WAVE MAPS ON (1+2)-DIMENSIONAL CURVED SPACETIMES

CRISTIAN GAVRUS, CASEY JAO, AND DANIEL TATARU

ABSTRACT. In this article we initiate the study of $1+2$ dimensional wave maps on a curved space time in the low regularity setting. Our main result asserts that in this context the wave maps equation is locally well-posed at almost critical regularity.

As a key part of the proof of this result, we generalize the classical optimal bilinear L^2 estimates for the wave equation to variable coefficients, by means of wave packet decompositions and characteristic energy estimates. This allows us to iterate in a curved $X^{s,b}$ space.

CONTENTS

1. INTRODUCTION

Let the $(1+2)$-dimensional spacetime $\mathbb{R}_t \times \mathbb{R}^2_x$ be endowed with a Lorentzian metric

$$g = g_{\alpha\beta}(t,x)dx^\alpha dx^\beta,$$

so that the time slices $t = const$ are space-like. Here $x^0 = t$ and we adopt the standard convention of referring to spacetime coordinates by Greek indices and purely spatial coordinates by Roman indices.

Given a smooth Riemannian manifold (M,h) with uniformly bounded geometry, a *wave map* $u : (\mathbb{R}^{1+2}, g) \to (M,h)$ is formally a critical point of the Lagrangian

$$\mathcal{L}(u) = \frac{1}{2}\int_{\mathbb{R}^{1+2}} \langle du, du\rangle_{T^*(\mathbb{R}\times\mathbb{R}^2)\otimes u^*TM}\, \mathrm{dvol}_g,$$

where $u^*TM = \bigcup_x \{x\} \times T_{u(x)}M$ is the pullback of TM by u and u^*h is the pullback metric. In local coordinates on M this becomes

$$\mathcal{L}(u) = \frac{1}{2} \int g^{\alpha\beta}(z) h_{ij}(u(z)) \partial_\alpha u^i(z) \partial_\beta u^j(z) \sqrt{|g(z)|}\, \mathrm{d}z$$

One may think of wave maps as the hyperbolic counterpart of harmonic maps. The Euler-Lagrange equations take the following form, and we refer to [13, Chapter 8] for related computations:

$$\frac{1}{\sqrt{|g|}} \mathbf{D}_\alpha \left(\sqrt{|g|} g^{\alpha\beta} \partial_\beta u \right) = 0 \tag{1.1}$$

Here $\mathbf{D}$ denotes the pullback covariant derivative on u^*TM given by $\mathbf{D}_X V = \nabla_{u_* X} V$ where ∇ is the Levi-Civita connection on TM.

In coordinates, wave maps solve a coupled system of nonlinear wave equations. We review two useful settings for this problem:

Intrinsic formulation. Suppose the image of u is supported in the domain of a local coordinate patch of M. Then the wave maps equation (1.1) is written as

$$\tilde{\Box}_g u^i = -\Gamma^i_{jk}(u) g^{\alpha\beta} \partial_\alpha u^j \partial_\beta u^k \tag{1.2}$$

where $\tilde{\Box}_g$ is the Laplace-Beltrami wave operator

$$\tilde{\Box}_g u = |g|^{-\frac{1}{2}} \partial_\alpha \Big(|g|^{\frac{1}{2}} g^{\alpha\beta} \partial_\beta u \Big) \tag{1.3}$$

denoted this way in order to distinguish it from its principal part $\Box_g$ defined by

$$\Box_g u = g^{\alpha\beta} \partial_\alpha \partial_\beta u,$$

and Γ^i_{jk} are the Christoffel symbols on (M, h).

Extrinsic formulation. When the manifold M is isometrically embedded into an euclidean space $\mathbb{R}^m$ wave maps can be equivalently defined extrinsically; see for instance [28]. If M is compact, such an embedding exists for m large enough by Nash's theorem. The Lagrangian becomes

$$\mathcal{L}(u) = \frac{1}{2} \int g^{\alpha\beta} \langle \partial_\alpha u, \partial_\beta u \rangle \sqrt{|g|}\, \mathrm{d}z$$

Formal critical points satisfy $\tilde{\Box}_g u \perp T_u M$ and our equation takes the form

$$\tilde{\Box}_g u^i = -\mathcal{S}^i_{jk}(u) g^{\alpha\beta} \partial_\alpha u^j \partial_\beta u^k, \tag{1.4}$$

where $\mathcal{S}$ stands for the second fundamental form on M.

Initial data for the Cauchy problem are chosen such that

$$u(0, x) \in M, \ \partial_t u(0, x) \in T_{u(0,x)} M.$$

The initial data space can be viewed as an infinite dimensional manifold, in which the wave map evolution takes place.

The formulations above exhibit the presence of the *null form*

$$Q_g(u,v) = g^{\alpha\beta}\partial_\alpha u \partial_\beta v.$$

Its structure is so that it eliminates quadratic resonant interactions in the wave map evolution, and has played a key role in the low regularity well-posedness of wave maps in the flat case ([15], [19], [20], [39], [34]) as well as in other problems, especially in low dimensions. A simple way to see this cancellation is via the formula

$$2Q_g(u,v) = 2g^{\alpha\beta}\partial_\alpha u \partial_\beta v = \Box_g(uv) - u\Box_g v - v\Box_g u. \tag{1.5}$$

However in *Section* 6 we will obtain estimates for $Q_g(u,v)$ directly.

In both the intrinsic and the extrinsic formulations one can use the coordinates to define Sobolev spaces for the initial data sets. To understand the relevant range of Sobolev indices we recall that in the Minkovski space the wave maps system is invariant under the dimensionless *rescaling*

$$u \mapsto u(\lambda t, \lambda x), \qquad g \mapsto g(\lambda t, \lambda x)$$

which identifies the *critical* (scale-invariant) initial data space as $\dot{H}^1 \times L^2(\mathbb{R}^2)$. In addition, one may consider initial data spaces for more regular data, of the form

$$\mathcal{H}^s = \{(u_0, u_1) : \mathbb{R}^2 \to TM;\ \nabla u_0, u_1 \in H^{s-1}(\mathbb{R}^2)\}, \qquad s > 1.$$

It is this latter case which is considered in the present paper.

Several remarks are in order here. First of all, the constant functions are always acceptable states in the context of manifold valued maps u_0; this is why we include them all in our state space $\mathcal{H}^s$ above. Secondly, one may ask whether the definition of the above spaces is context dependent. We separately discuss the two formulations above.

The simplest set-up is in the case of the extrinsic formulation with (M,h) a compact manifold. There one can directly use the $H^{s-1}(\mathbb{R}^2)$ spaces for derivatives of $\mathbb{R}^m$ valued functions. For simplicity, this is the set-up we adopt here.

For the intrinsic formulation matters are not as straightforward. The fact that $s > 1$ heuristically insures that functions u_0 with $\nabla u_0 \in H^{s-1}$ are Hölder continuous. Thus locally such functions are in the domain of a single chart, and the H^{s-1} norms can be defined using local coordinates. Further, by Moser estimates the local spaces are algebraically independent of the choice of the local chart. Finally, the global spaces are obtained from the local ones by using a suitable collection of local charts. Of course, some care is required here in order to avoid a circular argument. Most generally such a construction would apply for manifolds (M,g) with uniformly bounded geometry.

Finally we remark that in the setting of low-regularity discontinuous solutions as in [43], defining the critical Sobolev spaces $\dot{H}^1_x(\mathbb{R}^2 \to M)$ is a delicate matter due to the need of having an appropriate topology on these spaces that is independent of the choice of isometric embeddings.

Throughout the paper we assume that

- The metric coefficients $g_{\alpha\beta}$ and $g^{\alpha\beta}$ are uniformly bounded.
- The surfaces $t = const$ are uniformly space-like,

$$g_{ij}v^i v^j \gtrsim |v|^2.$$

These conditions in turn imply the bound from below

$$-g^{00} \gtrsim 1.$$

Now we are ready to formulate our main result, which establishes well-posedness in the *energy-subcritical* regime $s > 1$, as a stepping stone toward the energy-critical problem on a curved background.

Theorem 1.1. *Let $M \subset \mathbb{R}^m$ be an embedded manifold. Assume that $\partial_{t,x}^2 g^{ij} \in L_t^2 L_x^\infty$. Then the Cauchy problem for the wave maps equation* (1.4) *is locally well-posed in $\mathcal{H}^s$ for $1 < s \leq 2$.*

Our result includes existence, uniqueness, locally Lipschitz dependence on the initial data and persistence of regularity as explained in Section 1.2.

While our main result above applies to the large data problem, the bulk of the paper is devoted to the small data problem, to which the large data case reduces after a suitable localization. To state the small data result, we replace the qualitative property of the metric $\nabla^2 g \in L^2 L^\infty$ with a quantitative version which applies in the unit time interval $t \in [0, 1]$:

$$\|\partial_{t,x} g^{\alpha\beta}\|_{L^\infty L^\infty} + \|\partial_{t,x}^2 g^{\alpha\beta}\|_{L^2 L^\infty} \leq \eta^2. \tag{1.6}$$

Here $\eta \ll 1$ is a fixed small parameter.

For the next result below, we work in a local patch in the intrinsic setting. We assume that 0 belongs to the range of local coordinates associated to this patch, and work with data (u_0, u_1) which is small in $H_x^s \times H_x^{s-1}$. Then u_0 is continuous and uniformly small. As long as this property persists, the solution will remain within the domain of the local patch. Hence its regularity can be measured in those coordinates. Our small data result is as follows:

Theorem 1.2. *Let $1 < s \leq 2$ and assume g satisfies* (1.6) *in the time interval $[0, 1]$. There exists $\epsilon > 0$ such that for any initial data set (u_0, u_1) satisfying*

$$\|(u_0, u_1)\|_{H^s \times H^{s-1}} \leq \epsilon \tag{1.7}$$

there exists a unique solution u to the wave maps problem (1.2) *with this data in the space $C([0, 1]; H^s) \cap C^1([0, 1]; H^{s-1})$, satisfying*

$$\|(u, \partial_t u)\|_{L^\infty(H^s \times H^{s-1})} \lesssim \|(u_0, u_1)\|_{H^s \times H^{s-1}}. \tag{1.8}$$

The solution has a Lipschitz continuous dependence on the initial data.

Here the uniqueness is interpreted in the classical sense if $s = 2$. For $1 < s < 2$, the $H^s \times H^{s-1}$ can be defined as the unique limits of $H^2 \times H^1$ solutions. Alternatively, one may prove the uniqueness property in the $X^{s,\theta}$ spaces, see the discussion below.

The higher Sobolev regularity is limited to H^2 given the regularity of the metric g. However, adding higher regularity to g correspondingly adds regularity to the class of regular solutions.

The objective of Section 1.2 will be to reduce our main result in Theorem 1.1 to the small data result in Theorem 1.2 by a standard scaling and finite speed of propagation argument.

A key role in our analysis is played by the spaces $X^{s,\theta}$ associated to the wave operator $\Box_g$. Indeed, our main well-posedness argument is phrased as a fixed point argument in $X^{s,\theta}$ where

$$1 < s \leq 2, \qquad \frac{1}{2} < \theta \leq \min\{1, s - \frac{1}{2}\}.$$

In particular one can rephrase the uniqueness property in our main theorems as an unconditional uniqueness in $X^{s,\theta}$, and the Lipschitz dependence as a Lipschitz dependence in $X^{s,\theta}$. These spaces are defined in Section 2, and the study of their linear and bilinear properties occupies much of the paper.

1.1. **Previous works.** Here we provide a brief survey of previous results on Wave Maps well-posedness, with an emphasis on variable coefficients.

The Cauchy problem on a flat background $(\mathbb{R}^{1+n}, -dt^2 + dx^2)$ is by now well understood. In view of the scaling symmetry $u(t,x) \mapsto u(\lambda t, \lambda x)$, the critical Sobolev space is $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n)$. Local wellposedness in $H^s \times H^{s-1}$ was established for all subcritical regularities $s > \frac{n}{2}$ by Klainerman-Machedon for $n \geq 3$ and Klainerman-Selberg when $n = 2$ [15], [19]. The much more delicate critical problem $s = \frac{n}{2}$ was solved for small data in dimension $n = 2$ by Tataru [39], [43], Tao [34] and Krieger [21], with further contributions in higher dimension by Klainerman-Rodnianski [17], and Shatah-Struwe [28], Nahmod-Stefanov-Uhlenbeck [27]. Further, when $n = 2$, the energy-critical problem in $H^1 \times L^2$ admits a global theory for large data as developed by Sterbenz-Tataru [32], [33], Krieger-Schlag [22] and Tao [36].

For wave maps with variable coefficients, Geba [9] established local wellposedness in the subcritical regime $s > \frac{n}{2}$ when $3 \leq n \leq 5$, building on previous work of Geba and Tataru [11]. More recently, Lawrie constructed global-in-time solutions on perturbations of $\mathbb{R}^{1+4}$ Minkowski space for small data in the critical space $H^2 \times H^1(\mathbb{R}^4)$, and Lawrie-Oh-Shahshashani obtained analogous small-data results on $\mathbb{R} \times \mathbb{H}^n$, $n \geq 4$ [24]. See also the recent work of Li-Ma-Zhao on the stability of harmonic maps $\mathbb{H}^2 \to \mathbb{H}^2$ under the wave map flow [25].

A key component in the study of wave maps in the Minkowski case at critical regularity is Tao's renormalization idea, first introduced in [34]. The subcritical problem considered in this article avoids the renormalization argument, which simplifies matters considerably. On the other hand, Strichartz estimates do not suffice to treat the full subcritical range $s > \frac{n}{2}$ in low dimensions, in particular when $n = 2, 3$. Indeed the null structure $Q_g(u,u)$ distinguishes the wave maps system from equations with a generic quadratic derivative nonlinearity $(\nabla u)^2$; as observed by Lindblad [26], the latter can be illposed in $H^s \times H^{s-1}(\mathbb{R}^3)$ for $s = 2 > \frac{3}{2}$. The previous works [15, 19] rely on $X^{s,b}$ spaces to exploit the null structure in lower dimensions. We pursue an analogous strategy using the variable-coefficient $X^{s,b}$-type spaces first introduced by Tataru [37] and further developed by by Geba and Tataru [44]. However, the two-dimensional case involves additional subtleties as we shall discuss shortly.

1.2. **Reduction of Theorem 1.1 to Theorem 1.2.** Let

$$(\tilde{u}_0, \tilde{u}_1) : \mathbb{R}^2_x \to M \times T_{\tilde{u}_0} M \subset \mathbb{R}^m \times \mathbb{R}^m$$

be an initial data such that

$$\|(\nabla \tilde{u}_0, \tilde{u}_1)\|_{H^{s-1}} \leq R.$$

We have a solution $\tilde{u}$ on a small time interval $[0,T]$ if the rescaled function

$$u(t,x) = \tilde{u}(Tt, Tx)$$

is a solution on $[0,1]$ using the rescaled metric $g(Tt,Tx)$. This now obeys (1.6), provided that T is small enough.

The data for the rescaled solution will satisfy the scale invariant bound

$$\|(u_0, u_1)\|_{\dot{H}^1 \times L^2} \leq R$$

as well as the homogeneous bound

$$\|(u_0, u_1)\|_{\dot{H}^s \times \dot{H}^{s-1}} \leq R T^{s-1}$$

We will choose T small enough so that

$$R T^{s-1} \ll \epsilon.$$

To obtain smallness of the full $H^s \times H^{s-1}$ norms we truncate the initial data, then we apply Theorem 1.2 and we glue those solutions using finite speed of propagation.

Let c be the largest speed of propagation and let $(y_j)_j$ be the centers of a family of balls such that the truncated cones

$$K_j = \{(t, x) \mid ct + |x - y_j| \leq c + 1, \ t \in [0, 1]\}$$

are finitely overlapping and cover $[0, 1] \times \mathbb{R}^2$.

Let χ_j be a smooth function which equals 1 on the ball $B_j = B_{y_j}(c+1)$ and is supported on $\tilde{B}_j = B_{y_j}(c+2)$. We denote by $B_j^t = B_{y_j}(c+1-ct)$, $\tilde{B}_j^t = B_{y_j}(c+2-ct)$ the corresponding balls at time t.

For every y_j we choose a local chart of M such that $u_0(y_j)$ corresponds to the origin. We localize around y_j viewing (by slight abuse of notation) u_0, u_1 as having their image in the chart and defining $u_0^j = \chi_j u_0$, $u_1^j = \chi_j u_1$. Since, $u_0^j(y_j) = 0$, by homogeneous Sobolev embedding and Morrey's inequality we deduce that smallness is retained locally by the inhomogeneous Sobolev norms:

$$\|(u_0^j, u_1^j)\|_{H^s \times H^{s-1}} \leq \epsilon. \tag{1.9}$$

One uses Moser estimates to pass from the extrinsic Sobolev spaces to the Sobolev norms defined using the patch coordinates.

By Theorem 1.2 we obtain a solution u^j to (1.2) which remains in the image of the chart on $[0, 1]$ by (1.8) and Sobolev embedding. Now viewing u^j as taking values in $M \subset \mathbb{R}^m$, it solves (1.4) and we restrict it to the truncated cone K_j.

To obtain a solution u on $[0, 1] \times \mathbb{R}^2$ defined by each u^j restricted to K_j we argue that any two of them must coincide on their common domain. For H^2 solutions this follows from the finite speed of propagation, which is then proved in a standard fashion. For rough solutions we use the well-posedness result from Theorem 1.2 to approximate them by H^2 solutions, and then we pass to the limit.

Now we show that $\nabla_{t,x} u(t) \in H^{s-1}(\mathbb{R}^2)$. First we note that, by (1.25) we have

$$\begin{aligned} \sum_j \|u_0^j\|_{H^s}^2 &\simeq \sum_j \|\chi_j u_0\|_{L^2(\tilde{B}_j)}^2 + \|\nabla_x(\chi_j u_0)\|_{H^{s-1}}^2 \lesssim \\ &\lesssim \sum_j \|u_0\|_{L^2(\tilde{B}_j)}^2 + \|\nabla_x u_0\|_{W^{s-1,2}(\tilde{B}_j)}^2 \lesssim \sum_j \|\nabla_x u_0\|_{W^{s-1,2}(\tilde{B}_j)}^2 \end{aligned} \tag{1.10}$$

since by Morrey's inequality and Sobolev embedding we have

$$\|u_0\|_{L^2(\tilde{B}_j)} \lesssim \|\nabla_x u_0\|_{L^{2+}(\tilde{B}_j)} \lesssim \|\nabla_x u_0\|_{W^{s-1,2}(\tilde{B}_j)}.$$

Moreover, by (1.26) and (1.24) we have

$$\text{RHS (1.10)} \lesssim \sum_j \|\nabla_x u_0\|^2_{L^2(\tilde{B}_j)} + |\nabla_x u_0|^2_{\dot{W}^{s-1,2}(\tilde{B}_j)} \lesssim \|u_0\|^2_{\dot{H}^1} + \|\nabla_x u_0\|^2_{\dot{H}^{s-1}} \lesssim R^2.$$

Similarly we obtain $\sum_j \|u_1^j\|^2_{H^{s-1}} \lesssim R^2$. For the solution at time t, using (1.27) we get

$$\|\nabla_{t,x} u(t)\|^2_{H^{s-1}} \lesssim \sum_j \|\nabla_{t,x} u(t)\|^2_{W^{s-1,2}(B_j^t)} \lesssim \sum_j \|u^j[t]\|^2_{H^s \times H^{s-1}} \lesssim \sum_j \|u_0^j, u_1^j\|^2_{H^s \times H^{s-1}} \lesssim R^2.$$

This proves that $\tilde{u}(t) \in \mathcal{H}^s$ for $t \in [0,T]$ and

$$\|\tilde{u}\|_{L^\infty \mathcal{H}^s([0,T]\times\mathbb{R}^2)} \leq C_R \|(\tilde{u}_0, \tilde{u}_1)\|_{\mathcal{H}^s}$$

with a constant C_R depending on R.

Now we address the locally Lipschitz dependence on the initial data. Let $(\tilde{v}_0, \tilde{v}_1) : \mathbb{R}^2_x \to M \times T_{\tilde{u}_0} M \subset \mathbb{R}^m \times \mathbb{R}^m$ be another initial data set such that

$$\|(\tilde{u}_0 - \tilde{v}_0, \tilde{u}_1 - \tilde{v}_1)\|_{(\dot{H}^s \cap \dot{H}^1 \cap L^\infty) \times H^{s-1}}$$

is small enough and let $\tilde{v}$ be the solution on $[0,T]$ with this data. Then using the argument above together with the Lipschitz dependence given by Theorem 1.2 of the local solutions we obtain

$$\|\tilde{u} - \tilde{v}\|_{L^\infty \mathcal{H}^s([0,T]\times\mathbb{R}^2)} \lesssim C_R \|(\tilde{u}_0 - \tilde{v}_0, \tilde{u}_1 - \tilde{v}_1)\|_{\mathcal{H}^s}.$$

Finally, we remark that assuming higher regularity for the metric g (such as $\partial^k g \in L^2 L^\infty$ and $g \in L^\infty H^{k-1}$ for $k \geq 3$, see the discussion below in the proof of Theorem 1.2) we have that $(\dot{H}^n \cap \dot{H}^1) \times H^{n-1}$ regularity of the initial data is maintained by the solution on $[0,T]$ for $n \leq k$.

Remark 1.3. If we assume the initial data $(\tilde{u}_0, \tilde{u}_1)$ to be only in $\dot{H}^s \times \dot{H}^{s-1}$, then we can still construct a solution, but it will be only *locally* in $H^s \times H^{s-1}$ at future times.

1.3. **Proof of Theorem 1.2.** Here we set up the fixed point argument which yields Theorem 1.2, and show that the proof reduces to four estimates described below.

As a preliminary step, we replace the metric $g^{\alpha\beta}$ by $\tilde{g}^{\alpha\beta} = (g^{00})^{-1} g^{\alpha\beta}$ in order to insure that $\tilde{g}^{00} = 1$; this is not crucial in the analysis, but yields some minor technical simplifications. The price to pay for this substitution is that we get another term in the equations,

$$\tilde{\Box}_{\tilde{g}} u^i = -\Gamma^i_{jk}(u) \tilde{g}^{\alpha\beta} \partial_\alpha u^j \partial_\beta u^k + \frac{5}{2} \partial^\alpha (\log |g^{00}|) \tilde{g}^{\alpha\beta} \partial_\beta u \tag{1.11}$$

The extra term on the right will easily be perturbative in our setting. Hence we will simply neglect it, drop the $\tilde{g}$ notation and simply assume that $g^{00} = -1$.

We now consider an initial data satisfying (1.7) and proceed to obtain the solution to (1.2) on $[0,1]$ by a fixed point argument, as $u = \Phi(u)$ for the functional

$$\Phi^i(u) := u^i_{lin} + \tilde{\Box}_g^{-1}\big(- \Gamma^i_{jk}(u) Q_g(u^j, u^k)\big) \tag{1.12}$$

where u_{lin} is the solution of the linear equation with $F = 0$

$$\tilde{\Box}_g u = F, \qquad u[0] = (u_0, u_1) \tag{1.13}$$

while $\tilde{\Box}_g^{-1} F$ is defined as the solution of (1.13) with $u[0] = (0,0)$.

This argument relies on the choice of two Banach spaces X (for the components of our solutions) and N (for the perturbative nonlinearity) such that Φ is a contraction on a small ball of X. Specifically,

$$X = X^{s,\theta}, \qquad N = X^{s-1,\theta-1}$$

which are defined in *Section* 2 with $\theta \in (1/2, 1),\ \theta < s - 1/2$.

The *linear mapping property* in Lemma 2.14 states that for solutions of (1.13) we have

$$\|u\|_X \lesssim \|(u_0, u_1)\|_{H^s \times H^{s-1}} + \|F\|_N \tag{1.14}$$

Having (1.14) we also need to know that the mapping written schematically as $\Gamma(u)Q_g(u,u) : X \to N$ holds, as well as

$$\|\Gamma(u)Q_g(u,u) - \Gamma(v)Q_g(v,v)\|_N \lesssim \epsilon \|u - v\|_X$$

for u, v in a ball of radius $C\epsilon$ in X. These two properties are now easily reduced to the estimates (1.15)-(1.19) below; we obtain the ϵ smallness for the difference since the nonlinearity is at least quadratic. We add that this contraction argument gives Lipschitz dependence on u_{lin}, therefore on the initial data by (1.14).

The building blocks for the iteration are the following nonlinear estimates:

Algebra property:	$\|u \cdot v\|_X \lesssim \|u\|_X \|v\|_X$	(1.15)
Product estimate:	$\|u \cdot F\|_N \lesssim \|u\|_X \|F\|_N$	(1.16)
Null form estimate:	$\|Q_g(u,v)\|_N \lesssim \|u\|_X \|v\|_X$	(1.17)
Moser estimate:	$\|\Gamma(u)\|_X \lesssim \|u\|_X (1 + \|u\|_X^M)$.	(1.18)

The proof of the *Algebra property* (1.15) occupies the main part of this paper, being the object of *Section 6* which is based on the results of *Sections 2-5* . All the other properties rely fundamentally on (1.15):

The *Product estimate* (1.16) follows easily in *Section 7* as a corollary of the estimates established in the proof of (1.15), using a duality argument.

We obtain the *Null form estimate* (1.17) as a consequence of the identity (1.5) together with (1.15), (1.16) and the linear bound $\|\Box_g u\|_N \lesssim \|u\|_X$ from Lemma 2.13.

The nonlinear *Moser estimate* (1.18) is proved in *Section 8*. For the purpose of the fixed point iteration argument based on (1.17) we may subtract a constant and assume that $\Gamma(0) = 0$. Using the *Algebra property* (1.15) we may subtract a polynomial from Γ and assume that $\partial^\alpha \Gamma(0) = 0$ for $|\alpha| \le C$. Modifying Γ outside a neighborhood of the origin, it suffices to prove (1.18) under the assumption that Γ and its derivatives are uniformly bounded. Moreover, using the fundamental theorem of calculus, when $\|u\|_X, \|v\|_X$ are bounded we obtain as a consequence

$$\|\Gamma(u) - \Gamma(v)\|_X \lesssim \|u - v\|_X. \tag{1.19}$$

Persistence of regularity. Assuming we control $k \geq 2$ derivatives of the metric, and thus for the rescaled metric on [0,1] we assume $\|\partial^k g\|_{L^2 L^\infty} \ll 1$ and $g \in L^\infty H^{k-1}$, we show that $H^\sigma \times H^{\sigma-1}$ regularity of the initial data is maintained in time for any $\sigma \in [s, k]$. Let $M \geq 1$ and

$$\|u[0]\|_{H^\sigma \times H^{\sigma-1}} \leq M, \qquad \|u[0]\|_{H^s \times H^{s-1}} \leq \epsilon.$$

We obtain our solution as a fixed point for Φ from (1.12), which is a contraction on a small ball of the space $Z = X^{\sigma,\theta} \cap X^{s,\theta}$ endowed with the norm

$$\|u\|_Z = \frac{\epsilon}{M}\|u\|_{X^{\sigma,\theta}} + \|u\|_{X^{s,\theta}}$$

provided ϵ is sufficiently small, independently of M. By Remark 2.15, the mapping property (1.14) holds also for $H^\sigma \times H^{\sigma-1}, X^{\sigma,\theta}$ and $X^{\sigma-1,\theta-1}$. The nonlinear estimates (1.15)-(1.18) are replaced by

$$\begin{aligned}
\|u \cdot v\|_{X^{\sigma,\theta}} &\lesssim \|u\|_{X^{\sigma,\theta}}\|v\|_{X^{s,\theta}} + \|u\|_{X^{s,\theta}}\|v\|_{X^{\sigma,\theta}} \\
\|Q_g(u,v)\|_{X^{\sigma-1,\theta-1}} &\lesssim \|u\|_{X^{s,\theta}}\|v\|_{X^{\sigma,\theta}} + \|u\|_{X^{\sigma,\theta}}\|v\|_{X^{s,\theta}} \\
\|u \cdot F\|_{X^{\sigma-1,\theta-1}} &\lesssim \|u\|_{X^{s,\theta}}\|F\|_{X^{\sigma-1,\theta-1}} + \|u\|_{X^{\sigma,\theta}}\|F\|_{X^{s-1,\theta-1}} \\
\|\Gamma(u)\|_{X^{\sigma,\theta}} &\lesssim \|u\|_{X^{\sigma,\theta}}
\end{aligned}$$

We refer to (6.1), (7.3) and Remark 7.2 for a discussion of these properties. We conclude that the unique solution obtained earlier in $X = X^{s,\theta}$ is also in $X^{\sigma,\theta}$ and $\|u\|_{X^{\sigma,\theta}} \lesssim M$.

In both fixed point arguments we have continuous dependence on the initial data, therefore the obtained solution is the unique strong limit of smooth solutions.

1.4. **Main ideas.** We now provide an outline of the main ideas of the paper which are the ingredients used to establish the building blocks (1.15)-(1.18) of our results.

Curved $X^{s,b}$ spaces. The classical $X^{s,b}$ spaces are multiplier weighted L^2-spaces on the Minkowski space-time adapted to the symbol of the wave operator $\Box$ similarly to how Sobolev spaces are associated to the Laplacian Δ, see [16] and the references therein. These are used to prove local well-posedness for constant coefficients wave maps above scaling: [15] ($d \geq 3$) and [19] ($d = 2$). For the history of classical $X^{s,b}$ spaces and applications we refer to the survey article [20].

A variable coefficient version of the $X^{s,b}$ spaces, defined in physical space, was first introduced[1] in [37], and then further developed in [11] in order to study semilinear wave equations on curved backgrounds with a generic quadratic derivative nonlinearity. These spaces were later utilized by Geba [9] in his treatment of energy-subcritical wave maps in dimensions $3 \leq d \leq 5$. We borrow this notion of curved $X^{s,b}$ spaces in the present paper and review the relevant definitions and lemmata in *Section 2*.

In two space dimensions, new techniques are required to establish $X^{s,b}$ estimates, in particular the crucial algebra property (1.15). Using Strichartz estimates as for the case $d \geq 3$ would incur an unacceptable loss of derivatives when $d = 2$. Instead we adapt some ideas of Tataru from energy-critical wave maps in Minkowski space; see in particular [39, Theorem 3].

[1]Other variable-coefficient $X^{s,b}$ constructions have been proposed, for instance via spectral theory on a smooth compact manifold [3].

Our analysis involves bilinear angular decompositions, wave packets, and energy estimates along certain null hypersurfaces.

Bilinear estimates in $X^{s,b}$ require optimal control of $\|u_\lambda \cdot v_\mu\|_{L^2([0,1]\times\mathbb{R}^2)}$ with bounds depending on the angular frequency localization of u_λ and v_μ. Decomposing the lower frequency function v_μ into *wave packets* concentrated on certain "tubes" T, by orthogonality it will suffice to obtain bounds for $\|u_\lambda\|_{L^2(T)}$. In effect, this strategy is a variable-coefficient variant of the traveling wave decomposition employed in [39]. Foliating a tube T into null surfaces Λ (see *Section 4.1*) we reduce to controlling

$$\int_\Lambda |u_\lambda|^2 \, \mathrm{d}\sigma.$$

We call these *characteristic energy estimates*. While they can be proved by Fourier analysis in the Minkowski case [39], the present context requires a physical space proof which is considerably more delicate, as integration by parts based on the energy-momentum tensor only controls tangential derivatives such as $\int_\Lambda |Lu_\lambda|^2 \, \mathrm{d}\sigma$.

The problem of "inverting the L" in a manner consistent with the angular separation is addressed using microlocal analysis in *Section 5*.

Wave packets analysis. A key technical device in this article consists in approximating solutions to the linear wave equations by square-summable superpositions of *wave packets*, which are localized in both space and frequency on the scale of the uncertainty principle, and propagate in spacetime along null hypersurfaces. Such representations of the wave group originate in the work of Cordoba-Fefferman [5]. For wave equations with $C^{1,1}$ coefficients, the first wave packet parametrix was given by Smith [29]. An alternate construction, based on the use of phase space (FBI) transforms was provided by Tataru [38, 40, 41] for metrics satisfying $\partial^2 g \in L^1 L^\infty$. Analogous constructions have since been employed for even rougher coefficients, as in the the study of quasilinear wave equations in Smith-Tataru [30].

In this article we use Smith's parametrix, which easily extends to rougher metrics satisfying $\partial^2 g \in L^1 L^\infty$. There are two reasons for that: (i) causality (the parametrix in [38, 40, 41] is developed in a microlocal space-time foliation, without reference to a time foliation) and (ii) localization scales (in slabs [29] vs. tubes [38, 40, 41]). Neither of these reasons is critical but taken together they do make a difference from a technical standpoint. *Section 4* isolates the essential properties of Smith's parametrix that we shall need, while specific implementation details are reviewed in *Appendix A*.

Another contribution of this paper is a *wave packet characterization* of functions in $X^{s,b}$ with low modulation, see *Proposition 4.7* and *Corollary 4.8*.

Null structures arise in equations from mathematical physics where they manifest through the vanishing of parallel interactions. The cancellations are realized in our setting expressing the null form $Q_g(u,v)$ in a suitable *null frame* $\{L, \underline{L}, E\}$:

$$2Q_g\big(u,v\big) = Lu \cdot \underline{L}v + \underline{L}u \cdot Lv - 2Eu \cdot Ev.$$

If L, E are tangent to the null hypersurfaces along which v propagates, then Lv and Ev are better than the tranverse derivative $\underline{L}v$. Hence there is a gain in $Q_g(u,v)$ over a generic quadratic term $\nabla u \cdot \nabla v$ if u and v propagate along nearby directions.

Estimates for null forms have played a key role in the low regularity well-posedness of wave maps in the flat case ([15], [19], [20], [39], [34]) and in other problems, especially in low dimensions. Several null form estimates for variable coefficients have been obtained before by Sogge [31], Smith-Sogge [7] and Tataru [42], under various assumptions on the metric.

The strategy outlined above of obtaining L^2 estimates using wave packets and characteristic energy estimates applies as well for bounding $\|Q_g(u_\theta, v_\theta)\|_{L^2([0,1]\times\mathbb{R}^2)}$ where the terms u_θ, v_θ have an angular separation of $\simeq \alpha$. A Whitney-type decomposition is used to reduce to this assumption.

The *Moser estimate* (1.18), required for non-analytic target manifolds, is proved in *Section 8* following the method of iterated multilinear paradifferential expansions introduced in [43].

1.5. **Notations and preliminaries.**

- We denote by τ and ξ the time, respectively the spatial Fourier variables. For $0 \neq \xi$, write $\widehat{\xi} := \xi/|\xi|$ for its projection to the (Euclidean) unit sphere.
- We denote the Cauchy data at time s by $v[s] = (v(s), \partial_t v(s))$ for any time s.
- Mixed Lebesgue norms shall be denoted by $\|u\|_{L^p_t L^q_x} := \big\| \|u(t)\|_{L^q_x} \big\|_{L^q_t}$; to unclutter the notation we often omit the subscripts. Also, when $p = q$ we write $L^p := L^p_t L^p_x$.
- In the context of wave packet sums we will denote by ℓ^2_T the following norm

$$\|c_T\|^2_{\ell^2_T} = \sum_{T\in\mathcal{T}_\lambda} |c_T|^2$$

- To define Littlewood-Paley decompositions we fix a partition of unity of the positive real line

$$1 = \sum_{\lambda\geq 1 \text{ dyadic}} s_\lambda(r) \tag{1.20}$$

where s_λ is supported on the interval $r \in [\lambda/2, 2\lambda]$ for $\lambda \geq 2$ and in $r \leq 2$ for $\lambda = 1$. Setting $P_\lambda(\xi) := s_\lambda(|\xi|)$ and $P_\lambda(\tau,\xi) := s_\lambda(|(\tau,\xi)|)$, we define Littlewood-Paley frequency decompositions on space and spacetime:

$$1_{\mathbb{R}^2} = \sum_\lambda P_\lambda(D_x), \quad 1_{\mathbb{R}\times\mathbb{R}^2} = \sum_\lambda P_\lambda(D_t, D_x).$$

Projections $P_{<\lambda}$, $P_{>\lambda}$ are defined in the usual manner, and we also set $P_{[\lambda_1,\lambda_2]} = \sum_{\lambda\in[\lambda_1,\lambda_2]} P_\lambda$. In this article, summation indices representing frequency or modulation are always understood to be dyadic. More generally, when the index set is not written explicitly it will always be clear from the context.

If s_λ is a compactly supported multiplier, $\tilde{s}_\lambda$ represents slightly wider version of s_λ so that $s_\lambda = \tilde{s}_\lambda s_\lambda$.

In this article, Littlewood-Paley decompositions are defined respect to the x variable with one main exception: the (dual) metric g shall always be mollified in spacetime, and for a frequency μ we write $g_{<\mu} := P_{<\mu}(D_t, D_x)g$.

- Pseudo-differential operators in this paper are defined using Kohn-Nirenberg quantization. When there is no danger of confusion, we sometimes denote both a symbol $\phi(x,\xi)$ and its corresponding operator $\phi(X, D)$ by the abbreviation ϕ.

- The global Sobolev spaces $\dot{H}^s$ and H^s are defined for $s \in \mathbb{R}$ using the Fourier transform by

$$\|u\|_{\dot{H}^s} = \| \, |\xi|^s \, \hat{u}\|_{L^2}, \qquad \|u\|_{H^s} = \|(1+|\xi|^2)^{s/2}\hat{u}\|_{L^2} \tag{1.21}$$

- We define the local Sobolev spaces $W^{s-1,2}(B)$, for a ball B of radius $\simeq 1$ or for $B = \mathbb{R}^2$. When $s-1 \in \{0,1\}$ we use the classical definition, while when $s-1 \in (0,1)$ we have the norm

$$\|v\|^2_{W^{s-1,2}(B)} = \|v\|^2_{L^2(B)} + |v|^2_{\dot{W}^{s-1,2}(B)} \tag{1.22}$$

where $|\cdot|_{\dot{W}^{s-1,2}(B)}$ denotes the Gagliardo seminorm

$$|v|^2_{\dot{W}^{s-1,2}(B)} = \int_B \int_B \frac{|v(x)-v(y)|^2}{|x-y|^{2s}} \, \mathrm{d}x \, \mathrm{d}y. \tag{1.23}$$

- When $B = \mathbb{R}^2$ one has

$$|v|_{\dot{W}^{s-1,2}(\mathbb{R}^2)} \simeq \|v\|_{\dot{H}^{s-1}}, \qquad \|v\|_{W^{s-1,2}(\mathbb{R}^2)} \simeq \|v\|_{H^{s-1}}. \tag{1.24}$$

When $\psi \in C^{0,1}(\tilde{B})$ is supported in a slightly smaller ball one has

$$\|\psi v\|_{H^{s-1}} \lesssim \|\psi v\|_{W^{s-1,2}(\tilde{B})} \lesssim \|v\|_{W^{s-1,2}(\tilde{B})}. \tag{1.25}$$

We refer to [6] for these properties. Clearly, when the balls $(B_j)_j$ are finitely overlapping one has

$$\sum_j |v|^2_{\dot{W}^{s-1,2}(B_j)} \lesssim |v|^2_{\dot{W}^{s-1,2}(\mathbb{R}^2)} \tag{1.26}$$

Conversely, assuming in addition that $(B_j)_j$ cover $\mathbb{R}^2$, by splitting the $\mathbb{R}^2 \times \mathbb{R}^2$ integral in (1.23) into regions where there exists j such that both $x, y \in B_j$ and regions where $|x-y| > \delta$ (where we bound $|v(x)-v(y)| \le |v(x)| + |v(y)|$ and use the $L^2(B_j)$ norms), we obtain

$$|v|^2_{\dot{W}^{s-1,2}(\mathbb{R}^2)} \lesssim \sum_j \|v\|^2_{W^{s-1,2}(B_j)} \tag{1.27}$$

Bounds on the metric. Observe that we have for any λ the pointwise estimate

$$|\partial^2 g_{<\lambda}| \lesssim M(\partial^2 g) \le M(\|\partial^2 g\|_{L^\infty_x}),$$
$$|\partial^k \partial^2 g_{<\lambda}| \lesssim \lambda^k M(\|\partial^2 g\|_{L^\infty_x}),$$

where M is the Hardy-Littlewood maximal function, and by Bernstein

$$\|\partial^k \partial^2 g_{<\lambda}\|_{L^\infty} \lesssim \lambda^k \lambda^{\frac{1}{2}} \|\partial^2 g\|_{L^\infty_x L^2_t} \le \lambda^{\frac{1}{2}+k} \|\partial^2 g\|_{L^2_t L^\infty_x}. \tag{1.28}$$

Similarly, from the bound

$$|g_\lambda| = |P_\lambda D^{-2} D^2 g| \lesssim \lambda^{-2} M(\|\partial^2 g\|_{L^\infty_x})$$

we have

$$\|g_\lambda\|_{L^2_t L^\infty_x} \lesssim \lambda^{-2} \|\partial^2 g\|_{L^2_t L^\infty_x}. \tag{1.29}$$

As a consequence of the above bounds on g, we recall the commutator estimate

$$\|[\Box_{g_{<\sqrt{\lambda}}}, P_\lambda] v\|_{L^2_x} \lesssim \lambda \|v\|_{L^2_x} + \|\partial_t v\|_{L^2_x} \tag{1.30}$$

which follows from

$$[\Box_{g_{<\sqrt{\lambda}}}, P_\lambda] = [g_{<\sqrt{\lambda}}, P_\lambda]\partial_{t,x}\partial_x, \qquad \|[g_{<\sqrt{\lambda}}, P_\lambda]\|_{L^2\to L^2} \lesssim \lambda^{-1}\|\nabla g\|_{L^\infty}.$$

Acknowledgments. The first author was partially supported by the Simons Foundation. The second author was partially supported by an NSF Postdoctoral Fellowship. The third author was partially supported by the NSF grant DMS-1800294 as well as by a Simons Investigator grant from the Simons Foundation.

2. Curved $X^{s,\theta}$ spaces

In this section we present and expand on the theory of $X^{s,\theta}$ spaces associated to wave operators with variable coefficients from [11] (see also [37]). An alternative definition of $X^{s,\theta}$ spaces in the context of time-independent coefficients on compact manifolds was proposed in [3] using spectral theory.

2.1. **Definitions.** We begin with our main building blocks, which are norms associated to a single frequency λ and modulation d. The reader is reminded that sums over such parameters are always understood as dyadic summations.

Definition 2.1. Let $s \in \mathbb{R}$, $\theta \in (0,1)$ and let I be a time interval.

(1) For dyadic $\lambda \geq d \geq 1$, the norm of $X^{s,\theta}_{\lambda,d}[I]$ is defined by

$$\|u\|^2_{X^{s,\theta}_{\lambda,d}[I]} = \lambda^{2s} d^{2\theta} \|u\|^2_{L^2(I\times\mathbb{R}^n)} + \lambda^{2s-2} d^{2\theta-2} \|\Box_{g_{<\sqrt{\lambda}}} u\|^2_{L^2(I\times\mathbb{R}^n)}.$$

When $I = [0,1]$ we drop the I and denote simply $X^{s,\theta}_{\lambda,d}$.

(2) The norms of $X^{s,\theta}_{\lambda,\leq h}, X^{s,\theta}_{\lambda,\leq h,\infty}$ are defined by

$$\|u\|^2_{X^{s,\theta}_{\lambda,\leq h}} = \inf\Big\{ \sum_{d=1}^{h} \|u_{\lambda,d}\|^2_{X^{s,\theta}_{\lambda,d}} \quad ; \quad u = \sum_{d=1}^{h} u_{\lambda,d} \Big\},$$

$$\|u\|^2_{X^{s,\theta}_{\lambda,\leq h,\infty}} = \inf\Big\{ \sup_{1\leq d\leq h} \|u_{\lambda,d}\|^2_{X^{s,\theta}_{\lambda,d}} \quad ; \quad u = \sum_{d=1}^{h} u_{\lambda,d} \Big\},$$

for dyadic $h \leq \lambda$. When $h = \lambda$ we simply write $X^{s,\theta}_{\lambda}$ for $X^{s,\theta}_{\lambda,\leq\lambda}$. We also use a similar definition for $X^{s,\theta}_{\lambda,[d_1,d_2]}$ for a restricted range of dyadic modulations $d_1 \leq d \leq d_2$.

(3) For $\delta > 0$ such that $D = \delta^{-1}$ is a dyadic integer and $|I| \simeq \delta$, we analogously define $X^{s,\theta}_{\lambda,\leq h}[I]$, $X^{s,\theta}_{\lambda,\leq h,\infty}[I]$, $X^{s,\theta}_{\lambda,[d_1,d_2]}[I]$ for $D \leq h \leq \lambda$, with summation, respectively supremum, taken over $D \leq d \leq \lambda$ or $d_1 \leq d \leq d_2$. When D is bounded by a universal constant, we may equivalently take the summation over $1 \leq d \leq \lambda$.

Remark 2.2. There is some flexibility in how to mollify the coefficients of $\Box_g$. The threshold $\sqrt{\lambda}$ is motivated by the hypothesis on $\partial^2 g$ which gives $\|g^{\alpha\beta}_{\geq\sqrt{\lambda}}\|_{L^2_t L^\infty_x} \lesssim \lambda^{-1}$, making $\Box_{g\geq\sqrt{\lambda}} P_\lambda$ effectively a first-order operator. Allowing higher frequencies would yield equivalent norms, and indeed a paradifferential-type cutoff $< \lambda$ would be more natural when considering merely Lipschitz coefficients.

The full iteration spaces are defined as follows.

Definition 2.3. Let $s \in \mathbb{R}$ and $\theta \in (0,1)$.

(1) A function $u \in L^2([0,1], H^s(\mathbb{R}^n))$ is said to be in $X^{s,\theta}$ if it has finite norm defined by

$$\|u\|^2_{X^{s,\theta}} = \inf \Big\{ \sum_{\lambda=1}^{\infty} \sum_{d=1}^{\lambda} \|u_{\lambda,d}\|^2_{X^{s,\theta}_{\lambda,d}} \quad ; \quad u = \sum_{\lambda=1}^{\infty} \sum_{d=1}^{\lambda} P_\lambda u_{\lambda,d} \Big\}$$

(2) A function $f \in L^2([0,1], H^{s+\theta-2}(\mathbb{R}^n))$ is said to be in $X^{s-1,\theta-1}$ if it has finite norm defined by

$$\|f\|^2_{X^{s-1,\theta-1}} = \inf \Big\{ \|f_0\|^2_{L^2 H^{s-1}} + \sum_{\lambda=1}^{\infty} \sum_{d=1}^{\lambda} \|f_{\lambda,d}\|^2_{X^{s,\theta}_{\lambda,d}};\ f = f_0 + \sum_{\lambda=1}^{\infty} \sum_{d=1}^{\lambda} \Box_{g<\sqrt{\lambda}} P_\lambda f_{\lambda,d} \Big\}$$

Remark 2.4. To motivate these definitions, note that the $X^{s,\theta}$ norm is obtained essentially by interpolation between $\|u\|_{L^2 H^s}$ and $\|u\|_{L^2 H^s} + \|\Box_g u\|_{L^2 H^{s-1}}$, see (2.12). Since in the variable-coefficient context modulation cannot be precisely interpreted in terms of localization in Fourier space, the advantage of this concrete definition is that we get access to the notion of modulation in a more robust way.

For negative modulation regularity, i.e. $X^{s-1,\theta-1}$, this particular definition makes it easy to invert $\Box_g$ from $X^{s-1,\theta-1}$ to $X^{s,\theta}$ (see (2.21) and (2.20)) while at the same time obtaining the duality relation that one expects by analogy from the classical flat spaces, see (2.13).

Remark 2.5. Roughly speaking, $X^{s,\theta}_{\lambda,d}$ will hold the portion of u at frequency λ and "modulation" d. If g is the Minkowski metric and we modify the above definition of $X^{s,\theta}$ by replacing $L^2([-1,1]\times\mathbb{R}^n)$ with $L^2(\mathbb{R}\times\mathbb{R}^n)$, then for $u_{\lambda,d}$ localized to frequency $|\xi| \simeq \lambda$ and modulation $\big||\tau| - |\xi|\big| \simeq d$ with $d \leq \lambda$ we have

$$\|u_{\lambda,d}\|_{X^{s,\theta}_\lambda} \simeq \lambda^s d^\theta \|u_{\lambda,d}\|_{L^2} \simeq \|u_{\lambda,d}\|_{X^{s,\theta}_{\lambda,d}}.$$

On the other hand, if $d \gg \lambda$,

$$\|u_{\lambda,d}\|_{X^{s,\theta}_\lambda} \simeq \lambda^{s+\theta-2} \|\Box u_{\lambda,d}\|_{L^2} \simeq \|\Box u_{\lambda,d}\|_{L^2 H^{s+\theta-2}}.$$

Remark 2.6. By [11, Corollary 2.5], in the definition of the $X^{s,\theta}$ and $X^{s-1,\theta-1}$ spaces one can replace the $X^{s,\theta}_{\lambda,d}$ norm by the norm of $\bar{X}^{s,\theta}_{\lambda,d}$ defined by

$$\|u\|^2_{\bar{X}^{s,\theta}_{\lambda,d}} = \lambda^{2s-2} d^{2\theta} \|\nabla_{t,x} u\|^2_{L^2} + \lambda^{2s-2} d^{2\theta-2} \|\Box_{g<\sqrt{\lambda}} u\|^2_{L^2}.$$

This is based on the estimate in [11, Lemma 2.4]:

$$\lambda^{s-1} d^\theta \|\nabla_{t,x} P_\lambda u\|_{L^2} + \lambda^{s-1} d^{\theta-1} \|\Box_{g<\sqrt{\lambda}} P_\lambda u\|_{L^2} \lesssim \|u\|_{X^{s,\theta}_{\lambda,d}}. \tag{2.1}$$

This bound shows that on frequency localized functions, the $X^{s,\theta}_{\lambda,d}$ and $\bar{X}^{s,\theta}_{\lambda,d}$ norms are comparable and also that in Definition 2.3 one can assume that $u_{\lambda.d}$ and $f_{\lambda,d}$ are localized at frequency λ. Moreover, based on (2.1) we have the straightforward embedding

$$X^{s-1,\theta-1} \subset L^2H^{s+\theta-2}. \tag{2.2}$$

2.2. **Basic properties.** Here we show how some properties known for the classical $X^{s,b}$ spaces generalize to variable coefficients. We begin with some properties for our frequency localized building blocks:

Proposition 2.7. *Let u be a λ-frequency-localized function on $[0,1]\times\mathbb{R}^n$. Then:*

(1) (Energy estimates) For $d \gtrsim |I|^{-1}$ and any v one has:

$$\lambda^{s-1}d^{\theta-\frac{1}{2}}\|\nabla_{t,x}u\|_{L^\infty L^2[I]} \lesssim \|u\|_{X^{s,\theta}_{\lambda,d}[I]} \tag{2.3}$$

$$\|v\|_{CH^s\cap C^1H^{s-1}} \lesssim \|v\|_{X^{s,\theta}} \quad \text{if} \quad \theta > \frac{1}{2}. \tag{2.4}$$

(2) (Time localization) Let $1 \le d_1 \le d_2 \le \lambda$ and let $\chi_{d_2}(t)$ be a bump function in time localized on the d_2^{-1} scale. Then

$$\chi_{d_2} : \tilde{S}_\lambda X^{1,\frac{1}{2}}_{\lambda,d_1} \to X^{1,\frac{1}{2}}_{\lambda,d_2} \qquad i.e. \qquad \|\chi_{d_2}u\|_{X^{1,\frac{1}{2}}_{\lambda,d_2}} \lesssim \|u\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}}. \tag{2.5}$$

(3) (Global extension) There exists a frequency localized extension of u from $[0,1]$ to $\tilde{u}$ supported in $(d^{-1}, 1+d^{-1})$ such that

$$d^\theta\|\nabla_{t,x}\tilde{u}\|_{L^2(\mathbb{R}\times\mathbb{R}^n)} + d^{\theta-1}\|\Box_{g_{<\sqrt{\lambda}}}\tilde{u}\|_{L^2(\mathbb{R}\times\mathbb{R}^n)} \lesssim \|u\|_{X^{1,\theta}_{\lambda,d}}.$$

(4) (Time orthogonality) Let $1 \le d \le d' \le d'' \le \lambda$. For smooth partitions of unity with respect to time intervals of length d^{-1}: $1 = \sum_j \chi^j_d(t)$, one has

$$\|u\|^2_{X^{s,\theta}_{\lambda,d'}} \simeq \sum_j \|\chi^j_d(t)u\|^2_{X^{s,\theta}_{\lambda,d'}} \tag{2.6}$$

$$\|u\|^2_{X^{s,\theta}_{\lambda,[d',d'']}} \simeq \sum_j \|\chi^j_d(t)u\|^2_{X^{s,\theta}_{\lambda,[d',d'']}} \tag{2.7}$$

(5) (Scaling) Set $u^\delta(t,x) = u(\delta t, \delta x)$ defined on $[0,1]$ where $\delta d \ge 1$. Then:

$$\|u^\delta\|_{X^{s,\theta}_{\delta\lambda,\delta d}} \simeq \delta^{s+\theta-\frac{n+1}{2}}\|u\|_{X^{s,\theta}_{\lambda,d}[0,\delta]} \tag{2.8}$$

where $X^{s,\theta}_{\delta\lambda,\delta d}$ is defined using the metric $g^\delta(t,x) = g(\delta t, \delta x)$.

Proof. **(1)** By energy estimates for the wave equation we obtain

$$\|\nabla_{t,x}u\|^2_{L^\infty L^2[I]} \lesssim |I|^{-1}\|\nabla_{t,x}u\|^2_{L^2[I]} + \|\nabla_{t,x}u\|_{L^2[I]}\|\Box_{g_{<\sqrt{\lambda}}}u\|_{L^2[I]}$$

which implies (2.3) by (2.1). From this, for $\theta > \frac{1}{2}$ and $I = [0,1]$ we sum over d and use the definition of $X^{s,\theta}$ to obtain the $L^\infty H^s \times L^\infty H^{s-1}$ bound in (2.4).

Now we prove that the map

$$t \in [0,1] \to H^s \times H^{s-1} \ni v[t]$$

is continuous when $v \in X^{s,\theta}$. By the Fundamental theorem of calculus and Cauchy-Schwarz in t, by summing modulations we have

$$\|v_\lambda(t+h) - v_\lambda(t)\|_{H^s} \leq C_\lambda \, |h|^{\frac{1}{2}} \, \|v_\lambda\|_{X^{s,\theta}_\lambda}$$

for any $t, t+h \in [0,1]$. Let $\epsilon > 0$ and define λ_ϵ such that

$$\sum_{\lambda > \lambda_\epsilon} \|v_\lambda\|^2_{L^\infty H^s} \leq \epsilon^2.$$

Then

$$\|v(t+h) - v(t)\|^2_{H^s} \lesssim \sum_{\lambda \leq \lambda_\epsilon} \|v_\lambda(t+h) - v_\lambda(t)\|^2_{H^s} + \epsilon^2 \leq \tag{2.9}$$

$$\leq |h|^{\frac{1}{2}} \, C_\epsilon \|v\|_{X^{s,\theta}_\lambda} + \epsilon^2 \lesssim \varepsilon^2. \tag{2.10}$$

for h small enough. The same argument applies for $\partial_t v$.

(2) By Hölder's inequality and (2.3) we have

$$\lambda d_2^{\frac{1}{2}} \|\chi_{d_2} u\|_{L^2} \lesssim \lambda d_2^{\frac{1}{2}} \|\chi_{d_2}\|_{L^2_t} \|u\|_{L^\infty L^2} \lesssim \|u\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}}.$$

For the term $\Box_{g<\sqrt{\lambda}}(\chi_{d_2} u)$ we consider

$$d_2^{-\frac{1}{2}} \|\chi_{d_2} \Box_{g<\sqrt{\lambda}} u\|_{L^2} \lesssim d_1^{-\frac{1}{2}} \|\Box_{g<\sqrt{\lambda}} u\|_{L^2} \lesssim \|u\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}}$$

$$d_2^{-\frac{1}{2}} \|\partial_t^2 \chi_{d_2} u\|_{L^2} \lesssim d_2^{-\frac{1}{2}} \|\partial_t^2 \chi_{d_2}\|_{L^2_t} \|u\|_{L^\infty L^2} \lesssim \|u\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}}$$

$$d_2^{-\frac{1}{2}} \|\partial_t \chi_{d_2} \partial_{t,x} u\|_{L^2} \lesssim d_2^{-\frac{1}{2}} \|\partial_t \chi_{d_2}\|_{L^2_t} \|\partial_{t,x} u\|_{L^\infty L^2} \lesssim \|u\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}}$$

(3) First we extend u by solving $\Box_{g_{<\sqrt{\lambda}}} u = 0$ outside $[0,1]$ and then we define $\tilde{u} = \chi \tilde{P}_\lambda u$ where $\chi(t)$ is smooth time cutoff supported in $(-d^{-1}, 1+d^{-1})$ equal to 1 on $[0,1]$. For $t \in [1, 1+d^{-1}]$

$$\|\nabla_{t,x} u(t)\|_{L^2_x} \lesssim \|\nabla_{t,x} u(t-d^{-1})\|_{L^2} + \|\Box_{g_{<\sqrt{\lambda}}} u\|_{L^1 L^2([1-d^{-1}, 1+d^{-1}] \times \mathbb{R}^n)}$$

$$\lesssim \|\nabla_{t,x} u(t-d^{-1})\|_{L^2_x} + d^{-1/2} \|\Box_{g_{<\sqrt{\lambda}}} u\|_{L^2}.$$

Therefore

$$\|\nabla_{t,x} \tilde{u}\|_{L^2([1,1+d^{-1}], \mathbb{R}^n)} \lesssim \|\nabla_{t,x} u\|_{L^2[0,1]} + d^{-1} \|\Box_{g_{<\sqrt{\lambda}}} u\|_{L^2[0,1]}.$$

We repeat this analysis on the interval $[-d^{-1}, 0]$. For the term $\Box_{g_{<\sqrt{\lambda}}} \tilde{u}$ outside of $[0,1]$ we write

$$\Box_{g_{<\sqrt{\lambda}}} \chi \tilde{P}_\lambda = \chi''(t) \tilde{P}_\lambda + 2\chi'(t) \partial_{t,x} \tilde{P}_\lambda + \chi [\Box_{g_{<\sqrt{\lambda}}}, \tilde{P}_\lambda] + \chi \tilde{P}_\lambda \Box_{g_{<\sqrt{\lambda}}}.$$

Then we use the already established L^2 bounds together with (1.30).

(4) See [11, (53)], which is an easy commutation argument (see also Remark 2.6).

(5) The equivalence follows from a change of variables. The only issue is that the scaling does not commute with taking the $P_{<\sqrt{\lambda}}$ localization for g. Instead we have $(g_{<\sqrt{\lambda \delta^{-1}}})^\delta = (g^\delta)_{<\sqrt{\delta\lambda}}$. It remains to estimate

$$\lambda^{-1} d^{-1} \|\Box_{g_{[\sqrt{\lambda}, \sqrt{\lambda\delta^{-1}}]}} u\|_{L^2} \lesssim \lambda^{-2} d^{-1} \|\partial^2 g_{[\sqrt{\lambda}, \sqrt{\lambda\delta^{-1}}]}\|_{L^2 L^\infty} \|\partial^2 u\|_{L^\infty L^2} \lesssim \|u\|_{X^{0,0}_{\lambda,d}[0,\delta]}.$$

□

We next drop the frequency localization. Then one may ask whether we could also discard the frequency localization in the coefficients. This is indeed the case for a restricted range of Sobolev indices. Precisely, we have

Proposition 2.8. *Let $0 \le s \le 2$ and $0 < \theta < 1$. Then $u \in X^{s,\theta}$ iff it admits a decomposition*

$$u = \sum_{d \ge 1} u_d$$

so that

$$\sum_d d^{2\theta} \|\nabla u_d\|^2_{L^2 H^{s-1}} + d^{2\theta-2} \|\Box_g u_d\|^2_{L^2 H^{s-1}} < \infty$$

with an equivalent norm given by

$$\|u\|^2_{X^{s,\theta}} = \inf_{u = \sum u_d} \sum_d \big[d^{2\theta} \|\nabla u_d\|^2_{L^2 H^{s-1}} + d^{2\theta-2} \|\Box_g u_d\|^2_{L^2 H^{s-1}} \big] \tag{2.11}$$

We note that the counterpart of this result for $X^{s-1,\theta-1}$ is also valid.

Proof. a) In one direction, given $u_{\lambda,d}$ we define

$$u_d = \sum_{\lambda > d} u_{\lambda,d}$$

and prove the appropriate bounds for u_d. Then by orthogonality

$$\|\nabla u_d\|^2_{L^2 H^{s-1}} \lesssim \sum_{\lambda > d} \|\nabla u_{\lambda,d}\|^2_{L^2 H^{s-1}}$$

$$\|\Box_g u_d\|^2_{L^2 H^{s-1}} \lesssim \sum_{\lambda > d} \|\Box_{g_{<\sqrt{\lambda}}} u_{\lambda,d}\|^2_{L^2 H^{s-1}} + \| \sum_{\lambda > d} \Box_{g_{\ge \sqrt{\lambda}}} u_{\lambda,d}\|^2_{L^2 H^{s-1}},$$

and the rightmost term on the second line is controlled by the estimates

$$\| \sum_{\lambda > d} P_{<\lambda} \Box_{g_{<\sqrt{\lambda}}} u_{\lambda,d}\|_{L^2 H^{s-1}} \lesssim \lambda^{s-1} \|\partial^2 g\|_{L^2 L^\infty} \|\nabla u_{\lambda,d}\|_{L^\infty L^2}$$

$$\|P_\mu \Box_{g_{\ge \sqrt{\lambda}}} u_{\lambda,d}\|_{L^2 H^{s-1}} \lesssim \big(\frac{\lambda}{\mu}\big)^{2-s} \lambda^{s-1} \|\partial^2 g\|_{L^2 L^\infty} \|\nabla u_{\lambda,d}\|_{L^\infty L^2}, \quad \mu > \lambda,$$

and the $X^{s,\theta}$ energy bound.

b) In the opposite direction, we define

$$u_{\lambda,d} = P_\lambda u_d, \qquad d < \lambda$$

and

$$u_{\lambda,\lambda} = \sum_{d > \lambda} P_\lambda u_d$$

and again prove the appropriate bounds. □

Continuing our global description of the $X^{s,\theta}$ spaces, for $s \in [0,2]$ we define the endpoints $X^{s,0}$ and $X^{s,1}$ with norms

$$\|u\|^2_{X^{s,0}} = \|\nabla u\|^2_{L^2 H^{s-1}}$$

respectively

$$\|u\|^2_{X^{s,1}} = \|\nabla u\|^2_{L^2 H^{s-1}} + \|\Box_g u\|^2_{L^2 H^{s-1}}$$

Then we can also describe the full family of $X^{s,\theta}$ spaces as follows:

Proposition 2.9. *For $0 \le s \le 2$ and $0 \le \theta < 1$ we have:*

(1) (Interpolation) The space $X^{s,\theta}$ can be described by interpolation as

$$X^{s,\theta} = [X^{s,0}, X^{s,1}]_\theta \tag{2.12}$$

(2) (Duality) For $\frac{1}{2} < \theta < 1$ we have

$$X^{s-1,\theta-1} = (X^{1-s,1-\theta} + L^2H^{2-s-\theta})'. \tag{2.13}$$

We remark that the first part of the proposition in particular shows that the spaces $X^{s,\theta}$ defined in [37] and in [44] coincide.

Proof. **(1)** Examining the equivalent definition of the $X^{s,\theta}$ spaces in Proposition 2.8, one immediately sees that it is nothing but the real $(\theta, 2)$ interpolation space between $X^{s,0}$ and $X^{s,1}$ constructed using the J-method. Since these are Hilbert spaces, the outcome coincides with the one provided by complex interpolation.

(2) This is proved in [11, Lemma 2.13]. □

It will be technically convenient at some junctures to view functions in $X^{s,\theta}[I]$ as restrictions of globally defined functions. For this purpose we assume that the metric g is extended to a global Lorentzian metric in $\mathbb{R}^{1+n}$. This can be taken constant outside a compact time interval.

Corollary 2.10. *In the definition of $X^{s,\theta}_\lambda$, we may assume in the decomposition $u_\lambda = \sum u_{\lambda,d}$ that $\operatorname{supp} u_{\lambda,d} \subset (-d^{-1}, 1+d^{-1}) \times \mathbb{R}^n$ and take all spacetime norms over $\mathbb{R}^{1+n}$. The same holds for other intervals.*

We also have

Lemma 2.11. *Let $I = [0, \delta]$ and $D = \delta^{-1}$. If $v[0] = (0, 0)$, then*

$$\|\tilde{P}_\lambda v\|_{X^{0,\frac{1}{2}}_{\lambda,D}[I]} \lesssim \lambda^{-1} D^{-\frac{1}{2}} \|\Box_{g_{<\sqrt{\lambda}}} \tilde{P}_\lambda v\|_{L^2[I]}. \tag{2.14}$$

Proof. Since $v[0] = (0, 0)$, the estimate follows from energy estimates. □

2.3. **Linear mapping properties.** Consider the linear problem

$$\Box_g v = f, \qquad v[0] = (v_0, v_1). \tag{2.15}$$

Lemma 2.12. *Let $0 < s < 3$ and $0 < \theta \le 1$. Then the linear equation (2.15) is well-posed in $H^s \times H^{s-1}$ and*

$$\|v\|_{X^{s,\theta}} \lesssim \|(v_0, v_1)\|_{H^s \times H^{s-1}} + \|f\|_{L^2H^{s-1}}. \tag{2.16}$$

Proof. This is proved in [11, Lemma 2.11]. For the sake of completeness we recall the argument. We write

$$v = \sum_{\lambda=1}^{\infty} P_\lambda \tilde{P}_\lambda v$$

and apply Definitions (2.3) and (2.1):

$$\begin{aligned}\|v\|^2_{X^{s,\theta}} &\lesssim \sum_{\lambda=1}^{\infty} \|\tilde{P}_\lambda v\|^2_{X^{s,\theta}_{\lambda,1}} \\ &\lesssim \|v\|^2_{L^2H^s} + \sum_{\lambda=1}^{\infty} \lambda^{2(s-1)} \|\Box_{g_{<\sqrt{\lambda}}} \tilde{P}_\lambda v - \tilde{P}_\lambda \Box_g v\|^2_{L^2} + \|f\|^2_{H^{s-1}}.\end{aligned}$$

For the first term we use energy estimates, while for the second write

$$\Box_{g_{<\sqrt{\lambda}}} \tilde{P}_\lambda - \tilde{P}_\lambda \Box_g = [\Box_{g_{<\sqrt{\lambda}}}, \tilde{P}_\lambda] - \tilde{P}_\lambda \Box_{g_{>\sqrt{\lambda}}}$$

For the commutator use (1.30), while for the last term use the fixed-time bound

$$\sum_{\lambda=1}^{\infty} \lambda^{2(s-1)} \|\tilde{P}_\lambda (g_{>\sqrt{\lambda}} u)\|^2_{L^2} \lesssim M(\|\partial^2 g\|_{L^\infty})^2 \|u\|^2_{H^{s-2}}. \tag{2.17}$$

which follows from Littlewood-Paley arguments (see [11, Lemma 2.9]). □

Next, we recall the following mapping and its proof from [9, Proposition 3.1]:

Lemma 2.13. *For $0 < s < 3$ and $0 \le \theta \le 1$, the operator $\Box_g$ maps*

$$\Box_g : X^{s,\theta} \to X^{s-1,\theta-1}.$$

Proof. Let $u \in X^{s,\theta}$. Using a decomposition from Definition (2.3) we write

$$\Box_g u = \sum_{\lambda=1}^{\infty} \sum_{d=1}^{\lambda} \Box_{g_{<\sqrt{\lambda}}} P_\lambda u_{\lambda,d} + \sum_{\lambda=1}^{\infty} g_{>\sqrt{\lambda}} P_\lambda \partial^2 u_\lambda$$

Using the definition of $X^{s-1,\theta-1}$ the first term is clearly bounded and we place the second term in L^2H^{s-1}. We do this using the following fixed-time bound

$$\|\sum_{\lambda=1}^{\infty} g_{>\sqrt{\lambda}} P_\lambda f_\lambda\|^2_{H^{s-1}} \lesssim \sum_{\lambda=1}^{\infty} \lambda^{2(s-2)} \|f_\lambda\|^2_{L^2} \tag{2.18}$$

which is dual to (2.17) (see [11, Lemma 2.9]), together with (2.4). □

Finally, we want to extend and refine the linear bound (2.16) to apply to the Laplace-Beltrami operator $\tilde{\Box}_g$ given by (1.3) and to source terms in $X^{s-1,\theta-1}$.

Lemma 2.14. *Let $1 \le s \le 2$ and $\frac{1}{2} < \theta < 1$. The solution of the linear equation*

$$\tilde{\Box}_g u = F, \qquad u[0] = (u_0, u_1) \tag{2.19}$$

satisfies

$$\|u\|_{X^{s,\theta}} \lesssim \|(u_0, u_1)\|_{H^s \times H^{s-1}} + \|F\|_{X^{s-1,\theta-1}}. \tag{2.20}$$

Proof. **(1)** We first recall from [11, Lemma 2.12] that

$$\Box_g^{-1} : X^{s-1,\theta-1} \to X^{s,\theta} \tag{2.21}$$

where $\Box_g^{-1} f$ is the solution w of the inhomogeneous equation (2.15) with $w[0] = (0,0)$. To prove this, assume $f \in X^{s-1,\theta-1}$, which gives a representation

$$f = f_0 + \sum_{\lambda=1}^{\infty} \sum_{d=1}^{\lambda} \Box_{g<\sqrt{\lambda}} P_\lambda f_{\lambda,d} = f_0 + \sum_{\lambda=1}^{\infty} \Box_{g<\sqrt{\lambda}} P_\lambda f_\lambda.$$

The term

$$u = \sum_{\lambda=1}^{\infty} \sum_{d=1}^{\lambda} P_\lambda f_{\lambda,d}$$

is in $X^{s,\theta}$ by definition. Thus it remains to write $\Box_g^{-1} f = u + v$ and show that $v \in X^{s,\theta}$ using (2.16). The function v solves

$$\Box_g v = f_0 - \sum_{\lambda=1}^{\infty} g_{>\sqrt{\lambda}} P_\lambda \partial^2 f_\lambda, \qquad v[0] = -u[0].$$

The contribution of the initial data $-u[0]$ is controlled by (2.16), (2.4), while $f_0 \in L^2 H^{s-1}$. For the rest of the inhomogenity one uses (2.18) and (2.3).

(2) Now we return to the proof of (2.20). We write

$$\tilde{\Box}_g - \Box_g = h^\alpha \partial_\alpha$$

where

$$\|h^\alpha\|_{L^2 L^\infty} + \|\nabla_x h^\alpha\|_{L^2 L^\infty} \ll 1 \tag{2.22}$$

and we will treat $h^\alpha \partial_\alpha u$ as a perturbation. A function u solves (2.19) if $u = \Phi(u)$ where $\Phi(v) := S(u_0, u_1) + \Box_g^{-1}(F - h^\alpha \partial_\alpha v)$. Here $S(v_0, v_1)$ denotes the solution of (2.15) with $f = 0$. To show that $\Phi : X^{s,\theta} \to X^{s,\theta}$ and that Φ is a contraction on $X^{s,\theta}$, considering (2.21), (2.16) and (2.4) it remains to check that

$$\|h^\alpha \partial_\alpha v\|_{L^2 H^{s-1}} \ll \|\nabla_{t,x} v\|_{L^\infty H^{s-1}} \lesssim \|v\|_{X^{s,\theta}}. \tag{2.23}$$

Using (2.22) we obtain this bound first for $s = 1$ and $s = 2$ and by interpolation also for $1 \le s \le 2$. □

Remark 2.15 (Higher regularity). Let $k \ge 3$. Assuming control of k derivatives of the metric g, we obtain the previous properties for a wider range of s. Thus, [11][Lemma 2.9] will hold for $-k+2 < s < k+1$, which implies that (2.16), (2.21) and Lemma 2.13 extend to this range of s.

Assuming $\|\partial^k g\|_{L^2 L^\infty} \ll 1$ (since g is a rescaled metric) we extend (2.20) to $s \in [1, k]$. We also refer to [29][Theorem 4.7] for the fact that the parametrix representation with $H^s \times H^{s-1}$ bounds extends to $s \in [1, k]$ under the assumption $g \in L^\infty H^{k-1}$.

2.4. $\tilde{X}^{s,\theta}$ **spaces.** In the proof of the Moser estimate (1.18) it will be useful to have the following modification of the $X^{s,\theta}_\lambda$ norms.

Definition 2.16. Let $s \in \mathbb{R}$, $\theta \in (0,1)$ and dyadic $\lambda \geq 1$.

(1) The norm of $\tilde{X}^{s,\theta}_{\lambda,\lambda}$ is defined for λ-frequency localized functions u by

$$\|u\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} = \lambda^{s+\theta}\|\nabla_{t,x}u\|_{L^2} + \lambda^{s+\theta-\frac{1}{2}}\|\nabla_{t,x}u\|_{L^\infty L^2}.$$

For $1 \leq d < \lambda$ the norm of $\tilde{X}^{s,\theta}_{\lambda,d}$ is defined by $\|u\|_{\tilde{X}^{s,\theta}_{\lambda,d}} = \|u\|_{X^{s,\theta}_{\lambda,d}}$.

(2) The norm of $\tilde{X}^{s,\theta}_{\lambda}$ is defined for λ-frequency localized functions u by

$$\|u\|^2_{\tilde{X}^{s,\theta}_\lambda} = \inf\ \Big\{ \sum_{d=1}^{\lambda/2} \|u_{\lambda,d}\|^2_{X^{s,\theta}_{\lambda,d}} + \|u_{\lambda,\lambda}\|^2_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} \quad ; \quad u = \sum_{d=1}^{\lambda/2} u_{\lambda,d} + u_{\lambda,\lambda} \Big\},$$

(3) A function $u \in L^2([0,1], H^s)$ is said to be in $\tilde{X}^{s,\theta}$ if it has finite norm defined by

$$\|u\|^2_{\tilde{X}^{s,\theta}} = \inf\ \Big\{ \sum_{\lambda=1}^{\infty} \|u_\lambda\|^2_{\tilde{X}^{s,\theta}_\lambda} \quad ; \quad u = \sum_{\lambda=1}^{\infty} P_\lambda u_\lambda \Big\}$$

where the u_λ can be assumed wlog to be λ-frequency localized.

Clearly $X^{s,\theta} \subset \tilde{X}^{s,\theta}$. The only difference between the two spaces occurs at high modulations, where we discarded the terms $\lambda^{s+\theta-2}\|\Box_{g_{<\sqrt{\lambda}}} u_{\lambda,\lambda}\|_{L^2}$, making the norm $\tilde{X}^{s,\theta}$ smaller. This will be useful in the proof of the Moser estimate (1.18), see Remark 8.2. We can recover the $X^{s,\theta}$ bound if we control high modulations through $\Box_g$:

Lemma 2.17. *If $f \in \tilde{X}^{s,\theta}$ and $\Box_g f \in L^2H^{s+\theta-2}$ then $f \in X^{s,\theta}$ and*

$$\|f\|_{X^{s,\theta}} \lesssim \|f\|_{\tilde{X}^{s,\theta}} + \|\Box_g f\|_{L^2H^{s+\theta-2}}.$$

The proof is straightforward using definitions 2.1, 2.3, 2.16 and is omitted.

2.5. **Half-waves norms.** For microlocal analysis purposes, we would like an equivalent definition of the $X^{s,\theta}_{\lambda,d}$ norms in terms of half-waves. We factor the symbol of $\Box_{g_{<\sqrt{\lambda}}}$ as

$$\tau^2 - 2g^{0j}_{<\sqrt{\lambda}}\tau\xi_j - g^{ab}_{<\sqrt{\lambda}}\xi_a\xi_b = (\tau + a^+)(\tau + a^-),$$

where

$$a^\pm = -g^{0j}_{<\sqrt{\lambda}}\xi_j \mp \sqrt{(g^{0j}_{<\sqrt{\lambda}}\xi_j)^2 + g^{ab}_{<\sqrt{\lambda}}\xi_a\xi_b} = -g^{0j}_{\sqrt{\lambda}}\xi_j \mp a.$$

Write

$$A^\pm = -g^{0j}_{<\sqrt{\lambda}}(t,x)D_j \mp a(t,x,D).$$

Note that the frequency subscripts are omitted from a and A because a is not exactly localized at frequencies $< \sqrt{\lambda}$. However one does have

$$P_{\geq\lambda}(D_t, D_x)a \in \lambda^{-N}S^1 \text{ for any } N,$$

where S^1 is the classical symbol class, which is more than enough decay for our purposes.

Lemma 2.18. *In the definition of $X^{s,\theta}_{\lambda,d}$, $\|\Box_{g_{<\sqrt{\lambda}}} P_\lambda u\|_{L^2}$ may be replaced by $\|(D_t + A^-)(D_t + A^+)P_\lambda u\|_{L^2}$ or $\|(D_t + A^+)(D_t + A^-)P_\lambda u\|_{L^2}$.*

Proof. We have

$$\begin{aligned}&(D_t + A^-)(D_t + A^+)P_\lambda - (D_t + A^-)P_{>\lambda/8}(D_t + A^+)P_{>\lambda/8}P_\lambda\\ &= (D_t + A^-)P_{<\lambda/8}(D_t + A^+)P_\lambda = (D_t + A^-)P_{<\lambda/8}A^+P_\lambda\\ &= P_{<\lambda/8}(D_tA^+)P_\lambda + P_{<\lambda/8}A^+P_\lambda D_t + A^-P_{<\lambda/8}A^+P_\lambda,\end{aligned}$$

In view of the off-diagonal estimates

$$\|(P_{<\lambda/8} + P_{>8\lambda})AP_\lambda\|_{L^2\to L^2} = \|(P_{<\lambda/8} + P_{>8\lambda})A_{>\lambda/8}P_\lambda\|_{L^2\to L^2} = O(\lambda^{-\infty}),$$

where $A_{>\lambda/8} = (P_{>\lambda/8}(D_x)a)(t, x, D)$, and the pseudo-differential calculus on $\mathbb{R}^{1+n}$, we have

$$\begin{aligned}&\|(\Box_{g<\sqrt{\lambda}} - (D_t + A^-)(D_t + A^+))P_\lambda u\|_{L^2}\\ &\lesssim \|(\Box_{g<\sqrt{\lambda}} - (D_t + A^-)P_{>\lambda/8}(D_t + A^+)P_{>\lambda/8})P_\lambda u\|_{L^2} + O(\lambda^{-\infty})\|\nabla_{t,x}u\|_{L^2}\\ &\lesssim \|\nabla_{t,x}u\|_{L^2}.\end{aligned}$$

□

Just as in Minkowski space, for bilinear $X^{s,\theta}$ estimates it is helpful to split each factor into half-waves, approximately localized to positive or negative temporal frequencies (see the discussion in Step 1 of the proof of Prop. 6.2). An alternative construction of half-waves decompositions can be obtained using the $\pm$-wave packets from Section 4. For any frequency λ and modulation $d \in [1, \lambda]$, let

$$\|u\|_{X^{s,\theta,\pm}_{\lambda,d}} := \lambda^{s-1}d^\theta\big(\|\nabla_{t,x}u^\pm\|_{L^2} + \lambda d^{-1}\|(D_t \pm A^\pm)u^\pm\|_{L^2} + d^{-1}\|\Box_{g<\sqrt{\lambda}}u^\pm\|_{L^2}\big).$$

Proposition 2.19. *For any $u : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}$ satisfying $u = P_\lambda(D_x)u$, there is a decomposition $u = u^+ + u^-$ such that $u^\pm$ are also localized to spatial frequencies $\sim \lambda$ and*

$$\|u^\pm\|_{X^{s,\theta,\pm}_{\lambda,d}} \lesssim \|u\|_{X^{s,\theta}_{\lambda,d}}.$$

Proof. Write

$$u = P_{\geq\lambda/64}(D_t)u_\lambda + P_{<\lambda/64}(D_t)u_\lambda := u^+ + u^-;$$

note that we truncate smoothly D_t on the same scale as in D_x in order to make use of standard pseudo-differential calculus. The estimates will then follow from the ellipticity of $D_t + A^\mp$ on the microsupport of $u^\pm$.

We supply the details for u^+. Write $P^+(D_t) = P_{>-\lambda/64}(D_t)$ and $P^-(D_t) = P_{<-\lambda/64}(D_t)$. Introduce the slightly enlarged multiplier

$$T_\lambda = P_{\geq\lambda/128}(D_t)\widetilde{P}_\lambda(D_x).$$

Then

$$\|(1 - T_\lambda)(D_t + A^+)P^+(D_t)P_\lambda(D_x)\|_{L^2\to L^2} = \|(1 - T_\lambda)A^+_{>\lambda/64}P^+(D_t)P_\lambda(D_x)\|_{L^2\to L^2} = O(\lambda^{-\infty}),$$

where A^+_λ is mollified in (t, x). Now on the support of T_λ, the symbol $\tau + a^-$ is elliptic and belongs to $S^1_{1,\frac{3}{4}}(\mathbb{R}^{1+n})$, hence there is a parametrix $Q \in OPS^{-1}_{1,\frac{3}{4}}(\mathbb{R}^{1+n})$ such that $Q(D_t + A^-) + R = T_\lambda$ with $R \in S^{-\infty}(\mathbb{R}^{1+n})$.

Write

$$\begin{aligned}(D_t + A^+)u^+ &= Q(D_t + A^-)T_\lambda(D_t + A^+)u^+ + RT_\lambda(D_t + A^+)u^+ + (1 - T_\lambda)(D_t + A^+)u^+ \\ &= Q\tilde{T}_\lambda(D_t + A^-)T_\lambda(D_t + A^+)u^+ \\ &+ Q(1 - \tilde{T}_\lambda)(D_t + A^-)T_\lambda(D_t + A^+)u^+ + RT_\lambda(D_t + A^+)u^+ \\ &+ (1 - T_\lambda)(D_t + A^+)u^+.\end{aligned}$$

Therefore

$$\begin{aligned}\|(D_t + A^+)u^+\|_{L^2} \lesssim \lambda^{-1}\|(D_t + A^-)T_\lambda(D_t + A^+)u^+\|_{L^2} + \lambda^{-N}\|(D_t + A)T_\lambda(D_t + A^+)u^+\|_{L^2} \\ + \lambda^{-N}\|\nabla_{t,x}u\|_{L^2}.\end{aligned}$$

The main term is

$$\begin{aligned}(D_t + A^-)T_\lambda(D_t + A^+)u^+ = T_\lambda(D_t + A^-)(D_t + A^+)T'_\lambda u \\ + [D_t + A^-, T_\lambda]T'_\lambda(D_t + A^+)u + [D_t + A^-, T_\lambda][(D_t + A^+), T'_\lambda]u\end{aligned} \tag{2.24}$$

where $T'_\lambda = P_{>-\lambda/64}(D_t)\tilde{P}_\lambda(D_x)u$. The second and third terms are bounded by $\|\nabla_{t,x}u\|_{L^2}$ in view of the commutator estimate

$$[A, T_\lambda] : L^2 \to L^2, \quad A = A^\pm,$$

which can be verified by decomposing the input and output frequencies and exploiting orthogonality; we leave the details to the reader.

For the first term on the right side of (2.24), we use the proof of the the previous lemma and similar commutator estimates as above to obtain

$$\begin{aligned}&\|(D_t + A^-)(D_t + A^+)T'_\lambda u\|_{L^2} \\ &\leq \|\Box_{g<\sqrt{\lambda}}T'_\lambda u\|_{L^2} + \|[\Box_{g<\sqrt{\lambda}} - (D_t + A^-)(D_t + A^+)]T'_\lambda u\|_{L^2} \\ &\lesssim \|\Box_{g<\sqrt{\lambda}}u\|_{L^2} + \|[g^{0j}_{<\sqrt{\lambda}}, T'_\lambda]\partial_t\partial_j u\|_{L^2} + \|[g^{ab}_{<\sqrt{\lambda}}, T'_\lambda]\partial_a\partial_b u\|_{L^2} + \|\nabla_{t,x}u\|_{L^2} \\ &\lesssim \|\nabla_{t,x}u\|_{L^2} + \|\Box_{g<\sqrt{\lambda}}u\|_{L^2}.\end{aligned}$$

This last estimate also yields the desired bound for $\|\Box_{g<\sqrt{\lambda}}u^+\|_{L^2}$, and altogether we obtain

$$\|(D_t + A^+)u^+\|_{L^2} \lesssim \lambda^{-1}\|\nabla_{t,x}u\|_{L^2} + \lambda^{-1}\|\Box_{g<\sqrt{\lambda}}u\|_{L^2}.$$

□

If u is compactly supported in time, we may smoothly truncate the half-waves in time to obtain:

Corollary 2.20. *Suppose $u \in P_\lambda(D_x)u$ has Fourier transform supported in $[\lambda/2, 2\lambda]$ and is supported in $[-c, c] \times \mathbb{R}^n$. Then there exists a decomposition*

$$u = u^+ + u^-$$

where $u^\pm$ are supported in $[-2c, 2c]\times\mathbb{R}^n$, and satisfy the estimates of the previous proposition.

3. Null frames

3.1. **Null foliations.** In this section we construct the null "hyperplanes" along which wave packets propagate. Factor the principal symbol of $\Box_g$ as $(\tau + a^+)(\tau + a^-)$. For each direction $\theta \in S^1$ and sign $\pm$, we construct optical functions $\Phi_\theta^\pm$ as solutions to the eikonal equation

$$\partial_t \Phi_\theta^\pm + a^\pm(t, x, \partial_x \Phi_\theta^\pm) = 0, \quad \Phi_\theta^\pm(0, x) = \langle x, \theta \rangle.$$

By the standard theory of Hamilton-Jacobi equations, for small η these admit classical solutions on the spacetime slab $[-10, 10] \times \mathbb{R}^2$.

Recall that Φ is constructed via the Hamilton flow for $a^\pm$ defined by

$$\dot{x} = a_\xi^\pm(t, x, \xi), \quad \dot{\xi} = -a_x^\pm(t, x, \xi). \tag{3.1}$$

Solutions to this systems are called bicharacteristics. Write $t \mapsto (x_t^\pm, \xi_t^\pm)$ for the solution initialized at (x_0, ξ_0). The map $(x_0, \xi_0) \mapsto (x_t^\pm, \xi_t^\pm)$ is 1-homogeneous in the second variable. Moreover, a routine linearization argument reveals that

Lemma 3.1.

$$\frac{\partial(x_t^\pm, \xi_t^\pm)}{\partial(x_0, \xi_0)} = \begin{pmatrix} I + O(\eta) & B(t) \\ O(\eta) & I + O(\eta) \end{pmatrix},$$

where the matrix $B(t)$ has norm $O(t)$.

Proof. The linearized system is

$$\dot{y} = a_{\xi x} y + a_{\xi\xi} \zeta,$$
$$\dot{\zeta} = -a_{xx} y - a_{x\xi} \zeta.$$

The half-wave symbols a inherit the derivative bounds on the metric:

$$\|\partial_{t,x} \partial_\xi^k a(t, x, \xi)\|_{L^\infty_{t,x}} + \|\partial_{t,x}^2 \partial^k a_\xi(t, x, \xi)\|_{L^2_t L^\infty_x} \lesssim \eta \quad \text{for bounded } \xi,$$

hence Gronwall yields the preliminary estimate $|(y(t), \zeta(t))| \lesssim |(y(0), \zeta(0))|$.

Consider initial data $y(0) = I$, $\zeta(0) = 0$. Then

$$|\zeta(t)| \lesssim \|\partial_x^2 a\|_{L^2 L^\infty} \sqrt{t} + \int_0^t |\zeta(s)|\, ds,$$

so $|\zeta(t)| \lesssim \eta\sqrt{t}$. Substituting this into the equation for y, we obtain

$$|y(t) - I| \lesssim \eta.$$

Now consider initial data $y(0) = 0, \eta(0) = I$. Then $|y(t)| \lesssim t$, and

$$|\zeta(t) - I| \lesssim \int_0^t |a_{xx}| s\, ds + \int_0^t |a_{x\xi}(s)\zeta(s)|\, d\xi \lesssim \eta.$$

□

Hence we may parametrize the graph of the $\pm$ flow map at time t by $(x_t^\pm, \xi_0) \mapsto (x_0, \xi_t^\pm)$, via the diffeomorphisms

$$(x_t^\pm, \xi_0) \mapsto (x_0, \xi_0) \mapsto (x_t^\pm, \xi_t^\pm) \mapsto (x_0, \xi_t^\pm).$$

A short computation then yields

$$\frac{\partial(x_0,\xi_t^\pm)}{\partial(x_t^\pm,\xi_0)} = \frac{\partial(x_0,\xi_t^\pm)}{\partial(x_t,\xi_t^\pm)} \cdot \frac{\partial(x_t^\pm,\xi_t^\pm)}{\partial(x_0,\xi_0)} \cdot \frac{\partial(x_0,\xi_0)}{\partial(x_t^\pm,\xi_0)} = \begin{pmatrix} I+O(\eta) & B(t) \\ O(\eta) & I+O(\eta) \end{pmatrix}. \tag{3.2}$$

We define $\xi_\theta^\pm(t,x)$ by the relation

$$\xi_{\xi_0}^\pm(t,x_t^\pm) := \xi_t^\pm(x_0,\xi_0), \tag{3.3}$$

and recall that the method of characteristics construction gives

$$\partial_x \Phi_\theta^\pm(t,x) = \xi_\theta^\pm(t,x). \tag{3.4}$$

As $g^{\alpha\beta}\partial_\alpha\Phi_\theta^\pm\,\partial_\beta\Phi_\theta^\pm = 0$, we obtain a foliation $\Lambda_\theta^\pm$ of $[-10,10]\times\mathbb{R}^2$ for each θ by the null "hyperplanes"

$$\Lambda_{\theta,h}^\pm := \{\Phi_\theta^\pm = h\}. \tag{3.5}$$

The regularity of these null surfaces is easy to compute:

Lemma 3.2. *The functions $\Phi_\theta^\pm$ have regularity $\partial_{t,x}^2\Phi_\theta^\pm = O(\eta)$.*

Proof. To simplify notation, we fix θ and focus on the $+$ case for the remainder of this section. We redenote $a = a^+$, $\Phi = \Phi_\theta^+$. In our new notation, Φ solves

$$\partial_t\Phi + a(t,x,\partial_x\Phi) = 0, \quad \Phi(0,x) = \langle x,\theta\rangle,$$

Differentiating the identity (3.4) gives $\partial_x^2\Phi = \partial_x\xi_\theta = O(\eta)$. The estimates involving time derivatives now follow by differentiating the equation:

$$\partial_t\partial_x\Phi + a_x + a_\xi\partial_x^2\Phi = 0,\ \partial_t\partial_x\Phi(0,x) = 0 \Rightarrow \partial_t\partial_x\Phi = O(\eta),$$
$$\partial_t^2\Phi + a_t + a_\xi\partial_t\partial_x\Phi = 0,\ \partial_t^2\Phi(0,x) = 0 \Rightarrow \partial_t^2\Phi = O(\eta).$$

□

Lemma 3.3 (Separation between null planes)**.** $\operatorname{dist}(\Lambda_{h_1,\theta},\Lambda_{h_2,\theta}) \sim |h_1-h_2|$. *More precisely, there exists constants $c_1,c_2>0$ such that for each $(t,x)\in\Lambda_{h_1}$, we have*

$$c_1|h_1-h_2| \le d((t,x),\Lambda_{h_2}) \le c_2|h_1-h_2|,$$

where d denotes Euclidean distance measured in the time slice $\{t\}\times\mathbb{R}^2$.

Proof. Without loss of generality assume $h_2 > h_1$. By the bounds on $\partial^2\Phi$, we have that $|\partial_x\Phi| = 1+O(\eta)$. The Euclidean gradient flow $\dot\gamma = \frac{\nabla_x\Phi}{|\nabla_x\Phi|}$ satisfies

$$\Phi(t,\gamma(s)) - \Phi(t,\gamma(0)) = \int_0^s |\nabla_x\Phi(t,\gamma(\tau))|\,d\tau \in [C_1 s, C_2 s]$$

for absolute constants C_1, C_2. Thus the Euclidean distance of Λ_{h_2} is at most $|h_1-h_2|/C_1$. If η is any other unit speed curve with $\eta(0) = (t,x)$, then

$$|\Phi(t,\eta(s)) - \Phi(t,\eta(0))| \le \int_0^s |\nabla_x\Phi(t,\theta(\tau))|\,d\tau \le C_2 s,$$

so the distance is at least $|h_1-h_2|/C_2$. □

The next two lemmas compare the bicharacteristics and null foliations for mollified and unmollified metrics.

Lemma 3.4. [44, Prop. 4.3] *If $\mathbb{R}^n \times S^{n-1} \ni (x,\xi) \mapsto (x_t,\xi_t),(x_{t,\lambda},\xi_{t,\lambda})$ are the + bicharacteristics for the metrics g and $g_{<\sqrt{\lambda}}$ with the same initial data, then*

$$|x_{t,\lambda} - x_t| \lesssim \lambda^{-1/2}, \quad |\xi_{t,\lambda} - \xi_t| \lesssim \lambda^{-1/2}.$$

Lemma 3.5. *(Foliations for frequency truncated metrics) Let Λ_h^λ be the foliation defined as before but replacing g with $g_{<\lambda^{\frac{1}{2}}}$. Then $\operatorname{dist}(\Lambda_{h,\theta}, \Lambda^\lambda_{h,\theta}) \lesssim \lambda^{-1}$.*

Proof. The optical function Φ^λ for Λ^λ satisfies the corresponding eikonal equation

$$\partial_t \Phi^\lambda + a_{<\lambda^{\frac{1}{2}}}(t,x,\partial_x\Phi^\lambda) = 0, \quad \Phi^\lambda(0,x) = \langle x,\theta\rangle,$$

where $a_{<\lambda^{\frac{1}{2}}} = a^+_{<\lambda^{\frac{1}{2}}}$ is the + half-wave symbol for the mollified metric (but is not itself exactly localized, due to the square root). The difference $\psi := \Phi - \Phi^\lambda$ solves the transport equation

$$\partial_t\psi + \langle v(t,x), \partial_x\psi\rangle = a_{<\lambda^{1/2}}(t,x,\partial_x\Phi^\lambda) - a(t,x,\partial_x\Phi^\lambda) = O(\lambda^{-1}), \quad \psi(0,x) = 0,$$
$$v = \int_0^1 a_\xi(t,x,\partial_x\Phi^\lambda + s\partial_x(\Phi - \Phi^\lambda))\,ds,$$

which may be integrated along characteristics to yield $|\psi(t,x)| \lesssim |t\lambda^{-1}|$. □

Certain estimates will be expressed in spacetime coordinates adapted to the foliation Λ_θ. The following construction is modeled on [30].

Let $\Phi = \Phi_\theta$ be the optical function for Λ_θ. We rotate the coordinates in x by setting

$$x_\theta = \langle x,\theta\rangle, \quad x'_\theta = \langle x,\theta^\perp\rangle,$$

where $\theta^\perp$ is clockwise rotation of θ by angle $\pi/2$. Then in the coordinates (t,x'_θ,x_θ), we have $\partial_{x_\theta}\Phi(0,\cdot) = 1$ and $\partial^2\Phi = O(\eta)$. Hence $\partial_{x_\theta}\Phi = 1 + O(\eta)$ for all $(t,x) \in [-10,10]\times\mathbb{R}^2$.

Provided η is sufficiently small, the global implicit function theorem lets us write

$$\Lambda_{h,\theta} = \{(t,x'_\theta,\psi_\theta(t,x'_\theta,h)\}$$

for some C^2 function ψ_θ. Then (t,x'_θ,h) also define coordinates on $[-10,10]\times\mathbb{R}^2$ via

$$(t,x'_\theta,h) \mapsto (t,x'_\theta,\psi_\theta(x'_\theta,h)) \in \Lambda_{h,\theta}.$$

Straightforward computations show that

$$\begin{aligned} \frac{\partial(x'_\theta,x_\theta)}{\partial(x'_\theta,h)} &= I + O(\eta), \\ \frac{\partial^2(x'_\theta,x_\theta)}{\partial(x'_\theta,h)^2} &= O(\eta). \end{aligned} \tag{3.6}$$

The variable h parametrizes the leaves of the foliation and is constant along geodesics.

3.2. **The null frame.** The future-pointing geodesic generator of Λ_θ is $L = -\nabla\Phi$. We complete this to a null frame by the following standard construction. Let $E = \langle e(t,x),\partial_x\rangle$ be a vector field tangent to the fixed-time slices of Λ_θ, defined concretely in terms of the rotated coordinates (x'_θ,x_θ) as

$$E = \frac{\tilde{E}}{\langle\tilde{E},\tilde{E}\rangle^{1/2}}, \quad \tilde{E}(t,x'_\theta,x_\theta) = \partial_{x'_\theta} + (\partial_{x'_\theta}\psi)\partial_{x_\theta}.$$

Then $\langle L, E\rangle = -d_{t,x}\Phi(E) = 0$; in fact one also has

$$0 = d_x\Phi(E) = \langle E, \xi_\theta(t,x)\rangle = \langle e(t,x), \xi_\theta(t,x)\rangle, \tag{3.7}$$

where $\xi_\theta(t,x)$ is the Fourier variable defined by (3.3). Finally, let $\underline{L}$ be a null vector field transversal to Λ_θ and satisfying $\langle \underline{L}, E\rangle = 0, \langle L, \underline{L}\rangle = -1$. The vector fields $\{L, \underline{L}, E\}$ form a null frame adapted to the foliation Λ_θ.

Lemma 3.6. *We have $L(t,x) = \sigma(t,x)[\partial_t + \langle a_\xi(t,x,\widehat{\xi_\theta}(t,x)), \partial_x\rangle]$ for some bounded function σ.*

Proof. The multiplicative factor reflects the difference between the null bicharacteristics of the half wave symbol and full wave symbol. Write $z = (t,x)$, $\zeta = (\tau,\xi)$, and let $p = g^{\alpha\beta}\zeta_\alpha\zeta_\beta = p^+p^- = (\tau + a^+)(\tau + a^-)$. Let $\gamma(s) = (z(s), \zeta(s))$ be a null bicharacteristic for p^+. Along this curve one has

$$\dot z = p^+_\zeta = (p^-)^{-1}p_\zeta, \quad \dot\zeta = -p^+_z = -(p^-)^{-1}p_z,$$

Therefore

$$\dot z^\alpha g_{\alpha\beta} = 2(p^-)^{-1}g^{\alpha\mu}\zeta_\mu g_{\alpha\beta} = 2(p^-)^{-1}\zeta_\beta = 2(p^-)^{-1}\partial_\beta\Phi(z(s)),$$

so

$$-2(p^-)^{-1}L = 2(p^-)^{-1}\nabla\Phi = \dot z = \partial_t + \langle a_\xi, \partial_x\rangle.$$

Finally, observe that $p^- = p^+ + p^- - p^+ = -2\sqrt{(g^{0j}\xi_j)^2 + g^{ab}\xi_a\xi_b}$ is bounded above and below along γ. □

4. Wave packets analysis

In this section we collect and generalize the salient features of Smith's wave packet parametrix [29]. More specific implementation details are recalled in the appendix.

4.1. Packets, tubes, and null surfaces.

We begin by clarifying our "wave packet" terminology. For $\omega \in S^{n-1}$, let $\gamma^\pm_\omega(t) = (x^\pm(t), \xi^\pm(t))$ be a bicharacteristic for half wave symbol $\tau + a^\pm$ with $\xi^\pm(0) = \omega$. Let $\omega(t) := \xi(t)/|\xi(t)|$ be the projection of ξ on the unit sphere.

Definition 4.1. A smooth function u at frequency λ is a normalized wave packet for the bicharacteristic $\gamma^\pm_\omega$ if:

- u is localized in phase space along $\gamma^\pm$: there exist constants C_N such that

$$|u(t,x)| \le C_N\lambda^{\frac34}(\lambda|\langle x - x^\pm(t), \omega^\pm(t)\rangle| + \lambda^{1/2}|(x - x^\pm(t)) \wedge \omega^\pm(t)|\rangle|)^{-N}, \tag{4.1}$$

  and similar estimates hold for $\big(\lambda\langle\omega^\pm(t), \partial_x\rangle\big)^a$ and $[\lambda^{1/2}(\omega^\pm(t)\wedge\partial_x)]^b$ applied to $u^\pm_\gamma$.
- Lu satisfies the same estimates with constant $C_N(t)$, where $|C_N(t)| \lesssim_N \|\partial^2 g(t)\|_{L^\infty}$, and L is the null generator for the null foliation $\Lambda^\pm_\omega$ defined in Section 3.1.

For each frequency $\lambda > 1$, let ω vary over a maximal collection $\Omega_{\lambda^{-1/2}}$ of unit vectors separated by at least $\lambda^{-\frac12}$. To each such ω we associate a lattice Ξ^ω_λ in the physical space $\mathbb{R}^n_x$ on the dual scale, i.e. spaced λ^{-1} in the ω direction and spaced $\lambda^{-\frac12}$ in directions in $\omega^\perp$. Let

$$\mathcal{T}_\lambda = \{(x,\omega) : \omega \in \Omega_{\lambda^{-1/2}},\ x \in \Xi^\omega_\lambda\}.$$

To each bicharacteristic $\gamma^\pm = (x^\pm(t), \xi^\pm(t))$ with $(x^\pm(0), \xi^\pm(0)) \in \mathcal{T}_\lambda$, we associate a spacetime "λ-tube"

$$T := \{(t,x) : \lambda|\langle x - x^\pm(t), \omega^\pm(t)\rangle| + \lambda^{1/2}|(x - x^\pm(t)) \wedge \omega^\pm(t)|\rangle| \leq 10\}, \tag{4.2}$$

on which packets for $\gamma^\pm$ concentrate. Each tube, say with initial data (x_0, ω), is a $\lambda^{-1/2}$ neighborhood of a ray, intersected with a λ^{-1} neighborhood of the null surface $\Lambda_{h,\omega}$ containing the ray.

It is suggestive to identify T with $\gamma^\pm$ and denote normalized packets by u_T. Let $\mathcal{T}_\lambda^\pm$ denote λ-tubes associated to the bicharacteristics $\gamma^\pm$ for the metric g, initialized in the lattice $\mathcal{T}_\lambda$.

By Lemmas 3.4 and 3.5, using the mollified metric $g_{<\sqrt{\lambda}}$ instead of g would yield an essentially equivalent family of tubes in the sense that each tube from one family intersects boundedly many tubes from the other family, which are in turn contained in the dilate of the first tube by a fixed factor.

The tubes $\mathcal{T}_{\lambda,\omega}^\pm$ associated to a given initial direction ω are finitely overlapping and admit a natural notion of distance which is convenient for expressing the decay of wave packets. Following Smith [29, Section 2], if $T_1, T_2 \in \mathcal{T}_{\lambda,\omega}^+$ are tubes with initial data $(x_j, \omega),\ j = 1, 2$, we define

$$d(T_1, T_2) := \lambda|\langle x_1 - x_2, \omega\rangle| + \lambda^{\frac{1}{2}}|(x_1 - x_2) \wedge \omega|.$$

Let $(t, x'_{1,\omega}(t), h_1)$ and $(t, x'_{2,\omega}(t), h_2)$ denote the corresponding rays in the foliation-adapted coordinates. It follows from (3.6) that each tube takes the form

$$T_j = \{|x'_\omega - x'_{j,\omega}(t)| \lesssim \lambda^{-\frac{1}{2}},\ |h - h_j| \lesssim \lambda^{-1}\},$$

and that $|x'_{1,\omega}(t) - x'_{2,\omega}(t)| + |h_1 - h_2| \sim |x'_{1,\omega}(0) - x'_{2,\omega}(0)| + |h_1 - h_2|$. Hence for $\lambda \gg 1$ we can also write

$$d(T_1, T_2) \sim \lambda|h_1 - h_2| + \lambda^{\frac{1}{2}}|x_{1,\omega'}(t) - x_{2,\omega'}(t)|.$$

If u_T is a packet with initial direction ω and T' is any other tube with initial direction ω, at time 0 the decay condition (4.1) translates to

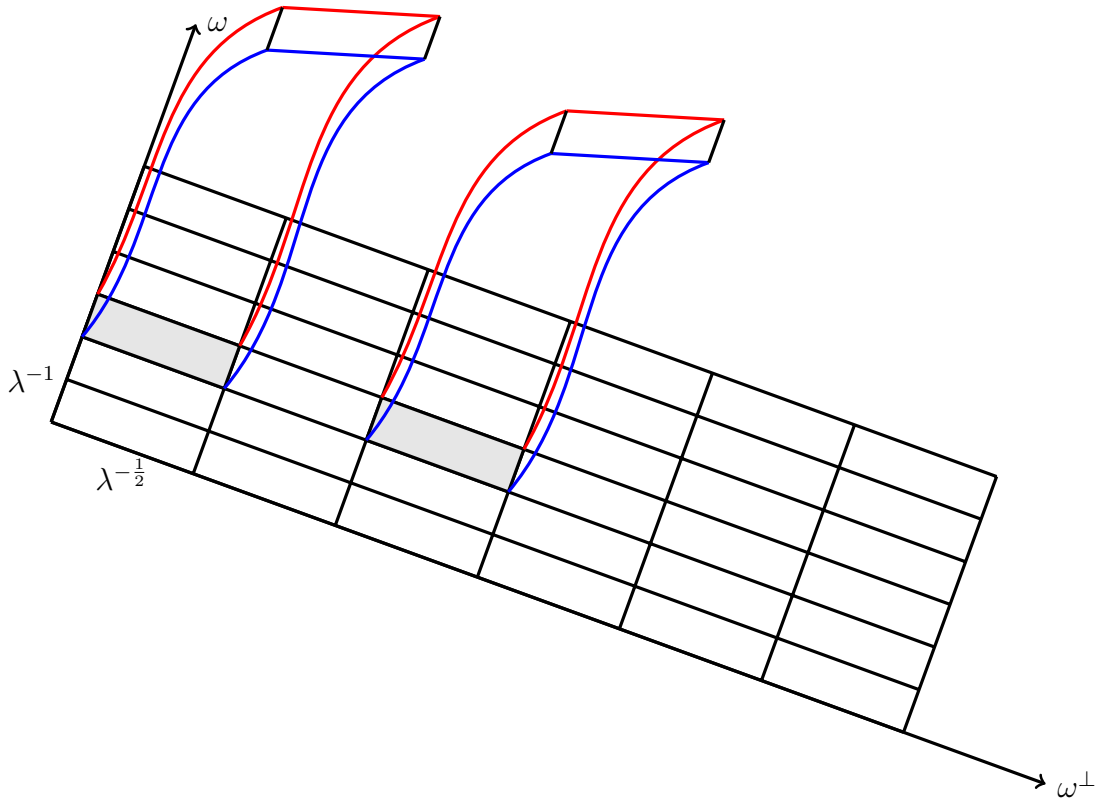
$$|u_T(0)|_{T'} \lesssim_N \lambda^{\frac{3}{4}}\langle d(T, T')\rangle^{-N}.$$

For $t \neq 0$, a similar bound holds but requires justification since the condition (4.1) is expressed in an orthogonal coordinate system whereas the null surfaces are curved.

Lemma 4.2. *Let $u = u_T$ be a frequency λ wave packet for the $+$ bicharacteristic (x_t^+, ξ_t^+) initialized at $(0, \omega)$. Let Λ_ω^λ be the $+$ null foliation with direction ω for $g_{<\sqrt{\lambda}}$, and denote by $\{L, \underline{L}, E\}$ the associated null frame. Then if $T' \in T_{\omega,\lambda}^+$ is any other tube, one has for all N*

$$|u_T|_{T'} \lesssim_N \lambda^{\frac{3}{4}}\langle d(T, T')\rangle^{-N}$$
$$|Eu_T|_{T'} \lesssim_N \lambda^{\frac{3}{4}+\frac{1}{2}}\langle d(T, T')\rangle^{-N}.$$

FIGURE 1. Tubes corresponding to the same direction ω are finitely overlapping.



More precisely, if $(t, x'_\omega(t), x_\omega(t))$ represents x_t^+ in the rotated coordinates and $\tilde{u}(t, x'_\omega, h) = u_T(t, x'_\omega, \psi(t, x'_\omega, h))$, $\widetilde{Eu}(t, x'_\omega, h) = (Eu_T)(t, x'_\omega, \psi(t, x'_\omega, h))$ represent u and Eu in the foliation adapted coordinates, then

$$|\tilde{u}| \lesssim_N \lambda^{\frac{3}{4}} \langle \lambda^{\frac{1}{2}} |x'_\omega - x'_\omega(t)| + |\lambda h| \rangle^{-N} \text{ for all } N.$$
$$|\widetilde{Eu}| \lesssim_N \lambda^{\frac{3}{4}+\frac{1}{2}} \langle \lambda^{\frac{1}{2}} |x'_\omega - x'_\omega(t)| + |\lambda h| \rangle^{-N} \text{ for all } N.$$

Proof. Modulo multiplicative factors of size $1 + O(\eta)$, we have

$$\omega(t) = -\partial_{x'_\omega} \psi(x'(t), 0) \partial_{x'_\omega} + \partial_{x_\omega},$$
$$E(t, x'_\omega, x_\omega(t, x'_\omega, h)) = \partial_{x'_\omega} + \partial_{x'_\omega} \psi(t, x'_\omega, h) \partial_{x_\omega}.$$

By a slight abuse of notation we also write $E = E(t, x'_\omega, h)$.

We express the wave packet decay condition (4.1) in terms of the coordinates (x'_ω, h). Let

$$y_\omega = \langle (x'_\omega - x'_\omega(t), \psi(t, x'_\omega) - \psi(t, x'_\omega(t), 0)), \omega(t) \rangle,$$
$$y'_\omega = \langle (x'_\omega - x'_\omega(t), \psi(t, x'_\omega) - \psi(t, x'_\omega(t), 0)), E(t) \rangle.$$

Then

$$y'_\omega = x'_\omega - x'_\omega(t) + [\psi(t, x'_\omega, h) - \psi(t, x'_\omega, 0) + \psi(t, x'_\omega, 0) - \psi(t, x'_\omega(t), 0)] \partial_x \psi(t, x'_\omega(t), 0),$$
$$y_\omega = \psi(t, x'_\omega, h) - \psi(t, x'_\omega, 0) + \psi(t, x'_\omega, 0) - \psi(t, x'_\omega(t), 0) - (x - x'_\omega(t)) \partial_{x'_\omega} \psi(t, x'_\omega(t), 0)$$

So

$$|y'_\omega| \geq |x'_\omega - x'_\omega(t)| - c(\eta t)^2|x'_\omega - x'_\omega(t)| - c\eta t h \geq \frac{1}{2}|x'_\omega - x'_\omega(t)|$$

whenever $\dfrac{C|x'_\omega - x'_\omega(t)|}{\eta t} \geq h$. Similarly,

$$|y_\omega| \geq (1 + O(\eta))h - c\eta t|x'_\omega - x'_\omega(t)| \geq \frac{1}{2}ch$$

if $|x'_\omega - x'_\omega(t)| \leq \frac{Ch}{\eta t}$. Thus

$$|\tilde{u}(t, x'_\omega, h)| \lesssim_N \begin{cases} \lambda^{\frac{3}{4}}\langle \lambda h\rangle^{-N}, & |x'_\omega - x'_\omega(t)| \leq \frac{\eta t h}{C}, \\ \lambda^{\frac{3}{4}}\langle \lambda^{\frac{1}{2}}|x'_\omega - x'_\omega(t)| + |\lambda h|^{-N}\rangle^{-N}, & \frac{\eta t h}{C} \leq |x'_\omega - x'_\omega(t)| \leq \frac{Ch}{\eta t}, \\ \lambda^{\frac{3}{4}}\langle \lambda^{\frac{1}{2}}|x'_\omega - x'_\omega(t)|\rangle^{-N}, & |x'_\omega - x'_\omega(t)| \geq \frac{Ch}{\eta t}, \end{cases}$$

but one may of course insert $\lambda^{\frac{1}{2}}|x'_\omega - x'_\omega(t)|$ and λh in the first and third regimes respectively.

To verify the bound on $\widetilde{Eu}$, write

$$E(t, x'_\omega, h)u = E(t, x'_\omega(t), 0)u + [E(t, x'_\omega, h) - E(t, x'_\omega(t), 0)]u.$$

For the first term we use the hypothesis that $\lambda^{-1/2}(\omega(t) \wedge \partial_x)u$ satisfies the packet bounds (4.1). The second term can be written as

$$\big(\partial_{x'_\omega}\psi(t, x'_\omega, h) - \partial_{x'_\omega}\psi(t, x'_\omega(t), 0)\big)\partial_{x_\omega}u,$$

which is $O(\lambda|x_\omega - x'_\omega(t)| + \lambda|h|)$ times a normalized packet. □

Corollary 4.3. *Let u be a frequency λ wave packet for the bicharacteristic (x_t^+, ξ_t^+) of $g_{<\lambda^{\frac{1}{2}}}$ initialized at $(0, \omega)$. Let Λ_ω be the corresponding null foliation for the* untruncated *metric g. Then the previous estimates hold with Λ_ω^λ replaced by the corresponding foliation Λ_ω for the untruncated metric g.*

Proof. The above lemmas imply that the foliations for g and $g_{<\sqrt{\lambda}}$ are interchangeable as far as the bound with respect to h is concerned. Similarly, as shown in [44, Prop. 4.3] the bicharacteristics for g and $g_{<\sqrt{\lambda}}$ differ by $O(\lambda^{-1/2})$. □

To express the decay of packets more compactly, we introduce the weight

$$m_T(t, x) := \big(1 + \lambda|\langle x - x^\pm(t), \omega^\pm(t)\rangle| + \lambda^{1/2}|(x - x^\pm(t)) \wedge \omega^\pm(t)|\big),$$

write

$$\|u_T\|_{WP_T^N} := \|\lambda^{-\frac{3}{4}} m_T^N u_T\|_{L^\infty}, \tag{4.3}$$

and write $\|\cdot\|_{WP}$ to denote a generic WP_T^N norm with the understanding that the constants in bounds depend on N.

4.2. **Superpositions of wave packets.** The following statement summarizes how the different wave packets fit together in Smith's parametrix. Note that in the rest of the paper we shall only use the properties below, and not the specifics of Smith's construction.

Parametrix property 4.1. *Let $I = [0, \delta]$. For large enough dyadic $\lambda \geq 1$, the following properties hold:*

(1) For any initial data $(u_1, u_2) \in L^2 \times H^{-1}$ localized at frequency $\simeq \lambda$, there exists a λ-wave packet superposition which depends linearly on (u_1, u_2):

$$u = u^+ + u^-, \qquad u^\pm = \sum_{T \in \mathcal{T}_\lambda^\pm} c_T u_T \tag{4.4}$$

such that $(u(0), \partial_t u(0)) = (u_1, u_2)$ and

$$\sum_{\pm, T \in \mathcal{T}_\lambda^\pm} |c_T|^2 \approx \|(u_1, u_2)\|^2_{L^2 \times H^{-1}} \tag{4.5}$$

For any such decomposition one has

$$\|\Box_{g < \sqrt{\lambda}} u(t)\|_{L^2_x} \lesssim \lambda \|\partial^2 g(t)\|_{L^\infty_x} \|c_T\|_{\ell^2_T} \qquad \forall t \in I. \tag{4.6}$$

(2) Let $D = \delta^{-1}$. For any solution of the homogeneous problem

$$\Box_{g_{<\sqrt{\lambda}}} v = 0, \qquad v[0] = (v_1, v_2). \tag{4.7}$$

there exists u as above satisfying

$$\|\tilde{P}_\lambda (v - u)\|_{X^{0, \frac{1}{2}}_{\lambda, D}[I]} \lesssim \delta \lambda^{-1} \|(v_1, v_2)\|_{H^1 \times L^2} \tag{4.8}$$

$$\|c_T\|_{\ell^2_T} \lesssim \lambda^{-1} \|(v_1, v_2)\|_{H^1 \times L^2} \tag{4.9}$$

(3) Let $T, T' \in \mathcal{T}^\pm_{\lambda, \omega}$ be two tubes associated to the same direction ω. For $\nu \geq \lambda$, let the vector fields $L, \underline{L}, E$ associated to $\pm g_{\sqrt{\nu}}$ and to ω form a null frame as in section 3.2. Then one has:

$$\|u_T\|_{L^\infty(T')} \lesssim \frac{\lambda^{\frac{3}{4}}}{\langle d(T, T') \rangle^N} \tag{4.10}$$

$$\|L u_T\|_{L^2 L^\infty(T')} \lesssim \frac{\lambda^{\frac{3}{4}}}{\langle d(T, T') \rangle^N} \tag{4.11}$$

$$\|E u_T\|_{L^\infty(T')} \lesssim \lambda^{\frac{1}{2}} \frac{\lambda^{\frac{3}{4}}}{\langle d(T, T') \rangle^N} \tag{4.12}$$

$$\|\underline{L} u_T\|_{L^\infty(T')} \lesssim \lambda \frac{\lambda^{\frac{3}{4}}}{\langle d(T, T') \rangle^N} \tag{4.13}$$

Moreover, the same inequalities hold with u_T replaced by $P_\lambda u_T$.

(4) For any $t \in I$ and sign $\pm$, for $(c_T)_T \in \ell^2_T$ one has

$$\| \sum_{T \in \mathcal{T}^\pm_\lambda} c_T u_T(t) \|^2_{L^2_x} \lesssim \sum_{T \in \mathcal{T}^\pm_\lambda} |c_T|^2, \quad \| \sum_{T \in \mathcal{T}^\pm_\lambda} c_T u'_T(t) \|^2_{L^2_x} \lesssim \lambda \sum_{T \in \mathcal{T}^\pm_\lambda} |c_T|^2 \tag{4.14}$$

We note that property (2) follows from property (1) together (1.30), (4.6) and Lemma 2.11. In Section A we discuss these properties in the context of Hart Smith's wave packet parametrix from [29].

Remark 4.4. The decay properties (4.10) through (4.13) reduce certain L^2 bilinear estimates to the characteristic energy estimates in section 5. Here is a typical computation. Suppose $v = \sum_{T \in \mathcal{T}^+_{\lambda,\omega}} a_T v_T$ is a superposition of frequency λ wave packets for a given initial direction ω. Then by Schur's test we deduce

$$\begin{aligned}\|uv\|_{L^2}^2 &\lesssim \sum_T \|uv\|_{L^2(T)}^2 = \sum_T \sum_{T_1,T_2} a_{T_1}\overline{a_{T_2}}\langle uv_{T_1}, uv_{T_2}\rangle_{L^2(T)} \\ &\lesssim \sum_T \sum_{T_1,T_2} \lambda^{\frac{3}{2}} \|u\|_{L^2(T)}^2 d(T,T_1)^{-N} d(T,T_2)^{-N} |a_{T_1} a_{T_2}| \\ &\lesssim \big(\sup_T \|u\|_{L^2(T)}^2\big)\ \lambda^{\frac{3}{2}} \sum_T |a_T|^2,\end{aligned}$$

and the sup term is essentially an estimate for u over the null surfaces associated to the direction ω.

4.3. **Preliminaries to the general decomposition.** The next goal is to obtain a more general wave packet decomposition similar to (4.4) for functions in $X^{s,\theta}$ which are close to being solutions of (4.7) in the sense of having low modulation. To allow for the extra flexibility of having inhomogeneities $\Box_{g_{<\sqrt{\lambda}}} v$, the resulting decomposition (Prop. 4.7 and Cor. 4.8) will have coefficients $c_T(\cdot)$ that depend on time, which arise from Duhamel's formula

$$v = v_0 + \int_I 1_{t \geq s} v_s \, \mathrm{d}s$$

We first express the functions v_s in the following way:

Lemma 4.5. *For $s \in I = [0,\delta]$, let v_s be the solution of the equation*

$$\begin{cases} \Box_{g_{<\sqrt{\lambda}}} v_s = 0 & \text{on } I \times \mathbb{R}^2 \\ v_s[s] = (f_s, g_s). \end{cases}$$

where f_s, g_s are assumed to be localized at frequency $\simeq \lambda$. Then, there exists a wave packet superposition (initialized at $t = 0$)

$$u_s = \sum_{\pm, T \in \mathcal{T}^{\pm}_{\lambda}} c_{T,s} u_T$$

and a function w_s such that

$$\tilde{P}_\lambda v_s = \tilde{P}_\lambda (u_s + w_s)$$

where

$$w_s[s] = (0,0) \tag{4.15}$$

$$\|\tilde{\tilde{P}}_\lambda w_s\|_{X^{0,\frac{1}{2}}_{\lambda,D}[I]} \lesssim \delta \|(f_s,g_s)\|_{L^2 \times H^{-1}} \tag{4.16}$$

$$\|c_{T,s}\|_{\ell^2_T} \lesssim \|(f_s,g_s)\|_{L^2 \times H^{-1}}. \tag{4.17}$$

Proof. Denote $M = \|(f_s, g_s)\|_{L^2\times H^{-1}}$ and for brevity we will denote $\|\cdot\|_{X^{0,\frac12}_{\lambda,D}[I]}$ by $\|\cdot\|_X$. Note that $\|v_s[0]\|_{H^1\times L^2} \lesssim \lambda M$. We apply the Parametrix property 4.1 with $(v_1, v_2) = v_s[0]$ and we obtain a representation $\tilde P_\lambda v_s = \tilde P_\lambda u^1 + \tilde P_\lambda w^1$ where u^1 is a wave packet superposition. For $s = 0$ this is sufficient. For $s \neq 0$, even though $w^1[s] \neq (0,0)$, we have $\|\tilde{\tilde P}_\lambda w^1\|_X \lesssim \delta M$.

Now we iterate this construction. For $i \geq 1$ we write

$$\tilde P_\lambda w^i = \tilde P_\lambda(\tilde{\tilde P}_\lambda w^i - v^{i+1}) + \tilde P_\lambda v^{i+1}$$

where v^{i+1} solves the homogeneous equation

$$\Box_{g<\sqrt\lambda} v^{i+1} = 0, \qquad v^{i+1}[s] = \tilde{\tilde P}_\lambda w^i[s].$$

Assuming $\|\tilde{\tilde P}_\lambda w^i\|_X \lesssim \delta^i M$ we have $\|v^{i+1}\|_{L^\infty(H^1\times L^2)} \lesssim \lambda\delta^i M$. As before we use the Parametrix property 4.1 to write $\tilde P_\lambda v^{i+1} = \tilde P_\lambda(u^{i+1} + w^{i+1})$ with $\|\tilde{\tilde P}_\lambda w^{i+1}\|_X \lesssim \delta^{i+1} M$.

From the above we obtain $\tilde P_\lambda v_s = \tilde P_\lambda(u_s + w_s)$ by defining

$$u_s = \sum_{i\geq 1} u^i, \qquad w_s = \sum_{i\geq 1} \tilde{\tilde P}_\lambda w^i - v^{i+1}$$

Note that $w_s[s] = (0,0)$. Both series converge geometrically due to the powers of δ. □

Corollary 4.6. *Let $v \in X^{0,\frac12}_{\lambda,D}[I]$ for $I = [0,\delta]$ and let w be defined by*

$$w = w_0 + \int_I 1_{t\geq s} w_s \,\mathrm{d}s$$

where w_0, w_s are obtained from Lemma 4.5 applied with $(f_0, g_0) = \tilde P_\lambda v[0]$, respectively $(f_s, g_s) = (0, \Box_{g<\sqrt\lambda}\tilde P_\lambda v(s))$. Then,

$$\|\tilde{\tilde P}_\lambda w\|_{X^{0,\frac12}_{\lambda,D}[I]} \lesssim \delta \|\tilde{\tilde P}_\lambda v\|_{X^{0,\frac12}_{\lambda,D}[I]}$$

Proof. To ease notation we denote $\tilde{\tilde P}_\lambda$ by P. The inequality for w_0 follows immediately due to (2.3). Denoting by $\tilde w$ the integral term, we have $\tilde w[0] = (0,0)$ and, since $w_t[t] = (0,0)$, we have

$$\Box_{g<\sqrt\lambda}\tilde w(t) = \int_I 1_{t\geq s}\Box_{g<\sqrt\lambda} w_s \,\mathrm{d}s. \tag{4.18}$$

Note that by (4.16), (2.3) and Hölder's inequality we have

$$\|P\tilde w(t)\|_{L^2\times H^{-1}} \lesssim \delta \|Pv\|_{X^{0,\frac12}_{\lambda,D}[I]} \tag{4.19}$$

We write

$$\Box_{g<\sqrt\lambda} P\tilde w = P\Box_{g<\sqrt\lambda}\tilde w + [\Box_{g<\sqrt\lambda}, P]\tilde w$$

and apply (2.14):

$$\|\tilde{\tilde P}_\lambda w\|_{X^{0,\frac12}_{\lambda,D}[I]} \lesssim \lambda^{-1} D^{-\frac12}(\|P\Box_{g<\sqrt\lambda}\tilde w\|_{L^2[I]} + \|[\Box_{g<\sqrt\lambda}, P]\tilde w)\|_{L^2[I]})$$

The second term is estimated by (1.30) and (4.19). For the first term on the RHS we apply Minkowski's inequality to (4.18)

$$\lambda^{-1}D^{-\frac12}\|P\int_I 1_{t\geq s}\Box_{g_{<\sqrt\lambda}}w_s\,\mathrm{d}s\|_{L^2[I]}\lesssim\lambda^{-1}D^{-\frac12}\int_I\|P\Box_{g_{<\sqrt\lambda}}w_s\|_{L^2[I]}\,\mathrm{d}s$$
$$\lesssim\lambda^{-1}D^{-\frac12}\int_I\|\Box_{g_{<\sqrt\lambda}}Pw_s\|_{L^2[I]}+\|[P,\Box_{g_{<\sqrt\lambda}}]w_s\|_{L^2[I]}\,\mathrm{d}s\lesssim\int_I\|Pw_s\|_{X^{0,\frac12}_{\lambda,D}[I]}\,\mathrm{d}s$$
$$\lesssim\delta\int_I\lambda^{-1}\|\Box_{g_{<\sqrt\lambda}}\tilde P_\lambda v(s)\|_{L^2_x}\,\mathrm{d}s\lesssim\delta\|\tilde{\tilde P}_\lambda v\|_{X^{0,\frac12}_{\lambda,D}[I]}$$

where we have used (1.30), (4.16) and Hölder's inequality in s. □

4.4. **A wave packet characterization of the $X^{s,\theta}$ spaces.** With the previous preliminaries we are now ready to state our general decomposition (see also [42, Sec. 4]).

Proposition 4.7. *Let $I=[0,\delta]$, $\mu\geq D=\delta^{-1}$ and let $v\in X^{0,\frac12}_{\mu,D}[I]$ be localized at frequency $\simeq\mu$. We denote $M:=\|\tilde{\tilde P}_\mu v\|_{X^{0,\frac12}_{\mu,D}[I]}\lesssim\|v\|_{X^{0,\frac12}_{\mu,D}[I]}$.*

(1) Then, there exists a wave packet decomposition

$$P_\mu v(t)=P_\mu\sum_{\pm,T\in\mathcal{T}^\pm_\mu}a_T(t)u_T(t)\tag{4.20}$$

where the time-dependent coefficients satisfy, for all $t\in I$,

$$\|a_T\|_{\ell^2_T L^\infty_t}\lesssim M\tag{4.21}$$
$$\|a_T'(t)\|_{\ell^2_T}\lesssim\mu^{-1}\|\Box_{g_{<\sqrt\mu}}\tilde P_\mu v(t)\|_{L^2_x}$$
$$\|a_T'(t)\|_{L^2_t\ell^2_T}\lesssim D^{\frac12}M\tag{4.22}$$
$$\sum_{\pm,T\in\mathcal{T}^\pm_\mu}a_T'(t)u_T(t)=0.\tag{4.23}$$

(2) Conversely, if (4.21), (4.22), (4.23) hold for some M and

$$w=\sum_{\pm,T\in\mathcal{T}^\pm_\mu}a_T(t)u_T(t),$$

then $\|w\|_{X^{0,\frac12}_{\mu,D}[I]}\lesssim M$.

Proof. For brevity we will denote $\|\cdot\|_{X^{0,\frac12}_{\mu,D}[I]}$ by $\|\cdot\|_X$ and $P_\mu,\tilde P_\mu,\tilde{\tilde P}_\mu$ by $P,\tilde P,\tilde{\tilde P}$.

(1) Note that it suffices to prove that there exists a decomposition

$$Pv(t)=Pw(t)+P\sum_{\pm,T\in\mathcal{T}^\pm_\mu}a_T(t)u_T(t)$$

for some function w with bound $\|\tilde{\tilde{P}}w\|_X \lesssim \delta\|\tilde{\tilde{P}}v\|_X$ and $(a_T)_T$ satisfying the requirements above, since then the proposition follows by an iterative argument.

By Duhamel's formula we represent

$$\tilde{P}v = v_0 + \int_I 1_{t\geq s} v_s \,\mathrm{d}s$$

where v_0 and each v_s are solutions to the problems

$$\begin{cases} \Box_{g_{<\sqrt{\mu}}} v_0 = 0 \\ v_0[0] = \tilde{P}v[0] \end{cases} \qquad \begin{cases} \Box_{g_{<\sqrt{\mu}}} v_s = 0 \\ v_s[s] = (0, \Box_{g_{<\sqrt{\mu}}}\tilde{P}v(s)). \end{cases}$$

We apply Lemma 4.5 and obtain:

$$Pv_0 = Pu_0 + Pw_0, \qquad Pv_s = Pu_s + Pw_s$$
$$u_0 = \sum_{\pm, T\in\mathcal{T}_\mu^\pm} c_T u_T, \qquad u_s = \sum_{\pm, T\in\mathcal{T}_\mu^\pm} c_{T,s} u_T$$

with the bounds

$$\|c_T\|_{\ell_T^2} \lesssim \|\tilde{P}v[0]\|_{L^2\times H^{-1}} \lesssim M, \qquad \|c_{T,s}\|_{\ell_T^2} \lesssim \mu^{-1}\|\Box_{g_{<\sqrt{\mu}}}\tilde{P}v(s)\|_{L_x^2} \tag{4.24}$$

and $u_s(s) = v_s(s) = 0$. We write $Pv = P\tilde{P}v = Pu + Pw$ where

$$u = u_0 + \int_I 1_{t\geq s} u_s \,\mathrm{d}s, \qquad w = w_0 + \int_I 1_{t\geq s} w_s \,\mathrm{d}s$$

By Corollary 4.6 we have $\|\tilde{\tilde{P}}w\|_X \lesssim \delta\|\tilde{\tilde{P}}v\|_X$.

We obtain the representation

$$u(t) = \sum_{\pm, T\in\mathcal{T}_\mu^\pm} a_T(t) u_T(t)$$

where, for any sign $\pm$ and any $T \in \mathcal{T}_\mu^\pm$:

$$a_T(t) = c_T + \int_I 1_{t\geq s} c_{T,s} \,\mathrm{d}s \qquad |a_T|_{L_t^\infty} \lesssim |c_T| + \int_I |c_{T,s}| \,\mathrm{d}s$$

and $a_T'(t) = c_{T,t}$ for all $t \in I$. We have

$$\|a_T\|_{\ell_T^2 L_t^\infty} \lesssim \|c_T\|_{\ell_T^2} + \| \int_I |c_{T,s}| \,\mathrm{d}s\|_{\ell_T^2} \lesssim M + \int_I \|c_{T,s}\|_{\ell_T^2} \,\mathrm{d}s \lesssim$$
$$\lesssim M + \int_I \mu^{-1}\|\Box_{g_{<\sqrt{\mu}}}\tilde{P}v(s)\|_{L_x^2} \,\mathrm{d}s \lesssim M + \delta^{\frac{1}{2}}\mu^{-1}\|\Box_{g_{<\sqrt{\mu}}}\tilde{P}v\|_{L_{t,x}^2} \lesssim M.$$

This verifies (4.21). The next condition holds due to (4.24):

$$\|a_T'(t)\|_{\ell_T^2} = \|c_{T,t}\|_{\ell_T^2} \lesssim \mu^{-1}\|\Box_{g_{<\sqrt{\mu}}}\tilde{P}v(s)\|_{L_x^2}$$

which also gives (4.22) by integration.

Since $u_s(s) = v_s(s) = 0$ we have $\sum a_T'(s)u_T(s) = 0$ for any $s \in I$, obtaining (4.23), which completes the proof of the first part.

(2) By Hölder's inequality in time and (4.14) we have

$$D^{\frac{1}{2}}\|w\|_{L^2(I\times\mathbb{R}^2)} \lesssim \sum_{\pm} \| \sum_{T\in\mathcal{T}_\mu^\pm} a_T(t)u_T(t)\|_{L_t^\infty L_x^2} \lesssim \sum_{\pm} \|a_T\|_{L_t^\infty \ell_T^2} \lesssim \sum_{\pm} \|a_T\|_{\ell_T^2 L_t^\infty} \lesssim M$$

Now we consider the term $\Box_{g_{<\sqrt{\mu}}} w$ which we write as

$$\Box_{g_{<\sqrt{\mu}}} w = \sum_{\pm, T\in\mathcal{T}_\mu^\pm} \Big[a_T(t)\Box_{g_{<\sqrt{\mu}}} u_T(t) + a_T'(t) g\partial_{t,x} u_T(t) + \partial_t(a_T'(t)u_T(t))\Big]$$

For the last term we use (4.23). For the first term we use (4.6):

$$D^{-\frac{1}{2}}\| \sum a_T(t)\Box_{g_{<\sqrt{\mu}}} u_T(t)\|_{L^2(I\times\mathbb{R}^2)} \lesssim \lambda D^{-1}\|\partial^2 g\|_{L_t^2 L_x^\infty} \sum_{\pm} \|a_T\|_{L_t^\infty \ell_T^2} \lesssim \lambda M,$$

while for the second term we use (4.14) and (4.22):

$$\| \sum a_T'(t)\partial_{t,x} u_T(t)\|_{L^2(I\times\mathbb{R}^2)} \lesssim \mu \|a_T'(t)\|_{L_t^2\ell_T^2} \lesssim \mu D^{\frac{1}{2}} M,$$

which completes the proof. □

The previous proposition provides the main part of the wave packet decomposition. However, it does not provide control of the second derivatives (in time) of the coefficients. To remedy this we have the following corollary.

Corollary 4.8. *Under the assumptions and notations of Prop. 4.7:*
(1) There exists a decomposition

$$P_\mu v = v^+ + v^- + v_R$$

where

$$v^\pm = P_\mu \sum_{T\in\mathcal{T}_\mu^\pm} c_T(t)u_T(t) \tag{4.25}$$

such that

$$\|v^\pm\|_{X^{0,\frac{1}{2}}_{\mu,D}[I]} \lesssim M, \qquad \|v_R\|_{X^{0,\frac{1}{2}}_{\mu,D}[I]} \lesssim M, \tag{4.26}$$

$$\|v_R\|_{L^2[I]} \lesssim \mu^{-1} D^{\frac{1}{2}} M \tag{4.27}$$

and

$$\|c_T\|_{\ell_T^2 L_t^\infty} \lesssim M \tag{4.28}$$

$$\|c_T'(t)\|_{L_t^2\ell_T^2} \lesssim D^{\frac{1}{2}} M \tag{4.29}$$

$$\|c_T''(t)\|_{L_t^2\ell_T^2} \lesssim \mu D^{\frac{1}{2}} M. \tag{4.30}$$

(2) Let $T, T' \in \mathcal{T}_{\mu,\omega}^\pm$ be two tubes associated to the same direction ω. Then

$$\|\underline{L} P_\mu\big(c_T u_T\big)\|_{L^\infty(T')} \lesssim \ \mu\ \frac{\mu^{\frac{3}{4}}}{\langle d(T,T')\rangle^N}\, |c_T|_{L_t^\infty} \tag{4.31}$$

Corollary 4.9. *Moreover, for any function u_λ localized at frequency $\simeq \lambda \gg \mu$ we have:*

$$\|u_\lambda \cdot v_R\|_{X^{0,\frac14}_{\lambda,\mu}[I]} \lesssim \mu^{\frac34} \|u_\lambda\|_{X^{0,\frac12}_{\lambda,D}[I]} M \tag{4.32}$$

while for $\lambda \simeq \mu \gtrsim \eta$ we have

$$\|P_\eta(u_\lambda \cdot v_R)\|_{X^{0,\frac14}_{\eta,\eta}[I]} \lesssim \frac{\lambda}{\eta^{\frac14}} \|u_\lambda\|_{X^{0,\frac12}_{\lambda,D}[I]} M \tag{4.33}$$

Proof of Corollary 4.8. **(1)** The previous proposition provides a decomposition into $+$ and $-$ components and only condition (4.30) is missing on the coefficients. To gain it, we regularize the coefficients $a_T(t)$ in time on the μ^{-1} scale at the expense of introducing the remainder v_R which obeys the favorable L^2 estimate (4.27). We write and define

$$a_T(t) = a_T^{<\mu}(t) + a_T^{>\mu}(t), \qquad c_T(t) := a_T^{<\mu}(t)$$

The conditions (4.28), (4.29) are maintained from (4.21), (4.22), while (4.30) follows from (4.29).

The $v^\pm$ are defined by (4.25). We prove that $\|v^\pm\|_{X^{0,\frac12}_{\mu,D}[I]} \lesssim M$ and as a consequence we also obtain $\|v_R\|_{X^{0,\frac12}_{\mu,D}[I]} \lesssim M$. We have to place $v^\pm$ and $\Box_{g<\sqrt{\mu}} v^\pm$ in L^2. The fact that $D^{\frac12}\|v^\pm\|_{L^2[I]} \lesssim M$ and

$$D^{-\frac12}\mu^{-1} \| \sum c_T(t) \Box_{g<\sqrt{\mu}} u_T(t) \|_{L^2(I\times\mathbb{R}^2)} \lesssim M$$

$$D^{-\frac12}\mu^{-1} \| \sum c_T'(t) \partial u_T(t) \|_{L^2(I\times\mathbb{R}^2)} \lesssim M$$

follow like in the proof of part (2) of Prop. 4.7. What remains is

$$D^{-\frac12}\mu^{-1} \| \sum c_T''(t) u_T(t) \|_{L^2(I\times\mathbb{R}^2)} \lesssim D^{-\frac12}\mu^{-1} \|c_T''(t)\|_{L^2_t\ell^2_T} \lesssim M$$

For v_R we have

$$v_R = P_\mu \sum_{\pm, T\in\mathcal{T}_\mu^\pm} a_T^{>\mu}(t) u_T(t)$$

Since

$$\mu \|a_T^{>\mu}(t)\|_{L^2_t\ell^2_T} \lesssim \|a_T^{>\mu,\prime}(t)\|_{L^2_t\ell^2_T} \lesssim D^{\frac12} M$$

we obtain (4.27).

(2) To prove (4.31), for any $t \in I$ we write

$$\underline{L} P_\mu \big(c_T(t) u_T(t) \big) = c_T(t) \underline{L} P_\mu u_T(t) + c_T'(t) P_\mu u_T(t)$$

For the first term we use (4.13), while for the second we use (4.10) together with

$$|c_T'|_{L^\infty_t} \lesssim \mu\, |c_T|_{L^\infty_t} \tag{4.34}$$

which holds due to the time regularization done in Step (1). □

Proof of Corollary 4.9. Note that by Bernstein's inequality and (4.27), (4.26), (2.3) we have

$$\|v_R\|_{L^2L^\infty} \lesssim D^{\frac12} M, \qquad \|\Box_{g<\sqrt{\mu}} v_R\|_{L^2L^\infty} \lesssim \mu^2 D^{\frac12} M, \qquad \|v_R\|_{L^\infty_{t,x}} \lesssim \mu M.$$

For the L^2 part of (4.32) we have

$$\mu^{\frac{1}{4}}\|u_\lambda \cdot v_R\|_{L^2} \lesssim \mu^{\frac{1}{4}}\|u_\lambda\|_{L^\infty L^2}\|v_R\|_{L^2L^\infty} \lesssim \mu^{\frac{3}{4}}(D/\mu)^{\frac{1}{2}}\|u_\lambda\|_{X^{0,\frac{1}{2}}_{\mu,D}[I]}M$$

and recall that $D \leq \mu$. For $\Box_{g_{<\sqrt{\lambda}}}(u_\lambda \cdot v_R)$ we consider

$$\begin{aligned}
&\|\Box_{g_{<\sqrt{\lambda}}}u_\lambda \cdot v_R\|_{L^2} \lesssim \|\Box_{g_{<\sqrt{\lambda}}}u_\lambda\|_{L^2}\|v_R\|_{L^\infty} \lesssim \lambda D^{\frac{1}{2}}\mu\|u_\lambda\|_{X^{0,\frac{1}{2}}_{\mu,D}[I]}M\\
&\|u_\lambda \cdot \Box_{g_{<\sqrt{\mu}}}v_R\|_{L^2} \lesssim \|u_\lambda\|_{L^\infty L^2}\|\Box_{g_{<\sqrt{\mu}}}v_R\|_{L^2L^\infty} \lesssim \mu^2 D^{\frac{1}{2}}\|u_\lambda\|_{X^{0,\frac{1}{2}}_{\mu,D}[I]}M\\
&\|u_\lambda \cdot (\Box_{g_{<\sqrt{\lambda}}} - \Box_{g_{<\sqrt{\mu}}})v_R\|_{L^2} \lesssim \|u_\lambda\|_{L^2}\|\mu^{-1}\partial^2 v_R\|_{L^\infty} \lesssim \mu^2 D^{\frac{1}{2}}\|u_\lambda\|_{X^{0,\frac{1}{2}}_{\mu,D}[I]}M\\
&\|\partial u_\lambda \cdot \partial v_R\|_{L^2} \lesssim \|\partial u_\lambda\|_{L^\infty L^2}\|\partial v_R\|_{L^2L^\infty} \lesssim \lambda\mu D^{\frac{1}{2}}\|u_\lambda\|_{X^{0,\frac{1}{2}}_{\mu,D}[I]}M
\end{aligned}$$

These estimates combine to complete the proof of (4.32) and we turn to (4.33). Here we use Bernstein's inequality in the form $P_\eta : L^2L^1 \to \eta L^2$. We have

$$\eta^{\frac{1}{4}}\|P_\eta(u_\lambda \cdot v_R)\|_{L^2} \lesssim \eta^{\frac{1}{4}}\eta\|u_\lambda \cdot v_R\|_{L^2L^1} \lesssim \eta^{\frac{1}{4}}\eta\|u_\lambda\|_{L^\infty L^2}\|v_R\|_{L^2}$$

which clearly suffices using (4.27). For $\Box_{g_{<\sqrt{\eta}}}(u_\lambda \cdot v_R)$ we similarly consider

$$\begin{aligned}
&\|(\Box_{g_{<\sqrt{\lambda}}} - \Box_{g_{<\sqrt{\eta}}})P_\eta(u_\lambda \cdot v_R)\|_{L^2} \lesssim \eta^2\|u_\lambda \cdot v_R\|_{L^2L^1} \lesssim\\
&\qquad\qquad \lesssim \eta^2\|u_\lambda\|_{L^2}\|v_R\|_{L^\infty L^2} \lesssim \eta^2\|u_\lambda\|_{X^{0,\frac{1}{2}}_{\mu,D}[I]}M\\
&\eta\|\partial u_\lambda \cdot \partial v_R\|_{L^2L^1} \lesssim \eta\|\partial u_\lambda\|_{L^\infty L^2}\|\partial v_R\|_{L^2} \lesssim \eta\lambda D^{\frac{1}{2}}\|u_\lambda\|_{X^{0,\frac{1}{2}}_{\mu,D}[I]}M\\
&\eta\|\Box_{g_{<\sqrt{\lambda}}}u_\lambda \cdot v_R\|_{L^2L^1} \lesssim \eta\|\Box_{g_{<\sqrt{\lambda}}}u_\lambda\|_{L^2}\|v_R\|_{L^\infty L^2} \lesssim \eta\lambda D^{\frac{1}{2}}\|u_\lambda\|_{X^{0,\frac{1}{2}}_{\mu,D}[I]}M
\end{aligned}$$

and similarly for $u_\lambda \cdot \Box_{g_{<\sqrt{\lambda}}}v_R$. Each of the term above times $\eta^{-\frac{7}{4}}$ is $\lesssim \lambda/\eta^{\frac{1}{4}}$. □

Finally, we have the following time-dependent wave packets sums analogue of (4.10)-(4.13).

Lemma 4.10. *For a fixed ω and a dyadic frequency η let*

$$v^\omega(t) = \sum_{T\in\mathcal{T}^\pm_\eta,\omega_T=\omega} c_T(t)u_T(t), \qquad\qquad t \in I.$$

For $\nu \geq \eta$, let the vector fields $L, \underline{L}, E$ associated to $\pm g_{\sqrt{\nu}}$ and to ω form a null frame as in section 3. Then one has:

$$\|v^\omega\|_{L^\infty} \lesssim \eta^{\frac{3}{4}} \big(\sum_{T, \omega_T = \omega} |c_T|^2_{L^\infty_t} \big)^{\frac{1}{2}} \tag{4.35}$$

$$\|EP_\eta v^\omega\|_{L^\infty} \lesssim \eta^{\frac{5}{4}} \big(\sum_{T, \omega_T = \omega} |c_T|^2_{L^\infty_t} \big)^{\frac{1}{2}} \tag{4.36}$$

$$\|LP_\eta v^\omega\|_{L^2 L^\infty} \lesssim \eta^{\frac{3}{4}} \big(\sum_{T, \omega_T = \omega} |c_T|^2_{L^\infty_t} + |c_T'|^2_{L^2_t} \big)^{\frac{1}{2}} \tag{4.37}$$

$$\|\underline{L}P_\eta v^\omega\|_{L^\infty} \lesssim \eta^{\frac{7}{4}} \big(\sum_{T, \omega_T = \omega} |c_T|^2_{L^\infty_t} \big)^{\frac{1}{2}}. \tag{4.38}$$

Proof. These follow easily from (4.10)-(4.13). In the case of (4.37) and (4.38) we have to consider the case in which the ∂_t derivative from L or $\underline{L}$ falls on the coefficients $c_T(t)$. We write

$$L(P_\eta v^\omega)(t) = \sum_{T, \omega_T = \omega} c_T(t) L P_\eta u_T(t) + \sum_{T, \omega_T = \omega} c_T'(t) P_\eta u_T(t)$$

For the first sum we use (4.11), while for the second we use (4.10). The bound (4.38) follows from (4.31). □

5. Microlocalized characteristic energy estimates

The proof of the algebra property $X^{s,\theta} \cdot X^{s,\theta} \subset X^{s,\theta}$ relies on estimates for directionally localized half-waves along certain characteristic surfaces.

5.1. **Microlocalization.** For each dyadic number $\alpha \leq 1$, let $\tau + a^\pm_{<\alpha^{-1}}$ be the half-wave symbols for the operator $\Box_{g_{<\alpha^{-1}}}$ (note that $a^\pm_{<\alpha^{-1}}$ is not quite frequency-localized due to the square root) and let $\Phi_t^{\alpha,\pm}(x,\xi) = (x_t^{\alpha,\pm}, \xi_t^{\alpha,\pm})$ denote their Hamiltonian flows.

On one hand, a routine linearization argument as in the proof of Lemma 3.2 shows that the flow map satisfies

$$\frac{\partial^k (x_t^{\alpha,\pm}, \xi_t^{\alpha,\pm})}{\partial (x,\xi)^k} = O(\alpha^{1-|k|}), \quad |k| \geq 1. \tag{5.1}$$

On the other hand, in view of homogeneity the flow map is smoother in some directions than others. To capture this directional information, we consider a class of phase space metrics adapted to the wave equation developed in [10, 11]. For each $0 < \alpha < 1$, define $g := g_\alpha$ by

$$g_{\alpha,(x,\xi)}(y,\eta) = \frac{|\langle y, \xi \rangle|^2}{\alpha^4 |\xi|^2} + \frac{|y \wedge \xi|^2}{\alpha^2 |\xi|^2} + \frac{|\langle \eta, \xi \rangle|^2}{|\xi|^4} + \frac{|\eta \wedge \xi|^2}{\alpha^2 |\xi|^4}. \tag{5.2}$$

We recall

Lemma 5.1 ([11, Lemma 4.2])**.** *The flows $\Phi_t^{\alpha,\pm}$ are bi-Lipschitz and g_α-smooth.*

These metrics shall define the symbol classes $S(m,g)$ in which we work. At a point (x,ξ), the unit ball with respect to $g_{(x,\xi)}$ consists of a $\alpha^2 \times (\alpha)^{n-1}$ rectangle in the spatial variable with long side orthogonal to ξ, and a $|\xi| \times (\alpha|\xi|)^{n-1}$ rectangle in the frequency variable with long side parallel to ξ. One easily verifies that perturbing the basepoint (x,ξ) within this

unit ball yields comparable metrics. Appendix B collects the relevant facts and notation concerning pseudo-differential calculus at this generality.

Throughout this article, for any frequency $\mu \geq 1$ we write

$$\alpha_\mu := \mu^{-1/2}.$$

The parameter μ will eventually be the smaller of the two frequencies in products of the form $u_\lambda v_\mu \in X^{s,\theta}_\lambda \cdot X^{s,\theta}_\mu$, and in this context, α_μ represents the smallest angular scale in the bilinear decomposition of the product.

If $\alpha \in [\alpha_\mu, 1]$, write Ω_α for the collection of half-open dyadic intervals on $S^1 = \mathbb{R}/\mathbb{Z}$ of width α.

For $\theta \in \Omega_\alpha$, let $s^\alpha_\theta(\xi)$ be a 0-homogeneous function supported in the sector defined by $\widehat{\xi} := \xi/|\xi| \in C\theta$, where $C\theta$ denotes the dilate of the interval θ about its center by some (fixed) factor $C > 1$, and define time-dependent symbols $\phi^{\alpha,\pm}_\theta$ by transporting s^α_θ along the flows $\Phi^{\alpha_\mu,\pm}_t$:

$$\phi^{\alpha,\pm}_\theta(t,x,\xi) := s^\alpha_\theta \circ \Phi^{\alpha_\mu,\pm}_{-t}. \tag{5.3}$$

By the previous Lemma, we observe

Lemma 5.2. *The symbols $\phi^{\alpha,\pm}_\theta$ satisfy $\partial_{x,\xi}\phi^{\alpha,\pm}_\theta \in S((\alpha|\xi|)^{-1}, g_{\alpha_\mu})$.*

The notation $S(m,g)$ refers to the symbol classes in Definition B.1.

Remark 5.3. A very similar construction was proposed by Geba-Tataru [11, Section 4]. However, here the time dependent symbols are defined using the same flows $\Phi^{\alpha_\mu,\pm}_t$ for all angular widths $\alpha \geq \alpha_\mu$, whereas in that article the initial symbols s^α_θ are transported along the flows $\Phi^{\alpha,\pm}_t$. Consequently their symbols satisfy the better bounds $\partial\phi^{\alpha,\pm}_\theta \in S(\alpha|\xi|^{-1}, g_\alpha)$.

For each $\lambda \geq \mu$, define the symbols at frequency λ by

$$\phi^{\alpha,\pm}_{\theta,\lambda}(t,x,\xi) := P_{<\lambda/8}(D_x)\phi^{\alpha,\pm}_\theta(t,x,\xi)s_\lambda(\xi), \tag{5.4}$$

where $s_\lambda(\xi)$ is a smooth cutoff supported in the annulus $|\xi| \in [\lambda/4, 4\lambda]$ and equal to 1 for $|\xi| \in [\lambda/2, 2\lambda]$.

Lemma 5.4. *The symbols $\phi^{\alpha,\pm}_{\theta,\lambda}$ satisfy $\partial\phi^{\alpha,\pm}_{\theta,\lambda} \in S((\alpha\lambda)^{-1}, g_{\alpha_\mu})$, and also*

$$\{\tau + a^\pm_{<\alpha_\mu^{-1}}, \phi^{\pm,\alpha}_{\theta,\lambda}\} \in S(1, g_{\alpha_\mu}).$$

Proof. The fact that $\phi^{\alpha,\pm}_{\theta,\lambda} \in S(1,\alpha_\mu)$ is straightforward. Since $\phi^{\alpha,\pm}_\theta$ is transported along the Hamilton flow for $a_{<\alpha_\mu^{-1}}(t,x,\xi)$, the Poisson bracket takes the form

$$-(\partial_x a_{<\alpha_\mu^{-1}}\partial_\xi \tilde{s}_\lambda(\xi))\phi^{\alpha,\pm}_{\theta,\lambda}(t,x,\xi) + s_\lambda(\xi)[H_{a^\pm_{<\alpha_\mu^{-1}}}, P_{<\lambda/8}(D_x)]\phi^{\alpha,\pm}_\theta(t,x,\xi),$$

where $\tilde{s}_\lambda$ is a fattened version of s_λ and $H_a = \langle a_x, \partial_x\rangle - \langle a_\xi, \partial_\xi\rangle$ denotes the Hamiltonian vector field for a symbol a. The first term belongs to $S(\lambda^{-1}, g_{\alpha_\mu})$. Now for any k_1, k_2, one has

$$\partial_x^{k_1}\partial_\xi^{k_2}[\partial_\xi a^\pm_{<\alpha_\mu^{-1}}, P_{<\lambda/8}]\partial_x\phi^{\alpha,\pm}_\theta = \sum_{j_1=0}^{k_1}\sum_{j_2=0}^{k_2}[\partial_x^{j_1}\partial_\xi^{j_2}\partial_\xi a^\pm_{<\alpha_\mu^{-1}}, P_{<\lambda/8}]\partial_x^{k_1-j_1}\partial_\xi^{k_2-j_2}(\partial_x\phi^{\alpha,\pm}_\theta)$$

As $\partial_x\partial_\xi a^{\pm}_{<\alpha_\mu^{-1}}$ is g_{α_μ}-smooth, for any $y_1,\dots,y_{k_1},\eta_1,\dots,\eta_{k_2}$ with $g(y_j,0)=g(0,\eta_j)=1$,

$$\|[\langle y_1,\partial_x\rangle\cdots\langle y_{j_1},\partial_x\rangle\langle\eta_1,\partial_\xi\rangle\dots\langle\eta_{j_2},\partial_\xi\rangle\partial_\xi a^{\pm}_{<\alpha_\mu^{-1}},P_{<\lambda/8}]\|_{L^p\to L^p}\lesssim\lambda^{-1}.$$

Consequently

$$\Big|\prod_{j_1=1}^{k_1}\prod_{j_2=1}^{k_2}\langle y_{j_1},\partial_x\rangle\langle\eta_{j_2},\partial_\xi\rangle[\partial_\xi a^{\pm}_{<\alpha_\mu^{-1}},P_{<\lambda/8}]\partial_x\phi^{\alpha,\pm}_\theta(t,x,\xi)\Big|\lesssim\lambda^{-1}(\alpha|\xi|)^{-1},$$

therefore

$$[\partial_\xi a^{\pm}_{<\alpha_\mu^{-1}},P_{<\lambda/8}](\partial_x\phi^{\alpha,\pm}_\theta)\in S((\alpha\lambda|\xi|)^{-1},g_{\alpha_\mu}).$$

Similarly, as Bernstein implies that $\partial_x^2 a^{\pm}_{<\alpha_\mu^{-1}}\in S(\alpha_\mu^{-\frac12}|\xi|,g_{\alpha_\mu})$, we have

$$[\partial_x a^{\pm}_{<\alpha_\mu^{-1}},P_{<\lambda/8}](\partial_\xi\phi^{\alpha,\pm}_\theta\in S((\alpha_\mu)^{-\frac12}(\alpha\lambda)^{-1},g_{\alpha_\mu}).$$

□

For each x, the symbol $\phi^{\alpha,\pm}_\theta(t,x,\xi)$ is supported in a sector $|\hat\xi-\widehat{\xi^{\alpha_\mu}_\theta}(t,x)|\le c\alpha$, where $\hat\xi:=\xi/|\xi|$ and

$$(x,\theta)\mapsto\big((x^{\alpha_\mu}_\theta(t,x),\theta)\mapsto(x,\xi^{\alpha_\mu}_\theta(t,x))$$

parametrizes the graph of the canonical transformation $\Phi^{\alpha_\mu,\pm}_t$ (the dependence on $\pm$ is suppressed in the notation $\xi^{\alpha_\mu}_\theta(t,x)$). The mollified symbol $\phi^{\alpha,\pm}_{\theta,\lambda}$ is no longer sharply localized to the sector $|\hat\xi-\widehat{\xi^{\alpha_\mu}_\theta}(t,x)|\le c\alpha$, but we can write

$$\phi^{\alpha,\pm}_{\theta,\lambda}(t,x,\xi)=\phi^{\alpha,\pm}_{\theta,\lambda}\chi_{<2c\alpha}(|\hat\xi-\widehat{\xi^{\alpha_\mu}_\theta}(t,x)|)+r^{\alpha,\pm}_{\theta,\lambda},\tag{5.5}$$

where the first symbol has the same regularity as $\phi^{\alpha,\pm}_{\theta,\lambda}$ and $r^{\alpha,\pm}_{\theta,\lambda}=O(\lambda^{-\infty})$.

For each θ, let

$$m_\theta(t,x,\xi):=\langle\alpha^{-1}(|\widehat\xi-\widehat{\xi^{\alpha_\mu}_\theta}(t,x))\rangle^{-1}.$$

In the notation of Section B one has

$$\phi^{\alpha,\pm}_\theta,\ \phi^{\alpha,\pm}_{\theta,\lambda}\in S^1_\alpha(m^\infty_\theta,g_{\alpha_\mu}).$$

For future reference, we also record a technical lemma regarding the time-regularity of symbols.

Lemma 5.5. *There is a decomposition*

$$\partial_t\phi^{\alpha,\pm}_\theta=\psi_1+\psi_2,$$

where $\psi_1\in S(1,g_{\alpha_\mu})$ and ψ_2 satisfies the estimates

$$|\psi_2|\lesssim_N\alpha^{-1}m^N_\theta\ \text{ for all } N,$$
$$\partial_x\psi_2\in\alpha_\mu^{-\frac12}\alpha^{-1}S(m^\infty_\theta,g_{\alpha_\mu})+\alpha_\mu^{-1}S(m^\infty_\theta,g_{\alpha_\mu}),$$
$$\partial_\xi\psi_2\in\alpha^{-1}(\alpha_\mu|\xi|)^{-1}S(m^\infty_\theta,g_{\alpha_\mu});$$

in particular $\partial_t\phi^{\alpha,\pm}_\theta\in\alpha^{-1}S(1,g_{\alpha_\mu})$. Also,

$$\partial_t\{\tau+a^{\pm}_{<\alpha_\mu^{-1}},\phi^{\alpha,\pm}_{\theta,\lambda}\}\in S(\alpha_\mu^{-\frac12}(\alpha_\mu^2\lambda)^{-1}m^\infty_\theta,g_{\alpha_\mu}).$$

Proof. The definition of the symbol implies that

$$\partial_t \phi_\theta^{\alpha,\pm} = \langle a_\xi, \partial_x \rangle \phi_\theta^{\alpha,\pm} - \langle a_x, \partial_\xi \rangle \phi_\theta^{\alpha,\pm}.$$

The first term evidently belongs to $S(1, g_{\alpha_\mu})$. The second term is pointwise bounded by α^{-1} and satisfies

$$\partial_x(\langle a_x, \partial_\xi \rangle \phi_\theta^{\pm,\alpha}) = \langle a_{xx}, \partial_\xi \rangle \phi_\theta^{\alpha,\pm} + \langle a_x, \partial_\xi \rangle \partial_x \phi_\theta^{\alpha,\pm} \in \alpha_\mu^{-\frac{1}{2}} \alpha^{-1} S(m_\theta^\infty, g_{\alpha_\mu}) + \alpha_\mu^{-1} S(m_\theta^\infty, g_{\alpha_\mu}),$$
$$\partial_\xi(\langle a_x, \partial_\xi \rangle \phi_\theta^{\alpha,\pm}) = \langle a_{\xi x}, \partial_\xi \rangle \phi_\theta^{\alpha,\pm} + \langle a_x, \partial_\xi \rangle \partial_\xi \phi_\theta^{\alpha,\pm} \in \alpha_\mu^{-1} (\alpha|\xi|)^{-1} S(m_\theta^\infty, g_{\alpha_\mu}),$$

where Bernstein is used for pointwise bounds a_{xx} and further derivatives as in (1.28). These estimates are preserved by the mollifier $P_{<\lambda/8}(D_x)$.

The second claim is proved by inspecting the Poisson bracket estimates in the previous lemma:

$$\begin{aligned}\partial_t \{\tau + a^\pm_{<\alpha_\mu^{-1}}, \phi^{\alpha,\pm}_{\theta,\lambda}\} &= -(\partial_x \partial_t a_{<\alpha_\mu^{-1}} \partial_\xi \tilde{s}_\lambda(\xi)) \phi^{\pm,\alpha}_{\theta,\lambda} - (\partial_x a_{<\alpha_\mu^{-1}} \partial_\xi \tilde{s}_\lambda(\xi)) \partial_t \phi^{\alpha,\pm}_{\theta,\lambda}(t,x,\xi) \\ &\quad + s_\lambda(\xi)[H_{\partial_t a^\pm_{<\alpha_\mu^{-1}}}, P_{<\lambda/8}(D_x)] \phi_\theta^{\alpha,\pm} + s_\lambda(\xi)[H_{a^\pm_{<\alpha_\mu^{-1}}}, P_{<\lambda/8}(D_x)] \partial_t \phi_\theta^{\alpha,\pm} \\ &\quad \in S(\alpha_\mu^{-\frac{1}{2}} (\alpha_\mu^2 \lambda)^{-1}, g_{\alpha_\mu}),\end{aligned}$$

where we have used the Bernstein type estimates $\partial_x \partial_\xi \partial_t a^\pm_{<\alpha_\mu^{-1}} \in S(\alpha_\mu^{-\frac{1}{4}}, g_{\alpha_\mu})$, respectively $\partial_x^2 \partial_t a^\pm_{<\alpha_\mu^{-1}} \in S(\alpha_\mu^{-\frac{3}{2}} |\xi|, g_{\alpha_\mu})$. □

Proposition 5.6. *Suppose $\alpha \geq \alpha_\mu$. If u is a function at frequency $\lambda > \alpha^{-2}$, then*

$$\sum_{\theta \in \Omega_\alpha} \|\phi^{\pm,\alpha}_{\theta,\lambda}(t, X, D) u\|^2_{X_\pm} \sim \|u\|^2_{X_\pm},$$

where

$$\|u\|^2_{X_\pm} := \|u\|^2_{L^2_{t,x}} + \|(D_t + A^\pm) u\|^2_{L^2_{t,x}},$$

and $a^\pm$ are the half wave symbols for the mollified wave operator

$$p = \tau^2 - 2 g^{0j}_{<\sqrt{\lambda}} \tau \xi_j - g^{ab}_{<\sqrt{\lambda}} \xi_a \xi_b.$$

More precisely, we have

$$\sum_\theta \|\phi^{\pm,\alpha}_{\theta,\lambda}(t, X, D) u\|^2_{L^2} \sim \|u\|^2_{L^2},$$
$$\sum_\theta \|(D_t + A^\pm) \phi^{\pm,\alpha}_{\theta,\lambda}(t, X, D) u\|^2_{L^2} \sim \|(D_t + A^\pm) u\|^2_{L^2} + O(\|u\|^2_{L^2}).$$

In practice we shall usually need only the "$\lesssim$" direction. The following terminology will be convenient for describing the size of various operators.

Definition 5.7. We say a collection of operators $\chi_\theta : L^2_x \to L^2_x$ is *square-summable* with respect to θ if $\sum_\theta \|\chi_\theta u\|^2_{L^2_x} \lesssim \|u\|^2_{L^2_x}$; that is, if $\sum_\theta \chi_\theta^* \chi_\theta$ is bounded on L^2_x.

If the operators χ_θ depend on t, it will be clear from the context whether the implicit constants are uniform or merely square-integrable with respect to t; for the latter case we use the term "$L^2_{t,x}$-square-summable".

From the pseudo-differential calculus in Section B and the Cotlar-Stein almost orthogonality criterion, one immediately deduces

Lemma 5.8. *If symbols $\phi_\theta \in S(m_\theta^\infty, g_{\alpha_\mu})$ are supported in $|\xi| \sim \lambda \geq \alpha_\mu^{-2}$, then the operators $P_\lambda(D)\phi_\theta(X,D)$ are square-summable with respect to θ.*

Proof of Proposition. Consider first the L^2 component. The previous lemma already shows that

$$\sum_\theta \|\phi_{\theta,\lambda}^{\pm,\alpha}(t,X,D)u\|_{L^2}^2 \lesssim \|u\|_{L^2}^2.$$

For the other direction, we note first that if u is localized at frequency λ, then

$$\|u\|_{L^2} \sim \Big\| \sum_{\theta\in\Omega_\alpha} \phi_{\theta,\lambda}^{\pm,\alpha}(t,X,D)u \Big\|_{L^2}.$$

As $\phi_{\theta,\lambda}^{\pm,\alpha} \in S(1, g_{\alpha_\mu})$ and localized to frequency λ in both input and output, the direction "$\gtrsim$" follows directly from Lemma B.5. For the opposite direction, use the first order calculus (B.1) to write

$$u = P_\lambda(D)\Big(\sum_\theta \phi_{\theta,\lambda}^\alpha\Big)^{-1}(t,X,D)\sum_\theta \phi_{\theta,\lambda}^\alpha(t,X,D)u + P_\lambda(D)r(t,X,D)u,$$

where $r \in S((\alpha\lambda)^{-2}, g_{\alpha_\mu})$, and apply Lemma B.5 to obtain

$$\|u\|_{L^2} \lesssim \Big\|\sum_\theta \phi_{\theta,\lambda}^\alpha u\Big\|_{L^2} + (\alpha\lambda)^{-2}\|u\|_{L^2}.$$

(Of course there is nothing to prove if $\sum_\theta \phi_\theta^{\pm,\alpha} \equiv 1$ but recall that according to our construction of the symbols, for a fixed angular scale α the sum is in general merely bounded above and below.)

Hence we may write

$$\|u\|_{L^2}^2 \lesssim \sum_{\theta,\theta'} \langle \chi_\theta u, \chi_{\theta'} u\rangle,$$

The pseudo-differential calculus yields the estimates

$$\|\chi_\theta^* \chi_{\theta'}\|_{L^2\to L^2} \lesssim \langle d_\alpha(\theta,\theta')\rangle^{-N}, \quad d_\alpha(\theta,\theta') := |\alpha^{-1}(\theta-\theta')|.$$

Splitting

$$\sum_{\theta,\theta'}\langle \chi_\theta u, \chi_{\theta'} u\rangle = \sum_{d_\alpha(\theta,\theta')\leq M} \langle \chi_\theta u, \chi_{\theta'} u\rangle + \sum_{d_\alpha(\theta,\theta')>M} \langle \chi_\theta u, \chi_{\theta'} u\rangle,$$

for M large enough the second sum may be absorbed into the left side, while the remaining terms are handled by Cauchy-Schwarz.

Next consider the half-wave component. Without loss of generality we consider just the "+" case and set $\phi_{\theta,\lambda}^\alpha := \phi_{\theta,\lambda}^+$, $a := a^+$, $A := A^+$. Writing

$$(D_t + A)\phi_{\theta,\lambda}^\alpha = \phi_{\theta,\lambda}(D_t + A)\phi_{\theta,\lambda}^\alpha + [D_t + A, \phi_{\theta,\lambda}^\alpha],$$

it suffices by the first part and energy estimates to show that

$$\sum_\theta \|[D_t + A, \phi_{\theta,\lambda}^\alpha]u\|_{L^2}^2 \lesssim \|u\|_{L^\infty L^2}^2.$$

Consider the first the low-frequency portion A_μ, where A is the corresponding half-wave operator for the low-frequency metric $g_{<\sqrt{\mu}}$. By the second-order symbol expansion (B.1), the commutator $[D_t + A_\mu, \phi^\alpha_{\theta,\lambda}]$ has symbol

$$\frac{1}{i}\{\tau + a_\mu, \phi^\alpha_{\theta,\lambda}\} - \frac{1}{2}\int_0^1 r_s\, ds,$$

where

$$\begin{aligned} r_s(t,x,\xi) = \sum_{j,k} e^{is\langle D_y, D_\eta\rangle}\big[&\partial_{\eta_j}\partial_{\eta_k} a(x,\eta)\partial_{y_j}\partial_{y_k}\phi^\alpha_{\theta,\lambda}(y,\xi) \\ &- \partial_{\eta_j}\partial_{\eta_k}\phi^\alpha_{\theta,\lambda}(x,\eta)\partial_{y_j}\partial_{y_k} a(y,\xi)\big]\big|_{\substack{y=x\\ \eta=\xi}}. \end{aligned}$$

The Poisson bracket belongs to $S(m^\infty_\theta, g_{\alpha_\mu})$. Hence $P_\lambda(D)\{\tau + a_\mu, \phi^\alpha_{\theta,\lambda}\}(t,X,D)$ is square-summable by the pseudo-differential calculus. The other frequency outputs result from $\{P_{>\lambda/16}(D_x)a_\mu, \phi^\alpha_{\theta,\lambda}\}(t,X,D) \in \lambda^{-N}OPS(m^\infty_\theta, g_{\alpha_\mu})$, and are therefore square-summable by a brute force bound using Lemma B.6 and the triangle inequality.

To evaluate the remainder, note that as

$$\begin{aligned} &\partial^2_\xi a \in S(|\xi|^{-1}, g_{\alpha_\mu}), \quad \partial^2_x \phi^\alpha_{\theta,\lambda} \in \alpha_\mu^{-2} S(m^\infty_\theta, g_{\alpha_\mu}), \\ &\partial^2_x a \in L^2 S(|\xi|, g_{\alpha_\mu}), \quad \partial^2_\xi \phi^\alpha_{\theta,\lambda} \in (\alpha\lambda)^{-1}(\alpha_\mu\lambda)^{-1} S(m^\infty_\theta, g_{\alpha_\mu}), \end{aligned}$$

by Lemma B.3 one has

$$r_s \in (\alpha_\mu^2\lambda)^{-1} S(m^\infty_\theta, g_{\alpha_\mu}) + (\alpha^2\lambda)^{-\frac{1}{2}}(\alpha_\mu^2\lambda)^{-\frac{1}{2}} L^2 S(m^\infty_\theta, g_\alpha).$$

So $r_s(t,X,D)$ are $L^2_{t,x}$-square-summable by considering the operators $P_\lambda(D) r_s(t,X,D)$ and $(1 - P_\lambda(D)) r_s(t,X,D)$ separately as before.

It remains to show that

$$\sum_\theta \|[A - A_\mu, \phi^\alpha_{\theta,\lambda}]u\|^2_{L^2} \lesssim \|u\|^2_{L^2}.$$

From the computations

$$\begin{aligned} &\partial_\xi(a - a_\mu) \in S(\alpha_\mu^2, g_{\alpha_\mu}), \qquad \partial_x \phi^\alpha_{\theta,\lambda} \in S(1, g_{\alpha_\mu}), \\ &\partial_x(a - a_\mu) \in S(\alpha_\mu^2|\xi|, g_{\alpha_\mu}), \quad \partial_\xi \phi^\alpha_{\theta,\lambda} \in S((\alpha\lambda)^{-1}, g_{\alpha_\mu}) \end{aligned}$$

the formula

$$a \circ b = ab + \frac{1}{i}\int_0^1 e^{is\langle D_y, D_\eta\rangle}\langle \partial_\eta a(x,\eta), \partial_y b(y,\xi)\rangle\big|_{\substack{y=x\\ \eta=\xi}}\, ds,$$

and Lemma B.3, it follows that the symbol of $[A - A_\mu, \phi^\alpha_{\theta,\lambda}]$ belongs to $S(\alpha_\mu^2\alpha^{-1}, g_{\alpha_\mu})$. As before this implies that $[A - A_\mu, \phi^\alpha_{\theta,\lambda}] = \alpha_\mu^2\alpha^{-1}\chi_\theta$ for some square-summable χ_θ. □

5.2. **Characteristic energy estimates.** The purpose of this section is to prove energy estimates for directionally localized half-waves along certain null surfaces.

We begin by recalling the usual characteristic energy estimate for a general function v. For further details, see for instance Alinhac's book [1]. Suppose Ω is a spacetime domain whose boundary $\partial\Omega = \Lambda \cup \Sigma_- \cup \Sigma_+$ decomposes into a null hypersurface Λ and the time slices $\Sigma_\pm = \{t = t_\pm\} \cap \Omega$, where $t_- < t_+$. Let L be a geodesic generator for Λ which is extended to a null frame $\{L, \underline{L}, E\}$ on Ω, so that L, E are tangent to Λ.

By contracting the stress energy tensor

$$T = dv \otimes dv - \frac{1}{2} g^{-1}(dv, dv) g$$

with ∂_t and applying the divergence theorem, one controls on Λ the derivatives of v tangential to Λ:

$$\begin{aligned} \int_\Lambda |Lv|^2 + |Ev|^2 \, d\sigma &\lesssim \int_{\Sigma_\pm} |\nabla_{t,x} u|^2 \, dx + \int_\Omega |\Box u| |\partial_t u| \, dxdt + |\langle T, \pi^{(\partial_t)} \rangle| \, dxdt \\ &\lesssim (1 + |t_+ - t_-|) \|\nabla_{t,x} v\|_{L^\infty L^2}^2 + \|\Box v\|_{L^1 L^2}^2 \end{aligned} \tag{5.6}$$

where $\pi^{(Z)}(X, Y) = \langle \nabla_X Z, Y \rangle + \langle \nabla_Y Z, X \rangle$ is the deformation tensor of a vector field Z.

Suppose further that v is microlocalized such that Lv and Ev are smaller than a generic derivative ∇v. Then by contracting the stress energy tensor instead with L, one deduces

$$\int_\Lambda |Lv|^2 \, d\sigma \lesssim \int_{\Sigma_\pm} |Lv(t_\pm, x)|^2 + |E(t_\pm, x)|^2 \, dx + \int_\Omega |\Box v| |Lv| + |\langle T, \pi^{(L)} \rangle| \, dxdt. \tag{5.7}$$

This yields a tighter estimate for Lv along Λ since the worst component $T^{LL} = T_{\underline{L}\underline{L}} = |\underline{L} v|^2$ is paired with $\pi^{(L)}_{LL} = 0$.

Now factor the symbol

$$\tau^2 - 2 g^{0j}_{<\sqrt{\lambda}} \tau \xi_j - g^{jk}_{\sqrt{\lambda}} \xi_a \xi_b = (\tau + a^+)(\tau + a^-).$$

In the sequel we redenote $a := a^+$ and let $A = a(t, X, D)$ denote the corresponding half-wave operator.

For each direction θ, introduce the associated + null foliation $\Lambda_\theta = \Lambda^\lambda_\theta$, defined by the Hamiltonian flow for the half-wave symbol $\tau + a$, and let $\{L, \underline{L}, E\}$ denote the associated null frame. As before, we parametrize the graph of the flow $(x_0, \xi_0) \mapsto (x_t, \xi_t)$ by the variables (x_t, ξ_0), and write $\xi_{\xi_0}(t, x_t) := \xi_t(x_0, \xi_0)$. Combining (5.1) with $\alpha = \sqrt{\lambda}$ and the computation (3.2), one deduces

$$\frac{\partial^k (x_0, \xi_t)}{\partial (x_t, \xi_0)^k} = O(\lambda^{\frac{|k|-1}{2}}), \quad |k| \geq 1.$$

The operators L, E therefore belong to $OPS^1_{1, \frac{3}{4}}$ when restricted to input frequencies $\geq \lambda$. To obtain estimates for $\partial_t \xi_\theta$, we use the equation

$$\partial_t \xi_\theta = -\langle a_\xi(t, x, \xi_\theta), \partial_x \xi_\theta \rangle - a_x(t, x, \xi_\theta), \tag{5.8}$$

obtained by differentiating the definition with respect to t, to deduce

$$|\partial_x^k \partial_t \xi_\theta| \lesssim \lambda^{\frac{k}{2}} + \lambda^{\frac{k-1}{2}} \lambda^{\frac{1}{4}}, \quad k \geq 1. \tag{5.9}$$

where we used the Bernstein-type estimate (1.28) to bound a_{xx} and higher order x derivatives (one could alternatively replace the $\lambda^{\frac{1}{4}}$ by $M(\|\partial^2 g(t)\|_{L^\infty_x})$).

The main result of this section is

Proposition 5.9. *Let $u = P_\lambda(D_x) u$ be supported in $|t| \leq 2$ and satisfy*

$$\|\nabla_{t,x} u\|_{L^2} + \|\Box_{g_{<\sqrt{\lambda}}} u\|_{L^2} < \infty,$$

and let $u = u^+ + u^-$ be the half-wave decomposition from Corollary 2.20. Suppose $\phi^\alpha_{\theta,\lambda}$ is a pseudo-differential localization operator of the form (5.4), defined via the + flow of the metric $g_{<\sqrt{\mu}}$ with $\mu \le \lambda$. Assume that $\alpha \ge \alpha_\mu := \mu^{-1/2}$, and set $\beta := |\theta - \theta'|$.

(1) If $L = L^+_{\theta'}$ is the null generator for the foliation $\Lambda_{\theta'}$, then

$$\sup_h \int_{\Lambda_{h,\theta'}} |L\phi^\alpha_{\theta,\lambda} u^+|^2 \, d\sigma \lesssim (\alpha+\beta)^2 \|\chi^1_\theta u^+\|^2_{L^2}.$$

(2) If $\beta \ge C\alpha$ for sufficiently large $C > 0$, then for any $\varepsilon > 0$ one has

$$\sup_h \int_{\Lambda^\varepsilon_{h,\theta'}} |\phi^\alpha_{\theta,\lambda} u^+|^2 \, dxdt \lesssim (\varepsilon + (\alpha\lambda)^{-2})\beta^{-2} \|\chi^0_\theta u^+\|^2_{L^2},$$

where $\Lambda^\varepsilon_{h,\theta'}$ denotes an ε-neighborhood of $\Lambda_{h,\theta'}$, and χ^j_θ are operators such that

$$\sum_{\theta\in\Omega_\alpha} \|\chi^j_\theta u^+\|^2_{L^2} \lesssim \lambda^{2(j-1)} \big(\|\nabla u\|^2_{L^2} + \|\Box_{g_{<\sqrt{\lambda}}} u\|^2_{L^2} \big). \tag{5.10}$$

These are variable-coefficient analogues of the estimates on null hyperplanes proved in [39, Section 3], and shall play an essential role in the proof of the algebra property. The small angle case $\beta \sim \alpha$ arises when studying low-modulation outputs, while the transversal case $\beta \sim 1$ is used for high-modulation outputs.

In view of the relation $a^+(t,x,-\xi) = -a^-(t,x,\xi)$, the null foliations and generators are related by $\Lambda^-_{-\theta} = \Lambda^+_\theta$, $L^-_{-\theta} = -L^+_\theta$. Indeed the optical functions satisfy $\Phi^-_{-\theta} = -\Phi^+_\theta$. Consequently one has

Corollary 5.10. *Assume the setup of the previous proposition.*

(1) If $L = L^-_{-\theta'}$ is the null generator for the $-$ foliation $\Lambda^-_{-\theta'}$, then

$$\sup_h \int_{\Lambda^-_{h,-\theta'}} |L\phi^\alpha_{\theta,\lambda} u^+|^2 \, d\sigma \lesssim (\alpha+\beta)^2 \|\chi^1_\theta u^+\|^2_{L^2}.$$

(2) If $\beta \ge C\alpha$ for sufficiently large $C > 0$, then for any $\varepsilon > 0$ one has

$$\sup_h \int_{(\Lambda^-_{h,-\theta'})^\varepsilon} |\phi^\alpha_{\theta,\lambda} u^+|^2 \, dxdt \lesssim (\varepsilon + (\alpha\lambda)^{-2})\beta^{-2} \|\chi^0_\theta u^+\|^2_{L^2},$$

Later we shall consider bilinear estimates of the form $\|(\phi^\alpha_{\theta,\lambda} u) v_T\|_{L^2}$ and $\|Q(\phi_{\theta,\lambda} u, v_T)\|_{L^2}$, where v_T is a frequency μ wave packet concentrating in some tube $T \in \mathcal{T}^\pm_{\mu,\pm\theta'}$. As each T is contained in a union $\bigcup_{|h-h_0|\lesssim\mu^{-1}} \Lambda^\pm_{h,\pm\theta'}$ of null surfaces, we deduce

Corollary 5.11. *Assume the setup of the previous corollary, and let $T^\pm \in \mathcal{T}^\pm_{\pm\theta',\mu}$ be a frequency-μ "tube" with initial direction $\pm\theta'$.*

(1) If $L^\pm = L^\pm_{\pm\theta'}$ is the null generator for the foliation $\Lambda^\pm_{\pm\theta'}$, then

$$\|L^\pm \phi^\alpha_{\theta,\lambda} u^+\|_{L^2(T^\pm)} \lesssim \mu^{-\frac{1}{2}} (\alpha+\beta) \|\chi^1_\theta u^+\|_{L^2}.$$

(2) If $\beta \ge C\alpha$ for sufficiently large $C > 0$, then

$$\|\phi^\alpha_{\theta,\lambda} u^+\|_{L^2(T^\pm)} \lesssim \mu^{-\frac{1}{2}} \beta^{-1} \|\chi^0_\theta u^+\|_{L^2}.$$

Proof of Proposition 5.9, part (1). We begin by recording the symbol estimates that underpin the gain or loss in the angular separation. By a harmless abuse of notation we ignore the prefactor σ in L (see Lemma 3.6) and redenote

$$L = D_t + \partial_\xi a(t,x,\xi_{\theta'}(t,x)) \cdot D_x = D_t + \tilde{A}.$$

The difference between L and the half-wave operator $D_t + A$ is

$$(D_t + A) - L = (A - \tilde{A}),$$

whose symbol is

$$a - \tilde{a} = a(t,x,\xi) - \langle a_\xi(t,x,\xi_{\theta'}(t,x)), \xi\rangle \sim |\xi| \angle(\widehat{\xi}, \widehat{\xi_{\theta'}})^2. \tag{5.11}$$

Indeed, put $\omega := \widehat{\xi_{\theta'}(t,x)}$, and let $f(\xi) := a(t,x,\xi)$, which is 1-homogeneous and $f_{\xi\xi}(\xi)$ is positive-definite in the orthogonal complement of ξ. When $\angle(\widehat{\xi}, -\omega) \geq c \min_{|\omega|=1} f(\omega)$ for a small constant $c > 0$, one computes

$$\begin{aligned} f(\xi) - f_\xi(\omega)\cdot\xi &= |\xi|\langle f_\xi(\widehat{\xi}) - f_\xi(\omega), \widehat{\xi}\rangle \\ &= \int_0^1 |\xi| \langle f_{\xi\xi}(\omega + s(\widehat{\xi} - \omega))\,(\widehat{\xi} - \omega), \widehat{\xi} - \omega - s(\widehat{\xi} - \omega)\rangle\, ds \\ &= |\xi| \int_0^1 (1-s)\langle f_{\xi\xi}(\omega + s(\widehat{\xi} - \omega))\,(\widehat{\xi} - \omega), \widehat{\xi} - \omega\rangle\, ds \\ &\sim |\xi| \angle(\widehat{\xi}, \omega)^2 \end{aligned}$$

In the remaining region, where $\widehat{\xi}$ and ω are nearly antipodal,

$$\begin{aligned} f(\xi) - f_\xi(\omega)\cdot\xi &= f(-\xi) - f_\xi(\omega)\cdot(-\xi) - 2 f_\xi(\omega)\cdot\xi \\ &= 2|\xi|(f(\omega) - O(\angle(\widehat{\xi}, -\omega))) \\ &\gtrsim |\xi|. \end{aligned}$$

Further, in view of the identity (3.7), the symbol of E is 1-homogeneous and satisfies

$$e(t,x,\xi) = |\xi|\langle e(t,x), \widehat{\xi}\rangle = |\xi|\langle e(t,x), \widehat{\xi} - \widehat{\xi_{\theta'}}\rangle. \tag{5.12}$$

Recall that the symbol $\phi^{\alpha,+}_{\theta,\lambda}$ is essentially supported in the sector $\{\xi : \angle(\xi, \xi^\mu_\theta(t,x)) \lesssim \alpha\}$, where ξ^μ_θ is defined by the metric $g_{<\sqrt{\mu}}$ mollified at frequency μ. Precisely, we can decompose

$$\phi^{\alpha,+}_{\theta,\lambda} = \chi^\alpha_\theta + r_\theta,$$

where χ^α_θ is essentially $\phi^{\alpha,+}_\theta(t,x,\xi)s_\lambda(\xi)$ before mollification in the x variable–see (5.4)–and has the required support, while $\|r_\theta\|_{L^2\to L^2} = O(\lambda^{-\infty})$.

Thus since $|\xi_\theta(t,x) - \xi_{\theta'}(t,x)| \sim |\theta - \theta'|$ and $|\widehat{\xi_\theta} - \widehat{\xi^\mu_\theta}| \lesssim \mu^{-1/2} \leq \alpha$, one has

$$|(a - \tilde{a})\phi^\alpha_{\theta,\lambda}| \lesssim (\alpha + |\theta - \theta'|)^2\lambda, \quad |e\phi^\alpha_{\theta,\lambda}| \lesssim (\alpha + |\theta - \theta'|)\lambda. \tag{5.13}$$

On the other hand, if $|\theta - \theta'| \geq C\alpha$ for some large C such that $C\alpha$ dominates the angular width of the cutoff χ^α_θ, the symbol $a - \tilde{a}$ is then microelliptic:

$$|\widehat{\xi} - \widehat{\xi^\mu_\theta}| \lesssim \alpha \Rightarrow |(a - \tilde{a})| \sim (\alpha + |\theta - \theta'|)^2\lambda, \quad |e| \lesssim (\alpha + |\theta - \theta'|)\lambda. \tag{5.14}$$

Without loss of generality we prove the estimate on the surface $\Lambda_{0,\theta'}$. In the sequel we write $\Box := \Box_{g_{<\sqrt{\lambda}}}$.

We apply the estimate (5.7) to the spacetime region

$$\Omega = \bigcup_{h\le 0} \Lambda_{h,\theta'} \cap \{|t| \le 5\},$$

whose boundary is

$$\partial\Omega = (\Lambda_{0,\theta'} \cap \{|t| \le 5\}) \cup \{t = \pm 5\}.$$

In terms of the null frame, we have

$$\int_{\Lambda_{0,\theta'}} |Lv|^2 \, d\sigma \lesssim \sum_{\pm} \int |Lv(\pm 5, x)|^2 + |Ev(\pm 5, x)|^2 \, dx + \int_\Omega |\Box v Lv| \, dxdt + \int_\Omega |\langle T, \pi\rangle| \, dxdt,$$

where $\pi(X, Y) = \langle \nabla_X L, Y\rangle + \langle \nabla_Y L, X\rangle$ is the deformation tensor for L.

Put $v = \phi^\alpha_{\theta,\lambda} u^+$. The boundary terms vanish since the half-wave u^+ is assumed to be supported in $|t| \le 3$.

For the other terms, write

$$\langle T, \pi\rangle = T_{\underline{L}\underline{L}}\pi_{LL} + T_{LL}\pi_{\underline{L}\underline{L}} + 2T_{L\underline{L}}\pi_{\underline{L}L} + 2T_{\underline{L}E}\pi_{LE} + 2T_{LE}\pi_{\underline{L}E} + T_{EE}\pi_{EE}$$

The components of the deformation tensor are

$$\begin{aligned}
\pi_{LL} &= 2\langle \nabla_L L, L\rangle = 0,\\
\pi_{LE} &= \langle \nabla_L L, E\rangle + \langle \nabla_E L, L\rangle = 0,\\
\pi_{L\underline{L}} &= \langle \nabla_L L, \underline{L}\rangle + \langle \nabla_{\underline{L}} L, L\rangle = 0,\\
\pi_{\underline{L}E} &= O(1),\\
\pi_{\underline{L}\underline{L}} &= O(1),\\
\pi_{EE} &= O(1);
\end{aligned}$$

the last three are a consequence of the derivative estimates (3.2) for the optical function. Thus

$$|\langle T, \pi\rangle| \lesssim |LvLv| + |LvEv| + |EvEv|.$$

Altogether we obtain

$$\int_{\Lambda_{\theta,0}} |L\phi^\alpha_{\theta,\lambda} u^+|^2 \, d\sigma \lesssim \|\Box\phi^\alpha_{\theta,\lambda} u^+\|_{L^2} \|L\phi^\alpha_{\theta,\lambda} u^+\|_{L^2} + \|L\phi^\alpha_{\theta,\lambda} u^+\|^2_{L^2} + \|E\phi^\alpha_{\theta,\lambda} u^+\|^2_{L^2}.$$

The claim now follows from the next lemma. □

Lemma 5.12. *If $\{L, \underline{L}, E\}$ is the null frame for the foliation $\Lambda_{\theta'}$, then*

$$\begin{aligned}
L\phi^\alpha_{\theta,\lambda} &= (D_t + A)\phi^\alpha_{\theta,\lambda} + (\alpha + |\theta - \theta'|)^2 \lambda \chi_\theta,\\
E\phi^\alpha_{\theta,\lambda} &= (\alpha + |\theta - \theta'|)\lambda\chi_\theta,
\end{aligned}$$

where χ_θ are L^2_x-square-summable with constants uniform in time. Also

$$\sum_{\theta\in\Omega_\alpha} \|\Box\phi^\alpha_{\theta,\lambda} u^+\|^2_{L^2} \lesssim \|\nabla_{t,x} u\|^2_{L^2} + \|\Box u\|^2_{L^2}.$$

Proof. The second term on the right side of $L\phi^\alpha_{\theta,\lambda}$ is

$$(\tilde{A}-A)\phi^\alpha_{\theta,\lambda} = P_\lambda(\tilde{A}-A)\phi_{\theta,\lambda} + (\tilde{a}_{>\lambda/16} - a_{>\lambda/16})(t,X,D)\phi_{\theta,\lambda}.$$

Since the first term is localized in output frequency, the symbol estimates (5.13) and pseudo-differential calculus (Lemmas B.3 and B.5) imply that the first term is $\alpha^2\lambda\chi_\theta$ for some square-summable operator χ_θ. On the other hand, $a, \tilde{a} \in S^1_{1,3/4}$ in the region $|\xi| \geq \lambda$, one has $a_{>\lambda/16}, \tilde{a}_{>\lambda/16} \in \bigcap_N \lambda^{-N} S^1_{1,3/4}$, so

$$\|(\tilde{a}_{>\lambda/16} - a_{>\lambda/16})(t,X,D)P_\lambda(D)\|_{L^2\to L^2} \lesssim \lambda^{-N} \text{ for any } N.$$

The estimate for $E\phi^\alpha_{\theta,\lambda}$ is similar, writing $E\phi^\alpha_{\theta,\lambda} = P_\lambda(D)E\phi^\alpha_{\theta,\lambda} + R\phi^\alpha_{\theta,\lambda}$, where $\|R\|_{L^2\to L^2} = O(\lambda^{-N})$ for any N, and using (5.13) for the main term.

As $\phi^\alpha_{\theta,\lambda}$ are square-summable in θ, it suffices to prove that

$$\sum_\theta \|[\Box, \phi^\alpha_{\theta,\lambda}u^+\|^2_{L^2} \lesssim \|\nabla_{t,x}u\|^2_{L^2} + \|\Box u\|^2_{L^2}.$$

To this end we note first of all that

$$\sum_\theta \|[\Box - (D_t+A^-)(D_t+A^+)]\phi^\alpha_{\theta,\lambda}u^+\|^2_{L^2} \lesssim \sum_\theta \|\nabla_{t,x}\phi^\alpha_{\theta,\lambda}u^+\|^2_{L^2} \lesssim \|\nabla_{t,x}u\|^2_{L^2},$$

and write

$$[(D_t+A^-)(D_t+A^+), \phi^\alpha_{\theta,\lambda}] = (D_t+A^-)[D_t+A^+, \phi^\alpha_{\theta,\lambda}] + [D_t+A^-, \phi^\alpha_{\theta,\lambda}](D_t+A^+).$$

For the first term we split $A^+ = A_\mu + A^+ - A_\mu$, where A_μ is the corresponding + operator for the low-frequency metric $g_{<\sqrt{\mu}}$. Recall from the proof of Proposition 5.6 that

$$[D_t + A_\mu, \phi^\alpha_{\theta,\lambda}] = \frac{1}{i}\{\tau + a_\mu, \phi^\alpha_{\theta,\lambda}\}(t,X,D) + r(t,X,D),$$

where $\{\tau + a_\mu, \phi^\alpha_{\theta,\lambda}\} \in S(m^\infty_\theta, g_{\alpha_\mu})$, $r \in (\alpha^2_\mu\lambda)^{-1}L^2S(m^\infty_\theta, g_{\alpha_\mu})$, $P_{>\lambda/8}(D_x)r \in \lambda^{-N}L^2S(m^\infty_\theta)$ for any N, and $\partial_t r \in \mu S(m^\infty_\theta, g_{\alpha_\mu})$. The last claim uses the modified computations

$$\begin{aligned}
&\partial^2_\xi\partial_t a \in S(|\xi|^{-1}, g_{\alpha_\mu}), \quad \partial^2_x\partial_t\phi^\alpha_{\theta,\lambda} \in \alpha^{-4}_\mu S(m^\infty_\theta, g_{\alpha_\mu}),\\
&\partial^2_x\partial_t a \in S(\mu^{\frac{3}{4}}|\xi|, g_{\alpha_\mu}), \quad \partial^2_\xi\partial_t\phi^\alpha_{\theta,\lambda} \in \alpha^{-1}(\alpha_\mu\lambda)^{-2}S(m^\infty_\theta, g_{\alpha_\mu}).
\end{aligned}$$

Then

$$\begin{aligned}
(D_t+A^-)\{\tau+a_\mu, \phi^\alpha_{\theta,\lambda}\}(t,X,D) = (D_t\{\tau+a_\mu, \phi^\alpha_{\theta,\lambda}\})(t,X,D) + \{\tau+a_\mu, \phi^\alpha_{\theta,\lambda}\}(t,X,D)D_t\\
+ A^-\{\tau+a_\mu, \phi^\alpha_{\theta,\lambda}\}(t,X,D).
\end{aligned}$$

By Lemma 5.5 the first term belongs to $\alpha^{-1}OPS(m^\infty_\theta, g_{\alpha_\mu})$. Modulo a negligible remainder we may restrict each term to output frequency λ, and write

$$(D_t+A^-)\{\tau+a_\mu, \phi^\alpha_{\theta,\lambda}\}(t,X,D) = \lambda T^1_\theta + T^2_\theta D_t$$

where T^1_θ and T^2_θ are square-summable.

$$\begin{aligned}
(D_t+A^-)r(t,X,D) &= (D_t r)(t,X,D) + r(t,X,D)D_t + A^- r(t,X,D)\\
&= \mu T^1_\theta + f(t)T^2_\theta D_t + \lambda T^3_\theta,
\end{aligned}$$

where $f(t) = M(\|\partial^2 g\|_{L^\infty}) \in L^2_t$ and T^j_θ are square-summable. This is acceptable in view of the energy estimate $\|\nabla u\|_{L^\infty L^2} \lesssim \|\nabla u\|_{L^2} + \|\Box u\|_{L^2}$.

For the term $(D_t + A^-)[A - A_\mu, \phi^\alpha_{\theta,\lambda}]$, we simply recall from the proof of Proposition 5.6 that the commutator belongs to $\alpha_\mu^2 \alpha^{-1} OPS(m^\infty_\theta, g_{\alpha_\mu})$ and outputs essentially at frequency λ, so that $A^-[A - A_\mu, \phi^\alpha_{\theta,\lambda}] \in \lambda \chi_\theta$ for some square-summable χ_θ, and

$$D_t[A - A_\mu, \phi^\alpha_{\theta,\lambda}] = [\partial_t A - \partial_t A_\mu, \phi^\alpha_{\theta,\lambda}] + [A - A_\mu, D_t \phi^\alpha_{\theta,\lambda}] + [A - A_\mu, \phi^\alpha_{\theta,\lambda}] D_t,$$

is acceptable as well. This shows that

$$\sum_\theta \|(D_t + A^-)[D_t + A^+, \phi^\alpha_{\theta,\lambda}] u^+\|^2_{L^2} \lesssim \|\nabla u^+\|^2_{L^2} + \|\Box u^+\|^2_{L^2}.$$

Next we write

$$[D_t + A^-, \phi^\alpha_{\theta,\lambda}] = (D_t \phi^\alpha_{\theta,\lambda})(t, X, D) + [A^-, \phi^\alpha_{\theta,\lambda}] = \alpha^{-1} \chi_\theta$$

for some square summable χ_θ, so that

$$\sum_\theta \|[D_t + A^-, \phi^\alpha_{\theta,\lambda}](D_t + A^+) u^+\|^2_{L^2} \lesssim \alpha^{-2} \|(D_t + A^+) u^+\|^2_{L^2} \lesssim \|\nabla u^+\|^2_{L^2} + \|\Box u^+\|^2_{L^2},$$

Finally, recall from Proposition 2.19 that $\|\nabla u^+\|_{L^2} + \|\Box u^+\|_{L^2} \lesssim \|\nabla u\|_{L^2} + \|\Box u\|_{L^2}$. □

For future reference, we collect the key estimates for $\phi^\alpha_{\theta,\lambda} u$ in the following

Corollary 5.13. *If $\{L, \underline{L}, E\}$ is the null frame for the + foliation $\Lambda_{\theta'}$, and $|\theta - \theta'| \sim \alpha \geq \alpha_\mu$, then:*

$$\Big(\sum_\theta \|\phi^\alpha_{\theta,\lambda} u\|^2_{L^\infty L^2}\Big)^{\frac{1}{2}} \lesssim \|u\|_{X_+} \tag{5.15}$$

$$\Big(\sum_\theta \|\Box_{g_{<\sqrt{\lambda}}} \phi^\alpha_{\theta,\lambda} u\|^2_{L^2 L^2}\Big)^{\frac{1}{2}} \lesssim \|\nabla u\|_{L^2} + \|\Box_{g_{<\sqrt{\lambda}}} u\|_{L^2} \tag{5.16}$$

$$\Big(\sum_\theta \|E \phi^\alpha_{\theta,\lambda} u\|^2_{L^2_x}\Big)^{\frac{1}{2}} \lesssim \lambda \alpha \|u\|_{X_+} \tag{5.17}$$

$$\Big(\sum_\theta \|L \phi^\alpha_{\theta,\lambda} u\|^2_{L^2 L^2}\Big)^{\frac{1}{2}} \lesssim \lambda \alpha^2 \|u\|_{X_+} \tag{5.18}$$

$$\Big(\sum_\theta \|\nabla_{t,x} \phi^\alpha_{\theta,\lambda} u\|^2_{L^\infty L^2}\Big)^{\frac{1}{2}} \lesssim \lambda \|u\|_{X_+} \tag{5.19}$$

Remark 5.14. For split metrics $g^{0j} = 0$, the same estimates hold with the replacements $\Lambda_{\theta'} \to \Lambda^-_{-\theta'}$, $X_+ \to X_-$, and $\alpha \sim |\theta + \theta'|$.

Proof. (5.15) follows from the energy estimate on bounded time intervals $\|v\|_{L^\infty L^2} \lesssim \|v\|_{L^2} + \|(D_t + A^\pm) v\|_{L^2}$.

For the next estimate (5.19) we simply note that by Lemma (5.5), the commutator $[\nabla_{t,x}, \lambda^{-1} \phi^\alpha_{\theta,\lambda}] = \lambda^{-1}(\nabla_{t,x} \phi^\alpha_{\theta,\lambda})$ is square-summable, and apply (5.15).

Finally, (5.16), (5.17), (5.18) were proved in the preceding lemma. □

We turn to the L^2 estimate in Proposition 5.9, which is slightly more involved. Roughly speaking, the angular separation allows one to microlocally invert the vector field L for the

foliation $\Lambda_{\theta'}$ on the support of the cutoff $\phi^\alpha_{\theta,\lambda}$. A similar ellipticity argument was employed previously in [11, Lemma 5.2].

Proof of Prop. 5.9, part (2). Recall from (5.11), (5.14), that the difference $D_t + A - L = A - \tilde{A}$ satisfies

$$q := a - \tilde{a} \approx \beta^2\lambda \tag{5.20}$$

on a neighborhood of the support of χ^α_θ. We construct a microlocal inverse for $A - \tilde{A}$. Let $\tilde{\chi}^\alpha_\theta$ be a slightly wider version of χ^α_θ supported where (5.20) holds. Define

$$\tilde{l}(t,x,\xi) := q^{-1}\tilde{\chi}^\alpha_\theta,$$

and define

$$\tilde{L} := P_\lambda(D)\tilde{l}(t,x,D).$$

Lemma 5.15. *On the support of $\tilde{\chi}^\alpha_\theta$, one has*

$$|(\beta\lambda\partial_\xi)^m\langle\xi,\partial_\xi\rangle^{m'}q| + |(\beta_\lambda\partial_x)^k(\beta\partial_x)(\beta\lambda\partial_\xi)^m\langle\xi,\partial_\xi\rangle^{m'}q| \lesssim (\beta^2\lambda), \quad \alpha_\lambda := \lambda^{-\frac{1}{2}},$$
$$|(\alpha_\lambda\partial_x)^k(\beta\lambda\partial_\xi)^m\langle\xi,\partial_\xi\rangle^{m'}\partial_t q| \lesssim \beta^{-1}(\beta^2\lambda).$$

In conjunction with $\tilde{\chi}^\alpha_\theta \in S^1_\alpha(m^\infty_\theta, g_{\alpha_\mu})$, this quickly leads to

Corollary 5.16. *The symbol $\tilde{l}$ satisfies*

$$\tilde{l},\ \beta\partial_x\tilde{l},\ \alpha\lambda\partial_\xi\tilde{l} \in (\beta^2\lambda)^{-1}S(m^\infty_\theta, g_{\alpha_\lambda}),$$

where as before $m_\theta(t,x,\xi) := \langle\alpha^{-1}(|\widehat{\xi} - \widehat{\xi^\mu_\theta}(t,x))\rangle^{-1}$.

Remark 5.17. The worse regularity of $\tilde{l}$ with respect to ξ is due to the factor $\tilde{\chi}^\alpha_\theta$.

Proof of Lemma. For simplicity of notation we suppress the t, x variables in the arguments to a. We have

$$\partial^k_x q(\xi) = \partial^k_x a(\xi) - \langle\partial^k_x a_\xi(\xi_{\theta'}),\xi\rangle - \sum_{j\geq 1}\langle\partial^{k-j}_x a_{\xi\xi}(\xi_{\theta'})(\partial^j_x\xi_{\theta'}), \xi - |\xi|\widehat{\xi_{\theta'}}\rangle$$
$$- \sum_{j\geq 1}\big[\partial^{k-j}_x\partial^2_\xi a_\xi(\xi_{\theta'})B_2(\partial_x\xi_{\theta'}) + \cdots + \partial^{k-j}_x\partial^j_\xi a_\xi(\xi_{\theta'})B_j(\partial_x\xi_{\theta'})\big],$$

where $B_j(\partial_x\xi_{\theta'})$ denotes a j-linear quantity in $\partial_x\xi_{\theta'}$ and its higher x derivatives such that the total order of the derivatives equals j. This yields

$$|\partial_x q| \lesssim \beta^2\lambda + \beta\lambda,$$

and when $k \geq 2$

$$|\partial^k_x q| \lesssim \lambda^{\frac{k-2}{2}}f(t)\beta^2\lambda| + \sum_{j=1}^{k-2}\lambda^{\frac{k-j-2}{2}}f(t)(\lambda^{\frac{j-1}{2}}\beta\lambda + \lambda^{\max(\frac{j-2}{2},0)}\lambda)$$
$$+ \lambda^{\frac{k-1}{2}}\beta\lambda + \lambda^{\max(\frac{k-2}{2},0)}\lambda,$$

where $f := M(\|\partial^2 g\|_{L^\infty_x}) \in L^2_t$. Since the metric is localized to frequencies $< \sqrt{\lambda}$, we may replace $f(t)$ by the uniform bound $\lambda^{\frac{1}{4}}$ as in (1.28), so that the dominant term when $\lambda \geq \alpha^{-2}$ is $\alpha_\lambda^{1-k}\beta\lambda = \alpha_\lambda^{1-k}\beta^{-1}(\beta^2\lambda)$.

The estimates for the ξ derivatives follow easily from the explicit form of q, and similar considerations handle

$$\partial_t q = a_t(\xi) - \langle a_{t\xi}(\xi_{\theta'}), \xi\rangle - \langle a_{\xi\xi}(\xi_{\theta'})\partial_t \xi_{\theta'}, \xi\rangle.$$

□

Some basic properties of the parametrix $\tilde{L}$ are recorded in

Lemma 5.18. *The operator $\tilde{L}$ satisfies*

$$\|\tilde{L}\|_{L^2\to L^2} \lesssim (\beta^2\lambda)^{-1}$$
$$\|P_\lambda(D)[D_t + A, \tilde{L}]P_\lambda(D)\|_{L^2\to L^2} \lesssim (\beta^2\lambda)^{-1}$$
$$(A - \tilde{A})\tilde{L} - \tilde{\chi}_\theta^{\alpha,+}(t, X, D) = (\alpha\beta\lambda)^{-1}\Phi_\theta,$$

where the operators $P_\lambda(D)\Phi_\theta : L^2 \to L^2$ are square-summable in θ.

Proof. The first follows directly from the estimates for the symbol $\tilde{l}$ and Lemma B.5.

For the third estimate, use the first-order symbol expansion (B.1) and Lemmas 5.15, B.3 to see that

$$(A - \tilde{A})\tilde{L} = \tilde{\chi}_\theta^{\alpha,+}(t, X, D) + r(t, X, D), \quad r \in (\alpha\beta\lambda)^{-1}S(m_\theta^\infty, g_{\alpha_\mu}),$$

and apply Lemma B.5 to the remainder.

For the commutator estimate we essentially follow the proof of [11, Lemma 5.2]. From the second order symbol expansion (B.1), the symbol of the commutator is

$$\frac{1}{i}\{\tau + a, \tilde{l}\} - \frac{1}{2}\int_0^1 r_s(t, x, \xi)\, ds,$$
$$r_s(t, x, \xi) = \sum_{j,k} e^{is\langle D_y, D_\eta\rangle}[\partial_{\eta_j}\partial_{\eta_k} a(x, \eta)\partial_{y_j}\partial_{y_k} l(y, \xi) - \partial_{\eta_j}\partial_{\eta_k} l(x, \eta)\partial_{y_j}\partial_{y_k} a(y, \xi)]|_{\substack{y=x\\ \eta=\xi}}. \quad (5.21)$$

We claim that

$$r_s \in L^2 S((\beta^2\lambda)^{-1}(\alpha_\mu^2\lambda)^{-1}, g_{\alpha_\lambda}) + S(((\alpha_\mu^2\lambda)^{-1} + (\beta^2\lambda)^{-\frac{1}{2}})(\beta^2\lambda)^{-1}, g_{\alpha_\lambda}),$$

and is therefore acceptable. This would follow from Lemma B.3 and the symbol estimates

$$\begin{aligned} \partial_x^2\tilde{l} &\in S((\alpha_\mu^{-2} + \alpha_\lambda^{-1}\beta^{-1})(\beta^2\lambda)^{-1}, g_{\alpha_\lambda}), \qquad \partial_\xi^2 a \in S(|\xi|^{-1}, g_{\alpha_\lambda}) \\ \partial_\xi^2\tilde{l} &\in S((\alpha\lambda)^{-1}(\beta^2\lambda)^{-1}(\alpha_\mu\lambda)^{-1}, g_{\alpha_\lambda}), \qquad \partial_x^2 a \in L^2 S(|\xi|, g_{\alpha_\lambda}). \end{aligned} \quad (5.22)$$

The bounds for a follow directly from the hypotheses on the metric, so we consider next

$$\partial_x^2\tilde{l} = \partial_x^2\tilde{\chi}_\theta^\alpha q^{-1} + 2\partial_x\tilde{\chi}_\theta^\alpha\partial_x q^{-1} + \tilde{\chi}_\theta^\alpha\partial_x^2 q^{-1}.$$

From Lemma 5.15 one sees that on the support of $\tilde{\chi}_\theta^\alpha$,

$$\partial_x(q^{-1}) \in \beta^{-1}S((\beta^2\lambda)^{-1}, g_{\alpha_\lambda}), \quad \partial_x^2(q^{-1}) = \beta^{-1}\alpha_\lambda^{-1}S((\beta^2\lambda)^{-1}, g_{\alpha_\lambda}),$$

thus

$$\partial_x^2\tilde{l} \in S((\alpha_\mu^{-2} + \alpha_\lambda^{-1}\beta^{-1})(\beta^2\lambda)^{-1}, g_{\alpha_\lambda})$$

Similarly

$$\partial_\xi^2\tilde{l} = \partial_\xi^2\tilde{\chi}_\theta^\alpha q^{-1} + 2\partial_\xi\tilde{\chi}_\theta^\alpha\partial_\xi q^{-1} + \tilde{\chi}_\theta^\alpha\partial_\xi^2 q^{-1} \in S((\alpha\lambda)^{-1}(\alpha_\mu\lambda)^{-1}(\beta^2\lambda)^{-1}, g_{\alpha_\lambda})$$

as claimed.

It remains to estimate the Poisson bracket. Letting a_μ denote the half-wave symbol corresponding to the low-frequency metric $g_{<\sqrt{\mu}}$, we write

$$\{\tau + a, \tilde{l}\} = \tilde{\chi}_\theta^\alpha q^{-2}\{\tau + a, q\} + q^{-1}\{\tau + a_\mu, \tilde{\chi}_\theta^\alpha\} + q^{-1}\{a - a_\mu, \tilde{\chi}_\theta^\alpha\}. \tag{5.23}$$

The symbol q vanishes to second order on the submanifold

$$\{(t, x, \xi_\theta(t, x))\} \subset \mathbb{R}_t \times (T^*\mathbb{R}^2_x \setminus 0),$$

which is invariant under the Hamiltonian flow. Hence its derivative along the flow $\{\tau + a, q\}$ vanishes to second order as well, and is smooth and 1-homogeneous ξ. Therefore we see that

$$\{\tau + a, q\} \in S((\beta^2|\xi|), g_{\alpha_\lambda}).$$

The argument of Lemma 5.4 shows that $\{\tau + a_\mu, \tilde{\chi}_\theta^\alpha\} \in S(1, g_{\alpha_\mu})$.

Finally, by a direct computation $\{a - a_\mu, \tilde{\chi}_\theta^\alpha\} \in S((\alpha_\mu^2\lambda)(\alpha\lambda)^{-1}, g_{\alpha_\mu})$. Altogether we obtain

$$\{\tau + a, \tilde{l}\} \in S((\beta^2\lambda)^{-1}, g_{\alpha_\lambda}).$$

□

Later it will be useful to have a more computational proof of the Poisson bracket bound $\{\tau + a, q\} \in S(\alpha^2\lambda, g_{\alpha_\lambda})$. Using the equation (5.8) for $\partial_t\xi_{\theta'}$, we compute

$$\begin{aligned}
\{\tau + a, q\} &= \partial_t q + \langle a_\xi, \partial_x q\rangle - \langle a_x, \partial_\xi q\rangle \\
&= a_t(\xi) - \langle a_{t\xi}(\xi_{\theta'}), \xi\rangle + \big\langle\langle a_\xi(\xi), (a_x(\xi) - \langle a_{x\xi}(\xi_{\theta'}), \xi\rangle)\big\rangle \\
&+ |\xi|\langle a_{\xi\xi}(\xi_{\theta'})(\widehat{\xi} - \widehat{\xi_{\theta'}}), \langle a_\xi(\xi_{\theta'}) - a_\xi(\xi), \partial_x\rangle\widehat{\xi_{\theta'}}\rangle \\
&+ |\xi|\langle a_\xi(\xi) - a_\xi(\xi_0), a_x(\widehat{\xi_{\theta'}}) - a_x(\widehat{\xi})\rangle \\
&- |\xi|\langle a_{\xi\xi}(a_\xi(\widehat{\xi}) - a_\xi(\widehat{\xi_{\theta'}}) - a_{\xi\xi}(\widehat{\xi_{\theta'}})(\widehat{\xi} - \widehat{\xi_{\theta'}})\rangle,
\end{aligned}$$

and an inductive argument similar to the proof of Lemma 5.15 yields the estimates on the support of $\tilde{\chi}_\theta^\alpha$

$$\begin{aligned}
|(\alpha_\lambda\partial_x)^k(\beta\lambda\partial_\xi)^m\langle\xi, \partial_\xi\rangle^{m'}\{\tau + a, q\}| &\lesssim \beta^2\lambda, \\
|(\alpha_\lambda\partial_x)^k(\beta\lambda\partial_\xi)^m\langle\xi, \partial_\xi\rangle^{m'}\partial_t\{\tau + a, q\}| &\lesssim \alpha_\lambda^{-1}(\beta^2\lambda).
\end{aligned} \tag{5.24}$$

Using the expansion (5.23) and Lemma 5.15, one deduces that

$$\partial_t\{\tau + a, \tilde{l}\} \in \alpha_\lambda^{-1}S((\beta^2\lambda)^{-1}, g_{\alpha_\lambda}). \tag{5.25}$$

We continue the proof of the proposition. Write

$$\begin{aligned}
\phi_{\theta,\lambda}^\alpha u^+ &= P_\lambda(D)(D_t + A - L)\tilde{L}\phi_{\theta,\lambda}^\alpha u^+ + P_\lambda(D)[\tilde{\chi}_\theta^\alpha - (A - \tilde{A})\tilde{L}]\phi_{\theta,\lambda}^\alpha u^+ + P_\lambda(D)(1 - \tilde{\chi}_\theta^\alpha)\phi_{\theta,\lambda}^\alpha u^+ \\
&= P_\lambda(D)(D_t + A - L)\tilde{L}\phi_{\theta,\lambda}^\alpha u^+ + R_1u^+ + R_2u^+.
\end{aligned}$$

By the previous lemma and the pseudo-differential calculus, the second and third terms both take the form $(\lambda^{-1/2} + (\alpha^2\lambda)^{-1})\chi_\theta$ where χ_θ is square-summable. Consequently, we estimate

$$\|(R_ju^+)\|_{L^2(\Lambda_{0,\theta}^\varepsilon)} \lesssim (\alpha\beta\lambda)^{-1}\|\chi_\theta u^+\|_{L^2}.$$

Also,

$$P_\lambda(D_t + A)\tilde{L}\phi_{\theta,\lambda}^\alpha u^+ = \tilde{L}(D_t + A)\phi_{\theta,\lambda}^\alpha u^+ + P_\lambda(D)[D_t + A, \tilde{L}]\phi_{\theta,\lambda}^\alpha u^+.$$

By the previous lemma, this is bounded in L^2 by

$$\begin{aligned}&(\beta^2\lambda)^{-1}\|(D_t+A)\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}+(\beta^2\lambda)^{-1}\|\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}\\ &\lesssim(\beta^2\lambda)^{-1}(\|\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}+\|(D_t+A)\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}).\end{aligned}$$

which is sufficient.

The remaining term is

$$P_\lambda(D)L\tilde{L}\phi^\alpha_{\theta,\lambda}u^+=L\tilde{L}\phi^\alpha_{\theta,\lambda}u^++[P_\lambda,L]\tilde{L}\phi^\alpha_{\theta,\lambda}u^+.$$

As the commutator $[P_\lambda,L]=[P_\lambda,\tilde{A}]$ is bounded on L^2 at frequency λ,

$$\|([P_\lambda,L]\tilde{L}\phi^\alpha_{\theta,\lambda}u^+)\|_{L^2}\lesssim\|\tilde{L}\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}\lesssim(\alpha^2\lambda)^{-1}\|\phi^\alpha_{\theta,\lambda}u^+\|_{L^2},$$

which is acceptable.

The remaining term is

$$\|L\phi^\alpha_{\theta,\lambda}u^+\|^2_{L^2(\Lambda^\varepsilon_{0,\theta'})}\lesssim\int_{|h|\le\varepsilon}\|L\phi^\alpha_{\theta,\lambda}u^+\|^2_{L^2(\Lambda_{h,\theta'})}\,dh.$$

For each null surface $\Lambda_{h,\theta'}$, we have

$$\int_{\Lambda_{h,\theta'}}|L\tilde{L}\phi^\alpha_{\theta,\lambda}u^+|^2\,d\sigma\lesssim\int|\Box\tilde{L}\phi^\alpha_{\theta,\lambda}u^+||L\phi^\alpha_{\theta,\lambda}u^+|\,dxdt+\int|\langle T,\pi\rangle|\,dxdt.$$

For the second term, write

$$\begin{aligned}\langle T,\pi\rangle&=T_{\underline{L}\underline{L}}\pi_{LL}+T_{LL}\pi_{\underline{L}\underline{L}}+2T_{L\underline{L}}\pi_{\underline{L}L}+2T_{\underline{L}E}\pi_{LE}+2T_{LE}\pi_{\underline{L}E}+T_{EE}\pi_{EE}\\&\lesssim|L\tilde{L}\phi^\alpha_{\theta,\lambda}u^+L\tilde{L}\phi^\alpha_{\theta,\lambda}u^+|+|L\tilde{L}\phi^\alpha_{\theta,\lambda}u^+E\tilde{L}\phi^\alpha_{\theta,\lambda}u^+|+|E\tilde{L}\phi^\alpha_{\theta,\lambda}u^+E\tilde{L}\phi^\alpha_{\theta,\lambda}u^+|,\end{aligned}$$

so

$$\begin{aligned}&\|(L\tilde{L}\phi^\alpha_{\theta,\lambda}u^+)v_T\|^2_{L^2(\Lambda^\varepsilon_{0,\theta'})}\\&\lesssim\varepsilon\Big(\|\Box\tilde{L}\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}\|L\tilde{L}\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}+\|L\tilde{L}\phi^\alpha_{\theta,\lambda}u^+\|^2_{L^2}+\|E\tilde{L}\phi^\alpha_{\theta,\lambda}u^+\|^2_{L^2}\Big)\\&\lesssim\varepsilon(\beta^2\lambda)^{-2}\Big(\|\Box\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}\|L\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}+\|L\phi^\alpha_{\theta,\lambda}u^+\|^2_{L^2}+\|E\phi^\alpha_{\theta,\lambda}u^+\|^2_{L^2}\Big)\\&+\varepsilon\Big(\|[\Box,\tilde{L}]\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}\big((\beta^2\lambda)^{-1}\|L\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}+\|[L,\tilde{L}]\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}\big)\\&+\big((\beta^2\lambda)^{-1}\|\Box\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}+\|[\Box,\tilde{L}]\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}\big)\|[L,\tilde{L}]\phi^\alpha_{\theta,\lambda}u^+\|_{L^2}\\&+\|[L,\tilde{L}]\phi^\alpha_{\theta,\lambda}u^+\|^2_{L^2}+\|[E,\tilde{L}]\phi^\alpha_{\theta,\lambda}u^+\|^2_{L^2}\Big),\end{aligned}$$

and appeal to Lemmas 5.12 and 5.19. □

Lemma 5.19. *If $\{L,\underline{L},E\}$ is the null frame for the foliation $\Lambda_{\theta'}$, then*

$$\begin{aligned}&\|[E,\tilde{L}]\|_{L^2\to L^2}\lesssim\alpha^{-1}(\beta^2\lambda)^{-1},\\&\|[L,\tilde{L}]\|_{L^2\to L^2}\lesssim(\beta^2\lambda)^{-1},\\&\|[\Box,\tilde{L}]u\|_{L^2\to L^2}\lesssim(\beta^2\lambda)^{-1}(\|\nabla_{t,x}u\|_{L^2}+\lambda\|(D_t+A)u\|_{L^2})\ \text{if } u=P_\lambda(D_x)u.\end{aligned}$$

Proof. The estimate for $[E, \tilde{L}]$ is simplest, using the first order symbol expansion (B.1), the computations

$$\begin{aligned} &\partial_x e \in S(|\xi|, g_{\alpha_\mu}), && \partial_\xi e \in S(1, g_{\alpha_\mu}), \\ &\partial_x \tilde{l} \in S(\beta^{-1}(\beta^2\lambda)^{-1}, g_{\alpha_\mu}) && \partial_\xi \tilde{l} \in S((\alpha\lambda)^{-1}(\beta^2\lambda)^{-1}, g_{\alpha_\mu}), \end{aligned}$$

and Lemmas B.3 and B.5 (again considering the outputs at frequency λ and $\neq \lambda$ separately).

Next, write $[L, \tilde{L}] = [D + A, \tilde{L}] + [\tilde{A} - A, \tilde{L}]$. The first term was studied in the previous lemma, while the second again follows from first order calculus.

The commutator estimate for $\Box$ is proven similarly as the lemma in the previous section, differing mainly in the symbol estimates involved.

It remains to consider $[\Box, \tilde{L}]$. As usual we first replace $\Box$ with $(D_t + A^-)(D_t + A^+)$ at the cost of an acceptable error, and write

$$[(D_t + A^-)(D + A^+), \tilde{L}] = (D_t + A^-)[D_t + A^+, \tilde{L}] + [D_t + A^-, \tilde{L}](D_t + A^+). \tag{5.26}$$

We proceed similarly as in the proof of Lemma 5.18 but first gather symbol bounds for $\partial_t \tilde{l}$. This is a routine computation using Lemmas 5.5 and 5.15 which leads to $\partial_t \tilde{l} \in S(\alpha^{-1}(\beta^2\lambda)^{-1}, g_{\alpha_\lambda})$,

$$\begin{aligned} &\partial_x \partial_t \tilde{l} \in S((\alpha_\mu^{-2} + \alpha^{-1}\alpha_\lambda^{-1})(\beta^2\lambda)^{-1}, g_{\alpha_\lambda}), && \partial_x^2 \partial_t \tilde{l} \in S((\alpha_\mu^{-4} + \alpha^{-1}\alpha_\lambda^{-2})(\beta^2\lambda)^{-1}, g_{\alpha_\lambda}), \\ &\partial_\xi \partial_t \tilde{l} \in S((\alpha_\mu\alpha\lambda)^{-1}(\beta^2\lambda)^{-1}, g_{\alpha_\lambda}), && \partial_\xi^2 \partial_t \tilde{l} \in S((\alpha\lambda)^{-1}(\alpha_\mu^2\lambda)^{-1}(\beta^2\lambda)^{-1}, g_{\alpha_\lambda}); \end{aligned} \tag{5.27}$$

the factors of α arise from derivatives that land on the χ_θ^α factor. Also note that

$$\begin{aligned} &\partial_x \partial_t a \in f(t) S(|\xi|, g_{\alpha_\lambda}), && \partial_x^2 \partial_t a \in \alpha_\lambda^{-1} f(t) S(|\xi|, g_{\alpha_\lambda}), \\ &\partial_\xi \partial_t a \in S(1, g_{\alpha_\lambda}), && \partial_\xi^2 \partial_t a \in S(|\xi|^{-1}, g_{\alpha_\lambda}), \end{aligned} \tag{5.28}$$

where as usual $f(t) = M(\|\partial^2 g(t)\|_{L_x^\infty})$.

The second term on the right side of (5.26) is handled by the first order estimates

$$\|P_\lambda(D)(\partial_t \tilde{l})(t, X, D)\|_{L^2 \to L^2} \lesssim \alpha^{-1}(\beta^2\lambda)^{-1}, \quad \|[A^-, \tilde{L}]\|_{L^2 \to L^2} \lesssim \alpha^{-1}(\beta^2\lambda)^{-1}.$$

Also recall from the proof of Lemma 5.18 that the symbol of $[D_t + A, \tilde{L}]$ is

$$\frac{1}{i}\{\tau + a, \tilde{l}\} + r,$$

where

$$r \in f(t) S((\beta^2\lambda)^{-1}(\alpha_\mu^2\lambda)^{-1}, g_{\alpha_\lambda}) + S(((\alpha_\mu^2\lambda)^{-1} + (\beta^2\lambda)^{-\frac{1}{2}})(\beta^2\lambda)^{-1}, g_{\alpha_\lambda}).$$

Combining the estimates (5.22), (5.27), (5.28) with the explicit form (5.21) of the second order commutator expansion and Lemma B.3, one obtains $\partial_t r = r_1 + r_2 + r_3 + r_4$, where

$$\begin{aligned} &r_1 \in \alpha^{-1} f(t) S((\alpha_\mu\alpha_\lambda\lambda)^{-1}(\beta^2\lambda)^{-1}, g_{\alpha_\lambda}), && r_2 \in \alpha^{-1} f(t) S((\alpha_\mu^2\lambda)^{-1}(\beta^2\lambda)^{-1}, g_{\alpha_\lambda}), \\ &r_3 \in S((\alpha_\mu^2\lambda)^{-1} + (\beta^2\lambda)^{-\frac{1}{2}})(\beta^2\lambda)^{-1}, g_{\alpha_\lambda}), && r_4 \in S(((\alpha_\mu^{-2}(\alpha_\mu^2\lambda)^{-1} + \alpha^{-1})(\beta^2\lambda)^{-1}, g_{\alpha_\lambda}). \end{aligned} \tag{5.29}$$

These correspond to the pairings $\{\partial_\xi^2 \tilde{l},\ \partial_x^2 \partial_t a\}$, $\{\partial_\xi^2 \partial_t \tilde{l},\ \partial_x^2 a\}$, $\{\partial_x^2 \tilde{l},\ \partial_\xi^2 \partial_t a\}$, and $\{\partial_x^2 \partial_t \tilde{l},\ \partial_\xi^2 a\}$.

For the first term on the right side of of (5.26), we have

$$\big[(D_t + A^-), [D_t + A^+, \tilde{L}]\big] = \frac{1}{i}(D_t\{\tau + a^+, \tilde{l}\})(t, X, D) + (D_t r)(t, X, D) \\ + \frac{1}{i}[A^-, \{\tau + a^+, \tilde{l}\}(t, X, D)] + [A^-, r(t, X, D)].$$

The bounds (5.29) imply that

$$\|(D_t r)(t, X, D)u\|_{L^2} \lesssim \alpha^{-1}(\beta^2\lambda)^{-1}\|u\|_{L^\infty L^2} + \mu(\beta^2\lambda)^{-1}\|u\|_{L^2}$$

which is acceptable in view of the energy estimate $\|u\|_{L^\infty L^2} \lesssim \|u\|_{L^2} + \|(D_t + A^+)u\|_{L^2}$.

By first order estimates, recalling that $\{\tau + a^+, \tilde{l}\} \in S((\beta^2\lambda)^{-1}, g_{\alpha_\lambda})$ the last two terms of the commutator satisfy

$$\|[A^-, \{\tau + a^+, \tilde{l}\}(t, X, D)]\|_{L^2\to L^2} \lesssim \beta^{-2} \\ \|[A^-, r(t, X, D)]\|_{L^2\to L^2} \lesssim \beta^{-2} f(t)(\alpha_\mu^2\lambda)^{-1} + \beta^{-2}\big((\alpha_\mu^2\lambda)^{-1} + (\beta^2\lambda)^{-\frac{1}{2}}\big),$$

which is also acceptable by the energy estimate.

Finally, the Poisson bracket estimate (5.25) shows that

$$\|(D_t\{\tau + a^+, \tilde{l}\})(t, X, D)\|_{L^2\to L^2} \lesssim \beta^{-1}(\beta^2\lambda)^{-\frac{1}{2}}.$$

This completes the proof of the lemma. □

6. The algebra property (1.15)

In this section we prove the estimate (1.15).

Proposition 6.1. *Assume that $\theta > \frac{1}{2}$ and $s > \theta + \frac{1}{2}$. Then the space $X^{s,\theta}$ is an algebra. Moreover, for $\sigma > s$ we have*

$$\|u \cdot v\|_{X^{\sigma,\theta}} \lesssim \|u\|_{X^{\sigma,\theta}}\|v\|_{X^{s,\theta}} + \|u\|_{X^{s,\theta}}\|v\|_{X^{\sigma,\theta}} \tag{6.1}$$

The proof of this algebra property will be based on the estimates in the following two propositions. The next Proposition is the crux of our result and is the variable coefficients analogue of Theorem 3 from [39] since, due to the low modulations, we may think of $u_{\lambda,1}, v_{\mu,1}$ as being approximately free waves.

Proposition 6.2. *Let $u_{\lambda,1}, v_{\mu,1}$, $v_{\lambda',1}$ be functions localized at frequency $\simeq \lambda$, $\simeq \mu$, resp. $\simeq \lambda'$ and let $d_0 = \min(\mu, \frac{\lambda}{\mu})$. Then*

(1) In the high-low case $\mu \ll \lambda$ we have:

$$\|u_{\lambda,1} \cdot v_{\mu,1}\|_{X^{0,\frac{1}{4}}_{\lambda',\le\mu,\infty}} \lesssim \mu^{\frac{3}{4}}\|u_{\lambda,1}\|_{X^{0,\frac{1}{2}}_{\lambda,1}}\|v_{\mu,1}\|_{X^{0,\frac{1}{2}}_{\mu,1}} \tag{6.2}$$

(2) In the high-high to low case $\mu \lesssim \lambda \simeq \lambda'$ we have

$$\|P_\mu(u_{\lambda,1} \cdot v_{\lambda',1})\|_{X^{1,\frac{1}{4}}_{\mu,[d_0,\mu],\infty} + (\tilde{X}^{1,\frac{1}{4}}_{\mu,\mu} \cap \frac{\lambda}{\mu} X^{1,\frac{1}{4}}_{\mu,\mu})} \lesssim \frac{\mu^{\frac{3}{4}}}{\lambda}\|u_{\lambda,1}\|_{X^{1,\frac{1}{2}}_{\lambda,1}}\|v_{\lambda',1}\|_{X^{1,\frac{1}{2}}_{\lambda',1}} \tag{6.3}$$

Here, due to the selection of modulation 1 inputs, the index $\frac{1}{2}$ on the right is superfluous. We have kept it in order to make it easier to compare the above result with the next result.

One can easily go from modulation 1 to any larger modulations by a rescaling and time orthogonality argument. This is accomplished in the next proposition.

Proposition 6.3. *Let $u_{\lambda,d_1}, v_{\mu,d_2}$, v_{λ',d_2} be functions localized at frequency $\simeq \lambda$, $\simeq \mu$, respectively $\simeq \lambda'$ where $d_1, d_2 \geq 1$ and denote $d_{\max} = \max(d_1, d_2)$.*

(1) (Low modulations) For $1 \leq d_1, d_2 \leq \mu < \lambda \simeq \lambda'$, denoting $d_0 = \min\left(\frac{\lambda}{\mu}, \frac{\mu}{d_{\max}}\right)$

$$\|u_{\lambda,d_1} \cdot v_{\mu,d_2}\|_{X^{1,\frac{1}{2}}_{\lambda',[d_{\max},\mu]}} \lesssim \|u_{\lambda,d_1}\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}} \|v_{\mu,d_2}\|_{X^{1,\frac{1}{2}}_{\mu,d_2}} \tag{6.4}$$

$$\|P_\mu(u_{\lambda,d_1} \cdot v_{\lambda',d_2})\|_{X^{1,\frac{1}{2}}_{\mu,[d_{\max}d_0,\mu]}+(\tilde{X}^{1,\frac{1}{2}}_{\mu,\mu}\cap \frac{\lambda}{\mu}X^{1,\frac{1}{2}}_{\mu,\mu})} \lesssim \frac{\mu}{\lambda}\|u_{\lambda,d_1}\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}} \|v_{\lambda',d_2}\|_{X^{1,\frac{1}{2}}_{\lambda',d_2}} \tag{6.5}$$

(2) (High modulations) For $1 \leq d_2 \leq \mu \leq d_1 \leq \lambda \simeq \lambda'$ we have

$$\|u_{\lambda,d_1} \cdot v_{\mu,d_2}\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}} \lesssim \|u_{\lambda,d_1}\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}} \|v_{\mu,d_2}\|_{X^{1,\frac{1}{2}}_{\mu,d_2}} \tag{6.6}$$

For $1 \leq \mu \leq d_{\max} \leq \lambda \simeq \lambda'$ we have

$$\|P_\mu(u_{\lambda,d_1} \cdot v_{\lambda',d_2})\|_{\frac{\lambda}{\mu}X^{1,\frac{1}{2}}_{\mu,\mu}\cap \tilde{X}^{1,\frac{1}{2}}_{\mu,\mu}} \lesssim \frac{\mu}{\lambda}\|u_{\lambda,d_1}\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}} \|v_{\lambda',d_2}\|_{X^{1,\frac{1}{2}}_{\lambda',d_2}} \tag{6.7}$$

6.1. **Proof of Proposition 6.1.** The proof consists of a simple summation argument based on Proposition 6.3. Let $u^1, u^2 \in X^{s,\theta}$ and write

$$u^i = \sum_{\lambda \geq 1} P_\lambda u^i_\lambda, \qquad \|u^i\|^2_{X^{s,\theta}} \simeq \sum_{\lambda \geq 1} \|u^i_\lambda\|^2_{X^{s,\theta}_\lambda}, \qquad i \in \overline{1,2}.$$

By Remark 2.6 we may assume that the u^i_λ are localized in frequency. By the standard Littlewood-Paley decomposition we write

$$u^1 \cdot u^2 = \sum_{\lambda_1,\lambda_2,\lambda_3 \geq 1} P_{\lambda_3}(P_{\lambda_1} u^1_{\lambda_1} \cdot P_{\lambda_2} u^2_{\lambda_2}).$$

By splitting the sum into three terms corresponding to the three cases: $\lambda_1 \ll \lambda_2 \simeq \lambda_3$, $\lambda_2 \ll \lambda_1 \simeq \lambda_3$, $\lambda_3 \lesssim \lambda_1 \simeq \lambda_2$ we obtain $u^1 \cdot u^2 \in X^{s,\theta}$ from the following estimates, stated for any frequency localized functions u_λ, v_μ: Let $s = \theta + \frac{1}{2} + \varepsilon$ for $\varepsilon > 0$. For $\mu \ll \lambda$, $\lambda' \simeq \lambda$ we have

$$\|P_\lambda u_\lambda \cdot P_\mu v_\mu\|_{X^{s,\theta}_{\lambda'}} \lesssim \frac{1}{\mu^\varepsilon}\|u_\lambda\|_{X^{s,\theta}_\lambda}\|v_\mu\|_{X^{s,\theta}_\mu} \tag{6.8}$$

For $\mu \lesssim \lambda$, $\lambda' \simeq \lambda$ we have

$$\|P_\mu(P_\lambda u_\lambda \cdot P_{\lambda'} v_{\lambda'})\|_{X^{s,\theta}_\mu} \lesssim \frac{\mu^{s-1}}{\lambda^{s-1+\varepsilon}}\|u_\lambda\|_{X^{s,\theta}_\lambda}\|v_{\lambda'}\|_{X^{s,\theta}_{\lambda'}} \tag{6.9}$$

Here it is essential that we are in the subcritical case $s > 1$ which allows us to have the the power $\mu^{-\varepsilon}$ in (6.8). We write

$$u_\lambda = \sum_{d=1}^{\lambda} u_{\lambda,d} \qquad \|u_\lambda\|^2_{X^{s,\theta}_\lambda} \simeq \sum_{d=1}^{\lambda} \|u_{\lambda,d}\|^2_{X^{s,\theta}_{\lambda,d}}.$$

and the similar decompositions for $v_\mu, v_{\lambda'}$. Then, using (6.4), (6.6) we have

$$\begin{aligned}
\|P_\lambda u_\lambda \cdot P_\mu v_\mu\|_{X^{s,\theta}_{\lambda'}} &\lesssim \sum_{d_1,d_2\le\mu} \|P_\lambda u_{\lambda,d_1}\cdot P_\mu v_{\mu,d_2}\|_{X^{s,\theta}_{\lambda',[d_{\max},\mu]}} + \sum_{d_2\le\mu} \|P_\lambda u_{\lambda,\ge\mu}\cdot P_\mu v_{\mu,d_2}\|_{X^{s,\theta}_{\lambda',\ge\mu}}\\
&\lesssim \lambda^{s-1}\mu^{\theta-\frac12}\sum_{d_1,d_2\le\mu}\|u_{\lambda,d_1}\|_{X^{1,\frac12}_{\lambda,d_1}}\|v_{\mu,d_2}\|_{X^{1,\frac12}_{\mu,d_2}} + \lambda^{s-1}\|u_\lambda\|_{X^{1,\theta}_{\lambda,[\mu,\lambda]}}\sum_{d_2\le\mu}\|v_{\mu,d_2}\|_{X^{1,\frac12}_{\mu,d_2}}\\
&\lesssim \frac{\mu^{\theta-\frac12}}{\mu^{s-1}}\|u_\lambda\|_{X^{s,\theta}_\lambda}\|v_\mu\|_{X^{s,\theta}_\mu} + \frac{1}{\mu^{s-1}}\|u_\lambda\|_{X^{s,\theta}_\lambda}\|v_\mu\|_{X^{s,\theta}_\mu} \lesssim \frac{1}{\mu^\varepsilon}\|u_\lambda\|_{X^{s,\theta}_\lambda}\|v_\mu\|_{X^{s,\theta}_\mu}
\end{aligned}$$

In this argument we have used the factors $d_i^{\frac12-\theta}$ to get square sums in $d_i\le\mu$ while the modulation square summability of $X^{s,\theta}_{\lambda',\ge\mu}$ is inherited from $\|u_\lambda\|_{X^{1,\theta}_{\lambda,[\mu,\lambda]}}$ due to (6.6). The proof of (6.9) is similar, using (6.5), (6.7):

$$\begin{aligned}
\|P_\mu(P_\lambda u_\lambda\cdot P_{\lambda'}v_{\lambda'})\|_{X^{s,\theta}_\mu} &\lesssim \mu^{s-1}\mu^{\theta-\frac12}\sum_{d_1,d_2}\|P_\mu(P_\lambda u_{\lambda,d_1}\cdot P_{\lambda'}v_{\lambda',d_2})\|_{X^{1,\frac12}_\mu}\\
&\lesssim \mu^{s-1}\mu^{\theta-\frac12}\sum_{d_1,d_2}\|u_{\lambda,d_1}\|_{X^{1,\frac12}_{\lambda,d_1}}\|v_{\lambda',d_2}\|_{X^{1,\frac12}_{\lambda',d_2}}\\
&\lesssim \frac{\mu^{s-1}}{\lambda^{s-1}}\frac{\mu^{\theta-\frac12}}{\lambda^{s-1}}\|u_\lambda\|_{X^{s,\theta}_\lambda}\|v_{\lambda'}\|_{X^{s,\theta}_{\lambda'}} \lesssim \frac{\mu^{s-1}}{\lambda^{s-1+\varepsilon}}\|u_\lambda\|_{X^{s,\theta}_\lambda}\|v_{\lambda'}\|_{X^{s,\theta}_{\lambda'}}.
\end{aligned}$$

Finally, (6.1) also follows from (6.8), (6.9) by readjusting the weights.

6.2. **Proof of Proposition 6.2.** To be able to use the wave packet decomposition from Proposition 4.7 and Corollary 4.8 we need to be on a small interval such as $[k\delta,(k+1)\delta]$. We fix δ and without loss of generality we prove (6.2), (6.3) on $I=[0,\delta]$. We sum these bounds by brute force treating $1/\delta=D$ as a universal constant and then $\|v_\eta\|_{X^{0,\frac12}_{\eta,D}[I]}\simeq\|v_\eta\|_{X^{0,\frac12}_{\eta,1}[I]}$. Let $\eta\in\{\mu,\lambda'\}$ and we refer to $\eta=\mu$, (6.2) as Case 1, and to $\eta=\lambda'$, (6.3) as Case 2. For the remainder of this proof, we normalize the norms of the inputs as follows:

$$\|u_{\lambda,1}\|_{X^{0,\frac12}_{\lambda,1}[I]} = \|v_{\eta,1}\|_{X^{0,\frac12}_{\eta,1}[I]} = 1, \qquad \eta\in\{\mu,\lambda'\} \tag{6.10}$$

We first give an overview of the estimates needed to establish (6.2), (6.3).

Step 1. (Bilinear angular decomposition)

We may apply appropriate multipliers such that $P_\lambda u_{\lambda,1}=u_{\lambda,1}$, $P_\eta v_{\eta,1}=v_{\eta,1}$, $\eta\in\{\mu,\lambda'\}$. The terms where $\mu\simeq 1$ of Case 1 and $\lambda\simeq\lambda'\simeq\mu\simeq 1$ of Case 2 are easily treated by Hölder's inequality and the chain rule. Thus we may assume $\mu,\lambda'\gg 1$ are large enough and Corollary 4.8 is applicable, providing a decomposition

$$v_{\eta,1} = v^+ + v^- + v_R, \qquad \eta\in\{\mu,\lambda'\}.$$

The terms $u_{\lambda,1}\cdot v_R$ are estimated by (4.32) and (4.33) in Corollary 4.9. We collect

$$v^\pm = \sum_{\omega\in\Omega_{\alpha\eta}} v^{\omega,\pm}, \qquad v^{\omega,\pm} = P_\eta \sum_{T\in\mathcal{T}^\pm_\eta,\omega_T=\omega} c_T(t)u_T(t). \tag{6.11}$$

Next, by Proposition 2.19 and its corollary we decompose

$$u_{\lambda,1} = u^+ + u^-$$

Recalling [39, Theorem 3], if u and v are free waves on a Minkowski background, the modulation of the product uv depends on their relative positions on the null cone $\tau^2 = |\xi|^2$. Accordingly, we perform a bilinear angular decomposition of the products $u^{\pm_1}v^{\pm_2}$ with angular separation $\simeq \alpha$. The bilinear decomposition is nonstandard since the two factors are localized differently: the high-frequency input will be split pseudo-differentially, while for the low-frequency input we use a wave packet decomposition to leverage the characteristic energy estimates from the previous section.

Let Ω_α be a partition of the unit circle into angle α arcs and let $\alpha_\eta = \eta^{-\frac{1}{2}}$. At time $t = 0$, for any $\omega \in \Omega_{\alpha_\eta}$ we invoke the partition of unity from the Appendix - Proposition C.1:

$$1 = \sum_j \phi_{\theta_j}^{\alpha_j, k_j(\omega)}(\xi), \qquad t = 0, \qquad k_j(\omega) \in \{1, 2, 3, 4\}.$$

For each interval $\omega \in \Omega_{\alpha_\eta}$ and $(\theta, k) \in \Omega_\alpha \times [1, 4]$, define the relation $\omega \sim_\alpha (\theta, k)$ if the triple (α, θ, k) appears in the above partition of unity. Then $\omega \sim_\alpha (\theta, k)$ only if $|\theta - \omega| \sim \alpha$ or $|\theta - \omega| \lesssim \alpha_\eta$, and for each scale α there are at most $O(1)$ intervals $\theta \in \Omega_\alpha$ related to ω.

Let $\Phi_t^{\alpha_\eta, \pm}$ denote the Hamiltonian flows for the half-wave symbols $\tau + a^{\pm}_{<\alpha_\eta^{-1}}$ in the factorization $g^{\alpha\beta}_{<\sqrt{\eta}}\xi_\alpha\xi_\beta = (\tau + a^+_{<\alpha_\eta^{-1}})(\tau + a^-_{<\alpha_\eta^{-1}})$. Pulling back both sides by this flow as in (5.3) and mollifying in the x variable as in (5.4), we obtain a time-dependent partition of unity for functions localized at frequency λ

$$\begin{aligned} P_\lambda(\xi) &= \sum_j (P_{<\lambda/8}(D_x)\phi_{\theta_j}^{\alpha_j, \pm, k_j(\omega)})(t, x, \xi) P_\lambda(\xi) \\ &= \sum_j \phi_{\theta_j, \lambda}^{\alpha_j, \pm, k_j(\omega)}(t, x, \xi) \qquad (6.12) \\ &= \sum_j \tilde{\phi}_{\theta_j, \lambda}^{\alpha_j, \pm, k_j(\omega)}(t, x, \xi), \qquad (6.13) \end{aligned}$$

where $\tilde{\phi}_\theta(t, x, \xi) := \phi_\theta(t, x, -\xi)$ (recall that the multipliers s_λ are assumed radial). Note that in general $\tilde{\phi}_\theta^\pm \neq \phi_{-\theta}^\pm$.

For any signs $\pm_1, \pm_2$ let $\pm = \pm_1\pm_2$. One has

$$\begin{aligned} u^{\pm_1}v^{\pm_2} &= \sum_{\omega \in \Omega_{\alpha_\eta}} u^{\pm_1} v^{\omega, \pm_2} = \sum_\omega \sum_j (\phi_{\theta_j, \lambda}^{\alpha_j, \pm_1, k_j(\omega)} u^{\pm_1}) v^{\omega, \pm_2} \\ &= \sum_{\alpha \in [\alpha_\eta, 1]} \sum_{\theta \in \Omega_\alpha} \sum_{k=1}^4 \phi_{\theta, \lambda}^{\alpha, \pm_1, k} u^{\pm_1} \sum_{\omega \sim (\alpha, \pm\theta, k)} v^{\omega, \pm_2} \\ &= \sum_{\alpha \in [\alpha_\eta, 1]} \sum_{\theta \in \Omega_\alpha} \sum_{k=1}^4 (\phi_{\theta, \lambda}^{\alpha, \pm_1, k} u^{\pm_1}) v_{\pm\theta}^{\alpha, \pm_2, k}. \end{aligned}$$

Thus for $\alpha \not\simeq \alpha_\eta$, $v_{\pm\theta}^{\alpha,\pm_2,k}$ contains packets u_T corresponding to $T = (x_T, \omega_T) \in \mathcal{T}_\eta^{\pm_2}$ which have angular separation of α relative to $\pm\theta$: $\angle(\omega_T, \pm\theta) \simeq \alpha$:

$$v_\theta^{\pm_2,\alpha,k}(t) = P_\eta \sum_{T\in\mathcal{T}_\eta^{\pm_2},\omega_T\sim(\alpha,\pm\theta,k)} c_T(t)u_T(t). \tag{6.14}$$

Since the ω_T's are separated by $\simeq \eta^{-\frac{1}{2}}$, for every θ we have roughly $\frac{\alpha}{\eta^{-\frac{1}{2}}} = \alpha\eta^{\frac{1}{2}}$ directions ω which obey this condition.

The index k is a technical artifact of our construction and can be safely ignored. Hence we shall hereafter simply write

$$u^{\pm_1}v^{\pm_2} = \sum_{\alpha\in[\alpha_\eta,1]}\sum_{\theta\in\Omega_\alpha} (\phi_{\theta,\lambda}^{\alpha,\pm_1}u^{\pm_1})v_{\pm\theta}^{\alpha,\pm_2}. \tag{6.15}$$

A typical term $(\phi_{\theta,\lambda}^{\alpha,+}u)v_{\pm\theta}^{\alpha,\pm}$ intuitively involves waves propagating at relative angle α.

For studying nonresonant interactions of the form $P_\mu\big(P_\lambda u^\pm P_\lambda v^\pm\big)$, we need a modified decomposition using instead the partition (6.13):

$$u^{\pm_1}v^{\pm_2} = \sum_{\alpha\in[\alpha_\eta,1]}\sum_{\theta\in\Omega_\alpha} (\tilde{\phi}_{\theta,\lambda}^{\alpha,\pm_1}u^{\pm_1})v_{\theta}^{\alpha,\pm_2}. \tag{6.16}$$

Note that both terms of the form $(\phi_{\theta,\lambda}^{\alpha,+}u)v_{-\theta}^{\alpha,-}$ and $(\tilde{\phi}_{\theta,\lambda}^{\alpha,+}u)v_{\theta}^{+,\alpha}$ only involve interactions between pairs of frequencies (ξ_1,ξ_2) with $\angle(\xi_1,-\xi_2)\sim\alpha$.

Finally, we sometimes write $v_{\theta,\eta}^{\alpha,\pm_2}$ to clarify the frequency of the packets constituting $v_\theta^{\alpha,\pm_2}$.

Step 2. (Small angles interactions)

We first consider the minimal angle case consisting of the $\alpha \simeq \alpha_\eta = \eta^{-\frac{1}{2}}$ terms in (6.15) , which will mostly follow from Hölder's inequality.

Proposition 6.4. *Let $\lambda \gtrsim \eta \in \{\mu, \lambda'\}$. Under normalization* (6.10)*, on I, one has:*

$$\sum_{\theta\in\Omega_{\alpha_\eta}} \|\phi_{\theta,\lambda}^{\alpha_\eta,\pm_1}u^{\pm_1}\cdot v_\theta^{\alpha_\eta,\pm_2}\|_{L^2} \lesssim \eta^{\frac{3}{4}} \tag{6.17}$$

$$\sum_{\theta\in\Omega_{\alpha_\eta}} \|Q_{g_{<\sqrt{\lambda}}}\big(\phi_{\theta,\lambda}^{\alpha_\eta,\pm_1}u^{\pm_1}\cdot v_\theta^{\alpha_\eta,\pm_2}\big)\|_{L^2} \lesssim \lambda\eta^{\frac{3}{4}} \tag{6.18}$$

$$\sum_{\theta\in\Omega_{\alpha_\eta}} \|Q_{g_{<\sqrt{\lambda}}}\big(\phi_{\theta,\lambda}^{\alpha_\eta,\pm_1}u^{\pm_1}\cdot v_\theta^{\alpha_\eta,\pm_2}\big)\|_{L^2L^1} \lesssim \lambda \tag{6.19}$$

For any $\alpha \geq \alpha_\eta$, under (6.10)*, one has:*

$$\sum_{\theta\in\Omega_\alpha} \|\Box_{g_{<\sqrt{\lambda}}}\phi_{\theta,\lambda}^{\alpha,\pm_1}u^{\pm_1}\cdot v_\theta^{\alpha,\pm_2}\|_{L^2} + \|\phi_{\theta,\lambda}^{\alpha,\pm_1}u^{\pm_1}\cdot \Box_{g_{<\sqrt{\lambda}}}v_\theta^{\alpha,\pm_2}\|_{L^2} \lesssim \lambda\eta\alpha^{\frac{1}{2}} \tag{6.20}$$

$$\sum_{\theta\in\Omega_\alpha} \|\Box_{g_{<\sqrt{\lambda}}}\phi_{\theta,\lambda}^{\alpha,\pm_1}u^{\pm_1}\cdot v_\theta^{\alpha,\pm_2}\|_{L^2L^1} + \|\phi_{\theta,\lambda}^{\alpha,\pm_1}u^{\pm_1}\cdot \Box_{g_{<\sqrt{\lambda}}}v_\theta^{\alpha,\pm_2}\|_{L^2L^1} \lesssim \lambda \tag{6.21}$$

As a consequence we will obtain the following estimates in low modulation spaces, which take care of the terms in (6.15) with $\alpha \simeq \alpha_\eta$.

Corollary 6.5. *In Case 1, resp. Case 2, under normalization* (6.10)*, one has:*

$$\sum_{\theta\in\Omega_{\alpha_\eta}} \|\phi^{\alpha_\mu,\pm_1}_{\theta,\lambda} u^{\pm_1}\cdot v^{\alpha_\mu,\pm_2}_{\theta,\mu}\|_{X^{0,\frac14}_{\lambda',1}[I]} \lesssim \mu^{\frac34} \tag{6.22}$$

$$\sum_{\theta\in\Omega_{\alpha_\eta}} \|P_\mu\big(\phi^{\alpha_{\lambda'},\pm_1}_{\theta,\lambda} u^{\pm_1}\cdot v^{\alpha_{\lambda'},\pm_2}_{\theta,\lambda'}\big)\|_{X^{0,\frac14}_{\mu,d_0}[I]} \lesssim \frac{\lambda}{\mu^{\frac14}} \tag{6.23}$$

Step 3. (Non-resonant interactions)

We continue with the non-resonant parts of Case 2 ($\eta=\lambda'$), which include $u^+\cdot v^+$, $u^-\cdot v^-$ and the terms of $u^\pm\cdot v^\mp$ in (6.15) where $\alpha > \max(\frac{\mu}{\lambda},\alpha_\lambda)$. Here we have the following estimates, which are responsible for the $\frac{\lambda}{\mu}$ loss in the high modulation bound $X^{1,\frac14}_{\mu,\mu}$ in (6.3)[2]

Proposition 6.6. *Let $D\le\mu\ll\lambda\simeq\lambda'$ and let $\pm$ be a sign. Suppose $u^\pm=P_\lambda u^\pm$, $v=P_{\lambda'}v^\pm$. Then:*

$$\|P_\mu(u^\pm\cdot v^\pm)\|_{X^{1,\frac14}_{\mu,\mu}[I]} \lesssim \mu^{-\frac14}\|u^\pm\|_{X^{1,\frac12}_{\lambda,1}}\|v^\pm\|_{X^{1,\frac12}_{\lambda,1}}. \tag{6.24}$$

More precisely,

$$\mu^{\frac14}\|\nabla_{t,x}P_\mu(u^\pm\cdot v^\pm)\|_{L^2} \lesssim \mu^{-\frac14}\frac{\mu}{\lambda}\|u^\pm\|_{X^{1,\frac12}_{\lambda,1}}\|v^\pm\|_{X^{1,\frac12}_{\lambda,1}},$$
$$\mu^{\frac14-1}\|\Box_{g<\sqrt{\mu}}P_\mu(u^\pm\cdot v^\pm)\|_{L^2} \lesssim \mu^{-\frac14}\|u^\pm\|_{X^{1,\frac12}_{\lambda,1}}\|v^\pm\|_{X^{1,\frac12}_{\lambda,1}}.$$

For $\alpha>\max(\frac{\mu}{\lambda},\alpha_\lambda)$ one has:

$$\sum_{\theta\in\Omega_\alpha} \|P_\mu\big(\phi^{\alpha}_{\theta,\lambda}u^\pm\cdot v^{\mp,\alpha}_{-\theta}\big)\|_{X^{1,\frac14}_{\mu,\mu}[I]} \lesssim_N \mu^{-\frac14}\Big(\frac{\mu}{\lambda}\Big)\|u^\pm\|_{X^{1,\frac12}_{\lambda,1}}\|v^\mp\|_{X^{1,\frac12}_{\lambda,1}} \tag{6.25}$$

The bounds (6.24), (6.25) are used for the $X^{1,\frac14}_{\mu,\mu}$ part of (6.3). To prove the $\tilde X^{1,\frac14}_{\mu,\mu}$ part we also need $L^\infty L^2$ estimates, but these follow easily by Bernstein $P_\mu L^\infty L^1\to\mu L^\infty L^2$ and Hölder $L^\infty L^2\times L^\infty L^2\to L^\infty L^1$.

The proof of this proposition uses several technical lemmas, whose proofs are deferred to a later section. We begin with a pseudo-differential calculus estimate.

Lemma 6.7. *Let $u^\pm=P_\lambda u^\pm$, $v^\pm=P_\lambda v^\pm$, where $v=v^\pm=\sum_{\omega\in\Omega_{\alpha_\lambda}}v^{\omega,\pm}$ is a superposition of frequency λ packets. Consider the bilinear decompositions*

$$u^\pm v^\mp=\sum_\alpha\sum_\theta(\phi^{\alpha,\pm}_{\theta,\lambda}u)v^{\alpha,\mp}_{-\theta},\quad u^\pm v^\pm=\sum_\alpha\sum_\theta(\tilde\phi^{\alpha,\pm}_{\theta,\lambda}u)v^{\alpha,\pm}_{\theta}$$

as in (6.15), (6.16)*. If $\mu\ll\alpha\lambda$ and $\alpha\gg\lambda^{-\frac12}$, then there is a rapidly converging expansion*

$$P_\mu(\phi^\alpha_{\theta,\lambda}uv^\alpha_{-\theta,\lambda})=\sum_{j=1,2}\sum_{\vec k}(\alpha\lambda)^{-1}P_{\mu,\vec k}(\phi^{j,\alpha}_{\theta,\vec k}u\psi^{j,\alpha}_{-\theta,\vec k}v)+\tilde P_\mu(\phi^\alpha_{\theta,\lambda}ur^\alpha_{-\theta}),$$

[2]The loss of $\frac{\lambda}{\mu}$ in (6.3) caused by (6.24) is essentially due to our choice of spaces. Assuming constant coefficients, a ++ high-high to low $(\lambda,\lambda)\to\mu$ interaction would have output modulation λ, while we force our modulation weights to be at most equal to the frequency (μ). To compensate for this, we introduced the $\tilde X^{1,\frac14}_{\mu,\mu}$ norms which retain the expected $\frac{\mu}{\lambda}$ factor.

where $\phi^{j,\alpha}_{\theta,\vec{k}},\ \psi^{j,\alpha}_{-\theta,\vec{k}} \in \langle\vec{k}\rangle^{-N} S(1, g_{\alpha_\lambda})$ *are square-summable in* θ, $P_{\mu,\vec{k}}$ *are Fourier multipliers supported in* $|\xi| \lesssim \mu$ *with* $L^2 \to L^2$ *norm* $O(\langle k\rangle^{-N})$ *and* $r^\alpha_{-\theta} = \sum_{|\omega\mp\theta|\lesssim\alpha} \sum_T c_T r_T$ *is a superposition of packets* r_T *with* $\|r_T\|_{WP} = O((\alpha^2\lambda)^{-\infty})$ *uniformly in* T. *Here* WP *denotes any weighted norm* WP^N_T *as defined in* (4.3).

The next lemma gives a variable-coefficient version of some L^2 null form estimates considered by Foschi and Klainerman [8]. Observe that $\mu^{\frac12}$ on the right side is consistent with scaling.

Lemma 6.8. *For* $\mu \ll \lambda$ *and any pair of signs* $\pm_1, \pm_2 \in \{\pm\}$,

$$\|P_\mu Q_{g_{<\sqrt{\lambda}}}(u^{\pm_1}, v^{\pm_2})\|_{L^2} \lesssim \mu^{\frac12} \|u^{\pm_1}\|_{X^{1,\frac12}_{\lambda,1}} \|v^{\pm_2}\|_{X^{1,\frac12}_{\lambda,1}}, \tag{6.26}$$

$$\|P_\mu(u^{\pm_1} v^{\pm_2})\|_{L^2} \lesssim \mu^{\frac12} \|u^{\pm_1}\|_{X^{0,\frac12}_{\lambda,1}} \|v^{\pm_2}\|_{X^{0,\frac12}_{\lambda,1}}, \tag{6.27}$$

$$\|\partial_t P_\mu(u^{\pm_1} v^{\pm_2})\|_{L^2} \lesssim \mu^{\frac12}\lambda \|u^{\pm_1}\|_{X^{0,\frac12}_{\lambda,1}} \|v^{\pm_2}\|_{X^{0,\frac12}_{\lambda,1}}. \tag{6.28}$$

Finally, the last part of Proposition 6.6 utilizes

Lemma 6.9. *For* $\alpha > \max(\frac{\mu}{\lambda}, \lambda^{-\frac12})$ *in the bilinear decomposition,*

$$\sum_\theta \mu^{1+\frac12} \|P_\mu(\phi^{\alpha,\mp}_{\theta,\lambda} u v^{\alpha,\pm}_{-\theta})\|_{L^2} \lesssim_N [(\alpha\lambda)^{-1} + (\alpha^2\lambda)^{-N}] \Big(\frac{\mu}{\lambda}\Big)^2 \|u\|_{X^{1,\frac12}_{\lambda,1}} \|v\|_{X^{1,\frac12}_{\lambda,1}},$$

$$\sum_\theta \mu^{\frac12-1} \|P_\mu Q(\phi^{\alpha,\mp}_{\theta,\lambda} u, v^{\alpha,\pm}_{-\theta})\|_{L^2} \lesssim_N [\alpha^{\frac12}\mu^{-\frac12} + (\alpha^2\lambda)^{-N}] \Big(\frac{\lambda^{\frac12}}{\mu}\Big) \frac{\mu}{\lambda} \|u\|_{X^{1,\frac12}_{\lambda,1}} \|v\|_{X^{1,\frac12}_{\lambda,1}}.$$

Now we show how these claims imply Proposition 6.6. Consider (6.24). The preceding lemmas imply the L^2 bound in (6.24). The $\Box$ part of the norm is bounded by

$$\mu^{\frac14-1}\big(\|P_\mu \Box_{g_{<\sqrt{\lambda}}}(u^+_\lambda, v^+_\lambda)\|_{L^2} + \|[P_\mu, \Box_{g_{<\sqrt{\lambda}}}](u^+_\lambda v^+_\lambda)\|_{L^2} + \|(\Box_{g_{<\sqrt{\mu}}} - \Box_{g_{<\sqrt{\lambda}}}) P_\mu(u^+_\lambda v^+_\lambda)\|_{L^2}\big).$$

For the last term we use the estimate (1.29), Bernstein, and the energy estimate to infer

$$\begin{aligned}
\|(\Box_{g_{<\sqrt{\mu}}} - \Box_{g_{<\sqrt{\lambda}}}) P_\mu(u^+_\lambda v^+_\lambda)\|_{L^2} &\lesssim \|P_\mu \partial_t(u^+_\lambda v^+_\lambda)\|_{L^\infty L^2} + \mu\|P_\mu(u^+_\lambda v^+_\lambda)\|_{L^\infty L^2} \\
&\lesssim \mu\|\nabla_{t,x} u\|_{L^\infty L^2}\|v\|_{L^\infty L^2} + \mu\|u\|_{L^\infty L^2}\|\nabla_{t,x} v\|_{L^\infty L^2} \\
&\lesssim \frac{\mu}{\lambda}\|u^+_\lambda\|_{X^{1,\frac12}_{\lambda,1}}\|v^+_\lambda\|_{X^{1,\frac12}_{\lambda,1}}.
\end{aligned}$$

Writing the commutator as

$$[P_\mu, \Box_{g_{<\sqrt{\lambda}}}] = \big([P_\mu, g_{<\mu}] P_{<8\mu} + \sum_{\nu\in[\mu,\sqrt{\lambda}]} [P_\mu, g_\nu]\tilde{P}_\nu\big)(\partial_x^2 + \partial_x\partial_t),$$

one deduces as above that

$$\|[P_\mu, \Box_{g_{<\sqrt{\lambda}}}](u^+_\lambda v^+_\lambda)\|_{L^2} \lesssim \frac{\mu}{\lambda}\|u^+_\lambda\|_{X^{1,\frac12}_{\lambda,1}}\|v^+_\lambda\|_{X^{1,\frac12}_{\lambda,1}}.$$

For the first term, write

$$\begin{aligned}
\|P_\mu \Box_{g_{<\sqrt{\lambda}}}(u^+_\lambda v^+_\lambda)\|_{L^2} \lesssim \|P_\mu(\Box_{g_{<\sqrt{\lambda}}} u^+_\lambda v^+_\lambda)\|_{L^2} + \|P_\mu(u^+_\lambda \Box_{g_{<\sqrt{\lambda}}} v^+_\lambda)\|_{L^2} \\
+ \|P_\mu Q_{<\sqrt{\lambda}}(u^+_\lambda, v^+_\lambda)\|_{L^2}.
\end{aligned}$$

By Bernstein, Hölder, and energy estimates, the first two terms are bounded by

$$\mu\|\Box_{g_{<\sqrt{\lambda}}}u_\lambda^+ v_\lambda^+\|_{L^2L^1} + \mu\|u_\lambda^+\Box_{g_{<\sqrt{\lambda}}}v_\lambda^+\|_{L^2L^1} \lesssim \frac{\mu}{\lambda}\|u_\lambda^+\|_{X^{1,\frac{1}{2}}_{\lambda,1}}\|v_\lambda^+\|_{X^{1,\frac{1}{2}}_{\lambda,1}}.$$

The main null form term is estimated using Lemma 6.8. Collecting all the estimates we obtain the $\Box_{g_{<\sqrt{\mu}}}$ part of (6.24).

The estimate (6.25) is proved similarly. For the L^2 part we appeal to Lemma 6.9, while for the $\Box_{g_{<\sqrt{\mu}}}$ part we argue as above to first bound

$$\sum_\theta \mu^{\frac{1}{2}-1}\|(\Box_{g_{<\sqrt{\mu}}}P_\mu - P_\mu\Box_{g_{<\sqrt{\lambda}}})(\phi^\alpha_{\theta,\lambda}uv^\alpha_{-\theta,\lambda})\|_{L^2} \lesssim \mu^{-\frac{1}{2}}[\frac{\mu}{\lambda}+\lambda^{-\frac{3}{2}}]\|u\|_{X^{1,\frac{1}{2}}_{\lambda,1}}\|v\|_{X^{1,\frac{1}{2}}_{\lambda,1}},$$

and estimate $\|P_\mu\Box_{g_{<\sqrt{\lambda}}}(\phi^\alpha_{\theta,\lambda}uv^\alpha_{-\theta,\lambda})\|_{L^2}$ as before. Modulo the preceding lemmas, the proof of Proposition 6.6 is complete.

Step 4. (Almost resonant interactions)

We have arrived at the key part of the argument, involving (6.15) for any signs $\pm_1, \pm_2$ in Case 1, respectively the signs $\pm_1\pm_2 = -$ in Case 2, with the sum restricted to $\alpha_\lambda = \lambda^{-\frac{1}{2}} < \alpha \leq \frac{\mu}{\lambda}$. Thus, in Case 2 we may assume $\sqrt{\lambda} < \mu$. The remaining parts of Case 2 are non-resonant and were treated in Step 3.

To show that $u^{\pm_1}\cdot v^{\pm_2}$ lies in $X^{0,\frac{1}{4}}_{\lambda',\leq\mu,\infty}[I]$, respectively $X^{0,\frac{1}{4}}_{\mu,[d_0,\mu],\infty}[I]$ it suffices to use the decomposition (6.15) and for each α to define an appropriate modulation d such that the α-term is in $X^{0,\frac{1}{4}}_{\lambda',d}[I]$, resp. $X^{0,\frac{1}{4}}_{\mu,d}[I]$, which proceeds as follows:

Case 1. ($\eta = \mu$). Define[3] $d = \mu\alpha^2$. When α ranges from α_μ to 1, d ranges from D to μ and it suffices to prove, under (6.10):

$$\sum_{\theta\in\Omega_\alpha}\|\phi^{\alpha,\pm_1}_{\theta,\lambda}u^{\pm_1}\cdot v^{\alpha,\pm_2}_{\theta,\mu}\|_{X^{0,\frac{1}{4}}_{\lambda',\mu\alpha^2}[I]} \lesssim \mu^{\frac{3}{4}} \tag{6.29}$$

Case 2. ($\eta = \lambda' \simeq \lambda$). Define[3] $d = \frac{\lambda^2\alpha^2}{\mu}$. When α ranges from α_λ to $\frac{\mu}{\lambda}$, d ranges from $d_0 = \frac{\lambda}{\mu}$ to μ and it suffices to prove, under (6.10):

$$\sum_{\theta\in\Omega_\alpha}\|P_\mu\big(\phi^{\alpha,\pm_1}_{\theta,\lambda}u^{\pm_1}\cdot v^{\alpha,\pm_2}_{\theta,\lambda'}\big)\|_{X^{0,\frac{1}{4}}_{\mu,\frac{\lambda^2\alpha^2}{\mu}}[I]} \lesssim \frac{\lambda}{\mu^{\frac{1}{4}}} \tag{6.30}$$

Both cases are based on the following proposition, which incorporates the characteristic energy estimate from Section 5.2 and the uniform bounds on wave packets.

[3] The choice of d is motivated by the Fourier analysis of the constant coefficients case, see [39]

Proposition 6.10. *Let $\alpha > \alpha_\eta$. In decomposition* (6.15)*, under* (6.10) *we have*[4]

$$\sum_{\theta\in\Omega_\alpha} \|\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1}\cdot v^{\alpha,\pm_2}_{\theta}\|_{L^2[I]} \lesssim \frac{\eta^{\frac12}}{\alpha^{\frac12}} \tag{6.31}$$

$$\sum_{\theta\in\Omega_\alpha} \|Q_{g_{<\sqrt\lambda}}\big(\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1}\cdot v^{\alpha,\pm_2}_{\theta}\big)\|_{L^2[I]} \lesssim (\lambda\eta\alpha^2)\frac{\eta^{\frac12}}{\alpha^{\frac12}}. \tag{6.32}$$

The L^2 part of both (6.29) and (6.30) follows immediately from (6.31), by discarding the multiplier P_μ if needed. For the $\Box_{g_{<\sqrt\mu}}$ part of (6.30) we write

$$\Box_{g_{<\sqrt\mu}} P_\mu = P_\mu \Box_{g_{<\sqrt\lambda}} + \big(\Box_{g_{<\sqrt\mu}} - \Box_{g_{<\sqrt\lambda}}\big)P_\mu + [\Box_{g_{<\sqrt\lambda}}, P_\mu]. \tag{6.33}$$

For the first term, we discard P_μ and write

$$\begin{aligned}\Box_{g_{<\sqrt\lambda}}\big(\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1}\cdot v^{\alpha,\pm_2}_{\theta}\big) = 2\, Q_{g_{<\sqrt\lambda}}\big(\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1}\cdot v^{\alpha,\pm_2}_{\theta}\big) + \\ \Box_{g_{<\sqrt\lambda}}\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1}\cdot v^{\alpha,\pm_2}_{\theta} + \phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1}\cdot \Box_{g_{<\sqrt\lambda}} v^{\alpha,\pm_2}_{\theta}\end{aligned} \tag{6.34}$$

In both (6.29) and (6.30), the bound for the $Q_{g_{<\sqrt\lambda}}$ term follows from (6.32), while the bounds for the other terms follow from (6.20).

For the other two terms in (6.33), we note first of all that as

$$\mu^{-1}d^{\frac14-1}\|(\Box_{g_{<\sqrt\mu}} - \Box_{g_{<\sqrt\lambda}})P_\mu w\|_{L^2} \lesssim \mu^{-1}d^{\frac14-1}\|\nabla_{t,x}P_\mu w\|_{L^\infty L^2} \lesssim d^{-\frac12}\|P_\mu w\|_{X^{0,\frac14}_{\mu,\frac{\lambda^2\alpha^2}{\mu}}[I]},$$

the second term may be absorbed in the left side of (6.30) so long as $\mu \ll \lambda$, which guarantees that $d \gg 1$.

Also, since $\mu > \sqrt\lambda$ we may estimate

$$\|[\Box_{g_{<\sqrt\lambda}}, P_\mu]w\|_{L^2} \lesssim \mu\|P_{[\mu/2,2\mu]}w\|_{L^2} + \|\partial_t P_{[\mu/2,2\mu]}w\|_{L^2}, \tag{6.35}$$

and again use (2.1) to obtain

$$\begin{aligned}&\sum_{\theta\in\Omega_\alpha} \|P_\mu(\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1}\cdot v^{\alpha,\pm_2}_{\theta,\lambda'})\|_{X^{0,\frac14}_{\mu,\frac{\lambda^2\alpha^2}{\mu}}[I]} \\ &\lesssim \frac{\lambda}{\mu^{\frac14}} + \sum_{\theta\in\Omega_\alpha} \mu^{-1}d^{\frac14-1}\|\partial_t P_{[\mu/2,2\mu]}(\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1}\cdot v^{\alpha,\pm_2}_{\theta,\lambda'})\|_{L^2} \\ &\lesssim \frac{\lambda}{\mu^{\frac14}} + \sum_{\theta\in\Omega_\alpha} d^{-1}\|P_{[\mu/2,2\mu]}(\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1}\cdot v^{\alpha,\pm_2}_{\theta,\lambda'})\|\|_{X^{0,\frac14}_{\mu,\frac{\lambda^2\alpha^2}{\mu}}[I]}.\end{aligned} \tag{6.36}$$

The summation on the right side is handled by another perturbative argument. For fixed $\alpha > \lambda^{-\frac12}$ and a small constant $c > 0$, let M be the best constant in (6.30) which is uniform in $\mu \le c\lambda$. Invoking Proposition 6.6 and (6.29) for $\mu < \alpha\lambda$ and $\mu > c\lambda$, respectively, we have $M \lesssim 1 + d_{c\lambda}^{-1}M$, where $d_\mu := \lambda^2\alpha^2/\mu$, and the right term may be absorbed into the left side if c is sufficiently small.

[4] The estimate (6.32) shows that the effect of the null form $Q_{g_{<\sqrt\lambda}}$ is $\lambda\eta\alpha^2$ which is a familiar factor from the constant coefficients case, given the angular localization.

In what follows we complete the proof of Proposition 6.2 by establishing Proposition 6.4, Corollary 6.5, Proposition 6.10, Lemma 6.7 and Lemmas 6.8, 6.9.

6.2.1. **Proof of Proposition 6.10.** We successively consider the two estimates in the proposition.

(1) Using (5.10) and (4.28) the estimate (6.31) reduces to proving

$$\|\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1} \cdot v^{\alpha,\pm_2}_{\theta}\|_{L^2} \lesssim \frac{\eta^{\frac{1}{2}}}{\alpha^{\frac{1}{2}}} \|\chi^0_\theta u^{\pm_1}\|_{L^2} \big(\sum_{T, \omega_T \sim (\alpha, \pm\theta)} |c_T|^2_{L^\infty_t} \big)^{\frac{1}{2}} \tag{6.37}$$

for any $\theta \in \Omega_\alpha$. Recall

$$v^{\alpha,\pm_2}_{\theta} = \sum_{\omega, \omega \sim (\alpha, \pm\theta)} v^{\omega,\pm_2} \tag{6.38}$$

where $v^{\omega,\pm_2}$ is given by (6.11). Since there are $\simeq \alpha\eta^{\frac{1}{2}}$ directions ω in the (6.38) sum, (6.37) follows from summing the following with Cauchy-Schwarz in ω:

$$\|\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1} \cdot v^{\omega,\pm_2}\|_{L^2} \lesssim \frac{\eta^{\frac{1}{4}}}{\alpha} \|\chi^0_\theta u^{\pm_1}\|_{L^2} \big(\sum_{T, \omega_T = \omega} |c_T|^2_{L^\infty_t} \big)^{\frac{1}{2}} \tag{6.39}$$

Since for any ω we have

$$I \times \mathbb{R}^2 = \bigcup_{T, x_T \in \Xi^\omega_\eta} T$$

and the T's are finitely overlapping, we obtain (6.39) if we have

$$\|\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1} \cdot P_\eta c_T u_T\|_{L^2(T')} \lesssim \frac{1}{\langle d(T,T')\rangle^N} \frac{\eta^{\frac{1}{4}}}{\alpha} \|\chi^0_\theta u^{\pm_1}\|_{L^2} \, |c_T|_{L^\infty_t} \tag{6.40}$$

for any T, T' with $\omega_T = \omega_{T'} = \omega$. This follows from (4.10) and Corollary 5.11.

(2) The proof of (6.32) proceeds similarly by reducing to

$$\|Q_{g_{<\sqrt{\lambda}}}\big(\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1} \cdot v^{\omega,\pm_2}\big)\|_{L^2} \lesssim (\lambda\eta\alpha^2) \frac{\eta^{\frac{1}{4}}}{\alpha} \|\chi^1_\theta u^{\pm_1}\| \big(\sum_{T, \omega_T = \omega} |c_T|^2_{L^\infty_t} + |c'_T|^2_{L^2_t} \big)^{\frac{1}{2}} \tag{6.41}$$

for every ω, based on (4.28), (4.29), (5.10) with $j = 1$. Associated to $g_{\sqrt{\lambda}}$ and to ω we consider vector fields $L, \underline{L}, E$ which form a null frame as in section 3.2. Then we can express the null form as

$$2Q_{g_{<\sqrt{\lambda}}}\big(u, v\big) = Lu \cdot \underline{L}v + \underline{L}u \cdot Lv - 2Eu \cdot Ev.$$

For the term $L\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1} \cdot \underline{L} v^{\omega,\pm_2}$ we proceed as before, reducing to

$$\begin{aligned} \|L\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1} \cdot \underline{L} P_\eta c_T u_T\|_{L^2(T')} &\lesssim \|L\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1}\|_{L^2(T')} \|\underline{L} P_\eta \big(c_T u_T\big)\|_{L^\infty(T')} \\ &\lesssim \lambda\eta^{\frac{5}{4}}\alpha \frac{1}{\langle d(T,T')\rangle^N} \|\chi^1_\theta u^{\pm_1}\|_{L^2} \, |c_T|_{L^\infty_t} \end{aligned}$$

which holds due to (4.31) and Corollary 5.11.

The terms

$$\underline{L}\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1} \cdot Lv^{\omega,\pm_2} \qquad \text{and} \qquad E\phi^{\alpha,\pm_1}_{\theta,\lambda} u^{\pm_1} \cdot Ev^{\omega,\pm_2}$$

are easier, as here we obtain the corresponding part of (6.41) by Hölder using (4.37), (4.36), (5.17), (5.19) and using the fact that $\alpha \geq \eta^{-\frac{1}{2}}$.

6.2.2. **Proof of Proposition 6.4.** For (6.17) - (6.19) we may assume without loss of generality that $v_\theta^{\alpha_\eta,\pm_2} = v^{\omega,\pm_2}$.

(1) For (6.17) we use Hölder's inequality

$$\|\phi_{\theta,\lambda}^{\alpha_\eta,\pm_1} u^{\pm_1} \cdot v^{\omega,\pm_2}\|_{L^2} \lesssim \|\phi_{\theta,\lambda}^{\alpha_\eta,\pm_1} u^{\pm_1}\|_{L^2} \|v^{\omega,\pm_2}\|_{L^\infty}$$

Square summing this with (4.35) and Prop. 5.6 we obtain (6.17).

(2) By square summing and Cauchy-Schwarz we reduce (6.18) to

$$\|Q_{g_{<\sqrt{\lambda}}}\big(\phi_{\theta,\lambda}^{\alpha_\eta,\pm_1} u^{\pm_1} \cdot v^{\omega,\pm_2}\big)\|_{L^2} \lesssim \lambda\eta^{\frac{3}{4}} \|\chi_\theta^{\alpha_\eta} u^{\pm_1}\|_{L^2} \big(\sum_{T,\omega_T=\omega} |c_T|_{L_t^\infty}^2 + |c_T'|_{L_t^2}^2 \big)^{\frac{1}{2}}$$

for some square-summable $\chi_\theta^{\alpha_\eta}$. Associated to $g_{\sqrt{\lambda}}$ and to ω we consider vector fields $L, \underline{L}, E$ which form a null frame as in section 3.2. Then we express the null form as

$$2Q_{g_{<\sqrt{\lambda}}}\big(u,v\big) = Lu \cdot \underline{L}v + \underline{L}u \cdot Lv - 2Eu \cdot Ev$$

We use Hölder's inequality: $L^2L^2 \times L^\infty L^\infty \to L^2L^2$ and (5.18), (4.38) for $Lu\underline{L}v$; $L^\infty L^2 \times L^2L^\infty \to L^2L^2$ with (5.19), (4.37) for the second term; and $L^2L^2 \times L^\infty L^\infty \to L^2L^2$ with (5.17), (4.36) for the third term.

(3) For (6.19) we use Hölder $L^\infty L^2 \times L^2L^2 \to L^2L^1$ together with the null frame $L, \underline{L}, E$ associated to $g_{\sqrt{\lambda}}$ and to θ, using (5.17), (5.19), (5.18) and

$$\begin{aligned}
\|Ev^{\omega,\pm_2}\|_{L^\infty L^2[I]}^2 &\lesssim \sum_{T,\omega_T=\omega} \eta\, |c_T|_{L_t^\infty}^2 \\
\|Lv^{\omega,\pm_2}\|_{L^2[I]}^2 &\lesssim \sum_{T,\omega_T=\omega} |c_T|_{L_t^\infty}^2 + |c_T'|_{L_t^2}^2 \\
\|\underline{L}v^{\omega,\pm_2}\|_{L^\infty L^2[I]}^2 &\lesssim \sum_{T,\omega_T=\omega} \eta^2\, |c_T|_{L_t^\infty}^2 \,.
\end{aligned}$$

(4) The first part of (6.20) follows from $L^2 \times L^\infty \to L^2$ with Prop. 5.12, (4.28) and

$$\|v_\theta^{\alpha,\pm_2}\|_{L^\infty} \lesssim \eta\alpha^{\frac{1}{2}} \big(\sum_{T,\omega_T\sim(\alpha,\pm\theta)} |c_T|_{L_t^\infty}^2 \big)^{\frac{1}{2}}. \tag{6.42}$$

The second part of (6.20) follows from $L^\infty L^2 \times L^2L^\infty \to L^2L^2$ based on (5.15) and

$$\Big(\sum_{\theta\in\Omega_\alpha} \|\Box_{g_{<\sqrt{\lambda}}} v_\theta^{\alpha,\pm_2}\|_{L^2L^\infty}^2 \Big)^{\frac{1}{2}} \lesssim \eta^2\alpha^{\frac{1}{2}}. \tag{6.43}$$

This estimate is obtained by decomposing as in (6.38), using Cauchy-Schwarz in ω, using (4.28)-(4.30) and

$$\|\Box_{g_{<\sqrt{\lambda}}} v^{\omega,\pm_2}\|_{L^2L^\infty} \lesssim \eta^{\frac{7}{4}} \Big(\sum_{T,\omega_T=\omega} |c_T|^2_{L^\infty_t} + |c_T'|^2_{L^2_t} + \eta^{-2} |c_T''|^2_{L^2_t} \Big)^{\frac{1}{2}}$$

(5) Finally, the proof of (6.21) is similar to the proof of (6.20) in **(4)**, except that we use $L^2L^2 \times L^\infty L^2 \to L^2L^1$ for the first part and $L^\infty L^2 \times L^2L^2 \to L^2L^1$ for the second part. Here (6.42) is replaced by

$$\|v_\theta^{\alpha,\pm_2}\|_{L^\infty L^2} \lesssim \Big(\sum_{T,\omega_T \sim (\alpha,\pm\theta)} |c_T|^2_{L^\infty_t} \Big)^{\frac{1}{2}}$$

while (6.43) is replaced by

$$\Big(\sum_{\theta\in\Omega_\alpha} \|\Box_{g_{<\sqrt{\lambda}}} v_\theta^{\alpha,\pm_2}\|^2_{L^2L^2} \Big)^{\frac{1}{2}} \lesssim \eta.$$

6.2.3. **Proof of Corollary 6.5.** The estimate (6.22) follows immediately from (6.17), (6.34), (6.18) and (6.20) for $\alpha = \alpha_\mu$. For (6.23) we consider two cases:

(1) $\lambda^{\frac{1}{2}} \le \mu$ and $d_0 = \frac{\lambda}{\mu}$

(2) $\mu \le \lambda^{\frac{1}{2}}$ and $d_0 = \mu$.

The L^2 part of (6.23) also follows from (6.17) in both cases (in Case (2) we use $\mu^{\frac{1}{2}} \le \lambda^{\frac{1}{4}}$). For the $\Box_{g_{<\sqrt{\mu}}} P_\mu$ part of Case (1), since $\lambda^{\frac{1}{2}} \le \mu$ we may freely replace $\Box_{g_{<\sqrt{\mu}}} P_\mu$ by $\Box_{g_{<\sqrt{\lambda}}} P_\mu$. We estimate the $P_\mu \Box_{g_{<\sqrt{\lambda}}}$ by (6.34), (6.18) and (6.20) for $\alpha = \alpha_{\lambda'}$. Then we treat $[P_\mu, \Box_{g_{<\sqrt{\lambda}}}]$ by the argument used to prove (6.35), (6.36).

For the $\Box_{g_{<\sqrt{\mu}}} P_\mu$ part of Case (2) we denote $w = \phi_{\theta,\lambda}^{\alpha_{\lambda'},\pm_1} u^{\pm_1} \cdot v_{\theta,\lambda'}^{\alpha_{\lambda'},\pm_2}$ and write

$$\Box_{g_{<\sqrt{\mu}}} P_\mu w = P_\mu \Box_{g_{<\sqrt{\lambda}}} w - \sum_{\nu\in[\sqrt{\mu},\sqrt{\lambda}]} P_\mu\big(g_\nu \cdot \partial_x \partial_{t,x} P_{\lesssim \max(\nu,\mu)} w\big) + [\Box_{g_{<\sqrt{\mu}}}, P_\mu] w$$

We will use Bernstein's inequality $P_\mu : L^2L^1 \to \mu L^2L^2$. For the first term we invoke the L^2L^1 estimates (6.19) and (6.21) for $\alpha = \alpha_{\lambda'}$.

For the second term we discard the P_μ and write

$$\begin{aligned} &\|g_\nu \cdot \partial_x \partial_{t,x} P_{\lesssim \max(\nu,\mu)} w\|_{L^2L^1} \lesssim \frac{1}{\nu^2} \max(\nu,\mu) \|\partial^2 g_\nu\|_{L^2L^\infty} \times \\ &\times \Big(\|\partial_{t,x} \phi_{\theta,\lambda}^{\alpha_{\lambda'},\pm_1} u^{\pm_1}\|_{L^\infty L^2} \|v_{\theta,\lambda'}^{\alpha_{\lambda'},\pm_2}\|_{L^\infty L^2} + \|\phi_{\theta,\lambda}^{\alpha_{\lambda'},\pm_1} u^{\pm_1}\|_{L^\infty L^2} \|\partial_{t,x} v_{\theta,\lambda'}^{\alpha_{\lambda'},\pm_2}\|_{L^\infty L^2} \Big). \end{aligned}$$

which is more than enough. The third term is estimated similarly since it is of the form $\approx \mu^{-1} \nabla g \partial_x \partial_{t,x}$.

6.2.4. **Proof of Lemma 6.7.** We remark first of all that by standard pseudo-differential calculus arguments (see for example [12]), if $A(x,\xi) \in S(1, g_\alpha)$ with $\alpha \ge \alpha_\lambda := \lambda^{-1/2}$ and v_T is a frequency λ packet, then so is $A(X,D)v_T$.

From this one infers

Lemma 6.11. *Let v_T be a frequency λ packet with initial direction ω. Suppose $|\theta-\omega| \le \alpha/4$ and $\phi_\theta(t,x,\xi) = \chi\big(\frac{|\widehat{\xi}-\widehat{\xi_\theta}(t,x)|}{\alpha}\big)P_\lambda(\xi)$ where χ is smooth and supported in $[-1,1]$. Then*

$$\|(1-\phi_\theta(t,X,D))v_T\|_{WP} = O((\alpha^2\lambda)^{-\infty}).$$

Proof. Put $\psi = 1-\phi$. Without loss of generality assume v_T is centered at $x=0$. Let $\chi(x)$ be a bump function supported in ball $|x| \le \alpha/16$. Then $\chi\psi$ only admits input frequencies outside the sector $\angle(\xi,\omega) \lesssim \alpha$, so $\|\chi\psi(t,X,D)v_T\|_{WP} = O((\alpha^2\lambda)^{-\infty})$. On the other hand, the spatial decay of ψv_T by our initial remark implies that $\|(1-\chi)\psi(t,X,D)v_T\|_{WP} = O((\alpha^2\lambda)^{-\infty})$ as well. □

Using Lemma 6.11 we may write

$$v^\alpha_{-\theta,\lambda} = \psi^\alpha_{-\theta} v^\alpha_{-\theta,\lambda} + r^\alpha_\theta,$$

where $\psi^\alpha_{-\theta} = \chi\big(\frac{|\widehat{\xi}+\widehat{\xi_\theta}(t,x)|}{\alpha}\big)s_\lambda(\xi)$ for smooth cutoff function χ, and where $r^\alpha_\theta = \sum_T a_T r_T$ is a sum of packets with $\|r_T\|_{WP} = O((\alpha^2\lambda)^{-\infty})$.

It remains to decompose the psuedo-differential term $P_\mu(\phi^\alpha_{\theta,\lambda}u \cdot \psi^\alpha_{-\theta}v^\alpha_{-\theta,\lambda})$. We write $v := v^\alpha_{-\theta,\lambda}$ and omit the dependence on α, λ, and t in the other notations. For any $w \in L^2$,

$$\begin{aligned}
\langle P_\mu(\phi_\theta u \psi_{-\theta} v), w\rangle &= \int e^{i\langle x,\xi+\eta+\zeta\rangle}\phi_\theta(x,\xi)\psi_{-\theta}(x,\eta)s_\mu(\zeta)\widehat{u}(\xi)\widehat{v}(\eta)\widehat{w}(\zeta)\,d\xi d\eta d\zeta dx \\
&= i\int e^{i\langle x,\xi+\eta+\zeta\rangle}\langle \partial_x, F(\xi,\eta,\zeta)\rangle\phi_\theta(x,\xi)\psi_{-\theta}(x,\eta)s_\mu(\zeta) \\
&\times \widehat{u}(\xi)\widehat{v}(\eta)\widehat{w}(\zeta)\,d\xi d\eta d\zeta dx,
\end{aligned}$$

where $F = \frac{\xi+\eta+\zeta}{|\xi+\eta+\zeta|^2}$. As $\mu \ll \alpha\lambda$, the denominator of F is bounded below $|\xi+\eta+\zeta| \gtrsim \alpha\lambda$ on the support of the integrand.

Introducing slightly wider cutoffs $\tilde{\phi}_\theta$, $\tilde{\psi}_{-\theta}$ so that $\tilde{\phi}_{\pm\theta}\phi_{\pm\theta} = \phi_{\pm\theta}$, we separate variables, for instance using Fourier series expansion

$$\tilde{\phi}_\theta\tilde{\psi}_{-\theta}s_{<2\mu}F = (\alpha\lambda)^{-1}\sum_{\vec{k}} a_{\vec{k}}(x)e_{1,k_1}(\xi)e_{2,k_2}(\eta)e_{3,k_3}(\zeta),$$

where e_{1,k_1} and e_{3,k_3} are Fourier characters adapted to $\lambda \times (\alpha\lambda)$ rectangles and e_{2,k_2} are adapted to $\mu\times\mu$ rectangles. By the derivative bounds (5.1) with $\alpha = \sqrt{\lambda}$, the coefficients $a_{\vec{k}} = (\alpha\lambda)\langle\tilde{\phi}_\theta\tilde{\psi}_{-\theta}s_{<2\mu}F, e_{\vec{k}}\rangle$ satisfy

$$\partial_x^j a_{\vec{k}} = O(\lambda^{\frac{j-1}{2}}|\vec{k}|^{-N}),\ j \ge 1. \tag{6.44}$$

Consequently, the integral takes the form

$$\begin{aligned}
&(\alpha\lambda)^{-1}\sum_{\vec{k}}\int e^{i\langle x,\xi+\eta+\zeta\rangle}a_{\vec{k}}(x)\big[(\partial_x\phi_\theta)(x,\xi)\psi_{-\theta}(x,\eta) + \phi_\theta(x,\xi)(\partial_x\psi)(x,\eta)\big] \\
&\times s_\mu(\zeta)e_{1,k_1}(\xi)e_{2,k_2}(\eta)e_{3,k_3}(\zeta)\widehat{u}(\xi)\widehat{v}(\eta)\widehat{w}(\zeta)\,d\xi d\eta d\zeta dx \\
&= \Big\langle \sum_{j=1,2}\sum_{\vec{k}}(\alpha\lambda)^{-1}\phi^j_{\theta,\vec{k}}u\psi^j_{-\theta,\vec{k}}v,\ P_{\mu,\vec{k}}(D)w\Big\rangle,
\end{aligned}$$

where

$$\phi^1_{\theta,\vec{k}}(x,\xi) = |k|^{2N/3} a_{\vec{k}}(x)(\partial_x \phi_\theta) e_{1,k_1}(\xi), \quad \phi^2_{\theta,\vec{k}}(x,\xi) = |k|^{2N/3} a_{\vec{k}}(x) \phi_\theta e_{1,k_1}(\xi)$$
$$\psi^1_{-\theta,\vec{k}}(x,\eta) = |k|^{-N/3} \psi_{-\theta}(x,\eta) e_{2,k_2}(\eta), \quad \psi^2_{-\theta,\vec{k}}(x,\eta) = |k|^{-N/3} (\partial_x \psi_{-\theta})(x,\eta) e_{2,k_2}(\eta)$$
$$P_{\mu,\vec{k}}(\zeta) = |k|^{-N/3} s_\mu(\zeta) e_{3,k_3}(\zeta)$$

The ∂_x is harmless since $\phi_\theta,\ \psi_{-\theta}$ are Lipschitz in x, so we have a rapidly converging series expansion

$$P_\mu(\phi_\theta u \psi_{-\theta} v) = \sum_{j=1,2} \sum_{\vec{k}} (\alpha\lambda)^{-1} P_{\mu,\vec{k}}(D)(\phi^j_{\theta,\vec{k}} u \psi^j_{-\theta,\vec{k}} v)$$

where $\phi^j_{\theta,\vec{k}},\ \psi^j_{-\theta,\vec{k}} \in \langle \vec{k} \rangle^{-N} S(1, g_{\alpha_\lambda})$ and retain the angular localization of $\phi_\theta,\ \psi_{-\theta}$. Although these are not quite localized in output frequency, the bounds (6.44) imply rapid decay from frequency λ on the $\sqrt{\lambda}$ scale. By Lemma B.5 these tails are negligible for square summability in θ.

6.2.5. **Proof of Lemmas 6.8 and 6.9.** Both parts can be proved in parallel. Without loss of generality we consider just the $++$ and $+-$ interactions. Begin with the decompositions

$$u^+ v^- = \sum_{\lambda^{-1/2} \le \alpha \le 1} \sum_\theta (\phi^\alpha_{\theta,\lambda} u) v^\alpha_{-\theta,\lambda},\ u^+ v^+ = \sum_\alpha \sum_\theta (\tilde{\phi}^{\alpha,\pm}_{\theta,\lambda} u) v^{\alpha,\pm}_{\theta,\lambda}$$

where $v^\alpha_{\pm\theta,\lambda} = \sum_{|\omega \mp \theta| \sim \alpha} v^\omega$ is a sum of frequency λ packets. The arguments involved depend on the relation between α, μ/λ, $\lambda^{-1/2}$, and for each of the following cases we discuss the null form estimates (Cases 1a, 2a, 3a, 4) and the L^2 estimates (Cases 1b, 2b, 3b, 4) for both $++$ and $+-$ interactions. Note that Lemma 6.9 is covered by Cases 2 and 4 below. Denote $Q := Q_{g_{<\sqrt{\lambda}}}$.

Case 1: $\lambda^{-1/2} \le \alpha \le \mu/\lambda$.

Case 1a$++$: We can use Cauchy-Schwartz in ω to obtain

$$\|P_\mu Q(\tilde{\phi}^\alpha_{\theta,\lambda} u, v^\alpha_{\theta,\lambda})\|_{L^2} \lesssim (\alpha^2\lambda)^{\frac{1}{4}} \Big(\sum_{|\omega - \theta| \sim \alpha} \|P_\mu Q(\tilde{\phi}^\alpha_{\theta,\lambda} u, v^\omega)\|^2_{L^2} \Big)^{1/2}$$

For each ω we expand Q using the null frame adapted to the direction ω and discard the mollifier P_μ. For the term $\underline{L}\tilde{\phi}^\alpha_{\theta,\lambda} u L v^\omega$, Hölder, energy estimates, and the wave packet bounds (4.37) imply

$$\|\underline{L}\tilde{\phi}^\alpha_{\theta,\lambda} u L v^\omega\|_{L^2} \lesssim \|\underline{L}\tilde{\phi}^\alpha_{\theta,\lambda} u\|_{L^\infty L^2} \|L v^\omega\|_{L^2 L^\infty} \lesssim \lambda^{-\frac{1}{4}} \|\tilde{\phi}^\alpha_{\theta,\lambda} u\|_{X^{1,\frac{1}{2}}_{\lambda,1}} \|v^\omega\|_{X^{1,\frac{1}{2}}_{\lambda,1}}.$$

For the other terms we use the characteristic energy estimate (5.6) and the packet bounds (4.10) through (4.13) to get

$$\begin{aligned}
\|E\tilde{\phi}^\alpha_{\theta,\lambda} u E v^\omega\|^2_{L^2} &\lesssim \big(\sum_T |c_T|^2_{L^\infty_t} \lambda^{\frac{3}{2}+1}\big) \cdot \sup_T \|E\tilde{\phi}^\alpha_{\theta,\lambda} u\|^2_{L^2(T)} \\
&\lesssim \lambda^{-\frac{1}{2}} \|\tilde{\phi}^\alpha_{\theta,\lambda} u\|^2_{X^{1,\frac{1}{2}}_{\lambda,1}} \sum_{\omega_T=\omega} \lambda^2 |c_T|^2_{L^\infty_t}, \\
\|L\tilde{\phi}^\alpha_{\theta,\lambda} u \underline{L} v^\omega\|^2_{L^2} &\lesssim \|\sum_T c_T L\tilde{\phi}^\alpha_{\theta,\lambda} u \underline{L} u_T\|^2_{L^2} + \big\|L\tilde{\phi}^\alpha_{\theta,\lambda} u \sum_T c'_T v_T\big\|^2_{L^2} \\
&\lesssim \big(\sum_T |c_T|^2_{L^\infty} \lambda^{\frac{3}{2}+2}\big) \cdot \sup_T \|L\tilde{\phi}^\alpha_{\theta,\lambda} u\|^2_{L^2(T)} + \|L\tilde{\phi}^\alpha_{\theta,\lambda} u\|^2_{L^\infty L^2} \|\sum_T c'_T u_T\|^2_{L^2 L^\infty} \\
&\lesssim \lambda^{\frac{1}{2}} \|\tilde{\phi}^\alpha_{\theta,\lambda} u\|^2_{X^{1,\frac{1}{2}}_{\lambda,1}} \sum_{\omega_T=\omega} \lambda^2 |c_T|^2_{L^\infty_t} + \|\tilde{\phi}^\alpha_{\theta,\lambda} u\|^2_{X^{1,\frac{1}{2}}_{\lambda,1}} \sum_{\omega_T=\omega} \lambda^{\frac{3}{2}} |c'_T|^2_{L^2_t}
\end{aligned}$$

Overall

$$\|P_\mu Q(\tilde{\phi}^\alpha_{\theta,\lambda} u, v^\omega)\|_{L^2} \lesssim \lambda^{\frac{1}{4}} \|\tilde{\phi}^\alpha_{\theta,\lambda} u\|_{X^{1,\frac{1}{2}}_{\lambda,1}} \Big(\sum_{\omega_T=\omega} \lambda^2 |c_T|^2_{L^\infty_t} + \lambda |c'_T|^2_{L^2_t} \Big)^{\frac{1}{2}},$$

$$\|P_\mu Q(\tilde{\phi}^\alpha_{\theta,\lambda} u, v^\alpha_{-\theta,\lambda})\|_{L^2} \lesssim \mu^{\frac{1}{2}} \Big(\frac{\alpha\lambda}{\mu}\Big)^{\frac{1}{2}} \|\tilde{\phi}^\alpha_{\theta,\lambda} u\|_{X^{1,\frac{1}{2}}_{\lambda,1}} \Big(\sum_{|\omega-\theta|\sim\alpha} \sum_{\omega_T=\omega\sim\alpha} \lambda^2 |c_T|^2_{L^\infty_t} + \lambda |c'_T|^2_{L^2_t} \Big)^{\frac{1}{2}},$$

which can be summed in α when $\alpha \in [\lambda^{-1/2},\, \mu/\lambda]$.

Case 1a$+-$: In this case the waves $\phi^\alpha_{\theta,\lambda} u$ and v^ω propagate at small relative angles, so by Proposition 5.9 the characteristic energy estimate for the term $L\phi^\alpha_{\theta,\lambda} u \underline{L} v^\omega$ improves by a factor of α. The previous estimates are replaced by

$$\|P_\mu Q(\phi^\alpha_{\theta,\lambda} u, v^\omega)\|_{L^2} \lesssim \alpha\lambda^{\frac{1}{4}} \|\phi^\alpha_{\theta,\lambda} u\|_{X^{1,\frac{1}{2}}_{\lambda,1}} \Big(\sum_{\omega_T=\omega} \lambda^2 |c_T|^2_{L^\infty_t} + \lambda |c'_T|^2_{L^2_t} \Big)^{\frac{1}{2}},$$

$$\begin{aligned}
\|P_\mu Q(\phi^\alpha_{\theta,\lambda} u, v^\alpha_{-\theta,\lambda})\|_{L^2} &\lesssim \alpha\mu^{\frac{1}{2}} \Big(\frac{\alpha\lambda}{\mu}\Big)^{\frac{1}{2}} \|\phi^\alpha_{-\theta,\lambda} u\|_{X^{1,\frac{1}{2}}_{\lambda,1}} \Big(\sum_{|\omega+\theta|\sim\alpha} \sum_{\omega_T=\omega\sim\alpha} \lambda^2 |c_T|^2_{L^\infty_t} + \lambda |c'_T|^2_{L^2_t} \Big)^{\frac{1}{2}} \\
&\lesssim \mu^{\frac{1}{2}} \frac{\mu}{\lambda} \Big(\frac{\alpha\lambda}{\mu}\Big)^{\frac{3}{2}} \Big(\sum_{|\omega+\theta|\sim\alpha} \sum_{\omega_T=\omega\sim\alpha} \lambda^2 |c_T|^2_{L^\infty_t} + \lambda |c'_T|^2_{L^2_t} \Big)^{\frac{1}{2}}.
\end{aligned}$$

Case 1b$++$: For the L^2 estimate we also write

$$\|P_\mu(\tilde{\phi}^\alpha_{\theta,\lambda} u v^\alpha_{\theta,\lambda})\|_{L^2} \lesssim (\alpha^2\lambda)^{\frac{1}{4}} \Big(\sum_{|\omega-\theta|\sim\alpha} \|P_\mu(\tilde{\phi}^\alpha_{\theta,\lambda} u v^\omega)\|^2_{L^2} \Big)^{1/2}$$

For each ω, discard P_μ and use the pointwise bounds for the packets in v^ω and the characteristic L^2 estimate for $\phi^\alpha_{\theta,\lambda} u$ along characteristic surfaces for v^ω (the second part of Proposition 5.9 with $\beta \sim 1$).

$$\|\tilde{\phi}^\alpha_{\theta,\lambda} u v^\omega\|_{L^2} \lesssim \sup_T \|\tilde{\phi}^\alpha_{\theta,\lambda} u\|_{L^2(T)} \cdot \lambda^{\frac{1}{4}} \Big(\sum_T |c_T|^2_{L^\infty_t} \Big)^{\frac{1}{2}} \lesssim \lambda^{\frac{1}{4}} \|\chi^\alpha_\theta u\|_{L^2} \Big(\sum_T |c_T|^2_{L^\infty_t} \Big)^{\frac{1}{2}},$$

where χ^α_θ are square-summable in θ. Thus

$$\|\tilde{\phi}^\alpha_{\theta,\lambda} u v^\alpha_{\theta,\lambda}\|_{L^2} \lesssim (\alpha\lambda)^{\frac{1}{2}} \|\chi^\alpha_\theta u\|_{L^2} \Big(\sum_{|\omega-\theta|\sim\alpha} \sum_{\omega(T)=\omega} |c_T|^2_{L^\infty_t} \Big)^{\frac{1}{2}}$$
$$\lesssim \mu^{\frac{1}{2}} \Big(\frac{\alpha\lambda}{\mu}\Big)^{\frac{1}{2}} \|\chi^\alpha_\theta u\|_{L^2} \Big(\sum_{|\omega-\theta|\sim\alpha} \sum_{\omega(T)=\omega} |c_T|^2_{L^\infty_t} \Big)^{\frac{1}{2}}.$$

Case 1b$+-$: The argument of the $++$ case would yield a loss of α^{-1} due to small angles ($\beta = \alpha$ in Corollary 6.9):

$$\|\phi^\alpha_{\theta,\lambda} u v^\alpha_{-\theta,\lambda}\|_{L^2} \lesssim \mu^{\frac{1}{2}} \Big(\frac{\alpha\lambda}{\mu}\Big)^{\frac{1}{2}} \alpha^{-1} \|\chi^\alpha_\theta u\|_{L^2} \Big(\sum_{|\omega+\theta|\sim\alpha} \sum_{\omega(T)=\omega} |c_T|^2_{L^\infty_t} \Big)^{\frac{1}{2}}.$$

Instead we use the method of Case 3 below with a null foliation *tranverse* to $\phi^\alpha_{\theta,\lambda} u$ to estimate

$$\|P_\mu(\phi^\alpha_{\theta,\lambda} u v^\alpha_{-\theta,\lambda})\|_{L^2} \lesssim \mu \Big(\sum_j \|\phi^\alpha_{\theta,\lambda} u\|^2_{L^2(\Sigma_j)} \|v^\alpha_{-\theta,\lambda}\|_{L^\infty L^2(\Sigma_j)} \Big)^{\frac{1}{2}}$$
$$\lesssim \mu^{\frac{1}{2}} \|\chi^\alpha_\theta u\|_{L^2} \Big(\sum_{|\omega+\theta|\sim\alpha} \sum_{\omega_T=\omega} |c_T|^2_{L^\infty_t} \Big)^{\frac{1}{2}}.$$

Case 2: $\alpha > \mu/\lambda \geq \lambda^{-1/2}$.

Case 2a$++$: By Lemma 6.7, the estimates of the previous case hold with an additional factor of $(\alpha\lambda)^{-1} + (\alpha^2\lambda)^{-N}$ for any N:

$$\|P_\mu Q(\tilde{\phi}^\alpha_{\theta,\lambda} u, v^\alpha_{\theta,\lambda})\|_{L^2}$$
$$\lesssim [(\alpha\lambda)^{-1} + (\alpha^2\lambda)^{-N}] \mu^{\frac{1}{2}} \Big(\frac{\alpha\lambda}{\mu}\Big)^{\frac{1}{2}} \|\tilde{\phi}^\alpha_{\theta,\lambda} u\|_{X^{1,\frac{1}{2}}_{\lambda,1}} \Big(\sum_{|\omega-\theta|\sim\alpha} \sum_{\omega_T=\omega} \lambda^2 |c_T|^2_{L^\infty} + \lambda |c'_T|^2_{L^2_t} \Big)^{\frac{1}{2}},$$

which is summable in α.

Case 2a$+-$:

$$\|P_\mu Q(\phi^\alpha_{\theta,\lambda} u, v^\alpha_{-\theta,\lambda})\|_{L^2}$$
$$\lesssim [(\alpha\lambda)^{-1} + (\alpha^2\lambda)^{-N}] \mu^{\frac{1}{2}} \frac{\mu}{\lambda} \Big(\frac{\alpha\lambda}{\mu}\Big)^{\frac{3}{2}} \|\phi^\alpha_{\theta,\lambda} u\|_{X^{1,\frac{1}{2}}_{\lambda,1}} \Big(\sum_{|\omega+\theta|\sim\alpha} \sum_{\omega_T=\omega} \lambda^2 |c_T|^2_{L^\infty_t} + \lambda |c'_T|^2_{L^2_t} \Big)^{\frac{1}{2}}$$
$$\lesssim \mu^{\frac{1}{2}} \frac{\mu}{\lambda} \Big(\alpha^{\frac{1}{2}} \mu^{-\frac{1}{2}} \frac{\lambda^{\frac{1}{2}}}{\mu} + (\alpha^2\lambda)^{-N+\frac{3}{4}} \Big(\frac{\lambda^{\frac{1}{2}}}{\mu}\Big)^{\frac{3}{2}} \Big) \|\phi^\alpha_{\theta,\lambda} u\|_{X^{1,\frac{1}{2}}_{\lambda,1}} \Big(\sum_{|\omega+\theta|\sim\alpha} \sum_{\omega_T=\omega} \lambda^2 |c_T|^2_{L^\infty_t} + \lambda |c'_T|^2_{L^2_t} \Big)^{\frac{1}{2}}$$

Case 2b$++$: The same considerations as for the null form estimate yield

$$\|P_\mu(\tilde{\phi}^\alpha_{\theta,\lambda} u v^\alpha_{\theta,\lambda})\|_{L^2} \lesssim [(\alpha\lambda)^{-1} + (\alpha^2\lambda)^{-N}] \mu^{\frac{1}{2}} \Big(\frac{\alpha\lambda}{\mu}\Big)^{\frac{1}{2}} \|\chi^\alpha_\theta u\|_{L^2} \Big(\sum_{|\omega-\theta|\sim\alpha} \sum_{\omega_T=\omega} |c_T|^2_{L^\infty_t} \Big)^{\frac{1}{2}}.$$

Case 2b$-$:

$$\|P_\mu(\phi^\alpha_{\theta,\lambda} u v^\alpha_{-\theta,\lambda})\|_{L^2} \lesssim [(\alpha\lambda)^{-1} + (\alpha^2\lambda)^{-N}] \mu^{\frac{1}{2}} \|\chi^\alpha_{-\theta} u\|_{L^2} \Big(\sum_{|\omega-\theta|\sim\alpha} \sum_{\omega_T=\omega} |c_T|^2_{L^\infty_t} \Big)^{\frac{1}{2}}.$$

Case 3: $\mu \leq \lambda^{1/2}$, $\alpha \approx \lambda^{-1/2}$.

Case 3a++: A direct application of Bernstein and energy estimates would yield

$$\|P_\mu Q(\tilde{\phi}^\alpha_{\theta,\lambda} u, v^\alpha_{\theta,\lambda})\|_{L^2} \lesssim \mu \|\tilde{\phi}^\alpha_{\theta,\lambda} u\|_{X^{1,\frac{1}{2}}_{\lambda,1}} \|v^\alpha_{\theta,\lambda}\|_{X^{1,\frac{1}{2}}_{\lambda,1}}.$$

The additional $\mu^{-\frac{1}{2}}$ gain required is precisely what would result from a characteristic energy estimate over a μ^{-1} neighborhood of a null surface.

Using the null foliation Λ_h adapted to $v^\alpha_{-\theta,\lambda}$, we partition spacetime into "null slabs" of thickness μ^{-1}

$$\mathbb{R}^{1+2} = \bigcup_j \bigcup_{h \in [j\mu^{-1}, (j+1)\mu^{-1}]} \Lambda_h =: \bigcup_j \Sigma_j$$

Since the mollifier P_μ averages functions on the μ^{-1} spatial scale, we have roughly

$$\|P_\mu f\|_{L^2(\Sigma_j)} \text{“} \lesssim'' \mu \|f\|_{L^2 L^1(\Sigma_j)}. \tag{6.45}$$

This is not quite accurate since the kernel of P_μ is not compactly supported. However, by partitioning the kernel one can decompose $P_\mu = \sum_k P^k_\mu$ where for any function $f(x)$ and any set $K \subset \mathbb{R}^2$ one has

$$\|P^k_\mu f\|_{L^2(K)} \lesssim_N 2^{-kN} \mu \|f\|_{L^1(K + B(0, 2^k \mu^{-1}))} \text{ for any } N.$$

Hence in the sequel we shall ignore the imprecision in the above Bernstein estimate.

We write

$$\|P_\mu Q(\phi^\alpha_{\theta,\lambda} u, v^\alpha_{-\theta,\lambda})\|^2_{L^2} \lesssim \sum_j \|P_\mu Q(\phi^\alpha_{\theta,\lambda} u, v^\alpha_{-\theta,\lambda})\|^2_{L^2(\Sigma_j)}.$$

Using the "space-localized" Bernstein (6.45) and the null frame for Λ, we estimate

$$\begin{aligned}
\|P_\mu Q(\tilde{\phi}^\alpha_{\theta,\lambda} u, v^\alpha_{\theta,\lambda})\|_{L^2(\Sigma_j)} &\lesssim \mu \|Q(\tilde{\phi}^\alpha_{\theta,\lambda} u, v^\alpha_{\theta,\lambda})\|_{L^2 L^1(\Sigma_j)} \\
&\lesssim \mu \|L\tilde{\phi}^\alpha_{\theta,\lambda} u\|_{L^2(\Sigma_j)} \|\underline{L} v^\alpha_{\theta,\lambda}\|_{L^\infty L^2(\Sigma_j)} \\
&+ \mu \|E\tilde{\phi}^\alpha_{\theta,\lambda} u\|_{L^2(\Sigma_j)} \|E v^\alpha_{\theta,\lambda}\|_{L^\infty L^2(\Sigma_j)} \\
&+ \mu \|\underline{L}\tilde{\phi}^\alpha_{\theta} u\|_{L^\infty L^2(\Sigma_j)} \|L v^\alpha_{\theta,\lambda}\|_{L^2(\Sigma_j)}.
\end{aligned}$$

Apply the characteristic energy estimate (5.6) to $\phi^\alpha_{\theta,\lambda} u$ for the first two terms and to $v^\alpha_{-\theta,\lambda}$ for the third term, thus obtaining a factor of $\mu^{-\frac{1}{2}}$. Since each of the resulting three products remains localized to Σ_j in one factor, we may square-sum both sides in j to conclude that

$$\begin{aligned}
\|P_\mu Q(\tilde{\phi}^\alpha_{\theta,\lambda} u, v^\alpha_{\theta,\lambda})\|_{L^2} &\lesssim \mu^{\frac{1}{2}} (\|\nabla \tilde{\phi}^\alpha_{\theta,\lambda} u\|_{L^\infty L^2} + \|\Box u\|_{L^1 L^2})(\|\nabla v^\alpha_{\theta,\lambda}\|_{L^\infty L^2} + \|\Box v^\alpha_{-\theta,\lambda}\|_{L^1 L^2}) \\
&\lesssim \mu^{\frac{1}{2}} \|\tilde{\phi}^\alpha_{\theta,\lambda} u\|_{X^{1,\frac{1}{2}}_{\lambda,1}} \|v^\alpha_{-\theta,\lambda}\|_{X^{1,\frac{1}{2}}_{\lambda,1}}.
\end{aligned}$$

Case 3a$+-$: In this case we omit the spacetime partition and directly apply Bernstein, Hölder, and the following small-angle improvements of the above estimates:

$$\lambda\|L\phi^\alpha_{\theta,\lambda}u\|_{L^2} + \lambda^{\frac12}\|E\phi^\alpha_{\theta,\lambda}u\|_{L^2} \lesssim \|\phi^\alpha_{\theta,\lambda}u\|_{X^{1,\frac12}_{\lambda,1}} \text{ (by Lemma 5.12)},$$

$$\|\underline{L}v^\alpha_{-\theta,\lambda}\|_{L^\infty L^2} + \lambda^{\frac12}\|Ev^\alpha_{-\theta,\lambda}\|_{L^\infty L^2} \lesssim \|v^\alpha_{-\theta,\lambda}\|_{X^{1,\frac12}_{\lambda,1}} \text{ (by (4.12))},$$

$$\Big(\sum_j \|Lv^\alpha_{-\theta,\lambda}\|^2_{L^\infty L^2(\Sigma_j)}\Big)^{\frac12} \lesssim \lambda^{-1}\|v^\alpha_{-\theta,\lambda}\|_{X^{1,\frac12}_{\lambda,1}} \text{ by (4.11)}.$$

Consequently

$$\|P_\mu Q(\phi^\alpha_{\theta,\lambda}u, v^\alpha_{-\theta,\lambda})\|_{L^2} \lesssim \frac{\mu}{\lambda}\|\phi^\alpha_{\theta,\lambda}u\|_{X^{1,\frac12}_{\lambda,1}}\|v^\alpha_{-\theta,\lambda}\|_{X^{1,\frac12}_{\lambda,1}}.$$

Case 3b$+\pm$: For the L^2 estimate we also use (6.45) and invoking the second part of Corollary 5.10, with $\beta \sim 1$,

$$\begin{aligned}\|P_\mu(\phi^\alpha_{\theta,\lambda}uv^\alpha_{-\theta,\lambda})\|_{L^2(\Sigma_j)} &\lesssim \mu\|\phi^\alpha_{\theta,\lambda}u\|_{L^2(\Sigma_j)}v^\alpha_{-\theta,\lambda}\|_{L^\infty L^2(\Sigma_j)},\\ &\lesssim \mu^{\frac12}\|\chi^\alpha_\theta u\|_{L^2}\|v^\alpha_{-\theta,\lambda}\|_{L^\infty L^2(\Sigma_j)},\end{aligned}$$

which can then be square-summed in j and then in θ.

Case 4: $\mu \le \lambda^{-1/2}, \alpha > \lambda^{-1/2}$. Combine the arguments from Cases 2 and 3.

Finally, to deduce (6.28) we write $u_\lambda := u^{\pm_1}$, $v_\lambda := v^{\pm_2}$, and use the Leibniz rule to bound (6.28) by

$$\|P_\mu(\partial_t u_\lambda v_\lambda)\|_{L^2} + \|P_\mu(u\partial_t v_\lambda)\|_{L^2}.$$

To estimate the first term, decompose

$$u_\lambda = P_{\le 8\lambda}(D_t)u_\lambda + P_{>8\lambda}(D_t)u_\lambda = u_\lambda^{<\lambda} + u_\lambda^{>\lambda}$$

The high-frequency piece satisfies the elliptic estimate

$$\|\nabla_{t,x}u_\lambda^{>\lambda}\|_{L^2} \lesssim \lambda^{-1}(\|\nabla u_\lambda\|_{L^2} + \|\Box u_\lambda\|_{L^2}),$$

whose proof is similar to that of Proposition 2.19. Then by Bernstein and Hölder,

$$\begin{aligned}\|P_\mu(\partial_t u_\lambda^{>\lambda}v_\lambda)\|_{L^2} &\lesssim \mu\|\partial_t u_\lambda^{>\lambda}\|_{L^2}\|v_\lambda\|_{L^\infty L^2}\\ &\lesssim \mu\|u_\lambda\|_{X^{0,\frac12}_{\lambda,1}}\|v_\lambda\|_{X^{0,\frac12}_{\lambda,1}},\end{aligned}$$

which is more than acceptable. On the other hand, the estimate (6.27)

$$\|P_\mu(\partial_t u_\lambda^{<\lambda}v_\lambda)\|_{L^2} \lesssim \mu^{\frac12}\|\partial_t u_\lambda\|_{X^{0,\frac12}_{\lambda,1}}\|v_\lambda\|_{X^{0,\frac12}_{\lambda,1}},$$

and by a simple commutation argument $\|\partial_t u_\lambda^{<\lambda}\|_{X^{0,\frac12}_{\lambda,1}} \lesssim \lambda\|u_\lambda\|_{X^{0,\frac12}_{\lambda,1}}$.

To treat the term $\|P_\mu(u_\lambda\partial_t v)\|_{L^2}$, we repeat the proofs of Case 1b, 2b, 3b, 4 and use (4.13) as well as the Bernstein estimates (4.30), (4.34) for the wave packet coefficients.

6.3. **Proof of Proposition 6.3.** We succesively consider the bounds in the proposition:

(1) First we consider the low modulation cases (6.4) and (6.5).

Let $(\chi^j)_j$ be a partition of unity with respect to time intervals $(I_j)_j$ of length $\simeq d_{\max}^{-1}$ and let $(\tilde{\chi}^j)_j, (\tilde{\tilde{\chi}}^j)_j$ be similar families such that $\tilde{\chi}^j = 1$ on I_j and $\tilde{\tilde{\chi}} = 1$ on the support of $\tilde{\chi}^j$.

By rescaling from Prop. 6.2 using (2.8) we obtain the product estimates for frequency localized functions:

$$X^{1,\frac{1}{2}}_{\lambda,d_{\max}}[I_j] \cdot X^{1,\frac{1}{2}}_{\mu,d_{\max}}[I_j] \longrightarrow X^{1,\frac{1}{2}}_{\lambda',[d_{\max},\mu]}[I_j] \tag{6.46}$$

Note that the regularity of the metric's coefficients improves with the rescaling. Suppose, for instance, that $d_1 \leq d_2 = d_{\max}$. Using properties (2.6), (2.7), (2.5) we have

$$\begin{aligned}
\|u_{\lambda,d_1} \cdot v_{\mu,d_2}\|^2_{X^{1,\frac{1}{2}}_{\lambda',[d_{\max},\mu]}} &\lesssim \sum_j \|\chi^j(u_{\lambda,d_1} \cdot v_{\mu,d_2})\|^2_{X^{1,\frac{1}{2}}_{\lambda',[d_{\max},\mu]}} \\
&\lesssim \sum_j \|\tilde{\chi}^j u_{\lambda,d_1} \cdot \tilde{\chi}^j v_{\mu,d_2}\|^2_{X^{1,\frac{1}{2}}_{\lambda',[d_{\max},\mu]}[I_j]} \\
&\lesssim \sum_j \|\tilde{\tilde{\chi}}^j \tilde{\chi}^j u_{\lambda,d_1}\|^2_{X^{1,\frac{1}{2}}_{\lambda,d_{\max}}[I_j]} \|\tilde{\tilde{\chi}}^j \tilde{\chi}^j v_{\mu,d_2}\|^2_{X^{1,\frac{1}{2}}_{\mu,d_{\max}}[I_j]} \\
&\lesssim \|u_{\lambda,d_1}\|^2_{X^{1,\frac{1}{2}}_{\lambda,d_1}} \sum_j \|\tilde{\chi}^j v_{\mu,d_2}\|^2_{X^{1,\frac{1}{2}}_{\mu,d_2}} \lesssim \|u_{\lambda,d_1}\|^2_{X^{1,\frac{1}{2}}_{\lambda,d_1}} \|v_{\mu,d_2}\|^2_{X^{1,\frac{1}{2}}_{\mu,d_2}}.
\end{aligned}$$

The same argument gives (6.5).

Remark 6.12. Note that between the rescaled (6.2) and (6.46) we have used the factors $\left(\frac{d}{\mu}\right)^{\frac{1}{4}}$ to sum over d- which is the modulation of the output. If we choose to keep this factor, the same argument gives decompositions

$$P_{\lambda'}(u_{\lambda,d_1} \cdot v_{\mu,d_2}) = \sum_{d=d_{\max}}^{\mu} w_{\lambda',d}, \qquad P_{\mu}(u_{\lambda,d_1} \cdot v_{\lambda',d_2}) = \sum_{d=d_{\max}d_0/2}^{\mu/2} w_{\mu,d} + w_{\mu,\mu},$$

for $1 \leq d_1, d_2 \leq \mu$, under the assumptions of Prop 6.3 (1), such that

$$\|w_{\lambda',d}\|_{X^{1,\frac{1}{2}}_{\lambda',d}} \lesssim \left(\frac{d}{\mu}\right)^{\frac{1}{4}} \|u_{\lambda,d_1}\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}} \|v_{\mu,d_2}\|_{X^{1,\frac{1}{2}}_{\mu,d_2}} \tag{6.47}$$

$$\|w_{\mu,d}\|_{X^{1,\frac{1}{2}}_{\mu,d}} \lesssim \frac{\mu}{\lambda}\left(\frac{d}{\mu}\right)^{\frac{1}{4}} \|u_{\lambda,d_1}\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}} \|v_{\lambda',d_2}\|_{X^{1,\frac{1}{2}}_{\lambda',d_2}} \tag{6.48}$$

$$\|w_{\mu,\mu}\|_{\tilde{X}^{1,\frac{1}{2}}_{\mu,\mu}} \lesssim \frac{\mu}{\lambda} \|u_{\lambda,d_1}\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}} \|v_{\lambda',d_2}\|_{X^{1,\frac{1}{2}}_{\lambda',d_2}} \tag{6.49}$$

These will be useful in the proof of the Moser-type estimate.

(2) Recall that by Bernstein's inequality and the energy estimate (2.3) we have

$$\|v_\mu\|_{L^\infty_{t,x}} \lesssim \|v_\mu\|_{X^{1,\frac{1}{2}}_{\mu,d_2}}$$

We now prove (6.6). We have

$$\lambda d_1^{\frac{1}{2}} \|u_{\lambda,d_1} \cdot v_{\mu,d_2}\|_{L^2} \lesssim \left(\lambda d_1^{\frac{1}{2}} \|u_{\lambda,d_1}\|_{L^2}\right) \|v_{\mu,d_2}\|_{L^\infty} \lesssim \|u_{\lambda,d_1}\|_{X^{1,\frac{1}{2}}_{\lambda,d_1}} \|v_{\mu,d_2}\|_{X^{1,\frac{1}{2}}_{\mu,d_2}}.$$

For $\Box_{g_{<\sqrt{\lambda'}}}(u_{\lambda,d_1}\cdot v_{\mu,d_2})$ we consider

$$
\begin{aligned}
&d_1^{-\frac12}\|\Box_{g_{<\sqrt{\lambda}}}u_{\lambda,d_1}\cdot v_{\mu,d_2}\|_{L^2}\lesssim d_1^{-\frac12}\|\Box_{g_{<\sqrt{\lambda}}}u_{\lambda,d_1}\|_{L^2}\|v_{\mu,d_2}\|_{L^\infty}\\
&d_1^{-\frac12}\|(\Box_{g_{<\sqrt{\lambda'}}}-\Box_{g_{<\sqrt{\lambda}}})u_{\lambda,d_1}\cdot v_{\mu,d_2}\|_{L^2}\lesssim \lambda\|u_{\lambda,d_1}\|_{L^2}\|v_{\mu,d_2}\|_{L^\infty}\\
&d_1^{-\frac12}\|u_{\lambda,d_1}\cdot\Box_{g_{<\sqrt{\mu}}}v_{\mu,d_2}\|_{L^2}\lesssim d_1^{-\frac12}\|u_{\lambda,d_1}\|_{L^\infty L^2}\|\Box_{g_{<\sqrt{\mu}}}v_{\mu,d_2}\|_{L^2L^\infty}\\
&\qquad\qquad\lesssim \frac{\mu}{\lambda}\|\nabla u_{\lambda,d_1}\|_{L^\infty L^2}d_2^{-\frac12}\|\Box_{g_{<\sqrt{\mu}}}v_{\mu,d_2}\|_{L^2}\\
&d_1^{-\frac12}\|u_{\lambda,d_1}\cdot(\Box_{g_{<\sqrt{\lambda'}}}-\Box_{g_{<\sqrt{\mu}}})v_{\mu,d_2}\|_{L^2}\lesssim d_1^{-\frac12}\|u_{\lambda,d_1}\|_{L^2}\mu\|v_{\mu,d_2}\|_{L^\infty}\\
&d_1^{-\frac12}\|\partial u_{\lambda,d_1}\cdot\partial v_{\mu,d_2}\|_{L^2}\lesssim d_1^{-\frac12}\|\partial u_{\lambda,d_1}\|_{L^2}\|\partial v_{\mu,d_2}\|_{L^\infty}\\
&\qquad\qquad\lesssim \frac{\mu}{d_1}\|u_{\lambda,d_1}\|_{X^{1,\frac12}_{\lambda,d_1}}\|\partial v_{\mu,d_2}\|_{L^\infty L^2}
\end{aligned}
$$

Each of the five terms above is $\lesssim \|u_{\lambda,d_1}\|_{X^{1,\frac12}_{\lambda,d_1}}\|v_{\mu,d_2}\|_{X^{1,\frac12}_{\mu,d_2}}$.

We continue with the proof of (6.7). Here we will use Bernstein's inequality

$$\|P_\mu w\|_{L^2_{t,x}}\lesssim \mu\|w\|_{L^2L^1}.$$

Assume without loss of generality that $d_1=d_{\max}$. We have

$$
\begin{aligned}
\mu^{\frac52}\|u_{\lambda,d_1}\cdot v_{\lambda',d_2}\|_{L^2L^1}&\lesssim \mu^{\frac52}\|u_{\lambda,d_1}\|_{L^2}\|v_{\lambda',d_2}\|_{L^\infty L^2}\\
&\lesssim \Big(\frac{\mu}{\lambda}\Big)^2\Big(\frac{\mu}{d_{\max}}\Big)^{\frac12}\|u_{\lambda,d_1}\|_{X^{1,\frac12}_{\lambda,d_1}}\|v_{\lambda',d_2}\|_{X^{1,\frac12}_{\lambda',d_2}}.
\end{aligned}
\tag{6.50}
$$

The same argument applies for the $\nabla_{t,x}$ in the L^2 part of $\tilde{X}^{1,\frac12}_{\mu,\mu}$. For the $L^\infty L^2$ parts we use Bernstein, the chain rule and $L^\infty L^2\times L^\infty L^2\to L^\infty L^1$.

For $\Box_{g_{<\sqrt{\mu}}}(u_{\lambda,d_1}\cdot v_{\lambda',d_2})$ we split as before

$$
\begin{aligned}
&\|\Box_{g_{<\sqrt{\lambda}}}u_{\lambda,d_1}\cdot v_{\lambda',d_2}\|_{L^2L^1}\lesssim \|\Box_{g_{<\sqrt{\lambda}}}u_{\lambda,d_1}\|_{L^2}\|v_{\lambda',d_2}\|_{L^\infty L^2}\\
&\|u_{\lambda,d_1}\cdot\Box_{g_{<\sqrt{\lambda'}}}v_{\lambda',d_2}\|_{L^2L^1}\lesssim \|u_{\lambda,d_1}\|_{L^\infty L^2}\|\Box_{g_{<\sqrt{\lambda'}}}v_{\lambda',d_2}\|_{L^2}\\
&\|\partial u_{\lambda,d_1}\cdot\partial v_{\lambda',d_2}\|_{L^2L^1}\lesssim \|\partial u_{\lambda,d_1}\|_{L^2}\|\partial v_{\lambda',d_2}\|_{L^\infty L^2}\\
&\|(\Box_{g_{<\sqrt{\mu}}}-\Box_{g_{<\sqrt{\lambda}}})u_{\lambda,d_1}\cdot v_{\lambda',d_2}\|_{L^2L^1}\lesssim \frac{\lambda}{\mu}\|\partial u_{\lambda,d_1}\|_{L^2}\|v_{\lambda',d_2}\|_{L^\infty L^2}\\
&\|u_{\lambda,d_1}\cdot(\Box_{g_{<\sqrt{\mu}}}-\Box_{g_{<\sqrt{\lambda'}}})v_{\lambda',d_2}\|_{L^2L^1}\lesssim \|u_{\lambda,d_1}\|_{L^2}\frac{\lambda'}{\mu}\|\partial v_{\lambda',d_2}\|_{L^\infty L^2}
\end{aligned}
$$

Each of the five terms times $\mu^{\frac12}$ is $\lesssim \|u_{\lambda,d_1}\|_{X^{1,\frac12}_{\lambda,d_1}}\|v_{\lambda',d_2}\|_{X^{1,\frac12}_{\lambda',d_2}}$, which completes the proof.

7. The product estimate (1.16)

We now turn our attention to the proof of (1.16). We recall the notations $X=X^{s,\theta}$ and $N=X^{s-1,\theta-1}$ with $\theta=\frac12+\varepsilon'$ and $s=1+\varepsilon'+\varepsilon$. The duality property (2.13) states

$$N=(X^{-\varepsilon'-\varepsilon,\frac12-\varepsilon'}+L^2H^{\frac12-2\varepsilon'-\varepsilon})'.$$

Therefore, by duality, to obtain (1.16) it suffices to prove

$$\|u^1 \cdot u^2\|_{X^{-\varepsilon'-\varepsilon,\frac{1}{2}-\varepsilon'}+L^2H^{\frac{1}{2}-2\varepsilon'-\varepsilon}} \lesssim \|u^1\|_{X^{s,\theta}}\|u^2\|_{X^{-\varepsilon'-\varepsilon,\frac{1}{2}-\varepsilon'}+L^2H^{\frac{1}{2}-2\varepsilon'-\varepsilon}}.$$

We reduce this estimate to the following bounds:

$$\|u^1 \cdot u^2\|_{L^2H^{\frac{1}{2}-2\varepsilon'-\varepsilon}} \lesssim \|u^1\|_{X^{s,\theta}}\|u^2\|_{L^2H^{\frac{1}{2}-2\varepsilon'-\varepsilon}} \tag{7.1}$$

$$\|u^1 \cdot u^2\|_{X^{-\varepsilon'-\varepsilon,\frac{1}{2}-\varepsilon'}+L^2H^{\frac{1}{2}-2\varepsilon'-\varepsilon}} \lesssim \|u^1\|_{X^{s,\theta}}\|u^2\|_{X^{-\varepsilon'-\varepsilon,\frac{1}{2}-\varepsilon'}} \tag{7.2}$$

For both estimates we use the Littlewood-Paley trichotomy and reduce to estimates for terms $P_{\lambda_3}(u^1_{\lambda_1} \cdot u^2_{\lambda_2})$.

Be begin with (7.1). In the high-high to low case $\lambda_1 \simeq \lambda_2$ and in the low-high case $\lambda_1 \ll \lambda_2 \simeq \lambda_3$ we use Hölder's inequality $L^\infty L^\infty \times L^2L^2 \to L^2L^2$ together with Bernstein's inequality $\tilde{P}_{\lambda_1}L^\infty L^2 \to \lambda_1 L^\infty L^\infty$. In the high-low case $\lambda_2 \ll \lambda_1 \simeq \lambda_3$ we use Hölder's inequality $L^\infty L^2 \times L^2L^\infty \to L^2L^2$ together with Bernstein's inequality $\tilde{P}_{\lambda_2}L^2L^2 \to \lambda_2 L^2L^\infty$.

Now we turn to the proof of (7.2). We write

$$u^1_{\lambda_1} = \sum_{d_1=1}^{\lambda_1} u^1_{\lambda_1,d_1} \qquad\qquad \|u^1_{\lambda_1}\|^2_{X^{s,\theta}_{\lambda_1}} \simeq \sum_{d_1=1}^{\lambda_1} \|u^1_{\lambda_1,d_1}\|^2_{X^{s,\theta}_{\lambda_1,d_1}}.$$

and the similar decomposition of $u^2_{\lambda_2}$ relative to the space $X^{-\varepsilon'-\varepsilon,\frac{1}{2}-\varepsilon'}$.

In the low-high case $\lambda_1 \ll \lambda_2$ it suffices to prove

$$\|u^1_{\lambda_1} \cdot u^2_{\lambda_2}\|_{X^{0,\frac{1}{2}-\varepsilon'}_{\lambda_2}} \lesssim \frac{1}{\lambda_1^{\varepsilon}}\|u^1_{\lambda_1}\|_{X^{1+\varepsilon'+\varepsilon,\frac{1}{2}+\varepsilon'}_{\lambda_1}}\|u^2_{\lambda_2}\|_{X^{0,\frac{1}{2}-\varepsilon'}_{\lambda_2}}.$$

We estimate $u^1_{\lambda_1,d_1}u^2_{\lambda_2,d_2}$ in $X^{0,\frac{1}{2}-\varepsilon'}_{\lambda_2,[\max(d_1,d_2),\lambda_1]}$ for $d_1,d_2 \le \lambda_1$ using (6.4) and computing the weights we note that there is enough room to sum the modulations $\le \lambda_1$. For $\lambda_1 \le d_2 \le \lambda_2$ we use (6.6) instead and we obtain square-summability in d_2.

In the high-low case $\lambda_2 \ll \lambda_1$ we follow the same argument and here we have a better factor including a power of λ_2/λ_1

In the high-high to low case $\lambda_1 \simeq \lambda_2$ we have

$$\|P_{\lambda_3}(u^1_{\lambda_1} \cdot u^2_{\lambda_2})\|_{X^{-\varepsilon'-\varepsilon,\frac{1}{2}-\varepsilon'}_{\lambda_3}+L^2H^{\frac{1}{2}-2\varepsilon'-\varepsilon}} \lesssim \frac{1}{\lambda_3^{\varepsilon}}\|u^1_{\lambda_1}\|_{X^{1,\frac{1}{2}+\varepsilon'}_{\lambda_1}}\|u^2_{\lambda_2}\|_{X^{0,\frac{1}{2}-\varepsilon'}_{\lambda_2}}.$$

For $d_1,d_2 \le \lambda_3$ we estimate $P_{\lambda_3}(u^1_{\lambda_1,d_1}u^2_{\lambda_2,d_2})$ using (6.5): with appropriate weights the $X^{1.\frac{1}{2}}_{\mu,[d_{\max},\mu]}$ bound transfers to $X^{-\varepsilon'-\varepsilon,\frac{1}{2}-\varepsilon'}_{\lambda_3}$, while the $\tilde{X}^{1.\frac{1}{2}}_{\mu,\tilde{\mu}}$ one transfers to $L^2H^{\frac{1}{2}-2\varepsilon'-\varepsilon}$. For $\max(d_1,d_2) \ge \lambda_3$ we use Bernstein and (6.50) to place the output into $L^2H^{\frac{1}{2}-2\varepsilon'-\varepsilon}$.

Remark 7.1 (Higher regularity). Let $\sigma > s$. By applying the product estimate with a slightly lower $s' < s$ (say $s' = 1+\varepsilon'+\varepsilon/2$) and considering the Littlwood-Paley trichotomy in terms of $u_\mu F_\lambda$, $u_\lambda F_\mu$ ($\mu \ll \lambda$) and $P_\mu(u_\lambda F_{\lambda'})$ ($\mu \lesssim \lambda \simeq \lambda'$) we obtain inequalities of type

$$\|u_\mu F_\lambda\|_{X^{\sigma-1,\theta-1}} \simeq \lambda^{\sigma-s'}\|u_\mu F_\lambda\|_{X^{s'-1,\theta-1}} \lesssim \mu^{-\varepsilon/2}\|u_\mu\|_{X^{s,\theta}}\|F_\lambda\|_{X^{\sigma-1,\theta-1}}$$

$$\|P_\mu(u_\lambda F_{\lambda'})\|_{X^{\sigma-1,\theta-1}} \lesssim (\mu/\lambda)^{\sigma-s}\|u_\lambda\|_{X^{s,\theta}}\|F_{\lambda'}\|_{X^{\sigma-1,\theta-1}}$$

Putting all these estimates together with the similar one for $u_\lambda F_\mu$ we obtain

$$\|u \cdot F\|_{X^{\sigma-1,\theta-1}} \lesssim \|u\|_{X^{s,\theta}} \|F\|_{X^{\sigma-1,\theta-1}} + \|u\|_{X^{\sigma,\theta}} \|F\|_{X^{s-1,\theta-1}} \tag{7.3}$$

Remark 7.2 (Higher regularity, Corollaries of (7.3)). Assuming higher regularity of the metric g, by Remark 2.15, which gives $\|\Box_g u\|_{X^{\sigma-1,\theta-1}} \lesssim \|u\|_{X^{\sigma,\theta}}$ for $\sigma \leq k+1$ and the identity

$$2Q_g(u,v) = 2g^{\alpha\beta}\partial_\alpha u \partial_\beta v = \Box_g(uv) - u\Box_g v - v\Box_g u,$$

from (7.3) and (6.1) we obtain the null form bound

$$\|Q_g(u,v)\|_{X^{\sigma-1,\theta-1}} \lesssim \|u\|_{X^{s,\theta}} \|v\|_{X^{\sigma,\theta}} + \|u\|_{X^{\sigma,\theta}} \|v\|_{X^{s,\theta}}. \tag{7.4}$$

Furthermore, when u, v are bounded in $X^{s,\theta}$ we have the Moser estimates in the form $\|\Gamma(u)\|_{X^{\sigma,\theta}} \lesssim \|u\|_{X^{\sigma,\theta}}$ and

$$\|\Gamma(u) - \Gamma(v)\|_{X^{\sigma,\theta}} \lesssim \|u - v\|_{X^{\sigma,\theta}} + \|u - v\|_{X^{s,\theta}} (\|u\|_{X^{\sigma,\theta}} + \|v\|_{X^{\sigma,\theta}}) \tag{7.5}$$

Using this together with (7.3), (7.4), (1.17), (1.18) we obtain

$$\|\Gamma(u) Q_g(u,u)\|_{X^{\sigma-1,\theta-1}} \lesssim \|u\|^2_{X^{s,\theta}} \|u\|_{X^{\sigma,\theta}}. \tag{7.6}$$

Similarly one obtains a bound for $\Gamma(u)Q_g(u,u) - \Gamma(v)Q_g(v,v)$ in $X^{\sigma-1,\theta-1}$.

8. The Moser estimate (1.18)

In this section we prove the nonlinear estimate (1.18), which resembles the Moser-type estimates in the context of Sobolev spaces (sometimes referred to as Schauder estimates) that can be proved using paradifferential calculus and the chain rule (see [35, Lemma A.9]). Here we will use the iterated paradifferential expansion strategy from [43] to leverage the product estimates from *Section 6*.

Proposition 8.1. *Let F be a smooth bounded function with uniformly bounded derivatives with $\partial^{(j)} F(0) = 0$ for $|j| \leq C$. If $u \in X^{s,\theta}$ then $F(u) \in \tilde{X}^{s,\theta}$ and*

$$\|F(u)\|_{\tilde{X}^{s,\theta}} \lesssim \|u\|_{X^{s,\theta}} \big(1 + \|u\|^{15}_{X^{s,\theta}}\big) \tag{8.1}$$

Moreover, $F(u) \in X^{s,\theta}$ and

$$\|F(u)\|_{X^{s,\theta}} \lesssim \|u\|_{X^{s,\theta}} \big(1 + \|u\|^{15}_{X^{s,\theta}}\big) \tag{8.2}$$

Remark 8.2. The bound (8.1) should be seen as a key intermediate step in the proof of (8.2). The space $\tilde{X}^{s,\theta}$ is defined in Definition 2.16 and $X^{s,\theta} \subset \tilde{X}^{s,\theta}$. The only difference between the two spaces occurs at high modulations, where we discarded the terms $\lambda^{s+\theta-2}\|\Box_{g_{<\sqrt{\lambda}}} u_{\lambda,\lambda}\|_{L^2}$, making the norm $\tilde{X}^{s,\theta}$ smaller. This allows us to have the factor $(\lambda/\lambda_0)^s$ in Lemma 8.5, based on the better factors μ/λ for $\tilde{X}^{1,\frac{1}{2}}_{\mu,\mu}$ compared to $X^{1,\frac{1}{2}}_{\mu,\mu}$ in the estimates (6.5), (6.7), thus making (8.1) easier to prove.

Notation-wise, we focus the proof on the case of functions F of a scalar argument and note that it is easy to see that the same argument applies for the case of multivariate arguments $F(u_1, \dots, u_d)$.

8.1. **Reduction of** (8.2) **to** (8.1)**.** In light of Lemma 2.17 it suffices to show that $\Box_g F(u) \in L^2H^{s+\theta-2}$:

Lemma 8.3. *If $u \in X^{s,\theta}$ then*

$$\|\Box_g F(u)\|_{L^2H^{s+\theta-2}} \lesssim \|u\|_{X^{s,\theta}}\big(1+\|u\|^3_{X^{s,\theta}}\big)$$

Proof. We split $\Box_g F(u)$ into $F'(u)\Box_g u$ and $F''(u)Q_g(u,u)$. Let G be either one of the terms $F'(u), F''(u)$ and let f be $\Box_g u$ or $Q_g(u,u)$.

By Lemma 2.13, (1.17) and (2.16) we have

$$\|f\|_{L^2H^{s+\theta-2}} \lesssim \|u\|_{X^{s,\theta}}\big(1+\|u\|_{X^{s,\theta}}\big). \tag{8.3}$$

To place Gf in $L^2H^{s+\theta-2}$ we use a Littlewood decomposition and bound $P_\lambda(Gf_\nu)$ using (8.3). When $\nu \lesssim \lambda$ we place $G \in L^\infty$ and get the factor $(\nu/\lambda)^{2-s-\theta}$. When $\nu \gg \lambda$, for the term $P_\lambda(G_\nu f_\nu)$ we first use Bernstein $P_\lambda : L^2L^1 \to \lambda L^2L^2$, then we use the fact that $G \in L^\infty H^s$ by classical Moser/Schauder estimates (see [35, Lemma A.9]), obtaining the factor $(\lambda/\nu)^{s+\theta-1}$. □

We continue towards the proof of (8.1) by setting up some preliminaries.

8.2. **Iterated paradifferential expansions.** We write

$$F(u)-F(v) = (u-v)h(v,u)$$

and

$$h(v,u)-h(x,y) = (v-x)h_1(x,y,v,u)+(u-y)h_2(x,y,v,u)$$

and so on, where the h's are generic smooth functions with uniformly bounded derivatives. For $\nu < \mu \le \infty$ we may decompose

$$\begin{aligned} F(u_{<\mu}) &= F(u_{<\nu}) + \sum_{\lambda_0=\nu}^{\mu/2} F(u_{<2\lambda_0}) - F(u_{<\lambda_0}) \\ &= F(u_{<\nu}) + \sum_{\lambda_0=\nu}^{\mu/2} u_{\lambda_0} h(u_{<\lambda_0}, u_{<2\lambda_0}). \end{aligned}$$

Repeating the same argument for $h(u_{<\lambda_0}, u_{<2\lambda_0})$ and denoting

$$h(..,u_{<\lambda_1},..) = h_1(u_{<\lambda_1}, u_{<2\lambda_1}, u_{<2\lambda_1}, u_{<4\lambda_1}) + h_2(u_{<\lambda_1/2}, u_{<\lambda_1}, u_{<\lambda_1}, u_{<2\lambda_1})$$

we further decompose

$$F(u_{<\mu}) = F(u_{<\nu}) + \sum_{\lambda_0=\nu}^{\mu/2} u_{\lambda_0} h(u_1, u_{\le 2}) + \sum_{\lambda_0=\nu}^{\mu/2}\sum_{\lambda_1=2}^{\lambda_0/2} u_{\lambda_0}u_{\lambda_1} h(..,u_{<\lambda_1},..)$$

Iterating this argument we can write $F(u_{<\mu})$ as a sum of three types of terms:

(1) $F(u_{<\nu})$
(2) $u_{\lambda_0}h(u_1,u_{\le 2})$, $u_{\lambda_0}u_{\lambda_1}h(u_1,u_{\le 2},\dots),\dots,u_{\lambda_0}u_{\lambda_1}\cdots u_{\lambda_{N-1}}h(u_1,\dots,u_{\le C})$
(3) $u_{\lambda_0}u_{\lambda_1}\cdots u_{\lambda_N}h(u_{<\lambda_N/c},\dots,u_{<\lambda_N},\dots,u_{<c\lambda_N})$

for $\nu \le \lambda_0 < \mu$ and $\lambda_0 \ge \lambda_1 \ge \cdots \ge \lambda_N \ge 1$.

8.3. **Bilinear estimates involving $\tilde{X}$ spaces.** Here we supplement the estimates from Propositions 6.2 and 6.3 with some bounds in terms of the $\tilde{X}^{s,\theta}_{\mu,\mu}$ norms.

Lemma 8.4. *Let $\mu \le d < \lambda$ and $u_{\mu,\mu} \in \tilde{X}^{s,\theta}_{\mu,\mu}$. Then there exists a decomposition*

$$u_{\mu,\mu} = u^{<\lambda}_{\mu,\mu} + u^{>\lambda}_{\mu,\mu}$$

such that

$$\|u_{\lambda,d} \cdot u^{<\lambda}_{\mu,\mu}\|_{X^{s,\theta}_{\lambda,d}} \lesssim \|u_{\lambda,d}\|_{X^{s,\theta}_{\lambda,d}} \frac{1}{\mu^{\varepsilon}} \|u_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\mu,\mu}} \tag{8.4}$$

$$\|u_{\lambda,d} \cdot u^{>\lambda}_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} \lesssim \Big(\frac{\mu}{d}\Big)^{\theta} \|u_{\lambda,d}\|_{X^{s,\theta}_{\lambda,d}} \frac{1}{\mu^{\varepsilon}} \|u_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\mu,\mu}} \tag{8.5}$$

Moreover, one has

$$\|u_{\lambda,\le\mu} \cdot u^{<\lambda}_{\mu,\mu}\|_{X^{s,\theta}_{\lambda,\mu}} \lesssim \|u_{\lambda,\le\mu}\|_{X^{s,\theta}_{\lambda,\le\mu}} \frac{1}{\mu^{\varepsilon}} \|u_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\mu,\mu}} \tag{8.6}$$

$$\|u_{\lambda,\le\mu} \cdot u^{>\lambda}_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} \lesssim \|u_{\lambda,\le\mu}\|_{X^{s,\theta}_{\lambda,\le\mu}} \frac{1}{\mu^{\varepsilon}} \|u_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\mu,\mu}} \tag{8.7}$$

and

$$\|u_{\lambda,\lambda} \cdot u_{\mu}\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} \lesssim \|u_{\lambda,\lambda}\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} \frac{1}{\mu^{\varepsilon}} \|u_{\mu}\|_{\tilde{X}^{s,\theta}_{\mu}} \tag{8.8}$$

Proof. We define $u^{<\lambda}_{\mu,\mu}$ by averaging $u_{\mu,\mu}$ in time on the λ^{-1} scale.

(1) We begin with (8.4) and (8.5). Similarly to the proof of (6.6), for all terms occurring in each of the products we place the higher frequency term in L^2 and the lower frequency term in L^∞. For (8.4) this works because

$$\|\Box_{g_{<\sqrt{\lambda}}} u^{<\lambda}_{\mu,\mu}\|_{L^\infty} \lesssim \lambda\mu \frac{1}{\mu^{\varepsilon}} \|u_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\mu,\mu}},$$

while for (8.5) we use

$$\|u^{>\lambda}_{\mu,\mu}\|_{L^\infty} \lesssim \lambda^{-1} \|\partial_t u_{\mu,\mu}\|_{L^\infty} \lesssim \frac{\mu}{\lambda} \frac{1}{\mu^{\varepsilon}} \|u_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\mu,\mu}}, \tag{8.9}$$

and we recall that the $\tilde{X}^{s,\theta}_{\lambda,\lambda}$ norm does not contain $\Box_{g_{<\sqrt{\lambda}}}$ terms.

(2) In the case of (8.6) we use $L^\infty L^2 \times L^2 L^\infty \to L^2 L^2$ and Bernstein for the terms $u_{\lambda,\le\mu} \cdot u^{<\lambda}_{\mu,\mu}$, $\nabla_{t,x} u_{\lambda,\le\mu} \cdot \nabla_{t,x} u^{<\lambda}_{\mu,\mu}$ and $u_{\lambda,\le\mu} \cdot \Box_{g_{<\sqrt{\lambda}}} u^{<\lambda}_{\mu,\mu}$. We use $L^2L^2 \times L^\infty L^\infty \to L^2L^2$ for $\Box_{g_{<\sqrt{\lambda}}} u_{\lambda,\le\mu} \cdot u^{<\lambda}_{\mu,\mu}$. The place where we use the $^{<\lambda}$ smoothness is

$$\lambda^{-1}\mu^{\theta} \|\Box_{g_{<\sqrt{\lambda}}} u^{<\lambda}_{\mu,\mu}\|_{L^2} \lesssim \mu^{\theta} \|\nabla_{t,x} u_{\mu,\mu}\|_{L^2} \lesssim \frac{1}{\mu^{\varepsilon}} \|u_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\mu,\mu}}.$$

Finally, in the case of (8.7) we always place the λ-frequency terms in $L^\infty L^2$, using either $L^\infty L^2 \times L^2 L^\infty \to L^2 L^2$ or $L^\infty L^2 \times L^\infty L^\infty \to L^\infty L^2$ and Bernstein for the μ-frequency terms. This works because of (8.9) and

$$\|u^{>\lambda}_{\mu,\mu}\|_{L^2L^\infty} \lesssim \frac{\mu}{\lambda} \|\partial_t u^{>\lambda}_{\mu,\mu}\|_{L^2} \lesssim \frac{\mu}{\lambda} \frac{1}{\mu^{\theta+\varepsilon}} \|u_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\mu,\mu}}.$$

(3) The proof of (8.8) is straightforward by bounding the μ-frequency terms in L^∞. □

8.4. **Multilinear estimates.** The next order of business is to obtain effective bounds for the products $u_{\lambda_0}u_{\lambda_1}\cdots u_{\lambda_N}$ in the expansion from section 8.2, while the effect of multiplying by $h(..u_{<\lambda_N}..)$ is studied in the next subsection.

Lemma 8.5. *Let $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_N$ for $N \geq 1$ and let $u \in \tilde{X}^{s,\theta}$. For any $\lambda \lesssim \lambda_0$*

$$\|P_\lambda\big(u_{\lambda_0}\dots u_{\lambda_N}\big)\|_{\tilde{X}^{s,\theta}_\lambda} \lesssim \frac{\lambda^s}{\lambda_0^s}\|u_{\lambda_0}\|_{\tilde{X}^{s,\theta}_{\lambda_0}}\prod_{i=1}^N \frac{1}{\lambda_i^\varepsilon}\|u_{\lambda_i}\|_{\tilde{X}^{s,\theta}_{\lambda_i}} \tag{8.10}$$

More precisely, there exists a decomposition

$$P_\lambda\big(u_{\lambda_0}\dots u_{\lambda_N}\big) = \sum_{d=1}^{\lambda} v_{\lambda,d}$$

such that

$$\|v_{\lambda,d}\|_{\tilde{X}^{s,\theta}_{\lambda,d}} \lesssim \frac{\lambda^s}{\lambda_0^s}\prod_{i=1}^{\min(3,N)}\min\Big(1,\frac{d}{\min(\lambda,\lambda_i)}\Big)^{\frac14}\|u_{\lambda_0}\|_{\tilde{X}^{s,\theta}_{\lambda_0}}\prod_{i=1}^N \frac{1}{\lambda_i^\varepsilon}\|u_{\lambda_i}\|_{\tilde{X}^{s,\theta}_{\lambda_i}} \tag{8.11}$$

Moreover, if $\lambda > d > \lambda_1$ then we can replace $\|u_{\lambda_0}\|_{\tilde{X}^{s,\theta}_{\lambda_0}}$ by $\|u_{\lambda_0,d}\|_{X^{s,\theta}_{\lambda_0,d}}$.

Remark 8.6. We recall that for $1 \leq d < \lambda$ the norms of $\tilde{X}^{s,\theta}_{\lambda,d}$ and $X^{s,\theta}_{\lambda,d}$ coincide.

Corollary 8.7. *Under the assumptions of Lemma 8.5, for $N \geq 3$ and any $\gamma \leq \min(\lambda,\lambda_3)$, one has*

$$\|v_{\lambda,\leq\gamma}\|_{\tilde{X}^{s,\theta}_{\lambda,\gamma}} \lesssim \frac{\lambda^s}{\lambda_0^s}\|u_{\lambda_0}\|_{\tilde{X}^{s,\theta}_{\lambda_0}}\prod_{i=1}^N \frac{1}{\lambda_i^\varepsilon}\|u_{\lambda_i}\|_{\tilde{X}^{s,\theta}_{\lambda_i}} \tag{8.12}$$

Proof of Corollary 8.7. Using (8.11) with $N \geq 3$ to sum the norms of $\|v_{\lambda,d}\|_{L^2}$ and $\|\Box_{g_{<\sqrt{\lambda}}}v_{\lambda,d}\|_{L^2}$ over $d \in \overline{1,\gamma}$ we obtain a favorable factor, since $\big(\frac{d}{\lambda_{\min}}\big)^{\frac34}$ can be used to always have a positive power of d on the RHS. When $\gamma = \lambda$ we use energy estimates. □

Proof of Lemma 8.5. We prove the statement by induction with respect to N.

(1) For $N = 1$ we use Remark 6.12, Prop. 6.3 (2) and Lemma 8.4 to obtain the decomposition. We split

$$P_\lambda(u_{\lambda_0}u_{\lambda_1}) = \sum_{1\leq d_i\leq\lambda_i, i=0,1} P_\lambda(u_{\lambda_0,d_0}u_{\lambda_1,d_1})$$

where for both $\lambda_i,\ i = 0,1$ we write, by Definition 2.16,

$$u_{\lambda_i} = \sum_{d_i=1}^{\lambda_i/2} u_{\lambda_i,d_i} + u_{\lambda_i,\lambda_i} \qquad \|u_{\lambda_i}\|^2_{X^{s,\theta}_{\lambda_i}} \simeq \sum_{d_i=1}^{\lambda_i/2}\|u_{\lambda_i,d_i}\|^2_{X^{s,\theta}_{\lambda_i,d_i}} + \|u_{\lambda_i,\lambda_i}\|^2_{\tilde{X}^{s,\theta}_{\lambda_i,\lambda_i}}.$$

We obtain the desired estimate by a summation of d_0, d_1 argument as in the proof of Prop. 6.1, as follows:

(1) In the high-low case $\lambda \simeq \lambda_0 \gg \lambda_1$: for $d_0, d_1 \le d < \lambda_1$ we use (6.47) to obtain

$$\|v_{\lambda,d}\|_{X^{s,\theta}_{\lambda,d}} \lesssim \Big(\frac{d}{\lambda_1}\Big)^{\frac14} \frac{1}{\lambda_1^{\varepsilon}} \|u_{\lambda_0,\le d}\|_{X^{s,\theta}_{\lambda_0,\le d}} \|u_{\lambda_1,\le d}\|_{X^{s,\theta}_{\lambda_1,\le d}} \tag{8.13}$$

while when $\lambda_1 < d = d_0 < \lambda, \lambda_0$ we have

$$\|v_{\lambda,d}\|_{X^{s,\theta}_{\lambda,d}} \lesssim \frac{1}{\lambda_1^{\varepsilon}} \|u_{\lambda_0,d}\|_{X^{s,\theta}_{\lambda_0,d}} \|u_{\lambda_1}\|_{\tilde X^{s,\theta}_{\lambda_1}} \tag{8.14}$$

based on (6.6), (8.4) and defining $v_{\lambda,d} = P_\lambda(u_{\lambda_0,d} u_{\lambda_1,<\lambda_1}) + P_\lambda(u_{\lambda_0,d} u^{<\lambda_0}_{\lambda_1,\lambda_1})$, where $u^{<\lambda_0}_{\lambda_1,\lambda_1}$ is defined in Lemma 8.4.

When $d = \lambda_1$ we define $v_{\lambda,\lambda_1} = P_\lambda(u_{\lambda_0,\lambda_1} u_{\lambda_1,<\lambda_1}) + P_\lambda(u_{\lambda_0,\le\lambda_1} u^{<\lambda_0}_{\lambda_1,\lambda_1})$ and using (6.6), (8.6) we have

$$\|v_{\lambda,\lambda_1}\|_{X^{s,\theta}_{\lambda,\lambda_1}} \lesssim \frac{1}{\lambda_1^{\varepsilon}} \|u_{\lambda_0,\le\lambda_1}\|_{X^{s,\theta}_{\lambda_0,\le\lambda_1}} \|u_{\lambda_1}\|_{\tilde X^{s,\theta}_{\lambda_1}} \tag{8.15}$$

Finally, when $d = \lambda$ we define $v_{\lambda,\lambda} = P_\lambda(u_{\lambda_0,<\lambda_0} u^{>\lambda_0}_{\lambda_1,\lambda_1}) + \sum_{d\simeq\lambda_0} P_{\lambda_0}(u_{\lambda_0,d} u_{\lambda_1})$ and using (8.5), (8.7), (8.8) we have

$$\|v_{\lambda,\lambda}\|_{\tilde X^{s,\theta}_{\lambda,\lambda}} \lesssim \frac{1}{\lambda_1^{\varepsilon}} \|u_{\lambda_0}\|_{\tilde X^{s,\theta}_{\lambda_0}} \|u_{\lambda_1}\|_{\tilde X^{s,\theta}_{\lambda_1}} \tag{8.16}$$

(2) In the high-high to low case $\lambda \lesssim \lambda_0 \simeq \lambda_1$: for $d \in [1, \lambda)$, by (6.48) we have

$$\|v_{\lambda,d}\|_{X^{s,\theta}_{\lambda,d}} \lesssim \frac{\lambda^s}{\lambda_0^s} \Big(\frac{d}{\lambda}\Big)^{\frac14} \frac{1}{\lambda_1^{\varepsilon}} \|u_{\lambda_0,\le d}\|_{X^{s,\theta}_{\lambda_0,\le d}} \|u_{\lambda_1,\le d}\|_{X^{s,\theta}_{\lambda_1,\le d}} \tag{8.17}$$

while for $d = \lambda$, $v_{\lambda,\lambda}$ contains the contributions of $u_{\lambda_0,d_0} u_{\lambda_1,d_1}$ when $\max(d_0,d_1) \ge \lambda$ (use (6.7)) and the contributions when $d_0, d_1 \le \lambda$ given by (6.49) obtaining

$$\|v_{\lambda,\lambda}\|_{\tilde X^{s,\theta}_{\lambda,\lambda}} \lesssim \frac{\lambda^s}{\lambda_0^s} \frac{1}{\lambda_1^{\varepsilon}} \|u_{\lambda_0}\|_{\tilde X^{s,\theta}_{\lambda_0}} \|u_{\lambda_1}\|_{\tilde X^{s,\theta}_{\lambda_1}} \tag{8.18}$$

(2) Now we assume the statement holds for $N-1$ and prove it for N. Letting $v^{(N-1)} = u_{\lambda_0} \dots u_{\lambda_{N-1}}$, summing the induction hypothesis for $v^{(N-1)}_{\nu,d'}$ over $d' \le d$ for $d \le \min(\nu, \lambda_1)$ we obtain

$$\|v^{(N-1)}_{\nu,\le d}\|_{\tilde X^{s,\theta}_{\nu,\le d}} \lesssim \frac{\nu^s}{\lambda_0^s} \prod_{i=1}^{\min(3,N-1)} \min\Big(1, \frac{d}{\min(\nu,\lambda_i)}\Big)^{\frac14} \|u_{\lambda_0}\|_{\tilde X^{s,\theta}_{\lambda_0}} \prod_{i=1}^{N-1} \frac{1}{\lambda_i^{\varepsilon}} \|u_{\lambda_i}\|_{\tilde X^{s,\theta}_{\lambda_i}} \tag{8.19}$$

We consider three cases:

(1) $\lambda \ll \lambda_N$. We have

$$P_\lambda\big(u_{\lambda_0} \dots u_{\lambda_N}\big) = P_\lambda\big(\tilde P_{\lambda_N} v^{(N-1)} u_{\lambda_N}\big).$$

For $d < \lambda$ we apply (8.17) to $\tilde P_{\lambda_N} v^{(N-1)}$ and u_{λ_N} and then use (8.19) for $\nu \simeq \lambda_N$. When $d = \lambda$ we use (8.18) instead of (8.17), and then (8.10) for $N-1$ and P_ν for $\nu \simeq \lambda_N$, as no powers of d/λ are necessary.

(2) $\lambda \simeq \lambda_N$. We decompose

$$P_\lambda\big(u_{\lambda_0}\dots u_{\lambda_N}\big) = \sum_{\nu\lesssim\lambda} P_\lambda\big(P_\nu v^{(N-1)} u_{\lambda_N}\big)$$

For $d \le \nu$ we apply (8.13) and (8.15) for u_{λ_N} and $P_\nu v^{(N-1)}$ together with (8.19), while for $d > \nu$ we use (8.14), (8.16) together with (8.10) for $P_\nu v^{(N-1)}$. Summing the factors that contain ν we obtain

$$\sum_{\nu\lesssim\lambda} \frac{\nu^s}{\lambda_0^s}\frac{1}{\nu^\varepsilon}\min\Big(1,\frac{d}{\nu}\Big)^{\frac14\min(3,N)} \simeq \frac{\lambda^s}{\lambda_0^s}\frac{1}{\lambda^\varepsilon}\Big(\frac{d}{\lambda}\Big)^{\frac14\min(3,N)}$$

which is favorable.

(3) $\lambda \gg \lambda_N$. We write

$$P_\lambda\big(u_{\lambda_0}\dots u_{\lambda_N}\big) = P_\lambda\big(\tilde P_\lambda v^{(N-1)} u_{\lambda_N}\big).$$

For $d \le \lambda_N$ we apply (8.13) and (8.15) to $\tilde P_\lambda v^{(N-1)}$ and u_{λ_N} and then use (8.19) for $\nu \simeq \lambda$. In the case $\lambda_N < d < \lambda$ (8.14) applies with $\|v^{(N-1)}_{\nu,d}\|_{X^{s,\theta}_{\nu,d}}$ on the RHS, $\nu \simeq \lambda$, and this norm is estimated by the induction hypothesis. Finally, when $d = \lambda$ we use (8.16) and then (8.10) for $N-1$. □

For the sum of the output frequencies and modulations of the products we have:

Corollary 8.8. *For $N \ge 4$, let $\lambda_0 \ge \lambda_1 \ge \dots \ge \lambda_N$ and let $u \in X^{s,\theta}$. Denoting $w = u_{\lambda_0}\dots u_{\lambda_N}$ and $M = \|u_{\lambda_0}\|_{\tilde X^{s,\theta}_{\lambda_0}} \prod_{i=1}^N \frac{1}{\lambda_i^\varepsilon}\|u_{\lambda_i}\|_{\tilde X^{s,\theta}_{\lambda_i}}$, one has*

$$\|w\|_{L^2} + \lambda_N^{-\frac12}\|w\|_{L^\infty L^2} + \lambda_N^{-1}\|\nabla_{t,x} w\|_{L^2} + \lambda_N^{-\frac32}\|\nabla_{t,x} w\|_{L^\infty L^2} \lesssim \lambda_N^{-s-\theta} M$$

Proof. First consider $\lambda_N \lesssim \lambda$. By summing (8.11) and using also the argument of Cor. 8.7, together with (2.1), (2.3) we obtain

$$\|w_\lambda\|_{L^2} + \lambda_N^{-\frac12}\|w_\lambda\|_{L^\infty L^2} + \lambda^{-1}\|\nabla_{t,x} w_\lambda\|_{L^2} + \lambda_N^{-\frac12}\lambda^{-1}\|\nabla_{t,x} w_\lambda\|_{L^\infty L^2} \lesssim \lambda^{-s}\lambda_N^{-\theta} M \qquad (8.20)$$

Summing in $\lambda \ge \lambda_N/C$ one obtains

$$\|w_{\ge\lambda_N}\|_{L^2} + \lambda_N^{-\frac12}\|w_{\ge\lambda_N}\|_{L^\infty L^2} + \lambda_N^{-1}\|\nabla_{t,x} w_{\ge\lambda_N}\|_{L^2} + \lambda_N^{-\frac32}\|\nabla_{t,x} w_{\ge\lambda_N}\|_{L^\infty L^2} \lesssim \lambda_N^{-s-\theta} M$$

Now we consider frequencies $\ll \lambda_N$ and write

$$P_{\ll\lambda_N} w = \tilde P_{\lambda_N}\big(u_{\lambda_0}\dots u_{\lambda_{N-1}}\big) u_{\lambda_N}.$$

We place $u_{\lambda_N}, \nabla_{t,x} u_{\lambda_N}$ in L^∞ while for $\tilde P_{\lambda_N}\big(u_{\lambda_0}\dots u_{\lambda_{N-1}}\big)$ we use the argument of the previous step used to obtain (8.20). □

8.5. **Multiplicative properties.** Now that we have bounds for multilinear terms in $\tilde X^{s,\theta}_{\lambda,d}$, in order to make use of the expansion in terms $u_{\lambda_0}u_{\lambda_1}\cdots u_{\lambda_N} h(..u_{<\lambda_N}..)$ we need to investigate which multiplications by h leave the $\tilde X^{s,\theta}_{\lambda,d}$ spaces unchanged.

Definition 8.9. For $d \le \lambda$ the norm of $M_{\lambda,d}$ is defined by

$$\|h\|_{M_{\lambda,d}} = \|h\|_{L^\infty} + \frac1d\|\nabla_{t,x} h\|_{L^\infty} + \frac{1}{d\lambda}\|\Box_{g_{<\sqrt\lambda}} h\|_{L^\infty + d^{-\frac12}L^2 L^\infty}$$

and for $\mu \leq \lambda$ the norm of $N_{\lambda,\mu}$ is defined by

$$\|h\|_{N_{\lambda,\mu}} = \|h\|_{L^\infty} + \frac{1}{\lambda}\|\nabla_{t,x} h\|_{L^\infty} + \frac{1}{\lambda}\|\Box_{g_{<\sqrt{\lambda}}} h\|_{L^\infty L^2 + \mu^{-\frac{1}{2}} L^2}.$$

We also define the following version of the $\tilde{X}^{s,\theta}_{\lambda,\lambda}$ norm adapted for function not assumed to be frequency localized:

$$\|w\|_{Y^{s,\theta}_\lambda} = \lambda^{s+\theta}\big(\|w\|_{L^2} + \lambda^{-\frac{1}{2}}\|w\|_{L^\infty L^2} + \lambda^{-1}\|\nabla_{t,x} w\|_{L^2} + \lambda^{-\frac{3}{2}}\|\nabla_{t,x} w\|_{L^\infty L^2}\big).$$

Then we have the following multiplication properties.

Lemma 8.10. *For any function u localized at frequency $\simeq \lambda$ one has*

$$\|u \cdot h\|_{X^{s,\theta}_{\lambda,d}} \lesssim \|u\|_{X^{s,\theta}_{\lambda,d}} \|h\|_{M_{\lambda,d}}. \tag{8.21}$$

and similarly with $X^{s,\theta}_{\lambda,d}$ replaced by $\tilde{X}^{s,\theta}_{\lambda,d}$.

For any function v localized at frequency $\simeq \mu \leq \lambda$ one has

$$\|v \cdot h\|_{X^{s,\theta}_{\lambda,\mu}} \lesssim \frac{\lambda^s}{\mu^s}\|v\|_{X^{s,\theta}_{\mu,\mu}} \|h\|_{N_{\lambda,\mu}} \tag{8.22}$$

$$\|v \cdot h\|_{X^{s,\theta}_{\lambda,\mu}} \lesssim \frac{\lambda^s}{\mu^s}\big[\|v\|_{\tilde{X}^{s,\theta}_{\mu,\mu}} + \mu^\theta \lambda^{-1}\|\partial_t^2 v\|_{L^2 H^{s-1}}\big]\|h\|_{N_{\lambda,\mu}} \tag{8.23}$$

$$\|v \cdot h\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} \lesssim \lambda^{s+\theta}\|(v, \lambda^{-1}\nabla_{t,x} v)\|_{L^2 \cap \mu^{\frac{1}{2}} L^\infty L^2}\|h\|_{N_{\lambda,\mu}}. \tag{8.24}$$

For functions w not assumed to be frequency localized one has:

$$\|w \cdot h\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} \lesssim \|w\|_{Y^{s,\theta}_\lambda}\|h\|_{M_{\lambda,\lambda}}. \tag{8.25}$$

Proof. The proof is straightforward, based on Leibniz's rule, Hölder's inequality and the energy estimate (2.3). For (8.22) we also use Bernstein's inequality $L^2_x \to \mu L^\infty_x$ for v. □

This property is used in conjunction with the following lemma.

Lemma 8.11. *Let $\mu \leq d \leq \lambda$ and $c \simeq 1$. For any smooth, bounded function h with uniformly bounded derivatives one has:*

$$\|P_{\lesssim\mu} h(u_{<\mu})\|_{M_{\lambda,d}} \lesssim 1 + \|u\|^2_{X^{s,\theta}} \tag{8.26}$$

$$\|\tilde{P}_\lambda h(u_{<c\lambda})\|_{N_{\lambda,\mu}} \lesssim 1 + \|u\|^2_{X^{s,\theta}} \tag{8.27}$$

$$\|\tilde{P}_\lambda h(u_{<\mu})\|_{M_{\lambda,\lambda}} \lesssim \big(1 + \|u\|^2_{X^{s,\theta}}\big) \max_{0\leq k\leq 2} \|\tilde{P}_\lambda(\partial^k h)(u_{<\mu})\|_{L^\infty}, \ \mu \ll \lambda. \tag{8.28}$$

The same statement holds for multivariate functions $h(u_{<\mu/c}, \cdots, u_{<\mu}, \cdots, u_{<c\mu})$.

Proof. The same argument applies for the scalar or multivariate case; however, for brevity we will simply use the notation $h(u_{<\mu})$ for both cases.

We begin with the $\nabla_{t,x} h(\cdot)$ terms since the $\|h\|_{L^\infty}$ bound is clear. By the chain rule, for $\nu \in \{\mu, c\lambda\}$, we write schematically

$$\nabla_{t,x} h(u_{<\nu}) = \sum \partial h \cdot \nabla_{t,x} u_{<c_i\nu}.$$

We have $\|\partial h\|_{L^\infty} \lesssim 1$ and $\|\nabla_{t,x} u_{<c_i\nu}\|_{L^\infty} \lesssim \nu \|u\|_{X^{s,\theta}}$, which suffices for all three bounds. Next, we write

$$\Box_{g_{<\sqrt{\lambda}}} P h(u_{<\nu}) = [\Box_{g_{<\sqrt{\lambda}}}, P] h(u_{<\nu}) + P \sum g_{<\sqrt{\lambda}} \partial^2 h \nabla u_{<c_i\nu} \nabla u_{<c_j\nu} + P \sum \partial h \Box_{g_{<\sqrt{\lambda}}} u_{<c_i\nu}$$

In the case of (8.28), since $\mu \ll \lambda$ for $P = \tilde{P}_\lambda$, the only parts of $\partial^2 h, \partial h$ that contribute to this equality are $\tilde{P}_\lambda \partial^2 h, \tilde{P}_\lambda \partial h$ for a frequency λ multiplier $\tilde{P}_\lambda$. By Bernstein we have

$$\begin{aligned}
\|g \partial^2 h \nabla u_{<c_i\mu} \nabla u_{<c_j\mu}\|_{L^\infty} &\lesssim \mu^2 \|\partial^2 h\|_{L^\infty} \|\nabla u_{\lesssim\mu}\|^2_{L^\infty L^2} \lesssim \mu^2 \|\partial^2 h\|_{L^\infty} \|u\|^2_{X^{s,\theta}} \\
\|g \partial^2 h \nabla u_{<c_i\lambda} \nabla u_{<c_j\lambda}\|_{L^\infty L^2} &\lesssim \lambda \|\nabla u_{\lesssim\lambda}\|^2_{L^\infty L^2} \lesssim \lambda \|u\|^2_{X^{s,\theta}}
\end{aligned}$$

Similarly, for $P \in \{P_{\lesssim\mu}, \tilde{P}_\lambda\}$ by writing $[\Box_{g_{<\sqrt{\lambda}}}, P] = [g_{<\sqrt{\lambda}}, P] \partial_{t,x} \partial_x$ and using a standard commutator estimate we obtain

$$\begin{aligned}
\|[\Box_{g_{<\sqrt{\lambda}}}, P_{\lesssim\mu}] h(u_{<\mu})\|_{L^\infty} &\lesssim \mu^2 \|\partial_{t,x} u_{\lesssim\mu}\|^2_{L^\infty L^2} + \mu \|\partial_{t,x} \partial_x u_{\lesssim\mu}\|_{L^\infty L^2} \lesssim \mu^2 \big(1 + \|u\|^2_{X^{s,\theta}}\big) \\
\|[\Box_{g_{<\sqrt{\lambda}}}, \tilde{P}_\lambda] h(u_{<\mu})\|_{L^\infty} &\lesssim \mu^2 \|\tilde{P}_\lambda \partial^2 h\|_{L^\infty} \|\partial_{t,x} u_{\lesssim\mu}\|^2_{L^\infty L^2} + \mu \|\tilde{P}_\lambda \partial h\|_{L^\infty} \|\partial_{t,x} \partial_x u_{\lesssim\mu}\|_{L^\infty L^2} \\
&\lesssim \mu^2 \max\big(\|\tilde{P}_\lambda \partial h\|_{L^\infty}, \|\tilde{P}_\lambda \partial^2 h\|_{L^\infty}\big) \big(1 + \|u\|^2_{X^{s,\theta}}\big) \\
\|[\Box_{g_{<\sqrt{\lambda}}}, \tilde{P}_\lambda] h(u_{<c\lambda})\|_{L^\infty L^2} &\lesssim \lambda \|\partial_{t,x} u_{\lesssim\lambda}\|^2_{L^\infty L^2} + \|\partial_{t,x} \partial_x u_{\lesssim\lambda}\|_{L^\infty L^2} \lesssim \lambda \big(1 + \|u\|^2_{X^{s,\theta}}\big)
\end{aligned}$$

Using Definitions 2.1, 2.3 of the $X^{s,\theta}_{\lambda,d}, X^{s,\theta}$ spaces, the hypothesis $\partial^2 g \in L^2 L^\infty$, Bernstein and a summation argument in frequencies and modulations one obtains

$$\begin{aligned}
\|\Box_{g_{<\sqrt{\lambda}}} u_{<c\mu}\|_{L^2 L^\infty} &\lesssim \mu^{\frac{3}{2}} \|u\|_{X^{s,\theta}} \\
\|\Box_{g_{<\sqrt{\lambda}}} u_{<c\lambda}\|_{L^2} &\lesssim \lambda^{\frac{1}{2}} \|u\|_{X^{s,\theta}}
\end{aligned}$$

Using these we get

$$\begin{aligned}
\|(\tilde{P}) \partial h \Box_{g_{<\sqrt{\lambda}}} u_{<c\mu}\|_{L^2 L^\infty} &\lesssim \mu^{\frac{3}{2}} \|(\tilde{P}) \partial h\|_{L^\infty} \|u\|_{X^{s,\theta}} \\
\|\partial h \Box_{g_{<\sqrt{\lambda}}} u_{<c\lambda}\|_{L^2} &\lesssim \lambda^{\frac{1}{2}} \|u\|_{X^{s,\theta}}
\end{aligned}$$

Putting all the above inequalities together completes the proof of (8.26)-(8.28). □

The main application of the multiplication properties above is:

Lemma 8.12. *Let $\lambda \simeq \lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_N$ for $N \geq 3$ and $v = u_{\lambda_0} \dots u_{\lambda_N}$. For any smooth, bounded function h with uniformly bounded derivatives one has:*

$$\|P_\lambda\big(v \cdot P_{\lesssim\lambda_N} h(u_{<\lambda_N})\big)\|_{\tilde{X}^{s,\theta}_\lambda} \lesssim \|u_{\lambda_0}\|_{X^{s,\theta}_{\lambda_0}} \big(1 + \|u\|^2_{X^{s,\theta}}\big) \prod_{i=1}^N \frac{1}{\lambda_i^{\varepsilon/2}} \|u_{\lambda_i}\|_{X^{s,\theta}_{\lambda_i}} \tag{8.29}$$

Proof. We decompose

$$P_\lambda\big(v \cdot P_{\lesssim\lambda_N} h(u_{<\lambda_N})\big) = \sum_{\lambda' \simeq \lambda} P_\lambda\big(v_{\lambda'} \cdot P_{\lesssim\lambda_N} h(u_{<\lambda_N})\big) + \sum_{\mu \ll \lambda} P_\lambda\big(v_\mu \cdot \tilde{P}_\lambda P_{\lesssim\lambda_N} h(u_{<\lambda_N})\big)$$

where the second sum is non-zero only when $\lambda_N \simeq \lambda$, in which case we can redenote $\tilde{P}_\lambda P_{\lesssim\lambda_N}$ by $\tilde{P}_\lambda$. We separately consider the terms with:

(1) $\lambda' \simeq \lambda$. We use Lemma 8.5 and Corollary 8.7 to decompose $v_{\lambda'}$. Applying (8.21) together with (8.26) (with $\mu = \lambda_N, d = \lambda_N$, resp. $d > \lambda_N$) and (8.12), resp. (8.11) we obtain

$$\|v_{\lambda',\leq\lambda_N} \cdot P_{\lesssim\lambda_N} h(u_{<\lambda_N})\|_{X^{s,\theta}_{\lambda,\lambda_N}} \lesssim \|u_{\lambda_0}\|_{X^{s,\theta}_{\lambda_0}} \big(1 + \|u\|^2_{X^{s,\theta}}\big) \prod_{i=1}^N \frac{1}{\lambda_i^\varepsilon} \|u_{\lambda_i}\|_{X^{s,\theta}_{\lambda_i}}$$

$$\|v_{\lambda',d} \cdot P_{\lesssim\lambda_N} h(u_{<\lambda_N})\|_{X^{s,\theta}_{\lambda,d}} \lesssim \Big(\frac{d}{\lambda_1}\Big)^{\frac14} \|u_{\lambda_0}\|_{X^{s,\theta}_{\lambda_0}} \big(1 + \|u\|^2_{X^{s,\theta}}\big) \prod_{i=1}^N \frac{1}{\lambda_i^\varepsilon} \|u_{\lambda_i}\|_{X^{s,\theta}_{\lambda_i}}$$

for $d \in [\lambda_N, \lambda_1]$, while for $d \in (\lambda_1, \lambda)$ we have

$$\|v_{\lambda',d} \cdot P_{\lesssim\lambda_N} h(u_{<\lambda_N})\|_{X^{s,\theta}_{\lambda,d}} \lesssim \|u_{\lambda_0,d}\|_{X^{s,\theta}_{\lambda_0,d}} \big(1 + \|u\|^2_{X^{s,\theta}}\big) \prod_{i=1}^N \frac{1}{\lambda_i^\varepsilon} \|u_{\lambda_i}\|_{X^{s,\theta}_{\lambda_i}}$$

$$\|v_{\lambda',\lambda'} \cdot P_{\lesssim\lambda_N} h(u_{<\lambda_N})\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} \lesssim \|u_{\lambda_0}\|_{X^{s,\theta}_{\lambda_0}} \big(1 + \|u\|^2_{X^{s,\theta}}\big) \prod_{i=1}^N \frac{1}{\lambda_i^\varepsilon} \|u_{\lambda_i}\|_{X^{s,\theta}_{\lambda_i}}$$

We square-sum these bounds to obtain the conclusion for the $v_{\lambda'}$ terms.

(2) $\mu \ll \lambda$, in the case $\lambda_N \simeq \lambda_{N-1} \simeq \cdots \simeq \lambda_0 \simeq \lambda$. By (8.12) from Corollary 8.7 and (8.11) we have

$$\|v_{\mu,<\mu}\|_{X^{s,\theta}_{\mu,\mu}} + \|v_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\mu,\mu}} \lesssim \frac{\mu^s}{\lambda_0^s} \|u_{\lambda_0}\|_{X^{s,\theta}_{\lambda_0}} \prod_{i=1}^N \frac{1}{\lambda_i^\varepsilon} \|u_{\lambda_i}\|_{X^{s,\theta}_{\lambda_i}}.$$

Using this together with (8.22) and (8.27) (with $c\lambda = \lambda_N$) we obtain :

$$\|v_{\mu,<\mu} \cdot \tilde{P}_\lambda h(u_{<\lambda_N})\|_{X^{s,\theta}_{\lambda,\mu}} \lesssim \|u_{\lambda_0}\|_{X^{s,\theta}_{\lambda_0}} \big(1 + \|u\|^2_{X^{s,\theta}}\big) \prod_{i=1}^N \frac{1}{\lambda_i^\varepsilon} \|u_{\lambda_i}\|_{X^{s,\theta}_{\lambda_i}}.$$

Summing $\mu \leq \lambda/C$ gives a factor of $\log\lambda$ which is overpowered by $\frac{1}{\lambda^{\varepsilon/2}} \simeq \frac{1}{\lambda_i^{\varepsilon/2}}$.

We regularize $v_{\mu,\mu}$ in time on the λ^{-1} scale to split $v_{\mu,\mu} = v^{<\lambda}_{\mu,\mu} + v^{>\lambda}_{\mu,\mu}$. We have

$$\|\partial_t^2 v^{<\lambda}_{\mu,\mu}\|_{L^2 H^{s-1}} \lesssim \lambda \mu^{-\theta} \|v_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\mu,\mu}},$$

so we can use this together with (8.23). Finally, we have

$$\|(v^{>\lambda}, \lambda^{-1}\nabla_{t,x} v^{>\lambda})\|_{L^2\cap\mu^{\frac12}L^\infty L^2} \lesssim \lambda^{-1} \|\nabla_{t,x} v^{>\lambda}\|_{L^2\cap\mu^{\frac12}L^\infty L^2} \lesssim \lambda^{-1}\mu^{1-s-\theta} \|v_{\mu,\mu}\|_{\tilde{X}^{s,\theta}_{\mu,\mu}}$$

and we use this with (8.24) to conclude. □

8.6. **Nonlinear low frequency input → high frequency output estimates.**

Low → high frequency interactions do not occur in bilinear or multilinear expressions, thus one expects their effect to be under control for nonlinear interactions as well. We begin with L^∞ bounds.

Lemma 8.13. *Let $\lambda \geq \mu$ and let h be a smooth, bounded function h with uniformly bounded derivatives. Then, for any even $N \geq 2$*

$$\|P_\lambda h(u_{<\mu})\|_{L^\infty} \lesssim \Big(\frac{\mu}{\lambda}\Big)^N \|u\|_{X^{s,\theta}}\big(1 + \|u\|_{X^{s,\theta}}^{N-1}\big) \tag{8.30}$$

$$\|\nabla_{t,x} P_\lambda h(u_{<\mu})\|_{L^\infty} \lesssim \mu\Big(\frac{\mu}{\lambda}\Big)^N \|u\|_{X^{s,\theta}}\big(1 + \|u\|_{X^{s,\theta}}^{N}\big) \tag{8.31}$$

Proof. For any nonzero multi-index β we may use the chain rule to write

$$\partial_x^\beta h(u_{<\mu}) = \sum_{j=1}^{|\beta|} \sum_{\beta_1+\cdots+\beta_j=\beta} h^{(j)}(u_{<\mu})\partial_x^{\beta_1} u_{<\mu} \dots \partial_x^{\beta_j} u_{<\mu}, \qquad \beta_i \neq 0.$$

Using $\|\partial_x^\beta u_{<\mu}\|_{L^\infty} \lesssim \mu^{|\beta|}\|u\|_{X^{s,\theta}}$ we obtain

$$\|P_\lambda h(u_{<\mu})\|_{L^\infty} \lesssim \frac{1}{\lambda^N}\|\Delta^{N/2} P_\lambda h(u_{<\mu})\|_{L^\infty} \lesssim \Big(\frac{\mu}{\lambda}\Big)^N \|u\|_{X^{s,\theta}}\big(1 + \|u\|_{X^{s,\theta}}^{N-1}\big).$$

The same argument is used for $\nabla_{t,x} P_\lambda h(u_{<\mu})$. □

For very low frequency terms we have:

Lemma 8.14. *Let h be a smooth, bounded function with uniformly bounded derivatives such that $h(0) = 0$. Then, for all $\lambda \geq 1$ and even $N \geq 2$ one has*

$$\|\tilde{P}_\lambda h(u_1)\|_{X_{\lambda,\lambda}^{s,\theta}} \lesssim \frac{1}{\lambda^N}\|u_1\|_{X_1^{s,\theta}}\big(1 + \|u_1\|_{X_1^{s,\theta}}^{N+1}\big). \tag{8.32}$$

The same statement holds for $h(u_{\leq C})$ and for multivariate functions $h(u_1, u_{\leq 2}, \cdots, u_{\leq 2^k})$.

Proof. When $\lambda = 1$ we use the Lipschitz property $|h(u_1)| \lesssim |u_1|$ to control $\|h(u_1)\|_{L^2}$. For $\|\Box_{g_{<1}} \tilde{P}_1 h(u_1)\|_{L^2}$ we use the chain rule.

Now we assume $\lambda > 1$. We have

$$\|\tilde{P}_\lambda h(u_1)\|_{L^2} \lesssim \frac{1}{\lambda^{N+2}}\|\Delta^{\frac{N}{2}+1} h(u_1)\|_{L^2}$$

Using the chain rule we write

$$\Delta^{\frac{N}{2}+1} h(u_1) = \sum h^{(j)}(u_1)\partial_x^{\beta_1} u_1 \cdots \partial_x^{\beta_j} u_1$$

We use the L^2 norm for $\partial_x^{\beta_1} u_1$ and L^∞ for the other terms. The same type of argument applies for bounding $\|\Box_{g_{<\sqrt{\lambda}}} \tilde{P}_\lambda h(u_1)\|_{L^2}$. □

With all the preparations above we are ready to treat high frequency outputs in high modulation spaces $\tilde{X}_{\lambda,\lambda}^{s,\theta}$.

Proposition 8.15. *Let F be a smooth, bounded function with uniformly bounded derivatives such that $F^{(j)}(0) = 0$ for $j \leq 4$. Then, for any even $N \geq 2$ and any $\mu \ll \lambda$ one has*

$$\|P_\lambda F(u_{<\mu})\|_{\tilde{X}_{\lambda,\lambda}^{s,\theta}} \lesssim \Big(\frac{\mu}{\lambda}\Big)^N \big(1 + \|u\|_{X^{s,\theta}}^{N+8}\big) \sup_{\nu \leq \mu} \Big(\frac{\nu}{\mu}\Big)^N \|u_\nu\|_{X_\nu^{s,\theta}}. \tag{8.33}$$

We same result applies for multivariate functions $P_\lambda h(u_{<\mu/c}, \dots, u_{<\mu}, \dots, u_{<c\mu})$.

Proof. We use the iterated paradifferential expansion from Section 8.2 with $\nu = 2$ to express $F(u_{<\mu})$ as a sum of terms of type:

(1) $F(u_1), u_{\lambda_0} h(u_1, u_{\leq 2}), \dots, u_{\lambda_0} u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} h(u_1, \dots, u_{\leq C})$ for $\lambda_3 \leq \lambda_2 \leq \lambda_1 \leq \lambda_0 \leq \mu \ll \lambda$. Using (8.32), (8.8) we get

$$\begin{aligned}
\|P_\lambda F(u_1)\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} &\lesssim \frac{1}{\lambda^N}\big(1 + \|u\|^{N+1}_{X^{s,\theta}}\big)\|u_1\|_{X^{s,\theta}_1} \\
\|u_{\lambda_0}\tilde{P}_\lambda h(u_1, u_{\leq 2})\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} &\lesssim \|\tilde{P}_\lambda h(u_1, u_{\leq 2})\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} \frac{1}{\lambda_0^\varepsilon}\|u_{\lambda_0}\|_{X^{s,\theta}_{\lambda_0}} \\
&\lesssim \frac{1}{\lambda^N}\frac{1}{\lambda_0^\varepsilon}\|u_{\lambda_0}\|_{X^{s,\theta}_{\lambda_0}}\big(1 + \|u\|^{N+1}_{X^{s,\theta}}\big)\|u_{\leq 2}\|_{X^{s,\theta}} \\
&\dots \\
\|u_{\lambda_0} u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} \tilde{P}_\lambda h(u_1, \dots, u_{\leq C})\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} &\lesssim \frac{1}{\lambda^N}\big(1 + \|u\|^{N+1}_{X^{s,\theta}}\big)\prod_{i=0}^{3}\frac{1}{\lambda_i^\varepsilon}\|u_{\lambda_i}\|_{X^{s,\theta}_{\lambda_i}}\|u_{\leq C}\|_{X^{s,\theta}}
\end{aligned} \tag{8.34}$$

Then we may sum $\lambda_0, \dots, \lambda_3$.

(2) $u_{\lambda_0} u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} u_{\lambda_4} h(u_{<\lambda_4/c}, \dots, u_{<\lambda_4}, \dots, u_{<c\lambda_4})$ for $\lambda_4 \leq \dots \leq \lambda_0 \leq \mu \ll \lambda$.

By abuse of notation we denote

$$h(u_{<\lambda_4}) = h(u_{<\lambda_4/c}, \dots, u_{<\lambda_4}, \dots, u_{<c\lambda_4}), \qquad w = u_{\lambda_0} u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} u_{\lambda_4}$$

Then

$$P_\lambda\big(w \cdot h(u_{<\lambda_4})\big) = P_\lambda\big(w \cdot \tilde{P}_\lambda h(u_{<\lambda_4})\big)$$

We discard the P_λ and use (8.25):

$$\|P_\lambda\big(w \cdot h(u_{<\lambda_4})\big)\|_{\tilde{X}^{s,\theta}_{\lambda,\lambda}} \lesssim \|w\|_{Y^{s,\theta}_\lambda}\|\tilde{P}_\lambda h(u_{<\lambda_4})\|_{M_{\lambda,\lambda}}$$

By Corollary 8.8 we obtain

$$\|w\|_{Y^{s,\theta}_\lambda} \lesssim \Big(\frac{\lambda}{\lambda_4}\Big)^{s+\theta}\|u_{\lambda_0}\|_{X^{s,\theta}_{\lambda_0}}\prod_{i=1}^{4}\frac{1}{\lambda_i^\varepsilon}\|u_{\lambda_i}\|_{X^{s,\theta}_{\lambda_i}}$$

Using (8.28) together with (8.30) (for $h, \partial h$ and $\partial^2 h$) we obtain

$$\|\tilde{P}_\lambda h(u_{<\lambda_4})\|_{M_{\lambda,\lambda}} \lesssim \Big(\frac{\lambda_4}{\lambda}\Big)^{N+2}\|u\|_{X^{s,\theta}}\big(1 + \|u\|^{N+3}_{X^{s,\theta}}\big)$$

We put these together and sum over all but λ_4. It remains to bound

$$\sum_{\lambda_4 \leq \mu} \log\big(\frac{\mu}{\lambda_4}\big)\Big(\frac{\lambda_4}{\lambda}\Big)^{N+2-s-\theta}\|u_{\lambda_4}\|\big(1 + \|u\|^{N+8}_{X^{s,\theta}}\big) \lesssim$$

$$\lesssim \Big(\frac{\mu}{\lambda}\Big)^N\Big[\sup_{\lambda_4 \leq \mu}\Big(\frac{\lambda_4}{\mu}\Big)^N\|u_{\lambda_4}\|_{X^{s,\theta}_{\lambda_4}}\Big]\sum_{\lambda_4 \leq \mu}\Big(\frac{\lambda_4}{\lambda}\Big)^{2-s-\theta-\varepsilon}\big(1 + \|u\|^{N+8}_{X^{s,\theta}}\big)$$

which completes the proof. □

8.7. **Conclusion - Proof of Proposition 8.1.** It remains to prove (8.1). We use the iterated paradifferential expansion from section 8.2 with $\mu = \infty$ and $\nu = \lambda/C$ to express $F(u)$ as a sum of terms of type:

(1) $F(u_{\ll\lambda})$
(2) $u_{\lambda_0} h(u_1, u_{\leq 2}),\ u_{\lambda_0} u_{\lambda_1} h(u_1, u_{\leq 2}, \dots), \dots, u_{\lambda_0} u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} h(u_1, \dots, u_{\leq C})$
(3) $u_{\lambda_0} u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} u_{\lambda_4} h(u_{<\lambda_4/c}, \dots, u_{<\lambda_4}, \dots, u_{<c\lambda_4})$

for $\lambda \lesssim \lambda_0$ and $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_4 \geq 1$.

The term $\|P_\lambda F(u_{\ll\lambda})\|_{\tilde{X}^{s,\theta}_\lambda}$ is estimated by (8.33). Note that the RHS is square-summable in λ.

For the terms in (2) we use (8.10), for $\eta \lesssim \lambda_0$

$$\|P_\lambda\big[\Pi_i u_{\lambda_i} P_\eta h(u_1, ..)\big]\|_{\tilde{X}^{s,\theta}_\lambda} \lesssim \frac{\lambda^s}{\lambda_0^s} \|u_{\lambda_0}\|_{\tilde{X}^{s,\theta}_{\lambda_0}} \big\|P_\eta h(u_1, ..)\big\|_{\tilde{X}^{s,\theta}_\eta} \prod_{i\neq 0} \frac{1}{\lambda_i^\varepsilon} \|u_{\lambda_i}\|_{\tilde{X}^{s,\theta}_{\lambda_i}}$$

Then we apply Lemma 8.14 and sum in $\eta, \lambda_i, i \geq 1$. We have square-summability in λ due to the factors $\lambda^s \lambda_0^{-s} \|u_{\lambda_0}\|_{\tilde{X}^{s,\theta}_{\lambda_0}}$.

We continue with the term (3). For $\eta \gg \lambda_4$ we use (8.10) and Prop. 8.15 for

$$P_\lambda\big[u_{\lambda_0} u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} u_{\lambda_4} P_\eta h(u_{<\lambda_4/c}, \dots, u_{<\lambda_4}, \dots, u_{<c\lambda_4})\big],$$

and then sum $\eta \geq C\lambda_4$ to obtain

$$\|P_\lambda\big[\Pi_i u_{\lambda_i} P_{\gg\lambda_4} h(..u_{<\lambda_4}..)\big]\|_{\tilde{X}^{s,\theta}_\lambda} \lesssim \frac{\lambda^s}{\lambda_0^s} \|u_{\lambda_0}\|_{X^{s,\theta}_{\lambda_0}} \big(1 + \|u\|^{11}_{X^{s,\theta}}\big) \prod_{i=1}^4 \frac{1}{\lambda_i^\varepsilon} \|u_{\lambda_i}\|_{X^{s,\theta}_{\lambda_i}}.$$

It remains to consider

$$P_\lambda\big[u_{\lambda_0} u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} u_{\lambda_4} P_{\lesssim\lambda_4} h(u_{<\lambda_4/c}, \dots, u_{<\lambda_4}, \dots, u_{<c\lambda_4})\big]$$

For $\lambda_0 \simeq \lambda$ we use Lemma 8.12. Now we assume $\lambda_0 \gg \lambda$, in which case $\lambda_1 \simeq \lambda_0$ and

$$P_\lambda\big[\Pi_i u_{\lambda_i} P_{\lesssim\lambda_4} h(..u_{<\lambda_4}..)\big] = P_\lambda\big[u_{\lambda_0} \tilde{P}_{\lambda_0}\big(u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} u_{\lambda_4} P_{\lesssim\lambda_4} h(..u_{<\lambda_4}..)\big)\big]$$

We bound this by first using the estimate (8.10) for the two terms and then Lemma 8.12 applied to $\tilde{P}_{\lambda_0}\big(u_{\lambda_1} u_{\lambda_2} u_{\lambda_3} u_{\lambda_4} P_{\lesssim\lambda_4} h(..u_{<\lambda_4}..)\big)$. This concludes the proof of Proposition 8.1. □

Appendix A. Smith's wave packets

We briefly review Smith's wave packet parametrix construction for wave equations

$$(\partial_t^2 - g^{ab}(t,x)\partial_a\partial_b)u = 0,\ u[0] = (u_0, u_1)$$

with $C^{1,1}$ coefficients [29]. This construction works as well for coefficients g^{ij} satisfying $\partial^2 g \in L^2 L^\infty$. Begin with a partition of unity on $\mathbb{R}^n_\xi$:

$$1 = |h_1(\xi)|^2 + \sum_{\lambda\geq 2} \sum_\omega |h^\omega_\lambda(\xi)|^2$$

where, for each frequency $\lambda \geq 2$, the ω's are summed over $\simeq \lambda^{\frac{n-1}{2}}$ directions that are uniformly separated on the unit sphere. The smooth functions $h^\omega_\lambda(\xi)$ vanish outside the annular sectors

$$\{\ |\xi| \simeq \lambda,\ |\xi/|\xi| - \omega| \lesssim \lambda^{-\frac{1}{2}}\ \}$$

and satisfy the natural bounds

$$\left|(\omega\cdot\nabla_\xi)^j\partial_\xi^\alpha\ h_\lambda^\omega(\xi)\right|\lesssim\lambda^{-j-\frac{|\alpha|}{2}}.$$

Thus each $h_\lambda^\omega(\xi)$ is supported in a rectangle of size $\simeq(\lambda^{\frac12})^{n-1}\times\lambda$ oriented in the ω direction. To each of them we associate a lattice Ξ_λ^ω in the physical space $\mathbb{R}^n_x$ on the dual scale, i.e. spaced λ^{-1} in the ω direction and spaced $\lambda^{-\frac12}$ in directions in $\omega^\perp$. For each λ we denote

$$\mathcal{T}_\lambda=\{T=(x,\omega)\mid x\in\Xi_\lambda^\omega\}$$

and to each T we will associate a "tube" and a function φ_T defined by

$$\widehat{\varphi_T}(\xi)=(2\pi)^{-\frac n2}\lambda^{-\frac{n+1}{4}}e^{-ix\cdot\xi}h_\lambda^\omega(\xi).$$

This function is concentrated in phase space around $(x,\lambda\omega)$. Thus, we have

$$\left|(\omega\cdot\nabla_y)^j(\omega^\perp\cdot\nabla_y)^\alpha\ \varphi_T(y)\right|\lesssim\lambda^{j+\frac{|\alpha|}{2}+\frac{n+1}{4}}\frac{1}{(1+\lambda\left|\omega\cdot(y-x)\right|+\lambda\left|y-x\right|^2)^N}$$

These properties imply that for any sum we have

$$\int_{\mathbb{R}^n}\Big|\sum_{\lambda,T\in\mathcal{T}_\lambda}d_T\varphi_T\Big|^2\,\mathrm{d}y\lesssim\sum_{\lambda,T\in\mathcal{T}_\lambda}|d_T|^2\,.$$

The family $(\varphi_T)_{T,\lambda}$ is used to decompose arbitrary functions into highly localized components. Indeed, for any $f\in L^2(\mathbb{R}^n)$ we have

$$f=\sum_{\lambda,T\in\mathcal{T}_\lambda}c_T\varphi_T,\qquad c_T:=\int_{\mathbb{R}^n}\overline{\varphi_T}f\,\mathrm{d}y$$

In addition, we have

$$\int_{\mathbb{R}^n}|f|^2\ \mathrm{d}y=\sum_{\lambda,T\in\mathcal{T}_\lambda}|c_T|^2\,.$$

Note that if f is localized at frequencies $\simeq\lambda$, then it's decomposition only contains terms φ_T with $T\in\mathcal{T}_{\lambda'}$ where $\lambda'\simeq\lambda$. By abuse of notation we will sometimes write simply the sum over $T\in\mathcal{T}_\lambda$. The decomposition is, in general, not unique since the φ_T's are not linearly independent.

Let $T=(x_T,\omega_T)\in\mathcal{T}_\lambda$ and fix a sign $\pm$. Let $(x_T(t),\omega_T(t))$ be the projection to $S^*(\mathbb{R}^n)$ of the bicharacteristic initialized at (x_T,ω_T). In other words, we set $\omega_T(t)=\frac{\xi(t)}{|\xi(t)|}$ and then $(x_T(t),\omega_T(t))$ solves

$$\begin{cases}\frac{\mathrm{d}x}{\mathrm{d}t}=\pm\partial_\xi a(t,x,\omega)\\ \frac{\mathrm{d}\omega}{\mathrm{d}t}=\mp\partial_x a(t,x,\omega)\pm\langle\omega,\partial_x a(t,x,\omega)\rangle\omega,\end{cases}\qquad a(t,x,\xi):=(g^{ab}_{<\lambda^{1/2}}\xi_a\xi_b)^{1/2}.\tag{A.1}$$

Define the orthogonal matrix $\Theta^\pm(t)$ by the ODE

$$\dot{\Theta^\pm}=\mp\Theta^\pm[\omega\otimes a_x(t,x,\omega)-(a_x)(t,x,\omega)\otimes\omega],$$

where $v \otimes w$ is the linear map $x \mapsto v\langle w, x\rangle$; by construction, $\Theta^{\pm}(t)\omega_T(t) = \omega_T$. Set

$$u_T^{\pm}(t,y) = \varphi_T(\Theta^{\pm}(t)(y - x_T(t)) + x_T),$$
$$v_T^{\pm}(t,y) = \frac{1}{a(0,x_T,\omega_T)}\psi_T(\Theta^{\pm}(t)(y - x_T^{\pm}(t)) + x_T),$$

where

$$\widehat{\psi_T}(\xi) := -\lambda\langle\omega_T,\xi\rangle^{-1}\widehat{\varphi}_T(\xi).$$

A parametrix for initial data $u[0] := (u_0, u_1)$ is then given by

$$\begin{aligned} w(t) &:= \frac{1}{2}\sum_{\lambda}\sum_{T\in\mathcal{T}_\lambda}(u_T^+(t) + u_T^-(t))\langle\varphi_T, u_0\rangle_{L_x^2} \\ &\quad + \frac{1}{2}\Big[\sum_{\lambda\le\lambda_0}\sum_{T\in\mathcal{T}_\lambda}\lambda^{-1}(v_T^+(t) - v_T^-(t))\langle\varphi_T,(I+E)u_1\rangle_{L_x^2}\Big] + tP_{<\lambda_0}(I+E)u_1 \\ &=: \mathbf{c}(t,0)u_0 + \mathbf{s}(t,0)u_1, \end{aligned}$$

where $\lambda_0 \gg 1$ is an absolute constant, and the operator E satisfies $\|E\|_{L^2\to L^2} \lesssim \lambda_0^{-\frac{1}{2}}$ and ensures that $\partial_t\mathbf{s}|_{t=0} = I$. The operators $\mathbf{c}(t,0)$, $\mathbf{s}(t,0)$ are approximations of the usual wave propagators $\cos t\sqrt{-\Delta}$, $(\sqrt{-\Delta})^{-1}\sin t\sqrt{-\Delta}$.

The functions $u_T^{\pm}$, $\lambda^{-1}v_T^{\pm}$ then satisfy the definitions of the packets in Section 4, and shall hereafter be denoted generically by u_T. If (u_0,u_1) is localized at frequency $\lambda \gg \lambda_0$, then the frequency-λ part of the solution u to $\Box_g u = 0$ is approximated in the sense of (4.4), (4.5), (4.6), (4.7), by retaining just the terms in w with frequencies comparable to λ; see [29, Theorem 4.3].

As

$$\langle\Theta^{\pm}(t)(y - x_T(t)),\omega_T\rangle = \langle y - x_T(t),\omega_T(t)\rangle,$$

for any $N \ge 0$ we have

$$\begin{aligned} &\big|(\omega_T(t)\nabla_y)^j(\omega_T(t)^{\perp}\nabla_y)^{\alpha}\ u_T(t,y)\big| \\ &\lesssim \lambda^{j+\frac{|\alpha|}{2}+\frac{n+1}{4}}\frac{1}{(1+\lambda\,|\omega_T(t)\cdot(y-x_T(t))| + \lambda\,|y-x_T(t)|^2)^N}. \end{aligned} \tag{A.2}$$

Thus, $u_T(t,\cdot)$ are concentrated on spatial rectangles of size $\lambda^{-1}\times(\lambda^{\frac{1}{2}})^{n-1}$ that get rotated according to $\omega_T(t)$ as time evolves and are centered around $x_T(t)$. By slight abuse of notation, we will also denote by T this space-time region, called a tube, where u_T is concentrated. For fixed ω and a sign $\pm$, corresponding to the lattice Ξ_λ^ω we obtain a family $\mathcal{T}_{\lambda,\omega}^{\pm}$ of spacetime tubes which are finitely overlapping.

We introduce the null foliation Λ_θ with direction θ associated to the metric $g_{<\lambda^{1/2}}$, and construct a null frame $\{L,\underline{L},E\}$ as before.

The following computation is a variation of the proof of [29, Lemma 3.4].

Lemma A.1. *Lu_T satisfies the decay estimate* (4.1).

Proof. By Lemma 3.6, we can replace L with the operator $\partial_t + \langle a_\xi(t,y,\widehat{\xi}(t,y)),\partial_y\rangle$.

Assume WLOG that $x_T = 0$. Then we have

$$(\partial_t + \langle a_\xi(t,x(t),\widehat{\xi}(t)),\partial_y\rangle)u_T(t,y) = \langle\widehat{\xi}(t), y-x(t)\rangle\langle a_x(t,x(t),\widehat{\xi}(t)),\partial_y\rangle u(t,y)$$
$$- \langle\widehat{\xi}(t)), y-x(t)\rangle\langle\widehat{\xi}(t),\partial_y\rangle u(t,y).$$

Therefore

$$(\partial_t + \langle a_\xi(t,y,\widehat{\xi}(t,y)),\partial_y\rangle)u_T = \langle a_\xi(t,y,\widehat{\xi}(t,y)) - a_\xi(t,x(t),\widehat{\xi}(t,x(t))),\partial_y\rangle u_T$$
$$- \langle a_x(t,x(t),\widehat{\xi}(t,x(t))), y-x(t)\rangle\langle\widehat{\xi}(t,x(t)),\partial_y\rangle u_T$$
$$+ \langle\widehat{\xi}(t), y-x(t)\rangle\langle a_x(t,x(t),\widehat{\xi}(t)),\partial_y\rangle u_T.$$

The third term is $\mu|\langle y-x(t),\widehat{\xi}(t)\rangle|$ times an order 0 packet, therefore is also order 0.

To estimate the first two terms, for simplicity of notation we drop the dependence in t and write $x := x(t)$, $\xi(y) := \xi(t,y)$, $a(y,\widehat{\xi}(y)) := a(t,y,\widehat{\xi}(t,y))$. The first two terms can then be written as

$$\langle a_\xi(y,\widehat{\xi}(y)) - a_\xi(x,\widehat{\xi}(x)) - \langle a_x(x,\widehat{\xi}(x)), y-x\rangle\widehat{\xi}(x),\partial_y\rangle u_T$$
$$= \langle a_\xi(y,\widehat{\xi}(y)) - a_\xi(y,\widehat{\xi}(x)),\partial_y\rangle u_T$$
$$+ \langle a_\xi(y,\widehat{\xi}(x)) - a_\xi(x,\widehat{\xi}(x)) - \langle a_x(x,\widehat{\xi}(x)), y-x\rangle\widehat{\xi}(x),\partial_y\rangle u_T.$$

For the first term, write $a_\xi(y,\widehat{\xi}(y)) - a_\xi(y,\widehat{\xi}(x)) = a_{\xi\xi}(y,\widehat{\xi}(x))[\widehat{\xi}(y)-\widehat{\xi}(x)] + O(|\widehat{\xi}(y)-\widehat{\xi}(x)|^2)$. The remainder is $O(|y-x|^2)$, while the linear term has size $O(|y-x|)$ and is orthogonal to $\widehat{\xi}(x)$ by homogeneity:

$$\langle a_{\xi\xi}(y,\widehat{\xi}(x))[\widehat{\xi}(y)-\widehat{\xi}(x)],\widehat{\xi}(x)\rangle = \langle\widehat{\xi}(y)-\widehat{\xi}(x), a_{\xi\xi}(y,\widehat{\xi}(x))\widehat{\xi}(x)\rangle = 0.$$

Therefore

$$\langle a_\xi(y,\widehat{\xi}(y)) - a_\xi(y,\widehat{\xi}(x)),\partial_y\rangle u_T = \mu^{1/2}|y-x|v_T + \mu|y-x|^2 w_T,$$

where v_T and w_T satisfy order 0 decay, hence is also order 0. Finally, the remaining term is acceptable since

$$\langle\widehat{\xi}(x), a_\xi(y,\widehat{\xi}(x)) - a_\xi(x,\widehat{\xi}(x)) - \langle a_x(x,\widehat{\xi}(x)), y-x\rangle\widehat{\xi}(x)\rangle$$
$$= a(y,\widehat{\xi}(x)) - a(x,\widehat{\xi}(x)) - \langle a_x(x,\widehat{\xi}(x)), y-x\rangle$$
$$= O(\|\partial^2 g(t)\|_{L^\infty_x}|y-x|^2),$$

while the component orthogonal to $\widehat{\xi}(x)$ has size $O(|y-x|)$. □

A.1. **General metrics.** Smith's construction adapts, with minor modifications, to more general equations of the form

$$g^{\alpha\beta}\partial_\alpha\partial_\beta u = 0,\ u[0] = (u_0,u_1).$$

Here we will harmlessly assume that $g^{00} = 1$.

For each frequency $\lambda \gg 1$, factor the principal symbol of the mollified operator $g^{\alpha\beta}_{<\lambda^{1/2}}\partial_\alpha\partial_\beta$ as $(\tau - a^+)(\tau - a^-)$, where

$$a^\pm(t,x,\xi) := g^{0b}_{<\lambda^{1/2}}\xi_b \pm \left[(g^{0b}_{<\lambda^{1/2}}\xi_b)^2 + g^{ab}_{<\lambda^{1/2}}\xi_a\xi_b\right]^{1/2}.$$

are smooth convex/concave symbols in view of the hyperbolicity condition.

Then the analogue of (A.1) is

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = \partial_\xi a^\pm(t,x,\omega) \\ \dfrac{\mathrm{d}\omega}{\mathrm{d}t} = -\partial_x a^\pm(t,x,\omega) \pm \langle \omega, \partial_x a^\pm(t,x,\omega)\rangle \omega, \\ \dfrac{\mathrm{d}\Theta^\pm}{dt} = -\Theta^\pm[\omega \otimes a_x^\pm(t,x,\omega) - (a_x^\pm)(t,x,\omega)\otimes\omega], \end{cases} \tag{A.3}$$

and the analogues of $u_T^\pm$ and $v_T^\pm$ are furnished by the following construction.

Lemma A.2. *Fix (x_T,ω_T), and let φ_T be as before. Then for all sufficiently large frequencies $\lambda \gg 1$, there exist functions $\varphi_T^\pm, \psi_T^\pm$ with similar properties as φ_T such that*

$$u_T^\pm(t,y) := \varphi_T^\pm(\Theta^\pm(t)(y - x_T(t))), \quad v_T^\pm(t,y) := \psi_T^\pm(\Theta^\pm(t)(y-x_T(t))),$$

satisfy the same estimates as before, and

$$\begin{cases} u_T^+(0) + u_T^-(0) = \varphi_T + \lambda^{-1/2}\widetilde{\varphi_T} \\ \partial_t u_T^+(0) + \partial_t u_T^-(0) = \lambda^{-1/2}\widetilde{\varphi_T} \end{cases} \qquad \begin{cases} v_T^+(0) - v_T^-(0) = \lambda^{-1/2}\widetilde{\varphi_T} \\ \partial_t v_T^+(0) - \partial_t v_T^-(0) = \varphi_T + \lambda^{-1/2}\widetilde{\varphi_T}, \end{cases},$$

where $\tilde{\varphi_T}$ denotes generic function with similar smoothness and decay as φ_T.

Proof. Without loss of generality consider the first system. Using equations (A.3), and writing $\Phi_T = (\varphi_T^+, \varphi_T^-)^*$, $F_T = (\varphi_T, 0)^*$, the system takes the form

$$[M_T(D) + R_T(X,D)]\Phi_T = F_T + O(\lambda^{-1/2}).$$

where

$$M_T(D) = \begin{pmatrix} 1 & 1 \\ -\langle a_\xi^+(0,z_T), \partial_x\rangle & -\langle a_\xi^-(0,z_T),\partial_x\rangle \end{pmatrix}, \quad z_T = (x_T,\omega_T)$$
$$R_T(X,D) = \begin{pmatrix} 0 & 0 \\ \langle \dot\Theta^+(0)(x-x_T),\partial_x\rangle & \langle \dot\Theta^-(0)(x-x_T),\partial_x\rangle \end{pmatrix}.$$

When ξ is restricted to a small sector centered at ω_T the main term is elliptic:

$$|\det M_T(\xi)| = |\langle (a_\xi^+(0,z_T) - a_\xi^-(0,z_T)), \xi\rangle| \sim |\xi|(a^+ - a^-)(0,z_T) \sim |\xi|.$$

Further, in view of spatial localization the operator $R_T(X,D)$ has order $\frac{1}{2}$ when acting on functions of the form $\widetilde{\varphi_T}$. Thus it suffices to set $\Phi_T := M_T(D)^{-1}F_T$. □

For each frequency $\lambda > 1$ and time t, let $\Phi_\lambda(t) : L^2 \times \lambda^{-1}L^2 \to L^2 \times \lambda^{-1}L^2$ denote the operator

$$\Phi_\lambda(t)\begin{pmatrix} f \\ g \end{pmatrix} := \sum_{T\in\mathcal{T}_\lambda} \Phi_T(t)\begin{pmatrix} f_T \\ g_T\end{pmatrix},$$
$$\Phi_T(t) := \begin{pmatrix} u_T^+ + u_T^- & v_T^+ - v_T^- \\ \partial_t u_T^+ + \partial_t u_T^- & \partial_t v_T^+ - \partial_t v_T^- \end{pmatrix}(t), \quad \begin{pmatrix} f_T \\ g_T\end{pmatrix} = \begin{pmatrix} \langle f,\varphi_T\rangle \\ \langle g,\varphi_T\rangle\end{pmatrix},$$

and set

$$\Phi(t) := \sum_{\lambda>\lambda_0} \Phi_\lambda(t) P_\lambda + \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} P_{\leq\lambda_0}.$$

Then as $\|\Phi(0) - I\|_{L^2\times H^{-1}\to L^2\times H^{-1}} = O(\lambda_0^{-1/2})$, by choosing λ_0 sufficiently large the operator $\Phi(0)$ is invertible on $L^2 \times H^{-1}$. Hence

$$\widetilde{\Phi}(t) := \Phi(t)\Phi(0)^{-1}$$

is a parametrix in the sense of Property 4.1.

Appendix B. Microlocal analysis tools

B.1. **Symbols and phase space metrics.** We begin by briefly reviewing the framework of Hormander metrics; for further details consult [12, Section 18.4].

Let g be a slowly varying metric on phase space $W := T^*\mathbb{R}^n = \mathbb{R}^n_x \times (\mathbb{R}^n)^*_\xi$. This induces norms on tensors in the usual manner; in particular, if $l \in T_zW \times \cdots T_zW \to \mathbb{R}$ is k-linear for some $z \in (x, \xi)$, set

$$|l_z|^g_k := \sup_{0\neq t_j\in T_zW} \frac{|l_z(t_1, \ldots, t_k)|}{\prod_{j=1}^k g_z(t_j)^{1/2}}.$$

Definition B.1. If m is a slowly varying function on W, write $S(m, g)$ for the space of functions u on W such that

$$\sup_z |\nabla^k u|^g_k(z)/m(z) \leq C_k \quad \text{for all } z \in W.$$

Definition B.2. A map $\chi : W \to W$ is g-smooth if the pullback $\chi^* S(1, g) \subset S(1, g)$.

By the chain rule and induction, this definition is equivalent to requiring that

$$|D^k\chi(z; t_1, \ldots, t_k)|_{g_{\chi(z)}} \leq C_k \prod_{j=1}^k g_z(t_j)^{\frac{1}{2}} \text{ for all } t_j \in T_zW, \text{ uniformly in } z.$$

B.2. **Pseudo-differential calculus.** Suppose A be the quadratic form on W given by $A(y, \eta) = \langle y, \eta\rangle$, and for $\alpha \leq 1$ let g_α denote the phase space metric (5.2). Let U_α denote the phase space region

$$U_\alpha := \{(x, \xi) : |\xi| \geq \alpha^{-2}\}$$

In the language of Hormander 18.4, the phase space metric g is *A-temperate* in U_α in the sense that there exist constants C, N

$$g_w(t) \leq C g_z(t)(1 + g^A_w(z - w))^N, \text{ for all } z, w \in U_\alpha,$$

where for a general quadratic form A one has

$$g^A_w(A\zeta) := \sup_{\eta\in \mathrm{Im}(A)} \frac{|\langle \eta, \zeta\rangle|^2}{g_w(\eta)}$$

and $g^A_w(\beta) := \infty$ for $\beta \notin \mathrm{Im}(A)$. In the present case,

$$g^A_{(x,\xi)}(y, \eta) = g^{-1}_{(x,\xi)}(\eta, y) = \alpha^4|\langle \eta, \widehat{\xi}\rangle|^2 + \alpha^2|\eta \wedge \widehat{\xi}|^2 + |\xi|^2|\langle y, \widehat{\xi}\rangle|^2 + \alpha^2|\xi|^2|y \wedge \widehat{\xi}|^2,$$

where we write $\widehat{\xi} := \xi/|\xi|$.

Similarly, a slowly varying function m is *A-temperate* with respect to $z \in U_\alpha$ if

$$m(w) \leq Cm(z)(1 + g_w^A(z - w))^N \text{ for all } w \in U_\alpha.$$

For a symbol a, we define the corresponding pseudo-differential operator using right-quantization

$$a(X, D) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{i\langle x-y,\xi\rangle} a(x, \xi)\, d\xi,$$

and write $OPS(m, g)$ for the quantizations of symbols in $S(m, g)$.

Recall that if a and b are symbols, then formally

$$a(X, D)b(X, D) = (a \circ b)(X, D),$$

where

$$a \circ b(x, \xi) := e^{i\langle D_\eta, D_y\rangle}[a(x, \eta)b(y, \xi)]|_{\substack{y=x\\ \eta=\xi}}.$$

In particular, we have the first order and second order symbol expansions:

$$\begin{aligned} a \circ b &= ab + \frac{1}{i}\int_0^1 r_{1,s}\, ds \\ &= ab + \frac{1}{i}a_\xi b_x - \frac{1}{2}\int_0^1 r_{2,s}\, ds, \end{aligned} \tag{B.1}$$

where

$$r_{j,s}(x, \xi) = e^{is\langle D_y, D_\eta\rangle}\langle \partial_y, \partial_\eta\rangle^j [a(x, \eta)b(y, \xi)]_{\substack{y=x\\ \eta=\xi}}.$$

The remainder can be estimated as in [12, Section 18.4]. For a real parameter $t \geq 0$, define

$$a \circ_t b(x, \xi) := e^{it\langle D_\eta, D_y\rangle}[a(x, \eta)b(y, \xi)]|_{\substack{y=x\\ \eta=\xi}}.$$

Lemma B.3. *Suppose m_1, m_2 are slowly varying and A-temperate in the region U_α. If $a \in S(m_1, g_\alpha)$, $b \in S(m_2, g_\alpha)$ are symbols supported in U_α, then*

$$a \circ_t b \in S(m_1 m_2, g_\alpha)$$

with constants uniform in $0 \leq t \leq 1$.

Proof sketch. Follow the proof of Theorem 18.4.10 and Proposition 18.5.2 in Hormander [12]. The main point is that estimates for the Gauss transform $e^{itA(D)}$ only improve when $t \leq 1$ since $g^{tA} = t^{-2}g^A \geq g^A$.

Let B_t : denote the quadratic form on $W \oplus W$ defined by $B_t(x, \eta), (y, \xi)) := t\langle y, \eta\rangle$, then one needs to verify that

- The metric $G := g_\alpha \oplus g_\alpha$ is slowly varying and B_t-temperate with respect to the diagonal (x, ξ, x, ξ) in $U_\alpha \times U_\alpha$, and $G \leq G^{B_t}$.
- The weight function $M := m_1 \otimes m_2$ is slowly varying and temperate with respect to B_t at the diagonal where $|\xi| \geq \alpha^{-2}$.

□

We hereafter denote by $S(m, g_\alpha)$ those symbols supported in the region $|\xi| \geq \alpha^{-2}/8$. The corresponding operators $a(X, D) \in OPS(m, g_\alpha)$ accept input frequencies $\gtrsim \alpha^{-2}$. The previous lemma shows that $OPS(m_1, g_\alpha) \cdot OPS(m_2, g_\alpha) \subset OPS(m_1 m_2, g_\alpha)$.

In our applications we encounter a slightly more complicated class of symbols associated to two angular scales. For $\beta \geq \alpha$, let

$$S^1_\beta(m, g_\alpha) := \{\phi \in S(m, g_\alpha) : \partial_x \phi \in S(m, g_\alpha),\ \partial_\xi \phi \in S((\beta|\xi|)^{-1} m, g_\alpha)\}.$$

In view of Lemma B.3 and the identities

$$\partial_x(a \circ b) = (\partial_x a) \circ b + a \circ (i\partial_x b),$$
$$\partial_\xi(a \circ b) = (i\partial_\xi a) \circ b + a \circ (\partial_\xi b),$$

one sees that

$$OPS^1_\beta(m_1, g_\alpha) \cdot OPS^1_\beta(m_2, g_\alpha) \subset OPS^1_\beta(m_1 m_2, g_\alpha).$$

However, we do not quite have the usual pseudo-differential calculus since $OPS(m, g_\alpha)$ is not closed under taking adjoints. One remedy is to consider the subclass of operators in $OPS(m, g_\alpha)$ which output at frequencies $\geq \alpha^{-2}$.

Lemma B.4. *Suppose m is slowly varying and temperate with respect to g_α. If $\phi \in S(m, g_\alpha)$ satisfies $\phi(X, D) = S_{\geq \alpha^{-2}}(D)\phi(X, D)S_{\geq \alpha^{-2}}(D)$ then*

$$\phi^*(x, \xi) = [e^{i\langle D_y, D_\eta\rangle}\phi(y, \eta)]_{\substack{y=x \\ \eta=\xi}} \in S(m, g_\alpha)$$

and is supported in $|\xi| \geq \alpha^{-2}/8$.

This leads to the main L^2 estimate:

Lemma B.5. [10, Theorem 4.8] *If $a \in S(1, g_\alpha)$ in addition satisfies*

$$|S_{<\lambda}(D_x + \xi)a(x, \xi)| \leq C_N \Big(\frac{\lambda}{\langle\xi\rangle}\Big)^N, 1 \leq \lambda \leq \langle\xi\rangle,$$

then $a(X, D)$ is continuous on L^2. In particular, the conclusion holds for operators of the form

$$a(X, D) = S_{>\lambda/8}(D)a(X, D)S_\lambda(D), \quad \lambda \geq \alpha^{-2}.$$

Proof. For the last statement, note that since

$$\widehat{a(X, D)u}(\eta) = \int \widehat{a}(\eta - \xi, \xi)\widehat{u}(\xi)\, d\xi,$$

the hypothesis implies that $S_{<\lambda/8}(D_x + \xi)a(x, \xi) = 0$. □

We also need bounds for certain pseudo-differential operators which are strongly localized in input frequency but do not quite satisfy the hypotheses of the previous lemma. For a direction field $x \mapsto \Theta(x) \in S^{n-1}$ on $\mathbb{R}^n$, let

$$m = m_\Theta(x, \xi) := \langle\alpha^{-1}(\widehat{\xi} - \Theta(x))\rangle^{-1}, \quad \widehat{\xi} := \xi/|\xi|.$$

To express the angular localization of various symbols it will be convenient to introduce the following

Notation. For a weight function m, write $S(m^\infty, g_\alpha) := \bigcap_N S(m^N, g_\alpha)$.

Lemma B.6. *Suppose $x \mapsto \Theta(x) \in S^{n-1} \subset \mathbb{R}^n$ is Lipschitz, and let $\phi \in S(m^\infty, g_\alpha)$ be supported in an annulus $|\xi| \sim \lambda \geq \alpha^{-2}$. Then $\phi(X, D)$ is bounded on L^p for all $1 \leq p \leq \infty$.*

Proof. We use Schur's test on the kernel $K(x,y) = (2\pi)^{-n} \int e^{i\langle x-y,\xi\rangle} \phi(x,\xi)\, d\xi = \mathcal{F}_2^{-1}\phi(x, x-y)$ of $\phi(X,D)$. For fixed x, $\xi \mapsto \phi(x,\xi)$ is a Schwartz function with height 1 and adapted to the sector $|\widehat{\xi} - \Theta(x)| \lesssim \alpha$, $|\xi| \sim \lambda$. Therefore

$$|K(x,y)| \lesssim \lambda(\alpha\lambda)^{n-1}\langle \lambda|\langle x-y, \Theta(x)\rangle| + \alpha\lambda|x-y\wedge\Theta(x)|\rangle^{-N} \text{ for any } N,$$

so

$$\sup_x \int |K(x,y)|\, dy < \infty.$$

To evaluate

$$\sup_y \int |K(x,y)|\, dx,$$

decompose $\phi = \sum \phi_\nu^\omega$, where ω ranges over a partition of the annulus $|\xi| \sim \lambda$ into $\lambda \times (\alpha\lambda)^{n-1}$ sectors, and for each ω, ν ranges over a partition of space into parallel $\alpha^2 \times (\alpha)^{n-1}$ parallelpipeds R_ν^ω with orientation ω. The kernel decomposes as

$$K(x,y) = \sum_\omega \sum_\nu \mathcal{F}_2^{-1}\phi_\nu^\omega(x, x-y) = \sum_\omega \sum_\nu K_\nu^\omega(x,y).$$

Then

$$|K_\nu^\omega(x,y)| \lesssim 1_{R_\nu^\omega}(x)\langle \alpha^{-1}(\omega - \Theta(x_\nu^\omega))\rangle^{-N} \lambda(\alpha\lambda)^{n-1}\langle \lambda|\langle x-y,\omega\rangle| + \alpha\lambda|x-y\wedge\omega|\rangle^{-N}.$$

For a constant c depending on the Lipschitz constant of Θ,

$$\begin{aligned}
\int |K(x,y)|\, dx &\lesssim \sum_{|x_\nu^\omega - y| \leq c(|\omega-\Theta(y)|+\alpha)} \langle \alpha^{-1}(\omega - \Theta(y))\rangle^{-N} \\
&\times \int_{R_\nu^\omega} \lambda(\alpha\lambda)^{n-1}\langle \lambda|\langle x-y,\omega\rangle| + \alpha\lambda|x-y\wedge\omega|\rangle^{-2N}\, dx \\
&+ \sum_{|x_\nu^\omega - y| > c(|\omega-\Theta(y)|+\alpha)} \langle \alpha^{-1}(\omega - \Theta(x_\nu^\omega))\rangle^{-N}\langle \alpha\lambda|\omega - \Theta(y)|\rangle^{-N} \\
&\times \int_{R_\nu^\omega} \lambda(\alpha\lambda)^{n-1}\langle \lambda|\langle x-y,\omega\rangle| + \alpha\lambda|x-y\wedge\omega|\rangle^{-N}\, dx \\
&\lesssim \sum_\omega \langle \alpha^{-1}(\omega - \Theta(y))\rangle^{-N} < \infty.
\end{aligned}$$

□

Appendix C. An angular partition of unity

Throughout this discussion we fix a minimum angular scale $\alpha_\mu \ll 1$.

Choose $0 \le \eta \in C_0^\infty\big((-1,1)\big)$ with $\int \eta = 1$, and let $\eta_h = h^{-1}\eta(h^{-1}\cdot)$. For each $\theta \in \Omega_\alpha$, with $\theta = [a.b)$, define the scale functions

$$h_1(\xi) = \begin{cases} \frac{1}{8}\alpha, & \xi \le \frac{a+b}{2}, \\ \frac{1}{16}\alpha, & \xi > \frac{a+b}{2} \end{cases} \qquad h_2(\xi) = \begin{cases} \frac{1}{16}\alpha, & \xi \le \frac{a+b}{2}, \\ \frac{1}{8}\alpha, & \xi > \frac{a+b}{2} \end{cases}$$

$$h_3(\xi) = \begin{cases} \frac{1}{16}\alpha, & \xi \le \frac{a+b}{2}, \\ \frac{1}{16}\alpha, & \xi > \frac{a+b}{2} \end{cases} \qquad h_4(\xi) = \begin{cases} \frac{1}{8}\alpha, & \xi \le \frac{a+b}{2}, \\ \frac{1}{8}\alpha, & \xi > \frac{a+b}{2} \end{cases}$$

and for $k = 1, 2, 3, 4$, define $\phi_\theta^{\alpha,k}$ by mollifying the characteristic functions of θ on a position-dependent scale:

$$\phi_\theta^{\alpha,k}(\xi) := 1_\theta * \eta_{h_k(\xi)}(\xi).$$

Then each $\phi_\theta^{\alpha,k}$ is smooth on the α scale and supported in a $\frac{1}{8}|\theta|$ neighborhood of the interval θ, and there exist $c_1, c_2 > 0$ such that

$$c_1 \le \sum_{\theta \in \Omega_\alpha} \phi_\theta^{\alpha,k(\theta)} \le c_2 \tag{C.1}$$

for any sequence $k(\theta) \in \{1, 2, 3, 4\}$.

Proposition C.1. *Fix a small dyadic number $\alpha_\mu < 1/4$. For each $\omega \in \Omega_{\alpha_\mu}$, there exist dyadic intervals $\theta_0 = \omega, \theta_1, \dots$ on the circle from the families Ω_α with, $|\theta_j - \omega| \sim \alpha_j$ when $j \ge 1$, such that we have the (fixed time) partition of unity*

$$1 = \sum_j \phi_{\theta_j}^{\alpha_j, k(\omega)},$$

where $k(\omega) \in \{1, 2, 3, 4\}$, and the scales α_j satisfy

- *$\frac{1}{2}\alpha_{j-1} \le \alpha_j \le 2\alpha_{j-1}$, and*
- *at most $O(1)$ consecutive α_j's are equal.*

The families Ω_α are independent of ω.

Before the proof, we consider

Lemma C.2. *Given $\omega \in \Omega_{\alpha_0}$, there exists a sequence $\omega = \theta_0, \theta_1, \theta_2, \dots$ of dyadic intervals in $\mathbb{R}$ with the following properties:*

- *θ_j is right-adjacent to θ_{j-1}.*
- *$2|\theta_{j-1}| \le |\theta_j| \le |\theta_{j+1}|$ for all j.*
- *If $|\theta_{j-1}| = |\theta_j|$, then $|\theta_{j+1}| = 2|\theta_j|$.*
- *$\sum_{j=1}^{J} |\theta_j| \le 4|\theta_J|$ for all J.*

The analogous statement with "right" and "left" interchanged also holds.

Proof. The idea of the construction is to "double whenever possible while moving right."

Using the usual tree terminology, define the sequence θ_j inductively as follows:

- If θ_j is the left-child of its parent, let θ_{j+1} be its sibling.
- Else let θ_{j+1} be the right neighbor of θ_j's parent.

Then $|\theta_{j+1}| = |\theta_j|$ in the first case and $|\theta_{j+1}| = 2|\theta_j|$ in the second. Since each interval has only two children, there cannot be more than two consecutive intervals of the same width.

The inequality $\sum_{j=1}^{J} |\theta_j| \le 4|\theta_J|$ is verified inductively. Assume it holds for all smaller J. If $|\theta_{J+1}| = 2|\theta_J|$, then

$$\sum_{j=1}^{J+1} |\theta_j| \le 4|\theta_J| + 2|\theta_J| \le 4|\theta_{J+1}|,$$

while if $|\theta_{J+1}| = |\theta_J|$, then $|\theta_J| = 2|\theta_{J-1}|$ and we have

$$\sum_{j=1}^{J+1} |\theta_j| \le 4|\theta_{J-1}| + 2|\theta_{J-1}| + 2|\theta_{J-1}| \le 4|\theta_{J+1}|.$$

□

Proof of Prop. C.1. Starting with ω, we construct dyadic intervals $\theta_1^r, \theta_2^r, \dots$ and $\theta_1^l, \theta_2^l, \dots$ to the right and left of ω, respectively, according to the lemma. Choose J^r and J^l maximal such that

$$\sum_{j=1}^{J^r} |\theta_j^r| \le \frac{1}{2} - |\omega|, \quad \sum_{j=1}^{J^l} |\theta_j^l| \le \frac{1}{2} - |\omega|.$$

By the lemma one must have

$$\frac{1}{8} - \frac{|\omega|}{4} \le |\theta_{J^r}^r|, |\theta_{J^l}^l| \le \frac{1}{2} - |\omega|;$$

that is,

$$|\theta_{J^r}^r|, |\theta_{J^l}^l| \in \left\{\frac{1}{16}, \frac{1}{8}, \frac{1}{4}\right\}.$$

Assume with loss of generality that $|\theta_{J^l}^l| \le |\theta_{J^r}^r|$.

- If $|\theta_{J^l}^l| = \frac{1}{16}$, $|\theta_{J^r}^r| = \frac{1}{4}$, then $\theta_{J^r}^r$ and $\theta_{J^l}^l$ are separated modulo 1 by less than 10 dyadic intervals of width $\frac{1}{16}$ (the worst case being when $|\theta_{J^l+1}^l| = \frac{1}{8}$ and $|\theta_{J^r+1}^r| = \frac{1}{2}$). We replace $\theta_{J^r}^r$ by its two children $\theta_{J^r}^{r,1}$, $\theta_{J^r}^{r,2}$ of width $\frac{1}{8}$, and reindex the intervals by defining $\theta_{J^r}^r := \theta_{J^r}^{r,1}$, $\theta_{J^r+1}^r := \theta_{J^r}^{r,2}$, and replacing J^r by $J^r + 1$.
- If $|\theta_{J^l}^l|$ and $|\theta_{J^r}^r|$ differ by at most one dyadic scale, then they are separated modulo 1 by at less than 6 dyadic intervals of width $|\theta_{J^l}^l|$.

Let $\theta_1', \dots, \theta_n' \in \Omega_{|\theta_{J^l}^l|}$, $n < 10$ be the intervening intervals. Projecting the intervals to $\mathbb{R}/\mathbb{Z}$, we relabel the sequence

$$\theta_0, \theta_1^r, \dots, \theta_{J^r}^r, \theta_1', \dots, \theta_n', \theta_{J^l}^l, \dots, \theta_1^l$$

as

$$\theta_0, \theta_1, \dots, \theta_{M-1},$$

and write

$$\alpha_j := |\theta_j|.$$

Then, interpreting the indices modulo M, we have

- $\frac{1}{2}\alpha_{j-1} \le \alpha_j \le 2\alpha_{j-1}$, and
- at most $O(1)$ consecutive α_j's are equal.

For each $\theta_j = [a_j, b_j) \pmod 1$, define the scales

$$h_j(\xi) = \begin{cases} \frac{\alpha_j}{8}, & \alpha_j \le \alpha_{j+1} \\ \frac{\alpha_j}{16}, & \alpha_j > \alpha_{j+1} \end{cases}$$

and define $\phi_{\theta_j}^{\alpha_j}$ by mollifying the characteristic functions of θ_j near the junction points.

$$\phi_{\theta_j}^{\alpha_j}(\xi) = \begin{cases} 1_{\theta_j} * \eta_{h_{j-1}}(\xi), & a_j - \frac{\alpha_j}{8} \le \xi \le \frac{a_j+b_j}{2}, \\ 1_{\theta_j} * \eta_{h_j}(\xi), & \frac{a_j+b_j}{2} < \xi \le b_j + \frac{\alpha_j}{8}, \\ 0, & \text{otherwise} \end{cases}$$

This function has one of the four forms asserted in the proposition, and

$$1 = \sum_{j=0}^{M-1} \phi_{\theta_j}^{\alpha_j}$$

since the same is true for the sum of characteristic functions 1_{θ_j} and the mollification scale is the same near the boundary between θ_j and θ_{j+1}. □

References

[1] S. Alinhac. *Geometric analysis of hyperbolic differential equations: an introduction*, volume 374 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2010.

[2] J.-M. Bony. Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires. *Ann. Sci. École Norm. Sup. (4)*, 14(2):209–246, 1981.

[3] N. Burq, P. Gérard, and N. Tzvetkov. Bilinear eigenfunction estimates and the nonlinear Schrödinger equation on surfaces. *Inventiones mathematicae*, 159(1):187–223, Jan 2005.

[4] T. Candy and S. Herr. On the Division Problem for the Wave Maps Equation. *ArXiv e-prints*, July 2018.

[5] A. Córdoba and C. Fefferman. Wave packets and Fourier integral operators. *Comm. Partial Differential Equations*, 3(11):979–1005, 1978.

[6] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bulletin des Sciences Mathématiques*, 136(5):521–573, 2012.

[7] H. F Smith and C. Sogge. Null form estimates for (1/2,1/2) symbols and local existence for a quasilinear Dirichlet-wave equation. *Annales Scientifiques de l'École Normale Supérieure*, 33:485–506, 05 2000.

[8] D. Foschi and S. Klainerman. Bilinear space-time estimates for homogeneous wave equations. *Ann. Sci. École Norm. Sup. (4)*, 33(2):211–274, 2000.

[9] D.-A. Geba. A remark on an equation of wave maps type with variable coefficients. *Mathematical Research Letters*, 16(3):395–404, 2009.

[10] D.-A. Geba and D. Tataru. A phase space transform adapted to the wave equation. *Communications in Partial Differential Equations*, 32(7):1065–1101, 2007.

[11] D.-A. Geba and D. Tataru. Gradient nlw on curved background in 4+ 1 dimensions. *International Mathematics Research Notices*, 2008, 2008.

[12] L. Hörmander. *The analysis of linear partial differential operators. III*. Classics in Mathematics. Springer, Berlin, 2007. Pseudo-differential operators, Reprint of the 1994 edition.

[13] J. Jost and J. Jost. *Riemannian geometry and geometric analysis*, volume 42005. Springer, 2008.

[14] S. Klainerman. A commuting vectorfields approach to strichartz-type inequalities and applications to quasi-linear wave equations. *International Mathematics Research Notices*, 2001(5):221–274, 2001.

[15] S. Klainerman and M. Machedon. Space-time estimates for null forms and the local existence theorem. *Communications on Pure and Applied Mathematics*, 46(9):1221–1268, 1993.

[16] S. Klainerman and M. Machedon. Estimates for null forms and the spaces $H_{s,\delta}$. *International Mathematics Research Notices*, 1996(17):853–865, 1996.

[17] S. Klainerman and I. Rodnianski. On the global regularity of wave maps in the critical Sobolev norm. *Internat. Math. Res. Notices*, (13):655–677, 2001.
[18] S. Klainerman, I. Rodnianski, and T. Tao. A physical space approach to wave equation bilinear estimates. *Journal d'Analyse Mathématique*, 87(1):299–336, 2002.
[19] S. Klainerman and S. Selberg. Remark on the optimal regularity for equations of wave maps type. *Communications in Partial Differential Equations*, 22(5-6):99–133, 1997.
[20] S. Klainerman and S. Selberg. Bilinear estimates and applications to nonlinear wave equations. *Communications in Contemporary Mathematics*, 4(02):223–295, 2002.
[21] J. Krieger. Global regularity of wave maps from $\mathbb{R}^{2+1}$ to $\mathbb{H}^2$. Small energy. *Comm. Math. Phys.*, 250(3):507–580, 2004.
[22] J. Krieger and W. Schlag. *Concentration compactness for critical wave maps*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2012.
[23] A. Lawrie. The cauchy problem for wave maps on a curved background. *Calculus of Variations and Partial Differential Equations*, 45(3-4):505–548, 2012.
[24] A. Lawrie, S.-J. Oh, and S. Shahshahani. The cauchy problem for wave maps on hyperbolic space in dimensions $d \geq 4$. *International Mathematics Research Notices*, page rnw272, 2016.
[25] Z. Li, X. Ma, and L. Zhao. Asymptotic stability of small energy harmonic maps under the wave map on 2d hyperbolic space. *arXiv preprint arXiv:1703.05207*, 2017.
[26] H. Lindblad. Counterexamples to local existence for semi-linear wave equations. *Amer. J. Math.*, 118(1):1–16, 1996.
[27] A. Nahmod, A. Stefanov, and K. Uhlenbeck. On the well-posedness of the wave map problem in high dimensions. *Comm. Anal. Geom.*, 11(1):49–83, 2003.
[28] J. M. I. Shatah and M. Struwe. *Geometric wave equations*, volume 2. American Mathematical Soc., 1998.
[29] H. F. Smith. A parametrix construction for wave equations with $c^{1,1}$ coefficients. *Annales de l'institut Fourier*, 48(3):797–836, 1998.
[30] H. F. Smith and D. Tataru. Sharp local well-posedness results for the nonlinear wave equation. *Annals of mathematics*, pages 291–366, 2005.
[31] C. D. Sogge. On local existence for nonlinear wave equations satisfying variable coefficient null conditions. *Communications in partial differential equations*, 18(11):1795–1821, 1993.
[32] J. Sterbenz and D. Tataru. Energy dispersed large data wave maps in $2+1$ dimensions. *Comm. Math. Phys.*, 298(1):139–230, 2010.
[33] J. Sterbenz and D. Tataru. Regularity of wave-maps in dimension 2+1. *Comm. Math. Phys.*, 298(1):231–264, 2010.
[34] T. Tao. Global regularity of wave maps. II. Small energy in two dimensions. *Comm. Math. Phys.*, 224(2):443–544, 2001.
[35] T. Tao. *Nonlinear dispersive equations*, volume 106 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006. Local and global analysis.
[36] T. Tao. Global regularity of wave maps III-VI. *ArXiv e-prints*, 2008.
[37] D. Tataru. The X^s_θ and unique continuation for solutions to the semilinear wave equation. *Communications in Partial Differential Equations*, 21(5-6):841–887, 1996.
[38] D. Tataru. Strichartz estimates for operators with nonsmooth coefficients and the nonlinear wave equation. *Amer. J. Math.*, 122(2):349–376, 2000.
[39] D. Tataru. On global existence and scattering for the wave maps equation. *Amer. J. Math.*, 123(1):37–77, 2001.
[40] D. Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. II. *Amer. J. Math.*, 123(3):385–423, 2001.
[41] D. Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. III. *J. Amer. Math. Soc.*, 15(2):419–442, 2002.
[42] D. Tataru. Null form estimates for second order hyperbolic operators with rough coefficients. *Contemporary Mathematics*, 320:383–410, 2003.

[43] D. Tataru. Rough solutions for the wave maps equation. *American Journal of Mathematics*, 127(2):293–377, 2005.

[44] D. Tataru and D.-A. Geba. Dispersive estimates for wave equations. *Comm. Partial Differential Equations*, 30(4-6):849–880, 2005.

Department of Mathematics, Johns Hopkins University
E-mail address: cgavrus1@jhu.edu

Department of Mathematics, University of California, Berkeley
E-mail address: cjao@math.berkeley.edu

Department of Mathematics, University of California, Berkeley
E-mail address: tataru@math.berkeley.edu