DISPERSIVE ESTIMATES FOR WAVE EQUATIONS WITH ROUGH COEFFICIENTS

DANIEL TATARU AND DAN-ANDREI GEBA

Abstract. We obtain a multiscale wave packet representation for the fundamental solution of the wave equation whose coefficients satisfy $\partial^2 g \in L^1_t L^\infty_x$. This leads to pointwise and weighted L^p bounds on the fundamental solution and also to a proof of dispersive estimates for such operators.

1. Introduction

The kernel of the forward fundamental solution for the wave operator $\Box$ in $\mathbb{R}^{n+1}$ has the form

$$K(t,x) = c_n 1_{t>0} \begin{cases} (t^2 - x^2)_+^{-\frac{n-1}{2}} & n \text{ even} \\ \\ \delta^{(\frac{n-3}{2})}(t^2 - x^2) & n \text{ odd} \end{cases}$$

In odd dimension it is supported on the characteristic cone

$$K = \{t = |x|\}$$

In even dimension it is supported inside the cone, but its singular support is still on the cone.

Solutions for the wave equation with smooth compactly supported initial data cannot stay concentrated for a long time. Instead they will spread out along the cone and decay at a rate of $t^{-\frac{n-1}{2}}$. One often refers to this as the dispersive phenomena. Gaining a good understanding of it is of special interest in the study of semilinear and quasilinear wave equations, as it is the main factor which limits the strength of nonlinear interactions.

A quantitative way of measuring the dispersion of waves for the homogeneous wave equation

$$\Box u = 0 \quad u(0) = u_0 \quad \partial_t u(0) = u_1 \tag{1}$$

in $\mathbb{R}^{n+1}$ takes advantage of the pointwise decay of the fundamental solution. Then one easily obtains a dispersive inequality roughly of the type

$$\|\nabla u(t)\|_{L^\infty} \lesssim t^{-\frac{n-1}{2}} \| |D_x|^{\frac{n+1}{2}} \nabla u(0)\|_{L^1} \tag{2}$$

To make this estimate correct one can either replace L^∞ by the BMO space, or L^1 by the Hardy norm H_1, or restrict the Fourier transform of the initial data (u_0, u_1) to a dyadic shell $\{2^k \leq |\xi| \leq 2^{k+1}\}$ for some fixed $k \in \mathbb{Z}$.

Interpolating between (2) and the L^2 conservation of energy

$$\|\nabla u(t)\|_{L^2} = \|\nabla u(0)\|_{L^2} \tag{3}$$

we obtain the intermediate bounds

$$\|\nabla u(t)\|_{L^p} \lesssim t^{-\frac{n-1}{2}(\frac{1}{p'}-\frac{1}{p})} \| |D_x|^{\frac{n+1}{2}(\frac{1}{p'}-\frac{1}{p})} \nabla u(0)\|_{L^{p'}}, \quad 2 \leq p \leq \infty \tag{4}$$

If one simply considers wave equations with energy data without any additional decay at spatial infinity then it becomes impossible to obtain any uniform rate of decay in time for the solutions. Instead, in order to measure the dispersion one needs to average in time and obtain bounds which say that the solution cannot stay concentrated for too long. These bounds are called Strichartz estimates. Applied to the solutions for the homogeneous wave equation (1) they have the form

$$\| |D_x|^{\frac{n+1}{2}(\frac{1}{p}-\frac{1}{2})} \nabla u\|_{L^q(L^p)} \lesssim \|\nabla u(0)\|_{L^2} \tag{5}$$

where (see [4], [6])

$$\frac{2}{q} + \frac{n-1}{p} = \frac{n-1}{2} \qquad 2 \leq p, q \leq \infty, \qquad (q,p,n) \neq (2, \infty, 3)$$

In this article we investigate the variable coefficient version of the dispersive inequalities (4). We consider the Cauchy problem

$$P(t,x,\partial) u = 0 \quad u(0) = u_0 \quad \partial_t u(0) = u_1 \tag{6}$$

in $[0,1] \times \mathbb{R}^n$ where the second order operator

$$P(t,x,\partial) = g^{\alpha\beta}(t,x)\partial_\alpha \partial_\beta + l^\alpha(t,x)\partial_\alpha + m(t,x) \tag{7}$$

is hyperbolic with respect to time. We assume that the matrices $(g^{\alpha\beta}(t,x))$, $(g_{\alpha\beta}(t,x))^{-1}$ are uniformly bounded, have signature $(1,n)$ and the surfaces $\{t = \textit{constant}\}$ are uniformly space-like. Our main result is the following:

Theorem 1.1. *Let $\epsilon > 0$ be sufficiently small. Assume that the coefficients of P satisfy*

$$\|\partial^2 g\|_{L^1(0,1;L^\infty(\mathbb{R}^n))} \leq \epsilon, \quad l, \partial l, m \in L^1(0,1; L^\infty(\mathbb{R}^n)) \tag{8}$$

Then for $t \in [0,1]$ the solution u to (6) satisfies

$$\| |D_x|^{-s} \nabla u(t)\|_{L^p} \lesssim t^{-\frac{n-1}{2}(\frac{1}{p'}-\frac{1}{p})} \| |D_x|^{\frac{n+1}{2}(\frac{1}{p'}-\frac{1}{p})-s} \nabla u[0]\|_{L^{p'}} \tag{9}$$

provided that

$$2 \le p \le \infty \quad (n = 2), \qquad 2 \le p < \frac{2(n-1)}{n-3} \quad (n \ge 3) \tag{10}$$

and

$$n(\frac{1}{2} - \frac{1}{p}) \le s \le \frac{3}{2} - \frac{1}{p}$$

The main part of the proof is to obtain a multiscale wave packet decomposition for solutions to the wave equation. This decomposition may also be of independent interest.

We note that results which are similar in spirit are independently obtained in ongoing work of Smith using a different approach, based on representing the wave evolution in a single frequency dependent wave packet basis.

The exponent $p = \frac{2(n-1)}{n-3}$ is exactly the one appearing in the endpoint Strichartz estimates. For this exponent there is a logarithmic divergence in our estimates. One can remedy it and even increase the range of p by improving the regularity of the coefficients. We note that our proof makes it easy to obtain many such variations of the above result.

Remark 1.2. *The only role of the small constant ϵ is to prevent caustics. One can easily replace the assumption that ϵ is small with an assumption that no caustics occur in the time interval* $[0, 1]$.

Remark 1.3. *The upper limit for s in the theorem is determined by the regularity of g. However, the lower limit only has to do with the regularity of the lower order terms. For instance if in addition we assume*

$$\partial^2 l \in L^1 L^n, \qquad \partial m \in L^1 L^n \tag{11}$$

then we obtain the extended range

$$n\left(\frac{1}{2} - \frac{1}{p}\right) - 2 \le s \le \frac{3}{2} - \frac{1}{p}$$

This can be further improved if the coefficients g, l, m have additional Sobolev regularity on the same scale.

Remark 1.4. *Using the multiscale wave packet decomposition of the fundamental solution one can also obtain weighted pointwise bounds for it. The weights depending on the distance to the characteristic cone, more precisely the variable coefficient analogue of the $t^2 - x^2$ function. This is perhaps less interesting since it matches the constant coefficient bounds only in dimension $n = 2$. Similarly, by adding up our wave* *packet bounds in a different way on can also obtain weighted L^p bounds for the fundamental solution.*

Remark 1.5. *The above result applies equally to systems with diagonal principal part and coupling only in the lower order terms.*

One can compare the above result with similar results obtained earlier for the Strichartz estimates. For operators with smooth coefficients these were proved in [5], [8]. The first result for low regularity coefficients was obtained in [10]. In this article Smith first introduced the idea of constructing wave packet parametrices for the wave equation with C^2 coefficients. He also showed that the Strichartz estimates hold for such operators in low dimension $n = 2, 3$.

Smith's result was extended to higher dimensions in [13]. The $g \in C^2$ condition was later relaxed to $\partial^2 g \in L^1 L^\infty$ in [15]. We also refer the reader to [2], [1] and [7] for related work. On the other hand, counterexamples in [11], [12] show that the Strichartz estimates do not hold in general for operators with C^s coefficients with $s < 2$. However, they do hold for certain operators with lower regularity coefficients arising in the study of quasilinear wave equations, see [9].

The result we prove here is stronger, since a classical argument shows that the dispersive estimates (4) imply the Strichartz estimates (5). Interestingly enough, the range of p that we can allow in Theorem 1.1 is exactly the range which is needed in order to prove the dispersive estimates (modulo the endpoint $p = \frac{2(n-1)}{n-3}$ in dimension $n \geq 4$). However, we do not know if this range for p is sharp.

One should note also that the Strichartz estimates are considerably more robust than the dispersive estimates, in that it suffices to prove them for a convenient parametrix and then iterate. By contrary, as far as we know, no such method can be used in order to prove the dispersive estimates.

Our strategy for the proof is as follows. The initial step is to use a Littlewood-Paley decomposition to reduce the problem to frequency localized estimates. To achieve this we prove in Section 3 that the leakage from one frequency to other frequencies is negligible. This is a very robust argument which requires only $\partial g \in L^1 L^\infty$ for the regularity of the coefficients.

The main step in the proof is to obtain a multiscale wave packet decomposition for a frequency localized part of the fundamental solution.

We briefly describe the Hamilton flow for the operator P in Section 4; especially, we consider the geometry of the characteristic cones and show that caustics do not occur in the time interval we consider.

In Section 5 we introduce the wave packets and show that they are coherent along the wave flow. In Section 6 we show that if the coefficients of P are frequency localized on some scale then we have an almost orthogonal decomposition of waves into wave packets on a related scale. Using a paradifferential type calculus, this leads to the multiscale wave packet decomposition for the P waves in Section 7. The dispersive estimates are obtained in Section 8 following a brief analysis of the Hamilton flow, precisely of the characteristic cones.

The simplest wave packets in our construction are sharply localized on the scale of the uncertainty principle as in [10], [9]. As the spatial scales increase we give up the sharp localization but we gain better L^2 bounds in the fundamental solution representation.

2. Notations

We adopt the usual convention that Roman alphabet letters stand for indices in the range $\{1, \dots, n\}$, while the Greek ones denote indices in the range $\{0, \dots, n\}$. We also use the common summation convention.

The partial differentiation is denoted by ∂, where ∂_0 is the time derivative and ∂_i are the spatial derivatives. ∇ stands for the space time gradient.

For the initial data in (6) we occasionally use the short notation

$$u[0] = (u(0), \partial_t u(0))$$

We consider a spatial Littlewood-Paley decomposition in the Fourier space,

$$1 = \sum_{k=0}^{\infty} S_k(D)$$

where for $k > 0$ we have

$$S_k(\xi) = \phi(2^{-k}|\xi|)$$

with ϕ supported in $\{\frac{1}{2} \leq |\xi| \leq 2\}$. Set

$$S_{<k} = \sum_{h=0}^{k-1} S_h, \qquad S_{>k} = \sum_{h>k} S_h, \qquad S_{[j,k]} = \sum_{j<h<k} S_h$$

We say that a function u is localized at frequency 2^k if its Fourier transform is supported in the annulus $\{2^{k-1} \leq |\xi| \leq 2^k\}$.

For the the paradifferential type calculus we also need to truncate the coefficients of P in frequency. However, for them it is more convenient to use a space-time truncation. Thus we denote a similar space-time Littlewood-Paley decomposition by

$$1 = \sum_{k=0}^{\infty} \tilde{S}_k(D_{x,t})$$

We assume that the two decompositions are chosen so that the difference of their symbols is supported in

$$\text{supp } (S_k(\xi) - \tilde{S}_k(\tau, \xi)) \subset \{|\xi| < 2^{k+1}, \ |\tau| > 2^{k+10}\}$$

Given P in (7) we define the modified operators

$$P_{<k} = (S_{<k} g^{\alpha\beta})\partial_\alpha \partial_\beta + (S_{<k} l^\alpha)\partial_\alpha + (S_{<k} m)$$

Similarly we define P_k, $P_{[j,k]}$, $\tilde{P}_{<k}$, $\tilde{P}_k$, $\tilde{P}_{[j,k]}$, The last three truncations require the coefficients to be defined globally in time. Hence we assume they have been extended to functions with similar properties in all of $\mathbb{R}^{n+1}$.

In our analysis we consider the equation (6) with P replaced by $\tilde{P}_{<h}$. This new equation is denoted $(6)_{<h}$ and is used for solutions which are essentially localized at a frequency 2^k with $k/2 < h < k$.

For various parts of our analysis we use weaker or stronger regularity of the coefficients of P. Thus we introduce the following two sets of assumptions[1]:

$$\partial g \in L^1 L^\infty, \qquad l \in L^1 L^\infty, \qquad m \in L^1 L^n \tag{12}$$

$$\partial^2 g \in L^1 L^\infty, \qquad l, \partial l \in L^1 L^\infty, \qquad m \in L^1 L^\infty \tag{13}$$

3. Dyadic frequency localization

Assume that the initial data for the solution u to (6) is frequency localized at frequency 2^k. Is the solution u also frequency localized to a similar region ? This is definitely not the case in general for operators with variable coefficients. However, if the coefficients satisfy (12) and are smooth on a scale larger than 2^{-k} then we are able to obtain a very convenient estimate of the tail that spills into the other frequencies.

Proposition 3.1. *For each $M > 0$ there exists $C > 0$ so that for each $0 < h < k$ and each solution u to (6) in $[0, 1]$ whose initial data (u_0, u_1) is frequency localized in $\{|\xi| \approx 2^k\}$ and whose coefficients satisfy*

$$\begin{cases} \|\nabla g\|_{L^1L^\infty} \leq M \quad \textit{and} \quad \|\partial_x^a \nabla g\|_{L^1L^\infty} \leq c_a 2^{|a|h} \quad \textit{for} \quad |a| \geq 1 \\ \\ \|\partial_x^a l\|_{L^1L^\infty} + \|\partial_x^a m\|_{L^1L^n} \leq c_a 2^{|a|h} \quad \textit{for} \quad |a| \geq 0 \end{cases} \tag{14}$$

[1] A slight variation is required for (12) in the case $n = 2$, namely $m \in L^1 L^{2+\delta}$ with $\delta > 0$

we have

$$\sum_{|b|\leq N} 2^{-|b|k}\|\partial_x^b \nabla S_{>k+C} u(t)\|_{L^2} + \|\nabla S_{<k-C} u(t)\|_{L^2} \lesssim 2^{N(h-k)}\|\nabla u(0)\|_{L^2}$$

for all $t \in [0,1]$, $N \geq 0$.

In this article we use the above proposition for operators whose coefficients are frequency localized. Precisely, we note that if the coefficients of P satisfy (12) then the coefficients of $\tilde{P}_{<h}$ satisfy (14) with $M = \|\nabla g\|_{L^1 L^\infty}$. As a consequence of this and a simple commutator estimate we obtain the following corollary:

Corollary 3.2. *Suppose that the coefficients of P satisfy* (12). *Then there exists $C = C(\|\nabla g\|_{L^1L^\infty}) > 0$ so that for $1 < h < k - 2C$ and for each initial data $u[0]$ which is frequency localized in $\{|\xi| \approx 2^k\}$ the solution u to $(6)_{<h}$ satisfies*

$$\|\tilde{P}_{<h} S_{[k-C,k+C]} u\|_{L^1(0,1;L^2)} \lesssim 2^{N(h-k)}\|\nabla u(0)\|_{L^2}$$

We emphasize again that only an L^1L^∞ bound on ∂g is needed; thus this result is much more robust than the more precise parametrix representations which we obtain later on.

Proof of Proposition 3.1: Without any restriction in generality we assume that $g^{00} = -1$. We introduce the L^1 functions

$$M(t) = \|\nabla g(t)\|_{L^\infty}, \qquad \mu_0(t) = \|\nabla g(t)\|_{L^\infty} + \|l(t)\|_{L^\infty} + \|m(t)\|_{L^n}$$

$$\mu(t) = \sup_{|a|\leq Z} 2^{-|a|h}(\|(\partial_x^a \nabla g)(t)\|_{L^\infty} + \|(\partial_x^a l)(t)\|_{L^\infty} + \|(\partial_x^a m)(t)\|_{L^n})$$

where Z is a sufficiently large integer.

We define the energy functional,

$$E(u(t)) = \|\partial_t u - g^{0j}\partial_j u\|_{L^2}^2 + \langle \tilde{g}^{ij}\partial_i u, \partial_j u\rangle + \|u\|_{L^2}^2$$

where $\tilde{g}^{ij}$ is the positive definite quadratic form

$$\tilde{g}^{ij} = g^{ij} + g^{0i}g^{0j}, \qquad i,j = \overline{1,n}$$

A routine computation shows that for all solutions to (6) we have

$$\frac{d}{dt}E(u(t)) \lesssim \mu_0(t) E(u(t))$$

which by Gronwall's inequality implies the energy inequality

$$E(u(t)) \lesssim E(u(0)), \qquad t \in [0,1]$$

To prove the proposition we seek to obtain a weighted version of the above energy estimate. Given a symbol $q(t,\xi)$ we define the modified energy functional

$$E_q(u(t)) = \|q(t,D)\partial_t u - g^{0j} q(t,D)\partial_j u\|_{L^2}^2 + \langle \tilde g^{ij} q(t,D)\partial_i u, q(t,D)\partial_j u\rangle \\ +\|q(t,D)u\|_{L^2}^2$$

For an appropriate choice of q we prove that the time derivative of $E_q(u(t))$ is bounded from above,

$$\frac{d}{dt} E_q(u(t)) \lesssim \mu(t) E_q(u(t)) \tag{15}$$

which by Gronwall's inequality yields

$$E_q(u(t)) \lesssim E_q(u(0)), \qquad t \in [0,1]$$

This implies the conclusion of the proposition provided that

$$\begin{array}{ll} q(0,\xi) = 1 & 2^{k-1} < |\xi| < 2^{k+1} \\ q(t,\xi) > 2^{N(k-h)} & |\xi| < 2^{k-C} \\ q(t,\xi) > |\xi|^N 2^{-Nh} & |\xi| > 2^{k+C} \end{array} \tag{16}$$

It remains to find a multiplier q so that both (15) and (16) hold. For the energy estimate (15) we compute

$$\frac{d}{dt}E_q(u(t)) = \langle q(t,D)\partial_t^2 u - g^{0j} q(t,D)\partial_t\partial_j u, (q(t,D)\partial_t - g^{0j}q(t,D)\partial_j)u\rangle \\ + \langle \tilde g^{ij} q(t,D)\partial_t\partial_i u, q(t,D)\partial_j u\rangle \\ + \langle (q_t(t,D)\partial_t - g^{0j}q_t(t,D)\partial_j)u, (q(t,D)\partial_t - g^{0j}q(t,D)\partial_j)u\rangle \\ + \langle \tilde g^{ij} q_t(t,D)\partial_i u, q(t,D)\partial_j u\rangle + O(\mu(t))E_q(u(t))$$

In the first term we substitute $\partial_t^2 u$ from the equation, commute the coefficients $g^{\alpha\beta}$, l^α and m with $q(t,D)$ and integrate by parts as in the usual energy estimates. This cancels the second term and, except for the commutators, everything else can be included in the last term. Denoting

$$C = [q(t,D), g^{i\alpha}]\partial_i\partial_\alpha + [q(t,D), l^\alpha]\partial_\alpha + [q(t,D), m]$$

we obtain

$$\frac{d}{dt}E_q(u(t)) = \langle Cu, (q(t,D)\partial_t - g^{0j}q(t,D)\partial_j)u\rangle \\ + \langle (q_t(t,D)\partial_t - g^{0j}q_t(t,D)\partial_j)u, (q(t,D)\partial_t - g^{0j}q(t,D)\partial_j)u\rangle \\ + \langle \tilde g^{ij} q_t(t,D)\partial_i u, q(t,D)\partial_j u\rangle + O(\mu(t))E_q(u(t))) \\ = \quad I + II + III + IV$$

To continue we need some assumptions on q. We impose some polynomial limit on its growth at infinity,

$$2^{-Nk}(|\xi| + 2^k)^N \lesssim q(t,\xi) \lesssim 2^{-Nh}(|\xi| + 2^k)^N \tag{17}$$

To insure that $II + III$ is essentially non-positive we need q to decrease in time,

$$q_t(t,\xi) \leq 0 \tag{18}$$

Finally, in order to have a reasonable calculus we ask that

$$|\partial_\xi^a \partial_t^b q(t,\xi)| \leq c_a \mu(t)^b q(t,\xi)[2^h(|\xi| + 2^k)]^{\frac{b-|a|}{2}}, \qquad b = 0,1 \tag{19}$$

We denote $v = q(t,D)\nabla u$ and we write

$$II + III = \langle A(t,x)\frac{q_t}{q}(t,D)v, v\rangle$$

where A is a positive definite matrix which has the same regularity as the coefficients g^{ij}. We claim that by Gårding's inequality we can conclude that

$$II + III \leq c_{23}\langle \frac{q_t}{q}(t,D)v, v\rangle + O(1)\|v\|_{L^2}^2, \qquad c_{23} > 0$$

The formulation of Gårding's inequality which we use follows [14]:

Lemma 3.3. *Let $\lambda > 0$ and $r(x,\xi)$ be a symbol which satisfies $\Re r \geq 0$ and*

$$|\partial_x^a \partial_\xi^b r(x,\xi)| \leq c_{ab}\lambda^{|a|-|b|}, \qquad |a| + |b| \geq 2$$

Then

$$\Re\langle R(x,D)u, u\rangle \geq -c\|u\|_{L^2}^2$$

By (14), (18) and (19) we can use the above result with

$$\lambda = [2^h(|\xi| + 2^k)]^{\frac{1}{2}}$$

It remains to estimate I. For this we consider the commutators in C, which we write as

$$C = [q, g^{i\alpha}]q^{-1}\partial_i q\partial_\alpha + [q, l^\alpha]q^{-1}q\partial_\alpha + [q, m]q^{-1}|D_x|^{-1}q|D_x|$$

The last two terms are easy to estimate. By (14), (17) and (19) we have the fixed time bound

$$\|[q, l^\alpha]q^{-1}\|_{L^2\to L^2} \lesssim \mu(t)$$

Intuitively this is because l^α are smooth on the 2^{-h} spatial scale while the symbol $q(t,\xi)$ varies little on the dual 2^h frequency scale. Similar considerations also show that $[q, m]q^{-1}|D_x|^{-1}$ satisfies a similar bound.

For the first term we write

$$[q(t,D), g^{i\alpha}]q^{-1}(t,D)\partial_i = (\partial_x g^{i\alpha})\frac{\partial_\xi q}{q}(t,D)D_i + R(t,x,D)$$

Using again (14), (17) and (19) it follows that the symbol of the remainder R satisfies

$$|\partial_x^a \partial_\xi^b r(t,x,\xi)| \lesssim \mu(t) 2^{|a|h} [2^h(|\xi| + 2^k)]^{-\frac{|b|}{2}}$$

Hence

$$\|R(t,x,D)\|_{L^2 \to L^2} \lesssim \mu(t)$$

For the first expression we note that

$$|(\partial_x g^{i\alpha})\frac{\partial_\xi q}{q}(t,\xi)\xi_i| \lesssim M(t)|\frac{\partial_\xi q}{q}(t,\xi)||\xi_i|$$

On the other hand by (14), (19) we obtain the bounds

$$|\partial_x^a \partial_\xi^b [(\partial_x g^{i\alpha})\frac{\partial_\xi q}{q}(t,\xi)\xi_i]| \lesssim [2^{\frac{3h}{4}}(|\xi| + 2^k)^{\frac{1}{4}}]^{|a|-|b|}, \qquad |a| + |b| \geq 2$$

Hence can use the above form of the sharp Gårding inequality combined with a Littlewood-Paley decomposition in ξ to obtain

$$I \leq M(t)\langle c_1 |\partial_\xi q|(t,D)|D|\nabla u, q(t,D)\nabla u\rangle + O(\mu(t))\|q(t,D)\nabla u\|_{L^2}^2$$

Summing up the bounds for I and $II + III$ we obtain

$$\frac{d}{dt}E_q(u(t)) \leq \langle (c_1 M(t)|\xi||q_\xi| + c_{23} q_t)\nabla u, q\nabla u\rangle + O(\mu(t))E_q(u(t))$$

Thus, in order to derive the estimate (15) it suffices to insure that

$$c_1 M(t)|\xi||q_\xi| + c_{23} q_t \lesssim O(\mu(t))q \tag{20}$$

It remains to construct a symbol q which satisfies (16), (17), (18), (19) and (20). We consider a smooth cutoff function χ which equals 0 in $[-\infty, 1]$ and 1 in $[2, \infty]$. We also set

$$\phi(t) = \left(\int_0^t M(s)ds\right)\left(\int_0^1 M(s)ds\right)^{-1}, \qquad 0 \leq \phi \leq 1$$

Then we define the multiplier $q(t,D)$ whose symbol is

$$q(t,\xi) = 2^{\psi^-(t,|\xi|) + \psi^+(t,|\xi|)}$$

where

$$\psi^-(t,r) = N(k-h)\chi(-\ln r + k - C\phi(t)),$$
$$\psi^+(t,r) = N(\ln r - h)\chi(\ln r - k - C\phi(t))$$

It is easy to verify that q has the desired properties for large C (precisely, for $C \gg M$).

□

Proof of Corollary 3.2: Using the equation $(6)_{<h}$ satisfied by u we can write:

$$\tilde{P}_{<h} S_{[k-C,k+C]} u = -\tilde{P}_{<h} S_{\leq k-C} u - \tilde{P}_{<h} S_{\geq k+C} u$$
$$= -[\tilde{P}_{<h}, S_{\leq k-C}] u - [\tilde{P}_{<h}, S_{\geq k+C}] u$$

Decomposing u in frequency and taking advantage of the fact that $h < k - 2C$, we finally obtain that:

$$\tilde{P}_{<h} S_{[k-C,k+C]} u = -[\tilde{P}_{<h}, S_{\leq k-C}] S_{[k-C-1,k-C+2]} u - \\ - [\tilde{P}_{<h}, S_{\geq k+C}] S_{[k+C-2,k+C+1]} u$$

Taking the $L^1 L^2$ norm and using a standard commutator estimate at fixed time it follows that:

$$\|\tilde{P}_{<h} S_{[k-C,k+C]} u\|_{L^1 L^2} \lesssim \\ (\|\partial g_{<h}\|_{L^1 L^\infty} + 2^{-k} \|\partial l_{<h}\|_{L^1 L^\infty} + 2^{-2k} \|\partial m_{<h}\|_{L^1 L^\infty}) \\ (\|\nabla S_{[k-C-1,k-C+2]} u\|_{L^\infty L^2} + \|\nabla S_{[k+C-2,k+C+1]} u\|_{L^\infty L^2})$$

For the first factor we use (12) and the frequency localization. For the second we use Proposition 3.1 after slightly readjusting C. The conclusion of the corollary follows.

□

4. The Hamilton flow.

Given coefficients g^{ij} as in (13) with $g^{00} = -1$ we decompose the principal symbol

$$p(t, x, \tau, \xi) = \tau^2 - 2g^{0j} \tau \xi_j - g^{ij} \xi_i \xi_j$$

of $P(t, x, \partial)$ as

$$p(t, x, \tau, \xi) = (\tau + a^+(t, x, \xi))(\tau + a^-(t, x, \xi))$$

where the symbols $a^+(t, x, \xi), a^-(t, x, \xi)$ are real, homogeneous and nowhere equal, satisfying

$$a^+(t, x, \xi) = -a^-(t, x, -\xi)$$

They are smooth in ξ and have the same regularity as the coefficients g^{ij}, namely

$$\| |\xi|^{|b|-1} \partial^2_{x,t} \partial^b_\xi a^\pm(t, x, \xi)\|_{L^1_t L^\infty_{x,\xi}} \leq c_b$$

To fix the signs we assume that $a^+ > a^-$. Then a^+ is convex while a^- is concave.

For $(x, \xi) \in T^*\mathbb{R}^n$ at time 0 we denote the Hamilton flows of each of the two factors by $\chi^\pm_t$,

$$\chi^\pm_t(x, \xi) = (x^\pm_t, \xi^\pm_t)$$

The $''+''$ and the $''-''$ flows can be identified by

$$\chi_t^-(x,-\xi) = (x_t^+, -\xi_t^+)$$

and are solutions for the ordinary differential system

$$\left\{ \begin{array}{l} \frac{d}{dt}x_t^\pm = \partial_\xi a^\pm(t, x_t^\pm, \xi_t^\pm) \\ \\ \frac{d}{dt}\xi_t^\pm = -\partial_x a^\pm(t, x_t^\pm, \xi_t^\pm) \end{array} \right. \qquad \left\{ \begin{array}{l} x_0^\pm = x \\ \\ \xi_0^\pm = \xi \end{array} \right. \tag{21}$$

We denote by $\gamma^\pm(x,\xi)$ the two trajectories of the Hamilton flow starting at (x,ξ).

The Hamilton flow is homogeneous in the sense that

$$\chi_t^\pm(x,\lambda\xi) = (x_t^\pm, \lambda\xi_t^\pm)$$

Hence, to study it, without any restriction in generality we can restrict our attention to $\xi \in S^{n-1}$. As a consequence of standard theory of ordinary differential equations one obtains

Proposition 4.1. *Assume that the coefficients $g^{\alpha\beta}$ have regularity $\partial^2 g^{\alpha\beta} \in L^1(0,1;L^\infty(\mathbb{R}^n)$. Then the maps $\chi_t^\pm$ are Lipschitz continuous from $\mathbb{R}^n \times S^{n-1}$ into $\mathbb{R}^n \times (\mathbb{R}^n \setminus \{0\})$.*

In the study of the fundamental solution for the wave equation an important role is played by the characteristic cones. Given $x \in \mathbb{R}^n$, the forward characteristic cone originating at x is obtained as the union of the bicharacteristics

$$K = \{(t, x_t^+), \quad t \in [0,1], \ \xi \in S^{n-1}\}$$

In general one expects the dispersive estimates to hold for as long (in time) as the cones remains nondegenerate. The degeneracies of the cone occur in general after a finite time and are called caustics.

An easy way to insure that caustics do not occur in a finite time interval is to assume that the coefficients g^{ij} have small first and second derivatives.

Proposition 4.2. *Assume that the coefficients $g^{\alpha\beta}$ satisfy*

$$\|\nabla g\|_{L^1(0,1;L^\infty(\mathbb{R}^n)} + \|\nabla^2 g\|_{L^1(0,1;L^\infty(\mathbb{R}^n)} \leq \epsilon$$

for ϵ sufficiently small. Then for each fixed $x \in \mathbb{R}^n$ and $t \in [0,1]$ the map

$$S^{n-1} \ni \xi \to x_t^\pm$$

is an $O(\epsilon)$ Lipschitz perturbation of the map

$$S^{n-1} \ni \xi \to x + t\partial_\xi a^\pm(0,x,\xi)$$

Proof. After a linear change of variable we can assume that $g^{\alpha\beta}(0,x)$ corresponds to the d'Alembertian,

$$g^{\alpha\beta}(0,x) = \mathrm{diag}(-1,1,\cdots,1)$$

Then $a^+(0,x,\xi) = |\xi|$, and using the equations (21) we successively obtain

$$\xi_t^\pm = \xi + O(\epsilon t)$$

and

$$x_t^\pm = x \pm t\xi + O(\epsilon t)$$

To evaluate the derivatives

$$X = \frac{\partial x_t^\pm}{\partial \xi}, \qquad \Xi = \frac{\partial \xi_t^\pm}{\partial \xi}$$

we consider the linearization of the Hamilton flow,

$$\begin{cases} \frac{d}{dt}X = a^\pm_{\xi x}(t,x_t^\pm,\xi_t^\pm)X + a^\pm_{\xi\xi}(t,x_t^\pm,\xi_t^\pm)\Xi \\ \\ \frac{d}{dt}\Xi = -a^\pm_{xx}(t,x_t^\pm,\xi_t^\pm)X - a^\pm_{x\xi}(t,x_t^\pm,\xi_t^\pm)\Xi \end{cases} \qquad \begin{cases} X(0) = 0 \\ \\ \Xi(0) = I_n \end{cases}$$

This is a linear system with integrable coefficients. A routine analysis leads to

$$\Xi(t) = I_n + O(\epsilon t), \qquad X(t) = t a^\pm_{\xi\xi}(0,x,\xi) + O(\epsilon t)$$

This concludes the proof.

□

In the sequel we apply the above results to the operators $\tilde{P}_{<h}$ as h varies. Then it is important to know that the corresponding Hamilton flows are not too far apart.

Proposition 4.3. *Assume that the coefficients $g^{\alpha\beta}$ have regularity $\partial^2 g^{\alpha\beta} \in L^1(0,1;L^\infty(\mathbb{R}^n))$. Let*

$$\mathbb{R} \times S^{n-1} \ni (x,\xi) \to (x^\pm_{t,h},\xi^\pm_{t,h})$$

be the Hamilton flow for $\tilde{P}_{<h}$. Then

$$|x^\pm_{t,h} - x^\pm_t| \lesssim 2^{-h}, \qquad |\xi^\pm_{t,h} - \xi^\pm_t| \lesssim 2^{-h}, \qquad |(x^\pm_{t,h} - x^\pm_t)\xi^\pm_t| \lesssim 2^{-2h}$$

Proof. We denote

$$\mu(t) = \|\partial^2 g(t)\|_{L^\infty}$$

Then we have

$$|(g - \tilde{g}_{<h})(t)| \lesssim 2^{-2h}\mu(t), \qquad |(\nabla(g - \tilde{g}_{<h}))(t)| \lesssim 2^{-h}\mu(t)$$

which further yield

$$|\partial_x(a^\pm - a^\pm_{<h})(t,x,\xi)| \lesssim 2^{-h}\mu(t)|\xi|$$

$$|\partial_\xi(a^\pm - a^\pm_{<h})(t,x,\xi)| \lesssim 2^{-2h}\mu(t)$$

Subtracting the two Hamilton flow equations (21) we have

$$\frac{d}{dt}(x^\pm_{t,h} - x^\pm_t) \lesssim 2^{-2h}\mu(t) + O(|x^\pm_{t,h} - x^\pm_t| + |\xi^\pm_{t,h} - \xi^\pm_t|)$$

$$\frac{d}{dt}(\xi^\pm_{t,h} - \xi^\pm_t) \lesssim 2^{-h}\mu(t) + O(|x^\pm_{t,h} - x^\pm_t|\mu(t) + |\xi^\pm_{t,h} - \xi^\pm_t|)$$

Using Gronwall's inequality yields the first two relations in the proposition. For the third we need to be more precise. Using the homogeneity of the symbols $a^\pm$ we obtain $\xi \cdot a^\pm_\xi = a^\pm$. Then we compute

$$\begin{aligned}\frac{d}{dt}[(x^\pm_{t,h} - x^\pm_t)\xi_t] &= ((a^\pm_{<h})_\xi(x^\pm_{t,h},\xi^\pm_{t,h}) - a^\pm_\xi(x^\pm_t,\xi^\pm_t))\xi_t - (x^\pm_{t,h} - x^\pm_t)a^\pm_x(x^\pm_t,\xi^\pm_t)\\ &= a^\pm(x^\pm_{t,h},\xi^\pm_{t,h}) - a^\pm(x^\pm_t,\xi^\pm_t) - (\xi^\pm_{t,h} - \xi^\pm_t)a^\pm_\xi(x^\pm_{t,h},\xi^\pm_{t,h})\\ &\quad -(x^\pm_{t,h} - x^\pm_t)a^\pm_x(x^\pm_t,\xi^\pm_t)\\ &= O(\mu(t))(|x^\pm_{t,h} - x^\pm_t|^2 + |\xi^\pm_{t,h} - \xi^\pm_t|^2)\end{aligned}$$

which yields the third relation in the proposition. □

5. Generalized wave packets

In this section we define a generalized notion of wave packets, which are functions which are phase space localized in certain neighborhoods of bicharacteristic rays for $\tilde{P}_{<h}$. Then we use energy estimates to show that initial data with an appropriate phase space localization yields solutions which are generalized wave packets. We assume that

$$2^{\frac{k}{2}} \leq 2^h \ll 2^k$$

Consider a fixed bicharacteristic ray

$$t \to \gamma(t) = (x^\pm_t, \xi^\pm_t)$$

for $\tilde{P}_{<h}$. One can easily compute its regularity using (21), namely

$$\|\partial^a_t x^\pm_t\|_{L^1} \leq c_a 2^{h\max\{0,|a|-3\}}, \tag{22}$$

$$\|\partial^a_t \xi^\pm_t\|_{L^1} \leq c_a 2^{h\max\{0,|a|-2\}}|\xi^\pm_t|. \tag{23}$$

Then we take $d(t) = (d_0(t), \cdots, d_{n-1}(t))$ to be a family of normalized first order differential operators chosen so that at time t we have

(i) d_0 is the flow direction,

$$d_0 = \partial_t + a^\pm_\xi \partial_x$$

(ii) $(d_1(t), \cdots, d_{n-1}(t))$ are a basis for the plane $x \cdot \xi^\pm_t = 0$.

These operators can be chosen so that their coefficients have a similar regularity to ξ_t, namely,

$$\|\partial_t^a d_k^j\|_{L^1} \leq c_a 2^{h\max\{0,|a|-2\}}, \qquad j=\overline{0,n}, \quad k=\overline{0,n-1}$$

For a spatial vector y at time t we denote by y_n its normal component with respect to the same plane,

$$y_n = \frac{\xi_t^\pm \cdot y}{|\xi_t^\pm|}$$

Now we define wave packets as approximate solutions to the wave equation which are localized near a bicharacteristic on an appropriate scale.

Definition 5.1. *Let N be a large integer and $(x,\xi)\in T^*\mathbb{R}^n$ with $|\xi|\approx 2^k$.*

a) A function u is called a frequency 2^k and localization 2^{-h} wave packet of $\pm$ type centered at $\gamma^\pm(x,\xi)$ for $\tilde{P}_{<h}$ if it is localized at frequency 2^k and satisfies the bounds

$$\|(x-x_t^\pm)^a(x-x_t^\pm)_n^b d^\rho \nabla u\|_{L^\infty L^2} \leq 2^{k|\rho|+(h-k)(|a|+2b+|\rho|)}, \tag{24}$$

$$\|(x-x_t^\pm)^a(x-x_t^\pm)_n^b d^\rho \tilde{P}_{<h} u\|_{L^1 L^2} \leq 2^{k|\rho|+(h-k)(|a|+2b+|\rho|)}. \tag{25}$$

for $|a|+2b+|\rho|\leq N$.

b) An initial data set (u_0,u_1) is called a frequency 2^k and localization 2^{-h} wave packet initial data of $\pm$ type centered at (x,ξ) for $\tilde{P}_{<h}$ if the corresponding solution satisfies (24) *at time $t=0$.*

c) A function u is called a frequency 2^k and localization 2^{-h} exact wave packet of $\pm$ type centered at $\gamma^\pm(x,\xi)$ for $\tilde{P}_{<h}$ if it is localized at frequency 2^k and satisfies the bound (24) together with

$$\|\tilde{P}_{<h} u\|_{L^1L^2} \leq 2^{N(h-k)} \tag{26}$$

Thus our packet are localized on the $2^{2(h-k)} \times (2^{h-k})^{n-1}$ spatially, respectively $2^k \times (2^h)^{n-1}$ in frequency. This is on the scale of the uncertainty principle precisely when $h=k/2$.

Our main energy estimate asserts that if the initial data is localized in the phase space then the solution remains localized along the corresponding bicharacteristic.

Proposition 5.2. *Assume that the coefficients of P satisfy* (13). *Let u be a frequency 2^k function so that (25) holds and that (24) holds at time 0. Then (24) holds at all times $t\in[0,1]$ with an additional constant.*

An immediate consequence of this result and of Corollary 3.2 is the following

Corollary 5.3. *Assume that the coefficients of P satisfy* (13) *and that* $u[0]$ *and* $\tilde{P}_{<h}u$ *are localized at frequency* 2^k*. If (25) holds and (24) holds at time* 0 *then (24) holds for* $S_{[k-C,k+C]}u$ *at all times.*

Proof. **The model case.** Our idea is to reduce the problem to a canonical case, where the analysis is considerably simpler. We begin by describing this case. First we want the geodesic γ to be straight. Precisely, we assume that

$$x_t^{\pm} = 0, \qquad \xi_t^{\pm} = (0, 2^k), \qquad d_i = \partial_i \tag{27}$$

This requires that the coefficients $g^{\alpha\beta}$ satisfy

$$g^{in}(t,0) = 0, \quad i = \overline{1,n}, \qquad \partial_x g^{nn}(t,0) = 0 \tag{28}$$

Then we want the coefficients to satisfy (12) and be smooth on the 2^{-h} scale with sharp spatial frequency localization. We consider a function $\mu \in L^1(0,1)$. For $g^{\alpha\beta}$ we must have

$$|\partial^a g^{\alpha\beta}| \leq c_a \mu(t) 2^{(|a|-1)h}, \qquad |a| \geq 1 \tag{29}$$

For l and m we need

$$|\partial^a l^\alpha| \leq c_a \mu(t) 2^{|a|h}, \qquad \|\partial^a m(t)\|_{L^n} \leq c_a \mu(t) 2^{|a|h} \tag{30}$$

A consequence of (28) and (29) is the bound

$$|g^{in}| \lesssim \mu(t)|x| \tag{31}$$

Furthermore, (28) suggests that for g^{nn} we should ask for a stronger bound, namely

$$|g^{nn}| \lesssim \mu(t)(|x|^2 + |x_n|), \qquad |\nabla g^{nn}| \lesssim \mu(t)|x| \tag{32}$$

In addition, we ask for a similar stronger bound for its higher order derivatives, namely

$$|\partial^a g^{nn}| \leq c_a \mu(t) 2^{(|a|-1)h}(|x| + 2^{-h}) \tag{33}$$

Reduction to the model case. We do a change of coordinates in order to straighten the bicharacteristic $t \to (x_t^{\pm}, \xi_t^{\pm})$. By (23) we can choose a matrix valued function $Q(t)$ so that

$$Q(t)\xi_t^{\pm} = (0, 2^k), \qquad \|\partial^a Q(t)\|_{L^1} \leq c_a 2^{h\max\{0,|a|-2\}} \tag{34}$$

Then the linear transformation

$$x := Q(t)^{-1}(x - x_t^{\pm})$$

achieves the first two conditions in (27).

The coefficients of P in the new coordinates involve the Jacobian of the change of coordinates. Hence they will retain their spatial regularity. However, we loose control of the second order time derivative in the terms which involve derivatives of $Q(t)$. Fortunately these terms have a special form, in that they contain a factor of x. Hence the new coefficients have the form

$$g(t,x) = g_0(t,x) + a(t)\, x\, g_1(t,x)$$

where $\partial^2 g_0, \partial^2 g_1 \in L^1L^\infty$ while a is of class $W^{1,1}$ in t and smooth on the 2^{-h} scale. Then the bound (29), (33) easily follow, while (32) is a consequence of (28).

A similar issue arises with the first order terms, which in the new coordinates have the form

$$l(t,x) = l_0(t,x) + a(t)\, x\, l_1(t,x)$$

where $\partial l_0, \partial l_1 \in L^1L^\infty$ while a is of class L^1 in t and smooth on the 2^{-h} scale. But this is still in L^1L^∞.

Finally, we investigate the change in the operators d. In the new coordinates, we obtain d of the form

$$d_0(t) + xd_1(t)$$

where d_0 retains the same regularity and properties as before, while d_1 is only of class $W^{1,1}$ and it is not necessarily tangent to $\{x_n = 0\}$. However, because of the balance between the powers of x and of d in the definition of wave packets we can neglect this second term. Furthermore, d is a basis in the space generated by $\partial_0, \cdots, \partial_{n-1}$ therefore we can substitute it by $\partial_0, \cdots, \partial_{n-1}$.

The energy estimates in the model case. We first recall he standard energy estimate for the wave equation, which we write in the form

$$\|\nabla u(t)\|_{L^2} \lesssim \|\nabla u(0)\|_{L^2} + \int_0^t \|Pu(s)\|_{L^2} ds \tag{35}$$

This requires only the weaker assumption (12) on the coefficients. We also note that we do not need to include the L^2 norm of u in the energy estimates, due to the fact that u is localized at frequency 2^k. This implies the bound

$$\|u(t)\|_{L^2} \lesssim 2^{-k}\|\nabla u(t)\|_{L^2} \tag{36}$$

The energy estimate implies the conclusion of the theorem directly in the case when $N = 0$. To prove the theorem for a larger N we will argue by induction with respect to N. Given multiindices (a, b, c) we denote

$$\rho(a,b,c) = k|c| + (h-k)(|a| + 2b + |c|)$$

Let $J_N, J_{\leq N}$ be the sets of indices

$$J_N = \{(a,b,c); |a|+2b+|c| = N\}, \quad J_{\leq N} = \{(\tilde{a},\tilde{b},\tilde{c}); |\tilde{a}|+2\tilde{b}+|\tilde{c}| \leq N\}$$

The key fixed time estimate which we will inductively prove is

$$\|P(x^a x_n^b d^c u)\|_{L^2} \lesssim 2^{\rho(a,b,c)} \sum_{J_{\leq N}} 2^{-\rho(\tilde a,\tilde b,\tilde c)} \Big(\mu(t)\|\nabla x^{\tilde a} x_n^{\tilde b} d^{\tilde c} u\|_{L^2} + \|x^{\tilde a} x_n^{\tilde b} d^{\tilde c} Pu\|_{L^2}\Big) \tag{37}$$

for all indices $(a,b,c) \in J_N$. If this holds then the energy estimate (35) together with Gronwall's lemma yield

$$\|\nabla x^a x_n^b d^c u\|_{L^\infty L^2} \lesssim 2^{\rho(a,b,c)} \sum_{J_{\leq N}} 2^{-\rho(\tilde a,\tilde b,\tilde c)} \|x^{\tilde a} x_n^{\tilde b} d^{\tilde c} Pu\|_{L^1 L^2}, \quad (a,b,c) \in J_N$$

which implies the conclusion of the theorem.

We note that by (36) the summand in (37) can be expanded to include expressions of the form

$$2^k \|x^{\tilde a} x_n^{\tilde b} d^{\tilde c} u\|_{L^2}$$

Because of this, it also follows that the order of operators $\nabla x^{\tilde a} x_n^{\tilde b} d^{\tilde c}$ in the first summand is not important, we can commute them in any possible way. These two observations are used repeatedly in the sequel without further comment.

Induction: The case $N = 0$. As mentioned before, this follows directly from (35).

Induction: The case $N = 1$. Strictly speaking this is not necessary, but we do it in order to better illustrate the idea. We write

$$P(xu) = xPu + [P,x]u, \qquad P(du) = dPu + [P,d]u$$

Then it suffices to estimate the second right hand side terms by

$$\|[P,x]u\|_{L^2} \lesssim \mu(t)(\|x\nabla u\|_{L^2} + \|du\|_{L^2} + \|u\|_{L^2}) \tag{38}$$

$$\|[P,d]u\|_{L^2} \lesssim \mu(t)(2^k\|x\nabla u\|_{L^2} + \|d\nabla u\|_{L^2} + 2^h(\|\nabla u\|_{L^2} + \|u\|_{L^2})) \tag{39}$$

For (38) we have

$$[P, x_j] = g^{\alpha j}\partial_\alpha + l^j$$

The first term contains a d unless $\alpha = n$, in which case we use (31).

The commutator in (39) is written as

$$[P,d] = -(dg^{\alpha\beta})\partial_\alpha\partial_\beta - (dl^\alpha)\partial_\alpha - dm$$

For the first term we use the Lipschitz bound on g unless $\alpha = \beta = n$, when we need (32). For the second and the last we use the bound (30). This yields

$$\|[P,d]u\|_{L^2} \lesssim \mu(t)(\|d\nabla u\|_{L^2} + \|x\partial_n^2 u\|_{L^2} + 2^h\|\nabla u\|_{L^2} + 2^h\|u\|_{L^2})$$

Finally, in the second term we use the frequency localization of u to replace a ∂_n by 2^k.

The induction step: We assume that we know the estimate (37) for $N-1$ and we prove it for N. We have

$$P(x^a x_n^b d^c u) = x^a x_n^b d^c P u + [P, x^a x_n^b d^c] u$$

so it remains to estimate the commutator.

We begin with the first order term in P namely $l^\alpha \partial_\alpha$; the zero order term is somewhat simpler. We compute

$$[l^\alpha \partial_\alpha, x^a x_n^b d^c] u = l x^{a-1} x_n^b d^c u + l x^a x_n^{b-1} d^c u + \sum_{1 \leq |e|}^{e \leq c} (d^e l) x^a x_n^b d^{c-e} \nabla u$$

We bound l by $\mu(t)$ in the L^2 estimate for the first two terms. For the last sum we simply use the bound (30) for the derivatives of l.

It remains to consider the second order terms in P. We write

$$\begin{aligned}
[g^{\alpha\beta}\partial_\alpha\partial_\beta, x^a x_n^b d^c] &= g^{ij} x^{a-2} x_n^b d^c + g^{in} x^{a-1} x_n^{b-1} d^c + g^{nn} x^a x_n^{b-2} d^c \\
&\quad + g^{i\gamma} x^{a-1} x_n^b \partial_\gamma d^c + g^{n\gamma} x^a x_n^{b-1} \partial_\gamma d^c + x^a x_n^b [g^{\alpha\beta}\partial_\alpha\partial_\beta, d^c] \\
&= E_1 + E_2 + E_3 + E_4 + E_5 + E_6
\end{aligned}$$

For the first term we use the frequency localization of u and the induction hypothesis to write:

$$\|E_1 u\|_{L^2} \lesssim \|x^{a-2} x_n^b d^c u\|_{L^2} \lesssim 2^{-k} \|\nabla (x^{a-2} x_n^b d^c u)\|_{L^2}$$

which suffices since $k < 2h$.

For the second term we use in addition (31) to obtain

$$\|E_2 u\|_{L^2} \lesssim \mu(t) \|x^a x_n^{b-1} d^c u\|_{L^2} \lesssim \mu(t) 2^{-k} \|\nabla (x^a x_n^{b-1} d^c u)\|_{L^2}$$

For the third term we need the more precise bound (32):

$$\begin{aligned}
\|E_3 u\|_{L^2} &\lesssim \mu(t)(\|x^{a+2} x_n^{b-2} d^c u\|_{L^2} + \|x^a x_n^{b-1} d^c u\|_{L^2}) \\
&\lesssim \mu(t) 2^{-k} (\|\nabla (x^{a+2} x_n^{b-2} d^c u)\|_{L^2} + \|\nabla (x^a x_n^{b-1} d^c u)\|_{L^2})
\end{aligned}$$

For the fourth term we have to split the discussion in two cases. The first one is when $\gamma \neq n$. It follows that

$$\|E_4^\gamma u\|_{L^2} \lesssim \|x^{a-1} x_n^b d^{c+1} u\|_{L^2} \lesssim 2^{-k} \|\nabla (x^{a-1} x_n^b d^{c+1} u)\|_{L^2}$$

The case $\gamma = n$ requires (31):

$$\|E_4^n u\|_{L^2} \lesssim \mu(t) \|x^a x_n^b d^c \nabla u\|_{L^2}$$

The fifth term will be the one which creates the most problems. We have to discuss it in terms of the possible values of γ. We start with $\gamma \in \{1 \dots n-1\}$, in which case we use (31) to obtain:

$$\|E_5^\gamma u\|_{L^2} \lesssim \mu(t) \|x^{a+1} x_n^{b-1} d^{c+1} u\|_{L^2} \lesssim \mu(t) 2^{-k} \|\nabla (x^{a+1} x_n^{b-1} d^{c+1} u)\|_{L^2}$$

For $\gamma = n$ we take advantage of (32):

$$\|E_5^n u\|_{L^2} \lesssim \mu(t)(\|x^a x_n^b d^c \nabla u\|_{L^2} + \|x^{a+2} x_n^{b-1} d^c \nabla u\|_{L^2})$$

The case when $\gamma = 0$ is the difficult one because for g^{0n} we do not have a nice estimate like in the case of the other g^{in} coefficients. So if we estimate $E_5^0 u$ directly using just $\|g^{0n}\|_{L^\infty} \lesssim 1$, we obtain

$$\|E_5^0 u\|_{L^2} \lesssim 2^{-k} \|\nabla(x^a x_n^{b-1} d^{c+1} u)\|_{L^2}$$

which is obviously not good enough.

On the other hand, due to the frequency localization we have

$$\begin{aligned} \|E_5^0 u\|_{L^2} &\lesssim 2^{-k} \|\nabla(g^{0n} x^a x_n^{b-1} d^c \partial_t) u\|_{L^2} \\ &\lesssim 2^{-k} \left(\|g^{0n} \partial_t \partial_n x^a x_n^{b-1} d^c u\|_{L^2} + \|g^{0n} d x^a x_n^{b-1} d^{c+1} u\|_{L^2} \right. \\ &\qquad \left. + \|(\nabla g^{0n}) x^a x_n^{b-1} d^{c+1} u\|_{L^2} \right) \end{aligned}$$

In the last two terms we can simply use the uniform bound on g^{0n} and the $\mu(t)$ bound on ∇g^{0n}. For the first term we use the form of the operator P to write

$$g^{0n} \partial_t \partial_n = P - \sum_{(\alpha,\beta) \neq (0,n)} g^{\alpha\beta} \partial_\alpha \partial_\beta - \sum_\alpha l^\alpha \partial_\alpha - m$$

Hence, using the pointwise bounds for g^{in}, g^{nn}, l and m in (31), (32), (30) we can estimate it by

$$\begin{aligned} &\lesssim \ 2^{-k} \|P x^a x_n^{b-1} d^c u\|_{L^2} + \mu(t) 2^{-k} \left(\|d^2 x^a x_n^{b-1} d^c u\|_{L^2} \right. \\ &+ \ \|x d \partial_n x^a x_n^{b-1} d^c u\|_{L^2} + \|x_n \partial_n^2 x^a x_n^{b-1} d^c u\|_{L^2} \\ &+ \ \left. \|x^2 \partial_n^2 x^a x_n^{b-1} d^c u\|_{L^2} + \|\nabla x^a x_n^{b-1} d^c u\|_{L^2} + \|x^a x_n^{b-1} d^c u\|_{L^2} \right) \end{aligned}$$

For the first norm we use the induction hypothesis, while the rest are all acceptable terms.

Finally we come to E_6, which we write in the form:

$$E_6 = -\sum_{f \neq 0} x^a x_n^b (d^f g^{\alpha\beta}) d^{c-f} \partial_\alpha \partial_\beta u = \sum_{f \neq 0} E_6^{\alpha\beta f}$$

Suppose $\alpha \neq n$ (the same applies by symmetry if $\beta \neq n$). Then by (29) we have

$$\|E_6^{\alpha\beta f} u\|_{L^2} \lesssim \mu(t) 2^{h(|f|-1)} \|x^a x_n^b d^{c-f+1} \nabla u\|_{L^2}$$

It remains to consider the case $\alpha = \beta = n$. If $|f| = 1$ then we use (32) and the frequency localization to obtain

$$\begin{aligned} \|E_6^{\alpha\beta f} u\|_{L^2} &\lesssim \ \mu(t) \|x^{a+1} x_n^b d^{c-1} \partial_n^2 u\|_{L^2} \\ &\lesssim \ \mu(t)(2^k \|x^{a+1} x_n^b d^{c-1} \nabla u\|_{L^2} + \|x^{a+1} x_n^{b-1} d^{c-1} \nabla u\|_{L^2}) \end{aligned}$$

where the second term appears only if $b \geq 1$. If $|f| \geq 2$ then we need the higher order bounds (33):

$$\begin{aligned}\|E_6^{\alpha\beta f} u\|_{L^2} &\lesssim \mu(t)(2^{h(|f|-2)}\|x^a x_n^b d^{c-f}\partial_n^2 u\|_{L^2} + 2^{h(|f|-1)}\|x^{a+1}x_n^b d^{c-f}\partial_n^2 u\|_{L^2}) \\ &\lesssim \mu(t)2^{h(|f|-2)}(2^k\|x^a x_n^b d^{c-f}\nabla u\|_{L^2} + \|x^a x_n^{b-1} d^{c-f}\nabla u\|_{L^2}) \\ &+ \mu(t)2^{h(|f|-1)}(2^k\|x^{a+1}x_n^b d^{c-f}\nabla u\|_{L^2} + \|x^{a+1}x_n^{b-1}d^{c-f}\nabla u\|_{L^2})\end{aligned}$$

□

6. Waves as superpositions of wave packets.

The aim of this section is to show that if the coefficients are localized at frequency $< 2^h$ then any initial data which is localized at frequency 2^k is a square summable superposition of frequency 2^k and localization 2^{-h} wave packet initial data.

We first introduce a discrete decomposition of the frequency 2^k part of the phase space

$$\mathbb{R}^n \times \{2^{k-1} < |\xi| < 2^{k+1}\}$$

We begin with a locally finite covering of the annulus in the Fourier space in sectors of size

$$2^k \times (2^h)^{n-1}$$

For each such sector we consider a locally finite covering of the physical space by parallelepipeds of size

$$2^{2(h-k)} \times (2^{h-k})^{n-1}$$

with respect to the dual set of directions. Thus we obtain a discrete locally finite phase space covering

$$\mathbb{R}^n \times \{2^{k-1} < |\xi| < 2^{k+1}\} \subset \bigcup_{j \in J_h} R_j$$

For $j \in J_h$ we denote by (x_j, ξ_j) the center of R_j.

Proposition 6.1. *Assume that the coefficients of P satisfy* (13). *Let $k/2 \leq h < k - C$. Then for any frequency 2^k initial data $u[0]$ for $\tilde{P}_{<h}$ there exist $(k, h, \pm)$ wave packet initial data $u_j^\pm[0]$ centered at (x_j, ξ_j) and coefficients $a_j^\pm$ so that*

$$u[0] = \sum_{\pm} \sum_{j \in J_h} a_j^\pm u_j^\pm[0], \qquad \sum_{\pm} \sum_{j \in J_h} |a_j^\pm|^2 \lesssim \|\nabla u(0)\|_{L^2}^2$$

Combining this result with Corollary 5.3 we obtain

Proposition 6.2. *Assume that the coefficients of P satisfy* (13). *Let $k/2 \leq h < k - C$. Given a frequency 2^k initial data $u[0]$ for $\tilde{P}_{<h}$ there exist $(k, h, \pm)$ exact wave packets $u_j^\pm$ and coefficients $a_j^\pm$ so that the solution u satisfies*

$$S_{[k-C,k+C]}u = \sum_{\pm} \sum_{j \in J_h} a_j^\pm u_j^\pm$$

with

$$\sum_{\pm} \sum_{j \in J_h} |a_j^\pm|^2 \lesssim \|\nabla u(0)\|_{L^2}^2$$

Proof. We first note that if the coefficients of P satisfy (13) then their frequency localizations satisfy at time 0 the following pointwise bounds:

$$|\partial^a \nabla g_{<h}| + |\partial^a l_{<h}| + 2^{-h}|\partial^a m_{<h}| \leq c_a 2^{|a|h} \tag{40}$$

This is all that is needed for this proof.

STEP 1: Phase space localization. Some care is required here, since we want the pieces to be localized sharply in frequency, while on the other hand the spatial decomposition depends on the frequency. To achieve this we proceed as follows:

(a1) We localize in frequency with respect to the $2^k \times (2^h)^{n-1}$ sector decomposition of the 2^k annulus.

(a2) We split each frequency localized piece into spatially localized pieces on the corresponding $2^{2(h-k)} \times (2^{h-k})^{n-1}$ spatial scale, at the expense of losing the sharp frequency localization.

(a3) We relocalize in frequency, obtaining rapidly decreasing tails in neighboring spatial regions.

At the end of this process, we obtain a decomposition

$$u[0] = \sum_{j \in J} a_j u_j[0], \qquad \sum_{j \in J} |a_j|^2 \lesssim \|\nabla u(0)\|_{L^2}^2$$

where for each j the L^2 normalized initial data $u_j[0]$ is localized in R_j, (x_j, ξ_j), with sharp frequency localization. More precisely, for $u_j[0]$ we have:

$$\|(x - x_j)^a (x - x_j)_n^b \partial'^c u_j[0]\|_{H^1 \times L^2} \leq c_{a,b,c} 2^{k|c| + (h-k)(|a|+2|b|+|c|)} \tag{41}$$

where ∂' stands for spatial derivatives in directions normal to ξ_j.

STEP 2: The $\pm$ decomposition. Here we consider the above initial data $u_j[0]$ and show that we can decompose it into two components corresponding to the two possible travel directions. The main result is contained in the next Lemma.

Lemma 6.3. *Let $u[0]$ be an initial data localized near (x_0, ξ_0), with sharp frequency localization, which satisfies (41). Then there is a decomposition*

$$u[0] = u^+[0] + u^-[0]$$

so that the corresponding solutions u^+ respectively u^- satisfy (24) at time 0.

Proof. We choose coordinates so that $x_0 = 0$ and $\xi_0 = (0, \xi_{0n})$. We consider the two possible initial travel directions $l^\pm = l^\pm(x_0, \xi_0)$, normalized in the form

$$l^\pm = \partial_t + q^\pm, \qquad q^\pm = a^\pm_\xi(0, x_0, \xi_0)\partial_x$$

Then it is easy to see that (24) for $u^\pm$ at time 0 takes the form

$$\|x'^a x_n^b \partial'^\rho (l^\pm)^\mu u^\pm[0]\|_{H^1 \times L^2} \leq 2^{k(|\rho|+\mu)+(h-k)(|a|+2b+|\rho|+\mu)}. \tag{42}$$

$$|a| + 2b + |\rho| + \mu \leq N$$

The product l^+l^- serves as a good approximation for $\tilde{P}_{<h}$ near (x_0, ξ_0). Precisely, set

$$R = \tilde{P}_{<h} - l^+l^- = R_1(t, x, \partial_x)\partial_t + R_2(t, x, \partial_x) + M(t, x, \partial)$$

where R_1, R_2 have orders 1 and 2, while M contains the lower order terms in P and has order 1. Then

$$R_1(0, x_0, \xi_0) = 0, \qquad R_2(0, x_0, \xi_0) = 0 \tag{43}$$

This will allow us to substitute $\Box_g$ with l^+l^- modulo a lower order term. Such a substitution may seem too rough, but it turns out to suffice for the $\pm$ split of the initial data.

We seek $u^\pm[0]$ as a partial sum of the formal series

$$u^\pm[0] = \sum_{j=0}^{N} u^{\pm,j}[0]$$

where

$$\|x'^a x_n^b \partial'^\gamma u^{\pm,j}[0]\|_{H^1 \times L^2} \leq c_{a,b,\gamma} 2^{k|\gamma|+(h-k)(|a|+2|b|+|\gamma|+j)} \tag{44}$$

We denote by $u^{\pm,j}$ the corresponding solutions for the wave equation and set

$$u^{\pm,(j)} = u^{\pm,0} + \cdots + u^{\pm,j}$$

We claim we can construct the $u^{\pm,j}[0]$'s inductively so that the solutions $u^{\pm,(j)}$ satisfy (42) for $\mu \leq j + 1$.

For $j = 0$ we solve

$$(u_0, u_1) = (u_0^{+,0}, u_1^{+,0}) + (u_0^{-,0}, u_1^{-,0})$$

and

$$u_1^{+,0} - q^+ u_0^{+,0} = 0, \qquad u_1^{-,0} - q^+ u_0^{-,0} = 0$$

Thus we get a system with 4 equations and four unknowns, governed by the matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & q^+ & 0 \\ 0 & 1 & 0 & q^- \end{pmatrix}$$

This has determinant $q^+ - q^-$, which is an elliptic multiplier on the frequency support of of u_0, u_1. Hence we can solve it and find $(u_0^{\pm,0}, u_1^{\pm,0})$ which satisfy (44).

To obtain (42) for $\mu = 1$ we still need to compute the regularity of $l^{\pm}\partial_t u^{\pm,0}$. We have

$$\partial_t l^{\pm} u^{\pm,0} = -q^{\mp} l^{\pm} u^{\pm,0} - (R_1 \partial_t + R_2 + M) u^{\pm,0}$$

The first term is zero by construction, while for the second we gain either a factor of x' or a factor of ∂' due to the properties of the remainder in (43).

For larger j we argue inductively. First observe that we can use the equation to convert time derivatives into spatial derivatives, which have size $O(2^k)$ because of the frequency localization. Then the contribution of $u^{\pm,j}$ to $(l^{\pm})^i u^{\pm,(j)}$ in (42) is trivially estimated using (44) provided that $i \leq j$. Hence in the induction step it suffices to show that we can find packets $u^{\pm,j+1}[0]$ so that (42) holds for $\mu = j+2$. We have

$$(l^{\pm})^{j+2} u^{\pm,(j+1)}[0] = (l^{\pm})^{j+2} u^{\pm,(j)}[0] + (l^{\pm})^{j+2} u^{\pm,j+1}[0] \tag{45}$$

We bound the first expression using the induction hypothesis. For a solution u we can write

$$\begin{aligned} (l^+)^{j+2} u &= (q^- - q^+)(l^+)^{j+1} u + (l^+)^j (R_1 \partial_t + R_2 + M) u \\ &= (q^- - q^+ + R_1)(l^+)^{j+1} u + (R_1 q^+ + R_2 + M)(l^+)^j u \\ &+ \sum_{|\alpha| \geq 1} (\partial^\alpha g_{<h}) \partial_x^2 (l^+)^{j-|\alpha|} \end{aligned}$$

We apply this to $u^{+,(j)}$. Using the induction hypothesis, (43) for the second term and and the C^1 bound (40) on g for the third term we obtain

$$\| x'^a x_n^b \partial'^\gamma (l^+)^{j+2} u^{+,(j)}[0] \|_{H^1 \times L^2} \leq c_{a,b,\gamma} 2^{k|\gamma| + (h-k)(|a|+2|b|+|\gamma|+j+1)} \tag{46}$$

On the other hand, for the second expression in (45) we write

$$\begin{aligned} (l^+)^{j+2} u^{+,j+1} &= (q^+ - q^-)^{j+1} l^+ u^{+,j+1} + \partial^j l^+ l^- u^{+,j+1} \\ &= (q^+ - q^-)^{j+1} l^+ u^{+,j+1} + \partial^j (R_1 \partial_t + R_2) u^{+,j+1} \end{aligned}$$

In the second part we have the additional gain of either an x' or a ξ' factor in R_1 and R_2. Then the fact that its Cauchy data at time 0 satisfies the correct bounds (42) follows directly from (44). Hence we can neglect it, and conclude that it suffices to choose $(u_0^{\pm,j+1}, u_1^{\pm,j+1})$ so that at time 0 we have

$$\begin{aligned} u_0^{+,j+1} + u_0^{-,j+1} &= 0 \\ u_1^{+,j+1} + u_1^{-,j+1} &= 0 \\ (q^+ - q^-)^j l^+ u^{+,j+1} &= -(l^+)^{j+1} u^{+,(j)} \\ (q^- - q^+)^j l^- u^{-,j+1} &= -(l^-)^{j+1} u^{-,(j)} \end{aligned}$$

This is an elliptic system which we can solve to obtain $(u_0^{\pm,j+1}, u_1^{\pm,j+1})$ satisfying (44). □

□

7. A multiscale wave packet decomposition

Here we consider frequency localized solutions for the wave equation with coefficients satisfying (13) and show that they can be represented as superpositions of wave packets which are localized on multiple scales.

Proposition 7.1. *a) Assume that the coefficients of P satisfy* (13) *and C is sufficiently large. Given a initial data $u[0]$ for $\tilde{P}_{<k-3C}$ which is localized at frequency 2^k, the solution u to* $(6)_{<k-3C}$ *satisfies*

$$\|\partial^a u\|_{L^1L^2} \lesssim 2^{k(|a|-1)} \|\nabla u(0)\|_{L^2}, \quad |a| \geq 0 \tag{47}$$

b) The frequencies away from 2^k in u satisfy the better bound

$$\|\partial^a (1 - S_{[k-2C+2,k+2C-2]})u\|_{L^1L^2} \lesssim 2^{k(|a|-2)} \|\nabla u(0)\|_{L^2}, \quad |a| \geq 1 \tag{48}$$

c) The frequency 2^k part of u can be represented as

$$S_{[k-2C,k+2C]} u = \sum_{\frac{k}{2} \leq h \leq k} 2^{k-2h} \sum_{j \in J_h} a_{h,j}^{\pm} v_{h,j}^{\pm}$$

where $v_{h,j}^{\pm}$ are $(k, h, \pm)$ type wave packets for $\tilde{P}_{<h}$ and

$$\sum_{j \in J_h} \sum_{\pm} |a_{h,j}^{\pm}|^2 \lesssim \|\nabla u(0)\|_{L^2}^2$$

Proof. a) For $a = 1$ this follows directly by energy estimates. For $a > 1$ we differentiate the equation $|a|-1$ times and then use energy estimates. In the case $a = 0$ we need an additional low frequency bound, which follows from part (b).

b) For $\frac{k}{2} \leq h \leq k - 3C$ we denote by u_h the solution to the Cauchy problem

$$\tilde{P}_{<h} u_h = 0 \quad u_h[0] = u[0]$$

The differences

$$v_h = u_{h+1} - u_h$$

solve

$$\tilde{P}_{<h} v_h = -\tilde{P}_h u_{h+1}, \qquad v_h[0] = 0.$$

We decompose v_h into two parts,

$$v_h = v_{h,0} + v_{h,1}$$

which solve

$$\tilde{P}_{<h} v_{h,0} = -\tilde{P}_h S_{[k-C,k+C]} u_{h+1}, \qquad v_{h,0}[0] = 0$$

respectively

$$\tilde{P}_{<h} v_{h,1} = -\tilde{P}_h (1 - S_{[k-C,k+C]}) u_{h+1}, \qquad v_{h,1}[0] = 0.$$

Then for the exact solution u we have

$$u = u_{\frac{k}{2}} + \sum_{\frac{k}{2} < h < k-3C} (v_{h,0} + v_{h,1})$$

The regularity (13) of the coefficients implies that $S_h g$ has size 2^{-2h} and $S_h l$ has size 2^{-h} in $L^1 L^\infty$. Then for the first component we have the trivial estimate

$$\|\tilde{P}_{<h} v_{h,0}\|_{L^1 L^2} \lesssim 2^{k-2h} \|\nabla u(0)\|_{L^2}$$

By energy estimates this gives

$$\|\nabla v_{h,0}\|_{L^\infty L^2} \lesssim 2^{k-2h} \|\nabla u(0)\|_{L^2}$$

On the other hand by Proposition 3.1 and Duhamel's formula we obtain a better estimate at frequencies away from 2^k, namely

$$\|\nabla^a (1 - S_{[k-2C+2,k+2C-2]}) v_{h,0}\|_{L^\infty L^2} \lesssim 2^{k(|a|-2)} 2^{N(h-k)} \|\nabla u(0)\|_{L^2}, \ |a| \geq 1$$

After summation in h this gives the corresponding part of (48). Integrating it in time one also obtains the weaker L^2 bound

$$\|S_{<k-2C+2} v_{h,0}\|_{L^\infty L^2} \lesssim 2^{-k} 2^{N(h-k)} \|\nabla u(0)\|_{L^2} \tag{49}$$

which after summation in h yields the corresponding part of (47) for $a = 0$.

Next we turn our attention to $v_{h,1}$. By Proposition 3.1 the part of u_h which is away from frequency 2^k has size

$$\|\nabla^a (1 - S_{[k-C,k+C]}) u_{h+1}\|_{L^\infty L^2} \lesssim 2^{k(|a|-1)} 2^{(N+2)(h-k)} \|\nabla u(0)\|_{L^2}, \quad |a| \geq 1$$

We claim that the same happens for $a = 0$; only the low frequency bound is nontrivial, so we need to show that

$$\|S_{<k-2C+2}u_{h+1}\|_{L^\infty L^2} \lesssim c_N 2^{-k} 2^{(N+2)(h-k)} \|\nabla u(0)\|_{L^2} \tag{50}$$

We use induction with respect to h. The implicit constants in our bounds depend on N; to avoid any danger of confusion, above and in the following argument we make this dependence explicit.

If $h = \frac{k}{2}$ then (50) follows easily by integrating the previous bound and readjusting N. For the induction step we need a bound for the difference $v_h = u_{h+1} - u_h$. The $v_{h,0}$ part was already estimated in (49), it remains to consider $v_{h,1}$. We rewrite the $v_{h,1}$ equation as

$$\begin{aligned}(\tilde{P}_{<h} + S_h m)v_{h,1} &= -(\tilde{P}_h - S_h m)(1 - S_{[k-C,k+C]})u_{h+1} \\ &\quad -(S_h m)(1 - S_{[k-C,k+C]})u_h\end{aligned}$$

The first part contains derivatives so we estimate it using the $|a| \geq 1$ bound. Hence we obtain

$$\|(\tilde{P}_{<h} + S_h m)v_{h,1}\|_{L^1L^2} \lesssim c_N 2^{-k} 2^{(N+2)(h+1-k)} \|\nabla u(0)\|_{L^2} + \|u_h\|_{L^\infty L^2}$$

It is important here that the secont term contains no N dependent constant. By energy estimates this gives

$$\|v_{h,1}\|_{L^\infty L^2} \lesssim c_N 2^{-k} 2^{(N+2)(h+1-k)} \|\nabla u(0)\|_{L^2} + \|u_h\|_{L^\infty L^2}$$

Summing up all bounds we have proved that

$$\|u_{h+1}\|_{L^\infty L^2} \lesssim c_N 2^{-k} 2^{(N+2)(h+1-k)} \|\nabla u(0)\|_{L^2} + \|u_h\|_{L^\infty L^2}$$

which suffices for the induction step provided that c_N is chosen sufficiently large.

The regularity (13) of the coefficients implies that $S_h g$ has size 2^{-2h} and $S_h l$ has size 2^{-h} in $L^1 L^\infty$ therefore

$$\|\tilde{P}_{<h} v_{h,1}\|_{L^1L^2} \lesssim 2^{-k} 2^{N(h-k)} \|\nabla u(0)\|_{L^2}$$

which leads to

$$\|\nabla v_{h,1}\|_{L^\infty L^2} \lesssim 2^{-k} 2^{N(h-k)} \|\nabla u(0)\|_{L^2}$$

Higher order derivatives of $v_{h,1}$ are estimated in a similar manner by differentiating the $v_{h,1}$ equation. Then the corresponding part of the bound (48) follows by summation in h.

c) Using the above bounds for $v_{h,1}$ and some simple commutations we obtain

$$\|\tilde{P}_{<h} S_{[k-2C,k+2C]} v_{h,1}\|_{L^1L^2} \lesssim 2^{-k} 2^{N(h-k)} \|\nabla u(0)\|_{L^2}$$

Hence the contributions of $S_{[k-2C,k+2C]}v_{h,1}$ to $S_{[k-2C,k+2C]}u$ is negligible, i.e. it can be included into a single wave packet $v^{\pm}_{k-3C,j}$.

On the other hand by Proposition 6.2 we have the wave packet decomposition for the frequency 2^k part of u_h,

$$S_{[k-C,k+C]}u_h = \sum_{j\in J_h}\sum_{\pm} a^{\pm}_{h,j}u^{\pm}_{h,j}$$

The functions $u^{\pm}_{h,j}$ satisfy the bounds (24). This implies that

$$\|(x-x^{\pm}_t)^a(x-x^{\pm}_t)^b_n d^\rho \tilde{P}_h u^{\pm}_{h,j}\|_{L^1L^2} \leq 2^{k-2h}2^{k\rho+(h-k)(|a|+2b+|\rho|)}, \tag{51}$$

We define the functions $v^{\pm}_{h,j}$ as

$$v^{\pm}_{h,j} = 2^{2h-k}S_{[k-2C,k+2C]}w^{\pm}_{h,j}$$

where $w^{\pm}_{h,j}$ solves

$$\tilde{P}_{<h}w^{\pm}_{h,j} = -\tilde{P}_h u^{\pm}_{h,j}$$

Then

$$S_{[k-2C<\cdot<k+2C]}v_{h,1} = 2^{k-2h}\sum_{j\in J_h} v^{\pm}_{h,j}$$

Furthermore, by Corollary 5.3 and (51) we conclude that $v^{\pm}_{h,j}$ is an $(k,h,\pm)$ wave packet associated to the same $\tilde{P}_{<h}$ bicharacteristic as $u^{\pm}_{h,j}$.

□

8. Frequency localized dispersive estimates.

Proposition 8.1. *Assume that the coefficients of P satisfy* (13). *Let u be the solution for $\tilde{P}_{<k-3C}$ with initial data $u[0]$ localized at frequency 2^k. Then the following estimates hold:*

$$\|\nabla S_{[k-2C,k+2C]}u(t)\|_{L^p} \lesssim 2^{k(n+1)(\frac{1}{2}-\frac{1}{p})}t^{-(n-1)(\frac{1}{2}-\frac{1}{p})}\|\nabla u(0)\|_{L^{p'}} \tag{52}$$

for p as in (10).

Proof. The bound is trivial for $p=2$. It also follows from the $p=2$ case by Sobolev embeddings provided that $|t|<2^{-k}$. For larger t we rescale and reduce the problem to the case $t=1$.

For $t=1$ we use the multiscale decomposition in the previous section. For each of the functions $S_{[k-2C,k+2C]}v_h$ we prove an $L^2\to L^2$ and an $L^1\to L^\infty$ bound. The $L^2\to L^2$ bound is the trivial one obtained from energy estimates for v_h, namely

$$\|\nabla S_{[k-2C,k+2C]}v_h(1)\|_{L^2} \lesssim 2^{k-2h}\|\nabla u(0)\|_{L^2} \tag{53}$$

For the $L^1\to L^\infty$ bound it suffices to consider an initial data $u[0]$ which is an L^1 normalized bump function on the 2^{-k} scale concentrated at a point, say $x=0$. Then the wave packet decomposition for the initial data contains only packets spatially centered at 0. There must be $2^{\frac{(n-1)(k-h)}{2}}$ such packets $u^{\pm}_{h,j}[0]$ whose coefficients $a^{\pm}_{h,j}$ have size $2^{\frac{k}{2}+(n-1)\frac{h}{2}}$. This leads to the $(k,h,\pm)$ wave packet decomposition for u_h.

The $(k,h,\pm)$ wave packet decomposition for $S_{[k-2C,k+2C]}v_h$ is similar and involves the same packets, except that we gain an additional 2^{k-2h} factor. Each packet $v^{\pm}_{h,j}$ is normalized in L^2 and frequency localized on a $2^k \times (2^h)^{n-1}$ scale, therefore by Sobolev embeddings

$$\|(x - x_j^t)^a (x - x_j^t)_n^b \nabla v^{\pm}_{h,j}(t)\|_{L^\infty} \lesssim 2^{(h-k)(|a|+2|b|)} 2^{\frac{k}{2}+(n-1)\frac{h}{2}}$$

Thus for $\nabla v_{h,j}$ we get an $2^{\frac{k}{2}+(n-1)\frac{h}{2}}$ uniform bound, plus rapid decay outside the parallelepiped $R_j(t)$ centered at $x_j(t)$ and of size $2^{2(h-k)} \times (2^{h-k})^{n-1}$.

Observe that at time 1 the rectangles $R_j(1)$ are essentially disjoint. This is due to Proposition 4.2 since the points $x_j(1)$ are reached on the bicharacteristics starting at $(0,\xi_j)$ where the directions ξ_j have angular separation 2^{h-k}.

Then for the sum $S_{[k-2C,k+2C]}v_h(1)$ we obtain

$$\|\nabla S_{[k-2C,k+2C]}v_h(1)\|_{L^\infty} \lesssim 2^{\frac{k}{2}+(n-1)\frac{h}{2}} 2^{k-2h} 2^{\frac{k}{2}+(n-1)\frac{h}{2}} = 2^{2k+(n-3)h}$$

Hence

$$\|\nabla S_{[k-2C,k+2C]}v_h(1)\|_{L^\infty} \lesssim 2^{2k+(n-3)h} \|\nabla u(0)\|_{L^1} \tag{54}$$

Interpolating between (53) and (54) we obtain

$$\|\nabla S_{[k-2C,k+2C]}v_h(1)\|_{L^p} \lesssim 2^{k(2-\frac{2}{p})} 2^{h(n-3-(n-1)\frac{2}{p})} \|\nabla u(0)\|_{L^{p'}} \tag{55}$$

The conclusion of the proposition follows after summation with respect to h in the range $\frac{k}{2} \leq h \leq k - 3C$.

□

9. Conclusion

We denote by $u_k, \tilde{u}_k$ the solutions to

$$P_{<k-3C} u_k = 0, \qquad u_k[0] = S_k u[0]$$

respectively

$$\tilde{P}_{<k-3C} \tilde{u}_k = 0, \qquad \tilde{u}_k[0] = S_k u[0]$$

The results in the previous two sections apply to the function $\tilde{u}_k$. The next result asserts assert that u_k is close to $\tilde{u}_k$ and has similar properties.

Proposition 9.1. *Assume that the coefficients of P satisfy* (13). *Then*

$$\|\nabla(u_k - \tilde{u}_k)\|_{L^\infty L^2} \lesssim 2^{-k}\|S_k u[0]\|_{H^1\times L^2}$$

In addition, if C is large enough then

$$\|P_{<k-3C} S_{[k-2C,k+2C]} u_k\|_{L^2} \lesssim \mu(t) 2^{-k}\|S_k u[0]\|_{H^1\times L^2} \tag{56}$$

Using the functions u_k we can construct a good approximate solution to (6).

Proposition 9.2. *Assume that the coefficients of P satisfy* (13).
a) For $s = 0$ we have

$$\|P\sum_{k=1}^{\infty} S_{[k-2C,k+2C]} u_k\|_{L^1 H^s} \lesssim \|\nabla u(0)\|_{H^{s-1}} \tag{57}$$

If in addition (11) *is valid then the above estimate holds for* $0 \leq s \leq 2$.
b) The equation (6) *is well-posed in $H^s \times H^{s-1}$, $s \in [0,1]$. If in addition* (11) *is valid then the well-posedness range extends to* $s \in [0,3]$.

A straightforward consequence of the above propositions is

Corollary 9.3. *Let $s = 0$. Then the solution u to* (6) *can be represented as*

$$u = v + \sum_{k=1}^{\infty} S_{[k-2C,k+2C]}\tilde{u}_k \tag{58}$$

where

$$\|\nabla v\|_{L^\infty(H^s)} \lesssim \|\nabla u(0)\|_{H^{s-1}}$$

If in addition (11) *is valid then the same result holds for* $0 \leq s \leq 2$.

This allows us to quickly conclude the proof of our main theorem. Precisely, the dispersive estimates for $\sum S_{[k-2C,k+2C]}\tilde{u}_k$ follow from Theorem 8.1 in the previous section combined with Littlewood-Paley theory. On the other hand in the bounds for v we have gained one derivative, therefore we can simply use Sobolev embeddings to obtain the desired L^p estimates.

Proof of Proposition 9.1: We write an equation for the difference,

$$P_{<k-3C}(u_k - \tilde{u}_k) = (\tilde{P}_{<k-3C} - P_{<k-3C})\tilde{u}_k, \qquad (u_k - \tilde{u}_k)[0] = 0$$

Then we need to show that

$$\|(\tilde{P}_{<k-3C} - P_{<k-3C})\tilde{u}_k\|_{L^1 L^2} \lesssim 2^{-k}\|S_k u[0]\|_{H^1\times L^2}$$

But this follows from the estimate (47) for $\tilde{u}_k$, since by (13) the coefficients of $\tilde{P}_{<k-3C} - P_{<k-3C}$ satisfy the pointwise bounds

$$\|(\tilde{S}_{<k-3C} - S_{<k-3C})g\|_{L^1L^\infty} \lesssim 2^{-2k}$$

$$\|(\tilde{S}_{<k-3C} - S_{<k-3C})l\|_{L^1L^\infty} \lesssim 2^{-k}$$

$$\|(\tilde{S}_{<k-3C} - S_{<k-3C})m\|_{L^1L^\infty} \lesssim 1$$

The proof of the second part of the proposition is identical to the proof of Corollary 3.2. The function $\mu(t)$ can be used because all commutators are estimated at fixed time.

□

Proof of Proposition 9.2: a) We need to bound the function

$$\sum_{k\geq 1} P_{<k-3C}S_{[k-2C,k+2C]}u_k + \sum_{k\geq 1} P_{[k-3C,k+3C]}S_{[k-2C,k+2C]}u_k$$
$$+ \sum_{k\geq 1} P_{>k+3C}S_{[k-2C,k+2C]}u_k$$

in $L^\infty H^s$. The first sum is easily estimated using (56) and orthogonality.

In the second sum we have two comparable frequencies interacting, and the product contains frequencies which are similar or lower. We claim that (13) suffices in order to estimate it as in (57) for all $s \geq 0$. From (13) we obtain

$$\|S_{[k-3C,k+3C]}g\|_{L^1L^\infty} \lesssim 2^{-2k}, \quad \|S_{[k-3C,k+3C]}l\|_{L^1L^\infty} \lesssim 2^{-k},$$
$$\|S_{[k-3C,k+3C]}m\|_{L^1L^\infty} \lesssim 1$$

which leads to the trivial bound

$$\|P_{[k-3C,k+3C]}S_{[k-2C,k+2C]}u_k\|_{L^\infty L^2} \lesssim 2^{-k}\|\nabla S_{[k-2C,k+2C]}u_k\|_{L^\infty L^2}$$

This suffices for $s > 0$. In the case $s = 0$ we make the additional observation that at fixed time the expressions

$$\sum_{k=1}^{\infty}(S_{[k-3C,k+3C]}g^{i\alpha})\partial_i\partial_\alpha S_{[k-2C,k+2C]}u_k$$

$$\sum_{k=1}^{\infty}(S_{[k-3C,k+3C]}l^{\alpha})\partial_\alpha S_{[k-2C,k+2C]}u_k, \quad \sum_{k=1}^{\infty}(S_{[k-3C,k+3C]}m)S_{[k-2C,k+2C]}u_k$$

are combinations of paraproducts of $\partial^2 g^{i\alpha}$, ∂l^α respectively m with sub-sums of $|D_x|^{-1}\sum_{k=1}^{\infty}\nabla S_{[k-2C,k+2C]}u_k$ which are chosen so that the terms have disjoint Fourier support. Hence at each time t we can bound them in L^2 using the Coifman-Meyer paraproduct estimate [3] by

$$(\|\partial^2 g\|_{L^\infty} + \|\nabla l\|_{L^\infty} + \|m\|_{L^\infty})(\sum_{k=1}^{\infty} \||D_x|^{-1}\nabla S_{[k-2C,k+2C]}u_k\|_{L^2}^2)^{\frac{1}{2}}$$

It remains to consider the third sum. There the g factor is at a higher frequency, which is inherited by the product. In the case $s=0$ the above paraproduct argument still applies under the assumption that (13) holds.

In the case $s=2$ we use the stronger assumption (11) on the coefficients. We need to estimate

$$\|\partial_x^2 \sum_{k\geq 1} P_{>k+3C} S_{[k-2C,k+2C]}u_k\|_{L^1L^2}$$

The worst case is when the two derivatives in front apply to the coefficients of P. Then we need fixed time L^2 bounds for paraproducts of

$$\begin{array}{ccc} \partial_x^2 g & \text{with} & \sum_{k\geq 1} \partial_x \nabla S_{[k-2C,k+2C]}u_k \\ \partial_x^2 l & \text{with} & \sum_{k\geq 1} \nabla S_{[k-2C,k+2C]}u_k \\ \partial_x^2 m & \text{with} & \sum_{k\geq 1} S_{[k-2C,k+2C]}u_k \end{array}$$

Using the Coifman-Meyer paraproduct estimate we bound these paraproducts at fixed time by

$$\|\partial_x^2 g\|_{L^\infty}\|\partial_x\nabla S_{[k-2C,k+2C]}u_k\|_{L^2} + \|\partial_x^2 l\|_{L^n}\|\nabla S_{[k-2C,k+2C]}u_k\|_{L^{\frac{2n}{n-2}}}$$

$$+\|\partial_x^2 m\|_{L^{\frac{n}{2}}}\|\nabla S_{[k-2C,k+2C]}u_k\|_{L^{\frac{2n}{n-4}}}$$

and then use Sobolev embeddings for the second factor.

b) For $s=1$ the well-posedness follows directly from energy estimates using only (12). For $s=2,3$ we use the stronger condition (11) to differentiate the equation once, respectively twice and then use the energy estimates.

For $s=0$ we use an indirect argument. We seek the solution u of the form (58). Then v should solve

$$Pv = -P\sum_{k\geq 1} S_{[k-2C,k+2C]}u_k, \qquad v[0]=0$$

But by (a) the right hand side can be estimated in L^1L^2, therefore we can solve this equation using only the $s = 1$ result. This concludes the proof of the Proposition. □

References

[1] Hajer Bahouri and Jean-Yves Chemin. Equations d'ondes quasilineaires et effet dispersif. (Quasilinear wave equations and dispersive effect). *Int. Math. Res. Not.*, 1999(21):1141–1178, 1999.

[2] Hajer Bahouri and Jean-Yves Chemin. Equations d'ondes quasilineaires et estimations de Strichartz. (Quasilinear wave equations and Strichartz estimates). *Am. J. Math.*, 121(6):1337–1377, 1999.

[3] Ronald R. Coifman and Yves Meyer. *Au delà des opérateurs pseudo-différentiels*, volume 57 of *Astérisque*. Société Mathématique de France, Paris, 1978. With an English summary.

[4] J. Ginibre and G. Velo. Generalized Strichartz inequalities for the wave equation. *J. Funct. Anal.*, 133(1):50–68, 1995.

[5] L.V. Kapitanskij. Norm estimates in Besov and Lizorkin-Triebel spaces for the solutions of second-order linear hyperbolic equations. *J. Sov. Math.*, 56(2):2348–2389, 1991.

[6] Markus Keel and Terence Tao. Endpoint Strichartz estimates. *Amer. J. Math.*, 120(5):955–980, 1998.

[7] S. Klainerman. A commuting vectorfields approach to Strichartz-type inequalities and applications to quasi-linear wave equations. *Internat. Math. Res. Notices*, (5):221–274, 2001.

[8] Gerd Mockenhaupt, Andreas Seeger, and Christopher D. Sogge. Local smoothing of Fourier integral operators and Carleson-Sjölin estimates. *J. Amer. Math. Soc.*, 6(1):65–130, 1993.

[9] H. Smith and D. Tataru. Sharp local well-posedness results for the nonlinear wave equation. to appear, Annals of Math.

[10] Hart F. Smith. A parametrix construction for wave equations with $C^{1,1}$ coefficients. *Ann. Inst. Fourier (Grenoble)*, 48(3):797–835, 1998.

[11] Hart F. Smith and Christopher D. Sogge. On Strichartz and eigenfunction estimates for low regularity metrics. *Math. Res. Lett.*, 1(6):729–737, 1994.

[12] Hart F. Smith and Daniel Tataru. Sharp counterexamples for Strichartz estimates for low regularity metrics. *Math. Res. Lett.*, 9(2-3):199–204, 2002.

[13] Daniel Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. II. *Amer. J. Math.*, 123(3):385–423, 2001.

[14] Daniel Tataru. On the Fefferman-Phong inequality and related problems. *Comm. Partial Differential Equations*, 27(11-12):2101–2138, 2002.

[15] Daniel Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. III. *J. Amer. Math. Soc.*, 15(2):419–442 (electronic), 2002.

Department of Mathematics, Evans Hall, University of California at Berkeley, Berkeley, CA 94720-3840