ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Functional Analysis 233 (2006) 228–259

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

Sharp well-posedness and ill-posedness results for a quadratic non-linear Schrödinger equation

Ioan Bejenaru, Terence Tao*

Department of Mathematics, UCLA, Los Angeles, CA 90095-1555, USA

Received 11 August 2005; accepted 17 August 2005
Communicated by J. Bourgain
Available online 23 September 2005

Abstract

We establish that the quadratic non-linear Schrödinger equation

$$iu_t + u_{xx} = u^2,$$

where $u : \mathbb{R} \times \mathbb{R} \to \mathbb{C}$, is locally well-posed in $H^s(\mathbb{R})$ when $s \geqslant -1$ and ill-posed when $s < -1$. Previous work in [C. Kenig, G. Ponce, L. Vega, Quadratic forms for the 1-D semilinear Schrödinger equation, Trans. Amer. Math. Soc. 346 (1996) 3323–3353] had established local well-posedness for $s > -\frac{3}{4}$. The local well-posedness is achieved by an iteration using a modification of the standard $X^{s,b}$ spaces. The ill-posedness uses an abstract and general argument relying on the high-to-low frequency cascade present in the non-linearity, and a computation of the first non-linear iterate.

Keywords: Nonlinear Schrödinger equation; $X^{s,b}$ spaces; Ill-posedness; Picard iteration method

* Corresponding author.
E-mail addresses: bejenaru@math.ucla.edu (I. Bejenaru), tao@math.ucla.edu (T. Tao).

0022-1236/$ - see front matter
doi:10.1016/j.jfa.2005.08.004

1. Introduction

Consider the Cauchy problem

$$\begin{aligned} iu_t + \Delta u &= F(u), \\ u(0) &= f \in H_x^s(\mathbb{R}^n), \\ u &\in C_t^0 H_x^s([0,T] \times \mathbb{R}^n) \end{aligned} \tag{1}$$

for a semilinear Schrödinger equation on $[0,T] \times \mathbb{R}^n$ for some local time interval[1] $[0,T]$ with $T > 0$, where the initial data f are given in some Sobolev space[2] $H_x^s(\mathbb{R}^n)$, the solution u is complex-valued, and $F : \mathbb{C} \to \mathbb{C}$ is a power-type non-linearity (thus $|F(z)| \sim |z|^p$ for some exponent p, and similarly for derivatives). To fix conventions, we define the Sobolev space $H_x^s(\mathbb{R}^n)$ for any $s \in \mathbb{R}$ as the Banach space of distributions f for which the norm

$$\|f\|_{H_x^s(\mathbb{R}^n)} := \|\langle \xi \rangle^s \hat{f}\|_{L_\xi^2(\mathbb{R}^n)}$$

is finite, where $\langle \xi \rangle := (1 + |\xi|^2)^{1/2}$, and $\hat{f}$ is the Fourier transform

$$\hat{f}(\xi) := \int_{\mathbb{R}^n} e^{ix\cdot\xi} f(x)\, dx.$$

This particular problem has been studied extensively in the literature, for various values of n, s, and F, as it is a simple model for the more general Cauchy problem for non-linear dispersive equations. In the situation considered in this paper, the regularity $H_x^s(\mathbb{R}^n)$ is very low (in fact, s will be negative), so that the solutions to (1) cannot be interpreted in the classical sense; we will make sense of the equation for rough data later, but suffice it to say for now that we will be able to show that the rough $C_t^0 H_x^s$ solutions we construct will be strong limits in $C_t^0 H_x^s$ of smooth solutions.

If F is smooth, then one typically obtains a local well-posedness result[3] when s is large, but not when s is small. For instance, for the power-type semilinear equation

$$iu_t + \Delta u = \pm |u|^{p-1} u \tag{2}$$

[1] One can also consider the Cauchy problem backwards in time, on some interval $[-T, 0]$, but this backwards problem is equivalent to the forwards problem after applying the conjugation $u(t) \mapsto \overline{u(-t)}$, $F(z) \mapsto \overline{F(\overline{z})}$.

[2] We will subscript spatial function spaces by x and temporal function spaces by t; thus, for instance $C_t^0 H_x^s([0,T] \times \mathbb{R}^n)$ is the space of all functions $u(t,x)$ for which the map $t \mapsto u(t)$ is continuous into $H_x^s(\mathbb{R}^n)$, equipped with the norm $\sup_{0 \leqslant t \leqslant T} \|u(t)\|_{H_x^s(\mathbb{R}^n)}$.

[3] By this we mean that for any choice of initial data $u_0 \in H_x^s(\mathbb{R}^n)$ there exists a time $T > 0$, and a continuous solution map defined in a small ball in $H_x^s(\mathbb{R}^n)$ centered at u_0, and taking values in $C_t^0 H_x^s([-T,T] \times \mathbb{R}^n)$. Furthermore, when the data are restricted to a suitable smooth class (e.g. $H^s(\mathbb{R}^n)$ for $s > n/2$), then the solution map agrees with the standard (and unique) solutions that can be constructed for instance by the energy method.

for $p > 1$ and either choice of sign $\pm$, this equation is locally well-posed[4] when $s \geqslant \min(0, s_c)$, where the scaling regularity s_c is defined by

$$s_c = \frac{n}{2} - \frac{2}{p-1},$$

see for instance [3]. This condition is fairly sharp; if $s < \min(0, s_c)$, then the solution map is known not to be uniformly continuous from H^s to $C_t H^s([-T, T] \times \mathbb{R}^n)$, even if we make T small and restrict the data to a small ball around the origin; see [1,6,4]. The regularity $H^{s_c}(\mathbb{R}^n)$ is a natural limit to well-posedness as it is preserved by the scale invariance

$$u(t,x) \mapsto \lambda^{-2/(p-1)} u\left(\frac{t}{\lambda^2}, \frac{x}{\lambda}\right) \quad \text{for any } \lambda > 0 \tag{3}$$

of Eq. (2), while the regularity $L^2(\mathbb{R}^n)$ is another natural limit, as it is preserved by the Galilean invariance

$$u(t,x) \mapsto \exp(i(v \cdot x - |v|^2 t))u(t, x - vt) \quad \text{for any } v \in \mathbb{R}^n$$

of the same equation.

Thus it would seem that the local well-posedness theory for semilinear Schrödinger equations with power-type non-linearity is complete. However, it was observed in [5] that one can lower the regularity threshold for local well-posedness below $s = 0$ (i.e. below the Galilean threshold) by choosing a non-linearity which is not Galilean invariant. In particular, the one-dimensional quadratic semilinear Schrödinger equation[5]

$$\begin{aligned} iu_t + u_{xx} &= u^2, \\ u(0) &= f \in H_x^s(\mathbb{R}), \\ u &\in C_t^0 H_x^s([0,T] \times \mathbb{R}) \end{aligned} \tag{4}$$

was shown in [5] to be locally well-posed in $H_x^s(\mathbb{R})$ for all $s > -\frac{3}{4}$, by means of an iteration argument in the $X^{s,b}$ spaces; in contrast, with a quadratic non-linearity such as $|u|u$, the lowest Sobolev regularity for which one has well-posedness is $L_x^2(\mathbb{R})$ (see [10,3,1,6,4]). One should remark that these regularities are well above the scaling regularity, which in this case is $s_c = -\frac{3}{2}$; thus these results are subcritical with respect to scaling.

[4] If p is not an odd integer, then we also need the technical condition $p > \lfloor s \rfloor + 1$ to ensure that the non-linearity is at least as regular as the non-linearity.

[5] We shall only consider scalar solutions here for simplicity. However, one can extend the analysis here to finite-dimensional systems with a quadratic form non-linearity $Q(u,u)$ which is linear (as opposed to anti-linear) in both variables, without any difficulty. Also, there is no distinction between the $+u^2$ and $-u^2$ non-linearities, as can be seen by the transformation $u \mapsto -u$.

Paper [5] also considered other quadratic non-linearities such as $u\overline{u}$ and $\overline{u}^2$, obtaining similar results. However, we wish to focus on the u^2 non-linearity in (4) to point out one interesting feature of this equation, namely its complex analyticity in u. This manifests itself in a number of ways; in particular, this equation in the spacetime-frequency domain (τ, ξ) is almost entirely supported in the upper half-space $\tau > 0$. To see this heuristically, let us formally introduce the spacetime Fourier transform

$$\tilde{u}(\tau, \xi) := \int_{\mathbb{R}} \int_{\mathbb{R}} u(t, x) e^{i(t\tau + x\xi)}\, dt\, dx$$

(ignoring for now the issue of extending u globally in time); then (4) transforms (heuristically, at least) to the integral equation

$$\tilde{u}(\tau, \xi) = \delta(\tau - \xi^2)\hat{u}_0(\xi) + \frac{1}{\tau - \xi^2} \int \int_{\tau = \tau_1 + \tau_2;\, \xi = \xi_1 + \xi_2} \tilde{u}(\tau_1, \xi_1)\tilde{u}(\tau_2, \xi_2), \tag{5}$$

where δ is the Dirac delta function, and we will be deliberately vague about how to define in a distributional sense the operation of dividing by $\tau - \xi^2$. If one tries to solve Eq. (5) iteratively, viewing this equation as a way to obtain a new approximation of $\tilde{u}$ from an old one, starting from (say) the zero solution $\tilde{u} = 0$, we see that all the iterates $\tilde{u}$ are supported on the upper half-plane $\tau > 0$, and thus we expect the final solution also to do so.

This additional property of problem (4) suggests that perhaps some further improvement to $H_x^{-3/4}(\mathbb{R})$ and beyond is possible; for instance, (5) suggests that the solution is unlikely to be concentrated near the spacetime frequency origin $(\tau, \xi) = (0, 0)$. It is also similarly difficult for the iterates of the solution to return back to the parabola $\tau = \xi^2$, where the solution is expected to concentrate (in analogy with the linear solution).

A first step in this direction was made by Muramutu and Taoka [7], obtaining local well-posedness in the Besov space $B_2^{-3/4,1}(\mathbb{R})$, which is slightly stronger than $H_x^{-3/4}(\mathbb{R})$, by a refinement of the $X^{s,b}$ iteration method. However, it was shown in [5,8] that the key bilinear $X^{s,b}$ estimate needed to apply this method failed for $H_x^s(\mathbb{R})$ for any $s \leqslant -\frac{3}{4}$. Nevertheless, it turns out that we can use the additional information that $\tilde{u}$ concentrates on the upper half-plane to avoid most of the counterexamples in [5,8], and after modification of the $X^{s,b}$ spaces we can in fact avoid the remaining counterexamples also, to obtain our first main theorem:

Theorem 1 (*Local well-posedness in $H_x^{-1}(\mathbb{R})$*). *Let $r > 0$ be any radius, and let B_r be the ball*

$$B_r := B_{H_x^{-1}(\mathbb{R})}(0, r) := \{u_0 \in H_x^{-1}(\mathbb{R}) : \|u_0\|_{H_x^{-1}(\mathbb{R})} < r\}.$$

Then there exists a time $T > 0$ (in fact we obtain $T = \max(1, cr^{-1/2})$ for some absolute constant $c > 0$) and a map $f \mapsto u[f]$ which is continuous from B_r to

$C_t^0 H_x^{-1}([0,T]\times\mathbb{R})$, *such that the restriction of this map to* $B_r \cap H_x^s(\mathbb{R})$ (*with the* $H_x^s(\mathbb{R})$ *topology*) *maps continuously to* $C_t^0 H^s([0,T]\times\mathbb{R})$ *for any* $s \geqslant -1$. *Furthermore, if* f *lies in a smooth space, say* $B_r \cap H_x^3(\mathbb{R})$, *then,* $u[f]$ *lies in* $C_t^0 H_x^3 \cap C_t^1 H_x^1([0,T]\times\mathbb{R})$ *and solves Eq.* (4) *in the classical sense.*

We prove this theorem in Section 4, as a consequence of standard iteration machinery and a construction of a function space obeying certain linear and bilinear estimates. It is easy to establish (e.g. by energy methods) that classical solutions to (4) in $C_t^0 H_x^3 \cap C_t^1 H_x^1$ are unique, and so the solution map S given by the above theorem is the unique strong limit of smooth solutions in $C_t^0 H_x^{-1}$. The theorem also shows that if the solution blows up at some time T_*, then the $H_x^{-1}(\mathbb{R})$ norm of $u(t)$ must blow up at a rate of $c|T_* - t|^{-2}$ or greater as $t \to T_*$. The main novelty in the proof is a modification of the $X^{s,b}$ spaces in order to exploit the concentration of the solution in the upper half-plane $\tau > 0$, and also to deal with the failure of the $X^{s,b}$ norms to adequately control the behavior of this equation near the time axis $\xi = 0$.

Our second main result is that the threshold $s \geqslant -1$ is completely sharp.

Theorem 2 (*Ill-posedness below* $H_x^{-1}(\mathbb{R})$). *Let* $r > 0$ *be arbitrary, and let* T *and* $f \mapsto u[f]$ *be as in Theorem* 1. *Then the solution map* $f \mapsto u[f]$ *is discontinuous on* B_r (*with the* $H_x^s(\mathbb{R})$ *topology*) *to* $C_t^0 H_x^{-1}([0,T]\times\mathbb{R})$ (*with the* $C_t^0 H_x^{s'}([0,T]\times\mathbb{R})$ *topology*) *for any* $s < -1$ *and* $s' \in \mathbb{R}$.

This theorem will be proven by demonstrating a high-to-low frequency cascade in the first non-trivial iterate of an integral equation associated to (4); see Section 4. We will then invoke a rather general result (Proposition 1 below), which may be of independent interest, which shows that any non-linear evolution equation with polynomial non-linearity will be illposed whenever a high-to-low frequency cascade in one of its iterates can be established.

2. Reduction to an integral equation

We first give some very standard reductions for Theorem 1. The first is to use the scale invariance (3) to scale the radius r to be small. Indeed, if one defines

$$f^{(\lambda)}(x) := \frac{1}{\lambda^2} f\left(\frac{x}{\lambda}\right)$$

then a simple computation shows[6]

$$\|f^{(\lambda)}\|_{H_x^{-1}(\mathbb{R})} \leqslant C\lambda^{-1/2} \|f\|_{H_x^{-1}(\mathbb{R})}$$

[6] Here and in the sequel we use $C, c > 0$ to denote various positive absolute constants.

for $\lambda > 1$. We can thus rescale the initial data to be arbitrarily small in $H^s_x(\mathbb{R})$ norm. It thus suffices to prove Theorem 1 when $T = 1$ and r is sufficiently small (smaller than some absolute constant $c > 0$).

Next, we shall use Duhamel's formula to recast (4) in the integral form

$$u(t) = \exp(it\partial_{xx})f + \int_0^t \exp(i(t-s)\partial_{xx})(u(s)^2)\,ds, \tag{6}$$

where $\exp(it\partial_{xx})$ is the propagator for the free Schrödinger equation $iu_t + u_{xx} = 0$, which can be defined for instance using the spatial Fourier transform as

$$\widehat{\exp(it\partial_{xx})f}(\xi) := e^{-it\xi^2}\hat{f}(\xi).$$

Following Bourgain [2], it turns out to be convenient to replace the local-in-time integral equation (6) with a global-in-time truncated integral equation. Let $\eta : \mathbb{R} \to \mathbb{R}$ be a smooth bump function such that $\eta(t) = 1$ for $|t| \leqslant 1$ and $\eta(t) = 0$ for $|t| > 2$, and let $a(t) := \frac{1}{2}\,\mathrm{sgn}(t)\eta(t/5)$. Then observe that

$$\int_0^t g(s)\,ds = \eta(t)\int_{\mathbb{R}} a(s)g(s)\,ds + \int_{\mathbb{R}} a(t-s)g(s)\,ds$$

for all $0 \leqslant t \leqslant 1$ and any $g : \mathbb{R} \to \mathbb{R}$. Hence we can replace (6) on the time interval $0 \leqslant t \leqslant 1$ by the equation

$$\begin{aligned} u(t) = {}& \eta(t)\exp(it\partial_{xx})f + \eta(t)\exp(it\partial_{xx})\int_{\mathbb{R}} a(s)\exp(-is\partial_{xx})(u(s)^2)\,ds \\ & + \int_{\mathbb{R}} a(t-s)\exp(i(t-s)\partial_{xx})(u(s)^2)\,ds, \end{aligned} \tag{7}$$

in the sense that any classical (e.g. $C^0_t H^3_x(\mathbb{R} \times \mathbb{R})$ will do) global-in-time solution to (7) is also a classical solution to (6) and hence (4). Note that if $u \in C^0_t H^3_x$, one can easily use (6) or (7) and Sobolev embedding to conclude that $u \in C^1_t H^1_x$, and so one can make sense of (4) in a classical sense.

It remains to find global-in-time solutions to (7) for initial data f in B_r.

We will write (7) more abstractly as

$$u = L(f) + N_2(u, u), \tag{8}$$

where L is the linear operator

$$L(f)(t) := \eta(t)\exp(it\partial_{xx})f \tag{9}$$

and N_2 is the bilinear operator

$$N_2(u,v)(t) := \eta(t)\exp(it\partial_{xx})\int_{\mathbb{R}} a(s)\exp(-is\partial_{xx})(u(s)v(s))\,ds + \int_{\mathbb{R}} a(t-s)\exp(i(t-s)\partial_{xx})(u(s)v(s))\,ds. \tag{10}$$

The subscript 2 denotes the fact that this operator is quadratic. We now pause to systematically develop the well-posedness, persistence of regularity, and ill-posedness theory for such an abstract type of operator. This theory is mostly standard, but the material on ill-posedness may be of independent interest.

3. Abstract well-posedness and ill-posedness theory

In this section we shall consider an abstract semilinear evolution equation with a k-linear non-linearity for some $k \geqslant 2$. Specifically, we consider the abstract equation

$$u = L(f) + N_k(u,\ldots,u), \tag{11}$$

where the initial data f take values in some data space D, the solution u takes values in some solution space S, the linear operator $L : D \to S$ is densely defined, and the k-linear operator $N_k : S \times \cdots \times S \to S$ is also densely defined.

Definition 1 (*Quantitative well-posedness*). Let $(D, \|\|_D)$ be a Banach space of initial data, and $(S, \|\|_S)$ be a Banach space of spacetime functions. We say that Eq. (11) is quantitatively well posed in the spaces D, S if one has estimates of the form[7]

$$\|L(f)\|_S \leqslant C\|f\|_D$$

and

$$\|N_k(u_1,\ldots,u_k)\|_S \leqslant C\|u_1\|_S \ldots \|u_k\|_S$$

for all $f \in D$, $u_1,\ldots,u_k \in S$ and some constant $C > 0$.

Note that once Eq. (11) is quantitatively well-posed, one can extend the densely defined operators L and N_k to all of D and $S \times \cdots \times S$, respectively, in a unique continuous fashion.

If $(D, \|\|_D)$ is a Banach space, we use $B_D(0,r) := \{f \in D : \|f\|_D < r\}$ to denote the usual open ball of radius r around the origin. The standard well-posedness result

[7] We adopt the convention that if X is a Banach space, then $\|u\|_X$ denotes the norm of u in X, and that $\|u\|_X = \infty$ if $u \notin X$.

for such equations is that quantitative well-posedness implies analytic well-posedness. More precisely:

Theorem 3 (*Standard well-posedness theorem*). *Suppose Eq.* (11) *is quantitatively well posed in the spaces* D, S. *Then there exist constants* $C_0, \varepsilon_0 > 0$ *such that for all* $f \in B_D(0, \varepsilon_0)$, *there exists a unique solution* $u[f] \in B_S(0, C_0\varepsilon_0)$ *to Eq.* (11). *More specifically*, *if we define the non-linear maps* $A_n : D \to S$ *for* $n = 1, 2, \ldots$ *by the recursive formulae*

$$
\begin{aligned}
A_1(f) &:= L(f) \\
A_n(f) &:= \sum_{n_1,\ldots,n_k \geqslant 1: n_1+\cdots+n_k=n} N_k(A_{n_1}(f), \ldots, A_{n_k}(f)) \quad \text{for } n > 1,
\end{aligned}
$$

then we have the homogeneity property

$$A_n(\lambda f) = \lambda^n A_n(f) \quad \text{for all } \lambda \in \mathbb{R}, n \geqslant 1, \text{ and } f \in D \tag{12}$$

(*so in particular* $A_n(0) = 0$) *and the Lipschitz bound*

$$\|A_n(f) - A_n(g)\|_S \leqslant \|f - g\|_D C_1^n (\|f\|_D + \|g\|_D)^{n-1} \tag{13}$$

for some $C_1 > 0$, *all* $f, g \in D$, *and all* $n \geqslant 1$. *In particular we have*

$$\|A_n(f)\|_S \leqslant C_2^n \|f\|_D^n \tag{14}$$

for some $C_2 > 0$. *Furthermore*, *we have the absolutely convergent* (*in S*) *power series expansion*

$$u[f] = \sum_{n=1}^{\infty} A_n(f) \tag{15}$$

for all $f \in B_D(0, \varepsilon_0)$.

Thus for instance, if $k = 2$, then

$$
\begin{aligned}
A_1(f) &= L(f), \\
A_2(f) &= N_2(Lf, Lf), \\
A_3(f) &= N_2(Lf, N_2(Lf, Lf)) + N_2(N_2(Lf, Lf), Lf), \\
&\vdots
\end{aligned}
$$

whereas if $k = 3$, then

$$\begin{aligned} A_1(f) &= L(f), \\ A_2(f) &= 0, \\ A_3(f) &= N_3(Lf, Lf, Lf), \\ A_4(f) &= 0. \\ &\vdots \end{aligned}$$

In general, one can express $A_n(f)$ as a sum over k-ary trees with n nodes, but we will not need such an explicit representation here.

Proof. We shall be somewhat brief here since this theorem is well-known. For any fixed $f \in B_D(0, \varepsilon_0)$, one can easily verify from the quantitative well-posedness hypothesis that the map $u \mapsto L(f) + N_k(u, \dots, u)$ will be a contraction from $B_S(0, C_0\varepsilon_0)$ to $B_S(0, C_0\varepsilon_0)$ if C_0 is sufficiently large and ε_0 is sufficiently small (depending on C_0). The contraction mapping theorem then gives the existence and uniqueness of the map $f \mapsto u[f]$.

Now we start proving the power series expansion. One can easily verify (12) by induction; in fact, we easily verify that

$$A_n(f) = M_n(f, \dots, f)$$

for some n-linear map $M_n : D \times \cdots \times D \to S$. One can also inductively obtain an estimate of the form

$$\|A_n(f)\|_S \leqslant (C_3\|f\|_D)^n \tag{16}$$

for some large constant $C_3 > 0$ (depending of course on the quantitative well-posedness constants and on k); note this already gives (14). We remark that for the purposes of proving (16), it is actually slightly easier for inductive purposes to establish a slight stronger upper bound of $(C_3\|f\|_D)^n/(C_4 n^{C_5})$, where C_4 and C_5 are somewhat large but not as large as C_3.

Now we prove (13). By symmetry we may take $\|f\|_D \leqslant \|g\|_D$, and by scaling we can take $\|g\|_D \leqslant 1$. We can of course assume that $f \neq g$. Write $t := \|f - g\|_D$, and write $f = g + th$; thus $0 < t \leqslant 2$ and $\|h\|_D = 1$. It then suffices to show that

$$\|A_n(g + th) - A_n(g)\|_S \leqslant t C_1^n. \tag{17}$$

The non-linear operator $A_n(f)$ can be written as $M_n(f, \dots, f)$ for some n-linear operator M_n, which implies that for fixed g, h, the function $s \mapsto A_n(g + sh) - A_n(g)$ is a polynomial of degree at most n in s with a zero constant term; thus,

$$A_n(g+sh) - A_n(g) = \sum_{j=1}^{n} F_j s^j \tag{18}$$

for some $F_j \in S$. From (14) and the triangle inequality we also have

$$\|A_n(g+sh) - A_n(g)\|_S \leqslant (4C_2)^n$$

for all $|s| \leqslant 1$ (say) and some constant $C > 0$. Thus we have

$$\left\| \sum_{j=1}^{n} F_j s^j \right\|_S \leqslant (4C_2)^n$$

for all $|s| \leqslant 1$. Using the Lagrange interpolation formula to recover $F_1, \ldots, F_n$ from various sample points of $\sum_{j=1}^{n} F_j s^j$, we conclude that

$$\|F_j\|_S \leqslant C^n$$

for all $1 \leqslant j \leqslant n$ and some $C > 0$. Inserting this back into (18) we obtain (17). (Note that if we allow s to be complex, one could also proceed using the Cauchy integral formula instead of the Lagrange interpolation formula.)

From (14) we see that the series $\sum_{n=1}^{\infty} A_n(f)$ is absolutely convergent in S for ε small enough. If for any integer $K \geqslant 1$, we let u_K be the partial sum $u_K := \sum_{n=1}^{K} A_n(f)$, one can easily verify a formula of the form

$$L(f) + N_k(u_K, \ldots, u_K) = u_K + \sum_{K < n \leqslant nK} A_{n,K}(f),$$

where $A_{n,K}(f)$ is a non-linear expression of f which consists of some subset of the terms used to form $A_n(f)$. One can then again use induction to obtain estimates of the form

$$\|A_{n,K}(f)\|_S \leqslant (C\|f\|_S)^n$$

for some $C > 0$, and hence we see that if $f \in B_D(0, \varepsilon_0)$ for ε_0 sufficiently small, that

$$\|L(f) + N_k(u_K, \ldots, u_K) - u_K\|_S \leqslant (C\varepsilon_0)^n.$$

Using the contraction mapping principle again, we see that u_K converges to $u[f]$ in S norm, and we obtain (15). $\square$

From (15), (13) one can verify that the map $f \mapsto u[f]$ is continuous (in fact Lipschitz continuous) from $B_D(0, \varepsilon_0)$ to $B_S(0, C_0\varepsilon_0)$ for ε_0 small enough. In fact this Lipschitz continuity can also be read off directly from the contraction mapping theorem.

Now, we investigate continuity in both finer and coarser topologies. The basic result for finer topologies is as follows:

Theorem 4 (*Standard persistence of regularity theorem*). *Suppose Eq.* (11) *is quantitatively well posed in the spaces D, S, and let $f \mapsto u[f]$ be the solution map from $B_D(0, \varepsilon_0)$ to $u[f] \in B_S(0, C_0\varepsilon_0)$ constructed in Theorem* 3. *Suppose we are given spaces $(D', \|\|_{D'})$ and $(S', \|\|_{S'})$ obeying the estimates*

$$\|L(f)\|_{S'} \leqslant C\|f\|_{D'}$$

and

$$\|N_k(u_1, \dots, u_k)\|_{S'} \leqslant C\sum_{j=1}^{k}\|u_j\|_{S'}\prod_{1\leqslant i\leqslant k; i\not\Leftrightarrow j}\|u_i\|_S.$$

Then, if ε_0 is sufficiently small, the solution map is also continuous from $B_D(0, \varepsilon_0) \cap D'$ (in the D' topology) to $B_S(0, C_0\varepsilon_0) \cap S'$ (in the S' topology).

Proof. If $f \in B_D(0, \varepsilon_0) \cap D'$ for a suitably small ε_0, then the above estimates easily imply that $u \mapsto L(f) + N_k(u, \dots, u)$ will be a contraction in the S' norm from $B_S(0, C_0\varepsilon_0) \cap S'$ to $B_S(0, C_0\varepsilon_0) \cap S'$, and the claim follows. □

Now we turn to coarser topologies. The basic result here is that if the map $f \mapsto \sum_{n=1}^{\infty} A_n(f)$ is continuous in a coarse topology, then each component $f \mapsto A_n(f)$ of the series is also continuous in this coarse topology.

Proposition 1. *Suppose that Eq.* (11) *is quantitatively well-posed in the Banach spaces D and S, with a solution map $f \mapsto u[f]$ from a ball B_D in D to a ball B_S in S. Suppose that these spaces are then given other norms D' and S', which are weaker than D and S in the sense that*

$$\|f\|_{D'} \leqslant C\|f\|_D, \quad \|u\|_{S'} \leqslant C\|u\|_S$$

for some absolute constant C. (Note that D is unlikely to be complete in the D' norm, and similarly for S and S'.) Suppose that the solution map $f \mapsto u[f]$ is continuous from $(B_D, \|\|_{D'})$ (i.e. the ball B_D equipped with the D' topology) to $(B_S, \|\|_{S'})$. Then for each n, the non-linear operator $A_n : D \to S$ is continuous from $(B_D, \|\|_{D'})$ to $(S, \|\|_{S'})$.

Proof. We induct on n, assuming that for all $n' < n$ the operator $A_{n'} : D \to S$ has already been shown to be continuous from $(D, \|\|_{D'})$ to $(S, \|\|_{S'})$.

Let f_m be a sequence in B_D which converges to $f \in B_D$ in the D' topology, thus, $\|f_m - f\|_{D'} \to 0$. Our task is to show that $\|A_n(f_m) - A_n(f)\|_{S'} \to 0$.

Now let $0 < \lambda \leqslant 1$ be a small number to be chosen later. By hypothesis, the map $f \mapsto u[f]$ is continuous from $(B_D, \|\|_{D'})$ to $(B_S, \|\|_{S'})$, and hence

$$\lim_{m\to\infty} \|u[\lambda f_m] - u[\lambda f]\|_{S'} = 0.$$

Expanding the power series and using homogeneity, we have

$$\lim_{m\to\infty} \left\| \sum_{n'=1}^{\infty} \lambda^{n'} (A_{n'}(f_m) - A_{n'}(f)) \right\|_{S'} = 0.$$

By the induction hypothesis we already have

$$\lim_{m\to\infty} \left\| \sum_{n'<n} \lambda^{n'} (A_{n'}(f_m) - A_{n'}(f)) \right\|_{S'} = 0,$$

so on subtracting and then dividing by λ^n, we conclude that

$$\lim_{m\to\infty} \left\| \sum_{n'\geqslant n} \lambda^{n'-n} (A_{n'}(f_m) - A_{n'}(f)) \right\|_{S'} = 0$$

and hence by the triangle inequality

$$\limsup_{m\to\infty} \|A_n(f_m) - A_n(f))\|_{S'} \leqslant \sum_{n'>n} \lambda^{n'-n} \sup_m \|A_{n'}(f_m) - A_{n'}(f)\|_{S'}.$$

Using (13), we conclude

$$\limsup_{m\to\infty} \|A_n(f_m) - A_n(f))\|_{S'} \leqslant \sum_{n'>n} \lambda^{n'-n} (C\varepsilon)^{n'}.$$

The right-hand side is convergent for λ small enough. Taking $\lambda \to 0$, we conclude

$$\limsup_{m\to\infty} \|A_n(f_m) - A_n(f))\|_{S'} = 0$$

and the claim follows. $\square$

This proposition gives us a way to *disprove* well-posedness in coarse topologies, simply by establishing that at least one of the operators A_n is discontinuous. This tends to be the case if A_n contains a significant "high-to-low frequency cascade", and we shall exploit this (in the $n = 2$ case) to prove Theorem 2.

4. Reduction to function spaces

We can now reduce Theorems 1 and 2 to the construction of a certain pair $\mathcal{S}^s(\mathbb{R}\times\mathbb{R})$ and $\mathcal{N}^s(\mathbb{R}\times\mathbb{R})$ of function spaces for each $s \in \mathbb{R}$, the verification of certain estimates for these spaces, and the verification of a certain bad behavior of the quadratic operator A_2. The precise statements are as follows: define a *bump function* to be a smooth compactly supported function $t \mapsto \eta(t)$ of the time variable $t \in \mathbb{R}$.

Proposition 2 (*Function spaces*). *For any $s \in \mathbb{R}$ there exists a Banach space $\mathcal{S}^s(\mathbb{R}\times\mathbb{R})$ (the "solution space" at regularity $H^s_x(\mathbb{R})$) and a Banach space $\mathcal{N}^s(\mathbb{R}\times\mathbb{R})$ (the "nonlinearity space" at regularity $H^s_x(\mathbb{R})$), with the following properties*:

(i) (*Density*). *The Schwartz functions on $\mathbb{R}\times\mathbb{R}$ are dense in $\mathcal{S}^s(\mathbb{R}\times\mathbb{R})$ and in $\mathcal{N}^s(\mathbb{R}\times\mathbb{R})$.*

(ii) (*Nesting*). *If $s \leqslant s'$ and $u \in \mathcal{S}^{s'}(\mathbb{R}\times\mathbb{R})$, then*

$$\|u\|_{\mathcal{S}^s(\mathbb{R}\times\mathbb{R})} \leqslant \|u\|_{\mathcal{S}^{s'}(\mathbb{R}\times\mathbb{R})}.$$

Similarly, if $F \in \mathcal{N}^{s'}(\mathbb{R}\times\mathbb{R})$ then

$$\|F\|_{\mathcal{N}^s(\mathbb{R}\times\mathbb{R})} \leqslant \|F\|_{\mathcal{N}^{s'}(\mathbb{R}\times\mathbb{R})}.$$

(iii) (*Energy estimate*). *If $u \in \mathcal{S}^s(\mathbb{R}\times\mathbb{R})$, then*[8]

$$\|u\|_{C^0_t H^s_x(\mathbb{R}\times\mathbb{R})} \leqslant C_s \|u\|_{\mathcal{S}^s(\mathbb{R}\times\mathbb{R})}.$$

(iv) (*Homogeneous estimate*). *If $u_0 \in H^s_x(\mathbb{R})$, $u(t) = \exp(it\partial_{xx})u_0$, and $\eta(t)$ is a bump function, then $\eta u \in \mathcal{S}^s(\mathbb{R}\times\mathbb{R})$ and*

$$\|\eta u\|_{\mathcal{S}^s(\mathbb{R}\times\mathbb{R})} \leqslant C_{\eta,s} \|u_0\|_{H^s_x(\mathbb{R})}.$$

(v) (*Dual homogeneous estimate*). *If $F \in \mathcal{N}(\mathbb{R}\times\mathbb{R})$, and $\eta(t)$ is a bump function, then*

$$\left\| \int_{\mathbb{R}} \operatorname{sgn}(s)\eta(s) \exp(-is\partial_{xx}) F(s)\, ds \right\|_{H^s_x(\mathbb{R})} \leqslant C_{\eta,s} \|F\|_{\mathcal{N}^s(\mathbb{R}\times\mathbb{R})}.$$

[8] We use C_s to denote a positive constant—which can vary from line to line—that can depend on s. Similarly if we subscript C by other parameters.

(vi) (*Inhomogeneous estimate*). *If* $F \in \mathcal{N}^s(\mathbb{R}\times\mathbb{R})$, *and* $\eta(t)$ *is a bump function*, *then*

$$\left\| \int_{\mathbb{R}} \operatorname{sgn}(t-s)\eta(t-s)\exp(-i(t-s)\partial_{xx})F(s)\,ds \right\|_{\mathcal{S}^s(\mathbb{R}\times\mathbb{R})} \leqslant C_{\eta,s}\|F\|_{\mathcal{N}^s(\mathbb{R}\times\mathbb{R})}.$$

(vii) (*Non-linear estimate*). *If* $u, v \in \mathcal{S}^s(\mathbb{R}\times\mathbb{R})$ *for some* $s \geqslant -1$, *then*

$$\|uv\|_{\mathcal{N}^s(\mathbb{R}\times\mathbb{R})} \leqslant C_s(\|u\|_{\mathcal{S}^s(\mathbb{R}\times\mathbb{R})}\|v\|_{\mathcal{S}^{-1}(\mathbb{R}\times\mathbb{R})} + \|u\|_{\mathcal{S}^{-1}(\mathbb{R}\times\mathbb{R})}\|v\|_{\mathcal{S}^s(\mathbb{R}\times\mathbb{R})}).$$

Here we define the non-linear operation $(u, v) \mapsto uv$ *first for Schwartz functions, and then extend it to the general case by density.*

We shall prove this proposition in later sections. Assuming it for now, we see from (iv) to (vii) and (9), (10) that Eq. (8) is quantitatively well posed in the spaces $H_x^{-1}(\mathbb{R})$, $\mathcal{S}^{-1}(\mathbb{R}\times\mathbb{R})$, and also one has persistence of regularity (in the sense of Theorem 4 for the spaces $H_x^s(\mathbb{R})$, $\mathcal{S}^s(\mathbb{R}\times\mathbb{R})$ for any $s \geqslant -1$. Combined with (ii) and (iii), we thus establish Theorem 1 (using the reductions in Section 2).

The derivation of Theorem 2 from this proposition is almost as immediate.

Proof of Theorem 2. Fix $s < -1$ and s'; we may rescale T to equal 1. Suppose for contradiction that solution map $f \mapsto u[f]$ is continuous on B_r (with the $H_x^s(\mathbb{R})$ topology) to $C_t^0 H_x^{-1}([0,1]\times\mathbb{R})$ (with the $C_t^0 H_x^{s'}$ topology). Applying Proposition 1, we conclude that the quadratic operator

$$A_2 : f \mapsto N_2(Lf, Lf)$$

(restricted of course to $[0,1]\times\mathbb{R}$) is continuous from $H_x^s(\mathbb{R})$ to $C_t^0 H_x^{s'}([0,1]\times\mathbb{R})$. In particular, this would imply a bound of the form

$$\sup_{0\leqslant t\leqslant 1} \|A_2(f)(y)\|_{H^{s'}(\mathbb{R})} \leqslant C\|f\|^2_{H^s(\mathbb{R})}$$

for some constant C. The left-hand side can be expanded by (9), (10) as

$$\sup_{0\leqslant t\leqslant 1} \left\| \int_0^t \exp(i(t-t')\partial_{xx})((\exp(it'\partial_{xx})f)^2)\,dt' \right\|_{H^{s'}(\mathbb{R})},$$

which, after taking Fourier transforms, becomes

$$\sup_{0\leqslant t\leqslant 1} \left\| \langle\xi\rangle^{s'} \int_0^t \int_{\mathbb{R}} \exp(-i(t-t')\xi^2)\exp(it'(\xi_1^2 + (\xi-\xi_1)^2)) \times \hat{f}(\xi_1)\hat{f}(\xi-\xi_1)\,d\xi_1\,dt' \right\|_{L^2_\xi(\mathbb{R})}.$$

Now let $N > 100$ be a large number, and set

$$\hat{f} := N^{-s} 1_{[-10,10]}(|\xi| - N).$$

Since $\|f\|_{H^s_x(\mathbb{R})} = N^{-s}$, we conclude that

$$\begin{aligned}\Bigg\| \langle\xi\rangle^{s'} \int_0^t \int_{\mathbb{R}} & \exp(-i(t-t')\xi^2) \exp(it'(\xi_1^2 + (\xi-\xi_1)^2)) \\ & \times \hat{f}(\xi_1)\hat{f}(\xi-\xi_1)\, d\xi_1\, dt' \Bigg\|_{L^2_\xi(\mathbb{R})} \leqslant C \end{aligned} \tag{19}$$

for all $0 \leqslant t \leqslant 1$. Now set $t := 1/100N^2$ and localize to the region where $-1 \leqslant \xi \leqslant 1$. One can verify that

$$\mathrm{Re}(\exp(-i(t-t')\xi^2) \exp(it'(\xi_1^2 + (\xi-\xi_1)^2)) > 1/2,$$

whenever $0 \leqslant t' \leqslant t$ and ξ_1 resides in the support of *f*; hence, we obtain

$$\begin{aligned}&\|\langle\xi\rangle^{s'} \int_0^t \int_{\mathbb{R}} \exp(-i(t-t')\xi^2) \exp(it'(\xi_1^2 + (\xi-\xi_1)^2) \hat{f}(\xi_1)\hat{f}(\xi-\xi_1)\, d\xi_1\, dt'\|_{L^2_\xi(\mathbb{R})} \\ &\geqslant cN^{-2s-2}\end{aligned}$$

for some $c > 0$. But this contradicts (19) for sufficiently large *N*, since $s < -1$. This proves Theorem 2. □

It is instructive to compute the spacetime Fourier transform of $A_2(f)$. A computation shows that it has a significant component near the time-frequency axis $\xi = 0$; indeed, it has a magnitude comparable to N^{-2s-1} on a rectangle $\{\xi = O(1), \tau = 2N^2 + O(N)\}$. This is already enough to cause it to leave the $X^{s,b}$ space whenever $s < b - \frac{5}{4}$, which explains why the $X^{s,b}$ method ceases to work well when $s < -\frac{3}{4}$. However, this will be fixed by replacing this space with an L^1_τ-based space near the time axis.

5. Taking the spacetime Fourier transform

To complete the proof of Theorems 1 and 2, we need to build spaces $\mathcal{S}^s(\mathbb{R} \times \mathbb{R})$ and $\mathcal{N}^s(\mathbb{R} \times \mathbb{R})$ which obey the properties in Proposition 2. To abbreviate the notation we shall now omit the domain $\mathbb{R} \times \mathbb{R}$ from these spaces. For $s > \frac{1}{2}$ one could use the energy spaces $\mathcal{S}^s := C^0_t H^s_x$, $\mathcal{N}^s := L^1_t H^s_x$, while for $s \geqslant 0$ one could use Strichartz spaces such as $\mathcal{S}^s := C^t_0 H^s_x \cap L^4_t L^\infty_x$, $\mathcal{N}^s := L^1_t H^s_x$ (other choices are available; see [10,3]). For $s > -\frac{3}{4}$, it was shown in [5] that one could use the spaces $\mathcal{S}^s := X^{s,b}$ and $\mathcal{N}^s := X^{s,b-1}$ for any $b > \frac{1}{2}$.

In order to construct spaces which work all the way down to $s \geqslant -1$, we have to modify the $X^{s,b}$ spaces somewhat. The precise modification is somewhat complicated, so for now we shall continue to work abstractly to avoid being bogged down in details. We shall require a space W, constructed by the following proposition. Call a function on $\mathbb{R} \times \mathbb{R}$ *reasonable* if it lies in $L_t^\infty L_x^\infty(\mathbb{R} \times \mathbb{R})$ and has compact support. For any $s, b \in \mathbb{R}$, we define $\hat{X}^{s,b}$ to be the closure of the reasonable functions via the norm

$$\|f\|_{\hat{X}^{s,b}} := \|\langle\xi\rangle^s \langle\tau - \xi^2\rangle^b f\|_{L_\tau^2 L_\xi^2}.$$

These are the Fourier transforms of the usual $X^{s,b}$ spaces.

Proposition 3 (*Construction of main space*). *There exists a Banach space W, which is the closure of the reasonable functions in $\mathbb{R} \times \mathbb{R}$ by some norm $\|\,\|_W$, with the following properties for all reasonable f, g:*

- (*Monotonicity*). *If $|f| \leqslant |g|$ pointwise, then $\|f\|_W \leqslant \|g\|_W$. In particular, we have $\|f\|_W = \||f|\|_W$ (so the W norm depends only on the magnitude of the function, and not the phase).*
- (*H^{-1} Energy estimate*). *We have*

$$\|\langle\xi\rangle^{-1} f\|_{L_\xi^2 L_\tau^1} \leqslant C\|f\|_W. \tag{20}$$

- (*Homogeneous H^{-1} solution estimate*). *We have*

$$\|f\|_W \leqslant C\|f\|_{\hat{X}^{-1,100}}. \tag{21}$$

- (*Bilinear estimate*). *We have*

$$\|\langle\tau - \xi^2\rangle^{-1} f * g\|_W \leqslant \|f\|_W \|g\|_W, \tag{22}$$

where of course $f * g$ denotes spacetime convolution

$$f * g(\tau, \xi) := \int_{\mathbb{R}} \int_{\mathbb{R}} f(\tau_1, \xi_1) g(\tau_2, \xi_2)\, d\tau_1\, d\xi_1$$

using the convention

$$(\tau_1, \xi_1) + (\tau_2, \xi_2) = (\tau, \xi). \tag{23}$$

The space W has been designed with the scaling of H_x^{-1}, as this is the most important regularity in our argument. A good candidate to keep in mind for W is the space $\hat{X}^{-1,b}$ for some $b > \frac{1}{2}$; this turns out to only obey the first three properties required (and enough of the fourth property that one can establish local existence in $s > -\frac{3}{4}$ rather than $s \geqslant -1$); our final version of W shall be a modification of $\hat{X}^{-1,b}$.

We prove this proposition in the remainder of the paper. For now, let us assume it, and use it to prove Proposition 2. We will take $\mathcal{S}^s$ and $\mathcal{N}^s$ to be the closure of the Schwartz functions under the norms

$$\|u\|_{\mathcal{S}^s} := \|\langle\xi\rangle^{s+1}\tilde{u}\|_W; \quad \|F\|_{\mathcal{N}^s} := \left\| \frac{\langle\xi\rangle^{s+1}}{\langle\tau - \xi^2\rangle}\tilde{F} \right\|_W, \tag{24}$$

where $\tilde{u}(\tau, \xi)$ denotes the spacetime Fourier transform.

By construction, the density and nesting properties of $\mathcal{S}^s$ and $\mathcal{N}^s$ required for Proposition 2 are immediate. To prove the energy estimate (iii), it suffices by the usual limiting arguments to show that

$$\sup_t \|u(t)\|_{H^s_x} \leqslant C_s\|u\|_{\mathcal{S}^s}$$

when u is Schwartz. From Fourier inversion and the triangle inequality we have

$$\|u(t)\|_{H^s_x} = \left\| \int_{\mathbb{R}} \langle\xi\rangle^s e^{it\tau}\tilde{u}(\tau, \xi)\, d\tau \right\|_{L^2_\xi} \leqslant \|\langle\xi\rangle^s\tilde{u}\|_{L^2_\xi L^1_\tau}$$

and the claim now follows from (20).

To prove the homogeneous estimate (iv), observe that if $u(t) = \exp(it\partial_{xx})u_0$, then

$$\widetilde{\eta u}(\tau, \xi) = \hat{\eta}(\tau - \xi^2)\hat{u}_0(\xi)$$

and in particular from the rapid decrease of $\hat{\eta}$ we see that

$$\|\eta u\|_{\hat{X}^{s,100}} \leqslant C_\eta\|u_0\|_{H^s_x}$$

(say). Thus the claim follows from (21).

To prove the dual homogeneous estimate (v), we apply the Fourier transform and Parseval's identity in space to write

$$\left\| \int_{\mathbb{R}} \operatorname{sgn}(s)\eta(s)\exp(-is\partial_{xx})F(s)\, ds \right\|_{H^s_x} = \sqrt{2\pi} \left\| \int_{\mathbb{R}} \langle\xi\rangle^s \tilde{F}(\tau, \xi)\widehat{\operatorname{sgn}\eta}(\tau - \xi^2)\, d\tau \right\|_{L^2_\xi}.$$

A simple integration by parts shows that

$$|\widehat{\operatorname{sgn}\eta}(\tau)| \leqslant C_\eta\langle\tau\rangle^{-1}. \tag{25}$$

Inserting this bound and then using (20), (24) we obtain

$$\left\| \int_{\mathbb{R}} \operatorname{sgn}(s)\eta(s) \exp(-is\partial_{xx}) F(s)\, ds \right\|_{H_x^s} \leqslant C_\eta \| \langle \tau - \xi^2 \rangle^{-1} \langle \xi \rangle^s |\tilde{F}| \|_W \leqslant C_\eta \|F\|_{\mathcal{N}^s}$$

as desired.

To establish the inhomogeneous estimate, observe that the spacetime Fourier transform of $\int_{\mathbb{R}} \operatorname{sgn}(t-s)\eta(t-s) \exp(-i(t-s)\partial_{xx}) F(s)\, ds$ at (τ, ξ) is simply $\widehat{\operatorname{sgn}\eta}(\tau - \xi)\tilde{F}(\tau, \xi)$. Using (25) and (24) we obtain

$$\left\| \int_{\mathbb{R}} \operatorname{sgn}(t-s)\eta(t-s) \exp(-i(t-s)\partial_{xx}) F(s)\, ds \right\|_{\mathcal{S}^s} \leqslant C_\eta \|F\|_{\mathcal{N}^s}$$

as desired.

Finally, we consider the non-linear estimate. Let us write $s = -1+\delta$. Taking Fourier transforms, we have

$$\|uv\|_{\mathcal{N}^{-1+\delta}} = \left\| \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\langle \xi \rangle^\delta}{\langle \tau - \xi^2 \rangle} \tilde{u}(\tau_1, \xi_1) \tilde{v}(\tau_2, \xi_2)\, d\xi_1\, d\tau_1 \right\|_W ,$$

where (τ_2, ξ_2) is defined by convention (23). We then estimate

$$\langle \xi \rangle^\delta \leqslant C_s (\langle \xi_1 \rangle^\delta + \langle \xi_2 \rangle^\delta)$$

and so by symmetry it would suffice to show that

$$\left\| \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\langle \xi_1 \rangle^\delta}{\langle \tau - \xi^2 \rangle} |\tilde{u}(\tau_1, \xi_1)| |\tilde{v}(\tau_2, \xi_2)|\, d\xi_1\, d\tau_1 \right\|_W \leqslant \|u\|_{\mathcal{S}^{-1+\delta}} \|v\|_{\mathcal{S}^{-1}}.$$

Writing $|\tilde{u}(\tau, \xi)| = \langle \xi \rangle^{-\delta} f(\tau, \xi)$ and $|\tilde{v}(\tau, \xi)| = g(\tau, \xi)$, the claim then follows from (22).

It remains to prove Proposition 3. For now, let us formulate an important consequence of convention (23) which is essential to the argument. Observe that (23) implies the algebraic identity

$$(\tau - \xi^2) = (\tau_1 - \xi_1^2) + (\tau_2 - \xi_2^2) - 2\xi_1\xi_2$$

and so by the triangle inequality we have the fundamental *resonance estimate*

$$\max(\langle \tau - \xi^2 \rangle, \langle \tau_1 - \xi_1^2 \rangle, \langle \tau_2 - \xi_2^2 \rangle) \geqslant 2^{-5} \langle \xi_1 \xi_2 \rangle. \tag{26}$$

(The constant 2^{-5} is very conservative, but its exact value is not important here.) Thus if both input frequencies ξ_1 and ξ_2 are large, then it is not possible for all three of (τ, ξ^2), (τ_1, ξ_1^2), and (τ_2, ξ_2^2) to lie close to the parabola.

6. Description of the space W

We are now ready to construct W and establish all the desired properties except for the bilinear estimate (22), which will be straightforward but will require some effort. As mentioned earlier, the space $\hat{X}^{-1,b}$ is the model candidate for W. However, we need to make three modifications to this space in order for it to be viable for us all the way down to the endpoint $s = -1$.

It will not be surprising that the geometry of the parabola $\tau = \xi^2$ plays a crucial role. We shall rely in particular on localizations to the spatial annuli

$$A_j := \{(\tau, \xi) \in \mathbb{R} \times \mathbb{R} : 2^j \leqslant \langle \xi \rangle < 2^{j+1}\}$$

for $j \geqslant 0$, as well as the parabolic neighborhoods

$$B_d := \{(\tau, \xi) \in \mathbb{R} \times \mathbb{R} : 2^d \leqslant \langle \tau - \xi^2 \rangle < 2^{d+1}\}$$

for $d \geqslant 0$. Thus the sets $A_j \cap B_d$ for $j, d \geqslant 0$ partition frequency space, and we have[9]

$$\|f\|_{\hat{X}^{s,b}} \approx \left(\sum_j \sum_d 2^{2sj} 2^{2bd} \|f\|^2_{L^2_\xi L^2_\tau (A_j \cap B_d)} \right)^{1/2}. \tag{27}$$

We also use the variant sets

$$A_{\leqslant j} := \bigcup_{j' \leqslant j} A_{j'}, \quad B_{\leqslant d} := \bigcup_{d' \leqslant d} B_{d'}$$

and similarly define $A_{\geqslant j}$, $A_{>j}$, $B_{\geqslant d}$, $B_{>d}$, etc.

A natural candidate for the space W is then the Besov endpoint $\hat{X}^{-1,1/2,1}$ of (27), defined by

$$\|f\|_{\hat{X}^{-1,1/2,1}} := \left(\sum_j 2^{-j} \left(\sum_d 2^{d/2} \|f\|_{L^2_\xi L^2_\tau (A_j \cap B_d)} \right)^2 \right)^{1/2}. \tag{28}$$

[9] All sums and unions involving j and d shall be over the non-negative integers unless otherwise mentioned.

This type of space has appeared previously in endpoint theory (see for instance [9]); we shall need this Besov refinement in order to handle the $s = -1$ endpoint without encountering logarithmic divergences (in particular, to handle the "parallel interaction" case when the non-linearity interacts with two components of the solution u with the same high frequency). The relationship between this space and the $\hat{X}^{s,b}$ spaces is provided by the following easy lemma.

Lemma 1. *Suppose that f is supported on $B_{\geqslant d}$ for some $d \geqslant 0$ (this condition is vacuous when $d = 0$). Then we have*

$$\|f\|_{\hat{X}^{-1,b}} \leqslant C_b 2^{-(1/2-b)d} \|f\|_{\hat{X}^{-1,1/2,1}}, \tag{29}$$

whenever $b < 1/2$, and

$$\|f\|_{\hat{X}^{-1,1/2,1}} \leqslant C_b 2^{-(b-1/2)d} \|f\|_{\hat{X}^{-1,b}}, \tag{30}$$

whenever $b > 1/2$.

Proof. We may easily restrict f to a single annulus A_j, since the general case then follows by square-summing. The claim then follows by decomposing further into $A_j \cap B_{d'}$ for $d' \geqslant d$ and using Cauchy–Schwarz. $\square$

By using $\hat{X}^{-1,1/2,1}$ instead of $\hat{X}^{-1,b}$, we will be able to handle parallel interactions. However, as essentially observed in [7], this Besov refinement is not sufficient by itself even to handle the endpoint $s = -\frac{3}{4}$, because of a divergence at the time axis $\tau = 0$. To handle these divergences we need a somewhat different norm $\| \|_Y$, defined as

$$\|f\|_Y := \|\langle\xi\rangle^{-1} f\|_{L^2_\xi L^1_\tau} + \|f\|_{L^2_\xi L^2_\tau} \tag{31}$$

and then form the sum space $Z := \hat{X}^{-1,1/2,1} + Y$ in the usual fashion,

$$\|f\|_Z := \inf\{\|f_1\|_{\hat{X}^{-1,1/2,1}} + \|f_2\|_Y : f_1 \in \hat{X}^{-1,1/2,1};\ f_2 \in Y;\ f = f_1 + f_2\}.$$

It is easy to verify that this is a Banach space, with the Schwartz functions being dense. Clearly we have

$$\|f\|_Z \leqslant \|f\|_{\hat{X}^{-1,1/2,1}}, \|f\|_Y$$

and conversely, to prove any linear estimate of the form $\|Tf\|_Z \leqslant M\|f\|_Z$ where Z is a Banach space and $M > 0$, it suffices to prove the separate estimates $\|Tf\|_Z \leqslant M \|f\|_{\hat{X}^{-1,1/2,1}}$, $\|Tf\|_Z \leqslant M\|f\|_Y$. For instance, we have the following basic estimates.

Proposition 4. *For any reasonable f, we have*

$$\|\langle\xi\rangle^{-1} f\|_{L^2_\xi L^1_\tau} \leqslant C\|f\|_Z. \tag{32}$$

If furthermore f is supported on $A_j \cap B_{\geqslant d}$ *for some* $j, d \geqslant 0$, *then we have*

$$\|f\|_{L^2_\xi L^2_\tau} \leqslant C(1 + 2^j 2^{-d/2})\|f\|_Z, \tag{33}$$

$$\|f\|_{L^1_\xi L^1_\tau} \leqslant C 2^{3j/2}\|f\|_Z, \tag{34}$$

$$\|f\|_{L^1_\xi L^2_\tau} \leqslant C(2^{j/2} + 2^{3j/2} 2^{-d/2})\|f\|_Z. \tag{35}$$

Proof. Observe that (32), (33) follow immediately from (31) if the right-hand sides were replaced by $C\|f\|_Y$, so it suffices to establish these estimates with the right-hand side of $C\|f\|_{\hat{X}^{-1,1/2,1}}$. For (33), this follows from (29). As for (34), we may reduce to a single annulus A_j (after square-summing in *j*) and reduce it to show that

$$\|f\|_{L^2_\xi L^1_\tau(A_j)} \leqslant C \sum_d 2^{d/2}\|f\|_{L^2_\xi L^2_\tau(A_j \cap B_d)}.$$

But this follows from the triangle inequality and Hölder's inequality, since for each fixed ξ and fixed B_d, the τ variable varies over a set of measure $O(2^d)$.

Finally, (34), (35) follow, respectively, from (32), (33) and Hölder in the ξ variable (which varies over a set of measure $O(2^j)$). $\square$

Observe that $\hat{X}^{-1,1/2,1}$ and Y are both monotonous in the sense of Proposition 3. Also the two spaces $\hat{X}^{-1,1/2,1}$ and Y paste together well along the fuzzy boundary $\langle\tau - \xi^2\rangle \approx \langle\xi\rangle^2$. A formalization of this heuristic is as follows:

Lemma 2 (*Pasting lemma*). *Let f be a reasonable function. If f is supported on* $\bigcup_j A_j \cap B_{\geqslant 2j-100}$, *then*

$$\|f\|_Y \leqslant C\|f\|_Z. \tag{36}$$

Conversely, if f is supported on $\bigcup_j A_j \cap B_{\leqslant 2j+100}$, *then*

$$\|f\|_{\hat{X}^{-1,1/2,1}} \leqslant C\|f\|_Z. \tag{37}$$

Proof. Let us first establish (36). It clearly suffices to show that

$$\|f\|_Y \leqslant C\|f\|_{\hat{X}^{-1,1/2,1}}$$

on this domain. Partitioning dyadically into the A_j, and then square-summing in j, it suffices to show that

$$\|f 1_{A_j}\|_Y \leqslant C\|f 1_{A_j}\|_{\hat{X}^{-1,1/2,1}}$$

for each j. From (29) we already have

$$\|f 1_{A_j}\|_{L^2_\xi L^2_\tau} \leqslant C\|f 1_{A_j}\|_{\hat{X}^{-1,1/2,1}},$$

while from (32) we have

$$\|\langle\xi\rangle^{-1} f 1_{A_j}\|_{L^2_\xi L^1_\tau} \leqslant C\|f 1_{A_j}\|_{\hat{X}^{-1,1/2,1}}$$

and the claim follows.

Now we establish (37). By arguing as before it suffices to show that

$$\|f 1_{A_j}\|_{\hat{X}^{-1,1/2,1}} \leqslant C\|f 1_{A_j}\|_Y.$$

But we have

$$\begin{aligned}\|f 1_{A_j}\|_{\hat{X}^{-1,1/2,1}} &\leqslant C \sum_{d \leqslant 2j+100} 2^{-j} 2^{d/2} \|f\|_{L^2_\xi L^2_\tau(A_j \cap B_d)} \\ &\leqslant C\|f\|_{L^2_\xi L^2_\tau(A_j)} \\ &\leqslant C\|f 1_{A_j}\|_Y\end{aligned}$$

as desired. $\square$

The space Z is a candidate for W, as it is able to cope with two of the dangerous quadratic interactions in the equation (namely the parallel interactions, and the interactions which output near the time axis). However, there is a third type of interaction which could cause trouble, when a solution component near the parabola $\{\tau = \xi^2\}$ interacts with a component near the reflected parabola $\{\tau = -\xi^2\}$ to create a large contribution near the frequency origin. The use of the space Z does not prohibit either component from occurring. However, as mentioned in the introduction, the solution should remain in the upper half-plane $\tau > 0$. To exploit this we shall introduce a weight

$$w(\tau, \xi) := \min(1, -\tau)^{10} \tag{38}$$

to localize to the upper half-plane, and define W to be the space

$$\|f\|_W := \|wf\|_Z \tag{39}$$

as discussed in the previous section.

The monotonicity of W is clear. Claim (20) follows immediately from (32) (since $w \geqslant 1$), while (21) follows from (30):

$$\|f\|_W \leqslant \|wf\|_{\hat{X}^{-1,1/2,1}} \leqslant C\|wf\|_{\hat{X}^{-1,90}} \leqslant C\|f\|_{\hat{X}^{-1,100}},$$

where we use the crude estimate $w(\tau,\xi) \leqslant C\langle \tau - \xi^2\rangle^{10}$.

It remains to show (22). Applying (39) and monotonicity, were reduced to showing that

$$\left\| \frac{w}{\langle \tau - \xi^2\rangle}\left(\frac{f}{w} * \frac{g}{w}\right)\right\|_Z \leqslant C\|f\|_Z\|g\|_Z \tag{40}$$

for all non-negative reasonable f, g.

Fix f, g. Observe from (23) that

$$w(\tau,\xi) \leqslant C w(\tau_1,\xi_1) w(\tau_2,\xi_2),$$

this basically reflects the fact that in order for $\tau = \tau_1 + \tau_2$ to be negative, at least one of τ_1, τ_2 has to be even more negative. This gives us the very handy pointwise estimate

$$\frac{w}{\langle \tau - \xi^2\rangle}\left(\frac{f}{w} * \frac{g}{w}\right) \leqslant \frac{C}{\langle \tau - \xi^2\rangle}(f * g), \tag{41}$$

which we shall rely upon in most cases (except for one special high–high interaction where we must utilize the localizing weight w more carefully). As one example of this, we present a simple case of (40):

Lemma 3. *For any non-negative reasonable f, g, we have*

$$\left\| \frac{w}{\langle \tau - \xi^2\rangle}\left(\frac{f}{w} * \frac{g}{w}\right)\right\|_Z \leqslant C\|f\|_Y\|g\|_Y.$$

Proof. Write $h := f * g$. From Young's inequality we have

$$\|h\|_{L^\infty_\tau L^\infty_\xi} \leqslant \|f\|_{L^2_\tau L^2_\xi}\|g\|_{L^2_\tau L^2_\xi} \leqslant \|f\|_Y\|g\|_Y$$

so by (41) and monotonicity it suffices to show that

$$\left\| \frac{1}{\langle \tau - \xi^2\rangle}\right\|_{\hat{X}^{-1,1/2,1}} \leqslant C.$$

But this is easily verified. $\square$

From this lemma, Lemma 2, and monotonicity, we thus see that to prove (40) we may thus assume that at least one of non-negative reasonable f, g lies near the parabola, or more precisely

$$\text{At least one of } f, g \text{ is supported in } \bigcup_j A_j \cap B_{<2j-100}. \tag{42}$$

The next step is dyadic decomposition. Observe the localization property

$$\|f\|_Z \sim \left(\sum_j \|1_{A_j} f\|_Z^2\right)^{1/2}, \tag{43}$$

which follows from the L^2_ξ nature of both $\hat{X}^{-1,1/2,1}$ and Y. We therefore split $f = \sum_{j_1} f_{j_1}$ and $g = \sum_{j_2} g_{j_2}$, where f_{j_1} and g_{j_2} are the, restrictions of f, g to A_{j_1}, A_{j_2}, respectively. Thus

$$\begin{aligned}\left\|\frac{w}{\langle\tau-\xi^2\rangle}\left(\frac{f}{w}*\frac{g}{w}\right)\right\|_Z &\leqslant C\left(\sum_j\left\|1_{A_j}\frac{w}{\langle\tau-\xi^2\rangle}\left(\frac{f}{w}*\frac{g}{w}\right)\right\|_Z^2\right)^{1/2}\\ &= C\left(\sum_j\left\|\sum_{j_1,j_2}1_{A_j}\frac{w}{\langle\tau-\xi^2\rangle}\left(\frac{f_{j_1}}{w}*\frac{g_{j_2}}{w}\right)\right\|_Z^2\right)^{1/2}.\end{aligned}$$

In order for the inner summand to be non-zero, it must be possible to find $(\tau, \xi) \in A_j$, $(\tau_1, \xi_1) \in A_{j_1}$, $(\tau_2, \xi_2) \in A_{j_2}$ obeying (23). This forces one of the following (overlapping) cases to hold:

- (High–low interaction). $|j - j_1| \leqslant 10$ (which implies $j_2 \leqslant j + 11$);
- (Low–high interaction). $|j - j_2| \leqslant 10$ (which implies $j_1 \leqslant j + 11$);
- (High–high interaction). $j < j_1 - 10, j_2 - 10$ (which implies $|j_1 - j_2| \leqslant 1$).

The former two cases are symmetric. Thus to prove (40) it suffices (again using (43)) to verify the high–low estimate

$$\begin{aligned}&\left(\sum_j\left\|\sum_{|j_1-j|\leqslant 10}\sum_{j_2\leqslant j+11}1_{A_j}\frac{w}{\langle\tau-\xi^2\rangle}\left(\frac{f_{j_1}}{w}*\frac{g_{j_2}}{w}\right)\right\|_Z^2\right)^{1/2}\\ &\quad\leqslant C\left(\sum_{j_1}\|f_{j_1}\|_Z^2\right)^{1/2}\left(\sum_{j_2}\|g_{j_2}\|_Z^2\right)^{1/2}\end{aligned}$$

and the high–high estimate

$$\left(\sum_j \left\| \sum_{j_1,j_2>j+10:|j_1-j_2|\leqslant 1} 1_{A_j} \frac{w}{\langle \tau-\xi^2\rangle}\left(\frac{f_{j_1}}{w} * \frac{g_{j_2}}{w}\right)\right\|_Z^2\right)^{1/2}$$

$$\leqslant C\left(\sum_{j_1}\|f_{j_1}\|_Z^2\right)^{1/2}\left(\sum_{j_2}\|g_{j_2}\|_Z^2\right)^{1/2}. \tag{44}$$

Consider the high–low estimate first. We use (41) to drop the weights w. Since for any j there are only $O(1)$ values of j_1 which contribute, we can use Schur's test and reduce to showing

$$\left\| \sum_{j_2\leqslant j+11} \frac{1_{A_j}}{\langle \tau-\xi^2\rangle}(f_{j_1} * g_{j_2})\right\|_Z \leqslant C\|f_{j_1}\|_Z \left(\sum_{j_2}\|g_{j_2}\|_Z^2\right)^{1/2},$$

whenever $|j_1-j|\leqslant 10$. By the triangle inequality, and estimating the Z norm by the $\hat{X}^{-1,1/2,1}$ norm, it thus suffices to establish

$$\left\|\frac{1_{A_j}}{\langle \tau-\xi^2\rangle}(f_{j_1} * g_{j_2})\right\|_{\hat{X}^{-1,1/2,1}} \leqslant C(2^{-j_2/10}+2^{-(j-j_2)/10})\|f_{j_1}\|_Z\|g_{j_2}\|_Z, \tag{45}$$

whenever $|j_1-j|\leqslant 10$ and $j_2\leqslant j+11$.

We shall prove (45) in Section 8. We leave this for now and turn to the high–high estimate (44). Here we cannot afford to discard the weights w. By the triangle inequality in l^2, we can bound the left-hand side of (44) by

$$\sum_{j_1,j_2:|j_1-j_2|\leqslant 1}\left(\sum_{j<j_1-10,j_2-10}\left\|1_{A_j}\frac{w}{\langle \tau-\xi^2\rangle}\left(\frac{f_{j_1}}{w} * \frac{g_{j_2}}{w}\right)\right\|_Z^2\right)^{1/2}.$$

By Schur's test again, it thus suffices to show that

$$\left(\sum_{j<j_1-10,j_2-10}\left\|1_{A_j}\frac{w}{\langle \tau-\xi^2\rangle}\left(\frac{f_{j_1}}{w} * \frac{g_{j_2}}{w}\right)\right\|_Z^2\right)^{1/2} \leqslant C\|f_{j_1}\|_Z\|g_{j_2}\|_Z,$$

whenever $|j_1 - j_2| \leqslant 1$. Applying (43) once more, and estimating the Z norm by the Y norm,[10] we can simplify this slightly as

$$\left\| 1_{A_{\leqslant j_1 - 9}} \frac{w}{\langle \tau - \xi^2 \rangle} \left(\frac{f_{j_1}}{w} * \frac{g_{j_2}}{w} \right) \right\|_Y \leqslant C \| f_{j_1} \|_Z \| g_{j_2} \|_Z. \tag{46}$$

This estimate shall be proven in Section 9.

Thus to conclude the proof of Proposition 3 (and hence Proposition 2) it suffices to prove (45) and (46). In many cases (basically, when at least two of $f, g, f * g$ are far from the parabola), these inequalities can be established through Young's inequality, Proposition 4, and the resonance estimate (26). However, when two of $f, g, f * g$ are close to the parabola we need a further (standard) bilinear estimate, to which we now turn.

7. Bilinear estimates near the parabola

We give a standard bilinear estimate.

Proposition 5 (*Bilinear estimate*). *Let f, g be test functions supported on A_{j_1} and A_{j_2}, respectively. Suppose also that there is $D \geqslant 0$ such that $|\xi_1 - \xi_2| \geqslant D$ whenever (τ_1, ξ_1) lies in the support of f and (τ_2, ξ_2) lies in the support of g (this hypothesis is vacuous if $D = 0$). Then*

$$\| f * g \|_{L^2_\xi L^2_\tau} \leqslant C 2^{j_1 + j_2} \langle D \rangle^{-1/2} \| f \|_{\hat{X}^{-1,1/2,1}} \| g \|_{\hat{X}^{-1,1/2,1}}.$$

Proof. Let f_{d_1} be the restriction of f to B_{d_1}, and similarly let g_{d_2} be the restriction of g to B_{d_2}. By (28) we have

$$\| f \|_{\hat{X}^{-1,1/2,1}} = 2^{-j_1} \sum_{d_1 \geqslant 0} 2^{d_1/2} \| f_{d_1} \|_{L^2_{\tau_1} L^2_{\xi_1}}$$

$$\| g \|_{\hat{X}^{-1,1/2,1}} = 2^{-j_2} \sum_{d_2 \geqslant 0} 2^{d_2/2} \| g_{d_2} \|_{L^2_{\tau_2} L^2_{\xi_2}}$$

and so by the triangle inequality it suffices to show that

$$\| f_{d_1} * g_{d_2} \|_{L^2(\mathbb{R} \times \mathbb{R})} \leqslant C 2^{(d_1 + d_2)/2} (2^{d_1/2} + 2^{d_2/2} + D)^{-1/2} \| f_{d_1} \|_{L^2_{\tau_1} L^2_{\xi_1}} \| g_{d_2} \|_{L^2_{\tau_2} L^2_{\xi_2}} \tag{47}$$

for each $d_1, d_2 \geqslant 0$. Note that this is slightly stronger than we need, as $(2^{d_1/2} + 2^{d_2/2} + D)^{-1/2}$ is better than $\langle D \rangle^{-1/2}$, but we shall use this improvement in Corollary 1 below.

[10] This reflects the fact that it is very difficult for the high–high interaction to return to the parabola $\tau = \xi^2$, especially given our use of the weight w to localize to the upper half-plane $\tau > 0$.

Fix d_1, d_2; we may take $d_1 \geqslant d_2$ by symmetry. From Cauchy–Schwarz we have

$$\begin{aligned}
&\| f_{d_1} * g_{d_2} \|^2_{L^2(\mathbb{R}\times\mathbb{R})} \\
&\quad = \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\int_{B_{d_1} \cap ((\tau,\xi)-B_{d_2}); |\xi_1-\xi_2| \geqslant D} f_{d_1}(\tau_1, \xi_1) g_{d_2}(\tau_2, \xi_2)\, d\tau_1\, d\xi_1 \right)^2 d\tau\, d\xi \\
&\quad \leqslant \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} f_{d_1}(\tau_1, \xi_1)^2 g_{d_2}(\tau_2, \xi_2)^2\, d\tau_1\, d\xi_1\, d\tau\, d\xi \\
&\qquad \times \sup_{\tau,\xi} |\{(\tau_1, \xi_1) \in B_{d_1} \cap ((\tau, \xi) - B_{d_2}) : |\xi_1 - \xi_2| \geqslant D\}| \\
&\quad = \| f_{d_1} \|^2_{L^2_{\tau_1} L^2_{\xi_1}} \| g_{d_2} \|^2_{L^2_{\tau_2} L^2_{\xi_2}} \\
&\qquad \times \sup_{\tau,\xi} |\{(\tau_1, \xi_1) \in B_{d_1} \cap ((\tau, \xi) - B_{d_2}) : |\xi_1 - \xi_2| \geqslant D\}|,
\end{aligned}$$

where we use convention (23). Thus it suffices to show that

$$|\{(\tau_1, \xi_1) \in B_{d_1} \cap ((\tau, \xi) - B_{d_2}) : |\xi_1 - \xi_2| \geqslant D\}| \leqslant C 2^{d_1+d_2} / (2^{d_1/2} + D).$$

Observe that if (τ_1, ξ_1) lies in the above set, then $\tau_1 = \xi_1^2 + O(2^{d_1})$, $\tau_2 = \xi_2^2 + O(2^{d_2})$, and thus $\tau = \xi_1^2 + \xi_2^2 + O(2^{d_1})$. From the parallelogram identity

$$\xi_1^2 + \xi_2^2 = \tfrac{1}{2}(\xi^2 + (\xi_1 - \xi_2)^2),$$

we thus have

$$(\xi_1 - \xi_2)^2 = 2\tau - \xi^2 + O(2^{d_1}).$$

On the other hand, we have $|\xi_1 - \xi_2| \geqslant D$. Elementary algebra then shows that $\xi_1 - \xi_2$ is confined to a set of measure at most $O(2^{d_1}/(2^{d_1/2} + D))$. Thus, ξ_1 is also confined to a set of a similar measure. For fixed ξ_1 and ξ_2, τ_2 (and hence τ_1) is confined to an interval of length $O(2^{d_2})$. The claim then follows from Fubini's theorem. $\square$

We can dualize this to obtain

Corollary 1 (*Dual bilinear estimate*). *Let $D \geqslant 0$, and let $\Omega_1 \subseteq A_{j_1}$, $\Omega \subseteq A_j$ be regions such that $|\xi_1 + \xi| \geqslant D$ whenever $(\tau_1, \xi_1) \in \Omega_1$ and $(\tau, \xi) \in \Omega$. Then for any f supported in Ω_1, any test function g, and any $d \geqslant 0$, we have*

$$2^{-d/2} \| f * g \|_{L^2_\xi L^2_\tau(\Omega \cap B_d)} \leqslant C 2^{j_1} (2^{d/2} + D)^{-1/2} \| f \|_{\hat{X}^{-1,1/2,1}} \| g \|_{L^2_\xi L^2_\tau}.$$

Proof. We can take f, g to be non-negative. By duality we can write

$$2^{-d/2}\|f * g\|_{L^2(\Omega\cap B_d)} = \int_{\mathbb{R}\times\mathbb{R}} f * g(\tau,\xi)h(\tau,\xi)\,d\tau\,d\xi$$

for some non-negative h supported in $\Omega \cap B_d$ with $\|h\|_{L^2(\Omega\cap B_d)} = 2^{-d/2}$. We can then use the Fubini–Tonelli theorem, Cauchy–Schwarz, and Fubini–Tonelli again to write

$$\begin{aligned}
&\int_{\mathbb{R}\times\mathbb{R}} f * g(\tau,\xi)h(\tau,\xi)\,d\tau\,d\xi \\
&\quad = \int_{\mathbb{R}\times\mathbb{R}} \left(\int_{\mathbb{R}\times\mathbb{R}} f(\tau_1,\xi_1)h(\tau_1+\tau_2,\xi_1+\xi_2)\,d\tau_1\,d\xi_1\right) g(\tau_2,\xi_2)\,d\tau_2\,d\xi_2 \\
&\quad \leqslant \left(\int_{\mathbb{R}\times\mathbb{R}} \left(\int_{\mathbb{R}\times\mathbb{R}} f(\tau_1,\xi_1)h(\tau_1+\tau_2,\xi_1+\xi_2)\,d\tau_1\,d\xi_1\right)^2 d\tau_2\,d\xi_2\right)^{1/2} \\
&\qquad \times \|g\|_{L^2_\tau L^2_\xi(\mathbb{R}\times\mathbb{R})} \\
&\quad = \|f_- * h\|_{L^2(\mathbb{R}\times\mathbb{R})}\|g\|_{L^2_\tau L^2_\xi(\mathbb{R}\times\mathbb{R})},
\end{aligned}$$

where f_- is the reflection of f. On the other hand, by decomposing $f = \sum_{d_1} f_{d_1}$, where each f_{d_1} is supported on B_{d_1}, and using (47), (28) we have

$$\begin{aligned}
\|f_- * h\|_{L^2(\mathbb{R}\times\mathbb{R})} &\leqslant \sum_{d_1} 2^{(d_1+d)/2}\left(2^{d_1/2} + 2^{d/2} + D\right)^{-1/2}\|f_{d_1}\|_{L^2_{\tau_1} L^2_{\xi_1}} 2^{-d/2} \\
&\leqslant \left(2^{d/2} + D\right)^{-1/2} 2^{j_1}\|f\|_X
\end{aligned}$$

and the claim follows. □

8. High–low interactions

We now prove the high–low interaction estimate (45). Recall that we have $|j_1 - j| \leqslant 10$, $j_2 \leqslant j + 11$.

Let us first dispose of the easy case $j_2 = 0$. In this case we use (30) followed by Young's inequality and Proposition 4 to estimate

$$\begin{aligned}
\left\|\frac{1_{A_j}}{\langle\tau-\xi^2\rangle}(f_{j_1} * g_0)\right\|_{\hat{X}^{-1,1/2,1}} &\leqslant C2^{-j_1}\|f_{j_1} * g_0\|_{L^2_\xi L^2_\tau} \\
&\leqslant C2^{-j_1}\|f_{j_1}\|_{L^2_\xi L^2_\tau}\|g_0\|_{L^1_\xi L^1_\tau} \\
&\leqslant C\|f_{j_1}\|_Z\|g_0\|_Z,
\end{aligned}$$

which is acceptable. Thus, we may restrict our attention to the case $j_2 > 0$. Applying the resonance estimate (26), we obtain

$$\max(\langle \tau - \xi^2 \rangle, \langle \tau_1 - \xi_1^2 \rangle, \langle \tau_2 - \xi_2^2 \rangle) \geqslant 2^{-20} 2^{j+j_2}.$$

Thus we may restrict one of f_{j_1}, g_{j_2}, or $f_{j_1} * g_{j_2}$ to the region $B_{\geqslant j+j_2-20}$.

Let us first consider the case when the high-frequency input f_{j_1} is restricted to the region $B_{\geqslant j+j_2-20}$. We can split this case into two sub-cases, depending on whether we measure g_{j_2} using $\hat{X}^{-1,1/2,1}$ or using Y. If we use Y, then we use Hölder's inequality in τ, followed by Young's inequality and Proposition 4, (31) to conclude that

$$\begin{aligned}
\left\| \frac{1_{A_j}}{\langle \tau - \xi^2 \rangle} (f_{j_1} * g_{j_2}) \right\|_{\hat{X}^{-1,1/2,1}} &\leqslant C 2^{-j} \| f_{j_1} * g_{j_2} \|_{L^2_\xi L^\infty_\tau} \\
&\leqslant C 2^{-j} \| f_{j_1} \|_{L^2_\xi L^2_\tau} \| g_{j_2} \|_{L^1_\xi L^2_\tau} \\
&\leqslant C 2^{-j} 2^j 2^{-(j+j_2)/2} \| f_{j_1} \|_Z 2^{j_2/2} \| g_{j_2} \|_Y \\
&\leqslant C 2^{-j_2/10} \| f_{j_1} \|_Z \| g_{j_2} \|_Y,
\end{aligned}$$

which is acceptable. If we instead measure g_{j_2} using $\hat{X}^{-1,1/2,1}$, we decompose into the regions B_d and use Corollary 1 with $D = 0$ (and with f and g swapped) followed by Proposition 4 to estimate[11]

$$\begin{aligned}
\left\| \frac{1_{A_j}}{\langle \tau - \xi^2 \rangle} (f_{j_1} * g_{j_2}) \right\|_{\hat{X}^{-1,1/2,1}} &\leqslant C \sum_d 2^{-j} 2^{-d/2} \| f_{j_1} * g_{j_2} \|_{L^2_\xi L^2_\tau (A_j \cap B_d)} \\
&\leqslant C \sum_d 2^{-j} 2^{j_2} 2^{-d/4} \| g_{j_2} \|_{\hat{X}^{-1,1/2,1}} \| f_{j_1} \|_{L^2_\tau L^2_\xi} \\
&\leqslant C 2^{-j} 2^{j_2} \| g_{j_2} \|_{\hat{X}^{-1,1/2,1}} 2^{j_1} 2^{-(j+j_2)/2} \| f_{j_1} \|_Z \\
&\leqslant 2^{-(j-j_2)/10} \| f_{j_1} \|_Z \| g_{j_2} \|_{\hat{X}^{-1,1/2,1}},
\end{aligned}$$

which is acceptable. Thus, we may now restrict f_{j_1} to the region $B_{<j+j_2-20}$. From Lemma 2 we may now measure f_{j_1} in $\hat{X}^{-1,1/2,1}$ instead of Z.

We next consider the case when g_{j_2} is restricted to the region $B_{\geqslant j+j_2-20}$. We subdivide the domain $A_{j_1} \cap B_{<j+j_2-20}$ of f_{j_1} into disjoint slabs, where on each slab the frequency variable ξ_1 is localized to an interval I of length $2^{j_2}/100$, and write $f_{j_1} = \sum_I f_{j_1,I}$ accordingly. Because g_{j_2} is localized to A_{j_2}, we see that the functions

[11] Note that if j_2 is substantially smaller than j then we can take D as large as 2^j, which leads to much better estimates. However, the $j = j_2$ case contains the delicate "parallel interaction" in which no gain occurs.

$f_{j_1,I} * g_{j_2}$ have a finite overlap in the ξ variable. Thus by square-summing in I it would suffice to establish the estimate

$$\left\| \frac{1_{A_j}}{\langle \tau - \xi^2 \rangle} (f_{j_1,I} * g_{j_2}) \right\|_{\hat{X}^{-1,1/2,1}} \leqslant C(2^{-j_2/10} + 2^{-(j-j_2)/10}) \| f_{j_1,I} \|_{\hat{X}^{-1,1/2,1}} \| g_{j_2} \|_Z .$$

But from dyadic decomposition into B_d regions and Corollary 1 (which now applies with $D = 2^{j_2}/10$, say) followed by Proposition 4 we have

$$\begin{aligned}\left\| \frac{1_{A_j}}{\langle \tau - \xi^2 \rangle} (f_{j_1,I} * g_{j_2}) \right\|_{\hat{X}^{-1,1/2,1}} &\leqslant C \sum_d 2^{-j} 2^{-d/2} \| f_{j_1,I} * g_{j_2} \|_{L^2_\xi L^2_\tau (A_j \cap B_d)} \\ &\leqslant C \sum_d 2^{-j} 2^{j_1} (2^{d/2} + 2^{j_2})^{-1/2} \| f_{j_1,I} \|_{\hat{X}^{-1,1/2,1}} \| g_{j_2} \|_{L^2_\xi L^2_\tau} \\ &\leqslant C j_2 2^{-j_2/2} \| f_{j_1,I} \|_{\hat{X}^{-1,1/2,1}} \| g_{j_2} \|_Z ,\end{aligned}$$

which is acceptable.

The only remaining case is when we restrict the output $f_{j_1} * g_{j_2}$ to the region $B_{\geqslant j+j_2-20}$. By Proposition 4, it then suffices to show that

$$2^{-j} 2^{-(j+j_2)/2} \| f_{j_1} * g_{j_2} \|_{L^2_\xi L^2_\tau} \leqslant C(2^{-j_2/10} + 2^{-(j-j_2)/10}) \| f_{j_1} \|_{\hat{X}^{-1,1/2,1}} \| g_{j_2} \|_Z .$$

If we measure g_{j_2} in $\hat{X}^{-1,1/2,1}$, then from Proposition 5 we have

$$\begin{aligned}2^{-j} 2^{-(j+j_2)/2} \| f_{j_1} * g_{j_2} \|_{L^2_\xi L^2_\tau} &\leqslant C 2^{-j} 2^{-(j+j_2)/2} 2^{j_1} 2^{j_2} \| f_{j_1} \|_{\hat{X}^{-1,1/2,1}} \| g_{j_2} \|_{\hat{X}^{-1,1/2,1}} \\ &\leqslant C 2^{-(j-j_2)/10} \| f_{j_1} \|_{\hat{X}^{-1,1/2,1}} \| g_{j_2} \|_{\hat{X}^{-1,1/2,1}} ,\end{aligned}$$

which is acceptable. On the other hand, if we measure g_{j_2} in Y, then from Young's inequality, Proposition 4, and (31) we have

$$\begin{aligned}2^{-j} 2^{-(j+j_2)/2} \| f_{j_1} * g_{j_2} \|_{L^2_\xi L^2_\tau} &\leqslant C 2^{-j} 2^{-(j+j_2)/2} \| f_{j_1} \|_{L^2_\xi L^1_\tau} \| g_{j_2} \|_{L^1_\xi L^2_\tau} \\ &\leqslant C 2^{-j} 2^{-(j+j_2)/2} 2^{j_1} \| f_{j_1} \|_{\hat{X}^{-1,1/2,1}} 2^{j_2/2} \| g_{j_2} \|_Y \\ &\leqslant C 2^{-j_2/10} \| f_{j_1} \|_{\hat{X}^{-1,1/2,1}} ,\end{aligned}$$

which is also acceptable. This concludes the proof of (45).

9. High–high interactions

We now prove the high–high estimate (46). Recall that $|j_1 - j_2| \leqslant 1$ and we are operating under assumption (42). We may also assume $j_1 \geqslant 9$ since the claim is vacuous otherwise. We can assume that f_{j_1} is supported on one half-space, say $\{\xi > 0\}$, and that g_{j_2} is supported on the other half-space $\{\xi < 0\}$, since if they are both supported by the same half-space then their convolution will not intersect $A_{\leqslant j_1-9}$. In particular, we can ensure that the ξ-supports of f_{j_1} and g_{j_2} are separated by at least $2^{j_1}/10$.

We need some preliminary convolution estimates. From Proposition 5 (with $D = 2^{j_1}/10$) we have

$$\|f_{j_1} * g_{j_2}\|_{L^2_\xi L^2_\tau} \leqslant C 2^{j_1} 2^{j_2} 2^{-j_1/2} \|f_{j_1}\|_{\hat{X}^{-1,1/2,1}} \|g_{j_2}\|_{\hat{X}^{-1,1/2,1}},$$

while from Young's inequality and Proposition 4

$$\begin{aligned}\|f_{j_1} * g_{j_2}\|_{L^2_\xi L^2_\tau} &\leqslant \|f_{j_1}\|_{L^2_\xi L^2_\tau} \|g_{j_2}\|_{L^1_\xi L^1_\tau} \\ &\leqslant \|f_{j_1}\|_Y 2^{3j_2/2} \|g_{j_2}\|_Z\end{aligned}$$

and similarly

$$\|f_{j_1} * g_{j_2}\|_{L^2_\xi L^2_\tau} \leqslant 2^{-2j_1} 2^{3j_1/2} \|f_{j_1}\|_Z \|g_{j_2}\|_Y.$$

Putting all these estimates together, we obtain

$$\|f_{j_1} * g_{j_2}\|_{L^2_\xi L^2_\tau} \leqslant C 2^{2j_1} \|f_{j_1}\|_Z \|g_{j_2}\|_Z.$$

In a similar spirit, from Hölder's inequality, Young's inequality and Proposition 4 we have

$$\begin{aligned}\|\langle\xi\rangle^{-1} f_{j_1} * g_{j_2}\|_{L^2_\xi L^1_\tau} &\leqslant C \|f_{j_1} * g_{j_2}\|_{L^\infty_\xi L^1_\tau} \\ &\leqslant C \|f_{j_1}\|_{L^2_\xi L^1_\tau} \|g_{j_2}\|_{L^2_\xi L^1_\tau} \\ &\leqslant C 2^{j_1} \|f_{j_1}\|_Z 2^{j_2} \|g_{j_2}\|_Z,\end{aligned}$$

combining these estimates using (31) we obtain

$$\|f_{j_1} * g_{j_2}\|_Y \leqslant C 2^{2j_1} \|f_{j_1}\|_Z \|g_{j_2}\|_Z. \tag{48}$$

We now return to (46). First let us restrict $A_{\leqslant j_1-9}$ to the region $A_{\leqslant j_1-9} \cap B_{\geqslant 2j_1-10}$. In this case we discard the weights w to obtain

$$\left\| 1_{A_{\leqslant j_1-9} \cap B_{\geqslant 2j_1-10}} \frac{w}{\langle \tau - \xi^2 \rangle} \left(\frac{f_{j_1}}{w} * \frac{g_{j_2}}{w} \right) \right\|_Y \leqslant C 2^{-2j_1} \|f_{j_1} * g_{j_2}\|_Y \tag{49}$$

and (46) in this case follows from (48).

Thus it remains to consider the contribution in the domain $A_{\leqslant j_1-9} \cap B_{<2j_1-10}$ (i.e. the contribution near the frequency origin). We now finally invoke (42); we shall assume that f_{j_1} lies in $B_{\leqslant 2j_1-100}$ since the other case is almost identical (recall that $|j_1 - j_2| \leqslant 1$). In particular, we have $\tau_1 \geqslant 2^{2j_1}/10$ and $|\tau| \leqslant 2^{2j_1}/100$, which forces $\tau_2 \leqslant -2^{2j_1}/20$ by (23). Thus we now have a large weight on g_{j_2}: $w(\tau_2, \xi_2) \geqslant c2^{20j_1}$. On the other hand, we make the elementary observation that

$$\frac{w(\tau, \xi)}{\langle \tau - \xi^2 \rangle} \leqslant C2^{-2j_1}2^{20j_1}$$

for (τ, ξ) in $A_{\leqslant j_1-9} \cap B_{<2j_1-10}$. Thus, in this case we again have (49), and again (46) in this case follows from (48). This concludes the proof of (46), and thus of Propositions 3 and 2.

References

[1] B. Birnir, C. Kenig, G. Ponce, N. Svanstedt, L. Vega, On the ill-posedness of the IVP for the generalized Korteweg-de Vries and nonlinear Schrödinger equations, J. London Math. Soc. 53 (2) (1996) 551–559.

[2] J. Bourgain, Fourier restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations, Part I, Geometric and Funct. Anal. 3 (1993) 107–156.

[3] T. Cazenave, F.B. Weissler, The Cauchy problem for the critical nonlinear Schrödinger equation, Non. Anal. TMA 14 (1990) 807–836.

[4] M. Christ, J. Colliander, T. Tao, Low-regularity ill-posedness for nonlinear Schrödinger and wave equations, preprint.

[5] C. Kenig, G. Ponce, L. Vega, Quadratic forms for the 1-D semilinear Schrödinger equation, Trans. Amer. Math. Soc. 346 (1996) 3323–3353.

[6] C. Kenig, G. Ponce, L. Vega, On the ill-posedness of some canonical dispersive equations, Duke Math. J. 106 (2001) 617–633.

[7] T. Muramatu, S. Taoka, The initial value problem for the 1-D semilinear Schrödinger equation in Besov spaces, J. Math. Soc. Japan 56 (3) (2004) 853–888.

[8] K. Nakanishi, H. Takaoka, Y. Tsutsumi, Counterexamples to bilinear estimates related with the *KdV* equation and the nonlinear Schrödinger equation, Methods Appl. Anal. 8 (2001) 569–578.

[9] D. Tataru, Local and global results for wave maps *I*, Comm. Partial Differential Equations 23 (1998) 1781–1793.

[10] Y. Tsutsumi, L^2 solutions for nonlinear Schrodinger equations and nonlinear groups, Funk. Ekva. 30 (1987) 115–125.