Title
Quasilinear Schrödinger equations, II: Small data and cubic nonlinearities

Journal
Kyoto Journal of Mathematics, 54(3)

ISSN
2154-3321

Authors
Marzuola, Jeremy L
Metcalfe, Jason
Tataru, Daniel

Publication Date
2014

DOI
10.1215/21562261-2693424

Peer reviewed

QUASILINEAR SCHRÖDINGER EQUATIONS II: SMALL DATA AND CUBIC NONLINEARITIES

JEREMY L. MARZUOLA, JASON METCALFE, AND DANIEL TATARU

Abstract. In part I of this project we examined low regularity local well-posedness for generic quasilinear Schrödinger equations with small data. This improved, in the small data regime, the preceding results of Kenig, Ponce, and Vega as well as Kenig, Ponce, Rolvung, and Vega. In the setting of quadratic interactions, the (translation invariant) function spaces which were utilized incorporated an l^1 summability over cubes in order to account for Mizohata's integrability condition, which is a necessary condition for the L^2 well-posedness for the linearized equation. For cubic interactions, this integrability condition meshes better with the inherent L^2 nature of the Schrödinger equation, and such summability is not required. Thus we are able to prove small data well-posedness in H^s spaces.

1. Introduction

We shall examine local well-posedness for quasilinear Schrödinger equations with cubic interactions and a Cauchy datum in a low regularity Sobolev space. In particular, we examine

$$\begin{cases} i\partial_t u + g^{jk}(u, \nabla u)\partial_j\partial_k u = F(u, \nabla u), \quad u : \mathbb{R}\times\mathbb{R}^d \to \mathbb{C}^m \\ u(0, x) = u_0(x) \end{cases} \tag{1.1}$$

where

$$g : \mathbb{C}^m \times (\mathbb{C}^m)^d \to \mathbb{R}^{d\times d}, \qquad F : \mathbb{C}^m \times (\mathbb{C}^m)^d \to \mathbb{C}^m$$

are smooth functions which satisfy

$$g(y, z) = I_d + O(|y|^2 + |z|^2), \qquad F(y, z) = O(|y|^3 + |z|^3) \text{ near } (y, z) = (0, 0).$$

We also examine

$$\begin{cases} i\partial_t u + \partial_j g^{jk}(u)\partial_k u = F(u, \nabla u), \ u : \mathbb{R}\times\mathbb{R}^d \to \mathbb{C}^m \\ u(0, x) = u_0(x), \end{cases} \tag{1.2}$$

where g depends only on u. The latter class of equations can be obtained by differentiating the first equation. Indeed, for u solving (1.1), the vector $(u, \nabla u)$ solves an equation of the latter type. The latter is written in divergence form, which follows easily for this class of equations as the terms obtained when commuting g with the derivative can be absorbed into $F(u, \nabla u)$.

The case of small initial data and quadratic, rather than cubic, nonlinear terms, was considered in [10]. There, local well-posedness was established for data of Sobolev-type regularity $s > \frac{d}{2} + 3$ for (1.1) and $s > \frac{d}{2} + 2$ for (1.2). This represented a significant improvement in regularity over the previous results [4], [5, 6], though data of arbitrary size was examined there. A more complete history of such problems can be found in [10], [4], and [8].

In the quadratic case above, it is insufficient, however, to simply work in Sobolev spaces. Indeed, one encounters Mizohata's integrability condition [11, 12, 13], [7], [9] which says that for the linear equation

$$(i\partial_t + \Delta_g)v = A_i(x)\partial_i v,$$

a necessary condition for L^2 well-posedness is an integrability of the real part of A along the Hamiltonian flow of the leading order operator. The potentials encountered when linearizing, e.g., (1.1) with quadratic interactions do not necessarily satisfy such a condition even with arbitrarily high regularities. For this reason, the initial data spaces need to incorporate some decay. The translation invariant approach of [10] was to require a summability over cubes, which was inspired by the earlier work [1] on semilinear derivative Schrödinger equations.

In the case of only cubic and higher order interactions, the scenario is much simpler, which is what we shall demonstrate here. When linearizing (1.1) as stated, the potential that is encountered is $O(|u|^2 + |\nabla u|^2)$, for which integrability follows easily from energy estimates. As such, the additional l^1 summability which was previously required is no longer needed, and the initial datum can be simply taken to be a member of a Sobolev space. This notion was previously explored in [2, 3], and we seek to improve on that by considering rough initial data.

Our main result is precisely that (1.1) and (1.2) are locally well-posed for $u_0(x) \in H^s$.

Theorem 1. *a) Let $s > \frac{d+5}{2}$. Then there exists $\epsilon > 0$ sufficiently small such that for all initial data u_0 with*

$$\|u_0\|_{H^s} \leq \epsilon,$$

the equation (1.1) is locally well-posed in $H^s(\mathbb{R}^d)$ on the time interval $I = [0, 1]$.

b) The same result holds for the equation (1.2) with $s > \frac{d+3}{2}$.

In the theorem, well-posedness is taken to include the existence of a local solution, uniqueness, and continuous dependence on the initial datum.

The above theorem also holds for ultrahyperbolic operators, as in [5, 6], where $g(0)$ is of a different signature. Our method of proof of the local smoothing estimates uses a wedge decomposition, and the estimates are proved in each coordinate direction separately. Thus, trivial adjustments of the sign of the multiplier in the corresponding directions of opposite signature will permit said results.

This article is organized as follows. In Section 2, we set our notations and describe the function spaces in which we shall work. Section 3 is devoted to proving the required nonlinear estimates. Section 4 contains the proof of our main linear estimate, which is a variant of the classical local smoothing estimate for the Schrödinger equation. In Section 5 we prove Theorem 1.

Acknowledgments. The second author was supported in part by NSF grant DMS-1054289. The third author was supported in part by NSF grant DMS-0801261 as well as by the Miller Foundation.

2. Function Spaces and Notations

The function spaces and notations presented in this section are the same as those established in [10]. We repeat them for easy reference.

For a function $u(t, x)$ or $u(x)$, we let $\hat{u} = \mathcal{F}u$ denote the Fourier transform in the spatial variables x. A function u is said to be localized at frequency 2^i if $\hat{u}(t, \xi)$ is supported in $\mathbb{R} \times [2^{i-1}, 2^{i+1}]$. We shall use a Littlewood-Paley decomposition of the spatial frequencies,

$$\sum_{i=0}^{\infty} S_i(D) = 1,$$

where S_i localizes to frequency 2^i for $i > 0$ and to frequencies $|\xi| \leq 2$ for $i = 0$. Note that we are working in nonhomogeneous spaces. We set

$$u_i = S_i u, \quad S_{\leq N} u = \sum_{i=0}^{N} S_i u, \quad S_{\geq N} u = \sum_{i=N}^{\infty} S_i u.$$

For each $j \in \mathbb{N}$, we let $\mathcal{Q}_j$ denote a partition of $\mathbb{R}^d$ into cubes of side length 2^j, and we let $\{\chi_Q\}$ denote an associated partition of unity. For a translation invariant Sobolev-type space U, we set $l^p_j U$ to be the Banach space with associated norm

$$\|u\|^p_{l^p_j U} = \sum_{Q \in \mathcal{Q}_j} \|\chi_Q u\|^p_U$$

with the obvious modification for $p = \infty$.

Within such norms, the smooth partition of unity with compactly supported functions χ_Q can be replaced by cutoffs which are frequency localized as the associated Schwartz tails decay rapidly away from Q. Thus, the components can be taken to retain the frequency localization of the function u.

Motived by the well-known local smoothing estimates, we define

$$\|u\|_X = \sup_l \sup_{Q\in\mathcal{Q}_l} 2^{-\frac{l}{2}} \|u\|_{L^2L^2([0,1]\times Q)}.$$

Here and throughout, L^pL^q represents $L^p_tL^q_x$.

To measure the forcing terms, we use an atomic space Y satisfying $X = Y^*$, see [10]. A function a is an atom in Y if there is a $j \geq 0$ and a $Q \in \mathcal{Q}_j$ so that a is supported in $[0,1]\times Q$ and

$$\|a\|_{L^2([0,1]\times Q)} \lesssim 2^{-\frac{j}{2}}.$$

Then we define Y as linear combinations of the form

$$f = \sum_k c_k a_k, \quad \sum |c_k| < \infty, \quad a_k \text{ atoms}$$

with norm

$$\|f\|_Y = \inf\{\sum |c_k| \,:\, f = \sum_k c_k a_k,\, a_k \text{ atoms}\}.$$

For solutions which are localized to frequency 2^j, we shall encorporate the typical half derivative smoothing by working within

$$X_j = 2^{-\frac{j}{2}} X \cap L^\infty L^2$$

with norm

$$\|u\|_{X_j} = 2^{\frac{j}{2}} \|u\|_X + \|u\|_{L^\infty L^2}.$$

We incorporate the Sobolev regularity and the cube summability by defining

$$\|u\|^2_{l^pX^s} = \sum_{j\geq 0} 2^{2js} \|S_j u\|^2_{l^p_j X_j}.$$

For the purposes of this paper, we will be working primarily in l^2X^s, as opposed to our previous paper [10] where we had to work in l^1X^s due to incompatibilities with the Mizohata condition. On occasion we will also use the slightly larger space

$$\|u\|^2_{X^s} = \sum_j 2^{2js} \|S_j u\|^2_{X_j}.$$

We analogously define

$$Y_j = 2^{\frac{j}{2}} Y + L^1L^2$$

which has norm

$$\|f\|_{Y_j} = \inf_{f=2^{\frac{j}{2}}f_1+f_2} \|f_1\|_Y + \|f_2\|_{L^1L^2}$$

and

$$\|f\|^2_{l^pY^s} = \sum_j 2^{2js}\|S_jf\|^2_{l^p_jY_j}.$$

Here, we shall be working primarily within l^2Y^s, though we note that the Minkowski integral inequality gives

$$\|f\|_{l^2Y^s} \lesssim \|f\|_{Y^s}$$

where

$$\|f\|^2_{Y^s} = \sum_j 2^{2js}\|S_jf\|^2_{Y_j}.$$

Hence, the cube summation will largely be ignored for the forcing terms.

We also note that for any j, we have

$$\sup_{Q\in\mathcal{Q}_j} 2^{-\frac{j}{2}}\|u\|_{L^2L^2([0,1]\times Q)} \leq \|u\|_X,$$

hence

$$\|u\|_Y \lesssim 2^{\frac{j}{2}}\|u\|_{l^1_jL^2L^2}. \tag{2.1}$$

This bound will come in handy at several places later on.

In our multilinear estimates, we shall use frequency envelopes. Consider a Sobolev-type space U for which we have

$$\|u\|^2_U \sim \sum_{k=0}^{\infty} \|S_ku\|^2_U.$$

A frequency envelope for a function $u \in U$ is a positive l^2 sequence, $\{a_j\}$, with

$$\|S_ju\|_U \leq a_j\|u\|_U.$$

We shall only permit slowly varying frequency envelopes. Thus, we require $a_0 \approx 1$ and

$$a_j \leq 2^{\delta|j-k|}a_k, \quad j,k \geq 0, \quad 0 < \delta \ll 1. \tag{2.2}$$

The constant δ shall be chosen later and only depends on s and the dimension d. Such frequency envelopes always exist. E.g., one may choose

$$a_j = 2^{-\delta j} + \|u\|_U^{-1} \sup_k 2^{-\delta|j-k|}\|S_ku\|_U. \tag{2.3}$$

3. Multilinear and nonlinear estimates

This section contains our main multilinear estimates. The first proposition is essentially from [9], though it is somewhat simpler as the summability over cubes is easier.

Proposition 3.1. *Let $\sigma > \frac{d}{2}$, and $u, v \in l^2 X^\sigma$ with admissible frequency envelopes a_k, respectively b_k. Then the $l^2 X^\sigma$ spaces satisfy the algebra type property*

$$\|S_k(uv)\|_{l^2X^\sigma} \lesssim (a_k + b_k)\|u\|_{l^2X^\sigma}\|v\|_{l^2X^\sigma}, \tag{3.1}$$

as well as the Moser type estimate

$$\|S_kF(u)\|_{l^2X^\sigma} \lesssim a_k\|u\|_{l^2X^\sigma}(1 + \|u\|_{l^2X^\sigma})c(\|u\|_{L^\infty}). \tag{3.2}$$

for all smooth F with $F(0) = 0$.

Proof. To prove (3.1), we consider $\|S_k(S_iuS_jv)\|_{l^2_kX_k}$. We must have either $i \geq k-4$ or $j \geq k-4$; by symmetry we shall assume the former. Since $i \geq k-4$, we have $X_i \subset X_k$. And we naively use the fact that there are $\approx 2^{d(i-k)}$ cubes of sidelength 2^k contained in one of sidelength 2^i. Thus, by Bernstein's inequality, we have

$$\begin{aligned} \|S_k(S_iuS_jv)\|_{l^2_kX_k} &\lesssim 2^{\frac{d}{2}(i-k)}\|S_iu\|_{l^2_iX_i}\|S_jv\|_{L^\infty} \\ &\lesssim 2^{\frac{d}{2}(i-k)}2^{j\frac{d}{2}}\|S_iu\|_{l^2_iX_i}\|S_jv\|_{l^2_jL^\infty L^2}. \end{aligned} \tag{3.3}$$

It follows that

$$\|S_k(S_iuS_jv)\|_{l^2X^\sigma} \lesssim a_ib_j2^{i(\frac{d}{2}-\sigma)}2^{j(\frac{d}{2}-\sigma)}2^{k(\sigma-\frac{d}{2})}\|u\|_{l^2X^\sigma}\|v\|_{l^2X^\sigma}.$$

As $\sigma > \frac{d}{2}$, the estimate (3.1) now follows after summation in i, j. When $k-4 \leq i \leq k+4$, we simply use that the Cauchy-Schwarz inequality implies that $\sum_j 2^{j(\frac{d}{2}-\sigma)}b_j$ is $O(1)$. When $i > k+4$, we also use that $\sum_{i>k} 2^{i(\frac{d}{2}-\sigma)}a_i \lesssim 2^{k(\frac{d}{2}-\sigma)}a_k$ provided the δ in (2.2) satisfies $0 < \delta < \sigma - \frac{d}{2}$.

We now examine (3.2) and replace the discrete Littlewood-Paley decomposition by a continuous one

$$Id = S_0 + \int_0^\infty S_k\,dk.$$

Abusing notation, we will set $u_0 = S_0u$ for now. Then, using the Fundamental Theorem of Calculus, we can write

$$S_kF(u) = S_kF(u_0) + \int_0^\infty S_k(u_{k_1}F'(u_{<k_1}))\,dk_1. \tag{3.4}$$

For the first term of (3.4), we note that

$$\|\partial^\alpha F(u_0)\|_{l_0^2 X_0} \lesssim \|u_0\|_{l_0^2 X_0} c(\|u_0\|_{L^\infty}).$$

Thus,

$$\|S_k F(u_0)\|_{l_k^2 X_k} \lesssim 2^{k/2} \|S_k F(u_0)\|_{l_0^2 X_0} \lesssim 2^{-Nk} \|u_0\|_{l_0^2 X_0} c(\|u_0\|_{L^\infty})$$

for any N, from which the $l^2 X^\sigma$ bound follows easily.

We split the integral in the second term of (3.4) into two regions. For $k_1 \geq k-4$, we use that $X_{k_1} \subset X_k$ and have

$$\begin{aligned} 2^{k\sigma} \Big\| \int_{k-4}^\infty & S_k(u_{k_1} F'(u_{<k_1}))\, dk_1 \Big\|_{l_k^2 X_k} \\ & \lesssim \int_{k-4}^\infty 2^{-(k_1-k)(\sigma - \frac{d}{2})} \|u_{k_1}\|_{l^2 X^\sigma} \|F'(u_{<k_1})\|_{L^\infty}\, dk_1 \end{aligned}$$

by using (3.3). From this, the desired bound follows if δ in (2.2) is chosen small enough. This case is significantly simpler here than its analogous case in [10] because we are now dealing with l^2 sums rather than l^1.

For $k_1 < k-4$, we note that

$$S_k(u_{k_1} F'(u_{<k_1})) = S_k(u_{k_1} \tilde{S}_k F'(u_{<k_1}))$$

where $\tilde{S}_k$ also localizes to frequency 2^k and $S_k \tilde{S}_k = S_k$. The chain rule allows us to estimate

$$\|\tilde{S}_k F'(u_{<k_1})\|_{L^\infty} \lesssim 2^{-N(k-k_1)} c(\|u_{<k_1}\|_{L^\infty}).$$

Thus,

$$\begin{aligned} \|S_k(u_{k_1} F'(u_{<k_1}))\|_{l_k^2 X_k} & \lesssim 2^{\frac{k-k_1}{2}} \|u_{k_1}\|_{l_{k_1}^2 X_{k_1}} \|\tilde{S}_k F'(u_{<k_1})\|_{L^\infty} \\ & \lesssim 2^{-N(k-k_1)} \|u_{k_1}\|_{l_{k_1}^2 X_{k_1}} c(\|u_{<k_1}\|_{L^\infty}). \end{aligned}$$

In turn, this leads to

$$\|S_k(u_{k_1} F'(u_{<k_1}))\|_{l^2 X^\sigma} \lesssim 2^{-N(k-k_1)} a_{k_1} \|u\|_{l^2 X^\sigma} c(\|u\|_{L^\infty}),$$

where as N was arbitrary, we allowed its value to change from line to line, and after integration, (3.2) is proved. □

Replacing the bilinear estimates which were used in [10] are the following trilinear estimates. In particular, we note that here we only require $s > \frac{d+3}{2}$, which accounts for the improvement in regularity as compared to the quadratic interactions explored in [10].

Proposition 3.2. *Let $s > \frac{d+3}{2}$, and suppose that $u \in l^2 X^{\sigma-1}$, $v, w \in l^2 X^{s-1}$ with frequency envelopes $\{a_k\}$, $\{b_k\}$, and $\{c_k\}$ respectively. Then for $0 \leq \sigma \leq s$, we have*

$$\|S_k(uvw)\|_{l^2 Y^\sigma} \lesssim (a_k + b_k + c_k)\|u\|_{X^{\sigma-1}}\|v\|_{l^2 X^{s-1}}\|w\|_{l^2 X^{s-1}}. \tag{3.5}$$

If $0 \leq \sigma \leq s-1$ and if $u \in l^2 X^\sigma$, $v \in l^2 X^{s-2}$ and $w \in l^2 X^{s-1}$, then

$$\|S_k(uvw)\|_{l^2 Y^\sigma} \lesssim (a_k + b_k + c_k)\|u\|_{l^2 X^\sigma}\|v\|_{l^2 X^{s-2}}\|w\|_{l^2 X^{s-1}}. \tag{3.6}$$

Finally, for $0 \leq \sigma \leq s$ and $u \in l^2 X^{\sigma-2}$ and $v, w \in l^2 X^s$, and

$$\|S_k(uS_{\geq k-4}(vw))\|_{l^2 Y^\sigma} \lesssim (a_k + b_k + c_k)\|u\|_{l^2 X^{\sigma-2}}\|v\|_{l^2 X^s}\|w\|_{l^2 X^s}. \tag{3.7}$$

Proof. For the proof we need to establish trilinear estimates of the form

$$\|S_k(u_i v_j w_l)\|_{l^2_k Y_k} \lesssim c_{kijl}\|u_i\|_{X_i}\|v_j\|_{l^2_j X_j}\|w_l\|_{l^2_l X_l}. \tag{3.8}$$

for several different cases of frequency balances.

Case A. $|i-k| \leq 4$ **and** $j, l \leq k+4$**.** We shall assume, without loss, that $j \geq l$. Here, we measure the output in Y, use X for the highest frequency factor, and $L^\infty L^2$ for the lower frequency factors. Using (2.1) we have

$$LHS(3.8) \lesssim 2^{-k/2}\|S_k(u_i v_j w_l)\|_Y \lesssim 2^{-k/2}2^{j/2}\|S_k(u_i v_j w_l)\|_{l^1_j L^2 L^2}.$$

By Bernstein's inequality and using that $l < j$, this yields

$$\begin{aligned} LHS(3.8) &\lesssim 2^{-k/2}2^{j/2}2^{jd/2}\|v_j\|_{l^2_j L^\infty L^2}\|u_i w_l\|_{l^2_l L^2 L^2} \\ &\lesssim 2^{-k/2}2^{j\frac{d+1}{2}}2^{l\frac{d+1}{2}}\|u_i\|_X\|v_j\|_{l^2_l L^\infty L^2}\|w_l\|_{l^2_l L^\infty L^2}, \end{aligned}$$

which in turn gives

$$c_{kijl} \lesssim 2^{-k/2}2^{-i/2}2^{l\frac{d+1}{2}}2^{j\frac{d+1}{2}}. \tag{3.9}$$

To prove (3.5) for this type of interactions we use (3.9) and the frequency envelopes of u, v, and w to bound

$$\begin{aligned} &\|S_k(u_i v_j w_l)\|_{l^2 Y^\sigma} \\ &\qquad \lesssim a_i 2^{j\left(\frac{d+3}{2}-s\right)} b_j 2^{l\left(\frac{d+3}{2}-s\right)} c_l \|u\|_{X^{\sigma-1}}\|v\|_{l^2 X^{s-1}}\|w\|_{l^2 X^{s-1}}, \end{aligned}$$

which is easily summed over i, j, l in the appropriate ranges. The case where the same frequency balance holds but with the roles of i, j and l interchanged is similar but simpler.

Now consider (3.6). The worst case is $i \geq j \geq l$ if $\sigma \leq s-2$, respectively $j \geq i \geq l$ if $s-2 < \sigma \leq s-1$. In both cases (3.6) is weaker than (3.5).

Finally, for the estimate (3.7) we must have either $j > k-8$ or $l > k-8$, in which case (3.7) is also weaker than (3.5).

Case B. $i, j > k-4$**,** $|i-j| < 4$**,** $l < i+8$**.** Grouping terms in the following manner

$$\left|\int \overline{z_k} u_i v_j w_l \, dx \, dt\right| \lesssim \|v_j\|_{l_j^2 X_j} \|z_k u_i w_l\|_{l_j^2 Y_j}$$

and in view of the duality relation $l_j^2 X_j = l_j^2 Y_j^*$, from the bound (3.9) we directly obtain

$$c_{ijkl} \lesssim 2^{-\frac{i}{2}} 2^{-\frac{j}{2}} 2^{\frac{d+1}{2}k} 2^{\frac{d+1}{2}l}. \tag{3.10}$$

Then the estimates (3.5), (3.6) and (3.7) follow again by passing to frequency envelopes and summing over the appropriate ranges of i, j, l. Here, again, we have handled explicitly the worst case, and those cases where the roles of i, j, l are interchanged are simpler.

We remark that in the case of (3.7) we need to deal with the additional multiplier $S_{>k-4}$. However, all our function spaces are translation invariant, and the kernel of $S_{>k-4}$ is a bounded measure. Thus it can easily be disposed of. □

We shall also utilize the following bound on commutators when, e.g., we decompose into functions whose Fourier transforms are supported in wedges about the coordinate axes. The proof closely resembles that of [10], though care is taken to permit $s > \frac{d+3}{2}$ rather than $s > \frac{d}{2} + 2$.

Proposition 3.3. *We assume*

$$g^{kl} - \delta^{kl} = h^{kl}(w(t,x))$$

where $h(z) = O(|z|^2)$ *near* $|z| = 0$. *For* $s > \frac{d+3}{2}$ *and any multiplier* $B \in S^0$ *we have*

$$\|\nabla[S_{<k-4} g, B(D)]\nabla S_k u\|_{Y_k} \lesssim \|w\|_{l^2 X^s}^2 (1 + \|w\|_{l^2 X^s}) c(\|w\|_{L^\infty}) \|S_k u\|_{X_k}. \tag{3.11}$$

Proof. In [10, Proposition 3.2], it was shown that

$$\nabla[S_{<k-4} g, B(D)]\nabla S_k u = L(\nabla S_{<k-4} g, \nabla S_k u),$$

where L is a translation invariant operator satisfying

$$L(f,g)(x) = \int f(x+y) g(x+z) w(y,z) \, dy \, dz, \quad w \in L^1.$$

We shall not reproduce this proof here.

Given this representation, as we are working in translation invariant spaces, the bound (3.11) follows immediately from (3.2) and (3.5). □

4. Local Energy Decay

Here we consider the main linear estimate which we shall employ in order to prove Theorem 1. This is a variant on the well-known local smoothing estimates for the linear Schrödinger equation. The metric is assumed to be a composition

$$g^{kl} - \delta^{kl} = h^{kl}(w(t,x)) \tag{4.1}$$

where $h(z) = O(|z|^2)$ near $|z| = 0$. The focus, then, is on a frequency localized linear Schrödinger equation

$$(i\partial_t + \partial_k g^{kl}_{<j-4}\partial_l)u_j = f_j, \qquad u_j(0) = u_{0j}. \tag{4.2}$$

In the sequel, we shall pass to this setting.

The main estimate is:

Proposition 4.1. *Assume that the coefficients g^{kl} in (4.2) satisfy (4.1) with*

$$\|w\|_{l^2X^s} \ll 1 \tag{4.3}$$

for some $s > \frac{d+3}{2}$. Let u_j be a solution to (4.2) which is localized at frequency 2^j. Then the following estimate holds:

$$\|u_j\|_{l^2_jX_j} \lesssim \|u_{0j}\|_{L^2} + \|f_j\|_{l^2_jY_j}. \tag{4.4}$$

The proof is closely related to that given in [10] and uses a positive commutator argument. The multipliers will be first order differential operators with smooth coefficients which are localized at frequency $\lesssim 1$; precisely, for estimating waves at frequency 2^j we will use multipliers $\mathcal{M}$ that are differential operators having the form

$$i2^j\mathcal{M} = a^k(x)\partial_k + \partial_k a^k(x) \tag{4.5}$$

with uniform bounds on a and its derivatives.

A key spot where the proof here differs from that of [10] is in the following lemma which is used to dismiss the $g - I$ contribution to the commutator.

Lemma 4.2. *Let $A = \partial_k S_{<j-4}(h^{kl}(w(t,x)))\partial_l$ with h as above, and let $\mathcal{M}$ be as above. Suppose that $s > \frac{d+3}{2}$. Then we have*

$$|\langle[A,\mathcal{M}]u_j, u_j\rangle| \lesssim \|w\|^2_{l^2X^s}(1 + \|w\|_{l^2X^s})c(\|w\|_{L^\infty})\|u_j\|^2_{X_j}. \tag{4.6}$$

Proof. We note that we can abstractly write $[A,\mathcal{M}]$ as

$$i[A,\mathcal{M}] = 2^{-j}(\nabla(h\nabla a + a\nabla h)\nabla + \nabla h\nabla^2 a + h\nabla^3 a)$$

where h abbreviates $S_{<j-4}(h(w(t,x)))$. As the a factors are bounded and low frequency, the worst term to bound is

$$2^{-j}\langle a\nabla h\nabla u_j, \nabla u_j\rangle.$$

By duality, we only require

$$\|(\nabla h)\nabla u_j\|_{Y_j} \lesssim \|w\|^2_{l^2X^s}(1+\|w\|_{l^2X^s})c(\|w\|_{L^\infty})\|u_j\|_{X_j},$$

which is a conseqeunce of (3.2) and (3.5). □

We now continue the proof of the proposition. For $Q \in \mathcal{Q}_l$, $0 \le l \le j$ fixed, we begin by proving that a solution u_j to (4.2) satisfies

(4.7)
$$2^{j-l}\|u_j\|^2_{L^2(Q)} \lesssim \ \|u_j\|^2_{L^\infty L^2} + \|u_j\|_{X_j}\|f_j\|_{Y_j} + (2^{-j} + \|w\|^2_{l^2X^s})\|u_j\|^2_{X_j}.$$

We start by abstracting and studying solutions to

$$(D_t + A)u = f, \quad u(0) = u_0$$

where A is a self-adjoint operator and $D_t = \frac{1}{i}\partial_t$. First of all, we have an energy-type estimate:

$$\|u\|^2_{L^\infty_t L^2_x} \lesssim \|u_0\|^2_{L^2} + \|u\|_{X_j}\|f\|_{Y_j}. \tag{4.8}$$

See [10, Lemma 4.2]. For a self-adjoint multiplier $\mathcal{M}$, we also have

$$\frac{d}{dt}\langle u, \mathcal{M}u\rangle = -2\,\mathrm{Im}\langle (D_t + A)u, \mathcal{M}u\rangle + \langle i[A,\mathcal{M}]u,u\rangle. \tag{4.9}$$

Thus, for $u = u_j$ and $A = -\partial_k g^{kl}_{<j-4}\partial_l$, if we can provide a multiplier $\mathcal{M}$ so that

(1) $\|\mathcal{M}u\|_{L^2_x} \lesssim \|u\|_{L^2_x}$,
(2) $\|\mathcal{M}u\|_X \lesssim \|u\|_X$,
(3) $i\langle [A,\mathcal{M}]u,u\rangle \gtrsim 2^{j-\ell}\|u\|^2_{L^2L^2([0,1]\times Q)} - O(2^{-j} + \|w\|^2_{l^2X^s})\|u\|^2_{X_j}$,

(4.7) would follow.

We first do this when the Fourier transform of the solution u_j is restricted to a small angle

$$\operatorname{supp} \hat{u}_j \subset \{|\xi| \lesssim \xi_1\}. \tag{4.10}$$

Here, we have done so about the first coordinate axis, though the modifications to obtain the same about the other axes are trivial. We take, without loss of generality due to translation invariance, $Q = \{|x_j| \le 2^l : j = 1,\dots,d\}$, and we set m to be a smooth, bounded, increasing function such that $m'(s) = \psi^2(s)$ where ψ is a Schwartz function localized at frequencies $\lesssim 1$, and $\psi \sim 1$ for $|s| \le 1$. We rescale m and set $m_l(s) = m(2^{-l}s)$. Then, we fix

$$\mathcal{M} = \frac{1}{i2^j}\Big(m_l(x_1)\partial_1 + \partial_1 m_l(x_1)\Big).$$

Properties (1) and (2) are immediate due to the frequency localizations of $u = u_j$ and m_l as well as the boundedness of m_l. By Lemma

4.2, it suffices to consider property (3) for $A = -\Delta$. This yields

$$i2^j[-\Delta, \mathcal{M}] = -2^{-l+2}\partial_1\psi^2(2^{-l}x_1)\partial_1 + O(1),$$

and hence

$$i2^j\langle[-\Delta, \mathcal{M}]u, u\rangle = 2^{-l+2}\|\psi(2^{-l}x_1)\partial_1 u_j\|^2_{L^2L^2} + O(\|u_j\|^2_{L^2L^2}).$$

Utilizing our assumption (4.10), it follows that

$$2^{-l}2^j\|\psi(2^{-l}x_1)u_j\|^2_{L^2L^2} \lesssim i\langle[-\Delta, \mathcal{M}]u_j, u_j\rangle + 2^{-j}O(\|u_j\|^2_{L^2L^2}),$$

which yields (3) when combined with Lemma 4.2.

We proceed to reducing to the the case that (4.10) holds. We let $\{\theta_k(\omega)\}_{k=1}^d$ be a partition of unity,

$$\sum_k \theta_k(\omega) = 1, \quad \omega \in \mathbb{S}^{d-1},$$

where $\theta_k(\omega)$ is supported in a small angle about the kth coordinate axis. Then, we can set $u_{j,k} = \Theta_{j,k}u_j$ where

$$\mathcal{F}\Theta_{j,k}u = \theta_k\Big(\frac{\xi}{|\xi|}\Big) \sum_{j-1\le l\le j+1} \phi_l(\xi)\hat{u}(t,\xi).$$

Here $\phi_l(\xi)$ is the Fourier multiplier associated to $S_l(D)$. We see that

$$(D_t + A)u_{j,k} = -\Theta_{j,k}f_j + [A, \Theta_{j,k}]u_j.$$

The operators $\Theta_{j,k}$ are bounded on L^2 by Plancherel's theorem. And as the kernel has Schwartz-type decay outside a ball of radius 2^{-j}, it is easy to see that they are also bounded on X, and via duality, Y. Hence, by applying $\mathcal{M}$, suitably adapted to the correct coordinate axis, to $u_{j,k}$ and summing over k, we obtain

$$\begin{aligned}2^{j-l}\|u_j\|^2_{L^2(Q)} \lesssim \|u_j\|^2_{L^\infty L^2} + \|u_j\|_{X_j}(\|f_j\|_{Y_j} + \sum_k \|[A, \Theta_{j,k}]u_j\|_{Y_j})\\ + (2^{-j} + \|w\|^2_{l^2X^s})\|u_j\|^2_{X_j}.\end{aligned} \tag{4.11}$$

Upon estimating the commutator term with (3.11), (4.7) follows. If we then take the supremum over $Q \in \mathcal{Q}_l$ and subsequently over l, yields

$$\|u_j\|^2_{X_j} \lesssim \|u_j\|^2_{L^\infty L^2} + \|u_j\|_{X_j}\|f_j\|_{Y_j} + (2^{-j} + \|w\|^2_{l^2X^s})\|u_j\|^2_{X_j},$$

provided $\|w\|^2_{l^2X^s}$ is bounded. Using (4.3), we may bootstrap to obtain

$$\|u_j\|^2_{X_j} \lesssim \|u_j\|^2_{L^\infty L^2} + \|u_j\|_{X_j}\|f_j\|_{Y_j} \tag{4.12}$$

for j sufficiently large. On the other hand, for small j, we have the trivial estimate

$$\|u\|_{X_j} \lesssim \|u\|_{L^\infty L^2},$$

which yields the same.

By applying (4.8) to (4.12), we have shown

$$\|u_j\|_{X_j} \lesssim \|u_{0j}\|_{L^2} + \|f_j\|_{Y_j}. \tag{4.13}$$

We now finish the proof by incorporating the summation over cubes. We let $\{\chi_Q\}$ denote a partition via functions which are localized to frequencies $\lesssim 1$ which are associated to cubes Q of scale $M2^j$. We also assume $|\nabla^k \chi_Q| \lesssim (2^j M)^{-k}$, $k = 1, 2$. Thus,

$$(D_t + A)(\chi_Q u_m) = -\chi_Q f_j + [A, \chi_Q]u_j.$$

Applying (4.13) to $\chi_Q u_m$, we obtain

$$\sum_Q \|\chi_Q u_j\|^2_{X_j} \lesssim \|u_{0j}\|_2^2 + \sum_Q \|\chi_Q f_j\|^2_{Y_j} + \sum_Q \|[A, \chi_Q]u_j\|^2_{L^1L^2}. \tag{4.14}$$

But

$$\sum_Q \|[A, \chi_Q]u_j\|^2_{L^1L^2} \lesssim M^{-2} \sum_Q \|\chi_Q u_j\|^2_{L^\infty L^2} \lesssim M^{-2} \sum_Q \|\chi_Q u_j\|^2_{X_j}.$$

For M sufficiently large, we can bootstrap the last term of (4.14), and upon a straightforward transition to cubes of scale 2^j rather than $M2^j$, (4.4) follows.

5. Proof of Theorem 1

We prove part (b) of the theorem, from which part (a) follows immediately.

We first examine the linear equation

$$\begin{cases} (i\partial_t + \partial_k g^{kl} \partial_l)u + V\nabla u + Wu = H, \\ u(0) = u_0 \end{cases} \tag{5.1}$$

under the assumption that

$$g^{kl} - \delta^{kl} = h^{kl}(u(t,x))$$

where $h(z) = O(|z|^2)$ near $|z| = 0$. We shall also assume that there exist smooth functions $\tilde{V}$ and $\tilde{W}$ so that

$$V = \tilde{V}(\tilde{w}(t,x)), \qquad W = \tilde{W}(v(t,x), w(t,x), \tilde{w}(t,x))$$

where the behavior of $\tilde{V}$ and $\tilde{W}$ near zero is given by

$$\tilde{V}(z) = O(|z|^2), \qquad \tilde{W}(\phi, \psi, \theta) = \phi \cdot O(|\psi|) + O(|\theta|^2).$$

Our main estimate is the following:

Proposition 5.1. *a) Assume that the metric g and potentials V and W are as above, and they satisfy*

$$\|w\|_{l^2X^s} \ll 1, \quad \|\tilde{w}\|_{l^2X^{s-1}} \ll 1, \quad \|v\|_{l^2X^{s-2}} \ll 1 \qquad s > \frac{d+3}{2}.$$

Then the equation (5.1) is well-posed for initial data $u_0 \in H^\sigma$ with $0 \le \sigma \le s-1$, and we have the estimate

$$\|u\|_{l^2X^\sigma} \lesssim \|u_0\|_{H^\sigma} + \|H\|_{l^2Y^\sigma}. \tag{5.2}$$

b) Assume in addition that $W = 0$. Then the equation (5.1) is well-posed for initial data $u_0 \in H^\sigma$ with $0 \le \sigma \le s$, and the estimate (5.2) holds.

Proof. The existence part follows in a standard manner by approximating the metric, potentials and data by Schwartz functions and passing to the limit on a subsequence, provided that we have the uniform bound (5.2). Hence the remainder of the proof is devoted to establishing (5.2).

For u solving (5.1), we see that u_j solves

$$\begin{cases} (i\partial_t + \partial_k g^{kl}_{<j-4}\partial_l)u_j = G_j + H_j, \\ u_j(0) = u_{0j}, \end{cases}$$

where

$$G_j = -S_j\partial_k g^{kl}_{>j-4}\partial_l u - [S_j, \partial_k g^{kl}_{<j-4}\partial_l]u - S_j V\nabla u - S_j W u.$$

If we apply Proposition 4.1 to each of these equations, we see that

$$\|u\|^2_{l^2X^\sigma} \lesssim \|u_0\|^2_{H^\sigma} + \|H\|^2_{l^2Y^\sigma} + \sum_j \|G_j\|^2_{l^2Y^\sigma}.$$

We claim that

$$\sum_j \|G_j\|^2_{l^2Y^\sigma} \lesssim \|u\|^2_{l^2X^\sigma}\Big(\|w\|^4_{l^2X^s} + \|\tilde{w}\|^4_{l^2X^{s-1}} + \|v\|^4_{l^2X^{s-2}}\Big).$$

Indeed, for the first term in G_j we apply (3.2) and (3.7). The bound for the second term in G_j follows from (3.11) and (3.2). The bounds for the last two terms of G_j use (3.5) and (3.6) respectively in conjunction with (3.2). The differences in the range of permissible σ in (3.5) versus (3.6) accounts precisely for the difference in parts (a) and (b) of this proposition. □

5.1. **The iteration.** We now set up an iteration to solve (1.2). We set $u^{(0)} \equiv 0$ and recursively define $u^{(n+1)}$ to be the solution to

$$\begin{cases} (i\partial_t + \partial_j g^{jk}(u^{(n)})\partial_k)u^{(n+1)} = F(u^{(n)}, \nabla u^{(n)}), \\ \\ u^{(n+1)}(0,x) = u_0(x). \end{cases} \tag{5.3}$$

As $F(y,z) = O(|y|^3 + |z|^3)$ near the origin, it follows from (3.2) with $\sigma = s-1$ and (3.5) with $\sigma = s$ that

$$\|F(u^{(n)}, \nabla u^{(n)})\|_{l^2Y^s} \lesssim \|u^{(n)}\|^3_{l^2X^s}$$

provided $s > \frac{d+3}{2}$ and $\|u^{(n)}\|_{l^2X^s} = O(1)$. Using this in each application of Proposition 5.1, we see, via induction, that

$$\|u^{(n)}\|_{l^2X^s} \lesssim \|u_0\|_{H^s} \tag{5.4}$$

for each n provided that $\|u_0\|_{H^s}$ is sufficiently small.

We now seek to show that the iteration converges in l^2X^{s-1}. To this end, we note that $v^{(n+1)} = u^{(n+1)} - u^{(n)}$ solves

$$\begin{cases} (i\partial_t + \partial_j g^{jk}(u^{(n)})\partial_k)v^{(n+1)} = V_n \nabla v^{(n)} + W_n v^{(n)}, \\ v^{(n+1)}(0,x) = 0, \end{cases} \tag{5.5}$$

where

$$\begin{aligned} V_n &= V_n(u^{(n)}, \nabla u^{(n)}, u^{(n-1)}, \nabla u^{(n-1)}), \\ W_n &= h_1(u^{(n)}, \nabla u^{(n)}, u^{(n-1)}, \nabla u^{(n-1)}) + h_2(u^{(n)}, u^{(n-1)})\nabla^2 u^{(n)}. \end{aligned}$$

Here $V_n(z) = O(|z|^2)$, $h_1(z) = O(|z|^2)$, and $h_2(z) = O(|z|)$ near $|z| = 0$. By Proposition 5.1, we have

$$\|v^{(n+1)}\|_{l^2X^{s-1}} \lesssim \|V_n \nabla v^{(n)}\|_{l^2Y^{s-1}} + \|W_n v^{(n)}\|_{l^2Y^{s-1}}.$$

By (3.5) and (3.6) with $\sigma = s-1$, we have

$$\|V_n \nabla v^{(n)}\|_{l^2Y^{s-1}} \lesssim \Big(\|u^{(n)}\|_{l^2X^s} + \|u^{(n-1)}\|_{l^2X^s}\Big)^2 \|v^{(n)}\|_{l^2X^{s-1}},$$

$$\|W_n v^{(n)}\|_{l^2Y^{s-1}} \lesssim \Big(\|u^{(n)}\|_{l^2X^s} + \|u^{(n-1)}\|_{l^2X^s}\Big)^2 \|v^{(n)}\|_{l^2X^{s-1}}.$$

Thus, it follows from (5.4) that

$$\|v^{(n+1)}\|_{l^2X^{s-1}} \ll \|v^{(n)}\|_{l^2X^{s-1}}, \tag{5.6}$$

which implies that the iteration converges in l^2X^{s-1} to a function u satisfying

$$\|u\|_{l^2X^s} \lesssim \|u_0\|_{H^s}. \tag{5.7}$$

This establishes the existence of a solution.

We next consider the question of uniqueness. For two solutions $u^{(1)}$ and $u^{(2)}$ to (1.2), we consider $v = u^{(2)} - u^{(1)}$. Here, v solves

$$\begin{cases} (i\partial_t + \partial_j g^{jk}(u^{(2)})\partial_k)v = V\nabla v + Wv, \\ v(0,x) = u^{(2)} - u^{(1)}, \end{cases} \tag{5.8}$$

where

$$\begin{aligned} V &= V(u^{(1)}, \nabla u^{(1)}, u^{(2)}, \nabla u^{(2)}), \\ W_n &= h_1(u^{(2)}, \nabla u^{(2)}, u^{(1)}, \nabla u^{(1)}) + h_2(u^{(2)}, u^{(1)})\nabla^2 u^{(1)}. \end{aligned}$$

By Proposition 5.1, we have

$$\|u^{(2)} - u^{(1)}\|_{l^2X^{s-1}} \lesssim \|u^{(2)}(0) - u^{(1)}(0)\|_{H^{s-1}} \tag{5.9}$$

from which uniqueness follows.

In order to show continuity of the map $u_0 \to u$ from $H^s \to l^2X^s$, we seek to strengthen the preceding argument. To do so, we shall use frequency envelopes. Indeed, we first seek to strengthen (5.7) by showing that a frequency envelope for the data is also a frequency envelope for the solution.

Proposition 5.2. *Let $s > \frac{d+3}{2}$, and let $u \in l^2X^s$ be a small data solution to* (1.2)*, which satisfies* (5.7)*. If $\{a_j\}$ is an admissible frequency envelope for the initial data u_0 in H^s, then $\{a_j\}$ is also a frequency envelope for u in l^2X^s.*

Proof. We set

$$b_j = 2^{-\delta j} + \|u\|_{l^2X^s}^{-1} \max_k 2^{-\delta|j-k|}\|S_k u\|_{l^2X^s} \tag{5.10}$$

and note that $\{b_j\}$ is a frequency envelope for u in l^2X^s. We note that u_j solves

$$\begin{cases} (i\partial_t + \partial_k g^{kl}_{<j-4}\partial_l)u_j = S_j F(u, \nabla u) - S_j\partial_k g^{kl}_{>j-4}\partial_l u \\ \qquad\qquad\qquad\qquad\qquad -[S_j, \partial_k g^{kl}_{<j-4}\partial_l]u \\ u_j(0) = (u_0)_j. \end{cases} \tag{5.11}$$

Labelling the right side of the equation above f_j, we apply Proposition 5.1 and obtain

$$\|S_j u\|_{l^2X^s} \lesssim a_j\|u_0\|_{H^s} + \|f_j\|_{l^2Y^s}. \tag{5.12}$$

We bound f_j in a manner that is akin to the above. This yields

$$\|f_j\|_{l^2Y^s} \lesssim b_j\|u\|^3_{l^2X^s} c(\|u\|_{l^2X^s}). \tag{5.13}$$

Indeed, provided $s > \frac{d+3}{2}$, we can apply (3.2) ($\sigma = s-1$) and (3.5) ($\sigma = s$) to bound the first term. For the second term we use (3.2) ($\sigma = s$) and (3.7). For the last term in f_j, we apply (3.11).

Applying (5.13) in (5.12) we obtain

$$\|S_j u\|_{l^2X^s} \lesssim a_j \|u_0\|_{H^s} + b_j \|u\|^3_{l^2X^s} c(\|u\|_{l^2X^s}),$$

which yields

$$b_j \lesssim a_j \|u_0\|_{H^s} \|u\|^{-1}_{l^2X^s} + b_j \|u\|^2_{l^2X^s} c(\|u\|_{l^2X^s}).$$

Using the smallness of $\|u\|_{l^2X^s}$, we can bootstrap the second term in the right. Moreover, by the definition of l^2X^s, we have $\|u_0\|_{H^s} \lesssim \|u\|_{l^2X^s}$. Thus, the above implies $b_j \lesssim a_j$ as desired. □

We now proceed to show that the map $H^s \to l^2X^s$ given by $u_0 \mapsto u$ is continuous. Let $\{u_0^{(n)}\} \subset H^s$ be a sequence which converges to u_0 in H^s, and let $\{a_j^{(n)}\}$ and $\{a_j\}$ denote their respective frequency envelopes defined via (2.3). It follows, thus, that $a_j^{(n)} \to a_j$ in l^2. For any $\varepsilon > 0$, there is a N_ε so that

$$\|a_j^{(n)}\|_{l^2(j>N_\varepsilon)} \le \varepsilon, \quad \|a_j\|_{l^2(j>N_\varepsilon)} \le \varepsilon$$

uniformly in n. The preceding proposition then yields

$$\begin{aligned} \|u^{(n)}_{>N_\varepsilon}\|_{l^2X^s} &\le \varepsilon \|u^{(n)}\|_{l^2X^s} \le C\varepsilon \|u_0^{(n)}\|_{H^s} \\ \|u_{>N_\varepsilon}\|_{l^2X^s} &\le \varepsilon \|u\|_{l^2X^s} \le C\varepsilon \|u_0\|_{H^s} \end{aligned} \tag{5.14}$$

where in the last step we applied (5.7).

To compare $u^{(n)}$, where $u^{(n)}$ is the solution to (1.2) with datum $u_0^{(n)}$, to u, we use (5.14) for the high frequencies and (5.9) for the low frequencies. Indeed,

$$\begin{aligned} \|u^{(n)} - u\|_{l^2X^s} &\lesssim \|S_{<N_\varepsilon}(u^{(n)} - u)\|_{l^2X^s} + \|u^{(n)}_{>N_\varepsilon}\|_{l^2X^s} + \|u_{>N_\varepsilon}\|_{l^2X^s} \\ &\lesssim 2^{N_\varepsilon} \|S_{<N_\varepsilon}(u^{(n)} - u)\|_{l^2X^{s-1}} + \varepsilon \|u_0^{(n)}\|_{H^s} + \varepsilon \|u_0\|_{H^s} \\ &\lesssim 2^{N_\varepsilon} \|S_{<N_\varepsilon}(u_0^{(n)} - u_0)\|_{H^{s-1}} + \varepsilon \|u_0^{(n)}\|_{H^s} + \varepsilon \|u_0\|_{H^s}. \end{aligned}$$

Letting $n \to \infty$ yields

$$\limsup_{n\to\infty} \|u^{(n)} - u\|_{l^2X^s} \lesssim \varepsilon \|u_0\|_{H^s},$$

and subsequently letting $\varepsilon \to 0$ yields the desired result.

We finish by showing the analog of (5.7) for higher frequencies:

$$\|u\|_{l^2X^\sigma} \lesssim \|u_0\|_{H^\sigma}, \quad \sigma \ge s \tag{5.15}$$

assuming that $u_0 \in H^\sigma$. Differentiating the original equation yields

$$\begin{aligned} (i\partial_t + \partial_j g^{jk}(u)\partial_k)(\partial_l u) &= -(g^{jk})'(u)(\partial_j\partial_l u \partial_k u + \partial_l u \partial_j \partial_k u) \\ &\quad - (g^{jk})''(u)(\partial_j u \partial_l u \partial_k u) + (\nabla_{z_1} F)(u, \nabla u) \cdot \nabla \partial_l u + F_{z_0}(u, \nabla u)\partial_l u. \end{aligned}$$

For $v_1 = \nabla u$, we have

$$(i\partial_t + \partial_j g^{jk}(u)\partial_k)v_1 = G(u, \nabla u)\nabla v_1 + F_1(u, \nabla u)$$

where $G(z) = O(|z|^2)$ and $F_1(z) = O(|z|^3)$ near $z = 0$. By (3.2) and (3.5), we have

$$\|G(u, \nabla u)\nabla v_1\|_{l^2 Y^s} \lesssim \|u\|^2_{l^2 X^s}\|v_1\|_{l^2 X^s}, \quad \|F_1(u, \nabla u)\|_{l^2 Y^s} \lesssim \|u\|^3_{l^2 X^s}.$$

And by Proposition 5.1,

$$\|v_1\|_{l^2 X^s} \lesssim \|v_1(0)\|_{H^s} + \|u\|^3_{l^2 X^s},$$

and thus,

$$\|u\|_{l^2 X^{s+1}} \lesssim \|u(0)\|_{H^{s+1}} + \|u\|^3_{l^2 X^s}.$$

Letting $v_n = \nabla^n u$, we see that v_n solves

$$(i\partial_t + \partial_j g^{jk}(u)\partial_k)v_n = G(u, \nabla u)\nabla v_n + F_n(u, \nabla u, \dots, \nabla^n u)$$

with G as above. Arguing inductively, we obtain

$$\|v_n\|_{l^2 X^s} \lesssim \|v_n(0)\|_{H^s} + \|u\|^3_{l^2 X^{s+n-1}},$$

which gives

$$\|u\|_{l^2 X^{s+n}} \lesssim \|u(0)\|_{H^{s+n}} + \|u\|^3_{l^2 X^{s+n-1}}$$

from which the desired conclusion follows.

References

[1] I. Bejenaru and D. Tataru: *Large data local solutions for the derivative NLS equation*. J. Eur. Math. Soc. **10** (2008), 957–985.

[2] C. E. Kenig, G. Ponce, and L. Vega: *Small solutions to nonlinear Schrödinger equations*. Ann. Inst. H. Poincaré Anal. Non Linéaire **10** (1993), 255–288.

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

[4] C. Kenig, G. Ponce, and L. Vega: *The Cauchy problem for quasi-linear Schrödinger equations*. Invent. Math. **158** (2004), 343-388.

[5] C. Kenig, G. Ponce, C. Rolvung, and L. Vega: *The general quasilinear ultrahyperbolic Schrdinger equation*. Adv. Math. **196** (2005), 402-433.

[6] C. Kenig, G. Ponce, C. Rolvung, and L. Vega: *Variable coefficient Schrödinger flows for ultrahyperbolic operators*. Adv. Math. **206** (2006), 373-486.

[7] W. Ichinose: *On L^2 well-posedness of the Cauchy problem for Schrödinger type equations on a Riemannian manifold and Maslov theory*. Duke Math. J. **56** (1988), 549-588.

[8] F. Linares and G. Ponce: *Introduction to nonlinear dispersive equations. Universitext*, Springer, New York (2009).

[9] J. Marzuola, J. Metcalfe and D. Tataru: *Wave packet parametrices for evolutions governed by pdo's with rough symbols*. Proc. Amer. Math. Soc. **136** (2007), 597-604.

[10] J. Marzuola, J. Metcalfe, and D. Tataru: *Quasilinear Schrödinger equations I: Small data and quadratic interactions*. Preprint, 2011. (arXiv: 1106.0490)

[11] S. Mizohata: *Some remarks on the Cauchy problem*. J. Math. Kyoto Univ. **1** (1961), 109-127.

[12] S. Mizohata: *Sur quelques equations du type Schrödinger*. Journees Equations aux derivees partielles, Saint-Jean de Monts, 1981.

[13] S. Mizohata: *On the Cauchy Problem*, *Notes and Reports in Mathematics in Science and Engineering*, **3**, Science Press and Academic Press (1985).

Department of Mathematics, University of North Carolina-Chapel Hill, Phillips Hall, Chapel Hill, NC 27599-3250, USA
E-mail address: marzuola@email.unc.edu

Department of Mathematics, University of North Carolina-Chapel Hill, Phillips Hall, Chapel Hill, NC 27599-3250, USA
E-mail address: metcalfe@email.unc.edu

Mathematics Department, University of California, Evans Hall, Berkeley, CA 94720, USA
E-mail address: tataru@math.berkeley.edu