have estimate (74) for the expression  $T_2 - \rho^{-\frac{1}{2}}\alpha_0[-2Gu + \frac{4}{3}\dot{G}\dot{u} + \frac{4}{3}G^2]$ . Thus, we only need to prove the bound:

$$\begin{aligned} & \|\rho^{-\frac{1}{2}}\alpha_0[2Gu - \frac{4}{3}\dot{G}\dot{u} - \frac{4}{3}G^2] - \rho^{-1}\alpha_0^2(2u^3 - \frac{8}{3}\dot{u}^2u)\|_{L_y^\infty} \\ & \lesssim \rho^{-\frac{3}{2}}(\mathcal{Q}_2(\|u\|_{S(\rho)}, \|\rho(F, \dot{F})\|_{L_y^\infty}) + \|\rho(F, \dot{F})\|_{L_y^\infty}^2). \end{aligned}$$

Expanding the formula for  $G$  into the first three terms on the LHS, and canceling the  $O(\rho^{-1})$  terms, this estimate follows easily from inspection.

Case 3 (The  $S(\rho)$  estimate for  $T_3$  on line (73)). The estimate for all but the first two terms follows immediately from (56) and (57), and the estimate for  $\dot{u}$  on line (72). For the first two terms we need to expand the commutator with  $\partial_\rho^2$ . The important observation here is that by formula (55) when  $\partial_\rho$  falls on the kernel  $K({}^1D, {}^2D)$  it preserves its  $S^{a,b;c}$  class and adds an inverse power of  $\rho$ . Therefore, modulo a factor of  $\rho^{-\frac{3}{2}}$  or better, the errors are one of four types:

$$\begin{aligned} & K[S(\rho), S(\rho)] \text{ and } k \in S^{1,1;2}, & K[S(\rho), \dot{S}(\rho)] \text{ and } k \in S^{1,1;2}, \\ & K[\dot{S}(\rho), \dot{S}(\rho)] \text{ and } k \in S^{0,0;2}, & K[\dot{S}(\rho), \partial_\rho \dot{S}(\rho)] \text{ and } k \in S^{0,0;2}. \end{aligned}$$

All but the last case are covered by the general class of estimates (56)–(58). To handle the last case above we need to replace the generic marker  $\partial_\rho \dot{S}(\rho)$  with its actual value, i.e.  $\ddot{u}$ . Using the estimate for  $\ddot{u} + D_y^2u$  on line (72) we can trade the last expression on the previous line for a sum of:

$$K[\dot{S}(\rho), S(\rho)] \text{ and } k \in S^{0,2;2}, \quad K[\dot{S}(\rho), S(\rho)] \text{ and } k \in S^{0,0;2},$$

which are again of the form covered by (56)–(58).

### 7. THE CUBIC CORRECTION

Our main result here is:

**Theorem 7.1** (Cubic normal forms). *Let  $w$  be a sufficiently smooth and well localized solution to the equation:*

$$(75) \quad (\square_{\mathcal{H}} + 1)u = \rho^{-1}\beta_1u^3 + F.$$

*Then there exist nonlinear functionals  $\mathcal{N}_{cubic} = \mathcal{N}_{cubic}(u, \dot{u})$  and  $\mathcal{R}_{cubic} = \mathcal{R}_{cubic}(u, \dot{u}, F, \dot{F})$  such that the following algebraic equation holds:*

$$(\square_{\mathcal{H}} + 1)\mathcal{N}_{cubic} = \rho^{-1}\beta_1u^3 + \mathcal{R}_{cubic},$$

and one has the following estimates:

$$(76) \quad \begin{aligned} & \| \mathcal{N}_{cubic} \|_{S[1,T]} + \| \mathcal{R}_{cubic} \|_{N[1,T]} \\ & \lesssim \mathcal{Q}_2(\| u \|_{S[1,T]}, \| F \|_{N[1,T]}, \| \rho \dot{F} \|_{L^\infty(L^2_y)[1,T]}) + \mathcal{Q}_3(\| F \|_{N[1,T]}). \end{aligned}$$

We will construct the correction as a dyadic sum  $\mathcal{N}_{cubic} = \sum_{k>0} \mathcal{N}_k$  with dyadic errors  $\mathcal{R}_k$ . Our ansatz for each dyadic piece is:

$$(77) \quad \mathcal{N}_k = \frac{1}{\rho} (f_{1,k}(u_{<k})^3 + f_{2,k}(u_{<k})^2 \dot{u}_{<k} + f_{3,k} u_{<k} (\dot{u}_{<k})^2 + f_{4,k} (\dot{u}_{<k})^3).$$

The functions  $f_{i,k}$  will be of the following type.

**Definition 7.2** ( $S_N^{\frac{1}{2}}$  symbols). We say a symbol (function)  $f(\rho, y)$  is in  $S_N^{\frac{1}{2}}$  if it obeys the uniform fixed time bounds:

$$(78) \quad |\nabla^\alpha f| \lesssim \rho^{(\frac{1-|\alpha|}{2})_+}, \quad \text{where } \nabla f = (\dot{f}, D_y f), \quad \text{for all } |\alpha| \leq N.$$

We write the associated norm as  $\|f(\rho)\|_{S_N^{\frac{1}{2}}}$ .

The first estimate we will need for such coefficients is the following, which follows directly from (34) and the inclusion  $S(\rho) \subseteq \dot{S}(\rho)$ , as well as the inequality  $\| \dot{u} \|_{\dot{S}(\rho)} \lesssim \mathcal{Q}_1(\| u \|_{S(\rho)} + \| F \|_{N(\rho)})$  which follows easily from (13) and (12).

**Lemma 7.3** ( $S(\rho)$  space bound for the  $\mathcal{N}_k$ ). Let  $f_k \in S_2^{\frac{1}{2}}$  with  $k > 0$ , and let  $u_{<k}, v_{<k}, w_{<k}$  be sufficiently smooth and well localized. Then one has the product estimates:

$$(79) \quad \| \rho^{-1} f_k u_{<k} v_{<k} w_{<k} \|_{S(\rho)} \lesssim \rho^{-\delta} \| f_k(\rho) \|_{S_2^{\frac{1}{2}}} \| u_{<k} \|_A \| v_{<k} \|_B \| w_{<k} \|_C,$$

where  $A, B, C$  denotes any combination of the norms  $S(\rho)$  and  $\dot{S}(\rho)$ .

In light of (79), to prove Theorem 7.1 it suffices to show:

**Proposition 7.4** (Dyadic cubic normal forms). Let  $u$  be sufficiently smooth and well localized and solve equation (75). Then there exists functions  $f_{i,k}$  such that  $f_{i,k} \equiv 0$  for  $2^k \ll \rho^{\frac{1}{2}}$ , and if one defines  $\mathcal{N}_k$  as on line (77), one has:

$$(80) \quad (\square_{\mathcal{H}} + 1) \mathcal{N}_k - \rho^{-1} (\beta_1)_k (u_{<k})^3 = \mathcal{R}_k,$$

and one has the fixed time dyadic sum estimates:

$$(81)$$

$$\sum_k \| f_{i,k}(\rho) \|_{S_2^{\frac{1}{2}}} \lesssim \ln(\rho),$$

$$(82)$$

$$\sum_k \| \mathcal{R}_k(\rho) \|_{N(\rho)} \lesssim \mathcal{Q}_2(\| u(\rho) \|_{S(\rho)}, \| F(\rho) \|_{N(\rho)}, \| \rho \dot{F}(\rho) \|_{L^2_y}) + \mathcal{Q}_3(\| F \|_{N(\rho)}).$$

We'll postpone the demonstration of this last result until the next two subsections. For now we use it to conclude:

*Proof that Lemma 7.3 and Proposition 7.4 imply Theorem 7.1.* First define  $\mathcal{N}_{cubic} = \sum_k \mathcal{N}_k$  and then set  $\mathcal{R}_{cubic} = \rho^{-1} \beta_1 u^3 - \sum_k \rho^{-1} (\beta_1)_k (u_{<k})^3 + \sum_k \mathcal{R}_k$ . Thanks to (79) and (81) we have the bound for  $\mathcal{N}_{cubic}$  on line (76). On the other hand, to

prove (76) for  $\mathcal{R}_{cubic}$  we see that estimate (82) reduces our task to showing at fixed time:

$$\|\rho^{-1}(\beta_1 u^3 - \sum_{k \geq 0} (\beta_1)_k (u_{<k})^3)\|_{N(\rho)} \lesssim \|u(\rho)\|_{S(\rho)}^3.$$

For the  $L_y^2 \cap L_y^\infty$  part of the  $N(\rho)$  norm it suffices to treat the two terms separately, and one can sum using the  $L_y^\infty$  version of the coefficient bound (40).

For the  $H_y^1$  portion of the norm, expanding the dyadic sum, it suffices to show the fixed frequency bound:

$$\|(\beta_1)_k (u^3 - (u_{<k})^3)\|_{H_y^1} \lesssim 2^{-|k - \ln_2(\rho)|} \|u\|_{H_y^1} \|u\|_{L_y^\infty}^2.$$

This follows at once by writing  $u^3 - (u_{<k})^3 = (u_{\geq k})^3 + 3(u_{\geq k})^2 u_{<k} + 3u_{\geq k} (u_{<k})^2$ , and then using the Leibniz rule and again (40). If the derivative falls on  $P_k \beta_1$  we can move it to a high frequency factor via the simple estimate  $2^k \|u_{\geq k}\|_{L_y^2} \lesssim \|u\|_{H_y^1}$ .  $\square$

The remainder of this section is devoted to the proof of Proposition 7.4.

**7.1. Construction of the symbols.** We now do an algebraic calculation of the equations satisfied by the coefficients  $f_{i,k}$  in the expansion (77). For starters we drop the index  $k$  and look for asymptotic solutions to the equation:

$$(83) \quad (\square_{\mathcal{H}} + 1)\mathcal{N} = \rho^{-1} \beta_1 u^3.$$

In deciding what terms are principle, we employ the following simple rules:

- Any term containing a product involving  $D_y u$  is an error.
- Any term with a factor of  $\rho^{-2}$  is an error.
- Any derivative of an error remains an error.

Then lengthy but straightforward computation shows:

$$(84) \quad \begin{aligned} (\square_{\mathcal{H}} + 1)\mathcal{N} = \frac{1}{\rho} & \left[ (\square_{\mathcal{H}} f_1 - 2f_1 - 2\partial_\rho f_2 + 2f_3) \cdot u^3 \right. \\ & + (\square_{\mathcal{H}} f_2 - 6f_2 + 6\partial_\rho f_1 - 4\partial_\rho f_3 + 6f_4) \cdot \dot{u} u^2 \\ & + (\square_{\mathcal{H}} f_3 - 6f_3 + 6f_1 + 4\partial_\rho f_2 - 6\partial_\rho f_4) \cdot \dot{u}^2 u \\ & \left. + (\square_{\mathcal{H}} f_4 - 2f_4 + 2f_2 + 2\partial_\rho f_3) \cdot \dot{u}^3 \right] + \mathcal{R}, \end{aligned}$$

where the remainder term can be written as  $\mathcal{R} = -2\rho^{-1} \partial_\rho \mathcal{N} + \mathcal{R}_1 + \mathcal{R}_2 + \mathcal{R}_3 + \mathcal{R}_4$  with:

$$(85) \quad \mathcal{R}_1 = \frac{1}{\rho} \left[ 3f_1 u^2 + 5f_3 \dot{u}^2 - 4f_3 u^2 + 6f_2 u \dot{u} - 12f_4 u \dot{u} \right. \\ \left. + 4\partial_\rho f_3 u \dot{u} + 2\partial_\rho f_2 u^2 + 6\partial_\rho f_4 \dot{u}^2 \right] (\ddot{u} + u),$$

$$(86) \quad \mathcal{R}_2 = \frac{1}{\rho} (2f_3 u + 6f_4 \dot{u}) (\ddot{u} + u)^2 + \frac{1}{\rho} \left[ 2f_3 u \dot{u} + f_2 u^2 + 3f_4 \dot{u}^2 \right] \partial_\rho (\ddot{u} + u),$$

$$(87) \quad \mathcal{R}_3 = 2 \frac{1}{\rho} \left[ D_y f_1 D_y (u^3) + D_y f_2 D_y (u^2 \dot{u}) + D_y f_3 D_y (u \dot{u}^2) + D_y f_4 D_y (\dot{u}^3) \right],$$

$$(88) \quad \mathcal{R}_4 = \frac{1}{\rho} \left[ f_1 D_y^2 (u^3) + f_2 D_y^2 (u^2 \dot{u}) + f_3 D_y^2 (u \dot{u}^2) + f_4 D_y^2 (\dot{u}^3) \right].$$

Setting the principle terms on the RHS of (84) so as to match the RHS of (83) we have the system of equations:

$$\begin{aligned} E_1 : & \quad \square_{\mathcal{H}} f_1 - 2f_1 - 2\partial_{\rho} f_2 + 2f_3 = \beta_1, \\ E_2 : & \quad \square_{\mathcal{H}} f_2 - 6f_2 + 6\partial_{\rho} f_1 - 4\partial_{\rho} f_3 + 6f_4 = 0, \\ E_3 : & \quad \square_{\mathcal{H}} f_3 - 6f_3 + 6f_1 + 4\partial_{\rho} f_2 - 6\partial_{\rho} f_4 = 0, \\ E_4 : & \quad \square_{\mathcal{H}} f_4 - 2f_4 + 2f_2 + 2\partial_{\rho} f_3 = 0. \end{aligned}$$

To uncover the underlying structure we form the new quantities:

$$\begin{aligned} F_1 &= 3f_1 + f_3, & F_2 &= f_1 - f_3, \\ G_1 &= f_2 + 3f_4, & G_2 &= f_2 - f_4. \end{aligned}$$

and then we take the associated linear combinations of the equations  $E_i$  above which yields the system:

$$\begin{aligned} 3E_1 + E_3 : & \quad \square_{\mathcal{H}} F_1 - 2\partial_{\rho} G_1 = 3\beta_1, \\ E_2 + 3E_4 : & \quad \square_{\mathcal{H}} G_1 + 2\partial_{\rho} F_1 = 0, \\ E_1 - E_3 : & \quad \square_{\mathcal{H}} F_2 - 8F_2 - 6\partial_{\rho} G_2 = \beta_1, \\ E_2 - E_4 : & \quad \square_{\mathcal{H}} G_2 - 8G_2 + 6\partial_{\rho} F_2 = 0. \end{aligned}$$

We now complexify this system by introducing the quantities:

$$K_1^{\dagger} = F_1 + \sqrt{-1}G_1, \quad K_2^{\dagger} = F_2 + \sqrt{-1}G_2.$$

This allows us to write the last system of equations succinctly as:

$$\begin{aligned} \square_{\mathcal{H}} K_1^{\dagger} + 2i\partial_{\rho} K_1^{\dagger} &= 3\beta_1, \\ \square_{\mathcal{H}} K_2^{\dagger} - 8K_2^{\dagger} + 6i\partial_{\rho} K_2^{\dagger} &= \beta_1. \end{aligned}$$

Finally, we introduce the gauge-transformed quantities:

$$K_1 = e^{i\rho} K_1^{\dagger}, \quad K_2 = e^{3i\rho} K_2^{\dagger},$$

which allow us to rewrite the previous system of two complex equations as follows:

$$\begin{aligned} (89) \quad (\square_{\mathcal{H}} + 1)K_1 &= 3e^{i\rho} \beta_1, & & \text{(0 resonance equation),} \\ (90) \quad (\square_{\mathcal{H}} + 1)K_2 &= e^{3i\rho} \beta_1, & & \text{(\pm\sqrt{8} resonance equation).} \end{aligned}$$

These equations only need to be solved asymptotically, so our main result for symbols becomes:

**Proposition 7.5.** *For each  $k > 0$  there exists symbols  $K_{i,k}$  and errors  $\mathcal{E}_{i,k}$  such that:*

$$(91) \quad (\square_{\mathcal{H}} + 1)K_{1,k} - 3e^{i\rho} \beta_{1,k} = \mathcal{E}_{1,k}, \quad (\square_{\mathcal{H}} + 1)K_{2,k} - e^{3i\rho} \beta_{1,k} = \mathcal{E}_{2,k}.$$

Furthermore each error can be broken up into a sum  $\mathcal{E}_{i,k} = \mathcal{E}_{i,k}^{small} + \mathcal{E}_{i,k}^{smooth}$  and one has:

$$(92) \quad \|(e^{-i\rho} \mathcal{E}_{1,k}^{smooth}, e^{-3i\rho} \mathcal{E}_{2,k}^{smooth})\|_{H_y^1} \lesssim 2^{-\frac{1}{2}|k - \ln_2(\rho)|},$$

$$(93) \quad \|(e^{-3i\rho} K_{2,k}, \rho e^{-3i\rho} \mathcal{E}_{2,k}^{small})\|_{S_N^{\frac{1}{2}}} \lesssim 2^{-|k - \ln_2(\rho)|}.$$

In addition for the zero resonance symbol we have:

$$(94) \quad \begin{aligned} & \| (\partial_\rho^n D_y^m e^{-i\rho} K_{1,k}, \rho \partial_\rho^n D_y^m e^{-i\rho} \mathcal{E}_{1,k}^{small}) \|_{L^\infty} \\ & \lesssim \rho^{\left(\frac{1-n-m}{2}\right)_+} \begin{cases} 2^{-|k-\frac{1}{2}\ln_2(\rho)|} + 2^{-|k-\ln_2(\rho)|}, & (n, m) \neq (0, 1); \\ 2^{-(k-\ln_2(\rho))_+}, & (n, m) = (0, 1). \end{cases} \end{aligned}$$

In particular for the original  $f_{i,k}$  one has  $\|f_{i,k}\|_{S_2^{\frac{1}{2}}} \lesssim 2^{-(k-\ln_2(\rho))_+}$  which implies estimate (81).

*Proof.* We break the proof up into a series of steps.

*Step 1* (Time scale decomposition and definition of  $K_{i,k}$  and  $\mathcal{E}_{i,k}$ ). The way we solve the equations on line (91) is to use different information at different times. For a fixed value of  $k$  we define the smooth partition of unity in time:

$$1 = h_k^{smooth} + h_k^{elliptic} + h_k^{dispersive} + h_k^{high},$$

where the supports are defined as:

$$\begin{aligned} \text{supp}(h_k^{high}) &= \{\rho \ll 2^k\}, & \text{supp}(h_k^{dispersive}) &= \{\rho \approx 2^k\}, \\ \text{supp}(h_k^{elliptic}) &= \{2^k \ll \rho \leq 2^{2k}\}, & \text{supp}(h_k^{smooth}) &= \{\rho > 2^{2k}\}. \end{aligned}$$

We may assume these functions are chosen so that  $|\partial_\rho^n h_k^\bullet| \lesssim \rho^{-n}$ . Then define:

$$(95) \quad \mathcal{E}_{1,k}^{smooth} = -3e^{i\rho} h_k^{smooth} \beta_{1,k}, \quad \mathcal{E}_{2,k}^{smooth} = -e^{3i\rho} h_k^{smooth} \beta_{1,k},$$

$$(96) \quad \begin{aligned} K_{1,k} &= 3e^{i\rho} (1 - h_k^{smooth}) D_y^{-2} \beta_{1,k}, \\ K_{2,k}^{nondisp} &= e^{3i\rho} (h_k^{elliptic} + h_k^{high}) (D_y^2 - 8)^{-1} \beta_{1,k}. \end{aligned}$$

Here we define the operators  $D_y^{-2}$  and  $(D_y^2 - 8)^{-1}$  in terms of Fourier space division by their symbols which is a smooth operation because we have cut away from the 0 and  $\pm\rho\sqrt{8}$  frequencies of  $\beta_1$ .

We will set  $K_{2,k} = K_{2,k}^{nondisp} + K_{2,k}^{disp}$  where the first term on the RHS is defined above and the second will be produced by asymptotically solving:

$$(97) \quad (\square_{\mathcal{H}} + 1) K_2^{disp} \approx e^{3i\rho} h_k^{disp} P_{|k-\ln_2(\rho)| < C} \beta_1, \quad \text{and then setting } K_{2,k}^{disp} = \tilde{h}_k^{disp} P_k K_2^{disp}.$$

The equation on the LHS above will be solved by reverting back to the original  $(t, x)$  coordinates, and the needed estimates will then become a standard stationary phase calculation. Explicit details will be given shortly. For now we simply note that constant  $C > 0$  in the above formula is chosen sufficiently large so that  $h_k^{disp} P_k P_{|k-\ln_2(\rho)| < C} = h_k^{disp} P_k$ , and  $\tilde{h}_k^{disp}$  is another cutoff on  $\rho \approx 2^k$  such that  $\tilde{h}_k^{disp} h_k^{disp} = h_k^{disp}$ .

Finally, the “small” remainders are defined by:

$$(98) \quad \begin{aligned} \mathcal{E}_{1,k}^{small} &= (\square_{\mathcal{H}} + 1) K_{1,k} - 3e^{i\rho} (1 - h_k^{smooth}) \beta_{1,k}, \\ \mathcal{E}_{2,k}^{small} &= (\square_{\mathcal{H}} + 1) K_{2,k} - e^{3i\rho} (1 - h_k^{smooth}) \beta_{1,k}. \end{aligned}$$

In particular this gives the formulas on line (91). We now turn to estimating the quantities defined above.

*Step 2* (Estimate (92) for  $\mathcal{E}_{i,k}^{smooth}$ ). We need to show estimate (92) for the quantities on line (95). This follows at once from (40) with  $p = 2$  and  $m = 0, 1$ , and the restriction  $2^k \lesssim \rho^{\frac{1}{2}}$ .

*Step 3* (Estimate (94) for  $K_{1,k}$ ). For any of the time dependent frequency cutoffs we may safely ignore the commutators  $[\partial_\rho^n, h_k^\bullet]$  as they produce inverse powers of  $\rho^{-n}$  and also restrict the relation between  $\rho$  and  $2^k$  to a finite number of dyadic frequencies at a given time. Concentrating on  $D_y^{-2}\beta_{1,k}$  we directly have from (40) the following estimate for  $2^k \geq \rho^{\frac{1}{2}}$ :

$$(99) \quad \begin{aligned} \|\partial_\rho^n D_y^m D_y^{-2}\beta_{1,k}\|_{L_y^\infty} &\lesssim \rho^{-n-m+1} 2^{(m-1)k} 2^{-N(k-\ln_2(\rho))_+} \\ &\lesssim \rho^{-n} \begin{cases} \rho^{\frac{1}{2}} 2^{-|k-\frac{1}{2}\ln_2(\rho)|}, & m = 0; \\ 2^{-(k-\ln_2(\rho))_+}, & m = 1; \\ 2^{-|k-\ln_2(\rho)|}, & m \geq 2. \end{cases} \end{aligned}$$

Note that there is no dyadic gain for  $2^k \leq \rho$  in the middle case above. Together these imply (94) for  $e^{-i\rho}K_{1,k}$ .

*Step 4* (Estimate (93) for  $K_{2,k}$ ). We'll do this separately for  $K_{2,k}^{nondisp}$  and  $K_{2,k}^{disp}$ .

*Step 4a* (Contribution of “elliptic” and “high” time scales). Here we consider the quantity  $(D_y^2 - 8)^{-1}\beta_{1,k}$ , where the frequencies are restricted to the range  $2^k \gg \rho$  or  $2^k \ll \rho$ . With this restriction division by  $(D_y^2 - 8)^{-1}$  is a bounded smooth multiplier and so  $(D_y^2 - 8)^{-1}\beta_{1,k} = \tilde{\beta}_{1,k}$  for some new  $\tilde{\beta}_1 \in \mathcal{S}$ . Then by (40) we get:

$$(100) \quad \|\partial_\rho^n D_y^m (D_y^2 - 8)^{-1}\beta_{1,k}\|_{L_y^\infty} \lesssim \rho^{-n-m-1} 2^{(m+1)k} 2^{-N(k-\ln_2(\rho))_+} \lesssim \rho^{-n} 2^{-|k-\ln_2(\rho)|}.$$

*Step 4b* (Contribution of the “dispersive” time scale). Set  $\tilde{\beta}_1(x) = P_{|k|<C}\beta_1(x)$ . Then define the function  $K_2^{disp}$  according to the formula:

$$\begin{aligned} K_2^{disp}(\rho, y) &= \rho^{\frac{1}{2}} \chi(y) \tilde{K}_2^{disp}(\rho \cosh(y), \rho \sinh(y)), \\ \tilde{K}_2^{disp}(t, x) &= \int_1^t \frac{\sin((t-s)\langle D_x \rangle)}{\langle D_x \rangle} (e^{3is} h_k^{disp}(s) s^{-\frac{1}{2}} \tilde{\beta}_1) ds. \end{aligned}$$

Here  $\chi(y)$  is a cutoff of sufficiently large support that  $1 - |x/t| \leq \epsilon$  on the support of  $\chi'$  where  $\epsilon$  is the same as on line (41). With this definition we have  $h_{2^k}(s) := \chi(y)\rho^{\frac{1}{2}} h_k^{disp}(s) s^{-\frac{1}{2}}$  satisfies  $|\partial_s^n h_{2^k}| \lesssim 2^{-nk}$  when  $\rho \approx 2^k$ . Then from the stationary phase estimate (41) we have the two bounds:

$$(101) \quad |\partial_\rho^n D_y^m \tilde{h}_k^{disp} P_k K_2^{disp}| \lesssim 1,$$

$$(102) \quad \|\partial_\rho^n D_y^m \tilde{h}_k^{disp} P_k ((\square_{\mathcal{H}} + 1)K_2^{disp} - e^{3it} h_k^{disp}(t) \chi(y) \sqrt{\rho/t} \tilde{\beta}_1(x))\|_{L_y^\infty} \lesssim \rho^{-N}.$$

To see (102) notice that (6) gives

$$(\square_{\mathcal{H}} + 1)K_2^{disp} = [D_y^2, \chi] \rho^{\frac{1}{2}} \tilde{K}_2^{disp} + e^{3it} h_k^{disp}(t) \chi(y) \sqrt{\rho/t} \tilde{\beta}_1(x),$$

and by the support properties of  $\chi'(y)$  we have that  $[D_y^2, \chi] \rho^{\frac{1}{2}} \tilde{K}_2^{disp}$  is rapidly decaying thanks to the second term on RHS (41). Note also that (101) in particular implies (93) for  $K_{2,k}^{disp} = \tilde{h}_k^{disp} P_k K_2^{disp}$ .

*Step 5* (Estimate for the “small” remainders). It remains to prove bounds (93) and (94) for the two quantities defined on line (98). We'll do this separately for each term.

Step 5a (Estimate (94) for  $\mathcal{E}_{1,k}^{small}$ ). A quick calculation shows:

$$e^{-i\rho}(\square_{\mathcal{H}}+1)K_{1,k}-3(1-h_k^{smooth})\beta_{1,k} = (2i\partial_\rho+\partial_\rho^2+\frac{1}{4}\rho^{-2})[(1-h_k^{smooth})D_y^{-2}\beta_{1,k}].$$

Then estimate (94) for the RHS of this last line follows from (99).

Step 5b (Estimate (93) for  $\mathcal{E}_{2,k}^{small}$ ). Another quick calculation shows:

$$e^{-3i\rho}(\square_{\mathcal{H}}+1)K_{2,k}-(1-h_k^{smooth})\beta_{1,k} = T_1+T_2+T_3,$$

where:

$$\begin{aligned} T_1 &= (6i\partial_\rho+\partial_\rho^2+\frac{1}{4}\rho^{-2})[(h_k^{elliptic}+h_k^{high})(D_y^2-8)^{-1}\beta_{1,k}], \\ T_2 &= e^{-3i\rho}P_k[\partial_\rho^2,\tilde{h}_k^{disp}]K_2^{disp}+e^{-3i\rho}\tilde{h}_k^{disp}P_k[(\square_{\mathcal{H}}+1)K_2^{disp} \\ &\quad -e^{3it}h_k^{disp}(t)\tilde{\chi}(y)\tilde{\beta}_1(x)], \\ T_3 &= \tilde{h}_k^{disp}P_k[e^{3i\rho(\cosh(y)-1)}h_k^{disp}(\rho\cosh(y))\tilde{\chi}(y)\tilde{\beta}_1(\rho\sinh(y)) \\ &\quad -h_k^{disp}(\rho)P_{|k-\ln_2(\rho)|<C}\beta_1], \end{aligned}$$

where we are using the shorthand  $\tilde{\chi}(y) = \sqrt{\rho/t}\chi(y) = \text{sech}^{\frac{1}{2}}(y)\chi(y)$ . Indeed, using the change of coordinates (4) we see that the last term in  $T_2$  cancels the first term in  $T_3$ , so by choosing cutoffs with the property  $\tilde{h}_k^{disp}P_k h_k^{disp}P_{|k-\ln_2(\rho)|<C} = h_k^{disp}P_k$  we have  $T_2+T_3 = e^{-3i\rho}(\square_{\mathcal{H}}+1)K_{2,k}^{disp} - h_k^{disp}\beta_{1,k}$ . On the other hand the identity  $T_1 = e^{-3i\rho}(\square_{\mathcal{H}}+1)K_{2,k}^{nondisp} - (h_k^{elliptic}+h_k^{high})\beta_{1,k}$  follows from a simple direct calculation.

Estimate (93) for  $T_1$  and  $T_2$  (resp) follows directly from (100) and (101)–(102) (resp). To bound the last term we may rewrite it as  $T_3 = T_{31}+T_{32}$  where:

$$\begin{aligned} T_{31} &= \tilde{h}_k^{disp}P_k[e^{3i\rho(\cosh(y)-1)}h_k^{disp}(\rho\cosh(y))\tilde{\chi}(y)(\tilde{\beta}_1(\rho\sinh(y))-\tilde{\beta}_1(\rho y))], \\ T_{32} &= \tilde{h}_k^{disp}P_k[P_{|k-\ln_2(\rho)|<C}\beta_1(\rho y) \\ &\quad \cdot (e^{3i\rho(\cosh(y)-1)}h_k^{disp}(\rho\cosh(y))\tilde{\chi}(y)-h_k^{disp}(\rho))]. \end{aligned}$$

Notice that by definition of  $\tilde{\beta}_1(x)$  given in Step 4b above and rescaling we have  $P_{|k-\ln_2(\rho)|<C}\beta_1(\rho y) = \tilde{\beta}_1(\rho y)$ . For the term  $T_{31}$  we easily obtain estimate (93) by using estimate (9). To handle  $T_{32}$  it suffices to show for any  $\varphi \in \mathcal{S}$ :

$$\begin{aligned} \|\partial_\rho^n D_y^m [(e^{3i\rho(\cosh(y)-1)}-1)h_k^{disp}(\rho\cosh(y))\varphi(\rho y)\tilde{\chi}(y)]\|_{L_y^\infty} &\lesssim \rho^{-1}, \\ \|\partial_\rho^n D_y^m [(h_k^{disp}(\rho\cosh(y))\tilde{\chi}(y)-h_k^{disp}(\rho))\varphi(\rho y)]\|_{L_y^\infty} &\lesssim \rho^{-1}, \end{aligned}$$

which in both cases follows by inspection when  $|y| \gtrsim \rho^{-\frac{1}{2}}$  and Taylor expansions for  $|y| \ll \rho^{-\frac{1}{2}}$ . □

**7.2. Estimates for the remainders.** Our final task is to prove estimate (82). To do this we write line (84) above in a frequency localized form  $(\square_{\mathcal{H}}+1)\mathcal{N}_k = \mathcal{R}_k^{coef} + \mathcal{R}_k$  where:

$$\mathcal{R}_k^{coef} = \frac{1}{\rho}[\mathcal{F}_{1,k}u_{<k}^3 + \mathcal{F}_{2,k}\dot{u}_{<k}u_{<k}^2 + \mathcal{F}_{3,k}(\dot{u}_{<k})^2u_{<k} + \mathcal{F}_{4,k}(\dot{u}_{<k})^3],$$

with the  $\mathcal{F}_{i,k}$  denoting the errors generated by applying the equations for  $f_i$  on RHS (84) to the coefficients  $f_{i,k}$  defined above, and where  $\mathcal{R}_k = -2\rho^{-1}\partial_\rho\mathcal{N}_k + \mathcal{R}_{1,k} + \mathcal{R}_{2,k} + \mathcal{R}_{3,k} + \mathcal{R}_{4,k}$ , with the  $\mathcal{R}_{i,k}$  denoting the expressions on lines (85)–(88) with  $f_i$  replaced by  $f_{i,k}$  and  $u, \dot{u}$  replaced by  $u_{<k}, \dot{u}_{<k}$ . Each term will be estimated separately.

*Case 1* (Terms in  $\mathcal{R}_k^{coef}$ ). Recall that by the construction of the previous subsection the  $\mathcal{F}_{i,k}$  are complex linear combinations of the errors  $e^{-i\rho}\mathcal{E}_{1,k}$  and  $e^{-i3\rho}\mathcal{E}_{2,k}$  which satisfy the estimates on lines (92)–(94). In the case of the “smooth” errors we have access to a uniform  $H_y^1$  bound and (82) then follows from the Leibniz rule and Hölder’s inequality. In the case of the “small errors” we can use  $S_2^{\frac{1}{2}}$  coefficient estimates so that (82) follows from (79) and the embedding (10).

*Case 2* (The term  $\rho^{-1}\partial_\rho\mathcal{N}_k$ ). The estimate (82) in this case follows from (10) and the fact that (81) already implied the  $\sum_k \mathcal{N}_k \in S(\rho)$  estimate.

*Case 3* (The terms  $\mathcal{R}_{1,k}$ ). All of the terms here are schematically of the form:

$$\mathcal{R}_{1,k}^{schem} = \rho^{-1}(f_k, \dot{f}_k)(u_{<k}, \dot{u}_{<k})^2(-D_y^2u_{<k} + G_{<k}),$$

where  $G_{<k} = -\frac{1}{4}\rho^{-2}u_{<k} + \rho^{-1}P_{<k}(\beta_1u^3) + F_{<k}$ . We break this into two further subcases.

*Case 3a* (The term involving  $G_{<k}$ ). Here we use the immediate estimate:

$$(103) \quad \|G_{<k}\|_{L_y^2 \cap L_y^\infty} + \|\dot{G}_{<k}\|_{L_y^2} \lesssim \rho^{-1}(\mathcal{Q}_1(\|u\|_{S(\rho)}) + \|F\|_{N(\rho)} + \|\rho\dot{F}\|_{L_y^2}).$$

Then from (36) we get:

$$\|\rho^{-1}(f_k, \dot{f}_k)(u_{<k}, \dot{u}_{<k})^2G_{<k}\|_{N(\rho)} \lesssim \rho^{-\delta}\|f_k\|_{S_2^{\frac{1}{2}}}\|u\|_{S(\rho)}^2\|\rho G_{<k}\|_{L_y^2},$$

which is sufficient to sum and produce (82) for terms of this form.

*Case 3b* (The term involving  $D_y^2u_{<k}$ ). It remains to bound the portion of  $\mathcal{R}_{1,k}^{schem}$  which contains the factor  $D_y^2u_{<k}$ . In this case we have two derivatives to put on the coefficients  $f_k$  so by (35) we have:

$$(104) \quad \|\rho^{-1}(f_k, \dot{f}_k)(u_{<k}, \dot{u}_{<k})^2D_y^2u_{<k}\|_{N(\rho)} \lesssim \|D_y^2(f_k, \dot{f}_k)\|_{L_y^\infty}\|u\|_{S(\rho)}^3.$$

By estimates (93) and (94) the RHS above can be summed over all  $k > 0$ .

*Case 4* (The terms  $\mathcal{R}_{2,k}$ ). In this case the schematic form is:

$$\mathcal{R}_{2,k}^{schem} = \rho^{-1}f_k(u_{<k}, \dot{u}_{<k})(-D_y^2u_{<k} + G_{<k})^2 + \rho^{-1}f_k(u, \dot{u})^2\partial_\rho(-D_y^2u_{<k} + G_{<k}),$$

where  $G_{<k}$  is as above. Expanding the product in the first RHS term, and using a combination of estimates (35) and (36) (possibly by setting  $H_{<k+C} = G_{<k}^2$  in the latter case) we have:

$$\begin{aligned} \|\mathcal{R}_{2,k}^{schem}\|_{N(\rho)} &\lesssim \rho^{-\delta}\|f_k\|_{S_3^{\frac{1}{2}}}\left(\|u\|_{S(\rho)}^2\|\rho(G_{<k}, \dot{G}_{<k})\|_{L_y^2} + \|u\|_{S(\rho)}\|\rho G_{<k}\|_{L_y^2 \cap L_y^\infty}^2\right) \\ &\quad + \|(D_y^2f_k, D_y^4f_k)\|_{L^\infty}\|u\|_{S(\rho)}^3. \end{aligned}$$

Then using (103) above for  $G_{<k}$  we have (82) for these terms.

Case 5 (The terms  $\mathcal{R}_{3,k}$  and  $\mathcal{R}_{4,k}$ ). In these cases the two schematic forms are:

$$\begin{aligned}\mathcal{R}_{3,k}^{schem} &= \rho^{-1} D_y(f_k, \dot{f}_k) D_y[(u_{<k}, \dot{u}_{<k})^3], \\ \mathcal{R}_{4,k}^{schem} &= \rho^{-1} (f_k, \dot{f}_k) D_y^2[(u_{<k}, \dot{u}_{<k})^3].\end{aligned}$$

In both cases we can use estimate (35) to produce an estimate of the form (104) above.

#### ACKNOWLEDGEMENTS

This work began as a joint project with H. Lindblad, A. Soffer, and I. Rodnianski. Although all participants made important contributions to the project, the author assumes full responsibility for the correctness of the details in this paper. The author would also like to thank the anonymous referee for helpful comments.

For further exposition and an alternate presentation of the technical details, based on a previous draft of this paper, please refer to [L-S3].

#### REFERENCES

- [D1] Jean-Marc Delort, *Erratum: "Global existence and asymptotic behavior for the quasi-linear Klein-Gordon equation with small data in dimension 1" (French)* [Ann. Sci. École Norm. Sup. (4) **34** (2001), no. 1, 1–61; MR 1833089] (French, with English and French summaries), Ann. Sci. École Norm. Sup. (4) **39** (2006), no. 2, 335–345, DOI 10.1016/j.ansens.2006.01.001. MR2245535 (2007c:35109)
- [H1] Lars Hörmander, *Lectures on nonlinear hyperbolic differential equations*, Mathématiques & Applications (Berlin) [Mathematics & Applications], vol. 26, Springer-Verlag, Berlin, 1997. MR1466700 (98e:35103)
- [K1] Sergiu Klainerman, *Global existence of small amplitude solutions to nonlinear Klein-Gordon equations in four space-time dimensions*, Comm. Pure Appl. Math. **38** (1985), no. 5, 631–641, DOI 10.1002/cpa.3160380512. MR803252 (87e:35080)
- [L-S1] Hans Lindblad and Avy Soffer, *A remark on long range scattering for the nonlinear Klein-Gordon equation*, J. Hyperbolic Differ. Equ. **2** (2005), no. 1, 77–89, DOI 10.1142/S0219891605000385. MR2134954 (2006d:35143)
- [L-S2] Hans Lindblad and Avy Soffer, *A remark on asymptotic completeness for the critical nonlinear Klein-Gordon equation*, Lett. Math. Phys. **73** (2005), no. 3, 249–258, DOI 10.1007/s11005-005-0021-y. MR2188297 (2006i:35249)
- [L-S3] H. Lindblad and A. Soffer, *Scattering for the Klein-Gordon equation with quadratic and variable coefficient cubic nonlinearities*, Trans. Amer. Math. Soc., electronically published on December 3, 2014, DOI: <http://dx.doi.org/10.1090/S0002-9947-2014-06455-6>.
- [M-S] Nicholas Manton and Paul Sutcliffe, *Topological solitons*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 2004. MR2068924 (2006d:58020)
- [Sh] Jalal Shatah, *Normal forms and quadratic nonlinear Klein-Gordon equations*, Comm. Pure Appl. Math. **38** (1985), no. 5, 685–696, DOI 10.1002/cpa.3160380516. MR803256 (87b:35160)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA SAN DIEGO, LA JOLLA, CALIFORNIA 92093-0112

*E-mail address:* [jsterbenz@math.ucsd.edu](mailto:jsterbenz@math.ucsd.edu)