TESTING BY WAVE PACKETS AND MODIFIED SCATTERING IN NONLINEAR DISPERSIVE PDE'S

MIHAELA IFRIM AND DANIEL TATARU

Abstract. Modified scattering phenomena are encountered in the study of global properties for nonlinear dispersive partial differential equations in situations where the decay of solutions at infinity is borderline and scattering fails just barely. An interesting example is that of problems with cubic nonlinearities in one space dimension.

The method of testing by wave packets was introduced by the authors as a tool to efficiently capture the asymptotic equations associated to such flows, and thus establish the modified scattering mechanism in a simpler, more efficient fashion, and at lower regularity. In these expository notes we describe how this method can be applied to problems with general dispersion relations.

Contents

1. Introduction

Given a nonlinear perturbation of a linear partial differential equation, scattering theory aims to compare the long time dynamics of the nonlinear problem with the long time dynamics of the corresponding linear flow. This is particularly interesting in the context of dispersive equations, which have two key properties:

- a conservative nature at the L^2 level, with some energy that is either exactly conserved, or essentially conserved for small data.
- some form of uniform or averaged decay, whose effect is that the strength of the nonlinear interactions decays with time.

Whether scattering holds for a given problem, that depends on the relative strength of the nonlinearity on one hand, and on the dispersive effects on the other hand. If the nonlinearity is mild, then scattering holds, in the sense that, as time goes to infinity, the solutions to the nonlinear problem will approach solutions to the linear problem.

In this paper we are instead interested in the *modified scattering* phenomena. These are encountered in situations where the decay of solutions at infinity is borderline, and scattering fails just barely. Then one might expect that the nonlinear asymptotic behavior can be seen as some perturbation of the linear asymptotic behavior. Such dynamics are encountered for many classes of equations, and the modified scattering effects may vary from case to case; this may include for instance corrections to the velocity, amplitude or phase. The class of problems we are interested in here is that of dispersive problems with cubic nonlinearities in one space dimension. As we will see, for this class modified scattering means a phase correction on a logarithmic time scale.

2020 *Mathematics Subject Classification.* Primary: 35P25; Secondary: 76B15, 35Q55.
Key words and phrases. wave packets, modified scattering.

The method of testing by wave packets was introduced by the authors in the context of the cubic nonlinear Schrödinger flow (NLS) [8], and used later in water wave contexts [9], [10], [1] as a tool to efficiently capture the asymptotic equations associated to such flows, and thus establish the modified scattering mechanism. See also [5, 17, 6, 18] for further examples where this idea is used.

These notes, written by the authors for a summer school at MSRI [12] in 2020, and based on earlier set of notes prepared by the authors for an AMS meeting in Las Vegas in 2016, aim to describe how this method can be applied to problems with general dispersion relations. Notably, here we work with with minimal structure assumptions on the nonlinearity, which do not include a scaling symmetry or energy conservation.

1.1. **A model dispersive problem.** The model problem we consider here is a one dimensional evolution of the form

$$\begin{cases} i\partial_t u - A(D)u = Q(u,\bar{u},u) \\ u(0,x) = u_0(x), \end{cases} \tag{1.1}$$

for a complex valued function

$$u : \mathbb{R}\times\mathbb{R} \to \mathbb{C}.$$

Here we will make the following general assumptions:

(H1) *Real symbol.* The symbol $a(\xi)$ of the multiplier $A(D)$ is real and smooth; this guarantees that the L^2 norm is preserved for solutions to the corresponding linear flow.

(H2) *Dispersive character.* The group velocity depends on the frequency,

$$\mathbb{R} \ni \xi \to a'(\xi) \in \mathbb{R} \quad \text{is a diffeomorphism} \quad (i.e.\ \ a'' \neq 0).$$

(H3) *Cubic, translation and phase shift invariant nonlinearity.* The nonlinearity Q is defined by its smooth symbol q as follows:

$$\widehat{Q(u,\bar{u},u)}(\xi) = \frac{1}{2\pi}\int_{\xi=\xi_1-\xi_2+\xi_3} q(\xi_1,\xi_2,\xi_3)\hat{u}(\xi_1)\overline{\hat{u}(\xi_2)}\hat{u}(\xi_3)\,d\sigma,$$

with $d\sigma = d_{\xi_i}d_{\xi_j}$, $i \neq j = \overline{1,3}$ (where one needs to adjust the sign corresponding to the chosen (i,j) pair).

(H4) *Conservative nonlinearity.* The symbol associated to Q and computed on the diagonal must be real i.e.

$$q(\xi,\xi,\xi) \in \mathbb{R}.$$

For the Cauchy problem (1.1) we will ask the following question:

Question 1. *Assume that the initial data for the evolution* (1.1) *is small, localized and sufficiently smooth. Does this guarantee that we have global solutions with dispersive decay and modified scattering ?*

Our goal in this paper will be to show that the answer is affirmative, under minimal assumptions on the behaviour of the symbols $a(\cdot)$ and $q(\cdot,\cdot,\cdot)$ at infinity, and also under minimal regularity and decay assumptions for the initial data u_0.

1.2. **An overview of the paper.** To motivate the results, our exposition will begin with a brief discussion of linear dispersion in Section 2, which notably ends with the vector field bound in Proposition 2.1. The standard linear scattering mechanism described here serves as the basis for the nonlinear, modified scattering results which are the main goal of the paper.

In the following section, i.e. in Section 3, we provide a heuristic discussion of the modified scattering phenomena. The premise here is that, relative to the linear scattering mechanism, the nonlinear *asymptotic profile* is governed by an *asymptotic equation*. The objective is then to efficiently capture both of these objects in the analysis. We outline several ideas which have been used over the years, and finish with a brief introduction to wave packet testing.

At this point we are ready to present the main results of the paper, which in a nutshell assert that global solutions with modified scattering dynamics exist for the flow in (1.1) under suitable assumptions. For expository purposes, we will split the discussion in two parts:

(i) In Section 4 we consider cubic forms Q with compactly supported symbols. Then one may also assume that the solution u has a compactly supported Fourier transform, and no restriction is imposed on the behavior of a at infinity. In this setting the arguments are simpler, and we are able to present the main steps, namely the energy estimates and the wave packet testing, in a streamlined fashion, without distracting technicalities.

(ii) In Section 5 we consider cubic forms Q with bounded symbols, and, correspondingly, symbols a so that $a''(\xi) \approx |\xi|^\sigma$ at infinity for some real σ.

Then we show that, for initial data u_0 which is small in suitable spaces

$$u_0 \in H^{s_0}, \qquad xu_0 \in H^{s_1},$$

the solutions are global and their asymptotic behavior is still governed by the modified scattering mechanism. Anticipating the precise results in Section 5, we point out here that there are two qualitatively different scenarios:

(i) The generalized NLS case: $\sigma \geq -1$, where a is superlinear at infinity and thus we have infinite speed of propagation. This includes for instance

a) SQG type problems, $\sigma = -1$, where we take

$$s_0 = 0+, \qquad s_1 = 1.$$

b) NLS type problems, $\sigma = 0$, where we take

$$s_0 = -\frac{1}{2}+, \qquad s_1 = \frac{1}{2}.$$

c) KdV type problems, $\sigma = 1$, where we take

$$s_0 = -1+, \qquad s_1 = 0.$$

(ii) The generalized Klein-Gordon case: $\sigma < -1$, where a is linear at infinity and then we have finite speed of propagation in the high frequency limit. This includes for instance

a) gravity wave models, $\sigma = -3/2$, where we take

$$s_0 = 1+, \qquad s_1 = \frac{1}{2}.$$

b) Klein-Gordon models, $\sigma = -3$, where we take

$$s_0 = 1+, \qquad s_1 = 2.$$

While our results do allow for a full range of asymptotic behaviors for a, this is far from capturing a full range of problems, as Q is not in general bounded in many interesting models. We leave it for the interested reader to investigate more general situations.

Another line of investigation which is completely omitted in our discussion here is that of normal form methods, which in many instances allow one to expand the scope of this type of results to problems which also have nonresonant quadratic nonlinearities.

1.3. **Acknowledgements.** The first author was supported by a Luce Professorship, by the Sloan Foundation, and by an NSF CAREER grant DMS-1845037. The second author was supported by the NSF grant DMS-2054975 as well as by a Simons Investigator grant from the Simons Foundation. Some of this work was presented during an MSRI Graduate summer school in 2020. Other parts of the work were carried out while both authors were participating in the MSRI program "Mathematical problems in fluid dynamics" during Spring 2021.

The authors are very grateful to the anonymous referee for a thorough reading of the paper, which led to many improvements and corrections.

s:linear

2. Dispersive decay for the linear equation

In this section we consider the dispersive properties and the asymptotic behavior of the solutions to the associated linear problem,

linear

$$\begin{cases} i\partial_t u - A(D)u = 0 \\ u(0) = u_0. \end{cases} \tag{2.1}$$

To avoid distracting technicalities, we will assume that the initial data u_0 is frequency localized in a fixed compact set.

2.1. **The fundamental solution and dispersive decay.** Denoting by τ, respectively ξ, the time and the space Fourier variables, the symbol of the linear operator is

$$p(\tau,\xi) = -\tau - a(\xi), \qquad P = i\partial_t - A(D),$$

and its characteristic set is the graph of $-a$,

$$\text{char } P = \{\tau = -a(\xi)\}.$$

This is commonly referred to as the *dispersion relation*.

The associated Hamilton flow is

$$(x,\xi) \to (x + ta'(\xi), \xi).$$

In particular we note here the group velocity $a'(\xi)$, which depends on the spatial frequency ξ of the waves. We denote the range of admissible group velocities by

$$V = a'(\mathbb{R}).$$

Here we may distinguish two different scenarios, depending on the asymptotic behavior of a at infinity.

a) The generalized NLS case, where a has superlinear growth at $\pm\infty$, in which case $a' : \mathbb{R} \to \mathbb{R}$ is surjective so $V = \mathbb{R}$, i.e. waves propagate with all possible velocities.
b) The generalized Klein-Gordon case, where a has linear growth at $\pm\infty$, in which case the set V of possible group velocities is a bounded open interval.

Of course, one may also differentiate between the behavior of a at $+\infty$ and at $-\infty$, with obvious consequences.

The spatial Fourier transform of the fundamental solution $K(t,x)$ for (2.1) is given by

$$\hat{K}(t,\xi) = e^{-ita(\xi)},$$

which yields the following oscillatory integral representation for K:

$$K(t,x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{i(x\xi - ta(\xi))}\, d\xi.$$

By the assumption (H2), the phase function is nondegenerate, and has at most one critical point, namely the solution ξ of the equation

$$x - ta'(\xi) = 0.$$

Denoting $v = x/t$ as the velocity along a ray starting from the origin, this becomes

xiv (2.2) $$v = a'(\xi).$$

We denote the solution of (2.2) by ξ_v. In the generalized NLS case this solution exists for all real velocities v. But in the generalized Klein-Gordon case the critical point exists only for $v \in V$.

Assuming that $v \in V$, the asymptotics of the the fundamental solution along the ray $x = vt$ can be computed using the stationary phase method, see [Stein 20], which yields the asymptotic expansion

fundamental (2.3) $$K(t,vt) \approx \frac{1}{\sqrt{2\pi t|a''(\xi_v)|}} e^{it\phi(v)} e^{-\frac{i\pi\,\mathrm{sgn}(a'')}{4}} + O(t^{-1})$$

where the phase function $\phi(v)$ is given by

$$\phi(v) = v\xi_v - a(\xi_v).$$

This holds uniformly for v in compact subsets of V, with a more complex behavior at the endpoints and rapid decay along rays outside the closure of V in the generalized Klein-Gordon case.

Since by (H2) a is either convex or concave, this last expression also allows one to interpret ϕ as the Legendre transform of a, so that ϕ' and a' are inverse functions,

phi-prime (2.4) $$\xi_v = \phi'(v).$$

Equivalently, ϕ can be thought of as the solution to the eikonal equation

a-vs-phi (2.5) $$a(\phi_v) = \phi - v\phi_v.$$

More generally, one may apply the stationary phase method to compute the asymptotics for any solution u to the linear equation (2.1) with initial data u_0 with a smooth Fourier transform (which corresponds to a localized initial data), namely

r-scattering

$$u(t, x = vt) \approx \gamma(v) \frac{1}{\sqrt{t|a''(\xi_v)|}} e^{it\phi(v)} e^{-\frac{i\pi \operatorname{sgn}(a'')}{4}} + O(t^{-1}), \tag{2.6}$$

where the asymptotic profile $\gamma = \gamma(v)$ depends on the initial data in a straightforward fashion,

$$\gamma(v) = \hat{u}_0(\xi_v).$$

The expansion in (2.6) is uniform for v in a compact subset of V, or if $\hat{u}_0$ is compactly supported or at the very least has sufficient decay at infinity.

One may also think of the linear dispersive decay of solutions to (2.1) in a translation invariant fashion. This is described by the following result:

t:disp

Theorem 1. *Assume that the conditions (H1), (H2) hold. Then the following translation invariant decay estimates hold for frequency localized solutions to* (2.1)*:*

- *Dispersive bounds:*

dispersive

$$\|u(t)\|_{L^\infty} \lesssim t^{-\frac{1}{2}} \|u(0)\|_{L^1}. \tag{2.7}$$

- *Strichartz estimates:*

Strichartz

$$\|u\|_{L^4 L^\infty} \lesssim \|u(0)\|_{L^2}. \tag{2.8}$$

The dispersive bound follows from the pointwise decay of the (frequency localized) fundamental solution, see (2.3). The Strichartz estimates (2.8) can be seen as a direct consequence of the dispersive estimates (2.7) and Young's inequality, see [GV] and [KeTa].

Similar bounds hold for problems with unlocalized data, provided one adds appropriate multiplier weights depending on the asymptotic behavior of a'' at infinity. Some details are provided in the last section of the paper.

2.2. **Dispersion via energy estimates.** As noted above, the standard proof of the dispersive bound (2.7) is via the pointwise bounds for the frequency localized fundamental solution, which in turn follow from the method of stationary phase, see e.g. [Stein]. However, there is also an alternative, more robust approach via energy estimates and the vector field method.

Precisely, using the atomic structure of the L^1 space, it suffices to prove the $t^{-\frac{1}{2}}$ decay in (2.7) for initial data u_0 which is both frequency localized and in the Schwartz space. To measure the decay, we introduce the linear operator

def:L

$$L = x - tA'(D), \tag{2.9}$$

which is the pushforward of x along the corresponding linear flow. For Schwarz data, we control

$$\|u_0\|_{L^2} + \|Lu_0\|_{L^2} \lesssim 1. \tag{2.10}$$

Then we want to show that

isp-decay-lin

$$\|u(t)\|_{L^\infty} \lesssim t^{-\frac{1}{2}}. \tag{2.11}$$

To prove this, we observe that if u solves (2.1) then Lu also solves (2.1). Hence, using the conservation of the L^2 norm, it follows that

unif-en

$$\|u(t)\|_{L^2} + \|Lu(t)\|_{L^2} \lesssim 1. \tag{2.12}$$

Hence, one can think of the decay bound (2.11) as a consequence of a Sobolev-type interpolation bound, where the uniform norm for u is estimated in terms of the uniform energy bound in (2.12). For later use, we will state and prove a more general statement, where uniform dispersion is assumed globally and implicit constants are carefully controlled:

p:vf **Proposition 2.1.** *Assume that the symbol $a(\cdot)$ satisfies the bounds*

$$a'' \approx R, \qquad |a'''| \lesssim MR, \tag{2.13}$$

where M and R are positive constants. Then the following estimate holds for any frequency localized function u:

vf-Sobolev
$$\|u(t)\|_{L^\infty}^2 \lesssim \frac{1}{tR}(\|u\|_{L^2}\|Lu\|_{L^2} + M\|u\|_{L^2}^2). \tag{2.14}$$

In particular, this yields the following result in the context of Theorem 1.

p:vf-comp **Proposition 2.2.** *Assume that the conditions $(H1)$, $(H2)$ hold. Then the following estimate holds for any function u which is frequency localized in a fixed compact set:*

f-Sobolev-cor
$$\|u(t)\|_{L^\infty}^2 \lesssim t^{-1}(\|u\|_{L^2}\|Lu\|_{L^2} + \|u\|_{L^2}^2) \tag{2.15}$$

As discussed above, this in turn implies (2.7). Conversely, we remark that the above inequality can be obtained as a direct consequence of (2.7).

Proof of Proposition 2.1. We begin by observing that, by independently scaling the space and the time, both R and M can be seen as scaling parameters. Precisely, a linear change of coordinates $x \to Mx$ reduces the problem to the case when $M = 1$. Thus without any restriction in generality we will assume $R = 1$, $M = 1$.

A second observation is that we can regularize the symbol a on the $\sqrt{t}$ scale,

$$\tilde{a} = P_{<t^{\frac{1}{2}}}(D_\xi)a$$

The bound on a''' shows that

$$|a' - \tilde{a}'| \lesssim t^{-1}$$

so that a and $\tilde{a}$ can be used interchangeably in the Proposition. The advantage is that $\tilde{a}$ satisfies higher regularity bounds

a-high
$$|\tilde{a}^{(j)}| \lesssim t^{\frac{j-3}{2}}, \qquad j \geq 3. \tag{2.16}$$

From here on we will drop the $\tilde{a}$ notation and assume that a satisfies (2.16).

Next we introduce a secondary operator $\tilde{L}$, which can be used interchangeably with L in energy estimates. We recall that the symbol of L is

$$\ell(x, \xi) = x - ta'(\xi),$$

so its characteristic set is given by

$$\operatorname{char} L = \{a'(\xi) = x/t\}.$$

Using the property (2.4), this can be written as

$$\operatorname{char} L = \{\xi = \phi'(x/t)\}.$$

This leads us to define the operator

$$\tilde{L} := t(\partial_x - i\phi'(x/t))$$

which has symbol

$$\tilde{\ell}(x, \xi) = it(\xi - \phi'(x/t)),$$

and thus the key property that it has the same characteristic set as L. Since a' and ϕ' are inverse functions, it follows directly that

$$\phi'' \approx 1.$$

In addition, from (2.16) and differentiation rules one also obtains

phi-high
$$|\phi^{(j)}| \lesssim t^{\frac{j-3}{2}}, \qquad j \geq 3. \tag{2.17}$$

To compare the two operators L and $\tilde{L}$ we need the following Gärding type inequality

l:tL **Lemma 2.3.** *We have*

$$\|\tilde{L}u\|_{L^2} \lesssim \|Lu\|_{L^2} + \|u\|_{L^2}.$$

The conclusion of the lemma is a direct consequence of the corresponding symbol bound

$$|\tilde{\ell}(x,\xi)| \lesssim |\ell(x,\xi)|$$

via a semiclassical form of Gärding's inequality. Precisely, we will directly invoke [T-FF, 21, Theorem 3], with the semiclassical parameter $\mu = t^{-\frac{1}{2}}$. For convenience, we recall it here:

t:garding **Theorem 2.** [T-FF, 21, Theorem 3] *Let $\mu > 0$ and a_j, b be real symbols which satisfy*

a2 (2.18) $$|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \le c_{\alpha\beta} \mu^{\frac{|\alpha|-|\beta|}{2}}, \qquad |\alpha| + |\beta| \ge 2$$

so that $|b| \le \sum |a_j|$. Then

garding (2.19) $$\|B^w(x,D)u\|_{L^2} \lesssim \sum \|A_j^w(x,D)u\|_{L^2} + \|u\|_{L^2}\,.$$

The fact that this is stated in the Weyl calculus makes no difference in our context. We also refer the reader to [Delort-KG, 3], where a similar bound is derived using semiclassical calculus.

Now we return to the proof of Proposition 2.1, where as discussed above we may substitute L by $\tilde{L}$. Then all we need is a simple integration, based on the relation

$$\frac{d}{dx}|u|^2 = \frac{2}{t}\Im(\bar{u}\tilde{L}u).$$

This yields

$$\|u\|_{L^\infty}^2 \lesssim \frac{1}{t}\|u\|_{L^2}\|\tilde{L}u\|_{L^2},$$

thus completing the proof.

□

We conclude this section with one last observation, which is that, if u is assumed to be frequency localized in some compact interval I, then both the L^2 and the pointwise bounds for u are better outside a neighbourhood of the the velocity range $J = a'(I)$. The following result clarifies the proper localization scales. The analysis is identical to the left and to the right of J. Hence, in order to set the notations, we fix a frequency ξ_0, the corresponding group velocity $v_0 = a'(\xi_0)$ and the associated position at time t, $x_0 = tv_0$. We also assume without any loss of generality that a is convex.

p:vf-ell **Proposition 2.4.** *In the context of Proposition 2.1, assume in addition that u is frequency localized in $I = \{\xi < \xi_0\}$. Then we have better bounds for u outside $J = \{x < x_0\}$, as follows:*

a) L^2 bounds:

vf-ell-L2 (2.20) $$\|(x-x_0)_+ u\|_{L^2} \lesssim \|Lu\|_{L^2} + M\|u\|_{L^2},$$

b) L^∞ bounds

vf-ell-Linf (2.21) $$|u(x)|^2 \lesssim \frac{1}{|x-x_0|Rt}(\|Lu\|_{L^2} + M\|u\|_{L^2})^2, \qquad x > x_0.$$

We remark that the bounds in the proposition are only interesting in the region $\{x - x_0 > (Rt)^{\frac{1}{2}}\}$. Closer to x_0, and in effect in the full region $\{|x-x_0| \lesssim (Rt)^{\frac{1}{2}}\}$ they can be replaced by

vf-ell-near (2.22) $$|u(x)|^2 \lesssim \frac{1}{(Rt)^{\frac{3}{2}}}(\|Lu\|_{L^2} + M\|u\|_{L^2})^2, \qquad |x-x_0| \lesssim (Rt)^{\frac{1}{2}}.$$

Proof. We first note that R and M in the hypothesis of the proposition are scaling parameters, and we can simply set them equal to 1.

Let $\chi = \chi(x)$ be a spatial cutoff function supported outside J, smooth on the r scale and equal to 1 in $[x_0 + r, \infty)$, where $r > 0$ is a parameter which will be chosen later as $r = t^{\frac{1}{2}}$. We will establish the stronger L^2 bound

loc (2.23) $$\|(x-x_0)\chi u\|_{L^2} + t\|\chi(D-\xi_0)u\|_{L^2} \lesssim \|Lu\|_{L^2} + \|u\|_{L^2}$$

for u frequency localized in $\xi < \xi_0$. To start with, we verify that this implies the bounds in the proposition. This is immediate for $x > x_0 + r$, so we need to cover the remaining range. It suffices to show that

$$|u(x)| \lesssim t^{-\frac{3}{4}}(\|Lu\|_{L^2} + \|u\|_{L^2}), \qquad |x-x_0| \lesssim r,$$

which is in effect exactly the bound (2.22).

We already know this if $x-x_0 \approx r$. To capture the remaining range we recall that we can use the operators L and $\tilde{L}$ interchangeably in these bounds. Our starting point is the straightforward relation

$$\left|\frac{d}{dx}|u|\right| \lesssim \left|\frac{d}{dx}e^{-it\phi(x/t)}u\right| = \frac{1}{t}|\tilde{L}u|.$$

Then for x_1 in the full range $|x_1 - x_0| \lesssim r$ we write by the fundamental theorem of calculus

$$||u(x_1)| - |u(x)|| \lesssim t^{-1}\int_{x_1}^{x} |\tilde{L}u|dy \lesssim t^{-1}|x_1 - x|^{\frac{1}{2}}\|\tilde{L}u\|_{L^2} \lesssim t^{-\frac{3}{4}},$$

which suffices. It remains to prove (2.23).

First, using the form of the operator $\tilde{L}$, we write

$$-i\tilde{L} = t(D-\xi_0) - t(\phi'(x/t) - \phi'(x_0/t))$$

and estimate

$$t\|\chi(D-\xi_0)u\|_{L^2} \lesssim \|(x-x_0)\chi u\|_{L^2} + \|\tilde{L}u\|_{L^2}$$

and thus reduce the bound (2.23) to

loc-a (2.24) $$\|(x-x_0)\chi u\|_{L^2} \lesssim \|u\|_{L^2} + \|\tilde{L}u\|_{L^2}.$$

The next step is to discard the frequency localization, by adding a term on the right

loc-b (2.25) $$\|(x-x_0)\chi u\|_{L^2} \lesssim \|u\|_{L^2} + \|t(D-\xi_0)\eta(D-\xi_0)u\|_{L^2} + \|\tilde{L}u\|_{L^2},$$

where η selects the region $\xi - \xi_0 > \rho$, with the parameter $\rho > 0$ to be chosen later of size $\rho \approx t^{-\frac{1}{2}}$.

Here we use again the Gärding type inequality (2.19) with $\mu = t^{-\frac{1}{2}}$. At the symbol level, we need to verify that

loc-c (2.26) $$(x-x_0)\chi \lesssim t(\xi-\xi_0)\eta(\xi-\xi_0) + |x - ta'(\xi)|.$$

If $\xi > \xi_0 + \rho$ then $\eta = 1$ and the inequality is directly verified without χ. Else, for $x > x_0 + r$ and $\xi < \xi_0 + \rho$ we have

$$x - ta'(\xi) > x - x_0 - Ct\rho > r - Ct\rho, \qquad C = \sup_{\xi} a''(\xi),$$

which suffices provided that $r \geq 2Ct\rho$.

It remains to ensure that we have the correct symbol regularity, as required by Theorem 2 with $\mu = t^{-\frac{1}{2}}$. This is indeed the case provided that

$$r^{-1} \lesssim t^{-\frac{1}{2}}, \qquad \rho^{-1} \lesssim t^{\frac{1}{2}}.$$

To satisfy all of the above requirements it suffices to choose r and ρ so that

$$r \approx t^{\frac{1}{2}}, \qquad \rho \approx t^{-\frac{1}{2}}, \qquad r \geq 2Ct\rho.$$

Then the L^2 bound (2.25) follows from the symbol bound (2.26), and the proof of the proposition is complete. □

3. The asymptotic equation for the nonlinear problem

s:asymptotic

We begin the discussion by recalling the asymptotic behavior of solutions for the linear flow (2.1), namely

wish (3.1) $$u(t,x) \approx \frac{1}{\sqrt{t}}\gamma(v)e^{it\phi(v)},$$

and ask whether such a pattern is also possible for the nonlinear flow (1.1). This would require the cubic term in the equation to play a perturbative role near infinity.

However, a heuristic computation shows that this cannot happen. To see that, suppose, more generally, that for the solution u we have an asymptotic representation of the form

$$u(t,x) \approx \frac{1}{\sqrt{t}}\gamma(t,v)e^{it\phi(v)},$$

where γ is a smooth function of v, uniformly in t. Then, at (t,x), this solution has spatial frequency close to

$$\xi_v = \phi'(v).$$

Expanding the symbol for A in a Taylor series around $\xi = \xi_v$ we obtain

$$a(\xi) = a(\xi_v) + a'(\xi_v)(\xi - \xi_v) + \frac{1}{2}a''(\xi_v)(\xi - \xi_v)^2 + O(\xi - \xi_v)^3,$$

which at the operator level yields

$$A(D) = a(\xi_v) + a'(\xi_v)(i\partial_x - \xi_v) + \frac{i}{2}a''(\xi_v)\phi''(v) + O(t^{-2}) + O(t^{-1})(i\partial_x - \xi_v) + O(1)(i\partial_x - \xi_v)^2.$$

This further simplifies since ϕ' and a' are inverse functions, so $a''(\xi_v)\phi''(v) = 1$. Hence we obtain the semiclassical formula

$$A(D)u \approx t^{-\frac{1}{2}} e^{it\psi(v)} \left(a(\xi_v)\gamma(t,v) + it^{-1}a'(\xi_v)\gamma'(t,v) + \frac{i}{2}t^{-1} + O(t^{-2}) \right).$$

A similar but simpler computation shows that

$$Q(u,\bar{u},u) = t^{-\frac{1}{2}} e^{it\psi(v)} \left(t^{-1} q(\xi_v,\xi_v,\xi_v)\gamma(t,v)|\gamma(t,v)|^2 + O(t^{-2}) \right).$$

Finally, by chain rule,

$$i\partial_t u = t^{-\frac{1}{2}} e^{it\psi(v)} \left(-\gamma(t,v)(\phi(v) - v\phi'(v)) - it^{-1}v\gamma'(v) - \frac{i}{2}t^{-1} + i\gamma_t \right).$$

Substituting the last three relations into the equation, we cancel the leading terms using the eikonal equation (2.5) (which justifies the phase in our ansatz in the first place), and the γ' terms using (2.2). This leaves us with the relation

$$i\gamma_t = t^{-1} q(\xi_v,\xi_v,\xi_v)\gamma|\gamma|^2 + O(t^{-2}).$$

Since t^{-1} is not integrable at infinity, we see that it is not possible for the function γ to have a nontrivial limit at infinity, which justifies our earlier claim that an asymptotic behavior as in (3.1) cannot hold for the nonlinear evolution.

However, all is not lost. We can ensure that the last relation is satisfied if we allow a very mild dependence of γ on t, precisely if we set γ to satisfy the *asymptotic equation*

asymptotic-t

$$i\gamma_t = t^{-1} q(\xi_v,\xi_v,\xi_v)\gamma|\gamma|^2. \tag{3.2}$$

This is an ode which only has global solutions provided that

$$\Im q(\xi,\xi,\xi) \leq 0, \qquad \xi \in \mathbb{R}.$$

The case when $\Im q(\xi,\xi,\xi) < 0$ corresponds to a damping nonlinearity, and solutions for the asymptotic equation which decay to 0. The more interesting case, which we will refer to as the conservative case, is when q is real on the diagonal. In this case, the solutions to the asymptotic equation (3.2) have constant amplitude.

In all cases, we remark that the asymptotic equation can be converted into an autonomous evolution with an exponential substitution, $t = e^s$. Then (3.2) becomes

asymptotic-s

$$i\gamma_s = q(\xi_v,\xi_v,\xi_v)\gamma|\gamma|^2. \tag{3.3}$$

Hence, the objective of the analysis becomes to show that the solutions to (1.1) with small and localized data have the asymptotic behavior

u-asympt

$$u(t,x) \approx \frac{1}{\sqrt{t}}\gamma(\ln t, v)e^{it\psi(v)}, \tag{3.4}$$

where γ solves the asymptotic equation (3.3). This has solutions of the form

$$\gamma(s,v) = e^{-isq(\xi_v,\xi_v,\xi_v)|\gamma_0(v)|^2}\gamma_0(v), \tag{3.5}$$

depending on a function γ_0, which we will call the *scattering profile* for the solution u. We will refer to such an asymptotic behavior as *modified scattering*.

We remark that in this case we cannot expect γ to be uniformly regular in v as $s \to \infty$. However, this is harmless from the perspective of any asymptotic computation as above, as it only yields extra $\log t$ factors.

To summarize, we conclude that the objective of any asymptotic analysis for the equation (1.1) is two-fold:

a) Make a good choice for the profile γ, so that (3.4) holds.
b) Show that γ approximately solves the asymptotic equation (3.3).

The goal of these notes is to describe the method of *testing by wave packets*, introduced by the authors in the context of a model NLS problem in [8] and then applied to quasilinear water wave evolutions in [9], [10]. This method is applied here in combination with energy estimates, which also raise some interesting questions due to the generality of the model considered, i.e. without any direct conservation laws.

3.1. **Asymptotic equations in the NLS context.** To set the stage for the presentation of our method, we will begin by first describing several alternative ideas which were proposed over the years in the context of the cubic NLS problem in one space dimension,

NLS
$$iu_t - \frac{1}{2}\partial_x^2 u = \pm u|u|^2. \tag{3.6}$$

There one may take $a(\xi) = \frac{1}{2}\xi^2$, in which case $\phi(v) = \frac{1}{2}v^2$ and $\xi_v = v$.

A. Asymptotic equation in the Fourier space, introduced by Hayashi-Naumkin [7], and refined by Kato-Pusateri [13]. This is based on the idea that, taking a Fourier transform in an asymptotic formula like (3.4), one obtains a related asymptotic for $\hat{u}$, namely

$$\hat{u}(t,\xi) \approx \gamma(t,\xi) e^{-\frac{i}{2}t\xi^2}.$$

Defining γ by

$$\gamma(t,\xi) = e^{\frac{i}{2}t\xi^2} u(t,\xi),$$

one then seeks an asymptotic equation for the Fourier transform of the solutions,

$$\frac{d}{dt}\hat{u}(t,\xi) = -i\xi^2 \hat{u}(t,\xi) + \pm i t^{-1}\, \hat{u}(t,\xi)|\hat{u}(t,\xi)|^2 + O_{L^\infty}(t^{-1-\epsilon}),$$

where the first, respectively the second term on the right correspond to the linear, respectively the nonlinear part of the equation (1.1).

B. Asymptotic equation in the physical space, introduced by Lindblad-Soffer [16]; here the goal is to derive an asymptotic equation in the physical space along rays,

$$(t\partial_t + x\partial_x + \frac{1}{2})u(t,x) = \pm i t u(t,x)|u(t,x)|^2 + O_{L^\infty}(t^{-\epsilon}),$$

where the left hand side represents the linear contribution, while the right hand side represents the nonlinear contribution.

C. Nonlinear Fourier methods, developed by Deift-Zhou [2], who used complete integrability and inverse scattering to obtain long range asymptotics via the steepest descent method. Unfortunately, these ideas are only available in the integrable case.

D. The *wave packet testing* method, introduced by the authors in [8, 9, 10], starts from the observation that the methods described in A. and B. above lack balance when it comes to estimating the errors in the asymptotic equation. Working on the Fourier side, there are no linear errors but the nonlinear errors are large. On the physical side, there are no nonlinear errors, but instead the linear errors are large. This led to the idea of looking for a balanced way of defining the asymptotic profile γ, where the linear and nonlinear errors are smaller and comparable. This is achieved by testing the NLS solution with an approximate wave packet type linear wave $\mathbf{u}_v$,

$$\gamma(t,v) = \langle u, \mathbf{u}_v \rangle_{L^2},$$

where $\mathbf{u}_v$ is both spatially localized in a $t^{\frac{1}{2}}$ neighbourhood of the ray $x = vt$, and frequency localized in a dual $t^{-\frac{1}{2}}$ neighbourhood of the frequency $\xi_v = v$. This perfectly balances the linear and the nonlinear errors, and leads to results which are near optimal with respect to the regularity and decay of the initial data.

s:compact

4. Global solutions for small localized data: the model case

In order to avoid distracting technicalities, in our first result we will make the simplifying assumption

(H5) *Frequency localized nonlinearity.* The symbol q is compactly supported.

This assumption makes the behavior of a at infinity irrelevant. Using the operator L, we define the following weighted time dependent function space X:

$$\|u(t)\|_X^2 := \|u(t)\|_{L^2} + \|Lu(t)\|_{L^2}^2.$$

This will be used both for the initial data, and in order to measure the solution as it evolves in time. In particular, at time $t = 0$ the X norm measures the localization of the initial data u_0,

$$\|u_0\|_X \approx \|(1+x^2)^{\frac{1}{2}} u_0\|_{L^2}.$$

With these notations, our main result is as follows:

t:comp **Theorem 3.** *Assume that the conditions (H1-5) above are satisfied, and that the initial data for our equation* (1.1) *satisfies:*

small
$$\|u(0)\|_X \lesssim \epsilon \ll 1. \tag{4.1}$$

Then the solution exists globally in time, with energy bounds

energy
$$\|u(t)\|_X \lesssim \epsilon t^{C\epsilon^2}, \tag{4.2}$$

and pointwise decay

disp-decay
$$\|u(t)\|_{L^\infty} \lesssim \frac{\epsilon}{\sqrt{t}}. \tag{4.3}$$

The rest of this section contains the proof of this result, organized as follows. Section 4.1 provides the set-up for the main bootstrap argument. The energy estimates leading to the bound (4.2) are discussed in Subsection 4.2; this includes the energy bound in Proposition 4.1 and the vector field bound in Proposition 4.2. Thus we arrive at the main objective of the paper, namely the wave packet analysis, which is considered in Subsection 4.3.

Finally, in Subsection 4.6 we briefly discuss the inverse problem, which is to reconstruct a solution given its asymptotic profile.

s:boot

4.1. Overview of the proof: A bootstrap argument.

The starting point of the proof is to make a bootstrap assumption for the pointwise bound,

boot
$$\|u(t)\|_{L^\infty} \lesssim C\epsilon \langle t\rangle^{-\frac{1}{2}}. \tag{4.4}$$

Then the proof proceeds in two steps:

1. Energy estimates: Here the objective is to establish the energy bound

o:energy
$$\|u(t)\|_X \lesssim \langle t\rangle^{C^2\epsilon^2} \|u(0)\|_X. \tag{4.5}$$

This uses Gronwall's inequality in the equation for u, and then a cubic correction to Lu. We note that, by the vector field bound in Proposition 2.1, this yields

o:first-point
$$\|u(t)\|_{L^\infty} \lesssim t^{-\frac{1}{2}} \|u(t)\|_X \lesssim \epsilon \langle t\rangle^{C^2\epsilon^2}. \tag{4.6}$$

This step is carried out in Section 4.2.

2. Pointwise bounds: Here the goal is to improve the bootstrap assumption, and show that

o:point
$$\|u(t)\|_{L^\infty} \lesssim \epsilon \langle t\rangle^{-\frac{1}{2}}. \tag{4.7}$$

This step is carried out in Section 4.3, and uses the method of *testing with wave-packets* to produce an asymptotic profile $\gamma(t, v)$, which may be compared to the solution u using the bounds (4.5), respectively (4.6). Then it remains to prove an ϵ bound for γ, which is achieved by showing that γ is a good approximate solution for the asymptotic equation (3.2).

s:energy

4.2. **Energy estimates.** Our objective here is to prove energy estimates for u and Lu, i.e. the bound (4.5). In the case of u, we have the following straightforward result:

p:energy **Proposition 4.1.** *Assume that $u \in L^2$ solves* (1.1)*. Then*

ee
$$\frac{d}{dt}\|u\|_{L^2}^2 \lesssim \|u\|_{L^\infty}^2 \|u\|_{L^2}^2. \tag{4.8}$$

We note that by Gronwall's inequality, this gives

energy-first
$$\|u(t)\|_{L^2} \lesssim \epsilon \langle t \rangle^{C^2\epsilon^2} \tag{4.9}$$

which is the first half of (4.5).

Proof. Multiplying equation (1.1) with $\bar{u}$, integrating over x, (H1) assumption, and adding the complex conjugate counterpart we obtain

$$\frac{d}{dt}\|u\|_{L^2}^2 = \int \left[u_t \bar{u} + \bar{u}_t u\right] \, dx = -2\Im \int Q(u, \bar{u}, u)\bar{u} \, dx.$$

Thus,

$$\left|\frac{d}{dt}\|u\|_{L^2}^2\right| = \left|\int \left[u_t \bar{u} + \bar{u}_t u\right] \, dx\right| \lesssim \Im \int |Q(u, \bar{u}, u)\bar{u}| \, dx \lesssim \|u\|_{L^\infty}^2 \|u\|_{L^2}^2,$$

where for the last inequality we used the (H3) assumption on the nonlinearity Q and pulled out the $L^\infty-$norm of $|u|^2$. Here one could think of the bound for Q as a trilinear product bound, using for instance the idea of separation of variables discussed in a more general setting in Section 5.1.

□

The more delicate matter is the energy estimate for Lu, which solves the equation

$$(i\partial_t - A(D))Lu = LQ(u, \bar{u}, u). \tag{4.10}$$

The difficulty is that the source term on the right does not directly satisfy a perturbative bound, e,g, of the form

Lu-not
$$\|LQ(u, \bar{u}, u)\|_{L^2} \lesssim \|u\|_{L^\infty}^2 (\|Lu\|_{L^2} + \|u\|_{L^2}). \tag{4.11}$$

To address this issue, we will add a nonlinear correction to Lu, precisely

$$L^{NL}u = Lu + tC(u, \bar{u}, u),$$

where C is a well chosen trilinear form which has a smooth compactly supported symbol. Precisely, we have the following:

p:Lu **Proposition 4.2.** *There exists a smooth, compactly supported symbol $c(\cdot, \cdot, \cdot)$ with the property that the following estimate holds for solutions u to* (1.1)*:*

dt-Lu
$$\frac{d}{dt}\|L^{NL}u\|_{L^2} \lesssim \|u\|_{L^\infty}^2 \|L^{NL}u\|_{L^2} + \|u\|_{L^\infty}^2 \|u\|_{L^2} + t\|u\|_{L^\infty}^4 \|u\|_{L^2}. \tag{4.12}$$

In essence, the correction C will be chosen so that a modified version of (4.11) holds; precisely, that is the bound (4.17) in the proof below.

Given our bootstrap assumption (4.4) and the L^2 estimate (4.9) for u, Gronwall's inequality allows us to close the energy estimate for $L^{NL}u$ and obtain

energy-second
$$\|L^{NL}u(t)\|_{L^2} \lesssim \epsilon \langle t \rangle^{C^2\epsilon^2}. \tag{4.13}$$

Here we can use L^{NL} and L interchangeably since by (4.4) and the L^2 estimate (4.9) we have a good bound for the difference,

C-L2
$$\|tC(u, \bar{u}, u)\|_{L^2} \lesssim t\|u\|_{L^\infty}^2 \|u\|_{L^2} \lesssim \langle t \rangle^{C^2\epsilon^2}. \tag{4.14}$$

Hence, the second part of (4.5) also follows.

Proof. We write the equation for $L^{NL}u$ in the form

$$(i\partial_t - A(D))L^{NL}u = LQ(u,\bar u,u) + iC(u,\bar u,u) - tR_3(u,\bar u,u) - tR_5(u,\bar u,u,\bar u,u) := f,$$

where R_3 and R_5 are translation invariant multilinear forms with smooth compactly supported symbols, r_3, and r_5 respectively. Furthermore, the symbol of R_3 is given by

$$r_3(\xi_1,\xi_2,\xi_3) := c(\xi_1,\xi_2,\xi_3)(a(\xi_1) - a(\xi_2) + a(\xi_3) - a(\xi_1-\xi_2+\xi_3)).$$

To prove (4.12) it suffices to have the following bound for the above source term f:

f-L2 (4.15) $$\|f\|_{L^2} \lesssim \|u\|_{L^\infty}^2 \|Lu\|_{L^2} + \|u\|_{L^\infty}^2 \|u\|_{L^2} + t\|u\|_{L^\infty}^4 \|u\|_{L^2}.$$

Here we used (4.14) to replace $L^{NL}u$ by Lu in the right.

The terms $iC(u,\bar u,u)$ respectively $tR_5(u,\bar u,u,\bar u,u)$ can be directly estimated by the second, respectively the third term on the right in (4.15), without using any structural properties of the corresponding symbols. Hence it remains to consider the expression

$$LQ(u,\bar u,u) - tR_3(u,\bar u,u).$$

Our objective will be to choose the trilinear form C with the property that the bound (4.11) holds for the above expression. The choice of the symbol c is given by the following algebraic division Lemma:

l:division **Lemma 4.3.** *There exist smooth, compactly supported symbols c, c_1, c_2 and c_3 so that the following algebraic relation holds:*

division (4.16) $$\ell(x,\xi)q(\xi_1,\xi_2,\xi_3) - tc(\xi_1,\xi_2,\xi_3)(a(\xi_1) - a(\xi_2) + a(\xi_3) - a(\xi)) = \sum_{j=1}^{3} c_j(\xi_1,\xi_2,\xi_3)\ell(x,\xi_j)$$

whenever $\xi = \xi_1 - \xi_2 + \xi_3$.

We first use Lemma 4.3 to complete the proof of Proposition 4.2. The relation (4.16) translates into the following operator identity:

$$LQ(u,\bar u,u) - tR_3(u,\bar u,u) = C_1(Lu,\bar u,u) + C_2(u,\overline{Lu},u) + C_3(u,\bar u,Lu) + D(u,\bar u,u),$$

where D has symbol

$$d(\xi_i,\xi_2,\xi_3) = i(\partial_{\xi_1}c_1 - \partial_{\xi_2}c_2 + \partial_{\xi_3}c_3).$$

This directly implies the bound

Lu-not+ (4.17) $$\|LQ(u,\bar u,u) - tR_3(u,\bar u,u)\|_{L^2} \lesssim \|u\|_{L^\infty}^2(\|Lu\|_{L^2} + \|u\|_{L^2}),$$

and the proof of Proposition 4.2 is concluded.

□

It remains to prove Lemma 4.3.

Proof of Lemma 4.3. We start with some simplifications. Without any restriction in generality we can set $q = 1$, with the minor proviso that now we discard the compact support property for c and c_j. Secondly, we can separate the x and the t component of the above identity, and conclude that we need to satisfy two identities:

$$c_1 - c_2 + c_3 = 1,$$

and

$$c_1a_\xi(\xi_1) - c_2a_\xi(\xi_2) + c_3a_\xi(\xi_3) = a_\xi(\xi) + c(\xi_1,\xi_2,\xi_3)(a(\xi_1) - a(\xi_2) + a(\xi_3) - a(\xi)).$$

Simplifying further, we set $c_j = 1$ so that the first identity is trivially satisfied. We are left with

$$c(\xi_1,\xi_2,\xi_3) = \frac{a_\xi(\xi_1) - a_\xi(\xi_2) + a_\xi(\xi_3) - a_\xi(\xi)}{a(\xi_1) - a(\xi_2) + a(\xi_3) - a(\xi)},$$

where we need to show that the quotient is smooth.

Since a is strictly convex (or concave), it is easily seen that the denominator can only vanish on the set

$$\mathcal{D} = \{\xi_1 + \xi_3 = \xi_2 + \xi\}$$

if and only if

$$\{\xi_1,\xi_3\} = \{\xi,\xi_2\}.$$

We claim that the denominator admits a representation of the form

d4a (4.18) $$a(\xi_1) - a(\xi_2) + a(\xi_3) - a(\xi) = (\xi - \xi_1)(\xi - \xi_3) b(\xi_1, \xi_2, \xi_3)$$

with b smooth and nonzero.

We start with the standard representation

$$a(\xi_1) - a(\xi) = (\xi - \xi_1) a_1(\xi_1, \xi)$$

with smooth, symmetric a_1, and then write on $\mathcal{D}$

$$a(\xi_1) - a(\xi_2) + a(\xi_3) - a(\xi) = (\xi - \xi_1)(a_1(\xi_1, \xi) - a_1(\xi_3, \xi_2)).$$

Then we repeat the process for a_1 to pull out a factor of

$$\xi - \xi_3 = \xi_1 - \xi_2.$$

This yields a representation as in (4.18), with a smooth b. It remains to verify that b is nonzero, for which we compute b on the zero set. Suppose for instance that $\xi = \xi_1$ and $\xi_2 = \xi_3$. Then $a_1(\xi, \xi_1) = a'(\xi_1)$, and

$$b = \frac{a'(\xi) - a'(\xi_3)}{\xi - \xi_3},$$

which is nonzero due to the strict convexity (concavity) of a. We also remark that at the double zero, when all frequencies are equal, we have

$$b = a''(\xi) \neq 0.$$

Next we consider the numerator, for which we also have a representation

d4da (4.19) $$a_\xi(\xi_1) - a_\xi(\xi_2) + a_\xi(\xi_3) - a_\xi(\xi) = (\xi - \xi_1)(\xi - \xi_3) b_1(\xi_1, \xi_2, \xi_3).$$

Here b_1 is again smooth, but not necessarily nonzero.

Finally, we divide the expressions in (4.18) and (4.19) to obtain

$$c(\xi_1, \xi_2, \xi_3) = \frac{b_1(\xi_1, \xi_2, \xi_3)}{b(\xi_1, \xi_2, \xi_3)},$$

which is easily seen to be smooth as the denominator is nonzero.

This concludes the proof of the lemma. □

s:wp 4.3. **Wave packet testing.** Our objective here is to describe the method of testing by wave packets, and show how it applies to our problem in order to define the asymptotic profile γ and to show that is approximately solves the asymptotic equation, with the final objective of establishing the uniform bound for γ, which in turn implies the uniform bound for the solution u.

We will begin with a short description of wave packes on a fixed spatial scale, both for the linear and then for the nonlinear model. Then we discuss the wave packets on a time dependent scale, which are critical in our analysis here. Finally, we use these wave packets to construct the asymptotic profile γ, and prove that it has the desired properties.

4.3.1. *Linear wave packets on a fixed scale.* The idea here is to look for solutions to the linear equation (2.1) which are localized near a trajectory for the Hamilton flow,

$$(x_0, \xi_0) \to (x_0 + t a_\xi(\xi_0), \xi_0).$$

Most desirably, this localization should occur both in position and in frequency, on the sharp, uncertainty principle scale. The localization scales are denoted as follows:

$$(\delta x, \delta \xi), \qquad \delta x \cdot \delta \xi \approx 1 \ \ \text{(uncertainty principle)}$$

The first step is to choose these scales so that this localization is coherent up to a given time T. Heuristically, the varying group velocities within the $\delta\xi$ range leads to position variations for the Hamilton flow up to the time T, which are given by

$$\delta x = T a_{\xi\xi}(\xi_0) \delta \xi.$$

Matching this with the uncertainty principle relation, we obtain the localization scales adapted to the time scale T, namely

$$\delta x = T^{\frac{1}{2}} a_{\xi\xi}(\xi_0)^{\frac{1}{2}}, \qquad \delta \xi = T^{-\frac{1}{2}} a_{\xi\xi}^{-\frac{1}{2}}.$$

So far we have only looked at the coherence at the level of the Hamilton flow. Next, we ask whether one can realize this localization at the level of actual solutions. This leads to the so called wave packet solutions, which are approximately of the form

$$u(x,t) \approx \gamma\ \chi((\delta x)^{-1}(x - x_0 - ta_\xi(\xi_0)))e^{i(x\xi_0 - ta(\xi_0))}.$$

Here one can adopt two equally useful view points. On one hand, keeping the Schwartz function χ independent of t, one obtains an approximate solution to (2.1), with errors which are small[1] up to time T. On the other hand, one can start with a given Schwartz function χ at $t = 0$, and show that the representation above persists exactly with a time dependent Schwartz function χ which satisfies uniform bounds up to time T. This can be achieved via Fourier analysis, but also via energy estimates, using the operator L defined above, as well as its powers. This philosophy applies as well in variable coefficient case, see e.g. [KT][15].

One can think of general L^2 solutions to the linear flow (2.1) as linear, square summable superpositions of wave packets, which can be taken either relative to a discrete set of centers (x, ξ) (wave packet parametrices) or with respect to a continuous set of centers, akin to phase space transform methods[2]

Finally, we remind the reader that, under the name of Knapp counterexamples, wave packets have been used to show that Strichartz and the dispersive estimates are sharp.

4.3.2. *Nonlinear wave packet solutions on a fixed scale.* Here we switch to the nonlinear flow (1.1), and consider wave packet solutions, which are localized on scales similar to the ones above. The new factor here is the amplitude of the nonlinearity, which we denote by $\mathcal{M}$. Then the linear ansatz for wave packets is modified to

$$u(x,t) \approx \gamma(t, (\delta x)^{-1}(x - x_0 - ta_\xi(\xi_0)))e^{i(x\xi_0 - ta(\xi_0))},$$

where the modulation factor γ is taken to have size $\mathcal{M}$.

For functions u with wave packet localization near frequency ξ_0, it turns out that the nonlinearity is well approximated by

$$Q(u, \bar{u}, u) = q(\xi_0, \xi_0, \xi_0)|u|^2 u + O(\delta\xi|\mathcal{M}|^3),$$

where the error can be thought off as perturbative provided that the amplitude is small enough,

$$\delta\xi\mathcal{M}^2 \ll 1.$$

Assuming this is the case, the amplitude function γ should approximatively solve the asymptotic ode

$$i\dot{\gamma} = q(\xi_0, \xi_0, \xi_0)\gamma|\gamma|^2.$$

This in turn is conservative if $q(\xi_0, \xi_0, \xi_0)$ is real.

One should relate these heuristics with the idea of NLS approximation, which roughly asserts that solutions with this type of localization and amplitude are well approximated by solutions to a suitable NLS problem, obtained by replacing the symbol a with its quadratic approximation at ξ_0, and the cubic form Q by $q(\xi_0, \xi_0, \xi_0)|u|^2 u$. For more information on this we refer the reader to [NLS-a][11], [19] and further references therein.

4.3.3. *Linear wave packets with time dependent scale.* Working with packets with fixed scales is useful for the study of the local problem, but not so much for the global in time evolution. Because of this, we will now consider global in time approximate wave packet solutions for the linear problem (2.1). To understand their structure, we recall that the spatial scales associated to time scale t at velocity v and associated frequency $a'(\xi_v) = v$ are given by

$$\delta x = t^{\frac{1}{2}}[a_{\xi\xi}(\xi_v)]^{\frac{1}{2}}, \quad \delta\xi = t^{-\frac{1}{2}}[a_{\xi\xi}(\xi_v)]^{-\frac{1}{2}}.$$

We now replicate the previous wave packet ansatz, but do it globally in time, with a time dependent scale. Thus we define the linear wave packet $\mathbf{u}_v$ associated with velocity v by

wp-def

$$\mathbf{u}_v = a''(\xi_v)^{-\frac{1}{2}}\chi(y)e^{it\phi(x/t)}, \qquad y = \frac{x - vt}{t^{\frac{1}{2}}a''(\xi_v)^{\frac{1}{2}}}. \tag{4.20}$$

where χ is a compactly supported smooth function, which we normalize so that

$$\int \chi(y)dy = 1.$$

[1] Say in $L^1_t L^2_x$.

[2] a.k.a. the Bargman or the FBI transform, see e.g. [Pisa][22].

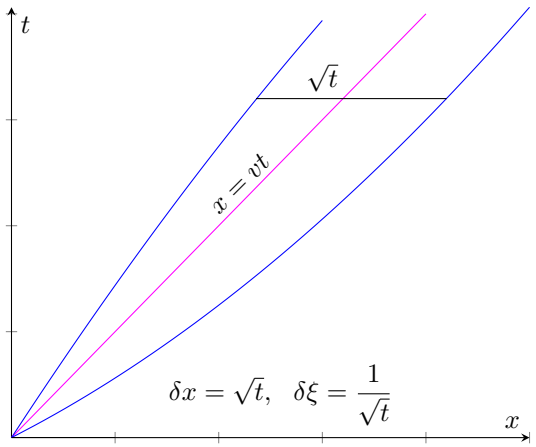

FIGURE 1. The support of a wave packet with velocity v.

This is a good approximate solution for the linear flow on dyadic time scales:

uu-approx
$$(i\partial_t - A(D))\mathbf{u}_v \approx O(t^{-1})\mathbf{u}. \tag{4.21}$$

However, we carefully remark that this is not a good approximate solution globally in time. Indeed, any global solution should disperse, rather than stay concentrated near a ray. As a corollary of this remark, we note that the above relation will still remain satisfied if we replace $\mathbf{u}_v$, say, by $t^\mu \mathbf{u}_v$. The choice we made above is for convenience only, and not at all intrinsic. If one wanted for instance to have solutions which stay bounded in L^2, then choosing $\mu = -\frac{1}{4}$ would be the appropriate choice.

However, there is one advantage for our normalization, which is seen when one attempts to gain a better understanding of the error term in the linear equation for $\mathbf{u}_v$. Precisely we have the following:

l:Puu
Lemma 4.4. *The wave packet $\mathbf{u}_v$ defined above solves a linear equation of the form*

wp-eqn
$$(i\partial_t - A(D))\mathbf{u}_v = t^{-\frac{3}{2}} L\mathbf{u}_v^I + t^{-\frac{3}{2}}\mathbf{r}_v, \tag{4.22}$$

where $\mathbf{u}_v^I$ and $\mathbf{r}_v$ have a wave packet form similar to $\mathbf{u}_v$.

Compared to (4.21), the lemma provides a more accurate description of the t^{-1} term. The function $\mathbf{u}_v^I$ here is quite explicit,

$$\mathbf{u}_v^I = -\frac{i}{2a''}(\chi'(y) - iy\chi(y))e^{it\phi(x/t)}.$$

This is not important later on, what matters is that the operator L is applied to it. The function $\mathbf{r}_v$ is less explicit but this is also not important as $\mathbf{r}_v$ only plays a perturbative role later on.

Proof. While not absolutely necessary, here it is helpful to simplify the problem using some simple linear transformations:

- Using a Galilean transformation $x - vt \to x$, the problem reduces to the case $v = 0$. Note that this changes a by a linear term.
- Using a spatial phase shift, $\mathbf{u}_v$ to $\mathbf{u}_v e^{-ix\xi_v}$, we can also ensure that $\xi_v = 0$. This translates a by ξ_v.
- Using a temporal phase shift, the problem reduces also to the case $a(0) = 0$.
- If $a''(0) < 0$, we can shift to $a''(0) > 0$ by replacing u with $\bar{u}$ (and thus $a(\xi)$ by $a(-\xi)$).

After these simplifications, we are now in the case when

$$v = 0, \quad \xi_v = 0, \quad a(0) = 0, \quad a'(0) = 0.$$

This in turn implies that

$$\phi(0) = 0, \quad \phi'(0) = 0, \quad \phi''(0) = \frac{1}{a''(0)}.$$

Now we finally compute the equation for $\mathbf{u}_0$. For this, we use the Taylor expansion of a at 0,

$$a(\xi) = \frac{1}{2}a''(0)\xi^2 + O(\xi^3).$$

The contribution of the ξ^3 sized error has size $t^{-\frac{3}{2}}$, and may be included into $\mathbf{r}_0$. Similarly, we have

$$\ell(x,\xi) = x - ta'(\xi) = x - t\xi a''(0) + O(t\xi^2),$$

where the contribution of the $t\xi^2$ tail also can be included into $\mathbf{r}_0$.

Since the v in the lemma was set to 0, in what follows we use the notation $v = x/t$. Hence we have

$$\begin{aligned}(i\partial_t - A(D))\mathbf{u}_0 &= (i\partial_t + \frac{1}{2}a''(0)\partial_x^2)\mathbf{u}_0 + t^{-\frac{3}{2}}\mathbf{r}_0\\ &= -\frac{i}{2}t^{-\frac{3}{2}}x(a''(0))^{-1}\chi'(y)e^{it\phi(v)} - (a''(0)^{-\frac{1}{2}}(\phi(v) - v\phi'(v))\chi(y)e^{it\phi(v)}\\ &\quad + \frac{1}{2}(a''(0))^{-\frac{1}{2}}\left(\frac{\chi''(y)}{t} + 2i(a'')^{\frac{1}{2}}\frac{\chi'(y)\phi'(v)}{t^{\frac{1}{2}}} - a''(0)\chi(y)\phi'^2(v) + ia''(0)\frac{\chi(y)\phi''(v)}{t}\right)e^{it\phi} + t^{-\frac{3}{2}}\mathbf{r}_0.\end{aligned}$$

Noting the leading order cancellation

$$\phi(v) - v\phi'(v) = -\frac{1}{2}a''(0)\phi'^2(v) + O(v^3),$$

where the last term only contributes to the error, we obtain

$$\begin{aligned}(i\partial_t - A(D))\mathbf{u}_0 &= -\frac{i}{2}t^{-\frac{3}{2}}(x + ita''(0)\partial_x)[a''(0)^{-1}\chi'(y)e^{it\phi(v)}] + \frac{i}{2}a''(0)^{\frac{1}{2}}\partial_x(\chi(y)\phi'(v))e^{it\phi(v)} + t^{-\frac{3}{2}}\mathbf{r}_0\\ &= -\frac{i}{2}t^{-\frac{3}{2}}(x + ita''(0)\partial_x)[a''(0)^{-1}\chi'(y)e^{it\phi(v)}] + \frac{i}{2}a''(0)^{\frac{1}{2}}(\partial_x - i\phi'(v))[\chi(y)\phi'(v)e^{it\phi(v)}] + t^{-\frac{3}{2}}\mathbf{r}_0.\end{aligned}$$

Since $\phi'(v) = (a''(0))^{-1}v + O(v^2)$, we can rewrite the second term on the right to get

$$\begin{aligned}(i\partial_t - A(D))\mathbf{u}_0 &= -\frac{i}{2}a''(0)^{-1}t^{-\frac{3}{2}}(x + ita''(0)\partial_x)\left[\chi'(y)e^{it\phi(v)} + iy\chi(y)e^{it\phi(v)}\right] + t^{-\frac{3}{2}}\mathbf{r}_0\\ &= -\frac{i}{2}a''(0)^{-1}t^{-\frac{3}{2}}L\left[(\chi'(y) + iy\chi(y))e^{it\phi(v)}\right] + t^{-\frac{3}{2}}\mathbf{r}_0\end{aligned}$$

as needed.

□

We also need to consider the v dependence of $\mathbf{u}_v$.

l:dv-uv **Lemma 4.5.** *The wave packet $\mathbf{u}_v$ defined above solves a linear equation of the form*

dv-wp (4.23) $$\partial_v\mathbf{u}_v = L\mathbf{u}_v^{II} + \mathbf{r}_v,$$

where $\mathbf{u}_v^{II}$ and $\mathbf{r}_v$ have a wave packet form similar to $\mathbf{u}_v$.

Here we have

$$\mathbf{u}_v^{II} = -i[a''(\xi_v)]^{-\frac{3}{2}}\chi(y)e^{it\phi(x/t)}.$$

Proof. Differentiating with respect to v yields

$$\partial_v\mathbf{u}_v = -t\partial_x[\chi(y)]e^{it\phi(x/t)} + \mathbf{r}_v,$$

and the first term on the right is similar to the second term on the right in the computation in the previous lemma. □

4.3.4. *Wave packet testing.* As described earlier, we will define our asymptotic profile function γ by

def-gamma (4.24) $$\gamma(t,v) = \langle u, \mathbf{u}_v\rangle_{L^2}.$$

Now our objective is two-fold:

- To show that γ provides a good approximation for u, in the sense of (3.4).
- To show that γ is an approximate solution for the asymptotic equation (3.2).

s:bd-gamma

4.4. **Bounds for γ.** Here we establish some base-line bounds for γ, using the energy estimates in Proposition 4.1:

p:gamma **Proposition 4.6.** *Assume that u satisfies the energy bounds in Proposition 4.1. Then γ satisfies*

$$\|\gamma\|_{L^2_v} + \|\partial_v \gamma\|_{L^2_v} \lesssim \epsilon t^{C^2\epsilon}. \tag{4.25}$$

Proof. If we bound $\mathbf{u}_v$ by

$$|\mathbf{u}_v(t,x_1)| \lesssim \frac{1}{1+t|v-v_1|^2}, \qquad x_1 = tv_1,$$

then the L^2 bound for γ can be interpreted as a convolution estimate, as

$$|\gamma(v)| \lesssim t|u(vt)| *_v \frac{1}{1+tv^2},$$

where the convolution kernel is integrable. By Young's inequality this yields

$$\|\gamma\|_{L^2_v} \lesssim \sqrt{t}\|u(vt)\|_{L^2_v} \lesssim \|u\|_{L^2_x}$$

as needed.

For the L^2 bound for $\partial_v\gamma$ we first apply Lemma 4.5. Then we obtain the convolution bound

$$|\partial_v\gamma(v)| \lesssim t(|Lu|+|u|)(vt) *_v \frac{1}{1+tv^2},$$

and then conclude as above. □

4.4.1. *Approximate profile.* Our goal here is to estimate the difference

$$r(t,x) = u(t,x) - \frac{1}{\sqrt{t}}\gamma(t,v)e^{it\phi(v)}, \qquad v = \frac{x}{t}$$

as follows:

p:u-gamma **Proposition 4.7.** *Assume that u satisfies the energy bounds in Proposition 4.1. Then the above error r satisfies the uniform bound*

$$\|r\|_{L^\infty} \lesssim \epsilon t^{-\frac{3}{4}+C^2\epsilon}, \tag{4.26}$$

and the L^2 bound

$$\|r\|_{L^2} \lesssim \epsilon t^{-\frac{1}{2}+C^2\epsilon}. \tag{4.27}$$

Proof. We represent

$$\sqrt{t}e^{-it\phi(v)}r(t,tv) = \langle u, \mathbf{w}_v\rangle,$$

where on the right we use the L^2_x pairing, with

$$\mathbf{w}_v = \sqrt{t}e^{it\phi}\delta_{x=vt} - \mathbf{u}_v.$$

Using the normalization $\int \chi = 1$ we rewrite $\mathbf{w}_v$ as

$$\mathbf{w}_v = t^{\frac{1}{2}}\partial_x(\chi_1(y)\operatorname{sgn}(y))e^{it\phi} = t^{\frac{1}{2}}(\partial_x - i\phi'(x/t))[\chi_1(y)\operatorname{sgn}(y)e^{it\phi}],$$

where χ_1 is

$$\chi_1(y) = \begin{cases} -\int_{-\infty}^y \chi(z)dz & y<0 \\ -\int_y^\infty \chi(z)dz & y>0, \end{cases}$$

which leads to

$$\sqrt{t}e^{-it\phi(v)}r(t,tv) = t^{-\frac{1}{2}}\langle \tilde{L}u, \mathbf{u}_v^{III}\rangle, \qquad \mathbf{u}_v^{III} = \chi_1(y)\operatorname{sgn}(y)e^{it\phi}.$$

Now $\mathbf{u}_v^{III}$ has the same size and localization as $\mathbf{u}$, so we can argue as in the proof of Proposition 4.6 that

$$|r(t,vt)| \lesssim |\tilde{L}u(t,vt)| *_v \frac{1}{1+tv^2}.$$

Then by Young's inequality we conclude that

$$\|r(t,x)\|_{L^2_x} = t^{\frac{1}{2}}\|r(t,vt)\|_{L^2_v} \lesssim \|\tilde{L}u(t,vt)\|_{L^2_v} = t^{-\frac{1}{2}}\|\tilde{L}u(t,x)\|_{L^2_x}$$

respectively

$$\|r\|_{L^\infty} \lesssim t^{-\frac{1}{4}} \|\tilde{L}u(t,vt)\|_{L^2_v} = t^{-\frac{3}{4}} \|\tilde{L}u\|_{L^2_x}.$$

Now we can conclude using Lemma 2.3. □

4.4.2. *The asymptotic equation for γ.* Here we prove the following:

p:gamma-er **Proposition 4.8.** *Assume that u satisfies the energy bounds*

$$\|u\|_X \lesssim \epsilon \langle t \rangle^{C^2\epsilon^2}.$$

Then γ solves the asymptotic equation

gamma-asympt
$$\dot{\gamma}(t,v) = iq(\xi_v,\xi_v,\xi_v)t^{-1}\gamma(t,v)|\gamma(t,v)|^2 + f(t,v), \tag{4.28}$$

where f satisfies the uniform bound

e:f-inft
$$\|f\|_{L^\infty} \lesssim \epsilon t^{-\frac{5}{4}+3C^2\epsilon}, \tag{4.29}$$

and the L^2 bound

e:f-l2
$$\|f\|_{L^2_v} \lesssim \epsilon t^{-\frac{3}{2}+3C^2\epsilon}. \tag{4.30}$$

Proof. We compute

$$\dot{\gamma}(t,v) = -i\langle (i\partial_t - A(D))u, \mathbf{u}_v\rangle + i\langle u, (i\partial_t - A(D))\mathbf{u}_v\rangle := I_1(t,v) + I_2(t,v).$$

For I_2 we use Lemma 4.4 to write

$$I_2(t,v) = it^{-\frac{3}{2}}(\langle Lu, \mathbf{u}^I_v\rangle + \langle u, \mathbf{r}_v\rangle).$$

This allow us to bound its size both in L^∞, using Hölder's inequality, and in L^2 via convolution bounds.

The expression I_1, on the other hand, has the form

$$I_1(t,v) = i\langle Q(u,\bar{u},u), \mathbf{u}_v\rangle.$$

Here we first use the bounds for r in Proposition 4.7 in order to substitute u with $t^{-\frac{1}{2}}\gamma(t,x/t)e^{it\phi(x/t)}$ modulo acceptable errors,

$$I_1(t,v) = it^{-\frac{3}{2}}\gamma(t,v)|\gamma(t,v)|^2\langle Q(\gamma(t,x/t)e^{it\phi}, \bar{\gamma}(t,x/t)e^{-it\phi}, \gamma(t,x/t)e^{it\phi}), \mathbf{u}_v\rangle + f,$$

where the error f is as in (4.29), (4.30).

Then we take advantage of the fact that the kernel of Q is localized on the unit scale in order to replace $\gamma(t,x/t)$ with $\gamma(t,v)$, again with acceptable errors, which are estimated using the bounds for $\partial_v\gamma$ in Proposition 4.6. Thus we get

$$I_1(t,v) = it^{-\frac{3}{2}}\gamma(t,v)|\gamma(t,v)|^2\langle Q(e^{it\phi}, e^{-it\phi}, e^{it\phi}), \mathbf{u}_v\rangle + f.$$

Finally, a semiclassical computation shows that

$$Q(e^{it\phi}, e^{-it\phi}, e^{it\phi}) = q(\xi_v,\xi_v,\xi_v)e^{it\phi} + O(t^{-1}),$$

so the desired asymptotic equation follows.

□

4.5. **Conclusion.** Here we show how to close the bootstrap argument, and prove that the global result follows as a consequence of the results in Propositions 4.7 and 4.28.

The bootstrap argument closes as follows:

- Proposition 4.2 gives the energy bounds on our corrected vector field L^{NL}, where we made use of the energy estimates on u obtained in (4.9) and of the bootstrap assumption expressed in (4.4) to get

$$\|L^{NL}u(t)\|_{L^2} \lesssim \epsilon \langle t \rangle^{C^2\epsilon^2}.$$

- We use the pointwise decay bound obtained in Proposition 4.7 for the difference between the asymptotic profile $\gamma(t,v)$ and solution u to conclude that the error term f in the asymptotic equation (4.28) for γ is acceptable, i.e. has better than t^{-1} decay, as stated in Proposition 4.8. Integrating (4.28) leads to a pointwise bound for γ:

$$|\gamma(t,v)| \leq |\gamma(1,v)| + \int_1^t |f(s,v)|\, ds.$$

Here, we use the energy bound (4.9) at time $t = 1$ and the pointwise bound on γ to conclude that

$$\|\gamma(1,v)\|_{L^\infty} \lesssim \epsilon,$$

as well as the pointwise bound on f given in Proposition 4.8, and get

$$|\gamma(t,v)| \leq |\gamma(1,v)| + \int_1^t |f(s,v)|\, ds \lesssim \epsilon.$$

- Lastly, from the above estimate and Proposition 4.7, it follows that the pointwise bound on u is

$$|u| \lesssim \epsilon t^{-\frac{1}{2}},$$

which, under the constraint $1 \ll C$, concludes the bootstrap argument.

s:complete

4.6. **Modified scattering and asymptotic completeness.** An immediate consequence of the approximate asymptotic equation (4.28) for γ is that, as $t \to \infty$, the function γ converges to a solution $\tilde{\gamma}$ to the exact asymptotic equation,

$$\dot{\tilde{\gamma}} = iq(\xi_v, \xi_v, \xi_v) t^{-1} \tilde{\gamma}(t,v) |\tilde{\gamma}(t,v)|^2,$$

which can be represented in the form

$$\tilde{\gamma}(t,v) = W(v) e^{iq(\xi_v,\xi_v,\xi_v)\ln t |W(v)|^2}.$$

We will refer to the function W as the asymptotic profile of the solution u, which is now asymptotically described as

$$u(t,x) \approx \frac{1}{\sqrt{t}} W(v) e^{iq(\xi_v,\xi_v,\xi_v)\ln t |W(v)|^2} e^{it\phi(x/t)}.$$

Then it is natural to consider the relation between the initial data u_0 and the asymptotic profile W, via bounds for the difference

scat-error

$$e(t,x) = u(t,x) - \frac{1}{\sqrt{t}} W(v) e^{iq(\xi_v,\xi_v,\xi_v)\ln t |W(v)|^2} e^{it\phi(x/t)}. \tag{4.31}$$

In order to avoid any discussion of the asymptotic behavior of a at infinity, here we choose some compact frequency interval I so that the the symbol q of the nonlinearity is supported in I^3, and assume that the initial data u_0 is frequency localized in I. Then the associated range of velocities is $J = a'(I)$.

t:scat-model

Theorem 4. *a) For each initial data u_0 satisfying the smallness condition* (4.1) *and which is frequency localized in I, there exists an asymptotic profile $W \in H^{1-C\epsilon}$, supported in $J = a'(I)$ and with the property that*

$$\|W\|_{H_v^{1-C\epsilon}} \lesssim \epsilon, \tag{4.32}$$

for which the above difference satisfies the L^2 bounds

$$\|e\|_{L_x^2} \lesssim \epsilon t^{-\frac{1}{2}+C\epsilon}, \tag{4.33}$$

as well as the L^∞ bounds

$$\|e\|_{L_x^\infty} \lesssim \epsilon t^{-\frac{3}{4}+C\epsilon}. \tag{4.34}$$

Furthermore, the map $u_0 \to W$ is injective.

b) For each W supported in J and satisfying

$$\|W\|_{H_v^{1+C\epsilon}} \lesssim \epsilon, \tag{4.35}$$

there exists an associated initial data u_0 satisfying the smallness condition (4.1) *and frequency localized in I so that W is the asymptotic profile of u_0.*

Often one refers to the first property as the scattering property (modified scattering) and the second as the existence of wave operators (modified wave operators in our context). Together, they are called the *asymptotic completeness* property.

We also remark on the slight imperfection in the above result, connected with the $\pm C\epsilon$ terms in the Sobolev indices. These are largely unavoidable due to the $\log t$ terms in the phase, though one might possibly replace small powers with logs.

Proof. The argument here repeats the one in [8], and is omitted. □

5. Global solutions for small localized data: the general case

s:general

Here we consider several possible extensions of our main result, where we drop the compact support assumption on the symbol of the nonlinearity Q. Then we can no longer work with frequency localized data, so instead we will have to assume a suitable Sobolev type regularity at infinity. Precisely, we will define the space X by

def:X
$$\|u\|_X^2 = \|\Lambda_0(D)u\|_{L^2} + \|\Lambda_1(D)Lu\|_{L^2} \tag{5.1}$$

with suitable multiplier weights Λ_0 and Λ_1. The question we ask is

Question 2. *Given the symbols a, q and the above space X, under what assumptions does a small initial data in X guarantee global solutions and modified scattering for the equation* (1.1) *?*

Here there are three high frequency properties that play a role, namely the behaviors of a, of q and of Λ_0, Λ_1, all of which will be assumed to be of symbol type. These need to be considered both at frequencies close to $+\infty$ and at $-\infty$, and the two regions are largely independent. For convenience only we will not differentiate between the two. We begin our discussion with several remarks, which will play a role both in terms of the model we consider (i.e. the choice of a and q) and the regularity level for the result (i.e. the choice of Λ_0 and Λ_1):

(i) The behavior of a and a'': The convexity (concavity) of a is associated to dispersion, and plays a critical role. To simplify the notations we will assume a is convex, $a'' > 0$, and also we will assume some polynomial behavior for a'' at infinity,

choose-a
$$a''(\xi) \approx \langle \xi \rangle^{\sigma}, \qquad |\xi| \to \infty, \qquad \sigma \in \mathbb{R}, \tag{5.2}$$

with symbol type bounds for higher derivatives,

choose-a-reg
$$|\partial^j a''(\xi)| \lesssim \langle \xi \rangle^{\sigma - j}, \qquad j \geq 2. \tag{5.3}$$

Here we distinguish two different scenarios:

- The *generalized Klein-Gordon* case, $\sigma < -1$, where a has linear behavior at infinity and the linear problem has finite speed of propagation in the high frequency limit. Here we could further distinguish the range $\sigma \in [-2, -1)$ where a does not have a linear asymptote. The exact Klein-Gordon problem corresponds to $\sigma = -3$.
- The *generalized NLS* case $\sigma \geq -1$, where a is superlinear at infinity and we have infinite speed of propagation. The NLS equation in particular corresponds to $\sigma = 0$, while mKdV type behavior is associated to $\sigma = 1$.

(ii) The NLS smallness condition: In the regime of balanced frequency interactions, our problem is well approximated by a cubic NLS problem. There solitons can occur in the focusing case, but not small solitons. To avoid such a scenario, a smallness condition is required. A straightforward scaling computation yields the relation

no-soliton
$$\Lambda_0(\xi)\Lambda_1(\xi) \gtrsim \frac{q(\xi,\xi,\xi)}{a''(\xi)}, \tag{5.4}$$

as necessary in order for scattering to hold.

(iii) The relative size of Λ_0, Λ_1: It is natural to expect the function space X in our result to be stable with respect to dyadic frequency localizations. Commuting x with localizations leads to the requirement

localization

$$\Lambda_1(\xi) \lesssim \langle \xi \rangle \Lambda_0(\xi). \tag{5.5}$$

The two norms in (5.1) will be close in scaling in the high frequency limit when we are close to equality in this relation.

(iv) The normalization of q: Here we observe that our problem admits the invariance

$$\Lambda_0 \to b\Lambda_0, \qquad \Lambda_1 \to b\Lambda_1, \qquad q(\xi_1,\xi_2,\xi_3) \to b^{-1}(\xi_1-\xi_2+\xi_3)b(\xi_1)b(\xi_2)b(\xi_3)q(\xi_1,\xi_2,\xi_3)$$

obtained via the substitution $u = B(D)v$. Because of this, we can normalize q at least in the region of balanced frequency interactions,

q-symbol

$$|q(\xi_1,\xi_2,\xi_3)| \lesssim 1 \quad \text{when} \qquad |\xi_1| \approx |\xi_2| \approx |\xi_3| \approx |\xi_1-\xi_2+\xi_3|. \tag{5.6}$$

(v) Semilinear vs. quasilinear: While the size of q in the balanced region contributes to resonant interactions, a large size in the imbalanced region may provide a quasilinear term, for which just looking at the size is not enough to even guarantee local well-posedness. In this article we will simply avoid this issue, and simply assume that q is bounded everywhere, with symbol type regularity separately in each component.

Based on the discussion above, for the results in this section we will consider the following set-up for the symbols a and q:

(a) The symbol a is smooth, convex, with a'' as in (5.2), and symbol type regularity.

(b) The symbol q is smooth, real on the diagonal, and has the form

$$q(\xi_1,\xi_2,\xi_3) = \mathbf{q}(\xi_1,\xi_2,\xi_3,\xi_1+\xi_3-\xi_2), \tag{5.7}$$

(i.e. the trace of $\mathbf{q}$ on the diagonal $\xi_1-\xi_2+\xi_3-\xi_4=0$), where $\mathbf{q}$ is bounded and with separate symbol type regularity in all variables,

$$|\partial_{\xi_1}^{\alpha_1}\partial_{\xi_2}^{\alpha_2}\partial_{\xi_3}^{\alpha_3}\partial_{\xi_4}^{\alpha_4}\mathbf{q}(\xi_1,\xi_2,\xi_3,\xi_4)| \lesssim \prod_{j=1}^{4}\langle \xi_j\rangle^{-\alpha_j}. \tag{5.8}$$

Here $\xi_1-\xi_2+\xi_3$ appears naturally as the output frequency in the trilinear interaction.

Now we turn our attention to the regularity required by our result, which is determined by the symbols Λ_0 and Λ_1. This will be chosen to be

choose-Lambda

$$\Lambda_0 = \langle \xi \rangle^{s_0}, \qquad \Lambda_1 = \langle \xi \rangle^{s_1}, \tag{5.9}$$

so that

$$\|u\|_X^2 = \|u\|_{H^{s_0}}^2 + \|Lu\|_{H^{s_1}}^2.$$

It remains to discuss the choice of s_0 and s_1, which we would like to have as low as possible. So far, the heuristics above indicate that the following two conditions, arising from (5.4) and (5.5), are required:

choose-s0

$$s_0 + s_1 \geq -\sigma, \qquad s_1 \leq s_0 + 1. \tag{5.10}$$

Within this range, we note the best case scenario

hoose-s-ideal

$$s_0 = -\frac{\sigma+1}{2}, \qquad s_1 = -\frac{\sigma-1}{2}. \tag{5.11}$$

Indeed, this would correspond to a scale invariant result in the pure power case in the high frequency limit[3]. We retain these values as an ideal but unreachable goal, and seek to at least get close to these values. In particular, it is helpful to allow for at least a small positive margin in the first inequality in (5.10), in order to be able to allow for the small power type growth in (4.2). Even with this proviso, we will only be able to get close to the ideal setting in (5.11) only for $\sigma = -3$ (i.e. exact Klein-Gordon) and for the restricted range $-1 \leq \sigma \leq 1$ (i.e. weak NLS).

[3]Here, if $\sigma < -1$, then we can normalize in a Galilean fashion to set $a'(\infty) = 0$ before scaling.

The conditions in (5.10) above are required by the behavior of balanced interactions. However, managing imbalanced frequency interactions imposes further restrictions, which will be reflected in the choices below. To summarize, we will distinguish several cases, where δ stands for a small positive constant:

(I) **Weak Klein-Gordon**, $\sigma < -3$. Then we set

$$s_0 = -\sigma - 2 \qquad s_1 = -\sigma - 1.$$

(II) **Intermediate Klein-Gordon**, $-3 \leq \sigma < -2$. Then we set

$$s_0 = 1 + \delta, \qquad s_1 = -\sigma - 1$$

(III) **Strong Klein-Gordon** $-2 \leq \sigma < -1$. Then we set

$$s_0 = -\sigma - 1 + \delta, \qquad s_1 = 1.$$

(IV) **Weak NLS**, $-1 \leq \sigma \leq 1$. Then we set

$$s_0 = -\frac{\sigma + 1}{2} + \delta, \qquad s_1 = -\frac{\sigma - 1}{2}.$$

(V) **Strong NLS** (or KdV+), $1 < \sigma$. Then we set

$$s_0 = -1, \qquad s_1 = 0.$$

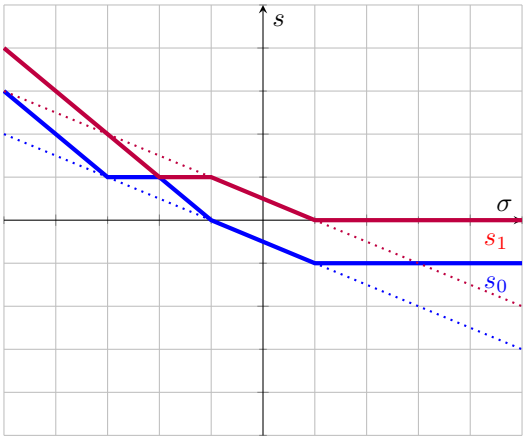

FIGURE 2. The Sobolev exponents s_0 and s_1 as a function of σ. Dotted lines indicate the best case scenario in (5.11).

Under these assumptions, we have

t:no-comp **Theorem 5.** *Assume that the symbols $a, q, \Lambda_0, \Lambda_1$ are as above, and that the initial data for our equation* (1.1) *satisfies:*

small+ (5.12) $$\|u(0)\|_X \lesssim \epsilon \ll 1.$$

Then the solution exists globally in time, with energy bounds

energy+ (5.13) $$\|u(t)\|_X \lesssim \epsilon t^{C\epsilon^2},$$

and pointwise decay

disp-decay+ (5.14) $$\|\langle D\rangle^{\frac{\delta}{4}} u(t)\|_{L^\infty} \lesssim \frac{\epsilon}{\sqrt{t}}.$$

Here δ is a small positive parameter, which depends on the choice of s_0 and s_1 above, and which can be taken to be exactly the one in the choice of s_0 in cases (II)-(III)-(IV) above.

Remark 5.1. *Our choice of exponents (s_0, s_1) in the cases (I)-(V) above was guided by the goal of coming as close as possible to the end-points of the two necessary conditions in* (5.10)*, giving priority to the first one. To minimize technicalities we have fixed the choice of some exponents rather than giving a range. We note however that increasing s_0 while keeping s_1 fixed is straightforward. This is also connected with the fact that we have simply assumed that the symbol for the cubic nonlinearity is bounded, rather than imposing various polynomial bounds. Many of the restrictions arising in the proofs are of technical nature rather than fundamental, and arise only in the study of unbalanced interactions, which is secondary to our main purpose. We leave it to the reader to explore other variants of these results, as needed.*

One may also supplement Theorem 5 with a matching result on modified scattering and asymptotic completeness, which exactly mirrors the result provided in Theorem 4 in the model case.

In order to best capture the behavior of the asymptotic profile W at high frequencies, it is best to parametrize W by ξ_v rather than by v. To account for this change, we define the asymptotic solution u_{asympt} associated to a profile W as

$$u_{asympt}(t,x) = \begin{cases} \frac{1}{\sqrt{t}} W(\xi_v) e^{iq(\xi_v,\xi_v,\xi_v)\ln t |W(\xi_v)|^2} e^{it\phi(v)} & \text{for } v \in a'(\mathbb{R}) \\ 0 & \text{for } v \notin a'(\mathbb{R}), \end{cases} \qquad x = vt,$$

where the second alternative occurs only in the generalized Klein-Gordon case (I)-(II)-(III).

Then we consider the relation between the initial data u_0 and the asymptotic profile W, via bounds for the difference

scat-error+ (5.15) $$e(t,x) = u(t,x) - u_{asympt}(t,x).$$

Theorem 6. *a) For each initial data u_0 satisfying the smallness condition* (5.12)*, there exists an asymptotic profile $W \in H^{1-C\epsilon}_{loc}(\mathbb{R})$ with the property that*

which-W (5.16) $$\|\langle\xi\rangle^{s_0+\frac{\sigma}{2}-C\epsilon^2} W\|_{L^2_\xi} + \|\langle\xi\rangle^{s_1+\frac{\sigma}{2}-C\epsilon^2} W\|_{H^{1-C\epsilon}_\xi} \lesssim \epsilon$$

for which the above difference satisfies the L^2 bounds

(5.17) $$\|e\|_{L^2_x} \lesssim \epsilon t^{-\delta_1}, \qquad \delta_1 > 0$$

as well as the L^∞ bounds

(5.18) $$\|e\|_{L^\infty} \lesssim \epsilon t^{-\frac{1}{2}-\delta_2}, \qquad \delta_2 > 0.$$

Furthermore, the map $u_0 \to W$ is injective.

b) For each W satisfying

which-W-in (5.19) $$\|\langle\xi\rangle^{s_0+\frac{\sigma}{2}+C\epsilon^2} W\|_{L^2} + \|\langle\xi\rangle^{s_1+\frac{\sigma}{2}-C\epsilon} W\|_{H^{1+C\epsilon}} \lesssim \epsilon$$

there exists an associated initial data u_0 satisfying the smallness condition (4.1) *so that W is the asymptotic profile of u_0.*

Just as in the case of Theorem 4, this result is also provided without proof. The proof follows again the same outline as in [8]. The exponents in (5.16), respectively (5.19) closely bracket the corresponding exponents in Lemma 5.9.

To avoid technicalities due to the many cases that would need to be considered, we do not attempt to specify exactly the positive constants δ_1 and δ_2 (which are independent of ϵ).

We remark that the choice of the exponents s_0 and s_1, and more precisely the second bounds in (5.10) guarantee that (5.16) satisfies the pointwise bound

$$|W(\xi)| \lesssim \epsilon \langle\xi\rangle^{-\frac{s_0+s_1}{2}+\sigma-C\epsilon^2}.$$

From here, the first bound in (5.10), if strict, guarantees that

$$\lim_{\xi\to\pm\infty} W(\xi) = 0.$$

This is particularly interesting in the generalized Klein-Gordon case $\sigma < 1$, where it implies that the asymptotic solution decays to zero at the edge of its support.

The proof of Theorem 5 follows the same outline as the proof of Theorem 3, using a bootstrap argument. The bootstrap assumption will be

p-decay+boot

$$\|\langle D\rangle^{\frac{\delta}{8}} u(t)\|_{L^\infty} \lesssim \frac{C\epsilon}{\sqrt{t}}. \tag{5.20}$$

Using the bootstrap assumption, we first prove the energy bound (5.13) with C replaced by C^2. By vector field bounds, the energy estimates will imply a pointwise estimate of the form

-decay+first

$$\|\langle D\rangle^{\frac{\delta}{2}} u(t)\|_{L^\infty} \lesssim \frac{\epsilon}{\sqrt{t}} t^{C^2\epsilon^2}, \tag{5.21}$$

which would give the bound (5.14) with an additional $t^{C^2\epsilon^2}$ loss, but also with a high frequency gain. To rectify that, we use our wave packet method to define a suitable asymptotic profile γ, which is then shown to be an approximate solution for the asymptotic equation. This will allow us to obtain pointwise bounds for the asymptotic profile without the loss, which are the transferred back to u. In the rest of the section, we successively discuss each of the steps of the proof, following the template of the model problem.

s:dyadic

5.1. **Dyadic decompositions.** Here we motivate and describe the dyadic decompositions that will be used in the sequel. In particular, these will turn out to depend on the ranges for γ.

1. The frequency decomposition. Here instead of the classical base 2 dyadic decomposition we will use narrower ranges,

$$\lambda = (1+\mu)^m, \qquad m \in \mathbb{N}, \qquad 0 < \mu \ll 1,$$

with the understanding that at frequencies $\lesssim 1$ we simply split into intervals of size μ. Here μ is a small universal parameter. The motivation for this choice is to allow for a clean classification of cubic interactions into balanced and unbalanced simply depending on the relative values of m.

We denote the corresponding frequency regions by $I_\lambda^\pm$. Here the $\pm$ signs stand for positive and negative frequencies, and will be at times omitted if they are not useful. We will also use an adapted partition of unity, again using the $\pm$ superscripts where needed.

$$1 = \sum_\lambda \nu_\lambda(\xi).$$

2. The velocity decomposition. At a given time t, we partition the spatial real axis corresponding to velocities associated to frequencies in I_λ. Precisely, we denote by $J_\lambda^\pm = a'(I_\lambda^\pm)$ the velocity ranges associated to frequencies in I_λ, and by $\tilde{J}_\lambda^\pm$ the corresponding spatial intervals, $\tilde{J}_\lambda^\pm = tJ_\lambda^\pm$. We can compute the size of these regions depending on the parameter γ,

$$|J_\lambda| \approx \lambda a''(\lambda), \qquad |\tilde{J}_\lambda| \approx t\lambda a''(\lambda),$$

where we simply denote $a''(\lambda) \approx |\lambda|^\sigma$. Within each interval J_λ, respectively $\tilde{J}_\lambda$ we will choose reference points v_λ, respectively x_λ.

Depending on the value of σ, we distinguish several scenarios:

a) The generalized NLS case, $\sigma \geq -1$. Here $\tilde{J}_\lambda$ are increasing in size with λ, and cover the entire real line (except for the degenerate case $\sigma = -1$ where they have equal size). In this case we have an associated spatial partition of unity

$$1 = \sum_\lambda \chi_\lambda^\pm(x), \qquad \operatorname{supp} \chi_\lambda^\pm \subset 2\tilde{J}_\lambda^\pm.$$

b) The strong Klein-Gordon case, $-2 \leq \sigma < -1$. Here $\tilde{J}_\lambda$ are decreasing in size with λ, but their sizes $t\lambda a''(\lambda)$ are large enough to dominate the associated uncertainty principle scale λ^{-1} as $\lambda \to \infty$. On the other hand, they do not cover the entire real line, only the range $\tilde{J}_{in} = (ta'(-\infty), ta'(+\infty))$. Thus we consider the partition of unity

$$1 = \sum_\lambda \chi_\lambda^\pm(x) + \chi_{out}(x),$$

where χ_{out} is the characteristic function of the outer region $\mathbb{R} \setminus J_{out}$.

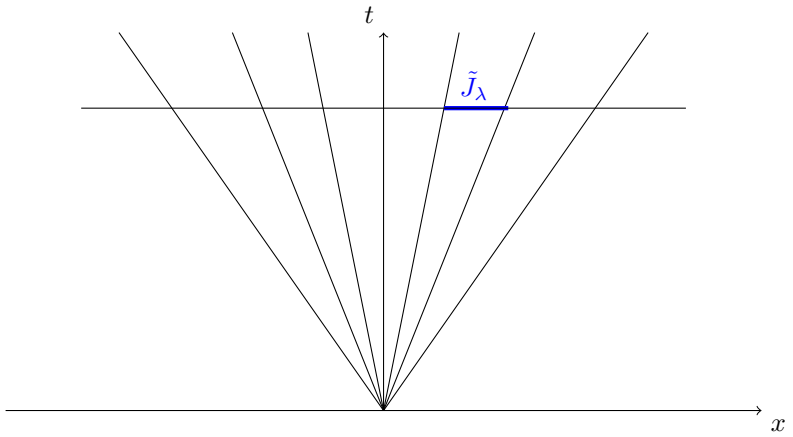

FIGURE 3. The velocity decomposition in Case (a), $\sigma \geq -1$: all group velocities are allowed.

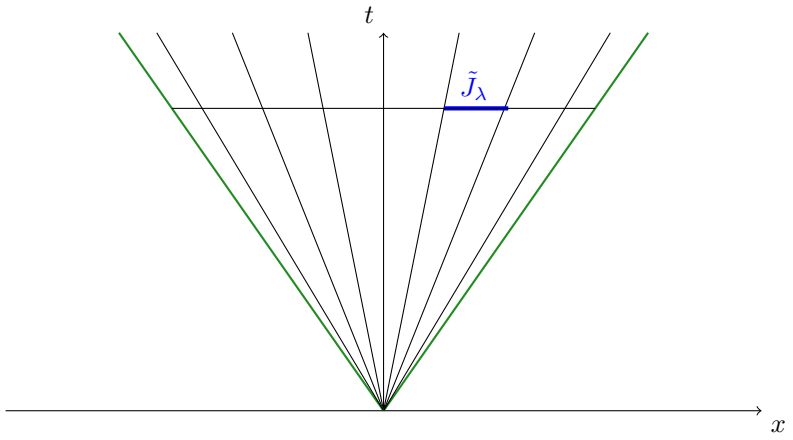

FIGURE 4. The velocity decomposition in Case (b), $-2 \leq \sigma < -1$: all dispersive waves are localized in an angle.

c) The Klein-Gordon case $\sigma < -2$. Here $\tilde{J}_\lambda$ are also decreasing in size with λ, but their sizes $t\lambda a''(\lambda)$ no longer dominate the associated uncertainty principle scale λ^{-1} as $\lambda \to \infty$. For this reason, based on this comparison we define the time dependent threshold λ_0 by

lambda0 (5.22)
$$t\lambda_0^2 a''(\lambda_0) = 1$$

and, depending on λ_0, we separate into low and high frequencies, and consider the partition of unity

$$1 = \sum_{\lambda > \lambda_0} \chi_\lambda^{\pm}(x) + \chi_{hi}(x) + \chi_{out}(x),$$

where χ_{hi} selects a region of size λ_0^{-1}. Here the intuition is that up to frequency λ_0 we see dispersive effects at time t, whereas above that we are simply solving a transport equation at leading order.

3. The decomposition of Q. For the trilinear form Q, it will be very useful to split it into a balanced and an unbalanced component,

$$Q(u, \bar{u}, u) = Q^{bal}(u, \bar{u}, u) + Q^{unbal}(u, \bar{u}, u),$$

depending on the size of the three interacting frequencies. Precisely, at the symbol level we set

$$\mathbf{q}^{bal}(\xi_1, \xi_2, \xi_3, \xi_4) = \chi_{\langle \xi_1 \rangle \approx \langle \xi_2 \rangle \approx \langle \xi_3 \rangle \approx \langle \xi_4 \rangle} \mathbf{q}(\xi_1, \xi_2, \xi_3, \xi_4).$$

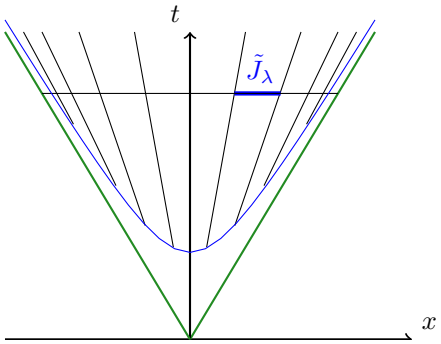

FIGURE 5. The velocity decomposition in Case (c), $\sigma < -2$: the dispersive region is above the blue curve.

Here the balanced part will play the leading role, and is the one responsible for the modified scattering behavior. The unbalanced, part, on the other hand, we will want to treat largely in a perturbative manner. However, some technical difficulties will have to be dealt with along the way.

From the perspective of the spatial Littlewood-Paley decomposition defined earlier, we will essentially think of the two components as combinations of dyadic frequency localizations. Precisely, given dyadic frequencies $\lambda_j = (1+c)^{m_j}$, we will call the quadruplet $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ diagonal if $\max\{|m_i - m_j|\} \leq 4$. We denote the diagonal set of frequencies by $\mathfrak{D}$. Then we will simply set

$$Q^{bal}(u, \bar{u}, u) = \sum_{\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathfrak{D}} P_{\lambda_4} Q(u_1, \bar{u}_2, u_3),$$

respectively

$$Q^{unbal}(u, \bar{u}, u) = \sum_{\lambda_1, \lambda_2, \lambda_3, \lambda_4 \notin \mathfrak{D}} P_{\lambda_4} Q(u_1, \bar{u}_2, u_3),$$

where for brevity we have denoted $u_i := P_{\lambda_j} u$. We remark that the P_{λ_4} projection can be omitted in the case when λ_4 is comparable to the highest frequency; this includes in particular the balanced case.

As a trilinear form applied to u, the symbol of the expression

$$P_{\lambda_4} Q(u_1, \bar{u}_2, u_3)$$

has the form

$$q_{\lambda_1, \lambda_2, \lambda_3, \lambda_4}(\xi_1, \xi_2, \xi_3) = \nu_{\lambda_1}(\xi_1) \nu_{\lambda_2}(\xi_2) \nu_{\lambda_3}(\xi_3) \nu_{\lambda_4}(\xi_4) \mathbf{q}(\xi_1, \xi_2, \xi_3, \xi_4), \qquad \xi_1 + \xi_3 = \xi_2 + \xi_4,$$

and can be thought of as the diagonal trace of a bump function on the rectangle $I_{\lambda_1} \times I_{\lambda_2} \times I_{\lambda_3} \times I_{\lambda_4}$. Using separation of variables on this product region, we can expand these localized symbols as rapidly convergent series

$$q_{\lambda_1, \lambda_2, \lambda_3, \lambda_4}(\xi_1, \xi_2, \xi_3) = \sum_{k=1}^{\infty} \nu_1^k(\xi_1) \nu_2^k(\xi_2) \nu_3^k(\xi_3) \nu_4^k(\xi_4) \qquad \xi_1 + \xi_3 = \xi_2 + \xi_4,$$

where the factors have decaying sizes

$$|\partial^l \nu_j^k| \lesssim k^{-N} \lambda_j^{-l}, \qquad l \leq N,$$

for a large N.

Since the dyadic multipliers ν_j^k are bounded in X, this will allow us to replace $Q_{\lambda_1, \lambda_2, \lambda_3, \lambda_4}$ in all X bounds with product type operators, precisely of the form

$$Q_{\lambda_1, \lambda_2, \lambda_3, \lambda_4}(u, \bar{u}, u) \approx P_{\lambda_4}(u_1 \bar{u}_2 u_3).$$

Furthermore, if $\lambda_1, \lambda_2, \lambda_3 \lesssim \lambda_4$ then we can further eliminate the outer projection P_{λ_4}. We will refer to this reduction, later in the paper, as *separation of variables*.

5.2. **The vector field bound.** Our primary goal here is to discuss the counterpart of the vector field estimate in Proposition 2.1. We will do this in a frequency localized setting, and also consider the better elliptic bounds outside the corresponding dyadic velocity range. Precisely, we have the following linear estimates:

p:vf-gen **Proposition 5.2.** *a) Let $\delta = s_0 + s_1 + \sigma > 0$. Then we have the uniform bound*

-point-lambda
$$\|\langle D\rangle^{\frac{\delta}{2}-} u\|_{L^\infty} \lesssim \|u\|_X. \tag{5.23}$$

b) We also have the dyadic elliptic bounds for a function u_λ localized at frequency λ, and $x_\lambda \in \tilde{J}_\lambda$:

vf-ell-ll2
$$\|(1-\chi_\lambda)(x-x_\lambda)u_\lambda\|_{L^2} \lesssim \lambda^{-s_1}\|u_\lambda\|_X, \tag{5.24}$$

respectively

vf-ell-llinf
$$|(1-\chi_\lambda)u_\lambda(x)| \lesssim \frac{\lambda^{-s_1+\frac{1}{2}}}{|x-x_\lambda|}\|u_\lambda\|_X. \tag{5.25}$$

Proof. Using a dyadic decomposition in frequency as described earlier in Section 5.1,

$$u = \sum_\lambda u_\lambda,$$

we first observe that we can localize the X bound and conclude that

$$\|u_\lambda\|_X \lesssim \|u\|_X.$$

This is where the condition $s_1 \le s_0 + 1$ is used.

The advantage is that for each λ, the size of a'' is essentially constant, and we may harmlessly extend a to have uniform convexity outside I_λ. Hence we will be able to apply directly the results in Propositions 2.1, 2.4, with the choice of parameters

$$R = a''(\lambda) \approx \lambda^\sigma, \qquad M \approx \lambda^{-1}.$$

a) Since we have $s_0 + s_1 \ge -\sigma + \delta$ as well as $s_1 \le s_0 + 1$, a direct application of Proposition 2.1 yields

$$\|u_\lambda\|_{L^\infty}^2 \lesssim \frac{1}{ta''(\lambda)}(\lambda^{-s_0-s_1} + \lambda^{-2s_0-1})\|u_\lambda\|_X^2 \lesssim \frac{1}{t\lambda^\delta}\|u_\lambda\|_X^2,$$

which immediately yields the bound (5.23).

We further remark that, in the context of the classification of cases in the previous subsection, in case (c), which is the Klein-Gordon case, it is also interesting to distinguish the low frequencies from the high frequencies, and replace the full dyadic decomposition of u by

$$u = \sum_{|\lambda|<\lambda_0} u_\lambda + u_{hi},$$

where the threshold λ_0 is as in (5.22). While the above argument applies in all cases, for high frequencies the desired bound also follows directly from Bernstein's inequality, completely neglecting the Lu_λ bound,

$$\|u_{hi}\|_{L^\infty} \lesssim \sum_{\lambda>\lambda_0} \lambda^{1+\delta}(a''(\lambda))^{\frac{1}{2}}\|u\|_{H^{s_0}} \lesssim \lambda_0^{-\delta} t^{-\frac{1}{2}}\|u\|_X.$$

This is consistent with the fact that in this regime our evolution is at leading order a transport equation, with negligible dispersion. Precisely, in this frequency range we can perturbatively replace the symbol a with its affine asymptotes as ξ approaches $\pm\infty$.

b) Here we similarly apply Proposition 2.4. The bound (5.24) is obtained directly from (2.20). For (5.25) we first use a multiplier $\tilde{P}_\lambda$ with slightly larger support to decompose

$$(1-\chi_\lambda)(x-x_\lambda)u_\lambda = \tilde{P}_\lambda(1-\chi_\lambda)(x-x_\lambda)u_\lambda + (1-\tilde{P}_\lambda)(1-\chi_\lambda)(x-x_\lambda)u_\lambda$$

The first term is localized at frequency λ, so we can estimate it using Bernstein's inequality and (5.24),

$$\|\tilde{P}_\lambda(1-\chi_\lambda)(x-x_\lambda)u_\lambda\|_{L^\infty} \lesssim \lambda^{\frac{1}{2}}\|(1-\chi_\lambda)(x-x_\lambda)u_\lambda\|_{L^2} \lesssim \lambda^{-s_1+\frac{1}{2}}\|u_\lambda\|_X.$$

In the second term the coefficient $(1-\chi_\lambda)(x-x_\lambda)$ must be localized at frequency at least λ,

$$(1-\tilde P_\lambda)(1-\chi_\lambda)(x-x_\lambda)u_\lambda = (1-\tilde P_\lambda) g_{\gtrsim\lambda} u_\lambda, \qquad g = (1-\chi_\lambda)(x-x_\lambda).$$

Then we estimate

$$\begin{aligned}
\|(1-\tilde P_\lambda)(1-\chi_\lambda)(x-x_\lambda)u_\lambda\|_{L^\infty} &\lesssim \|g_{\gtrsim\lambda}u_\lambda\|_{L^\infty}\\
&\lesssim \lambda^{\frac12}\|g_{\gtrsim\lambda}u_\lambda\|_{L^2} + \lambda^{-\frac12}\|\partial_x[g_{\gtrsim\lambda}u_\lambda]\|_{L^2}\\
&\lesssim (\lambda^{\frac12}\|g_{\gtrsim\lambda}\|_{L^\infty} + \lambda^{-\frac12}\|\partial_x g_{\gtrsim\lambda}\|_{L^\infty})\|u_\lambda\|_{L^2}\\
&\lesssim \lambda^{-s_0-\frac12}\|\partial_x g\|_{L^\infty}\|u_\lambda\|_X
\end{aligned}$$

which suffices since $\|\partial_x g\|_{L^\infty} \lesssim 1$ and $s_0 \geq s_1 - 1$.

5.3. **Bounds for Q and the energy estimate for u.** Here the first goal is to prove the following energy bound for the function u:

p:ee-u **Proposition 5.3.** *Assume that u is a solution to* (1.1)*, under the same assumptions as in Theorem 5. Then we have the bound*

$$\frac{d}{dt}\|u(t)\|_{H^{s_0}}^2 \lesssim \|\langle D\rangle^{\frac{\delta}{8}}u\|_{L^\infty}^2\|u\|_X^2. \tag{5.26}$$

We note that in many problems this bound is independent on the X norm, and has instead the form

$$\frac{d}{dt}\|u(t)\|_{H^{s_0}}^2 \lesssim \|\langle D\rangle^{\frac{\delta}{8}}u\|_{L^\infty}^2\|u\|_{H^{s_0}}^2. \tag{5.27}$$

This is the case if $s_0 \geq 0$ (see the proof below) but also if Q has additional structure.

Proof. Differentiating in time and using the equation (1.1), this reduces to the weighted inequality

q1
$$\|Q(u,\bar u,u)\|_{H^{s_0}} \lesssim \|\langle D\rangle^{\frac{\delta}{8}}u\|_{L^\infty}^2\|u\|_X. \tag{5.28}$$

Here we distinguish two cases depending on the sign of s_0:

i) $s_0 \geq 0$. Here we have the simpler bound

q1-easy
$$\|Q(u,\bar u,u)\|_{H^{s_0}} \lesssim \|\langle D\rangle^{\frac{\delta}{8}}u\|_{L^\infty}^2\|u\|_{H^{s_0}}, \tag{5.29}$$

which does not involve any control for Lu. Since Q satisfies the symbol bounds (5.6), this easily follows by a standard Littlewood-Paley decomposition with respect to all inputs and the output. The H^{s_0} factor on the right is always chosen to correspond to the highest frequency, and the δ exponent readily ensures dyadic summation. More precisely, writing

$$Q(u,\bar u,u) = \sum_{\lambda_1,\lambda_2,\lambda_3,\lambda_4} P_{\lambda_4}Q(u_{\lambda_1},\bar u_{\lambda_2},u_{\lambda_3}),$$

and relabeling increasing order $\{\lambda_1,\lambda_2,\lambda_3\} = \{\lambda_{lo},\lambda_{mid},\lambda_{hi}\}$ we must have either

(a) $\lambda_4 \approx \lambda_{hi}$, or
(b) $\lambda_4 < \lambda_{mid} \approx \lambda_{hi}$,

and correspondingly decompose $Q = Q_a + Q_b$.

For Q_a we may use orthogonality to estimate

$$\|Q_a(u,\bar u,u)\|_{H^{s_0}}^2 \lesssim \sum_{\lambda_{hi}}\left(\sum_{\lambda_{lo},\lambda_{mid}}\|u_{\lambda_{lo}}\|_{L^\infty}\|u_{\lambda_{mid}}\|_{L^\infty}\right)^2\|u_{\lambda_{hi}}\|_{H^{s_0}}^2,$$

where the inner sum is estimated by the L^∞ norm in (5.29).

For Q_b on the other hand we neglect orthogonality and estimate directly

$$\|Q_b(u,\bar u,u)\|_{H^{s_0}}^2 \lesssim \sum_{\lambda_{lo}}\sum_{\lambda_{mid}\approx\lambda_{hi}}\|u_{\lambda_{lo}}\|_{L^\infty}\|u_{\lambda_{mid}}\|_{L^\infty}\|u_{\lambda_{hi}}\|_{H^{s_0}},$$

where the summation with respect to the two indices is again guaranteed by the L^∞ norm in (5.29).

ii) $s_0 < 0$, which is needed only in the generalized NLS case $\sigma \geq -1$. In this case, the bound (5.29) applies only to the portion of Q where at least one of the three input frequencies, which we denote by λ_1, λ_2 and λ_3, is at most comparable to the output frequency λ_4.

Hence, from here on we assume that $\lambda_4 \ll \lambda_j$, $j = 1, 2, 3$. This guarantees that λ_1, λ_2 and λ_3 should all be distinct, and also the largest two should be comparable. Under these assumptions, it remains to prove the estimate

$$\begin{aligned} \lambda_4^{s_0} \|P_{\lambda_4} Q(u_{\lambda_1}, \bar{u}_{\lambda_2}, u_{\lambda_3})\|_{L^2} \lesssim \|u_{\lambda_1}\|_{L^\infty} \|u_{\lambda_2}\|_{L^\infty} \|u_{\lambda_3}\|_X + \|u_{\lambda_2}\|_{L^\infty} \|u_{\lambda_3}\|_{L^\infty} \|u_{\lambda_1}\|_X \\ + \|u_{\lambda_3}\|_{L^\infty} \|u_{\lambda_1}\|_{L^\infty} \|u_{\lambda_2}\|_X. \end{aligned} \tag{5.30}$$

Here we note that, since the two highest frequencies are comparable, the dyadic summation with respect to the four frequencies is straightforward using the δ factor, and (5.28) follows.

To prove the last bound, we retain the restrictions on λ_1, λ_2 and λ_3, but then harmlessly drop the projection P_{λ_4}. Then we can use separation of variables and reduce the problem to the product case, where it suffices to show that

$$\begin{aligned} \|u_{\lambda_1} \bar{u}_{\lambda_2} u_{\lambda_3}\|_{L^2} \lesssim \|u_{\lambda_1}\|_{L^\infty} \|u_{\lambda_2}\|_{L^\infty} \|u_{\lambda_3}\|_X + \|u_{\lambda_2}\|_{L^\infty} \|u_{\lambda_3}\|_{L^\infty} \|u_{\lambda_1}\|_X \\ + \|u_{\lambda_3}\|_{L^\infty} \|u_{\lambda_1}\|_{L^\infty} \|u_{\lambda_2}\|_X. \end{aligned} \tag{5.31}$$

Next, we separate the product with respect to dyadic velocity ranges. Since the λ's cannot be all equal, it suffices to estimate the expression

$$I = \|(1 - \chi_{\lambda_1}) u_{\lambda_1} \bar{u}_{\lambda_2} u_{\lambda_3}\|_{L^2}.$$

By Proposition 5.2 we have

$$\begin{aligned} I &\lesssim \|(1 - \chi_{\lambda_1}) u_{\lambda_1}\|_{L^2} \|u_{\lambda_2}\|_{L^\infty} \|u_{\lambda_3}\|_{L^\infty} \\ &\lesssim \sup_{\lambda \neq \lambda_1} \frac{\lambda_1^{-s_1}}{t|a'(\lambda_1) - a'(\lambda)|} \|u_{\lambda_1}\|_X \|u_{\lambda_2}\|_{L^\infty} \|u_{\lambda_3}\|_{L^\infty} \\ &\lesssim \frac{\lambda_1^{-s_1}}{t\lambda_1 |a''(\lambda_1)|} \|u_{\lambda_1}\|_X \|u_{\lambda_2}\|_{L^\infty} \|u_{\lambda_3}\|_{L^\infty}, \end{aligned}$$

where the λ dependent weight is maximized when λ is near λ_1. Then it suffices to check that

$$\frac{\lambda_1^{-s_1}}{t\lambda_1 |a''(\lambda_1)|} \lesssim 1.$$

Given the choice of s_1 and that $\sigma \geq -1$, this is true with a substantial gain. Thus the proof of the Proposition is complete. □

A second objective here is to show that, in the context of the balanced/unbalanced decomposition for the cubic nolinearity Q, we have a better bound for the unbalanced part. This bound will play a role in our wave packet testing in the next subsection, precisely in the estimate for the error in the asymptotic equation.

Proposition 5.4. *The unbalanced part Q^{unbal} of Q satisfies the better L^∞ bound*

Q-L2-from-Lh

$$\|\chi_\lambda P_\lambda Q^{unbal}(u, \bar{u}, u)\|_{L^\infty} \lesssim \frac{\lambda^{-\frac{\delta}{4}}}{t^{\frac{3}{2}+\frac{\delta}{4}}} \|u\|_X^3, \tag{5.32}$$

provided that either $\sigma \geq -2$ or $\{\sigma < -2$ and $t\lambda^{\sigma+2} \geq 1\}$.

We remark that, depending on σ and on the balance of the three frequencies, in some of the cases one can get a better asymptotic equation error bound by using L^2 estimates for Q^{unbal}. We do not pursue this here because it is not needed.

Remark 5.5. *This bound is needed in order to be able to control the contribution of Q^{unbal} to the error in the wave packet testing. Precisely, we will need to be able to verify that*

$$\langle Q^{unbal}(u, \bar{u}, u), \mathbf{u}_v \rangle \lesssim t^{-1-\delta}$$

for $v \in J_\lambda$, and $\lambda < \lambda_0$ in the case $\sigma < -2$. This requires the L^∞ bound

$$\|\chi_\lambda P_\lambda Q\|_{L^\infty} \lesssim t^{-\frac{3}{2}-}.$$

Proof. We first simply consider a triple product $u_1\bar{u}_2u_3$ where λ_1, λ_2 and λ_3 are not all equal, and estimate it within a dyadic velocity region A_λ. For that we apply (5.24) and (5.23) for a λ_j, say λ_3, which is away from λ. This yields

Q-2 (5.33) $$\|\chi_\lambda u_1\bar{u}_2u_3\|_{L^2} \lesssim \frac{\lambda_1^{-\frac{\delta}{2}}\lambda_2^{-\frac{\delta}{2}}\lambda_3^{-s_1}}{t^2|a'(\lambda_3)-a'(\lambda)|}\|u\|_X^3,$$

respectively

Q-inf (5.34) $$\|\chi_\lambda u_1\bar{u}_2u_3\|_{L^\infty} \lesssim \frac{\lambda_1^{-\frac{\delta}{2}}\lambda_2^{-\frac{\delta}{2}}\lambda_3^{-s_1+\frac{1}{2}}}{t^2|a'(\lambda_3)-a'(\lambda)|}\|u\|_X^3.$$

We complement these with the trivial bound

Q-inf+ (5.35) $$\|\chi_\lambda u_1\bar{u}_2u_3\|_{L^\infty} \lesssim \frac{\lambda_1^{-\frac{\delta}{2}}\lambda_2^{-\frac{\delta}{2}}\lambda_3^{-s_0+\frac{1}{2}}}{t}\|u\|_X^3.$$

To use these estimates we consider two scenarios:

(i) $\lambda_1 = \lambda$ and $\lambda_3 < \lambda_2 \ll \lambda$. Then we can separate variables to discard P_λ, and apply the above bounds (5.34) and (5.35). Now we examine the coefficient in (5.34) as a function of λ_3. For $\sigma \geq -1$ we get

$$\|\chi_\lambda u_1\bar{u}_2u_3\|_{L^\infty} \lesssim \lambda_1^{-\frac{\delta}{2}}\lambda_2^{-\frac{\delta}{2}}\lambda_3^{-s_1+\frac{1}{2}}\lambda^{-\sigma-1}t^{-2}\|u\|_X^3,$$

which suffices. For $\sigma < -1$ we get

$$\|\chi_\lambda u_1\bar{u}_2u_3\|_{L^\infty} \lesssim \lambda_1^{-\frac{\delta}{2}}\lambda_2^{-\frac{\delta}{2}}\lambda_3^{-s_1-\frac{1}{2}-\sigma}t^{-2}\|u\|_X^3.$$

This still suffices directly in the range $-\frac{3}{2} \leq \sigma \leq -1$, and after interpolation with (5.35) in the remaining range $\sigma < -\frac{3}{2}$. In all cases the summation in λ_3 and λ_2 is straightforward.

(ii) In the remaining case we must have at least two comparable high frequencies, say $\lambda_2, \lambda_3 \gtrsim \lambda$, one of which, say λ_3, is separated from λ. Then we replace the cutoff χ_λ by one with a double support, call it $\tilde{\chi}_\lambda$, which equals one on a comparably sized neighbourhood of the support of χ_λ. Precisely, we write

$$\chi_\lambda P_\lambda = \chi_\lambda P_\lambda \tilde{\chi}_\lambda + \chi_\lambda P_\lambda(1-\tilde{\chi}_\lambda).$$

The second term is easily taken care of by noting that

$$\|\chi_\lambda P_\lambda(1-\tilde{\chi}_\lambda)\|_{L^\infty\to L^\infty} \lesssim \frac{1}{(ta''(\lambda)\lambda^2)^N}$$

combined with the pointwise bound for each of the factors.

For the first term we apply (5.33), noting that the coefficient is nonincreasing in $\lambda_3 \gtrsim \lambda$. For $\sigma \geq -1$ we obtain

$$\|\tilde{\chi}_\lambda u_1\bar{u}_2u_3\|_{L^2} \lesssim \frac{\lambda_3^{-s_1}}{t^2\lambda_3^{1+\sigma}}\|u\|_X^3,$$

and conclude using Bernstein's inequality at frequency λ. For $\sigma < -1$ we obtain

$$\|\tilde{\chi}_\lambda u_1\bar{u}_2u_3\|_{L^2} \lesssim \frac{\lambda_3^{-s_1}}{t^2\lambda^{1+\sigma}}\|u\|_X^3,$$

Then we use Bernstein's inequality at frequency λ and interpolate with (5.35) as in case (i).

□

5.4. **The energy estimate for Lu.** Here the objective is to prove the energy estimate for Lu. As in the model case, this will be achieved via a cubic correction C so that we can obtain a favorable estimate for the nonlinear expression

$$L^{NL}u = Lu + tC(u,\bar{u},u).$$

Precisely, we will prove the following

p:ee-Lu **Proposition 5.6.** *There exists a trilinear, translation invariant correction C with the following properties*
(i) Uniform bound for C:

L2-C
$$\|C(u,\bar u,u)\|_{H^{s_1}} \lesssim \|\langle D\rangle^{\frac{\delta}{8}} u\|_{L^\infty}^2 \|u\|_{H^{s_0}}. \tag{5.36}$$

(ii) Energy bound for $L^{NL}u$,

L2-LNL
$$\frac{d}{dt}\|L^{NL}u\|_{H^{s_1}}^2 \lesssim \|u\|_X^2 \|\langle D\rangle^{\frac{\delta}{8}} u\|_{L^\infty}^2 + t^{-\frac12-\delta}\|u\|_X^3 \|\langle D\rangle^{\frac{\delta}{8}} u\|_{L^\infty}. \tag{5.37}$$

One immediate consequence of (5.36) combined with the bootstrap assumption (5.20) is the norm equivalence

same-X
$$\|u\|_X^2 \approx \|u\|_{H^{s_0}}^2 + \|L^{NL}u\|_{H^{s_1}}^2. \tag{5.38}$$

Using this property one easily sees that, combining the energy estimates for u and $L^{NL}u$ in Propositions 5.3, 5.6, and using the bootstrap assumption (5.20), we obtain by Gronwall's inequality the energy estimate in (5.13).

Proof. For the expression $w := L^{NL}u$ we have an equation of the form

$$(i\partial_t - A)w = LQ(u,\bar u,u) + tR_3(u,\bar u,u) + iC(u,\bar u,u) + tR_5(u,\bar u,u,\bar u,u),$$

where R_3 has symbol

$$r_3(\xi_1,\xi_2,\xi_3) = c(\xi_1,\xi_2,\xi_3)(a(\xi_1) - a(\xi_2) + a(\xi_3) - a(\xi_1-\xi_2+\xi_3)),$$

and R_5 is simply the quintilinear form arising from the time derivative of C.

The objective is then to choose the correction C so that (5.36) holds, and we can estimate the source terms in H^{s_1},

c2
$$\|LQ(u,\bar u,u) + tR_3(u,\bar u,u)\|_{H^{s_1}} \lesssim \|u\|_X \|\langle D\rangle^{\frac{\delta}{8}} u\|_{L^\infty}^2, \tag{5.39}$$

respectively

c3
$$\|R_5(u,\bar u,u,\bar u,u)\|_{H^{s_1}} \lesssim \|u\|_{H^{s_0}} \|\langle D\rangle^{\frac{\delta}{8}} u\|_{L^\infty}^4. \tag{5.40}$$

Here naively one may hope to use the same correction C as in the compact case, so that we have

$$LQ(u,\bar u,u) + tR_3(u,\bar u,u) = Q(Lu,\bar u,u) - Q(u,\overline{Lu},u) + Q(u,\bar u,Lu).$$

However, as it turns out, there are some difficulties which such a direct approach. Precisely, considering a full dyadic decomposition for Q, there are two interesting scenarios to consider:

a) Balanced interactions, where the three input frequencies and the output frequency are all comparable, say to a fixed frequency λ. Then the symbol c has similar support, symbol type regularity and size

$$|c(\xi_1,\xi_2,\xi_3)| \lesssim \lambda^{-1}|q(\xi_1,\xi_2,\xi_3)| \lesssim |\lambda|^{-1}.$$

In this case the bounds (5.36) and (5.39) are straightforward, nothing but a rescaled version of the corresponding bounds in the compact case. We still need to prove (5.40), which contains some unbalanced interactions, but this is not so difficult.

b) Unbalanced interactions, where, instead, the use of the correction C would cause trouble:

- The expression of C would be more complicated, which causes difficulties with (5.36) and (5.40).
- the bound (5.39) is unbalanced, which causes difficulties unless $s_1 = 0$ or we have a favourable frequency balance.

However, the redeeming feature in this case is that, in each dyadic velocity range, at least one of the three inputs must correspond to a different range of velocities, so the corresponding frequency localized operator L is elliptic there. It follows that the expression $LQ(u,\bar u,u)$ no longer needs to be corrected, and instead should be estimated directly, in an elliptic fashion.

To implement the heuristic strategy described above, we decompose Q into a balanced and an unbalanced component,

$$Q(u,\bar u,u) = Q^{bal}(u,\bar u,u) + Q^{unbal}(u,\bar u,u),$$

where at the symbol level we set

$$\mathbf{q}^{bal}(\xi_1,\xi_2,\xi_3,\xi_4) = \chi_{\langle\xi_1\rangle\approx\langle\xi_2\rangle\approx\langle\xi_3\rangle\approx\langle\xi_4\rangle}\mathbf{q}(\xi_1,\xi_2,\xi_3,\xi_4).$$

Then we choose the normal form correction $C = C^{bal}$ to account for the balanced term, where the corresponding errors are estimated as discussed above. On the other hand, the unbalanced term we simply treat perturbatively, without any correction.

A. The balanced term. To account for the balanced term, we follow the compact case and set

$$c^{bal}(\xi_1,\xi_2,\xi_3) = q^{bal}(\xi_1,\xi_2,\xi_3)\frac{a_\xi(\xi_1)-a_\xi(\xi_2)+a_\xi(\xi_3)-a_\xi(\xi)}{a(\xi_1)-a(\xi_2)+a(\xi_3)-a(\xi)}, \tag{5.41}$$

so that we have the algebraic relation

$$LQ^{bal}(u,\bar u,u) - tR_3^{bal}(u,\bar u,u) = Q^{bal}(Lu,\bar u,u) - C^{bal}(u,\overline{Lu},u) + C^{bal}(u,\bar u,Lu) + D(u,\bar u,u),$$

with

$$d(\xi_1,\xi_2,\xi_3) = i(\partial_{\xi_1}-\partial_{\xi_2}+\partial_{\xi_3})q^{bal}(\xi_1,\xi_2,\xi_3).$$

Then we have

Lemma 5.7. *The above correction C^{bal} satisfies the estimates* (5.36), (5.39) *and* (5.40).

Proof. As mentioned earlier, the proof of (5.36), (5.39) is simply a rescaled version of the similar argument in Section 4.2. As such, it is omitted and left as an exercise for the reader.

The bound (5.40), on the other hand, involves also some unbalanced interactions and deserves some separate attention. Localizing in frequency and separating variables, we split

$$C^{bal} = \sum_\lambda C^{bal}_\lambda,$$

where we can assume that the frequency λ portion C^{bal}_λ of C^{bal} has the form

$$C^{bal}_\lambda(u,\bar u,\bar u) = \lambda^{-1}u_\lambda\bar u_\lambda u_\lambda.$$

Then the corresponding component of R_5 has terms of the form

$$R_{5,\lambda}(u,\bar u,u,\bar u,u) = \lambda^{-1}u_\lambda\bar u_\lambda P_\lambda Q(u,\bar u,u).$$

Hence, we can bound it by

$$\|R_{5,\lambda}(u,\bar u,u,\bar u,u)\|_{H^{s_1}} \lesssim \lambda^{s_1-1}\|u_\lambda\|_{L^2}\|u_\lambda\|_{L^\infty}\|Q(u,u,u)\|_{L^\infty} \lesssim \|u_\lambda\|_{H^{s_0}}\|\langle D\rangle^{\frac{\delta}{8}}u\|^4_{L^\infty}.$$

□

B. The unbalanced term. This corresponds to the unbalanced component Q^{unbal} of Q. Here we set our correction to 0, so that R_3 and R_5 also vanish. Then it remains to prove that we have the following result:

l:LQunbal **Lemma 5.8.** *Assume that s_0, s_1 are chosen as in Theorem 5. Then for $\delta > 0$ we have the following L^2 type bound:*

c2+
$$\|LQ^{unbal}(u,\bar u,u)\|_{H^{s_1}} \lesssim \|u\|_X\|\langle D\rangle^{\frac{\delta}{8}}u\|^2_{L^\infty} + t^{-\frac12-\delta}\|u\|^2_X\|\langle D\rangle^{\frac{\delta}{8}}u\|_{L^\infty}. \tag{5.42}$$

Here the advantage is that we can choose which inner u we place the L on. Using the bound (5.42) in the lemma, the proof of the L^2 energy bound (5.37) for $L^{NL}u$ is concluded.

Proof. We localize the cubic expression Q in frequency to dyadic regions associated with input frequencies λ_1, λ_2, λ_3 and output frequency λ_4. Since the choice of the small parameter δ is flexible, the dyadic summation with respect to λ_1, λ_2, λ_3 and λ_4 is straightforward. For this reason, it suffices to prove the lemma in the case when λ_1, λ_2, λ_3 and λ_4 are fixed. To streamline notations, we will denote $u_j := P_{\lambda_j}u$ for $j = 1,2,3$.

In each such region, the nonlinear expression Q is essentially like a product, which then gets localized to the output frequency λ_4. Using separation of variables, we can reduce the problem to the case

$$Q^{unbal}(u_1,\bar u_2,u_3) = P_{\lambda_4}(u_1\bar u_2u_3),$$

where $\lambda_1, \lambda_2, \lambda_3$ are not all equal. This key property implies that, in a given a dyadic velocity range associated to a frequency λ, we must have at least one of the three L's act as an elliptic operator; our estimate is primarily based on this principle. We will further separate the problem into two cases, depending on the relative size of the three interacting frequencies λ_1, λ_2 and λ_3.

a) The *llh* case, where

$$\lambda_1 \leq \lambda_2 \ll \lambda_3,$$

or the symmetric case where λ_1 and λ_3 are interchanged. In this case we must have $\lambda_4 \approx \lambda_3$, and we can also use separation of variables to discard the P_{λ_4} projector. Further, it will be convenient to commute L inside, and write

$$L(u_1 \bar{u}_2 u_3) = u_1 \bar{u}_2 L u_3 + tR(u_1, \bar{u}_2, u_3),$$

where the symbol of R is

$$r(\xi_1, \xi_2, \xi_3) = a'(\xi_1 - \xi_2 + \xi_3) - a'(\xi_3).$$

This is a smooth symbol in all three variables on the corresponding dyadic scales, and has size

$$|r(\xi_1, \xi_2, \xi_3)| \lesssim \lambda_2 a''(\lambda_3).$$

The first term is estimated in a straightforward fashion,

$$\|u_1 \bar{u}_2 L u_3\|_{H^{s_1}} \lesssim \|u_1\|_{L^\infty} \|u_2\|_{L^\infty} \|L u_3\|_{H^{s_1}}.$$

For the second term we can use separation of variables to drop again the multipliers, and we are left with the task of estimating the expression

$$t\lambda_3^{s_1} \lambda_2 a''(\lambda_3) \|u_1 \bar{u}_2 u_3\|_{L^2}.$$

Finally, we use a spatial angular localization to separate into directions associated to a dyadic frequency λ. In this case we can consider a very simple separation, depending on whether the direction λ matches λ_3 or not, writing

$$u_1 \bar{u}_2 u_3 = \chi_{\lambda_3}(u_1 \bar{u}_2 u_3) + (1 - \chi_{\lambda_3})(u_1 \bar{u}_2 u_3).$$

For the first term we use (5.24) for u_2 in order to write

$$\begin{aligned} t\lambda_3^{s_1} \lambda_2 a''(\lambda_3) \|\chi_{\lambda_3}(u_1 \bar{u}_2 u_3)\|_{L^2} &\lesssim t\lambda_3^{s_1} \lambda_2 a''(\lambda_3) \|\chi_{\lambda_3} u_2\|_{L^2} \|u_1\|_{L^\infty} \|u_3\|_{L^\infty} \\ &\lesssim t\lambda_3^{s_1} \lambda_2 a''(\lambda_3) \frac{1}{\lambda_2^{s_1} t |a'(\lambda_2) - a'(\lambda_3)|} \|u_2\|_X \|u_1\|_{L^\infty} \|u_3\|_{L^\infty}. \end{aligned}$$

Here the coefficient on the right is nondecreasing in λ_2 in all cases (this corresponds to the restriction $s_1 \leq 1$ if $\sigma \geq -1$, respectively $s_1 \leq -\sigma$ if $\sigma < -1$, which are satisfied for our choice of exponents) and equals 1 if $\lambda_2 = \lambda_3$.

For the second term we instead use (5.24) for u_2 in order to write

$$\begin{aligned} t\lambda_3^{s_1} \lambda_2 a''(\lambda_3) \|(1 - \chi_{\lambda_3})(u_1 \bar{u}_2 u_3)\|_{L^2} &\lesssim t\lambda_3^{s_1} \lambda_2 a''(\lambda_3) \|u_1\|_{L^\infty} \|u_3\|_{L^\infty} \|(1 - \chi_{\lambda_3}) u_2\|_{L^2} \\ &\lesssim t\lambda_3^{s_1} \lambda_2 a''(\lambda_3) \frac{1}{\lambda_3^{s_1} t \lambda_3 |a''(\lambda_3)|} \|u_2\|_X \|u_1\|_{L^\infty} \|u_3\|_{L^\infty} \\ &\lesssim \|u_2\|_X \|u_1\|_{L^\infty} \|u_3\|_{L^\infty}. \end{aligned}$$

This concludes the proof of (5.42) in this case.

b) The *lhh* case,

$$\lambda_1 < \lambda_2 \approx \lambda_3, \qquad \lambda_4 \lesssim \lambda_2.$$

or permutations thereof. Here we have many subcases to consider. We first reduce their number by peeling off some of the easier ones.

A first argument we can apply is to simply write

LP3 (5.43) $$L P_{\lambda_4}(u_1 \bar{u}_2 u_3) = P_{\lambda_4}(L u_1 \bar{u}_2 u_3) + [x, P_{\lambda_4}](u_1 \bar{u}_2 u_3) + tR(u_1, \bar{u}_2, u_3),$$

where the commutator term is essentially of the form $\lambda_4^{-1} P_{\lambda_4}(u_1 u_2 u_3)$ and the remainder R arises from switching the argument of L, and has symbol

$$r(\xi_1, \xi_2, \xi_3) = a'(\xi_1) - a'(\xi_1 - \xi_2 + \xi_3).$$

This is a smooth symbol on the four associated dyadic scales, and of size

$$|r(\xi_1,\xi_2,\xi_3)| \lesssim |a'(\lambda_1) - a'(\lambda_4)|.$$

Here we can estimate the first term in (5.43) in H^{s_1} by

$$\|P_{\lambda_4}(Lu_1\bar{u}_2u_3)\|_{H^{s_1}} \lesssim \|Lu_1\|_{H^{s_1}}\|u_2\|_{L^\infty}\|u_3\|_{L^\infty},$$

provided that

$$\text{either} \quad \lambda_4 \lesssim \lambda_1 \quad \text{or} \quad \sigma \geq 1, \tag{5.44}$$

where the second condition ensures that $s_1 = 0$.

The second term in (5.43) is estimated by

$$\lambda_4^{-1}\|P_{\lambda_4}(u_1\bar{u}_2u_3)\|_{H^{s_1}} \lesssim \|u_1\bar{u}_2u_3\|_{H^{s_0}},$$

after which we can reuse the bounds in Proposition 5.3.

Finally, for the last term we separate variables, and it remains to estimate

$$\lambda_4^{s_1} t|a'(\lambda_1) - a'(\lambda_4)|\|u_1\bar{u}_2u_3\|_{L^2}.$$

We split the triple product with respect to angles,

$$u_1\bar{u}_2u_3 = \chi_{\lambda_3}u_1\bar{u}_2u_3 + (1-\chi_{\lambda_3})u_1\bar{u}_2u_3,$$

and estimate the two terms separately. For the first one we use (5.24) for u_1,

$$\begin{aligned}\lambda_4^{s_1} t|a'(\lambda_1) - a'(\lambda_4)|\|\chi_{\lambda_3}u_1\bar{u}_2u_3\|_{L^2} &\lesssim \lambda_4^{s_1} t|a'(\lambda_1) - a'(\lambda_4)|\|\chi_{\lambda_3}u_1\|_{L^2}\|u_2\|_{L^\infty}\|u_3\|_{L^\infty}\\ &\lesssim \lambda_4^{s_1} t|a'(\lambda_1) - a'(\lambda_4)|\frac{1}{\lambda_1^{s_1}t|a'(\lambda_1)-a'(\lambda_3)|}\|u_1\|_X\|u_2\|_{L^\infty}\|u_3\|_{L^\infty}.\end{aligned}$$

If $\sigma \geq -1$ then the coefficient equals

$$\frac{\lambda_4^{s_1}|a'(\lambda_1) - a'(\lambda_4)|}{\lambda_1^{s_1}\lambda_3^{\sigma+1}} \lesssim \frac{\lambda_4^{s_1}(\lambda_1+\lambda_4)^{\sigma+1}}{\lambda_1^{s_1}\lambda_3^{\sigma+1}} \leq 1.$$

Else, $\lambda_4 \lesssim \lambda_1$ therefore the coefficient equals

$$\frac{\lambda_4^{s_1}\lambda_4^{\sigma+1}}{\lambda_1^{s_1}\lambda_1^{\sigma+1}} \leq 1.$$

For the second one we use (5.24) for u_1,

$$\begin{aligned}\lambda_4^{s_1} t|a'(\lambda_1) - a'(\lambda_4)|\|(1-\chi_{\lambda_3})u_1\bar{u}_2u_3\|_{L^2} &\lesssim \lambda_4^{s_1} t|a'(\lambda_1) - a'(\lambda_4)|\|u_1\|_{L^\infty}\|u_2\|_{L^\infty}\|(1-\chi_{\lambda_3})u_3\|_{L^2}\\ &\lesssim \lambda_4^{s_1} t|a'(\lambda_1) - a'(\lambda_4)|\frac{1}{\lambda_3^{s_1}t\lambda_3 a''(\lambda_3)|}\|u_1\|_X\|u_2\|_{L^\infty}\|u_3\|_{L^\infty},\end{aligned}$$

and the coefficient is again easily verified to be ≤ 1 by considering the same two cases as above.

After this reduction, it remains to consider the case when

$$\lambda_1 \ll \lambda_4 \lesssim \lambda_2 \approx \lambda_3, \qquad \sigma < 1. \tag{5.45}$$

Here we separate the case $\sigma < -2$, where the threshold λ_0 plays a role. Precisely, if $\lambda_3 > \lambda_0$ then we can use (5.43) where r has size $\lambda_4^{\sigma+1}$ in order to write schematically

$$LP_{\lambda_4}(u_1\bar{u}_2u_3) = P_{\lambda_4}(u_1\bar{u}_2Lu_3) + \lambda_4^{-1}P_{\lambda_4}(u_1\bar{u}_2u_3) + t\lambda_4^{\sigma+1}P_{\lambda_4}(u_1\bar{u}_2u_3).$$

The first two terms are easy to estimate directly. So it remains to consider the third, where we estimate

$$\begin{aligned}t\lambda_4^{\sigma+1}\|P_{\lambda_4}(u_1\bar{u}_2u_3)\|_{H^{s_1}} &\lesssim \lambda_4^{s_1+\sigma+1}\|u_1\|_{L^\infty}\|u_2\|_{L^\infty}\|u_3\|_{L^2}\\ &\lesssim t\lambda_4^{s_1+\sigma+1}\lambda_3^{-s_0}\|u_1\|_{L^\infty}\|u_2\|_{L^\infty}\|u_3\|_X\\ &\lesssim t\lambda_3^{\sigma+2}\|u_1\|_{L^\infty}\|u_2\|_{L^\infty}\|u_3\|_X,\end{aligned}$$

where $t\lambda_3^{\sigma+2} \leq t\lambda_0^{\sigma+2} = 1$.

From here on, we will assume that $\lambda_3 < \lambda_0$ in the case $\sigma < -2$. Since $\lambda_1 \ll \lambda_4$, we can harmlessly move the P_{λ_4} projection to the product $\bar{u}_2 u_3$, and work with

$$v := u_1 P_{\lambda_4}(\bar{u}_2 u_3).$$

To simplify matters, we note that within the λ_4 frequency region we have

$$l(x,\xi) = x - x_4 + tO(\lambda_4^{\sigma+1}),$$

and similarly at the operator level we get

L-last
$$\|Lv\|_{H^{s_1}} \lesssim \lambda_4^{s_1}(\|(x-x_4)u_1 P_{\lambda_4}(\bar{u}_2 u_3)\|_{L^2} + t\lambda_4^{\sigma+1}\|u_1 P_{\lambda_4}(\bar{u}_2 u_3)\|_{L^2}). \tag{5.46}$$

We will rely on this bound for $\sigma \geq -1$. However, for $\sigma < -1$ we can process the first term further. We first move $x - x_4$ inside P_{λ_4} at the expense of a mild commutator term, which is schematically written as

$$(x-x_4)u_1 P_{\lambda_4}(\bar{u}_2 u_3) = u_1 P_{\lambda_4}((x-x_4)\bar{u}_2 u_3) + \lambda_4^{-1} u_1 P_{\lambda_4}(\bar{u}_2 u_3).$$

The contribution of the L^2 norm of the commutator term can be harmlessly included into the second RHS term in (5.46), using the upper bound $\lambda_4 \lesssim \lambda_0$ if $\sigma < -2$. On the other hand for the main term we can write

$$(x-x_4)u_3 = Lu_3 + tRu_3, \qquad |r| \approx \lambda_4^{\sigma+1}.$$

Since $\sigma < -1$, the contribution of the error term R can also be included into the second RHS term in (5.46). Finally, for the Lu_3 term we estimate directly

L-help
$$\lambda_4^{s_1}\|u_1 P_{\lambda_4} u_2 L u_3\|_{L^2} \lesssim \left(\frac{\lambda_4}{\lambda_3}\right)^{s_1} \|u_1\|_{L^\infty}\|u_2\|_{L^\infty}\|u_3\|_X, \tag{5.47}$$

which is an acceptable contribution. We arrive at the following simplification of (5.46),

L-last+
$$\|Lv\|_{L-last} \lesssim RHS(5.47) + t\lambda_4^{s_1+\sigma+1}\|u_1 P_{\lambda_4}(\bar{u}_2 u_3)\|_{L^2}, \qquad \sigma < -1. \tag{5.48}$$

At this point we would like to consider angular localizations for the triple product $v = u_1 P_{\lambda_4}(\bar{u}_2 u_3)$, centered on the angle associated to λ_3. This angular region has size $t\lambda_3 a''(\lambda_3) = t\lambda_3^{\sigma+1}$, whereas v has frequency λ_4. So, by the uncertainty principle, this localization is meaningful only if

t-uncert
$$t\lambda_3^{\sigma+1}\lambda_4 \gtrsim 1. \tag{5.49}$$

This constraint is nontrivial only if $\sigma < -1$. We dispense with the complementary range by estimating directly the second RHS term in (5.48) as follows:

direct
$$\begin{aligned} t\lambda_4^{s_1+\sigma+1}\|u_1 P_{\lambda_4}(\bar{u}_2 u_3)\|_{L^2} &\lesssim t\lambda_4^{s_1+\sigma+1}\lambda_3^{-s_0}\|u_1\|_{L^\infty}\|u_2\|_X\|u_3\|_{L^\infty} \\ &= (t\lambda_3^{\sigma+1}\lambda_4)\lambda_4^{s_1+\sigma}\lambda_3^{-s_0-\sigma-1}\|u_1\|_{L^\infty}\|u_2\|_X\|u_3\|_{L^\infty}, \end{aligned} \tag{5.50}$$

where all the factors on the right are $\lesssim 1$ given our choice of s_0 and s_1. We assume (5.49) from here on.

We are now ready to localize v using the angular cutoff χ_{λ_3} associated to frequency λ_3 waves. It is easier to first consider the contribution of $(1-\chi_{\lambda_3})v$. One difficulty we encounter is that we need to commute this localization with P_{λ_4},

$$(1-\chi_{\lambda_3})P_{\lambda_4} = (1-\chi_{\lambda_3})P_{\lambda_4}(1-\tilde{\chi}_{\lambda_3}) + R,$$

where the error R has size

$$\|R\|_{L^2\to L^2} \lesssim \left(\frac{1}{t\lambda_4\lambda_3^{\sigma+1}}\right)^N.$$

Here, if $\sigma \geq -1$ then we get t^{-N} and the R bound becomes straightforward. Otherwise (5.49) holds so we can simply add the R bound to the computation in (5.50).

Hence we are left with the bound for the contribution of the expression

$$u_1 P_{\lambda_4}(\bar{u}_2(1-\chi_{\lambda_3})u_3).$$

to either (5.46) (for $\sigma \geq -1$) or (5.48) (for $\sigma < -1$). This is

$$I_1 = \lambda_4^{s_1}(\|(x-x_4)u_1 P_{\lambda_4}(\bar{u}_2(1-\chi_{\lambda_3})u_3)\|_{L^2} + t\lambda_4^{\sigma+1}\|u_1 P_{\lambda_4}(\bar{u}_2(1-\chi_{\lambda_3})u_3)\|_{L^2}).$$

Here we harmlessly commute $x - x_4$ inside P_{λ_4}, modulo a mild error term which is controlled by the second term on the right. Then we use Proposition 2.4 to estimate

$$\|(1-\chi_{\lambda_3})(x-x_3)u_3\|_{L^2} \lesssim \lambda_3^{-s_1}\|u_3\|_X.$$

Bounding the other two factors in L^∞, this yields

$$I_1 \lesssim \lambda_4^{s_1}\lambda_3^{-s_1} \sup_{x\notin A_3} \frac{|x-x_4| + t\lambda_4^{\sigma+1}}{|x-x_3|}\|u_3\|_X\|u_1\|_{L^\infty}\|u_2\|_{L^\infty}.$$

The supremum is attained when x is closest to x_3, i.e. when $|x-x_3| \approx t\lambda_3 a''(\lambda_3)$, in which case we get the coefficient

$$\lambda_4^{s_1}\lambda_3^{-s_1}\frac{|a'(\lambda_4)-a'(\lambda_3)|}{\lambda_3 a''(\lambda_3)}.$$

If $\sigma \geq -1$ this gives

$$\lambda_4^{s_1}\lambda_3^{-s_1} \leq 1.$$

If $\sigma < -1$ we get instead

$$\lambda_4^{s_1+\sigma+1}\lambda_3^{-s_1-\sigma-1} \leq 1,$$

both of which suffice.

Finally, we consider the most difficult case, where we estimate the contribution of $\chi_{\lambda_3}v = \chi_{\lambda_3}u_1 P_{\lambda_4}(\bar{u}_2 u_3)$, namely

$$\begin{aligned} I_2 &= \lambda_4^{s_1}(\|(x-x_4)\chi_{\lambda_3}u_1 P_{\lambda_4}(\bar{u}_2u_3)\|_{L^2} + t\lambda_4^{\sigma+1}\|\chi_{\lambda_3}u_1P_{\lambda_4}(\bar{u}_2u_3)v\|_{L^2}) \\ &\approx t\lambda_4^{s_1}|a'(\lambda_4)-a'(\lambda_3)|\|\chi_{\lambda_3}u_1P_{\lambda_4}(\bar{u}_2u_3)\|_{L^2}. \end{aligned}$$

Here we can apply the bound (5.25) for u_1 to get

u1-low (5.51) $$\|\chi_{\lambda_3}u_1\|_{L^\infty} \lesssim \frac{\lambda_1^{-s_1+\frac{1}{2}}}{t|a'(\lambda_1)-a'(\lambda_3)|}\|u_1\|_X.$$

On the other hand, for $\bar{u}_2u_3$ we compute

$$t\partial_x(\bar{u}_2u_3) = \overline{\tilde{L}u_2}u_3 + \bar{u}_2\tilde{L}u_3,$$

which allows us to estimate

$$\|P_{\lambda_4}(\bar{u}_2u_3)\|_{L^2} \lesssim \frac{\lambda_3^{-s_1}}{t\lambda_4 a''(\lambda_3)}(\|u_2\|_X\|u_3\|_{L^\infty} + \|u_2\|_{L^\infty}\|u_3\|_X).$$

Combining the last two bounds, we arrive at

$$I_2 \lesssim \lambda_4^{s_1-1}|a'(\lambda_4)-a'(\lambda_3)|\frac{\lambda_1^{-s_1+\frac{1}{2}}}{|a'(\lambda_1)-a'(\lambda_3)|}\frac{\lambda_3^{-s_1}}{ta''(\lambda_3)}\|u_1\|_X(\|u_2\|_X\|u_3\|_{L^\infty} + \|u_2\|_{L^\infty}\|u_3\|_X).$$

Now we examine the coefficient in front. If $\sigma \geq -1$ then we obtain

$$\lambda_4^{s_1-1}\lambda_1^{-s_1+\frac{1}{2}}t^{-1}\lambda_3^{-s_1-\sigma} \leq \lambda_4^{-\frac{1}{2}}t^{-1},$$

which is more than sufficient.

However, if $\sigma < -1$ then we get instead

$$I_2 \lesssim \lambda_4^{s_1+\sigma}\lambda_1^{-s_1-\sigma-\frac{1}{2}}t^{-1}\lambda_3^{-s_1-\sigma}\|u_1\|_X(\|u_2\|_X\|u_3\|_{L^\infty} + \|u_2\|_{L^\infty}\|u_3\|_X),$$

which is unsatisfactory since the power of the high frequency λ_3 is positive. To rectify this, we use again (5.51) but estimate u_3 directly in L^2 to get

$$\begin{aligned} I_2 &\lesssim \lambda_4^{s_1}|a'(\lambda_4)-a'(\lambda_3)|\frac{\lambda_1^{-s_1+\frac{1}{2}}}{|a'(\lambda_1)-a'(\lambda_3)|}\lambda_3^{-s_0}\|u_1\|_X\|u_2\|_{L^\infty}\|u_3\|_{H^{s_0}} \\ &\lesssim \lambda_4^{s_1+\sigma+1}\lambda_1^{-s_1-\sigma-\frac{1}{2}}\lambda_3^{-s_0}\|u_1\|_X\|u_2\|_{L^\infty}\|u_3\|_{H^{s_0}}. \end{aligned}$$

This has a negative power of λ_3 but insufficient time decay. Combining the two bounds we arrive at

$$I_2 \lesssim \left(\frac{\lambda_4}{\lambda_1}\right)^{s_1+\sigma+\frac{1}{2}} \min\left\{t^{-1}\lambda_4^{\frac{1}{2}}\lambda_3^{-s_0}, \lambda_4^{-\frac{1}{2}}\lambda_3^{-s_1-\sigma}\right\}\|u_1\|_X(\|u_2\|_X\|u_3\|_{L^\infty} + \|u_2\|_{L^\infty}\|u_3\|_X).$$

Here the first exponent is negative $s_1 + \sigma + \frac{1}{2} < 0$, and thus favourable. In the second factor, balancing exactly at the middle would yield the factor

$$t^{-\frac{1}{2}} \lambda_3^{\frac{-s_0-s_1-\sigma}{2}},$$

with a favourable negative power of λ_3 but a marginally insufficient power of t. But unbalancing this slightly suffices in order to improve the power of t while maintaining a negative power for λ_3.

This concludes the proof of Lemma 5.8.

□

The proof of Proposition 5.6 is now also concluded. □

5.5. **Wave packets and the asymptotic profile.** For each admissible velocity $v \in J = a'(\mathbb{R})$ we define the associated wave packet $\mathbf{u}_v$ using the same formula (4.20) as in the model case. Then the associated asymptotic profile $\gamma(t, v)$ can be defined exactly as before, following (4.24) but as a function

domain-gamma (5.52) $$\gamma : D = J \times \mathbb{R}^+ \to \mathbb{C}.$$

If we consider velocities in the dyadic range $v \in J_\lambda$ then the spatial localization scale for the associated wave packet $\mathbf{u}_v$ is

$$\delta x \approx t^{\frac{1}{2}} (a''(\lambda))^{\frac{1}{2}}.$$

It is instructive to compare this scale with the size of the spatial region $\tilde{J}_\lambda$ associated to frequency λ, which is

$$|\tilde{J}_\lambda| = t\lambda a''(\lambda).$$

It is meaningful to define our asymptotic profile only if this dominates the wave packet scale,

$$\delta x \lesssim |\tilde{J}_\lambda|.$$

This is equivalent to

$$t^{\frac{1}{2}} (a'')^{\frac{1}{2}} \lesssim t\lambda a'' \quad \Longleftrightarrow \quad t \gtrsim (\lambda^2 a''(\lambda))^{-1}.$$

This is nontrivial only in the Klein-Gordon case $\sigma < -2$, where it can be rewritten in the form $\lambda \lesssim \lambda_0$, which is the same threshold we have encountered before. Hence, from here on, in the case $\sigma < -2$ we will restrict γ to a smaller set. Precisely, in this case we will redefine D as

ain-gamma-low (5.53) $$D = \bigcup_\lambda \{(v,t) \subset J \times \mathbb{R}^+; v \in J_\lambda, \;\; t \gtrsim (\lambda^2 a''(\lambda))^{-1}\}.$$

The first step in our study of the asymptotic profile γ is to obtain bounds for it in terms of the X norm of u.

l:gamma **Lemma 5.9.** *Let $t \geq 1$, and $u \in X$ be a function at time t. Then within D we have the bounds*

gamma-L2 (5.54) $$\|\gamma\|_{L^2_v(J_\lambda)} \lesssim (\lambda^{-s_0} + (ta''(\lambda)\lambda^2)^{-N}) \|u\|_X,$$

gamma-inf (5.55) $$\|\gamma\|_{L^\infty(J_\lambda)} \lesssim (\lambda^{-\frac{\delta}{2}} + (ta''(\lambda)\lambda^2)^{-N}) \|u\|_X,$$

gamma-dv (5.56) $$\|\partial_v \gamma\|_{L^2_v(J_\lambda)} \lesssim (\lambda^{-s_1-\sigma} + (ta''(\lambda)\lambda^2)^{-N}) \|u\|_X.$$

We remark that the term $(ta''(\lambda)\lambda^2)^{-N})$ is only relevant in the case $\sigma \leq -2$. Precisely, if $\sigma = -2$ then it gives t^{-N}, and if $\sigma < -2$ then it gives $(\lambda/\lambda_0)^N$.

Proof. Here we use the fact that, for $v \in J_\lambda$, our wave packet $\mathbf{u}_v$ is essentially localized at frequency λ. Precisely, we can represent it as

$$\mathbf{u}_v = (a''(\lambda))^{-\frac{1}{2}} \chi(v, y) e^{i\xi_v x}, \qquad y = \frac{x - vt}{\sqrt{ta''(\lambda)}},$$

with χ Schwartz in y, uniformly in v. This allows us to obtain favourable bounds for the portion of $\mathbf{u}_v$ away from frequency λ,

$$|\partial_x^k P_{\neq\lambda} \mathbf{u}_v| \lesssim_{k,N} (a''(\lambda))^{-\frac{1}{2}} (1 + |y|)^{-N} \lambda^k (ta''(\lambda)\lambda^2)^{-N},$$

where one can distinguish three separate cases:

a) $\sigma > -2$, where we get an arbitrarily large gain,

$$|\partial_x^k P_{\neq\lambda}\mathbf{u}_v| \lesssim_{k,N} \lambda^{-N}(1+|y|)^{-N}t^{-N}.$$

b) $\sigma = -2$, where we only have the gain in time,

$$|\partial_x^k P_{\neq\lambda}\mathbf{u}_v| \lesssim_{k,N} (a''(\lambda))^{-\frac{1}{2}}(1+|y|)^{-N}t^{-N}.$$

c) $\sigma < -2$, where the gain depends on the distance to λ_0,

$$|\partial_x^k P_{\neq\lambda}\mathbf{u}_v| \lesssim_{k,N} (a''(\lambda))^{-\frac{1}{2}}(1+|y|)^{-N}\lambda^k(\lambda_0/\lambda)^{-N}.$$

We now use this in order to prove the three bounds in the Lemma.

Proof of (5.54)*:* We separate frequencies λ and frequencies away from λ,

$$\gamma(t,v) = \langle u_\lambda, \mathbf{u}_v\rangle + \langle u, P_{\neq\lambda}\mathbf{u}_v\rangle.$$

For the first inner product we use Young's inequality to get

$$\|\langle u_\lambda, \mathbf{u}_v\rangle\|_{L^2_v} \lesssim \|u_\lambda\|_{L^2_x},$$

where we lose a $t^{\frac{1}{2}}$ factor from the L^1_x norm of $\mathbf{u}_v$ but we regain it from the change of coordinates from x to x/t. For the second product, on the other hand, we take advantage of the rapid decay in the above bounds for $P_{\neq\lambda}\mathbf{u}_v$. In the nontrivial range $\sigma \leq -2$ we have $s_0 > 0$, so the worst contribution comes from frequencies $\lesssim 1$ in u.

Proof of (5.55)*:* Here we use instead the pointwise bound (5.23) for u. The main contribution comes from u_λ via Young's inequality, while the other frequencies only contribute a rapidly decaying tail, as above.

Proof of (5.56)*:* We have

$$\gamma_v(t,v) = \langle u, \partial_v\mathbf{u}_v\rangle.$$

For $\partial_v\mathbf{u}_v$ we use the representation in Lemma 4.5, to write the above expression as

$$\gamma_v(t,v) = \langle Lu, \mathbf{u}_v^2\rangle + \langle u, \mathbf{r}_v\rangle.$$

As above, the leading contribution comes from u_λ where we can use directly Young's inequality. □

Next we compare the asymptotic profile with the exact solution, working in the same region.

l:u-gamma **Lemma 5.10.** *Suppose $v \in J_\lambda$, with the additional restriction $\lambda < \lambda_0$ in the case when $\sigma < -2$. Then we have*

$$|\gamma(t,v) - t^{\frac{1}{2}}u_\lambda(t,vt)e^{-it\phi(v)}| \lesssim (t\lambda^2 a''(\lambda))^{-\frac{1}{4}}\|u\|_X. \tag{5.57}$$

Proof. Since $\lambda < \lambda_0$, the contribution of $u_{\neq\lambda}$ to γ has size $(t\lambda^2 a''(\lambda))^{-N}$ and may be neglected.

Next we consider the contribution of u_λ to γ, which generates the error

$$r(t,v) = \langle \mathbf{u}_v, u_\lambda\rangle - t^{\frac{1}{2}}u_\lambda(t,vt)e^{-it\phi(v)}.$$

Here the scales are fixed, so we can directly apply the argument in Section 4.4, Proposition 4.7 to get the error bound

$$|r(t,v)| \lesssim \lambda^{-s_1}t^{-\frac{1}{4}}a''(\lambda)^{-\frac{3}{4}}\|u\|_X \lesssim t^{-\frac{1}{4}}\lambda^{-\frac{1}{2}}(a'')^{-\frac{1}{4}}\|u\|_X,$$

which is exactly as needed. □

Finally, we show that γ is a good approximate solution for the asymptotic equation,

ma-asympt-re
$$\dot\gamma(t,v) = iq(\xi_v,\xi_v,\xi_v)t^{-1}\gamma(t,v)|\gamma(t,v)|^2 + f(t,v), \tag{5.58}$$

where f satisfies favourable bounds:

l:asympt-err **Lemma 5.11.** *Suppose $v \in J_\lambda$ and that, in addition, $\lambda < \lambda_0$ if $\sigma < -2$. Then the error f satisfies the uniform bound*

$$|f(t,v)| \lesssim (t^{-1}(t\lambda^2 a'')^{-\frac{1}{4}} + t^{-1-\frac{\delta}{4}}\lambda^{-\frac{\delta}{4}})t^{C^2\epsilon^2}, \qquad \sigma \leq -2, \quad t\lambda^2 a'' \geq 1, \tag{5.59}$$

respectively

$$|f(t,v)| \lesssim t^{-1-\frac{\delta}{4}}\lambda^{-\frac{\delta}{4}}t^{C^2\epsilon^2}, \qquad \sigma > -2. \tag{5.60}$$

Proof. We can write

$$f(t,v) = \langle u, (i\partial_t - A(D))\mathbf{u}_v\rangle + \langle Q(u,\bar{u},u), \mathbf{u}_v\rangle := f_1 + f_2,$$

and estimate each term separately.

For f_1 we use Lemma 4.4 to write

$$(i\partial_t - A(D))\mathbf{u}_v = t^{-\frac{3}{2}}(L\mathbf{u}_v^1 + \mathbf{r}_v^1),$$

where

$$\mathbf{u}_v^1 \approx (a'')^{-\frac{1}{2}}\mathbf{u}_v, \qquad \mathbf{r}_v \approx \lambda^{-1}(a'')^{-\frac{1}{2}}\mathbf{u}_v.$$

Hence we can use Holder's inequality to bound

$$|f_1(t,v)| \lesssim t^{-\frac{5}{4}}(a'')^{-\frac{3}{4}}\lambda^{-s_1}\|u\|_X \lesssim t^{-1}(ta''\lambda^2)^{-\frac{1}{4}}\|u\|_X.$$

Next we consider f_2, where we use the balanced/unbalanced decomposition of Q. The contribution of the unbalanced part Q^{unbal} is placed in f using the bound and (5.32).

It remains to consider the balanced component of Q. Furthermore, in view of the frequency localization of $\mathbf{u}_v$ at frequency λ, it suffices to consider the balanced component of Q localized to frequency λ.

Here we go through two stages, exactly as in the similar argument in the model case:

a) Replace u_λ by $t^{-\frac{1}{2}}\gamma(t,v)\chi_\lambda e^{it\phi}$, with errors controlled by Lemmas 5.10 and 5.9.

b) Replace $Q_\lambda(\chi_\lambda e^{it\phi}, \chi_\lambda e^{-it\phi}, \chi_\lambda e^{it\phi})$ by $t^{-\frac{3}{2}}\chi_\lambda^3 e^{it\phi}q(\xi_v,\xi_v,\xi_v)$.

□

5.6. **Conclusion.** Our remaining objective is to recover our bootstrap assumption, and show that we have the better bound

$$\|\langle D\rangle^{\frac{\delta}{4}}u\|_{L^\infty} \lesssim \epsilon. \tag{5.61}$$

We consider separately each dyadic component u_λ, for which we seek to show that

want (5.62) $$\|u_\lambda\|_{L^\infty} \lesssim \epsilon t^{-\frac{1}{2}}\lambda^{-\frac{\delta}{3}}.$$

On the other hand, from the vector field bound (5.23) and the energy estimates (5.13) we have

have (5.63) $$\|u_\lambda\|_{L^\infty} \lesssim \epsilon t^{-\frac{1}{2}+C^2\epsilon^2}\lambda^{-\frac{\delta}{2}}.$$

Here ϵ is sufficiently small, so in particular we can assume that $\epsilon \ll \delta$. Hence, the desired conclusion (5.63) follows provided that $t \lesssim \lambda^N$, where the large constant N can be chosen arbitrarily. It remains to consider the complementary region $t \gtrsim \lambda^N$. We remark that in the case when $\sigma < -2$, this region lies entirely within D, so in particular it ensures that $\lambda < \lambda_0$.

We now divide and conquer depending on the spatial location:

a) Outside of the region $\tilde{J}_\lambda$, we use the elliptic bound (5.25). This yields

$$\|u_\lambda\|_{L^\infty(\tilde{J}_\lambda^C)} \lesssim \frac{\lambda^{-s_1+\frac{1}{2}}}{t\lambda a''(\lambda)}t^{C^2\epsilon^2} \lesssim \frac{1}{(t\lambda^2 a''(\lambda))^{\frac{1}{2}}}t^{C^2\epsilon^2},$$

which suffices if $t \gtrsim \lambda^N$.

b) It remains to bound $\chi_\lambda u_\lambda$. By Lemma 5.10, this is equivalent to showing that our asymptotic function γ satisfies a similar bound, namely

$$|\gamma(t,v)| \lesssim \epsilon\lambda^{-\frac{\delta}{2}}, \qquad v \in J_\lambda, \qquad t \gtrsim \lambda^N. \tag{5.64}$$

At this point it is natural to split into two cases:

A) $\sigma > -2$. Here we initialize γ at $t = 1$, and use the asymptotic equation (5.58) to bound γ at later times.

B) $\sigma \leq -2$. Here γ is restricted to the set D, so for each velocity $v \in J_\lambda$ we initialize at times where $t \approx \lambda^N$, using (5.63), and propagate the bound using the asymptotic equation(5.58).

□

References

AIT-global [1] Albert Ai, Mihaela Ifrim, and Daniel Tataru. Two dimensional gravity waves at low regularity II: Global solutions. *arXiv e-prints*, page arXiv:2009.11513, September 2020.

MR1207209 [2] P. Deift and X. Zhou. A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the MKdV equation. *Ann. of Math. (2)*, 137(2):295–368, 1993.

Delort-KG [3] Jean-Marc Delort. Semiclassical microlocal normal forms and global solutions of modified one-dimensional KG equations. *Ann. Inst. Fourier (Grenoble)*, 66(4):1451–1528, 2016.

GV [4] J. Ginibre and G. Velo. Smoothing properties and retarded estimates for some dispersive evolution equations. *Comm. Math. Phys.*, 144(1):163–188, 1992.

BHG [5] Benjamin Harrop-Griffiths. Long time behavior of solutions to the mKdV. *Comm. Partial Differential Equations*, 41(2):282–317, 2016.

r2 [6] Benjamin Harrop-Griffiths and Jeremy L. Marzuola. Small data global solutions for the Camassa-Choi equations. *Nonlinearity*, 31(5):1868–1904, 2018.

HN [7] Nakao Hayashi and Pavel I. Naumkin. Asymptotics for large time of solutions to the nonlinear Schrödinger and Hartree equations. *Amer. J. Math.*, 120(2):369–389, 1998.

NLS [8] Mihaela Ifrim and Daniel Tataru. Global bounds for the cubic nonlinear Schrödinger equation (NLS) in one space dimension. *Nonlinearity*, 28(8):2661–2675, 2015.

IT-global [9] Mihaela Ifrim and Daniel Tataru. Two dimensional water waves in holomorphic coordinates II: Global solutions. *Bull. Soc. Math. France*, 144(2):369–394, 2016.

IT-t [10] Mihaela Ifrim and Daniel Tataru. The lifespan of small data solutions in two dimensional capillary water waves. *Arch. Ration. Mech. Anal.*, 225(3):1279–1346, 2017.

NLS-approx [11] Mihaela Ifrim and Daniel Tataru. The NLS approximation for two dimensional deep gravity waves. *Sci. China Math.*, 62(6):1101–1120, 2019.

msri-summer [12] Mihaela Ifrim and Daniel Tataru. Introduction to water waves [virtual summer graduate school]. `https://www.msri.org/summer_schools/910`, August 2020.

KP [13] Jun Kato and Fabio Pusateri. A new proof of long-range scattering for critical nonlinear Schrödinger equations. *Differential Integral Equations*, 24(9-10):923–940, 2011.

KeTa [14] Markus Keel and Terence Tao. Endpoint Strichartz estimates. *Amer. J. Math.*, 120(5):955–980, 1998.

KT [15] Herbert Koch and Daniel Tataru. Dispersive estimates for principally normal pseudodifferential operators. *Comm. Pure Appl. Math.*, 58(2):217–284, 2005.

LS [16] Hans Lindblad and Avy Soffer. Scattering and small data completeness for the critical nonlinear Schrödinger equation. *Nonlinearity*, 19(2):345–353, 2006.

r1 [17] Grace Liu. Modified scattering for the cubic Schrödinger equation small data solution on product space. *SIAM J. Math. Anal.*, 51(5):4023–4073, 2019.

r3 [18] Mamoru Okamoto. Long-time behavior of solutions to the fifth-order modified KdV-type equation. *Adv. Differential Equations*, 23(9-10):751–792, 2018.

MR3409892 [19] Guido Schneider. Validity and non-validity of the nonlinear Schrödinger equation as a model for water waves. In *Lectures on the theory of water waves*, volume 426 of *London Math. Soc. Lecture Note Ser.*, pages 121–139. Cambridge Univ. Press, Cambridge, 2016.

Stein [20] Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

T-FF [21] Daniel Tataru. On the Fefferman-Phong inequality and related problems. *Comm. Partial Differential Equations*, 27(11-12):2101–2138, 2002.

Pisa [22] Daniel Tataru. Phase space transforms and microlocal analysis. In *Phase space analysis of partial differential equations. Vol. II*, Pubbl. Cent. Ric. Mat. Ennio Giorgi, pages 505–524. Scuola Norm. Sup., Pisa, 2004.

Department of Mathematics, University of Wisconsin, Madison
Email address: `ifrim@wisc.edu`

Department of Mathematics, University of California at Berkeley
Email address: `tataru@math.berkeley.edu`