J. Eur. Math. Soc. 10, 957–985

Ioan Bejenaru · Daniel Tataru

Large data local solutions for the derivative NLS equation

Received May 29, 2007

Abstract. We consider the derivative NLS equation with general quadratic nonlinearities. In [2] the first author has proved a sharp small data local well-posedness result in Sobolev spaces with a decay structure at infinity in dimension $n = 2$. Here we prove a similar result for large initial data in all dimensions $n \geq 2$.

1. Introduction

The general Cauchy problem for the semilinear Schrödinger equation has the form

$$\begin{cases} iu_t - \Delta u = P(u, \bar{u}, \nabla u, \nabla \bar{u}), & t \in \mathbb{R},\ x \in \mathbb{R}^n, \\ u(x, 0) = u_0(x), \end{cases} \tag{1}$$

where $u : \mathbb{R}^n \times \mathbb{R} \to \mathbb{C}^m$ and $P : \mathbb{C}^{2n+2} \to \mathbb{C}^m$. Assuming that P is a polynomial containing terms of order at least $\kappa \geq 2$ and higher, one is interested in studying the local well-posedness for this evolution in a suitable Sobolev space.

In the simpler case when P does not depend on ∇u, $\nabla \bar{u}$ this problem is rather well understood, and the Strichartz estimates for the linear Schrödinger equation play a proeminent role.

Here we are interested in nonlinearities which contain derivatives. This problem was considered in full generality in the work of Kenig–Ponce–Vega [8]. Due to the need to regain one derivative in multilinear estimates, they use in an essential fashion the local smoothing estimates for the linear Schrödinger equation. Their results make it clear that one needs to differentiate two cases.

If $\kappa \geq 3$ then they prove local well-posedness for initial data in a Sobolev space H^N with N sufficiently large. However, if quadratic nonlinearities are present, i.e. $\kappa = 2$, then the local well-posedness space also incorporates decay at infinity, namely

$$H^{N,N} = \{x^N u \in L^2 : D^N u \in L^2\}$$

I. Bejenaru: Department of Mathematics, University of California at Los Angeles, Los Angeles, CA 90095, USA; e-mail: bejenaru@math.ucla.edu

D. Tataru: Department of Mathematics, University of California at Berkeley, Berkeley, CA 94720, USA; e-mail: tataru@math.berkeley.edu

with N large enough. The need for decay is motivated by work of Mizohata [10]. He proves that a necessary condition for the L^2 well-posedness of the problem

$$\begin{cases} iu_t - \Delta u = b_1(x)\nabla u, & t \in \mathbb{R},\ x \in \mathbb{R}^n, \\ u(x,0) = u_0(x), \end{cases} \tag{2}$$

is the uniform bound

$$\sup_{x\in\mathbb{R}^n,\ \omega\in\mathbb{S}^{n-1},\ R>0} \left| \mathrm{Re} \int_0^R b_1(x + r\omega)\cdot\omega\, dr \right| < \infty. \tag{3}$$

The idea behind this condition is that $\mathrm{Re}\, b_1$ contributes to exponential growth of the solution along the Hamilton flow of the linear Schrödinger operator.

If $\kappa \geq 3$ then naively one needs such bounds for expressions which are quadratic in u. If $u \in H^N$ then $u^2 \in W^{N,1}$ and the above integrability can be gained. However, if $\kappa = 2$ then one would want similar bounds for linear expressions in u; but this cannot follow from square integrability, therefore one compensates by adding decay at infinity.

In this work we consider the case of quadratic nonlinearities, precisely the problem

$$\begin{cases} iu_t - \Delta u = B((u,\bar{u}),(\nabla u, \nabla\bar{u})), \\ u(x,0) = u_0(x). \end{cases} \tag{4}$$

Here B is a generic bilinear form which contains one differentiated and one undifferentiated factor, and also may have complex conjugates. Our results easily transfer via differentiation to the similar problem with two derivatives in the nonlinearity,

$$\begin{cases} iu_t - \Delta u = B((\nabla u,\nabla\bar{u}),(\nabla u, \nabla\bar{u})), \\ u(x,0) = u_0(x), \end{cases} \tag{5}$$

The initial data u_0 is assumed to be locally in Sobolev spaces H^s but with some additional decay at infinity. In what follows λ is a dyadic index. For a function u we consider a Littlewood–Paley decomposition in frequency

$$u = \sum_{\lambda\geq 1} u_\lambda, \qquad u_\lambda = S_\lambda u,$$

where all the frequencies smaller than 1 are included in u_1. Then following [2] we define the spaces $\mathcal{D}H^s$ by

$$\|u\|^2_{\mathcal{D}H^s} = \sum_{\lambda\geq 1} \lambda^{2s} \|u_\lambda\|^2_{\mathcal{D}L^2_\lambda},$$

where the dyadic norms $\mathcal{D}L^2_\lambda$ are defined in a manner somewhat similar to (3), namely

$$\|v\|_{\mathcal{D}L^2_\lambda} = \sup_{x_0\in\mathbb{R}^n} \sup_{\omega\in\mathbb{S}^{n-1}} \sum_{k\in\mathbb{N}} \|1_{\{|\lambda^{-1}(x-x_0)-k\omega|<1\}} u\|_{L^2}.$$

This definition is consistent with the speed of propagation properties for the linear Schrödinger equation. Waves with frequency λ have speed λ, and therefore move about λ within a unit time interval. Hence the linear Schrödinger equation is well-posed[1] in $\mathcal{D}H^s$.

If one considers the low regularity well-posedness for (4) and (5) in $\mathcal{D}H^s$ then a natural threshold is given by scaling, namely $s_c = n/2 - 1$ for (4), and $s_c = n/2$ for (5). However, it turns out that the obstruction identified in Mizohata's work is much stronger and leads to additional restrictions. In [3] Chihara obtains some better results on this problem, lowering the threshold for (5) to $s = n/2 + 4$. However, this is still far from optimal. Indeed, a sharp small data result is obtained by the first author in a recent paper:

Theorem 1 (Bejenaru [2]). *If $n = 2$, the equations* (4) *and* (5) *are locally well-posed for initial data which is small in $\mathcal{D}H^s$ for all $s > s_c + 1$.*

Although the above theorem was proved in dimension two, the same result can be derived in all dimensions $n \geq 2$. In addition, if some limited spherical symmetry is imposed on the data then the above exponents can be relaxed up to scaling (see [1]).

The goal of this paper is to obtain a local well-posedness result for the same initial data space as in [2], but for large initial data. Our main result is

Theorem 2. *Let $n \geq 2$ and $s > s_c + 1$. Then the equations* (4) *and* (5) *are locally well-posed for initial data in $\mathcal{D}H^s$. More precisely, there is $C > 0$ so that for each initial data $u_0 \in \mathcal{D}H^s$ there is a unique solution*

$$u \in C(0, T; \mathcal{D}H^s), \qquad T = e^{-C\|u_0\|_{\mathcal{D}H^s}}.$$

In addition, the solution has a Lipschitz dependence on the initial data.

We believe that this result is sharp for generic nonlinearities. However, there are special cases when one is able to obtain stronger results (see for instance [4]–[7], [12]). These results consider nonlinearities of the type $B((u, \bar{u}), \nabla \bar{u})$ and $B(\nabla \bar{u}, \nabla \bar{u})$ in various dimensions and show that the regularity threshold for the initial data can be lowered all the way to the scaling.

The step from small to large data is entirely nontrivial. The difficulty is related to the infinite speed of propagation for the linear Schrödinger equation. Precisely, a large low frequency component of the solution produces an exponentially large perturbation in the high frequency flow even for an arbitrarily short time. Thus one needs to add the low frequency part of the data to the linear equation and only then do a perturbative analysis for the high frequencies. A somewhat similar analysis has been carried out before for related problems (see [8], [3]); however, in both works the full initial data becomes part of the nonlinearity, leading to considerable technical difficulties. What also differentiates the present work is that the perturbed linear equation is very close to Mizohata's necessary condition.

Another interesting feature of this work is the choice of the function spaces for the perturbative analysis. It has been known for some time that the $X^{s,b}$ spaces are not good

[1] The operator norm of the evolution in $\mathcal{D}H^s$ will grow polynomially in time, though.

enough in order to study the local theory for derivative NLS equation. This is due to the need to regain one derivative in the bilinear estimates, which leads to logarithmic losses. In [2] the first author introduced a refinement of the $X^{s,b}$ spaces which removes this difficulty. In this article we provide an alternative modification of the $X^{s,b}$ spaces which seems better suited for the study of variable coefficient equations. This is based on a wave packet type decomposition of solutions on the uncertainty principle scale.

2. Scaling and the perturbative argument

Differentiating once the equation (5) we obtain an equation of type (4). Thus in what follows we restrict our analysis to (4). We set

$$M = \|u_0\|_{\mathcal{D}H^s}.$$

Our equation is invariant with respect to the scaling

$$u^\varepsilon(x,t) = \varepsilon u(\varepsilon x, \varepsilon^2 t), \qquad u_0^\varepsilon(x) = \varepsilon u_0(\varepsilon x). \tag{6}$$

It is natural to seek to decrease the size of the initial data by rescaling with a sufficiently small parameter ε. However, there is a difficulty arising from the fact that we are using inhomogeneous Sobolev spaces. This is why the result in [2] does not work for large initial data.

We split the rescaled initial data u_0^ε into low and high frequencies,

$$u_0^\varepsilon = u_{0,\le 1}^\varepsilon + u_{0,>1}^\varepsilon.$$

It is not too difficult to estimate their size:

Proposition 1. *Assume that $s > n/2$. Then the components $u_{0,\le 1}^\varepsilon$ and $u_{0,>1}^\varepsilon$ of u_0^ε satisfy the pointwise bounds*

$$\|u_{0,\le 1}^\varepsilon\|_{L^\infty} + \|u_{0,>1}^\varepsilon\|_{L^\infty} \lesssim \varepsilon M \tag{7}$$

and the estimates[2]

$$\|u_{0,\le 1}^\varepsilon\|_{\mathcal{D}L^2} \lesssim M, \qquad \|\nabla u_{0,\le 1}^\varepsilon\|_{\mathcal{D}L^2} \lesssim \max\{\varepsilon, \varepsilon^{s-n/2}\} M, \tag{8}$$

$$\|u_{0,>1}^\varepsilon\|_{\mathcal{D}H^s} \lesssim \varepsilon^{s-n/2} M. \tag{9}$$

The low frequency component is large but has the redeeming feature that it does not change much on the unit time scale; hence we freeze it in time modulo small errors. On the other hand, the high frequency component is small, therefore we can treat it perturbatively on a unit time interval. To write the equation for the function

$$v^\varepsilon = u^\varepsilon - u_{0,\le 1}^\varepsilon$$

[2] The constant in the second bound needs to be adjusted to $\varepsilon|\ln\varepsilon|$ if $s = n/2 + 1$.

we decompose the quadratic nonlinearity B depending on whether the gradient factor is complex conjugate or not,

$$B((u,\bar{u}),(\nabla u,\nabla\bar{u})) = B_0((u,\bar{u}),\nabla u) + B_1((u,\bar{u}),\nabla\bar{u}).$$

Then v^ε satisfies

$$\begin{cases} iv_t - \Delta v - Av = B((v,\bar{v}),(\nabla v,\nabla\bar{v})) + A_1 v + N(u^\varepsilon_{0,\le 1}), \\ v(0) = u^\varepsilon_{0,>1}, \end{cases} \tag{10}$$

where

$$Av = B_0((u^\varepsilon_{0,\le 1},\bar{u}^\varepsilon_{0,\le 1}),\nabla v)$$

is a linear term which is included in the principal part,

$$A_1 v = B((v,\bar{v}),(\nabla u^\varepsilon_{0,\le 1},\nabla\bar{u}^\varepsilon_{0,\le 1})) + B_1((u^\varepsilon_{0,\le 1},\bar{u}^\varepsilon_{0,\le 1}),\nabla\bar{v})$$

is a linear term which can be treated perturbatively, and

$$N(u^\varepsilon_{0,\le 1}) = B((u^\varepsilon_{0,\le 1},\bar{u}^\varepsilon_{0,\le 1}),(\nabla u^\varepsilon_{0,\le 1},\bar{\nabla} u^\varepsilon_{0,\le 1})) + \Delta u^\varepsilon_{0,\le 1}$$

represents the time independent contribution of the low frequency part of the data.

To solve this we need some Banach spaces X^s, $\mathcal{D}X^s$ for the solution v, and Y^s, $\mathcal{D}Y^s$ for the inhomogeneous term in the equation. These are defined in Section 3.

Instead of working with the linear Schrödinger equation we need to consider lower order perturbations of it of the form

$$\begin{cases} iv_t - \Delta v - a(t,x,D)v = f(x,t), \\ v(x,0) = g(x). \end{cases} \tag{11}$$

Given any $M \ge 1$ we introduce a larger class $\mathcal{C}_M$ of pseudodifferential operators so that for $a \in \mathcal{C}_M$ we can solve (11). To motivate the following definition, we note that the imaginary part of the symbol a can produce exponential growth in (11). We want to be able to control the growth in the phase space along the Hamilton flow of the linear Schrödinger equation, which leads to a (possibly large) bound on the integral of a along the flow. We also want the integral of a along the flow to give an accurate picture of the evolution. To ensure this we impose a smallness condition on the integral of derivatives of a along the flow. This motivates the following

Definition 1. *Let $M \ge 1$. The symbol $a(x,\xi)$ belongs to the class $\mathcal{C}_M$ if it satisfies the following conditions:*

$$M = \sup_{x,\xi} \int_0^1 |a(t,x+2t\xi,\xi)|\,dt < \infty, \tag{12}$$

$$\sup_{x,\xi} \int_0^1 |\partial_x^\alpha \partial_\xi^\beta a(t,x+2t\xi,\xi)|\,dt \le C_{\alpha,\beta}\delta, \quad |\alpha|+|\beta| \ge 1, \tag{13}$$

with $\delta e^M \ll 1$.

Then our main linear result has the form:

Theorem 3. *Let $s \in \mathbb{R}$ and $a \in \mathcal{C}_M$. Then the solution v of the equation* (11) *satisfies the estimate*

$$\|\chi_{[0,1]} v\|_{\mathcal{D}X^s} \lesssim e^M (\|f\|_{\mathcal{D}Y^s} + \|g\|_{\mathcal{D}H^s}). \tag{14}$$

To apply this theorem in our context we need to show that

Proposition 2. *For $u_0 \in \mathcal{D}H^s$ as above, set*

$$a(x,\xi) = B_0((u^\varepsilon_{0,\le 1}, \bar{u}^\varepsilon_{0,\le 1}), \xi).$$

If $\varepsilon \le e^{-CM}$ with C sufficiently large then $a \in \mathcal{C}_{c_n M}$ with c_n depending only on the dimension n.

In order to iteratively solve the equation (10) we want bounds for the right hand side terms. The main one is a bilinear estimate:

Theorem 4. *The following bilinear estimate holds:*

$$\|B((u,\bar{u}), (\nabla v, \nabla \bar{v}))\|_{\mathcal{D}Y^s} \lesssim \|u\|_{\mathcal{D}X^s} \|v\|_{\mathcal{D}X^s}, \qquad s > n/2. \tag{15}$$

For the linear term A_1 we use the following estimates:

Proposition 3. *Let A_1 be defined as above. Then*

$$\|A_1 v\|_{Y^s} \lesssim \|u^\varepsilon_{0,\le 1}\|_{L^\infty} \|v\|_{X^s}, \qquad \|A_1 v\|_{\mathcal{D}Y^s} \lesssim \|u^\varepsilon_{0,\le 1}\|_{L^\infty} \|v\|_{\mathcal{D}X^s}. \tag{16}$$

We also apply an estimate for the time independent term:

Proposition 4. *If $s > 1$ then*

$$\|B(u^\varepsilon_{0,\le 1}, \nabla u^\varepsilon_{0,\le 1}) + \Delta u^\varepsilon_{0,\le 1}\|_{\mathcal{D}H^s} \lesssim \min(\varepsilon, \varepsilon^{s-n/2}) M^2. \tag{17}$$

Using the results above we are able to conclude the proof of Theorem 2. We choose $\varepsilon = e^{-CM}$ with C large enough (depending on s). Then we solve the rescaled problem (10) on a unit time interval using the contraction principle in the space X^s.

We define the operator $\mathcal{T}_1$ by $w = \mathcal{T}_1 f$ to be the solution of the inhomogeneous Schrödinger equation with zero initial data:

$$\begin{cases} i w_t - \Delta w - A(x,D) w = f, \\ w(x,0) = 0. \end{cases} \tag{18}$$

We also denote by $\mathcal{T}_2 g$ the solution to the homogeneous equation

$$\begin{cases} i w_t - \Delta w - A(x,D) w = 0, \\ w(x,0) = g, \end{cases} \tag{19}$$

With these notations the equation (10) can be rewritten in the form

$$v = \mathcal{T} v, \qquad \mathcal{T} v = \mathcal{T}_2 u^{\varepsilon}_{0,>1} + \mathcal{T}_1 (B(v, \nabla v) + A_1 v + B(w^{\varepsilon}_0, \nabla w_0{}^{\varepsilon}) + \Delta w^{\varepsilon}_0).$$

We define the set

$$K = \{w \in \mathcal{D}X^s : \|w\|_{\mathcal{D}X^s} \leq \varepsilon^{\sigma}\}, \qquad 0 < \sigma < s - n/2,$$

and prove that $\mathcal{T} : K \to K$ and that $\mathcal{T}$ is a contraction on K. This give us the existence of a fixed point for $\mathcal{T}$ which is the solution of our problem in the interval [0, 1].

To prove the invariance of K under the action of $\mathcal{T}$ we use the results in Theorem 3:

$$\begin{aligned} \|\mathcal{T} v\|_{\mathcal{D}X^s} \lesssim \; & e^{c_n M}(\|u^{\varepsilon}_{0,>1}\|_{\mathcal{D}H^s} + \|B(v, \nabla v)\|_{\mathcal{D}Y^s} + \|A_1 v\|_{\mathcal{D}Y^s} \\ & + \|B(u^{\varepsilon}_{0,\leq 1} \nabla u^{\varepsilon}_{0,\leq 1}) + \Delta u^{\varepsilon}_{0,\leq 1}\|_{\mathcal{D}H^s}). \end{aligned}$$

Using the estimates (9) and (15)–(17) then yields

$$\|\mathcal{T} v\|_{\mathcal{D}X^s} \lesssim M^2 e^{c_n M}(\varepsilon^{s-n/2} + \|v\|^2_{\mathcal{D}X^s}).$$

If $v \in K$ then we use the smallness of ε to obtain

$$\|\mathcal{T} v\|_{\mathcal{D}X^s} \lesssim \varepsilon^{\sigma} M^2 e^{c_n M}(\varepsilon^{s-n/2-\sigma} + \varepsilon^{\sigma}) \leq \varepsilon^{\sigma},$$

which shows that $\mathcal{T} v \in K$.

To prove that $\mathcal{T}$ is a contraction we write

$$\mathcal{T} v_1 - \mathcal{T} v_2 = \mathcal{T}_1(B(v_1, \nabla(v_1 - v_2))) + \mathcal{T}_1(B(v_1 - v_2, \nabla v_2))$$

and estimate in a similar manner:

$$\begin{aligned} \|\mathcal{T} v_1 - \mathcal{T} v_2\|_{\mathcal{D}X^s} & \lesssim (\|v_1\|_{\mathcal{D}X^s} + \|v_2\|_{\mathcal{D}X^s})\|v_1 - v_2\|_{\mathcal{D}X^s} \\ & \lesssim 2\varepsilon^{\sigma} \|\chi_{[0,1]}(v_1 - v_2)\|_{\mathcal{D}X^s} < \frac{1}{2}\|\chi_{[0,1]}(v_1 - v_2)\|_{\mathcal{D}X^s}. \end{aligned}$$

3. The function spaces

Let $(x_0, \xi_0) \in \mathbb{R}^{2n}$. To describe functions which are localized in the phase space on the unit scale near (x_0, ξ_0) we use the space

$$H^{N,N}_{x^0,\xi^0} := \{f : \langle D - \xi^0\rangle^N f \in L^2, \; \langle x - x^0\rangle^N f \in L^2\}.$$

We work with the lattice $\mathbb{Z}^n$ both in the physical and Fourier space. We consider a partition of unity in the physical space,

$$\sum_{x_0 \in \mathbb{Z}^n} \phi_{x_0} = 1, \qquad \phi_{x_0}(x) = \phi(x - x_0),$$

where ϕ is a smooth bump function with compact support. We use a similar partition of unity on the Fourier side:

$$\sum_{\xi_0 \in \mathbb{Z}^n} \varphi_{\xi_0} = 1, \qquad \varphi_{\xi_0}(\xi) = \varphi(\xi - \xi_0).$$

Let H be a Hilbert space. Let $V^2 H$ be the space of right continuous H-valued functions on $\mathbb{R}$ with bounded 2-variation

$$\|u\|_{V^2 H}^2 = \sup_{(t_i) \in T} \sum_i \|u(t_{i+1}) - u(t_i)\|_H^2,$$

where T is the set of finite increasing sequences in $\mathbb{R}$. The V^2 spaces are close to the homogeneous Sobolev space $\dot{H}^{1/2}$ in the sense that

$$\dot{B}^{1/2}_{2,1} \subset V^2 \subset \dot{B}^{1/2}_{2,\infty}. \tag{20}$$

Let $U^2 H$ be the atomic space defined by the atoms:

$$u = \sum_i h_i \chi_{[t_i, t_{i+1})}, \qquad \sum_i \|h_i\|_H^2 = 1,$$

for some $(t_i) \in T$. We have the inclusion $U^2 H \subset V^2 H$ but actually these spaces are very close. There is also a duality relation between $V^2 H$ and $U^2 H$, namely

$$(DU^2 H)^* = V^2 H, \tag{21}$$

where $DU^2 H$ represents the space of derivatives of $U^2 H$ functions with the induced norm.

We can associate similar spaces to the Schrödinger flow by pulling back functions to time 0 along the flow, e.g.

$$\|u\|_{V^2_\Delta L^2} = \|e^{it\Delta} u\|_{V^2 L^2}.$$

This turns out to be a good replacement for the $X^{0,1/2}$ space associated to the Schrödinger equations. Such spaces originate in unpublished work of the second author on the wave-map equation, and have been succesfully used in various contexts so far (see [9] and [11]).

In the present paper we consider a wave packet type refinement of this structure. We begin with spatial localization, and introduce the space X of functions in $[0, 1] \times \mathbb{R}^n$ with norm

$$\|u\|_X^2 = \sum_{x_0 \in \mathbb{Z}^n} \|\phi_{x_0} e^{it\Delta} u\|^2_{V^2_t L^2_x}.$$

It is easy to see that this is a stronger norm than the $V^2_\Delta L^2$ norm,

$$\|u\|_{V^2_\Delta L^2} \lesssim \|u\|_X^2. \tag{22}$$

The X norm is usually applied to functions which are also frequency localized on the unit scale. This is consistent with the spatial localization. Precisely, we have the straightforward bound

$$\|\varphi_\xi(D) u\|_X \lesssim \|u\|_X \tag{23}$$

For a function $u : [0,1] \times \mathbb{R}^n \to \mathbb{C}$ we decompose

$$u = \sum_{\xi_0 \in \mathbb{Z}^n} u_{\xi_0}, \qquad u_{\xi_0} = \varphi_{\xi_0}(D)u,$$

and set

$$\|u\|_{X^s}^2 = \sum_{\xi_0 \in \mathbb{Z}^n} \langle \xi_0 \rangle^{2s} \|u_{\xi_0}\|_X^2.$$

An immediate consequence of (22) is that

$$\|u\|_{V_\Delta^2 L^2} \lesssim \|u\|_{X^0}^2. \tag{24}$$

We also denote by X_λ the subspace of functions in X^0 which are localized at frequency λ. It is easy to see that X^s has an l^2 dyadic structure,

$$\|u\|_{X^s}^2 \approx \sum_{\lambda \geq 1} \lambda^{2s} \|u_\lambda\|_{X_\lambda}^2$$

This is the most compact definition of X^s. Some equivalent formulations of the norm on X turn out to be more helpful in some estimates. We dedicate the next few paragraphs to such formulations. The first such formulation simply adds regularity and decay:

Proposition 5. *Let $N \in \mathbb{N}$. Then*

$$\|u_{\xi_0}\|_X^2 \approx \sum_{x_0 \in \mathbb{Z}^n} \|\phi_{x_0} e^{it\Delta} u_{\xi_0}\|_{V^2 H^{N,N}_{x_0,\xi_0}}^2. \tag{25}$$

Proof. Set $v_{\xi_0} = e^{it\Delta} u_{\xi_0}$. Then it suffices to show that

$$\sum_{x_0 \in \mathbb{Z}^n} \|\phi_{x_0} v_{\xi_0}\|_{V^2 H^{N,N}_{x_0,\xi_0}}^2 \lesssim \sum_{x_0 \in \mathbb{Z}^n} \|\phi_{x_0} v_{\xi_0}\|_{V^2 L^2}^2.$$

This in turn follows by summation from

$$\|\phi_{x_0} v_{\xi_0}\|_{V^2 H^{N,N}_{x_0,\xi_0}} \lesssim \sum_{y_0 \in \mathbb{Z}^n} \langle x_0 - y_0 \rangle^{-N} \|\phi_{y_0} v_{\xi_0}\|_{V^2 L^2}.$$

We can translate both in space and in frequency and reduce the problem to the case when $x_0 = 0$ and $\xi_0 = 0$. Then we need to show that

$$\|x^N \phi_0 v_0\|_{V^2 L^2} + \|D^N(\phi_0 v_0)\|_{V^2 L^2} \lesssim \sum_{y_0 \in \mathbb{Z}^n} \langle y_0 \rangle^{-N} \|\phi_{y_0} v_0\|_{V^2 L^2}.$$

The derivatives which fall on v_0 can be truncated at frequencies larger than 1. Then a more general formulation of the above bound is

$$\|\psi_0 \chi_0(D) v_0\|_{V^2 L^2} \lesssim \sum_{y_0 \in \mathbb{Z}^n} \langle y_0 \rangle^{-N} \|\phi_{y_0} v_0\|_{V^2 L^2}, \tag{26}$$

where both ψ_0 and χ_0 are bump functions concentrated at 0. We write

$$\psi_0 \chi_0(D) v_0 = \sum_{y_0 \in \mathbb{Z}^n} \psi_0 \chi_0(D) \phi_{y_0} v_0,$$

It remains to show that

$$\| \psi_0 \chi_0(D) \phi_{y_0} \|_{L^2 \to L^2} \lesssim \langle y_0 \rangle^{-N}.$$

which is straightforward since $\chi_0(D)$ has a bounded and rapidly decreasing kernel. □

The next equivalent definition of our function spaces relates the Schrödinger evolution to the associated Hamilton flow,

$$(x_0, \xi_0) \mapsto (x_0 - 2t\xi_0, \xi_0)$$

We begin by linearizing the symbol of $-\Delta$ near ξ_0,

$$\xi^2 = L_{\xi_0}(\xi) + O((\xi - \xi_0)^2), \qquad L_{\xi_0}(\xi) = \xi_0^2 + 2\xi\xi_0$$

The evolution generated by L is simply the transport along the Hamilton flow,

$$e^{itL_{\xi_0}} u(x) = e^{it\xi_0^2} u(x - 2t\xi_0).$$

Our next characterization of the frequency localized X norm asserts that we can replace $-\Delta$ by L_{ξ_0}.

Proposition 6. *We have*

$$\|u_{\xi_0}\|_X^2 \approx \sum_{x_0 \in \mathbb{Z}^n} \|\phi_{x_0} e^{-itL_{\xi_0}} u_{\xi_0}\|_{V^2 L^2}^2 \tag{27}$$

and

$$\|u_{\xi_0}\|_X^2 \approx \sum_{x_0 \in \mathbb{Z}^n} \|\phi_{x_0} e^{-itL_{\xi_0}} u_{\xi_0}\|_{V^2 H^{N,N}_{x_0,\xi_0}}^2. \tag{28}$$

Proof. Setting as before $v_{\xi_0} = e^{it\Delta} u_{\xi_0}$ we can write

$$e^{-itL_{\xi_0}} u_{\xi_0} = e^{it(D^2 - L_{\xi_0})} v_{\xi_0} = \chi_{\xi_0}(t, D) v_{\xi_0},$$

where the symbol $\chi(t, \xi)$ is a unit bump function in ξ around ξ_0 and smooth in t. There is also a similar formula with L_{ξ_0} and D^2 interchanged. Hence for (27) it suffices to show that

$$\sum_{x_0 \in \mathbb{Z}^n} \|\phi_{x_0} \chi_{\xi_0}(t, D) v_{\xi_0}\|_{V^2 L^2}^2 \lesssim \sum_{x_0 \in \mathbb{Z}^n} \|\phi_{x_0} v_{\xi_0}\|_{V^2 L^2}^2.$$

Without any restriction in generality we can take $\xi_0 = 0$. Using a Fourier series in t we can also replace $\chi_{\xi_0}(t, \xi)$ by an expression of the form $a(t)\chi_{\xi_0}(\xi)$ with a smooth. It remains to show that

$$\sum_{x_0 \in \mathbb{Z}^n} \|\phi_{x_0} \chi_{\xi_0}(D) v_{\xi_0}\|_{V^2 L^2}^2 \lesssim \sum_{x_0 \in \mathbb{Z}^n} \|\phi_{x_0} v_{\xi_0}\|_{V^2 L^2}^2.$$

But this follows as in Proposition 5 from (26).

Finally, the proof of (28) uses the same argument as in Proposition 25. □

For each ξ_0 we define the set T_{ξ_0} of tubes of the form

$$Q = \{(t, x) : |x - (x_Q - 2t\xi_0)| \leq 1\}, \qquad x_Q \in \mathbb{Z}^n.$$

They are the image of unit cubes centered at x_Q at time 0 along the L_{ξ_0} flow. These tubes generate the decomposition

$$[0, 1] \times \mathbb{R}^n = \bigcup_{Q \in T_{\xi_0}} Q.$$

Now we can state our last equivalent formulation of the X norm for functions localized at frequency ξ_0 in terms of a wave packet decomposition associated to tubes $Q \in T_{\xi_0}$:

Proposition 7. *Let $u_{\xi_0} \in X$. Then it can be represented as the sum of a rapidly convergent series*

$$u_{\xi_0}(t, x) = \sum_j e^{ix\xi_0} e^{it\xi_0^2} \sum_{Q \in T_{\xi_0}} a_Q^j(t) \chi^j(x - x_Q - 2t\xi_0) \tag{29}$$

with χ^j uniformly bounded in $H^{N,N}$, supported in $B(0, 2)$, and

$$\sum_{Q \in T_{\xi_0}} \|a_Q^j\|_{V^2}^2 \lesssim j^{-N} \|u_{\xi_0}\|_X.$$

Proof. The pull back to time 0 of the above representation using the L_{ξ_0} flow is

$$e^{-itL_{\xi_0}} u_{\xi_0}(t, x) = \sum_j e^{ix\xi_0} \sum_{x_0 \in \mathbb{Z}^n} a_Q^j(t) \chi^j(x - x_0).$$

We define $v_0 = e^{-ix\xi_0} e^{-itL_{\xi_0}} u_{\xi_0}$ and use (28) for the X_{ξ_0} norm. Then it suffices to show that for fixed x_0 we can represent

$$\phi_{x_0} v_0 = \sum_j a^j(t) \chi^j(x - x_0),$$

where χ^j are uniformly bounded in $H^{N,N}$ and

$$\|a^j\|_{V^2} \lesssim j^{-N} \|\phi_{x_0} v_0\|_{V^2 H_{x_0}^{CN,CN}}.$$

Without any restriction in generality we take $x_0 = 0$. We take χ^j to be the Hermite functions with the $H^{N,N}$ normalization (see e.g. [13]). Then the V^2 functions a^j are the Fourier coefficients of $\phi_{x_0} v_0$. They decay rapidly due to the additional regularity of $\phi_{x_0} v_0$.

Finally, to ensure that the χ_j's have compact support we can truncate the Fourier series outside the support of ϕ_{x_0}. □

We follow a similar path to define the Y^s structure:

$$\|u\|_{Y^s}^2 = \sum_{\xi_0 \in \mathbb{Z}^n} \langle \xi_0 \rangle^{2s} \|u_{\xi_0}\|_Y^2,$$

where Y is defined by

$$\|f\|_Y^2 = \sum_{x_0 \in \mathbb{Z}^n} \|\phi_{x_0} e^{-itD^2} f\|_{DU_t^2 L_x^2}^2.$$

All the equivalent definitions for the X have a counterpart for Y by simply replacing the V_t^2 structure by the DU_t^2 one.

There are two key relations between the X^s and Y^s spaces. The first one is concerned with solvability for the linear Schrödinger equation:

Proposition 8. *The solution u to the linear Schrödinger equation*

$$iu_t - \Delta u = f, \qquad u(0) = u_0,$$

satisfies

$$\|u\|_{X^s} \lesssim \|u_0\|_{H^s} + \|f\|_{Y^s}. \tag{30}$$

This is stated here only for the sake of completeness, as in the next section we prove a stronger estimate in Theorem 3.

The second relation is a duality relation:

Proposition 9. *We have the duality relation $(Y^s)^* = X^{-s}$.*

Proof. (a) We first verify that $X^{-s} \subset (Y^s)^*$. For this we need the bound

$$\left| \int_0^1 \int u \bar{f} \, dx \, dt \right| \lesssim \|u\|_{X^{-s}} \|f\|_{Y^s}.$$

We decompose

$$\int_0^1 \int u \bar{f} \, dx \, dt = \sum_{\xi \in \mathbb{Z}^n} \int_0^1 \int u_\xi \bar{f}_\xi \, dx \, dt.$$

Due to the definition of the X^s and Y^s norms, it suffices to show that

$$\left| \int_0^1 \int u_\xi \bar{f}_\xi \, dx \, dt \right| \lesssim \|u_\xi\|_X \|f_\xi\|_Y.$$

For this we write

$$\int_0^1 \int u_\xi \bar{f}_\xi \, dx \, dt = \int_0^1 \int e^{it\Delta} u_\xi \overline{e^{it\Delta} f_\xi} \, dx \, dt = \sum_{x_0 \in \mathbb{Z}^n} \int_0^1 \int \phi_{x_0} e^{it\Delta} u_\xi \overline{e^{it\Delta} f_\xi} \, dx \, dt$$

and use the duality relation (21) together with the definition of the X and Y norms.

(b) We now show that $(Y^s)^* \subset X^{-s}$. Let T be a bounded linear functional on Y^s. Then we have

$$|Tf| \lesssim \|f\|_{Y^s} \lesssim \Big(\sum_{\xi\in\mathbb{Z}^n} \langle\xi\rangle^{2s}\|f_\xi\|_Y^2\Big)^{1/2} \lesssim \Big(\sum_{x_0\in\mathbb{Z}^n}\sum_{\xi\in\mathbb{Z}^n} \langle\xi\rangle^{2s}\|\phi_{x_0}e^{it\Delta}f_\xi\|_{DU^2L^2}^2\Big)^{1/2}.$$

By the Hahn–Banach theorem we can extend T to a bounded linear functional on the space $l^2_{\langle\xi\rangle^s}DU^2L^2$. By (21) this implies that we can represent T in the form

$$Tf = \sum_{x_0\in\mathbb{Z}^n}\sum_{\xi\in\mathbb{Z}^n}\int_0^1\int v_{x_0,\xi}\overline{\phi_{x_0}e^{it\Delta}f_\xi}\,dx\,dt,$$

where

$$\sum_{x_0\in\mathbb{Z}^n}\sum_{\xi\in\mathbb{Z}^n}\langle\xi\rangle^{-2s}\|v_{x_0,\xi}\|_{V^2L^2}^2 \lesssim \|T\|_{(Y^s)^*}^2.$$

Due to the above representation we can identify T with the function

$$u_T = \sum_{x_0\in\mathbb{Z}^n}\sum_{\xi\in\mathbb{Z}^n}\phi_\xi(D)e^{-it\Delta}\phi_{x_0}v_{x_0,\xi}.$$

It remains to show that

$$\|u_T\|_{X^{-s}}^2 \lesssim \sum_{x_0\in\mathbb{Z}^n}\sum_{\xi\in\mathbb{Z}^n}\langle\xi\rangle^{-2s}\|v_{x_0,\xi}\|_{V^2L^2}^2.$$

This reduces to the fixed ξ bound

$$\Big\|\sum_{x_0\in\mathbb{Z}^n}\varphi_\xi(D)e^{-it\Delta}\phi_{x_0}v_{x_0,\xi}\Big\|_X^2 \lesssim \sum_{x_0\in\mathbb{Z}^n}\|v_{x_0,\xi}\|_{V^2L^2}^2$$

with a modified φ_ξ. Using the definition of the X norm we rewrite this as

$$\sum_{y_0\in\mathbb{Z}^n}\Big\|\sum_{x_0\in\mathbb{Z}^n}\phi_{y_0}\varphi_\xi(D)\phi_{x_0}v_{x_0,\xi}\Big\|_{V^2L^2}^2 \lesssim \sum_{x_0\in\mathbb{Z}^n}\|v_{x_0,\xi}\|_{V^2L^2}^2.$$

But this follows by Cauchy–Schwarz from the rapid decay

$$\|\phi_{y_0}\varphi_\xi(D)\phi_{x_0}\|_{L^2\to L^2} \lesssim \langle x_0-y_0\rangle^{-N}. \qquad \square$$

We still need to add the decay structure to the X^s and Y^s spaces. Given $\lambda \ge 1$, we roughly want to ask for l^1 summability of frequency λ norms along collinear cubes of size λ. We define the $\mathcal{D}X_\lambda$ norm by

$$\|u\|_{\mathcal{D}X_\lambda} = \sup_{x_0\in\mathbb{R}^n}\sup_{\omega\in\mathbb{S}^{n-1}}\sum_{k=-\infty}^{\infty}\|\chi(\lambda^{-1}(x-x_0)-k\omega)u\|_{X_\lambda}, \tag{31}$$

where χ is a compactly supported bump function. We note that this norm can only be meaningfully used for functions at frequency λ. Summing up with respect to λ we also set

$$\|u\|^2_{\mathcal{D}X^s} = \sum_{\lambda \geq 1} \|u_\lambda\|^2_{\mathcal{D}X_\lambda}.$$

In a completely similar way we can define $\mathcal{D}Y_\lambda$ and $\mathcal{D}Y^s$.

A useful tool in our analysis is a family of embeddings which correspond to the Strichartz estimates for the Schrödinger equation.

Proposition 10. *Let p and q be indices which satisfy*

$$\frac{2}{p} + \frac{n}{q} = \frac{n}{2}, \quad 2 < p \leq \infty, \quad 2 \leq q \leq \infty.$$

Then we have the embedding

$$X^0 \subset L^p_t L^q_x.$$

By (24) these embeddings are a direct consequence of the Strichartz estimates for the Schrödinger equation (see [9, proof of Proposition 6.2]). By Sobolev embeddings we also obtain bounds with larger p, q for frequency localized solutions.

Corollary 1. *Let p and q be indices which satisfy*

$$\frac{2}{p} + \frac{n}{q} \leq \frac{n}{2}, \quad 2 < p \leq \infty, \quad 2 \leq q \leq \infty.$$

Then we have the embedding

$$S_{<\mu} X^0 \subset \mu^{n/2-2/p-n/q} L^p_t L^q_x,$$

where $S_{<\mu}$ can be replaced with a multiplier localizing in frequency to an arbitrary cube of size μ.

In this article we use only the case $p = q$, more precisely

$$X^0 \subset L^{2(n+2)/n}, \quad S_{<\mu} X^0 \subset \mu^{n/2-1} L^{n+2}. \tag{32}$$

Finally, we introduce modulation localization operators, which we can define in two equivalent ways. The first method is as multipliers,

$$\widehat{M_{<\sigma} u} = s_{<\sigma}(\tau - \xi^2)\hat{u}.$$

The second is obtained by conjugation with respect to the Schrödinger flow,

$$e^{itD^2}(M_{<\sigma} u)(t) = s_{<\sigma}(D_t)(e^{itD^2} u(t)).$$

The operators $s_{<\sigma}(D_t)$ are bounded on V^2. On the other hand, for the remainder, by (20), we have a good L^2 bound:

$$\|s_{>\sigma}(D_t) a\|_{L^2} \lesssim \sigma^{-1/2} \|a\|_{V^2}.$$

By the second definition of the modulation localization operators above we obtain

Proposition 11. (a) *The operators $M_{<\sigma}$ are bounded on X^0.*
(b) *The following estimate holds:*

$$\|M_{>\sigma} u\|_{L^2} \lesssim \sigma^{-1/2} \|u\|_{X^0}.$$

We note that in fact, by using both inclusions in (20), one can relate our X^s spaces to the traditional $X^{s,b}$ spaces, namely

$$\dot{X}^{s,1/2,1} \subset X^s \subset \dot{X}^{s,1/2,\infty}. \tag{33}$$

4. Linear estimates

This section is devoted to the study of the linear equation (11) with $a \in \mathcal{C}_M$. Precisely, we aim to prove Theorem 3. We define

$$L_a = i\partial_t - \Delta - a(t, x, D).$$

We begin our analysis with a heuristic computation. Suppose we have a solution u^{x_0,ξ_0} to

$$(i\partial_t - \Delta - a(t, x, D))u = 0$$

which is localized on the unit scale near the bicharacteristic $t \mapsto (x_0 + 2t\xi_0, \xi_0)$. Then we can freeze the symbol of $a(t, x, D)$ along the ray and write

$$a(t, x, D)u = a(t, x_0 + 2\xi_0, \xi_0)u + \text{error}.$$

Thus u^{x_0,ξ_0} approximately solves

$$(i\partial_t - \Delta - a(t, x_0 + 2\xi_0, \xi_0))u_{x_0,\xi_0} \approx 0.$$

This implies that we can represent u^{x_0,ξ_0} in the form

$$u^{x_0,\xi_0}(t) \approx e^{\int_0^t a(s,x_0+2s\xi_0,\xi_0)\,ds} v^{x_0,\xi_0},$$

where v^{x_0,ξ_0} solves the equation

$$(i\partial_t - \Delta)v^{x_0,\xi_0} = 0.$$

Hence along each wave packet we can use the above exponential to approximately conjugate the variable coefficient equation to the flat flow.

Using this idea we produce wave packet approximate solutions for the equation (11). By orthogonality these combine into general approximate solutions to (11). The exact solutions are obtained via a Picard iteration.

We first consider the regularity of the exponential weight. By (12) we have

$$\int_0^1 |a(t, x_0 + 2t\xi_0, \xi_0)|\,dt \le M,$$

which implies a $W^{1,1}$ bound for the exponential,

$$\left\| \frac{d}{dt} e^{\int_0^t a(s,x_0+2s\xi_0,\xi_0)\,ds} \right\|_{L^1} \le e^M. \tag{34}$$

Since $W^{1,1} \subset V^2$ this yields a similar V^2 bound,

$$\left\| e^{\int_0^t a(s,x_0+2s\xi_0,\xi_0)\,ds} \right\|_{V^2} \le e^M. \tag{35}$$

We continue our analysis with the localized equation

$$L_a u = f_{x_0,\xi_0}, \quad u(0) = g_{x_0,\xi_0}. \tag{36}$$

According to the above heuristics, an approximate solution for this should be given by Duhamel's formula,

$$u^{x_0,\xi_0} = e^{\int_0^t a(s,x_0+2s\xi_0,\xi_0)\,ds} e^{-it\Delta} g_{x_0,\xi_0} + \int_0^t e^{\int_s^t a(\tau,x_0+2\tau\xi_0,\xi_0)\,d\tau} e^{-i(t-s)\Delta} f_{x_0,\xi_0}(s)\,ds.$$

We prove that this is indeed the case:

Proposition 12. *The function u^{x_0,ξ_0} is an approximate solution for* (36) *in the sense that*

$$\|e^{it\Delta} u^{x_0,\xi_0}\|_{V^2 H^{N,N}_{x_0,\xi_0}} \lesssim e^M (\|g_{x_0,\xi_0}\|_{H^{N,N}_{x_0}} + \|e^{it\Delta} f_{x_0,\xi_0}\|_{DU^2 H^{N,N}_{x_0}}) \tag{37}$$

and

$$\begin{aligned} &\|e^{it\Delta}(L_a u^{x_0,\xi_0} - f_{x_0,\xi_0})\|_{L^1 H^{N,N}_{x_0,\xi_0}} \\ &\qquad \lesssim \delta e^M (\|g_{x_0,\xi_0}\|_{H^{N+2n+1,N+2n+1}_{x_0,\xi_0}} + \|e^{it\Delta} f_{x_0,\xi_0}\|_{DV^2 H^{N+2n+1,N+2n+1}_{x_0,\xi_0}}). \end{aligned} \tag{38}$$

Proof. For the first bound we shorten the notation

$$a_0(s) := a(s, x_0 + 2s\xi_0, \xi_0)$$

and compute

$$e^{it\Delta} u^{x_0,\xi_0} = e^{\int_0^t a_0(s)\,ds} g_{x_0,\xi_0} + \int_0^t e^{\int_s^t a_0(\tau)\,d\tau} e^{is\Delta} f_{x_0,\xi_0}(s)\,ds.$$

The first term is estimated directly by (35). Setting

$$F(t) = \int_0^t e^{is\Delta} f_{x_0,\xi_0}(s)\,ds$$

and integrating by parts we write the second term in the form

$$F(t) - \int_0^t \frac{d}{ds} e^{\int_s^t a_0(\tau)\,d\tau} F(s)\,ds.$$

We have

$$\|F\|_{V^2 H^{N,N}_{x_0,\xi_0}} \lesssim \|e^{is\Delta} f_{x_0,\xi_0}(s)\|_{DU^2 H^{N,N}_{x_0,\xi_0}},$$

while the V^2 norm of the second part is controlled by its $W^{1,1}$ norm, namely

$$\begin{aligned}
&\left\| \int_0^t \frac{d}{ds} e^{\int_s^t a_(0\tau)\, d\tau} F(s)\, ds \right\|_{V^2 H^{N,N}_{x_0,\xi_0}} \\
&\quad \lesssim \left\| \frac{d}{dt} \int_0^t \frac{d}{ds} e^{\int_s^t a_0(\tau)\, d\tau} F(s)\, ds \right\|_{L^1 H^{N,N}_{x_0,\xi_0}} \\
&\quad \lesssim \int_0^1 |a_0(t)| \left(1 + \int_0^t \left| \frac{d}{ds} e^{\int_s^t a_0(\tau)\, d\tau} \right| ds \right) dt\, \|F\|_{L^\infty H^{N,N}_{x_0,\xi_0}} \\
&\quad \lesssim \int_0^1 |a_0(t)| \left(1 + \int_0^t |a_0(s)| e^{\int_s^t |a_0(\tau)|\, d\tau}\, ds \right) dt\, \|F\|_{L^\infty H^{N,N}_{x_0,\xi_0}} \\
&\quad \lesssim \int_0^1 |a_0(t)| e^{\int_0^t |a_0(\tau)|\, d\tau}\, dt\, \|F\|_{L^\infty H^{N,N}_{x_0,\xi_0}} \\
&\quad \lesssim e^M \|F\|_{L^\infty H^{N,N}_{x_0,\xi_0}} \lesssim e^M \|F\|_{V^2 H^{N,N}_{x_0,\xi_0}}.
\end{aligned}$$

This concludes the proof of (37).

It remains to prove (38). A direct computation yields

$$(i\partial_t - \Delta - a(t,x,D))u^{x_0,\xi_0} - f_{x_0,\xi_0} = b(t,x,D)u^{x_0,\xi_0},$$

where

$$b(t,x,\xi) = a(t, x_0 + 2t\xi_0, \xi_0) - a(t,x,\xi).$$

By (37) it suffices to show that

$$\|e^{it\Delta} b(t,x,D)u^{x_0,\xi_0}\|_{L^1 H^{N,N}_{x_0,\xi_0}} \lesssim \delta \|e^{it\Delta} u^{x_0,\xi_0}\|_{L^\infty H^{N+2n+1,N+2n+1}_{x_0,\xi_0}}$$

Since the flat Schrödinger flow has the mapping property

$$\|e^{it\Delta} f\|_{H^{N,N}_{x_0,\xi_0}} \approx \|f\|_{H^{N,N}_{x_0+2t\xi_0,\xi_0}},$$

this is equivalent to

$$\|b(t,x,D)u^{x_0,\xi_0}\|_{L^1 H^{N,N}_{x_0+2t\xi_0,\xi_0}} \lesssim \delta \|u^{x_0,\xi_0}\|_{L^\infty H^{N+3,N+3}_{x_0+2t\xi_0,\xi_0}}. \tag{39}$$

We begin with a straightforward consequence of the S_{00} calculus:

Lemma 1. *Let k be a nonnegative integer and c be a symbol which satisfies*

$$|\partial_x^\alpha \partial_\xi^\beta c(x,\xi)| \leq c_{\alpha,\beta} \langle (x - x_0, \xi - \xi_0) \rangle^k, \qquad |\alpha| + |\beta| \geq 0.$$

Then for all $N \in \mathbb{N}$ we have

$$\|c(x,D)u\|_{H^{N,N}_{x_0,\xi_0}} \lesssim \|u\|_{H^{N+k,N+k}_{x_0,\xi_0}}.$$

In order to use this lemma for the operator b above we need the following Sobolev embedding:

Lemma 2. *Let $R > 1$ and $c \in W^{2n,1}(B_R(x_0, \xi_0))$. Then*

$$|c(x,\xi)| \lesssim \sum_{0\le|\alpha|\le 2n} \|\partial^\alpha c\|_{L^1(B_R(x_0,\xi_0))}$$

and

$$|c(x,\xi) - c(x_0,\xi_0)| \lesssim \sum_{1\le|\alpha|\le 2n} \|\partial^\alpha c\|_{L^1(B_R(x_0,\xi_0))}.$$

As a consequence, we obtain

Lemma 3. *Let $c \in W^{2n,1}_{\mathrm{loc}}(\mathbb{R}^{2n})$. Then*

$$\frac{|c(x,\xi)|}{\langle (x-x_0, \xi-\xi_0)\rangle^{2n+1}} \lesssim \int_{\mathbb{R}^{2n}} \sum_{0\le|\alpha|\le 2n} \frac{|\partial^\alpha c(x,\xi)|}{\langle (x-x_0,\xi-\xi_0)\rangle^{2n+1}}\, dx\, d\xi$$

and

$$\frac{|c(x,\xi) - c(x_0,\xi_0)|}{\langle (x-x_0, \xi-\xi_0)\rangle^{2n+1}} \lesssim \int_{\mathbb{R}^{2n}} \sum_{1\le|\alpha|\le 2n} \frac{|\partial^\alpha c(x,\xi)|}{\langle (x-x_0,\xi-\xi_0)\rangle^{2n+1}}\, dx\, d\xi.$$

Applying these inequalities to the symbol b above we obtain pointwise bounds for b,

$$\frac{|b(t,x,\xi)|}{\langle (x-x_0-2t\xi_0, \xi-\xi_0)\rangle^{2n+1}} \lesssim \int_{\mathbb{R}^{2n}} \frac{\sum_{|\alpha|=1}^{2n} |\partial^\alpha a(t,x,\xi)|}{\langle (x-x_0-2t\xi_0,\xi-\xi_0)\rangle^{2n+1}}\, dx\, d\xi,$$

and also for its derivatives,

$$\frac{|\partial^k b(t,x,\xi)|}{\langle (x-x_0-2t\xi_0, \xi-\xi_0)\rangle^{2n+1}} \lesssim \int_{\mathbb{R}^{2n}} \frac{\sum_{|\alpha|=k}^{2n+k} |\partial^\alpha a(t,x,\xi)|}{\langle (x-x_0-2t\xi_0,\xi-\xi_0)\rangle^{2n+1}}\, dx\, d\xi.$$

Then by Lemma 1 it follows that

$$\|b(t,x,D)\|_{H^{N+2n+1,N+2n+1}_{x_0+2t\xi_0,\xi_0} \to H^{N,N}_{x_0+2t\xi_0,\xi_0}} \lesssim \int_{\mathbb{R}^{2n}} \frac{\sum_{|\alpha|=1}^{N_0} |\partial^\alpha a(t,x,\xi)|}{\langle (x-x_0,\xi-\xi_0)\rangle^{2n+1}}\, dx\, d\xi$$

with N_0 sufficiently large. Integrating in t and changing coordinates in the integral gives

$$\|b(t,x,D)\|_{L^\infty H^{N+2n+1,N+2n+1}_{x_0+2t\xi_0,\xi_0} \to L^1 H^{N,N}_{x_0+2t\xi_0,\xi_0}} \lesssim \int_{\mathbb{R}^{2n}} \int_0^1 \sum_{|\alpha|=1}^{N_0} \frac{|\partial^\alpha a(t,x+2t\xi,\xi)|}{\langle (x-x_0,\xi-\xi_0)\rangle^{2n+1}}\, dt\, dx\, d\xi.$$

By (13) we can bound each time integral by δ and the remaining weight is integrable in x and ξ. Hence (39) follows. □

Next we produce approximate solutions for the frequency localized data,

$$L_a u = f_{\xi_0}, \quad u(0) = g_{\xi_0}. \tag{40}$$

We denote the approximate solution by u^{ξ_0}; the notation u_{ξ_0} continues to be reserved for a frequency localized part of a function u.

Proposition 13. *There is an approximate solution u^{ξ_0} to the equation* (40)*, localized at frequency ξ_0, with $u(0) = g_{\xi_0}$, which satisfies the bounds*

$$\|u^{\xi_0}\|_X \lesssim e^M(\|g_{\xi_0}\|_{L^2} + \|f_{\xi_0}\|_Y) \tag{41}$$

and

$$\|S_\xi(L_a u^{\xi_0} - f_{\xi_0})\|_Y \lesssim \delta\langle\xi - \xi_0\rangle^{-N} e^M(\|g_{\xi_0}\|_{L^2} + \|f_{\xi_0}\|_Y). \tag{42}$$

Proof. We decompose

$$g_{\xi_0} = \sum_{x_0} g_{x_0,\xi_0}, \qquad f_{\xi_0} = \sum_{x_0} f_{x_0,\xi_0},$$

and solve the problem (36) for which we have the estimates (37) and (38). We define our approximate solution to be the sum of the approximate solutions u^{x_0,ξ_0}:

$$u^{\xi_0} = \sum_{x_0} u^{x_0,\xi_0}.$$

Using (37) we obtain

$$\begin{aligned}
\|u^{\xi_0}\|_X^2 &= \sum_{y_0} \|\phi_{y_0} e^{-it\Delta} u_{\xi_0}\|_{V^2L^2}^2 \lesssim \sum_{y_0}\Big(\sum_{x_0} \|\phi_{y_0} e^{-it\Delta} u^{x_0,\xi_0}\|_{V^2L^2}\Big)^2 \\
&\lesssim \sum_{y_0}\Big(\sum_{x_0} \langle x_0 - y_0\rangle^{-N} \|e^{-it\Delta} u^{x_0,\xi_0}\|_{V^2H^{N,N}_{x_0,\xi_0}}\Big)^2 \\
&\lesssim \sum_{x_0} \|e^{-it\Delta} u^{x_0,\xi_0}\|^2_{V^2L^{N,N}_{x_0,\xi_0}} \\
&\lesssim e^M \sum_{x_0} \|g_{x_0,\xi_0}\|^2_{H^{N+2n+1,N+2n+1}_{x_0,\xi_0}} + \|e^{-it\Delta} f_{x_0,\xi_0}\|^2_{DU^2H^{N+2n+1,N+2n+1}_{x_0,\xi_0}} \\
&\lesssim e^M(\|g_{\xi_0}\|_{L^2}^2 + \|f_{\xi_0}\|^2_{Y_{\xi_0}}).
\end{aligned}$$

We continue now with the estimates for the error, using (38) instead of (37). We have

$$\begin{aligned}
\|S_\xi(L_a u^{\xi_0} - f_{\xi_0})\|_Y^2 &= \sum_y \|\phi_y e^{it\Delta} S_\xi(L_a u^{\xi_0} - f_{\xi_0})\|^2_{DU^2L^2} \\
&\lesssim \sum_y\Big(\sum_{x_0} \|\phi_y S_\xi e^{it\Delta}(L_a u^{x_0,\xi_0} - f_{x_0,\xi_0})\|_{L^1L^2}\Big)^2 \\
&\lesssim \sum_y\Bigg(\sum_{x_0} \frac{\|e^{it\Delta}(L_a u^{x_0,\xi_0} - f_{x_0,\xi_0})\|_{L^1H^{N,N}_{x_0,\xi_0}}}{\langle(y - x_0 - 2t\xi_0, \xi - \xi_0)\rangle^N}\Bigg)^2 \\
&\lesssim \langle\xi - \xi_0\rangle^{n-2N} \sum_{x_0} \|e^{it\Delta}(L_a u^{x_0,\xi_0} - f_{x_0,\xi_0})\|^2_{L^1H^{N,N}_{x_0,\xi_0}}.
\end{aligned}$$

Then (42) follows from (38). ◻

The next stage is to consider data which is localized at frequency λ,

$$L_a u = f_\lambda, \quad u(0) = g_\lambda. \tag{43}$$

Summing up the frequency localized solutions we obtain as above a dyadic approximate solution:

Proposition 14. *There is an approximate solution u^λ to the equation* (43)*, localized at frequency λ, with $u(0) = g_\lambda$, which satisfies the bounds*

$$\|u^\lambda\|_{X_\lambda} \lesssim e^M(\|g_\lambda\|_{L^2} + \|f_\lambda\|_{Y_\lambda}) \tag{44}$$

and

$$\|S_\mu(L_a u^\lambda - f_\lambda)\|_{Y_\mu} \lesssim (\min\{\mu/\lambda, \lambda/\mu\})^N \delta e^M(\|g_\lambda\|_{L^2} + \|f_\lambda\|_{Y_\lambda}). \tag{45}$$

The construction of the functions u^λ involves only the constant coefficient Schrödinger flow at frequency λ. This has spatial speed of propagation λ, therefore it can spread by at most $O(\lambda)$ in a unit time interval. Thus the above construction can be trivially localized on the λ spatial scale, leading to the bounds

$$\|u^\lambda\|_{\mathcal{D}X_\lambda} \lesssim e^M(\|g_\lambda\|_{\mathcal{D}L^2_\lambda} + \|f_\lambda\|_{\mathcal{D}Y_\lambda}) \tag{46}$$

and

$$\|S_\mu(L_a u^\lambda - f_\lambda)\|_{\mathcal{D}Y_\mu} \lesssim (\min\{\mu/\lambda, \lambda/\mu\})^N \delta e^M(\|g_\lambda\|_{\mathcal{D}L^2_\lambda} + \|f_\lambda\|_{\mathcal{D}Y_\lambda}). \tag{47}$$

After an addition dyadic summation we obtain a global parametrix:

Proposition 15. *There is an approximate solution u to the equation* (11)*, with $u(0) = g$, which satisfies the bounds*

$$\|u\|_{\mathcal{D}X^s} \lesssim e^M(\|g\|_{\mathcal{D}H^s} + \|f\|_{\mathcal{D}Y^s}) \tag{48}$$

and

$$\|(i\partial_t - \Delta - a(t,x,D))u - f)\|_{\mathcal{D}Y^s} \lesssim \delta e^M(\|g\|_{\mathcal{D}H^s} + \|f\|_{\mathcal{D}Y^s}). \tag{49}$$

Of course the similar result without decay is also valid.

If $\delta \ll e^{-M}$ then the constant in (49) is less than 1. Then one can iterate to obtain an exact solution u to (11) which still satisfies (46). Theorem 3 follows.

5. Low frequency bounds

Here we prove Propositions 1–4.

Proof of Proposition 1. For simplicity we set $f = u_0$. The pointwise bounds (7) follow from Sobolev type estimates and scaling,

$$\|f^\varepsilon\|_{L^\infty} = \varepsilon\|f\|_{L^\infty} \lesssim \varepsilon\|f\|_{\mathcal{D}H^s}, \qquad s > n/2.$$

We now move to the L^2 bounds. The effect of scaling on the $\mathcal{D}L^2_\lambda$ norms is easy to compute,

$$\|f^\varepsilon\|_{\mathcal{D}L^2_\lambda} = \varepsilon^{1-n/2}\|f\|_{\mathcal{D}L^2_{\varepsilon\lambda}}. \tag{50}$$

Hence the first bound in (8) can be rewritten in the form

$$\varepsilon^{1-n/2}\|f_{<\varepsilon^{-1}}\|_{\mathcal{D}L^2_\varepsilon} \lesssim \|f\|_{\mathcal{D}H^s}. \tag{51}$$

We have

$$f_{\leq\varepsilon^{-1}} = \sum_{1\leq\lambda\leq\varepsilon^{-1}} f_\lambda.$$

For each such λ we have three relevant spatial scales,

$$\varepsilon \leq \lambda^{-1} \leq \lambda.$$

The middle one arises due to the uncertainty principle; namely, this is the scale on which f_λ is smooth. Then we can write the sequence of inequalities

$$\begin{aligned}\|f_\lambda\|_{\mathcal{D}L^2_\varepsilon} &\lesssim (\varepsilon\lambda)^{-1}\varepsilon^{n/2}\|f_\lambda\|_{\mathcal{D}L^\infty_{\lambda^{-1}}} \lesssim (\varepsilon\lambda)^{-1}\varepsilon^{n/2}\lambda^{n/2}\|f_\lambda\|_{\mathcal{D}L^2_{\lambda^{-1}}}\\ &\lesssim (\varepsilon\lambda)^{-1}\varepsilon^{n/2}\lambda^{n/2}\lambda\|f_\lambda\|_{\mathcal{D}L^2_\lambda} = \varepsilon^{n/2-1}\lambda^{n/2}\|f_\lambda\|_{\mathcal{D}L^2_\lambda}.\end{aligned}$$

All these bounds are obtained by comparing tubes with the same orientation ω. The first step uses Hölder's inequality to switch from the ε scale to the λ^{-1} scale; the $\mathcal{D}L^\infty$ norm is defined in the same way as the $\mathcal{D}L^2$ norm. The second takes advantage of the localization at frequency λ. Finally, the third step uses Hölder's inequality to switch from the λ^{-1} scale to the λ scale. Summing up with respect to λ we obtain (51),

$$\|f_{<\varepsilon^{-1}}\|_{\mathcal{D}L^2_\varepsilon} \lesssim \sum_{1\leq\lambda\leq\varepsilon^{-1}} \|f_\lambda\|_{\mathcal{D}L^2_\varepsilon} \lesssim \varepsilon^{n/2-1} \sum_{1\leq\lambda\leq\varepsilon^{-1}} \lambda^{n/2}\|f_\lambda\|_{\mathcal{D}L^2_\lambda} \lesssim \varepsilon^{n/2-1}\|f\|_{\mathcal{D}H^s}$$

for $s > n/2$.

Consider now the second part of (8). The analogue of (50) is

$$\|\nabla f^\varepsilon\|_{\mathcal{D}L^2_\lambda} = \varepsilon^{2-n/2}\|\nabla f\|_{\mathcal{D}L^2_{\varepsilon\lambda}},$$

and therefore we have to show that

$$\varepsilon^{2-n/2}\|\nabla f_{<\varepsilon^{-1}}\|_{\mathcal{D}L^2_\varepsilon} \lesssim \max\{\varepsilon, \varepsilon^{s-n/2}\}\|f\|_{\mathcal{D}H^s}.$$

We proceed as above:

$$\|\nabla f_{<\varepsilon^{-1}}\|_{\mathcal{D}L^2_\varepsilon} \lesssim \sum_{1\le\lambda\le\varepsilon^{-1}} \lambda\|f_\lambda\|_{\mathcal{D}L^2_\varepsilon} \lesssim \varepsilon^{n/2-1} \sum_{1\le\lambda\le\varepsilon^{-1}} \lambda^{n/2+1}\|f_\lambda\|_{\mathcal{D}L^2_\lambda}.$$

The bound for the last sum is straightforward and depends on the relative positions of s and $n/2+1$.

It remains to consider the high frequency estimate (9), for which we only need the scaling relation (50) and Hölder's inequality:

$$\begin{aligned}\|f^\varepsilon_{>1}\|^2_{\mathcal{D}H^s} &= \sum_{\lambda>1}\lambda^{2s}\|(f^\varepsilon)_\lambda\|^2_{\mathcal{D}L^2_\lambda} = \varepsilon^{2-n}\sum_{\lambda>1}\lambda^{2s}\|f_{\varepsilon^{-1}\lambda}\|^2_{\mathcal{D}L^2_{\varepsilon\lambda}}\\ &\lesssim \varepsilon^{2-n}\sum_{\lambda>1}\varepsilon^{-2}\lambda^{2s}\|f_{\varepsilon^{-1}\lambda}\|^2_{\mathcal{D}L^2_{\varepsilon^{-1}\lambda}} = \varepsilon^{2s-n}(\varepsilon^{-1}\lambda)^{2s} \lesssim \varepsilon^{2s-n}\|f\|^2_{\mathcal{D}H^s}. \quad \square\end{aligned}$$

Proof of Proposition 2. It suffices to consider the $\tilde a(t,x,\xi) = u^\varepsilon_{0,\le 1}\xi$ part of the symbol, since $\nabla u^\varepsilon_{0,\le 1}$ is obtained from $\tilde a$ by differentiation. The main ingredient of the proof is

Lemma 4. *Let $s > n/2$. Then*

$$\sup_{x_0\in\mathbb{R}^n}\sup_{|\omega_0|=1}\int_{-\infty}^{\infty}|f(x_0+t\omega_0)|\,dt \le \|f\|_{\mathcal{D}H^s}. \tag{52}$$

Proof. For each dyadic component f_λ of f we have

$$\begin{aligned}\int_{-\infty}^{\infty}|f_\lambda(x_0+t\omega_0)|\,dt &= \sum_{k\in\mathbb{Z}}\int_{\lambda k}^{\lambda(k+1)}|f_\lambda(x_0+t\omega_0)|\,dt\\ &\lesssim \lambda^{1/2}\sum_{k\in\mathbb{Z}}\Big(\int_{\lambda k}^{\lambda(k+1)}|f^2_\lambda(x_0+t\omega_0)|\,dt\Big)^{1/2}\\ &\lesssim \lambda^{n/2}\sum_{k\in\mathbb{Z}}\|1_{\{|\lambda^{-1}(x-x_0)-k\omega|<1\}}f_\lambda\|_{L^2} \lesssim \lambda^{n/2}\|f_\lambda\|_{\mathcal{D}L^2_\lambda}.\end{aligned}$$

At the second step we have used Hölder's inequality and at the third we have used the frequency localization. The summation with respect to λ gives the desired result. $\square$

We return to the proof of the proposition. A change of variables combined with Lemma 4 and the first part of (8) gives

$$\sup_{x,\xi}\int_0^1|\tilde a(t,x+2t\xi,\xi)|\,dt = \sup_{x,\xi}\int_0^{2|\xi|}|u^\varepsilon_{0,\le1}(x+t\omega_0)|\,dt \lesssim \|u^\varepsilon_{0,\le1}\|_{\mathcal{D}H^s} \lesssim M,$$

where $\omega_0 = \xi/|\xi|$.

For x derivatives of $\tilde a$ we use the second part of (8) instead:

$$\int_0^1|\partial_x\tilde a(x+2t\xi,\xi)|\,dt \lesssim \|\nabla u^\varepsilon_{0,\le1}\|_{\mathcal{D}H^s} \lesssim \max\{\varepsilon,\varepsilon^{s-n/2}\}M.$$

For ξ derivatives we use the pointwise bound (7):

$$\int_0^1 |\partial_\xi \tilde{a}(x+2t\xi,\xi)|\,dt = \int_0^1 |u^\varepsilon_{0,\le 1}(x+2t\xi)|\,dt \lesssim \varepsilon M.$$

Since $u^\varepsilon_{0,\le 1}$ is supported at frequencies $\lesssim 1$, we obtain similar bounds for higher order derivatives. □

Proof of Proposition 3. We neglect the gradients applied to $u^\varepsilon_{0,\le 1}$. Also for the estimate involving $\bar{v}$ we retain only the stronger bound[3] involving $\nabla\bar{v}$. Then we need to show that

$$\|u^\varepsilon_{0,\le 1} v\|_{Y^s} \lesssim \|u^\varepsilon_{0,\le 1}\|_{L^\infty}\|v\|_{X^s}, \tag{53}$$

$$\|u^\varepsilon_{0,\le 1} \nabla\bar{v}\|_{Y^s} \lesssim \|u^\varepsilon_{0,\le 1}\|_{L^\infty}\|v\|_{X^s}. \tag{54}$$

In (53) we use orthogonality with respect to unit frequency cubes to reduce it to

$$\|u^\varepsilon_{0,\le 1} v_{\xi_0}\|_{Y} \lesssim \|u^\varepsilon_{0,\le 1}\|_{L^\infty}\|v_{\xi_0}\|_{X}.$$

By duality this becomes

$$\left|\int u^\varepsilon_{0,\le 1} v_{\xi_0}\bar{w}_{\xi_0}\,dx\,dt\right| \lesssim \|u^\varepsilon_{0,\le 1}\|_{L^\infty}\|v_{\xi_0}\|_{X}\|w_{\xi_0}\|_{X}.$$

Hence it suffices to show that

$$\int |v_{\xi_0}|\,|w_{\xi_0}|\,dx\,dt \lesssim \|v_{\xi_0}\|_{X}\|w_{\xi_0}\|_{X}.$$

By orthogonality with respect to T_{ξ_0} tubes (see Proposition 7) it remains to verify that given a ξ tube Q we can integrate a bump function on Q,

$$\int 1_Q\,dx\,dt \lesssim 1,$$

which is trivial.

Similarly, (54) reduces to the frequency localized bound

$$\left|\int u^\varepsilon_{0,\le 1} v_{\xi_0} w_{-\xi_0}\,dx\,dt\right| \lesssim |\xi_0|^{-1}\|u^\varepsilon_{0,\le 1}\|_{L^\infty}\|v_{\xi_0}\|_{Y}\|w_{\xi_0}\|_{Y}.$$

We use a modulation decomposition of v_{ξ_0} and $w_{-\xi_0}$ at modulation $\xi^2/8$. Due to the Fourier localization, the integral corresponding to the low modulation parts vanishes,

$$\int u^\varepsilon_{0,\le 1} M_{<\xi_0^2/4} v_{\xi_0} M_{<\xi_0^2/4} w_{-\xi_0}\,dx\,dt = 0.$$

On the other hand, we use Proposition 11 to estimate

$$\begin{aligned}\left|\int u^\varepsilon_{0,\le 1} M_{>\xi_0^2/4} v_{\xi_0} w_{-\xi_0}\,dx\,dt\right| &\lesssim \|u^\varepsilon_{0,\le 1}\|_{L^\infty}\|M_{>\xi_0^2/4} v_{\xi_0}\|_{L^2}\|w_{-\xi_0}\|_{L^2}\\ &\lesssim |\xi_0|^{-1}\|u^\varepsilon_{0,\le 1}\|_{L^\infty}\|v_{\xi_0}\|_{X^0}\|w_{-\xi_0}\|_{X^0}.\end{aligned}$$ □

Proof of Proposition 4. This is a direct consequence of the estimates in (8) and (7). □

[3] This is because we do not differentiate between frequencies 1 and less than 1 in what follows.

6. The bilinear estimate

Here we prove Theorem 4. After a Littlewood–Paley decomposition it suffices to prove a bound for the high-low frequency interactions

$$\|u_\mu v_\lambda\|_{\mathcal{D}Y_\lambda} \lesssim \lambda^{-1}\mu^{n/2}(\ln \mu)^{1/2}\|u_\mu\|_{\mathcal{D}X_\mu}\|v_\lambda\|_{\mathcal{D}X_\lambda}, \quad 1 \le \mu \ll \lambda, \tag{55}$$

as well as[4] for high-high frequency interactions

$$\|S_\mu(u_\lambda v_\lambda)\|_{\mathcal{D}Y_\mu} \lesssim \lambda^{n-1}\mu^{-n/2}\|u_\lambda\|_{\mathcal{D}X_\lambda}\|v_\lambda\|_{\mathcal{D}X_\lambda}, \quad 1 \le \mu \lesssim \lambda, \tag{56}$$

and similar bounds where one or both of the factors are replaced by their complex conjugates.

Both bounds can be localized on the λ spatial scale. Thus the decay structures in the two $\mathcal{D}X_\lambda$ spaces can be factored out in the first bound, and neglected in the second. Then, by the duality result in Proposition 9, (55) can be rewritten as

$$\left|\int u_\mu v_\lambda \overline{w}_\lambda \, dx\, dt\right| \lesssim \lambda^{-1}\mu^{n/2}(\ln \mu)^{1/2}\|u_\mu\|_{\mathcal{D}X_\mu}\|v_\lambda\|_{X_\lambda}\|w_\lambda\|_{X_\lambda}. \tag{57}$$

On the other hand, in (56) we can replace the l^1 summation by an l^2 summation on the μ spatial scale by losing a $\lambda^{1/2}\mu^{-1/2}$ factor. Thus it remains to show that

$$\|S_\mu(u_\lambda v_\lambda)\|_{Y_\mu} \lesssim \lambda^{n-3/2}\mu^{-(n-1)/2}\|u_\lambda\|_{X_\lambda}\|v_\lambda\|_{X_\lambda}.$$

By the duality result in Proposition 9 this is equivalent to

$$\left|\int u_\lambda v_\lambda \overline{w}_\mu \, dx\, dt\right| \lesssim \lambda^{n-3/2}\mu^{-(n-1)/2}\|u_\lambda\|_{X_\lambda}\|v_\lambda\|_{X_\lambda}\|w_\mu\|_{X_\mu},$$

which is easily seen to be weaker than (57).

It remains to prove (57) where we allow $1 \le \mu \le \lambda$ in order to include both cases above, and where we allow any combination of complex conjugates. The seemingly large number of cases is reduced by observing that the bound rests unchanged if we conjugate the entire product. Hence we can assume without any restriction in generality that at most one factor is conjugated. Hence it suffices to consider the following three cases:

(i) The product $u_\mu v_\lambda \overline{w}_\lambda$ with $1 \le \mu \le \lambda$. This is the main case, where all three factors can simultaneously concentrate in frequency near the parabola.
(ii) The product $\overline{u}_\mu v_\lambda w_\lambda$ with $1 \le \mu \ll \lambda$. Because the high frequency factors cannot simultaneously concentrate on the parabola, the estimate turns essentially into a bilinear L^2 estimate.
(iii) The product $u_\mu v_\lambda w_\lambda$ with $1 \le \mu \lesssim \lambda$. This is very similar to the second case.

Case 1: Here we prove (57) exactly as stated, for $1 \le \mu \le \lambda$. This follows by summation with respect to ξ, η in the following result:

[4] Strictly speaking, we should consider products of the form $u_{\lambda_1} v_{\lambda_2}$ with $\lambda_1 \approx \lambda_2$, but this makes no difference.

Proposition 16. (a) *Let* $1 \leq \mu \leq \lambda$ *and* $\xi, \eta \in \mathbb{R}^n$ *with*

$$|\xi| \approx \lambda, \quad |\eta| \approx \mu, \quad |\xi + \eta| \approx \lambda.$$

Then

$$\left| \int S_\eta u S_\xi v \overline{S_{\xi+\eta} w} \, dx \, dt \right| \lesssim \frac{\|S_\eta u\|_{X_\eta} \|S_\xi v\|_{X_\xi} \|S_{\xi+\eta} w\|_{X_{\xi+\eta}}}{\mu^{1/2}(\lambda + \min\{|\xi \cdot \eta|, |\xi \wedge \eta|\})^{1/2}}. \tag{58}$$

(b) *Assume in addition that* $\mu \lesssim \lambda^{1/2}$. *Then*

$$\left| \int S_\eta u S_\xi v \overline{S_{\xi+\eta} w} \, dx \, dt \right| \lesssim \frac{\|S_\eta u\|_{\mathcal{D}_i X_\eta} \|S_\xi v\|_{X_\xi} \|S_{\xi+\eta} w\|_{X_{\xi+\eta}}}{\lambda^{1/2}\mu^{-1/2}(\lambda + |\xi \cdot \eta|)^{1/2}}. \tag{59}$$

We first show how to use the proposition to conclude the proof of (57). We decompose u_μ, v_λ and w_λ in unit frequency cubes,

$$u_\mu = \sum_{|\eta| \approx \mu} S_\eta u_\mu, \quad v_\lambda = \sum_{|\xi| \approx \lambda} S_\xi v_\lambda, \quad w_\lambda = \sum_{|\zeta| \approx \lambda} S_\zeta w_\lambda,$$

and use the corresponding decomposition of the integral in (57).

We consider two cases. If $\mu \geq \lambda^{1/2}$ then we use (58) to estimate

$$\left| \int u_\mu v_\lambda \overline{w}_\lambda \, dx \, dt \right| \lesssim \sum_{|\xi| \approx \lambda, \, |\eta| \approx \mu}^{|\xi+\eta| \approx \lambda} \frac{\|S_\eta u_\mu\|_{X_\eta} \|S_\xi v_\lambda\|_{X_\xi} \|S_{\xi+\eta} w_\lambda\|_{X_{\xi+\eta}}}{\mu^{1/2}(\lambda + \min\{|\xi \cdot \eta|, |\xi \wedge \eta|\})^{1/2}}.$$

By the Cauchy–Schwarz inequality this is bounded by

$$\left(\sum_{|\xi| \approx \lambda, \, |\eta| \approx \mu}^{|\xi+\eta| \approx \lambda} \|S_\eta u_\mu\|_{X_\eta}^2 \|S_{\xi+\eta} w_\lambda\|_{X_{\xi+\eta}}^2 \right)^{1/2} \left(\sum_{|\xi| \approx \lambda, \, |\eta| \approx \mu}^{|\xi+\eta| \approx \lambda} \frac{\|S_\xi v_\lambda\|_{X_\xi}^2}{\mu(\lambda + \min\{|\xi \cdot \eta|, |\xi \wedge \eta|\})} \right)^{1/2}$$

and further by

$$\|u_\mu\|_{X_\mu} \|v_\lambda\|_{X_\lambda} \|w_\lambda\|_{X_\lambda} \left(\sup_{|\xi| \approx \lambda} \sum_{|\eta| \approx \mu}^{|\xi+\eta| \approx \lambda} \frac{1}{\mu(\lambda + \min\{|\xi \cdot \eta|, |\xi \wedge \eta|\})} \right)^{1/2},$$

which gives (57) since

$$\sum_{|\eta| \approx \mu}^{|\xi+\eta| \approx \lambda} \frac{1}{\mu(\lambda + \min\{|\xi \cdot \eta|, |\xi \wedge \eta|\})} \approx \frac{\mu^n \ln \mu}{\mu^2 \lambda}$$

and $\mu^2 \geq \lambda$. We note that the bound improves as μ increases.

If $\mu \leq \lambda^{1/2}$ then the argument is similar but using (59) instead of (58). This concludes the proof of (57).

Proof of Proposition 16. We decompose each of the factors in wave packets,

$$S_\eta u = \sum_{P \in T_\eta} u_P, \quad S_\xi v = \sum_{Q \in T_\xi} v_Q, \quad S_{\xi+\eta} w = \sum_{R \in T_{\xi+\eta}} w_R.$$

We first prove a bound with Q, R fixed.

Lemma 5. *For ξ and η as above the following estimate holds:*

$$\left| \int \sum_{P \in T_\eta} u_P v_Q \overline{w}_R \, dx \right| \lesssim \Big(\sum_{P \in T_\eta}^{P \cap Q \cap R \neq \emptyset} \|u_P\|_{X_\eta}^2 \Big)^{1/2} \frac{\|v_Q\|_{X_\xi} \|w_R\|_{X_{\xi+\eta}}}{\mu^{1/2}(\lambda + |\xi \cdot \eta|)^{1/2}}, \tag{60}$$

Proof. By Proposition 7 we can assume without any restriction in generality that

$$u_P = a_P(t) e^{-it\eta^2} \chi_P(x - 2t\eta), \qquad v_Q = a_Q(t) e^{-it\xi^2} \chi_Q(x - 2t\xi),$$
$$\overline{w}_R = a_R(t) e^{it(\xi+\eta)^2} \chi_R(x - 2t(\xi + \eta))$$

with the χ's being unit bumps and the a's in V^2. Then the integral has the form

$$\int_{-1}^{1} e^{4it\xi\eta} a_Q(t) a_R(t) \sum_P a_P(t) \int_{\mathbb{R}^n} \chi_P(x - 2t\eta) \chi_Q(x - 2t\xi) \chi_R(x - 2t(\xi + \eta)) \, dx \, dt.$$

The tubes Q and R differ in speed by η, therefore they intersect in a time interval I of length at most μ^{-1}. The tubes P and R differ in speed by ξ, therefore they intersect in a time interval of length at most λ^{-1}. Thus there are about $\lambda\mu^{-1}$ tubes P which intersect both Q and R.

For fixed P the x integral above is a smooth bump function on a λ^{-1} interval. Thus we can express the above integral in the form

$$\int_I e^{4it\xi\eta} a_Q(t) a_R(t) \sum_P a_P(t) b_P(t) \, dt,$$

where b_P are smooth bump functions on essentially disjoint λ^{-1} intervals. For the sum with respect to P we can estimate

$$\Big\| \sum_P a_P(t) b_P(t) \Big\|_{L^2}^2 \lesssim \lambda^{-1} \sum \|a_P\|_{L^\infty}^2$$

and

$$\Big\| \sum_P a_P(t) b_P(t) \Big\|_{V^2}^2 \lesssim \sum \|a_P\|_{V^2}^2.$$

We consider two possibilities. If $|\xi \cdot \eta| \lesssim \lambda$ then we use Hölder's inequality to bound the integral by

$$\mu^{-1/2} \lambda^{-1/2} \sum_P \|a_P\|_{L^\infty} \|a_Q\|_{L^\infty} \|a_R\|_{L^\infty}.$$

If $|\xi \cdot \eta| \gtrsim \lambda$ then we use the algebra property for V^2. It remains to prove that

$$\left| \int_0^{\mu^{-1}} a(t) e^{it\sigma} \, dt \right| \lesssim \mu^{-1/2} |\sigma|^{-1/2} \|a\|_{V^2}, \qquad \sigma = 4\xi\eta.$$

After rescaling this becomes

$$\left| \int_0^1 a(t) e^{it\sigma} \, dt \right| \lesssim |\sigma|^{-1/2} \|a\|_{V^2}.$$

This follows by Hölder's inequality from the bound

$$\|S_{\geq \sigma} a(t)\|_{L^2} \lesssim \sigma^{-1/2} \|a\|_{V^2}. \qquad \square$$

Now we prove part (a) of the proposition. Three tubes P, Q, R contribute to the integral only if they intersect. We consider the intersection pattern of ξ and $\xi + \eta$ tubes. For any ξ tube Q, all $\xi + \eta$ tubes intersecting it are contained in a larger slab obtained by horizontally translating Q in the η direction,

$$H = 2(Q + \{0\} \times [-2\eta, 2\eta]).$$

We denote by $\mathcal{H}$ a locally finite covering of $[-1, 1] \times \mathbb{R}^n$ with such slabs. Then we use the lemma to bound the integral in (58) by

$$\mu^{-1/2}(\lambda + |\xi \cdot \eta|)^{-1/2} \sum_{H \in \mathcal{H}} \sum_{Q, R \subset H} \|u_Q\|_{X_\xi} \|u_R\|_{X_{\xi+\eta}} \Big(\sum_{P \in T_\eta}^{P \cap Q \cap R \neq \emptyset} \|u_P\|_{X_\eta}^2 \Big)^{1/2}.$$

Using Cauchy–Schwarz in the second sum with respect to (Q, R) we bound this by

$$\frac{N^{1/2}}{\mu^{1/2}(\lambda + |\xi \cdot \eta|)^{1/2}} \Big(\sum_{P \in T_\eta} \|u_P\|_{X_\eta}^2 \Big)^{1/2} \sum_{H \in \mathcal{H}} \Big(\sum_{Q \subset H} \|u_Q\|_{X_\xi}^2 \Big)^{1/2} \Big(\sum_{R \subset H} \|u_R\|_{X_{\xi+\eta}}^2 \Big)^{1/2},$$

where

$$N = \max_{H, P} |\{(Q, R) : Q, R \subset H, \ P \cap Q \cap R \neq \emptyset\}|.$$

An additional Cauchy–Schwarz allows us to estimate the above sum by

$$\frac{N^{1/2}}{\mu^{1/2}(\lambda + |\xi \cdot \eta|)^{1/2}} \Big(\sum_{P \in T_\eta} \|u_P\|_{X_\eta}^2 \Big)^{1/2} \Big(\sum_{Q} \|u_Q\|_{X_\xi}^2 \Big)^{1/2} \Big(\sum_{R} \|u_R\|_{X_{\xi+\eta}}^2 \Big)^{1/2}.$$

To conclude the proof it remains to establish a bound for N, namely

$$N \lesssim \frac{|\xi| \, |\eta|}{|\xi| + |\xi \wedge \eta|}.$$

Both $P \cap R$ and $P \cap Q$ have a λ^{-1} time length and are uniquely determined by this intersection up to finite multiplicity. It follows that

$$N \lesssim \lambda |I|,$$

where I is the time interval where H and P intersect. Given the definition of H it follows that I has the form

$$I = \{t : |x_0 + t\xi + s\eta| \leq 2 \text{ for some } s \in [-1, 1]\}.$$

Symmetrizing we can assume that $x_0 = 0$. Taking inner and wedge products with η it follows that t must satisfy

$$|t| \leq |\eta|^2 |\xi \cdot \eta|^{-1}, \qquad |t| \leq |\eta|\, |\xi \wedge \eta|^{-1},$$

which lead to the desired bound for N.

(b) We first note that both Q and R are contained in spatial strips of size $\mu \times \lambda$ oriented in the ξ direction. Hence by orthogonality it suffices to prove the estimate in a single such strip. Then we can take advantage of the l^1 summability in the $\mathcal{D}X_\mu$ norm to further reduce the estimate to the case when the η tubes are spatially concentrated in a single $\mu \times \mu$ cube Z.

A ξ or a $\xi + \eta$ tube needs a time of $\mu\lambda^{-1}$ to move through such a cube. On the other hand, the tubes Q and R need a larger time μ^{-1} to separate. Hence within Z we can identify the ξ tubes and the $\xi + \eta$ tubes. By orthogonality it suffices to consider a single Q and a single R. Consequently, the conclusion follows from the following counterpart of Lemma 5.

Lemma 6. *For ξ and η as in part* (b) *of the proposition the following estimate holds:*

$$\left| \int \sum_{P \in T_\eta}^{P \subset Z} u_P v_Q w_R \, dx \right| \lesssim \Big(\sum_{P \in T_\eta}^{P \cap Q \cap R \neq \emptyset} \|u_P\|_{X_\eta}^2 \Big)^{1/2} \frac{\|v_Q\|_{X_\xi} \|w_R\|_{X_{\xi+\eta}}}{\lambda^{1/2}\mu^{-1/2}(\lambda + |\xi \cdot \eta|)^{1/2}}. \tag{61}$$

The proof is almost identical with the proof of Lemma 5, the only difference is that we now have $|I| \approx \mu\lambda^{-1}$.

Case 2: Here we consider the product $\overline{u}_\mu v_\lambda w_\lambda$ with $1 \leq \mu \ll \lambda$, and prove a stronger bound, namely

$$\left| \int \overline{u}_\mu v_\lambda w_\lambda \, dx \, dt \right| \lesssim \lambda^{-1} \mu^{n/2-1} \|u_\mu\|_{X_\mu} \|v_\lambda\|_{X_\lambda} \|w_\lambda\|_{X_\lambda}. \tag{62}$$

We use the modulation localization operators $S_{<\lambda^2/100}$ to split each of the factors in two,

$$u_\mu = M_{<\lambda^2/100} u_\mu + (1 - M_{<\lambda^2/100} u_\mu)$$

etc. We observe that

$$\int \overline{M_{<\lambda^2/100} u_\mu} M_{<\lambda^2/100} v_\lambda M_{<\lambda^2/100} w_\lambda \, dx \, dt = 0$$

due to the time frequency localizations. Precisely, the first factor is frequency localized in the region $\{|\tau| < \lambda^2/50\}$, while the other two are frequency localized in the region $\{\tau > \lambda^2/8\}$. Hence it remains to consider the case when at least one factor has high modulation. For that factor we have a favorable L^2 bound as in Proposition 11,

$$\|M_{>\lambda^2/100} u_\mu\|_{L^2} \lesssim \lambda^{-1} \|u_\mu\|_{X_\mu},$$

and similarly for the other factors. Then it remains to prove that

$$\|\bar{u}_\mu v_\lambda\|_{L^2} \lesssim \mu^{n/2-1} \|u_\mu\|_{\mathcal{D}X_\mu} \|v_\lambda\|_{X_\lambda}$$

and

$$\|S_\mu(v_\lambda w_\lambda)\|_{L^2} \lesssim \mu^{n/2-1}\|v_\lambda\|_{X_\lambda}\|w_\lambda\|_{X_\lambda}.$$

For this we need the Strichartz estimates in Proposition 10 (see also (32)).

For the first bound we use the $L^{2(n+2)/n}$ estimate for v_λ, and the L^{n+2} estimate for $\bar{u}_\mu$.

For the second we first observe that by orthogonality it suffices to prove it when both v_λ and w_λ are frequency localized to cubes of size μ. Then we use the $L^{2(n+2)/n}$ estimate for both factors to derive L^4 bounds by Sobolev embeddings.

Case 3: Here we consider the product $u_\mu v_\lambda w_\lambda$ with $1 \le \mu \lesssim \lambda$. This is treated exactly as above, and an estimate similar to (62) is obtained. □

Acknowledgments. The first author was partially supported by MSRI for Fall 2005. The second author was partially supported by NSF grants DMS0354539 and DMS 0301122 and also by MSRI for Fall 2005.

References

[1] Bejenaru, I.: Quadratic nonlinear derivative Schrödinger equations—Part 1. Int. Math. Res. Papers **2006**, Art. ID 70630, 84 pp. Zbl 1123.35060 MR 2235496

[2] Bejenaru, I.: Quadratic nonlinear derivative Schrödinger equations—Part 2. Trans. Amer. Math. Soc., to appear

[3] Chihara, H.: Gain of regularity for semilinear Schrödinger equations. Math. Ann. **315**, 529–567 (1999) Zbl 0942.35161 MR 1731461

[4] Cohn, S.: Global existence for the nonresonant Schrdinger equation in two space dimensions. Canad. Appl. Math. Quart. **2**, 257–282 (1994) Zbl 0853.35113 MR 1318545

[5] Cohn, S.: Resonance and long time existence for the quadratic semilinear Schrödinger equation. Comm. Pure Appl. Math. **45**, 973–1001 (1992) Zbl 0774.34063 MR 1168116

[6] Grünrock, A.: On the Cauchy- and periodic boundary value problem for a certain class of derivative nonlinear Schrödinger equations. arXiv:math/0006195

[7] Hayashi, N., Naumkin, P.: A quadratic nonlinear Schrödinger equation in one space dimension. J. Differential Equations **186**, 165–185 (2002) MR 1941097

[8] Kenig, C. E., Ponce, G., Vega, L.: Smoothing effects and local existence theory for the generalized nonlinear Schrödinger equations. Invent. Math. **134**, 489–545 (1998) Zbl 0928.35158 MR 1660933

[9] Koch, H., Tataru, D.: Dispersive estimates for principally normal pseudodifferential operators. Comm. Pure Appl. Math. **58**, 217–284 (2005) Zbl 1078.35143 MR 2094851

[10] Mizohata, S.: On the Cauchy Problem. Notes and Reports in Mathematics in Science and Engineering 3, Science Press & Academic Press (1985) Zbl 0616.35002 MR 0860041

[11] Smith, H., Sogge, C.: On the L^p norm of spectral clusters for compact manifolds with boundary. Acta Math. **198**, 107–153 (2007) Zbl pre05166605 MR 2316270

[12] Stefanov, A.: On first-order perturbations of the Schrödinger equation with conjugation. Differential Integral Equations **18**, 997–1012 (2005) MR 2162984

[13] Taylor, M. E.: Partial Differential Equations II. Appl. Math. Sci. 116, Springer (1996) Zbl 0869.35003 MR 1395149