Title
Global parametrices and dispersive estimates for variable coefficient wave equations

Permalink
https://escholarship.org/uc/item/5vf8q5b6

Authors
Metcalfe, Jason
Tataru, Daniel

Publication Date
2007-07-09

Peer reviewed

GLOBAL PARAMETRICES AND DISPERSIVE ESTIMATES FOR VARIABLE COEFFICIENT WAVE EQUATIONS

JASON METCALFE AND DANIEL TATARU

ABSTRACT. In this article we consider variable coefficient time dependent wave equations in $\mathbb{R} \times \mathbb{R}^n$. Using phase space methods we construct outgoing parametrices and prove Strichartz type estimates globally in time. This is done in the context of C^2 metrics which satisfy a weak asymptotic flatness condition at infinity.

CONTENTS

1. INTRODUCTION

Begin with the constant coefficient wave equation in $\mathbb{R} \times \mathbb{R}^n$, $n \geq 2$,

$$\Box u = (\partial_t^2 - \Delta)u = 0, \qquad u(0) = u_0, \ \ u_t(0) = u_1.$$

On one hand the energy is preserved,

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

where ∇ stands for the space-time gradient of the solution. On the other hand there is pointwise decay of waves with localized initial data. Precisely, if we set $u_0 = 0$ then

$$(1) \qquad \||D_x|^{\frac{1-n}{2}} u(t)\|_{L^\infty} \lesssim t^{-\frac{n-1}{2}} \|u_1\|_{L^1}$$

for all initial data u_1 with a dyadic frequency localization. As a consequence of this one obtains the Strichartz estimates, which have the form

$$(2) \qquad \||D_x|^{-\rho} \nabla u\|_{L^p L^q} \leq \|\nabla u_0\|_{L^2} + \|u_1\|_{L^2}.$$

This holds for all pairs (ρ, p, q) satisfying the relations $2 \leq p \leq \infty$, $2 \leq q \leq \infty$ and

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

with the exception of the forbidden endpoint $(1, 2, \infty)$ in dimension $n = 3$. All (ρ, p, q) satisfying these relations are called, in the sequel, Strichartz pairs. If the equality holds in the second part of (3) then the corresponding pair is called a sharp Strichartz pair.

In the sequel, we shall not explicitly deal with the case of $q = \infty$, allowing us to freely use Littlewood-Paley theory. When $q = \infty$, we only obtain estimates with L^∞ replaced by the appropriate $B^{-\rho}_{\infty,2}$ Besov spaces. With additional work, possibly akin to the modifications given to obtain the $\tilde{X}$ estimate in (25), we believe that the proper estimate can be recovered. In what follows, we shall also concentrate our efforts on the cases that the Strichartz pairs are sharp. This can be done without loss of generality since the remaining estimates can then be recovered using Sobolev embeddings.

A straightforward consequence of (2) is an estimate for solutions to the inhomogeneous problem

$$\Box u = f, \qquad u(0) = 0, \qquad u_t(0) = 0,$$

namely

$$(4) \qquad \||D_x|^{-\rho} \nabla u\|_{L^p L^q} \leq \|f\|_{L^1 L^2}.$$

The simplest case of (4) is the well-known energy estimate

$$(5) \qquad \|\nabla u\|_{L^\infty L^2} \leq \|f\|_{L^1 L^2}.$$

However, there is a larger family of estimates for solutions to the inhomogeneous wave equation where we also vary the norms in the right hand side,

$$(6) \qquad \||D_x|^{-\rho} \nabla u\|_{L^p L^q} \leq \||D_x|^{\rho_1} f\|_{L^{p_1'} L^{q_1'}}.$$

This holds for all Strichartz pairs (ρ, p, q), (ρ_1, p_1, q_1).

2

Estimates of the above type were first proved in the constant coefficient case in [3], [31]. Further references can be found in a more recent expository article [10]. The endpoint estimate $(p, q) = (2, \frac{2(n-1)}{n-3})$ was only recently obtained in [13] $(n \geq 4)$.

In this article we are interested in the variable coefficient case of these estimates, where we replace $\Box$ by a second order hyperbolic operator of the form[1]

$$P(t, x, D) = D_\alpha a^{\alpha\beta}(t, x) D_\beta + b^\alpha(t, x) D_\alpha + c(t, x).$$

where $D_k = \partial_k / i$. Here the matrix $a^{\alpha\beta}$ is assumed to have signature $(n, 1)$, and the time slices are assumed to be space-like, i.e. $a^{00} < 0$. Thus we consider evolutions of the form

(7) $$Pu = f, \qquad u(0) = u_0, \qquad u_t(0) = u_1.$$

Locally in time this problem is well understood. If the coefficients are smooth then parametrices are obtained using Fourier integral operators, and the Strichartz estimates were established in [19]. Operators with $C^{1,1}$ coefficients were first considered in [25], where a wave packet parametrix is constructed in all dimensions and the Strichartz estimates are proved in low dimension $n = 2, 3$. An alternative parametrix construction, based on the FBI transform, was later obtained in [32], [33] [35]. There the Strichartz estimates are obtained first for $C^{1,1}$ coefficients and then for $\nabla^2 a \in L^1 L^\infty$. Below this regularity threshold for the coefficients the full Strichartz estimates are lost (see [26],[28]), and one only retains partial results (see [33],[35]).

Our goal here is to study the global in time behavior, which is a considerably more difficult problem. The present article is inspired by an earlier article of the second author [37] which deals with the same issues for the corresponding Schrödinger equation. There are many similarities between the two problems, but also differences. In what follows we try to discuss both problems in parallel.

The dynamics for high frequencies are closely related to the Hamilton flow dynamics, although perhaps less so than in the case of the local in time problems.

A first phenomena that one needs to consider is that of refocusing, which in general precludes the dispersive estimates (1) even if we restrict ourselves to coefficients $a^{\alpha\beta}$ which are sufficiently small, smooth, compactly supported perturbations of the (Minkowski) identity. This is because even a small perturbation of the flat metric suffices in order to refocus a group of Hamilton flow rays originating at the same point and thus produce caustics.

---

[1]Here we employ the summation convention where repeated indices are implicitly summed. Repeated Greek letters $\alpha, \beta, \ldots$ are summed from 0 to n, where $D_0 = D_t$, and repeated Latin indices $i, j, \ldots$ are summed from 1 to n.

At the parametrix level this is reflected in the fact that a good parametrix along a ray which crosses through a bounded region is very difficult to construct. This is why, following [37], we construct an outgoing parametrix, which only requires the analysis of the outgoing Hamilton flow. The price we pay is that our parametrix cannot evolve only forward in time; instead it must have a forward and a backward component.

In the case of the Schrödinger equation, this is seen on arbitrarily small time scales due to the infinite speed of propagation; for the wave equation, on the other hand, a large time scale is needed.

A second feature is related to the long time behavior of the bicharacteristics. In the flat case all bicharacteristics are straight so they escape to infinity both forward and backward in time. However, in the variable coefficient case, it is possible to have trapped rays, which are confined to a bounded spatial region. These correspond to singularities which are largely concentrated in a bounded region and may destroy not only the dispersive estimates (1) but also the Strichartz estimates in (2). On the positive side, the nonexistence of trapped rays is a more stable phenomena; in particular, it cannot happen for small perturbations. Again, this obstruction is seen even on short time scales for the Schrödinger equation, but only on large time scales for the wave equation.

The local in time problem for the Schrödinger equation has been previously considered by other authors. Stafillani and Tataru [29] study C^2 compactly supported perturbations of the flat metric. Robbiano and Zuily [21] consider smooth asymptotically flat nontrapping metrics in $\mathbb{R}^n$ of the short range type and use a parametrix which is a Fourier integral operator with complex phase, relying considerably on Sjöstrand's theory of the FBI transform. Hassell-Tao-Wunsch [11] have a more direct parametrix construction emulating the model of the constant coefficient fundamental solution, which applies to smooth asymptotically conic manifolds with short range scattering metrics, extended shortly afterward to long range scattering metrics.

The dynamics for low frequencies are even more delicate, and for now there seem to be only two cases where anything at all can be said. The first is for sufficiently small perturbations of the flat metric, which is the case studied in [37] and here. The second is for time independent operators, with suitable spectral assumptions; for the Schrödinger equation this problem is considered in [17] (see also [2] and [22]), while for the wave it will be explored in another forthcoming paper.

A key part of the global decay estimates are the local energy estimates, which measure the local averaged decay in the L^2 settings. In the simplest form (see e.g. [1], [12], [14], [16],

[18], [20], [27], [30]), they are stated as

$$\|\nabla u\|_{L^2(\mathbb{R}\times B(0,R))} \lesssim R^{\frac{1}{2}}\|\nabla u(0)\|_{L^2}$$

when $\square u = 0$. Heuristically this is a reflection of the fact that waves move at speed $O(1)$ and thus spend a time $O(R)$ within a bounded spatial ball of radius R. These are the counterpart of the so called local smoothing estimates for the Schrödinger equation. See, e.g., [23], [39], [4], [7], and [5]. A significant difference is that, in the case of the Schrödinger equation the speed is proportional to the frequency; therefore one also gains half a derivative in the estimates.

The local energy estimates provide us with a convenient space to place the errors in our parametrix and also with a simpler setup in which to measure the decay of low frequency waves. In a nutshell, one of our main results asserts that

$$\text{Local energy estimates} \quad \Longrightarrow \quad \text{Strichartz estimates.}$$

The most important part of the article is the outgoing parametrix construction, for which we are able to adapt the ideas in [37]. The parametrix construction in [37] is based on the use of a time dependent FBI transform. However it does not use Sjöstrand's theory [24]. Instead, it takes advantage of the simpler approach introduced by the second author in [32], [33].

For more information about phase space transforms, we refer to [9] and [6]. One of the main starting points in the phase space analysis of pde's is Fefferman's article [8].

Simplified presentations of localized wave packet type parametrix constructions are now available in [15], [36]. These apply to evolutions of the form

$$(D_t + a^w(t, x, D))u = 0, \qquad u(0) = u_0$$

on the unit time scale, for symbols a which satisfy a partial S^0_{00} type condition

$$|\partial_x^\alpha \partial_\xi^\beta a(t, x, \xi)| \le c_{\alpha\beta}, \qquad |\alpha| + |\beta| \ge 2.$$

In the finite time analysis in [15], [36] the evolution is turned into a transport equation in the phase space modulo small errors. These parametrices are often useful in rescaled forms. However due to their finite time horizon, they cannot be directly applied to obtain optimal results for metrics which are not compactly supported perturbations of the identity.

In the long time analysis in [37] a time dependent FBI transform is used instead. A second order term in an asymptotic expansion becomes nontrivial, and the equation turns into a degenerate parabolic evolution in the phase space. Bounds for this evolution are then

obtained using the maximum principle. Fortunately for us, the main step in the parametrix construction in [37] can be applied directly here for half-waves. See Theorem 29.

Even though our parametrix is very precise, there are still errors which need to be controlled and this is done using localized energy estimates. We prove such estimates in the case of small perturbations of the flat metric. For large perturbations nontrapping may fail, and thus the localized energy estimates may fail. A nontrapping assumption would help with the localized energy estimates at high frequencies, but not for the low frequencies. Here we avoid this problem by using the localized energy estimates as an assumption for large perturbations of the flat metric. In the case of the Schrödinger equation, the local smoothing estimates for large perturbations were considered in [17]. See, also, [22]. In a follow-up paper we will consider the same issue in the case of the wave equation.

Scaling plays an essential role in our analysis. Modulo rescaling and Littlewood-Paley theory all our analysis is reduced to waves which have fixed frequency of size $O(1)$. Since waves have a propagation speed of size $O(1)$, our study of outgoing waves can be largely localized to cones of the form $\{|x| \approx |t|\}$. Certainly the exact flow cannot have a precise localization of this type due to the uncertainty principle. To compensate for this we introduce an artificial damping term which produces rapid decay of waves which do not have the above localization. This allows us to restrict our attention to the above cone modulo rapidly decreasing errors.

In the present article we consider global in time parametrices and Strichartz estimates for $C^{1,1}$ metrics in $\mathbb{R}^n$ which satisfy a weak asymptotic flatness assumption. Due to the global nature of the result it is convenient to consider scale invariant assumptions on the coefficients. We denote

$$A_j = \mathbb{R} \times \{2^j \leq |x| \leq 2^{j+1}\}, \qquad A_{<j} = \mathbb{R} \times \{|x| \leq 2^j\}.$$

Following [37], we assume that

$$(8) \qquad \sum_{j \in \mathbb{Z}} \sup_{A_j} |x|^2 |\nabla^2 a(t,x)| + |x||\nabla a(t,x)| + |a(t,x) - M_{1+n}| \leq \epsilon$$

where M_{1+n} is the $(n+1) \times (n+1)$ matrix $\mathrm{diag}(-1, 1, \ldots, 1)$ and, for the lower order terms,

$$(9) \qquad \sum_{j \in \mathbb{Z}} \sup_{A_j} |x|^2 |\nabla b(t,x)| + |x||b(t,x)| \leq \epsilon$$

$$(10) \qquad \sup_{\mathbb{R} \times \mathbb{R}^n} |x|^2 |c(t,x)| \leq \epsilon.$$

6

In some special cases we will need to strengthen the last condition to

$$(11) \qquad \sum_{j\in\mathbb{Z}} \sup_{A_j} |x|^4 |c(t,x)|^2 \leq \epsilon.$$

If ϵ is small enough then (8) precludes the existence of trapped rays, while for arbitrary ϵ it restricts the trapped rays to finitely many dyadic regions.

Before we state our main results we need to introduce the function spaces for the localized energy estimates. We consider a dyadic partition of unity in frequency,

$$1 = \sum_{k=-\infty}^{\infty} S_k(D_x),$$

and for each $k \in \mathbb{Z}$ we measure functions of frequency 2^k using the norm

$$\|u\|_{X_k} = 2^{k/2}\|u\|_{L^2(A_{<-k})} + \sup_{j\geq -k} \||x|^{-\frac{1}{2}}u\|_{L^2(A_j)}.$$

To measure the regularity of solutions to the wave equation, we use the global norm

$$\|u\|_{X^s}^2 = \sum_{k=-\infty}^{\infty} 2^{2sk}\|S_k u\|_{X_k}^2, \qquad -\frac{n+1}{2} < s < \frac{n+1}{2}.$$

All Schwartz functions $u \in \mathcal{S}(\mathbb{R}\times\mathbb{R}^n)$ have finite X^s norm. This allows us to define the space X^s as the completion of $\mathcal{S}(\mathbb{R}\times\mathbb{R}^n)$ with respect to the X^s norm. Its structure is clarified by the next lemma:

Lemma 1. [37] *a) ($s = 0$) We have*

$$(12) \qquad \sup_j \||x|^{-\frac{1}{2}}u\|_{L^2(A_j)} \lesssim \|u\|_{X^0}.$$

b) If $0 < s < \frac{n-1}{2}$ then the following Hardy type inequality holds for all $u \in \mathcal{S}(\mathbb{R}\times\mathbb{R}^n)$:

$$(13) \qquad \||x|^{-\frac{1}{2}-s}u\|_{L^2} \lesssim \|u\|_{X^s}.$$

c) If $\frac{n-1}{2} \leq s < \frac{n+1}{2}$ then we have the weaker bound

$$(14) \qquad \sum_{j=-\infty}^{\infty} 2^{-(1+2s)j}\|u - \bar{u}_{A_{<j}}\|_{L^2(A_{<j})}^2 \lesssim \|u\|_{X^s}^2$$

where the time dependent function $\bar{u}_{A_{<j}}$ stands for the spatial averages of u in $\{|x| \leq 2^j\}$.

The proof of the lemma is similar to the special case $s = \frac{1}{2}$ considered in [37] and is omitted. From the lemma we conclude that if s is as in case (a,b), then one can think of X as a space of distributions. On the other hand if s is as in case (c), then X has a BMO type structure, i.e. X is a space of distributions modulo time dependent constants.

7

Controlling the constants is important, particularly when it comes to localizing parametrices in dyadic regions. This is why we introduce also a stronger norm which removes the BMO structure, namely

$$\|u\|_{\tilde{X}^s}^2 = \|u\|_{X^s}^2 + \||x|^{-\frac{1}{2}-s}u\|_{L^2}^2, \qquad 0 < s < \frac{n+1}{2}.$$

This coincides with the X^s norm for $0 < s < \frac{n-1}{2}$. To simplify the exposition we also set $\tilde{X}^0 = X^0$.

For the inhomogeneous term in the equation, on the other hand, we use the dual space $Y^s = (X^{-s})'$ with norm

$$\|f\|_{Y^s}^2 = \sum_{k=-\infty}^{\infty} 2^{2sk}\|S_k f\|_{X_k'}^2, \qquad -\frac{n+1}{2} < s < \frac{n+1}{2}.$$

As X^s is the completion of $\mathcal{S}(\mathbb{R} \times \mathbb{R}^n)$, for $s > -\frac{n+1}{2}$, the space Y^s is dense in $S'(\mathbb{R} \times \mathbb{R}^n)$. In addition,

$$(15) \qquad \|u\|_{Y^s} \lesssim \||x|^{\frac{1}{2}-s}u\|_{L^2}, \qquad \frac{1-n}{2} < s < 0$$

and

$$(16) \qquad \|u\|_{Y^0} \lesssim \sum_j \||x|^{\frac{1}{2}}u\|_{L^2(A_j)}.$$

Definition 2. *We say that the operator P satisfies the $\dot{H}^s$ localized energy estimates if for each initial data $(u_0, u_1) \in \dot{H}^{s+1} \times \dot{H}^s$ and each inhomogeneous term $f \in L^1\dot{H}^s + Y^s$, there exists a unique solution u to (7) with $\nabla u \in L^\infty \dot{H}^s \cap X^s$ which satisfies the bound*

$$(17) \qquad \|\nabla u\|_{L^\infty \dot{H}^s \cap X^s} \lesssim \|\nabla u(0)\|_{\dot{H}^s} + \|f\|_{L^1\dot{H}^s+Y^s}.$$

In this context the lower order terms can be often treated as negligible perturbations:

Lemma 3. *a) Let b be as in (9) and*

$$(18) \qquad |s| \leq 1, \qquad |s| < \frac{n-1}{2}.$$

Then

$$(19) \qquad \|b\nabla u\|_{Y^s} \lesssim \epsilon\|\nabla u\|_{X^s}.$$

b) Let $n \geq 3$, c be as in (10) and $-1 < s < 0$. Then

$$(20) \qquad \|cu\|_{Y^s} \lesssim \epsilon\|\nabla u\|_{X^s}.$$

c) Let $n \geq 4$, c be as in (11) and $s = -1, 0$. Then

$$(21) \qquad \|cu\|_{Y^s} \lesssim \epsilon\|\nabla u\|_{X^s}.$$

8

d) Let $n = 3$ and c be as in (11). Then

$$(22) \qquad \|cu\|_{Y^0} \lesssim \epsilon \|u\|_{\tilde{X}^1}.$$

The localized energy estimates hold under the assumption that the coefficients $a^{\alpha\beta}$ are a small perturbation of the Minkowski metric.

Theorem 4. *Assume that the coefficients $a^{\alpha\beta}$, b^α satisfy (8), (9) with an ϵ which is sufficiently small. Assume also that $c = 0$. Then the operator P satisfies the $\dot{H}^s$ localized energy estimates globally in time for s as in (18).*

A general coefficient b and a coefficient c can be dealt with perturbatively but only in dimension $n \geq 3$:

Corollary 5. *a) Let $n \geq 3$ and $a^{\alpha\beta}$, b^α and c as in (8), (9), (10) with an ϵ which is sufficiently small. Then the operator P satisfies the $\dot{H}^s$ localized energy estimates globally in time for*

$$-1 < s < 0.$$

b) Let $n \geq 4$ and $a^{\alpha\beta}$, b^α and c as in (8), (9), (11) with an ϵ which is sufficiently small. Then P satisfies the $\dot{H}^s$ localized energy estimates globally in time for $s = -1, 0$.

Once we have the local energy estimates, the next step is to construct an outgoing parametrix which has good time decay and suitable error bounds in the dual local energy spaces. The parametrix is constructed at first in the case of a small perturbation of the flat metric. This leads to our main scale invariant Strichartz estimate:

Theorem 6. *Assume that $c = 0$ and the coefficients $a^{\alpha\beta}$, b^α satisfy (8), (9) with an ϵ which is sufficiently small. Let (ρ_1, p_1, q_1) and (ρ_2, p_2, q_2) be two Strichartz pairs and s as in (18). Then the solution u to (7) satisfies*

$$(23) \qquad \|\nabla u\|_{|D_x|^{\rho_1-s}L^{p_1}L^{q_1} \cap X^s} \lesssim \|\nabla u(0)\|_{\dot{H}^s} + \|f\|_{|D_x|^{-\rho_2-s}L^{p_2'}L^{q_2'}+Y^s}.$$

A zero order term c can also be added to P subject to the conditions in Corollary 5.

If ϵ is large then any localized energy estimates require an additional nontrapping condition. Even then the nontrapping can at most guarantee local in time bounds. However, we can still prove a conditional result:

Theorem 7. *a) Assume that $c = 0$ and the coefficients $a^{\alpha\beta}$, b^α satisfy (8), (9). Then for every Strichartz pair (ρ, p, q) and s as in (18), we have*

$$(24) \qquad \||D_x|^{s-\rho}\nabla u\|_{L^pL^q} \lesssim \|\nabla u\|_{X^s \cap L^\infty \dot{H}^s} + \|Pu\|_{Y^s}.$$

In addition there is a parametrix K for P which satisfies

$$(25) \qquad \|\nabla K f\|_{|D_x|^{\rho_1-s} L^{p_1} L^{q_1} \cap X^s} + \|Kf\|_{\tilde{X}^{s+1}} + \|(PK-I)f\|_{Y^s} \lesssim \|f\|_{|D_x|^{-\rho_2-s} L^{p_2'} L^{q_2'}}$$

for any two Strichartz pairs (ρ_1, p_1, q_1) and (ρ_2, p_2, q_2). A zero order term c can also be added to P subject to the conditions in Corollary 5.

b) Assume that in addition the operator P satisfies the $\dot{H}^s$ localized energy estimates. Then the solution u to (7) satisfies the full Strichartz estimates in (23).

In applications one might be concerned that the condition (8) imposes the nontrivial restriction $a(t,0) = M_{1+n}$. This is true, but it is needed only because we are allowing the derivatives of the coefficients to be singular at 0. Otherwise, such a restriction is unnecessary:

Remark 8. *Assume that the the condition (8) on the coefficients $a^{\alpha\beta}$ is modified for $|x| < 1$ to*

$$\sup_{|x|<1} \left(|\nabla^2 a(t,x)| + |\nabla a(t,x)| + |a(t,x) - M_{1+n}| \right) \leq \epsilon,$$

and similarly for (9), (10), and (11). Assume also that for $k > 0$ the definition of the space X_k is changed to

$$\|u\|_{X_k} = \|u\|_{L^2(A_{<-0})} + \sup_{j \geq 0} \| |x|^{-\frac{1}{2}} u \|_{L^2(A_j)}.$$

Then the results in Theorems 4,6,7 remain valid. Their proofs are essentially identical with only a few obvious changes.

The paper is structured as follows. After introducing some notations in the next section and making a reduction to the case $a^{00} = -1$ in the third section, in Section 4 we consider the paradifferential calculus associated to our problem. More precisely, we show that without any loss we are allowed to mollify the coefficients $a^{\alpha\beta}$ on a suitable x dependent scale. This allows us to reduce our analysis to problems which are frequency localized in dyadic regions. We also prove the bound, Lemma 3, for the lower order terms.

Section 5 contains the proof of the localized energy estimates in Theorem 4. The main step of the proof is carried out in a frequency localized context and involves a Morawetz type multiplier technique.

After making a reduction to the half-wave operator in Section 6, we state our main result on the existence of frequency localized outgoing parametrices for half-wave equations, namely Proposition 15 in Section 7. Using this result we conclude the proof of Theorems 6,7.

The rest of the paper is devoted to the parametrix construction. This largely follows [37]. In Section 8 we introduce the pseudodifferential operators and the phase space transforms.

An important role is played by the conjugation of pdo's with respect to the phase space transform, for which we use some results from [34], [36]. In a first step, the parametrix is obtained in Section 9 in the case of evolutions governed by a pseudodifferential operator a^w whose symbol satisfies a suitable smallness condition uniformly in x; for this we are fortunately able to apply directly the result proved in [37]. This construction is then transferred in Section 10 to small perturbations of half-waves via conjugation with respect to the flat half-wave flow. Finally to arrive at the desired setup we need to insure that the parametrix is localized in outgoing propagation cones. This is done in the last section by means of choosing a suitable damping term in the equation.

2. Notations

We consider a smooth spatial Littlewood-Paley decomposition

$$1 = \sum_{j=-\infty}^{\infty} \chi_j(x) \qquad \text{supp } \chi_j \subset \{2^{j-1} < |x| < 2^{j+1}\}.$$

We also set

$$\chi_{<j} = \sum_{k<j} \chi_k.$$

Given ϵ as in (8), we can find a sequence $\epsilon_j \in l^1$ so that

(26) $$\sup_{A_j} |x|^2 |\nabla^2 a(t,x)| + |x||\nabla a(t,x)| + |a(t,x) - M_{1+n}| \leq \epsilon_j$$

and

$$\sum \epsilon_j \lesssim \epsilon.$$

Without any restriction in generality, we can assume that ϵ_j is slowly varying, say

(27) $$|\ln \epsilon_j - \ln \epsilon_{j-1}| \leq 2^{-10}.$$

We also choose a function ϵ in $\mathbb{R}^+$ with the property that

$$\epsilon_j < \epsilon(s) < 2\epsilon_j \qquad \text{for } 2^j < s < 2^{j+1},$$

and so that

$$|\epsilon'(s)| \leq 2^{-5} s^{-1} \epsilon(s).$$

This implies that

$$\int_0^{\infty} \frac{\epsilon(s)}{s} \, ds \approx \epsilon.$$

11

We also define $\epsilon_k(s)$ so that

$$\epsilon_k(s) \approx \epsilon_j, \quad s \approx 2^j, \quad j \geq -k$$

$$\epsilon_k(s) \approx \epsilon_{-k}, \quad s \leq 2^{-k}.$$

Note that

$$\epsilon_k(|x|) \approx \epsilon(2^{-k} + |x|).$$

We consider a frequency Littlewood-Paley decomposition

$$1 = \sum_{j=-\infty}^{\infty} S_j(D_x)$$

where

$$\text{supp } s_j \subset \{2^{j-1} < |\xi| < 2^{j+1}\}.$$

We also use the related notations $S_{<k}$, $S_{>k}$, etc.

We say that a function f is localized at frequency 2^k if $\hat{f}$ is supported in $\{2^{k-1} < |\xi| < 2^{k+1}\}$. An operator K is localized at frequency 2^k if for any f localized at frequency 2^k its image Kf is frequency localized in $\{2^{k-10} < |\xi| < 2^{k+10}\}$.

3. A MINOR SIMPLIFICATION

The aim of this section is to reduce the problem to the case when P has the form

$$P = -D_t^2 + 2D_i a^{i0} D_t + D_i a^{ij} D_j + b^\alpha D_\alpha + c,$$

and once this is accomplished, P will be taken to be of this form throughout the sequel. To arrange that $a^{00} = -1$ we multiply the operator P by $-(a^{00})^{-1}$ which satisfies the same bounds as a^{00}. This modifies the other coefficients

$$a^{\alpha\beta} \to -a^{\alpha\beta}(a^{00})^{-1}, \qquad b^j \to -b^j(a^{00})^{-1} + D_\alpha((a^{00})^{-1})a^{\alpha j},$$

$$b^0 \to -b^0(a^{00})^{-1} - (a^{00})^{-1}(D_t a^{00}) + D_j((a^{00})^{-1})a^{j0}, \qquad c \to -c(a^{00})^{-1},$$

and it is easy to verify that the assumptions (8), (9), (10), and (11) are left unchanged.

To express the second term in the form above we note that

$$D_t a^{0i} D_i = D_i a^{0i} D_t + (D_t a^{0i})D_i - (D_i a^{0i})D_t.$$

This changes the coefficients b^α but still within the allowed limits. Arguing similarly and picking up only lower order errors within the limits, we may assume that $a^{ij} = a^{ji}$. We also note that the coefficient c is not affected by these transformations.

To conclude our simplification, we need to verify that our function spaces are not affected by multiplication by $(a^{00})^{-1}$.

Lemma 9. *Let a be as in (8) and s as in (18). Then*

(28)
$$\|af\|_{Y^s} \lesssim \|f\|_{Y^s}.$$

In addition, for all Strichartz pairs (ρ, p, q), we have

(29)
$$\|af\|_{|D_x|^{-\rho-s}L^{p'}L^{q'}} \lesssim \|f\|_{|D_x|^{-\rho-s}L^{p'}L^{q'}}.$$

Proof. We write (28) in the dual form

$$|\langle af, u \rangle| \lesssim \|f\|_{Y^s}\|u\|_{X^{-s}}$$

and take a simultaneous Littlewood-Paley decomposition of the three factors a, f and u. Nontrivial output is obtained when the two larger frequencies are comparable. Hence there are three cases to consider. The trivial one is when the a factor has the low frequency. For the remaining two cases it suffices to prove the off-diagonal decay

$$|\langle S_k a S_k f, S_j u \rangle| \lesssim 2^{(s-\delta)(k-j)}\|S_k f\|_{X_k'}\|S_j u\|_{X_j}, \qquad j \le k,$$

respectively

$$|\langle S_k a S_j f, S_k u \rangle| \lesssim 2^{(-s-\delta)(k-j)}\|S_j f\|_{X_j'}\|S_k u\|_{X_k}, \qquad j \le k$$

for s as in (18). This follows from the definition of the X_k' and X_k norms combined with the bound on $S_k a$,

$$|S_k a(x)| \lesssim 2^{-2k}(2^{-2k} + |x|^2)^{-1},$$

and an uncertainty principle bound for the low frequency factor on the dual spatial scale,

$$\|S_j u\|_{L^2 L^\infty(A_{<-j})} \lesssim 2^{\frac{n-1}{2}j}\|S_j u\|_{X_j}, \qquad \|S_j f\|_{L^2 L^\infty(A_{<-j})} \lesssim 2^{\frac{n+1}{2}j}\|S_j f\|_{X_j'}.$$

The details are straightforward and are left for the reader.

We now prove (29). The time variable plays no role and is neglected in the sequel. We shall use the following variant of a Moser estimate:

$$\|fg\|_{\dot{W}^{s,p}} \lesssim \|f\|_{L^{q_1}}\|g\|_{\dot{W}^{s,q_2}} + \|g\|_{L^{r_1}}\|f\|_{\dot{W}^{s,r_2}}$$

with $s > 0$, $1 < p < \infty$, and

$$\frac{1}{p} = \frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{r_1} + \frac{1}{r_2}, \quad q_2, r_2 \in (1, \infty), \quad q_1, r_1 \in (1, \infty].$$

See, e.g., [38, §2.1, Proposition 1.1, p. 105]. We first assume that $s + \rho \ge 0$ and apply the above estimate to af. This yields

(30)
$$\|af\|_{\dot{W}^{\rho+s,q'}} \lesssim \|a\|_\infty \|f\|_{\dot{W}^{\rho+s,q'}} + \|a\|_{\dot{W}^{\rho+s,r_1}}\|f\|_{r_2}.$$

13

The first term on the right is trivially bounded by the right side of (29). For the second term on the right, we first pass from the Sobolev space to an appropriate Besov space, and then we use the following consequence of (8)

$$|S_l a(t, x)| \lesssim \begin{cases} 2^{-2m-2l}\epsilon_m & 2^m < |x| < 2^{m+1}, \; m+l \geq 0 \\ \epsilon_{-l} & |x| < 2^{-l}. \end{cases}$$

Indeed, we see that

$$\||D_x|^{\rho+s} a\|_{L^{r_1}} \lesssim \sum_k 2^{(\rho+s)k}\|S_k a\|_{L^{r_1}}$$

$$\lesssim \sum_k 2^{(\rho+s)k}\left[\sum_{m \geq -k} 2^{-2m-2k}\epsilon_m 2^{nm/r_1} + 2^{-nk/r_1}\epsilon_{-k}\right].$$

Using that $\rho = \frac{n+1}{4} - \frac{n+1}{2q}$ for a sharp Strichartz pair, we have that $q' < \frac{n}{\rho+s}$. Thus, we may choose $r_1 = \frac{n}{\rho+s}$. Substituting this in the previous calculation, noting that $\rho + s < 2$ for ρ a part of a sharp Strichartz pair as above and s as in (18), and using the summability of $\{\epsilon_m\}$, we have that

$$\|a\|_{\dot{W}^{\rho+s,r_1}} < \infty, \quad r_1 = \frac{n}{\rho+s}.$$

If we now apply Sobolev embeddings to the second factor, we see that the second term in the right of (30) is also bounded by the right side of (29). As we may use a dual argument if $s + \rho < 0$, this completes the proof. □

4. The paradifferential calculus

In order to reduce the problem to a frequency localized context and to simplify the parametrix construction it is convenient to localize the coefficients in frequency. This is somewhat more complicated than usual because the frequency localization scale needs to depend on the spatial scale.

It suffices to work with only the principal part of the operator P, which we denote by

$$P_a = -D_t^2 + 2D_i a^{i0} D_t + D_i a^{ij} D_j.$$

Given a frequency scale k we define the regularized coefficients

$$a_{(k)}^{i\beta} = \delta^{i\beta} + \sum_{l < k-4} (S_{<l}\chi_{<k-2l})S_l a^{i\beta}.$$

Correspondingly we define the mollified operators

$$P_{(k)} = -D_t^2 + 2D_i a_{(k)}^{i0} D_t + D_i a_{(k)}^{ij} D_j$$

14

which are used on functions of frequency 2^k. Roughly speaking, their coefficients are frequency localized in the region

$$|\xi| \ll 2^k(1 + 2^k|x|)^{-\frac{1}{2}}.$$

We also introduce a global mollified operator

$$\tilde{P} = \sum_{k=-\infty}^{\infty} P_{(k)} S_k.$$

Due to (26) and to the fact that the ϵ_j's are slowly varying, it follows that the dyadic parts of the coefficients will satisfy the bounds

$$(31) \qquad |S_l a^{i\beta}(t, x)| \lesssim \begin{cases} 2^{-2m-2l}\epsilon_m & 2^m < |x| < 2^{m+1},\ m+l \geq 0 \\ \epsilon_{-l} & |x| < 2^{-l}. \end{cases}$$

This also allows us to obtain bounds on the coefficients of $P_{(k)}$,

$$(32) \qquad \begin{aligned} |\partial^\alpha(a_{(k)}^{i\beta}(x) - \delta^{i\beta})| &\leq c_\alpha \epsilon_k(|x|)2^{|\alpha|k}(1 + 2^k|x|)^{-|\alpha|}, & |\alpha| \leq 2 \\ |\partial^\alpha a_{(k)}^{i\beta}(x)| &\leq c_\alpha \epsilon_k(|x|)2^{|\alpha|k}(1 + 2^k|x|)^{-1-\frac{|\alpha|}{2}}, & |\alpha| \geq 2. \end{aligned}$$

The main result of this section shows that we can freely replace P_a by $\tilde{P}$ in Theorems 4, 6, 7(a). It also shows that at frequency 2^k the operators $\tilde{P}$ and $P_{(k)}$ are interchangeable.

Proposition 10. *Assume that the coefficients a satisfy* (8). *Then*

$$(33) \qquad \|(\tilde{P} - P_{(k)})S_l u\|_{X'_k} \lesssim \epsilon\|S_l \nabla u\|_{X_k}, \qquad |l - k| \leq 2$$

$$(34) \qquad \|[P_{(k)}, S_k]u\|_{X'_k} \lesssim \epsilon\|\nabla u\|_{X_k}.$$

In addition, for s as in (18) *the following estimate holds:*

$$(35) \qquad \|(P_a - \tilde{P})u\|_{Y^s} \lesssim \epsilon\|\nabla u\|_{X^s}.$$

Proof. The proof is very similar to the analogous one in [37]. We begin with (35), and write

$$P_a - \tilde{P} = P_{low} + P_{mid} + P_{high}$$

15

with[2]

$$P_{low} = \sum_{k=-\infty}^{\infty} D_i\Big(\sum_{l<k-4}(S_{<l}\chi_{\geq k-2l})S_l a^{ij}\Big)D_j S_k + \sum_{k=-\infty}^{\infty} D_i\Big(\sum_{l<k-4}(S_{<l}\chi_{\geq k-2l})S_l a^{i0}\Big)D_t S_k$$

$$P_{mid} = \sum_{k=-\infty}^{\infty}\sum_{l=k-4}^{k+4} D_i(S_l a^{ij})D_j S_k + \sum_{k=-\infty}^{\infty}\sum_{l=k-4}^{k+4} D_i(S_l a^{i0})D_t S_k$$

$$P_{high} = \sum_{k=-\infty}^{\infty}\sum_{l>k+4} D_i(S_l a^{ij})D_j S_k + \sum_{k=-\infty}^{\infty}\sum_{l>k+4} D_i(S_l a^{i0})D_t S_k.$$

Let us examine in detail the second term in each of the expressions above. The bounds for the remaining terms follow from similar arguments.

For P_{low}, we notice that the output is at the same frequency 2^k as the input. Since the factor D_i contributes a factor of 2^k, it suffices to show

$$(36) \qquad \Big\|\sum_{l<k-4}(S_{<l}\chi_{\geq k-2l})S_l a^{i0} v\Big\|_{X'_k} \lesssim \epsilon 2^{-k}\|v\|_{X_k}.$$

Here, we shall use the bound

$$|S_{<l}\chi_{\geq k-2l}(x)| \leq \begin{cases} 2^{4l-4k}, & |x| < 2^{k-2l-2}, \\ 1, & |x| \geq 2^{k-2l-2}, \end{cases} \qquad l < k-4.$$

For $|x| \approx 2^m$, $m \geq -k$, we use this and (31) to see that

$$\Big|\sum_{l<k-4}(S_{<l}\chi_{\geq k-2l})S_l a^{i0}\Big| \lesssim \sum_{l=-\infty}^{-m-1} 2^{4l-4k}\epsilon_{-l} + \sum_{l=-m}^{\frac{k-m}{2}-1} \epsilon_m 2^{-2m-2l}2^{4l-4k} + \sum_{l=\frac{k-m}{2}}^{m-4} \epsilon_m 2^{-2m-2l}$$

$$\lesssim 2^{-m-k}\epsilon_m.$$

For $|x| < 2^{-k}$, the argument is easily modified to give

$$\Big|\sum_{l<k-4}(S_{<l}\chi_{\geq k-2l})S_l a^{i0}\Big| \lesssim \epsilon_{-k}, \qquad |x| < 2^{-k},$$

which, combined with the previous estimate, yields the desired bound (36).

For input frequency 2^k, P_{mid} permits output frequencies 2^h for all $h \leq k+4$. We take $l = k$ for simplicity of exposition, and consider separately low and high dimensions.

---

[2]To be explicit with the order of operations, the Littlewood-Paley projectors take precedence, followed by multiplication and differentiation using right associativity, and finally addition. Thus, for example, the first term in $P_{low}u$ is understood to be

$$\sum_{k=-\infty}^{\infty} D_i\Big[\Big(\sum_{l<k-4}(S_{<l}\chi_{\geq k-2l})(S_l a^{ij})\Big)\{D_j(S_k u)\}\Big].$$

In low dimension $n = 2, 3$ the bound for P_{mid} follows from the off-diagonal decay

$$\|S_h D_i (S_k a^{i0} D_t S_k u)\|_{X'_h} \lesssim \epsilon 2^{\frac{n+1}{2}(h-k)} \|D_t S_k u\|_{X_k}, \quad h \leq k + 2,$$

or more simply,

$$(37) \qquad \|S_h(S_k a^{i0} v)\|_{X'_h} \lesssim \epsilon 2^{\frac{n+1}{2}(h-k)} 2^{-h} \|v\|_{X_k}.$$

Writing

$$S_h(S_k a^{i0} v) = S_h(\chi_{<-k} S_k a^{i0} v) + \sum_{m \geq -k} S_h(\chi_m S_k a^{i0} v),$$

it is sufficient to show that

$$\|S_h(\chi_m S_k a^{i0} v)\|_{X'_h} \lesssim \epsilon_m 2^{\frac{n+1}{2}(h-k)} 2^{-h} \|v\|_{X_k}, \quad m \geq -k,$$
$$\|S_h(\chi_{<-k} S_k a^{i0} v)\|_{X'_h} \lesssim \epsilon_{-k} 2^{\frac{n+1}{2}(h-k)} 2^{-h} \|v\|_{X_k}.$$

For the former, we apply (31) and see that it suffices to show

$$(38) \qquad \|S_h(\chi_m v)\|_{X'_h} \lesssim 2^{\frac{n-1}{2}(h-k)} 2^{k + \frac{3m}{2}} \|v\|_{L^2(|x| \approx 2^m)}, \quad m + k \geq 0.$$

By interpolating the estimates

$$\|S_h(\chi_m v)\|_{L^2} \lesssim \|\chi_m v\|_{L^2}, \quad \|x S_h(\chi_m v)\|_{L^2} \lesssim 2^m \|\chi_m v\|_{L^2}, \quad m + h \geq 0,$$

we obtain

$$(39) \qquad \|S_h(\chi_m v)\|_{X'_h} \lesssim 2^{\frac{m}{2}} \|\chi_m v\|_{L^2}.$$

Recalling that we are in the case when $h < k - 2$, this yields (38) when $m + h \geq 0$.

For $m + h < 0$ we have improved bounds

$$\|S_h(\chi_m v)\|_{L^2} \lesssim 2^{\frac{n(m+h)}{2}} \|\chi_m v\|_{L^2}, \quad \|x S_h(\chi_m v)\|_{L^2} \lesssim 2^{-h} 2^{\frac{n(m+h)}{2}} \|\chi_m v\|_2, \quad m + h < 0,$$

which upon interpolation yields

$$\|S_h(\chi_m v)\|_{X'_h} \lesssim 2^{\frac{m}{2}} 2^{\frac{n-1}{2}(m+h)} \|\chi_m v\|_{L^2},$$

and implies (38). The bound when χ_m is replaced by $\chi_{<-k}$ is identical to the $m + h < 0$ argument above.

In high dimension $n \geq 4$ the bound (37) is replaced by

$$(40) \qquad \|S_h(S_k a^{i0} v)\|_{X'_h} \lesssim \epsilon 2^{2(h-k)} 2^{-h} \|v\|_{X_k}.$$

whose proof is similar. The only difference is that now the worst case is $m = -h$, whereas in low dimension the worst case is $m = -k$ ($n = 2$) respectively $-k \leq m \leq -h$ ($n = 3$).

17

It remains to consider P_{high} whose output is at frequencies 2^l with $l > k$ where 2^k is the input frequency. In low dimension $n = 2, 3$ it suffices to show that

$$\|S_l D_i(S_l a^{i0} D_t S_k u)\|_{X'_l} \lesssim \epsilon 2^{\frac{n-1}{2}(k-l)} \|D_t S_k u\|_{X_k}, \quad l > k + 4$$

or

(41) $$\|S_l(S_l a^{i0} S_k v)\|_{X'_l} \lesssim \epsilon 2^{\frac{n-1}{2}(k-l)} 2^{-l} \|S_k v\|_{X_k}, \quad l > k + 4$$

which follows by duality from (37).

In high dimension $n \geq 4$ instead of (41) we have

(42) $$\|S_l a^{i0} S_k v\|_{X'_l} \lesssim \epsilon 2^{k-l} 2^{-l} \|v\|_{X_k}$$

which is still sufficient except for the endpoint $s = 1$. At the endpoint we are left with no off-diagonal decay. To compensate for that we need a stronger version of (42), namely

(43) $$\|S_{>k+4} D^2_x a^{i0} S_k v\|_{Y^0} \lesssim \epsilon_{-k} 2^k \|v\|_{X_k}.$$

By (16) it suffices to show that

$$\sum_{j=-\infty}^{\infty} \||x|^{\frac{1}{2}} S_{>k+4} D^2_x a^{i0} S_k v\|_{L^2(A_j)} \lesssim \epsilon_{-k} 2^k \|v\|_{X_k}.$$

But this follows from the bound

$$|S_{>k+4} D^2_x a^{i0}| \lesssim \epsilon_k(|x|) |x|^{-2},$$

and in the case that $j < -k$, the following consequence of Bernstein's inequality

$$\|\chi_{<-k} S_k v\|_{L^2 L^\infty} \lesssim 2^{\frac{n-1}{2}k} \|S_k v\|_{X_k}.$$

The constant ϵ_{-k} is obtained since the worst case is when $j = -k$, with exponential decay away from it.

The estimate (33) follows from arguments similar to those used for P_{low}. To show (34), it would suffice to show

$$\|[S_k, a^{ij}_{(k)}]v\|_{X'_k} \lesssim \epsilon 2^{-k} \|v\|_{X_k}$$

and the equivalent statement with $j = 0$, which follows directly from (32) with $\alpha = 1$. □

In a similar manner we prove the bounds of Lemma 3 which show that in high dimension we can completely dispense with lower order terms.

Proof of Lemma 3. This is again quite similar to the related result from [37].

18

From (9) we may obtain the following bounds on the frequency localized pieces of the coefficients

$$(44) \qquad |S_k b(x)| \lesssim 2^{-k} \epsilon_k(|x|)(2^{-k} + |x|)^{-2}, \quad |S_{<k} b(x)| \lesssim \epsilon_k(|x|)(2^{-k} + |x|)^{-1}.$$

To prove (19), we first expand

$$b^\alpha \nabla_\alpha u = \sum_j (S_{<j-4} b^\alpha \nabla_\alpha) S_j u + \sum_j \sum_{|k-j| \leq 4} (S_k b^\alpha \nabla_\alpha) S_j u + \sum_j \sum_{k>j+4} (S_k b^\alpha \nabla_\alpha) S_j u.$$

The easiest case is the first term: the low-high interactions. Here the output is at the same frequency range as the input. Thus, it would suffice to show

$$\|(S_{<j-4} b) v\|_{X'_j} \lesssim \epsilon \|v\|_{X_j}.$$

As the $\epsilon(|x|)$ provides summability, this follows directly from (44).

The third case is the high-low interactions. Here, for input at frequency 2^j, the output is at frequency 2^k with $k > j + 4$. We, thus, measure the output in X'_k. In low dimension $n = 2, 3$ it suffices to show that

$$\|S_k b S_j v\|_{X'_k} \lesssim \epsilon 2^{\frac{n-1}{2}(j-k)} \|S_j v\|_{X_j}, \quad j + 4 < k.$$

This, however, is just a reformulation of (41) since b has exactly the same regularity as Da.

In high dimension $n \geq 4$ we have the similar relation

$$\|S_k b S_j v\|_{X'_k} \lesssim \epsilon 2^{j-k} \|S_j v\|_{X_j}, \quad j + 4 < k$$

which covers all cases but $s = 1$. For $s = 1$ we replace this with (43) with Db instead of $D^2 a$.

The remaining case, the high-high interactions, is dual to the previous case. This finishes the proof of (19).

Finally, (20), (21), and (22) follow directly from the embeddings (12), (13) and their duals (16), (15). $\qquad \square$

5. LOCALIZED ENERGY ESTIMATES

Here we prove Theorem 4. We can assume that P has the form in Section 3 with $c = 0$. Also due to (19) we can take $b = 0$.

The theorem is proved via a positive commutator method. Let $(\alpha_m)_{m \in \mathbb{Z}}$ be a positive slowly varying sequence with $\sum \alpha_m = 1$. Correspondingly we define the space $X_{k,\alpha}$ with norm

$$\|u\|^2_{X_{k,\alpha}} = 2^k \|u\|^2_{L^2(A_{<-k})} + \sum_{j \geq -k} \alpha_j \| |x|^{-\frac{1}{2}} u \|^2_{L^2(A_j)}$$

and the dual space

$$\|u\|_{X'_{k,\alpha}}^2 = 2^{-k}\|u\|_{L^2(A_{<-k})}^2 + \sum_{j \geq -k} \alpha_j^{-1} \||x|^{\frac{1}{2}} u\|_{L^2(A_j)}^2.$$

The key step in the proof of Theorem 4 is the following frequency localized estimate:

Proposition 11. *Assume that ϵ is sufficiently small. Then the bound*

$$\text{(45)} \qquad \|\nabla u\|_{L^\infty L^2 \cap X_{k,\alpha}} \lesssim \|\nabla u(0)\|_{L^2} + \|P_{(k)} u\|_{L^1 L^2 + X'_{k,\alpha}}$$

holds for all functions $u \in L^\infty L^2 \cap X_{k,\alpha}$ localized at frequency 2^k, uniformly with respect to all slowly varying sequences (α_m) (as defined in (27)) with

$$\text{(46)} \qquad \sum_{m=-k}^\infty \alpha_m = 1.$$

Proof. By rescaling, the problem reduces to the case when $k = 0$. We may without loss increase the sequence (α_m) so that it remains slowly varying and

$$\text{(47)} \qquad \begin{cases} \alpha_0 \approx 1, \\ \sum_{m>0} \alpha_m \approx 1, \\ \epsilon_m \leq \epsilon \alpha_m. \end{cases}$$

This is accomplished by redefining

$$\alpha_m := \alpha_m + \frac{\epsilon_m}{\epsilon} + 2^{-2^{-10}m}.$$

Since $\epsilon \ll 1$, we can fix another small parameter $\epsilon \ll \delta \ll 1$ so that

$$\epsilon_m \ll \delta \alpha_{m + \log_2 \delta}.$$

Associating to (α_m) a function $\alpha(s) = \alpha_0(s)$ whose definition is analogous to that of $\epsilon_0(s)$ in Section 2, we have, from the last property,

$$\epsilon_0(s) \lesssim \epsilon \alpha(s) \ll \delta \alpha(\delta s).$$

The proof has three ingredients, the first of which is the classical energy estimate. Since ϵ is small it follows that the ∂_t vector field is time-like, and the corresponding energy

$$E_0(u) = \frac{1}{2}\|D_t u\|^2 + \frac{1}{2}\langle a_{(0)}^{ij} D_j u, D_i u \rangle$$

is positive definite. Here and throughout $\langle \cdot, \cdot \rangle$ is the $L_x^2(\mathbb{R}^n)$ inner product. The time derivative of this energy is

$$\frac{d}{dt} E_0(u) = \Im \langle P_{(0)} u, D_t u \rangle + \langle (\partial_i a_{(0)}^{i0}) D_t u, D_t u \rangle + \frac{1}{2}\langle (\partial_t a_{(0)}^{ij}) D_j u, D_i u \rangle.$$

20

The second component of the proof is a Morawetz-type commutator estimate. Let $Q(x, D_x)$ be a spatially self-adjoint operator. On time slices we obtain

$$(48) \quad \frac{d}{dt}\left\{-2\Re\langle D_t u, Qu\rangle + 2\Re\langle a_{(0)}^{j0} D_j u, Qu\rangle\right\} = -2\Im\langle P_{(0)} u, Qu\rangle + \langle i[D_i a_{(0)}^{ij} D_j, Q]u, u\rangle$$
$$+ 2\Re\langle i[a_{(0)}^{i0} D_i, Q]u, D_t u\rangle - 2\Re\langle(\partial_i a_{(0)}^{i0}) D_t u, Qu\rangle + 2\Re\langle(\partial_t a_{(0)}^{i0}) D_i u, Qu\rangle.$$

The point here is that we seek to choose Q so that the commutator $[D_i a_{(0)}^{ij} D_j, Q]$ is positive on the characteristic set of the operator $P_{(0)}$.

Finally, to account for the elliptic region, i.e. away from the characteristic set of the operator $P_{(0)}$, we use a Lagrangian term. Precisely, for a real-valued, time-independent, scalar function $\psi(x)$, we compute

$$\frac{d}{dt}\Im\langle(-D_t + 2a_{(0)}^{0j} D_j)u, \psi u\rangle = \Re\langle P_{(0)} u, \psi u\rangle - \Re\langle a_{(0)}^{ij} D_j u, \psi D_i u\rangle + \Im\langle a_{(0)}^{ij} D_j u, (\partial_i \psi)u\rangle$$
$$+ 2\Im\langle(\partial_t a_{(0)}^{0j}) D_j u, \psi u\rangle - 2\Im\langle(\partial_j a_{(0)}^{0j}) D_t u, \psi u\rangle$$
$$+ \langle D_t u, \psi D_t u\rangle - 2\Re\langle a_{(0)}^{0j} D_j u, \psi D_t u\rangle.$$

We consider two additional small parameters δ_0 and δ_1 so that

$$\epsilon \ll \delta_1 \ll \delta \ll \delta_0 \ll 1$$

and define the modified energy

$$E(u) = E_0(u) - \delta_0 \Re\langle(-D_t + a_{(0)}^{j0} D_j)u, Qu\rangle - \delta_1 \Re\langle(-D_t + 2a_{(0)}^{0j} D_j)u, i\psi u\rangle.$$

Combining the last three relations we obtain

$$(49) \quad \begin{aligned} \frac{d}{dt}E(u) \ &+ \frac{\delta_0}{2}\langle i[D_i a_{(0)}^{ij} D_j, Q]u, u\rangle + \delta_1\langle D_t u, \psi D_t u\rangle \lesssim \Im\langle P_{(0)} u, (D_t + \delta_0 Q + i\delta_1 \psi)u\rangle \\ &+ \langle|\nabla a_{(0)}|\nabla u, \nabla u\rangle + \delta_0|\langle i[a_{(0)}^{j0} D_j, Q]u, D_t u\rangle| + \delta_0\langle|\nabla a_{(0)}||\nabla u|, |Qu|\rangle \\ &+ \delta_1\langle|a||\nabla_x u|, \psi|\nabla u|\rangle + \delta_1\langle|a||\nabla_x u|, |\nabla\psi||u|\rangle + \delta_1\langle|\nabla a_{(0)}||\nabla u|, |\psi||u|\rangle. \end{aligned}$$

We choose Q as in [37]. For convenience its properties are summarized in the following

Lemma 12. *There exists an operator Q of the form*

$$Q(x, D_x) = \delta(Dx\phi(\delta|x|) + \phi(\delta|x|)xD)$$

where ϕ has the properties
  (i) $\phi(s) \approx (1+s)^{-1}$ for $s > 0$ and $|\partial^k \phi(s)| \lesssim (1+s)^{-k-1}$ for $k \le 4$,
  (ii) $\phi(s) + s\phi'(s) \approx (1+s)^{-1}\alpha(s)$ for $s > 0$,
  (iii) $\phi(|x|)$ is localized at frequency $\ll 1$,
and which satisfies the bounds

$$\|Qu\|_{L^2} \lesssim \|u\|_{L^2}$$

$$\|Qu\|_{X_{0,\alpha}} \lesssim \|u\|_{X_{0,\alpha}}$$

$$\int_{\mathbb{R}} \langle i[D_i a^{ij}_{(0)} D_j, Q]u, u \rangle \, dt \gtrsim \delta \|u\|^2_{X_{0,\alpha}}$$

for all functions u localized at frequency 1.

The function $\psi(|x|)$ is chosen so that

$$\psi(s) \approx \frac{\alpha(s)}{1+s}, \quad |\psi'(s)| \ll \psi(s).$$

We first note that the above properties of Q and ψ insure that E is positive definite; specifically

$$E(u) \approx \|\nabla u\|^2_{L^2}$$

for all functions u at frequency 1. Moreover, upon integration in t, we can estimate

$$\int_{\mathbb{R}} \langle D_t u, \psi D_t u \rangle \, dt \gtrsim \|D_t u\|^2_{X_{0,\alpha}},$$

and thus, the integral of the left side of (49) is bounded below by

$$\sup_{t \in \mathbb{R}} E(u)(t) - E(u)(0) + \delta \|\nabla_x u\|^2_{X_{0,\alpha}} + \delta_1 \|\partial_t u\|^2_{X_{0,\alpha}}.$$

We now examine the right side of (49) after integration in t. Using (32), we have

$$\int \langle |\nabla a_{(0)}| \nabla u, \nabla u \rangle + |\langle i[a^{j0}_{(0)} D_j, Q]u, D_t u \rangle| + \langle |\nabla a_{(0)}| |\nabla u|, |Qu| + |\psi||u| \rangle \, dt \lesssim \epsilon \|\nabla u\|^2_{X_{0,\alpha}}.$$

Similarly, by our choice of ψ, we may find a constant $M > 0$ so that

$$\int C\delta_1 \langle |a||\nabla_x u|, \psi |\nabla u| \rangle + C\delta_1 \langle |a||\nabla_x u|, |\nabla \psi|u \rangle \, dt \leq \frac{\delta_1}{2} \|\partial_t u\|^2_{X_{0,\alpha}} + M\delta_1 \|\nabla_x u\|^2_{X_{0,\alpha}}$$

where C is the implicit constant in (49).

Using these bounds to estimate the right side of (49) and using Cauchy-Schwarz, we obtain

$$\|\nabla u\|^2_{L^\infty L^2} + \delta\delta_0 \|\nabla_x u\|^2_{X_{0,\alpha}} + \delta_1 \|\partial_t u\|^2_{X_{0,\alpha}} \lesssim \|\nabla u(0)\|^2_{L^2} + \delta_1^{-1} \|P_{(0)} u\|^2_{L^1 L^2 + X'_{0,\alpha}}$$

provided, say, $\delta\delta_0 > 2M\delta_1$. This concludes the proof of Proposition 11. $\square$

We conclude now the proof of Theorem 4. Let (β_m) be another slowly varying sequence with

$$\sum_m \beta_m = 1.$$

Applying Proposition 11 with α_m replaced by $\alpha_m + \beta_m$ we obtain the bound

$$\|\nabla u\|_{L^\infty L^2 \cap X_{k,\alpha+\beta}} \lesssim \|\nabla u(0)\|_{L^2} + \|P_{(k)} u\|_{L^1 L^2 + X'_{k,\alpha+\beta}}$$

for all u localized at frequency 2^k. This implies the weaker estimate

$$\|\nabla u\|_{L^\infty L^2 \cap X_{k,\alpha}} \lesssim \|\nabla u(0)\|_{L^2} + \|P_{(k)} u\|_{L^1 L^2 + X'_{k,\beta}}.$$

Since any l^1 sequence is dominated by a slowly varying l^1 sequence, we can drop the assumption that α and β are slowly varying. Then we maximize the left hand side with respect to $\alpha \in l^1$ and minimize the right hand side with respect to $\beta \in l^1$. This yields

$$(50) \qquad \|\nabla u\|_{L^\infty L^2 \cap X_k} \lesssim \|\nabla u(0)\|_{L^2} + \|P_{(k)}u\|_{L^1 L^2 + X'_k}.$$

For an arbitrary function $u \in X^s$, we apply this bound to $S_k u$. We have

$$P_{(k)} S_k u = S_k \tilde{P} u + [P_{(k)}, S_k] u + S_k (P_{(k)} - \tilde{P}) u.$$

The last two terms are frequency localized and can be estimated by (33) and (34),

$$\|[P_{(k)}, S_k] u + S_k (P_{(k)} - \tilde{P}) u\|_{X'_k} \lesssim \epsilon \sum_{|k-l| \leq 2} \|\nabla S_l u\|_{X_k}.$$

Then after summation we obtain

$$\|\nabla u\|^2_{L^\infty \dot{H}^s \cap X^s} \lesssim \sum_k 2^{2sk} \|\nabla S_k u\|^2_{L^\infty L^2 \cap X_k}$$

$$\lesssim \sum_k \left[2^{2sk} \|S_k \nabla u(0)\|^2_{L^2} + 2^{2sk} \|P_{(k)} S_k u\|^2_{L^1 L^2 + X'_k} \right]$$

$$\lesssim \|\nabla u(0)\|^2_{\dot{H}^s} + \sum_k \left[2^{2sk} \|S_k \tilde{P} u\|^2_{L^1 L^2 + X'_k} \right.$$

$$\left. + 2^{2sk} \|[P_{(k)}, S_k] u + S_k (P_{(k)} - \tilde{P}) u\|^2_{X'_k} \right]$$

$$\lesssim \|\nabla u(0)\|^2_{\dot{H}^s} + \|\tilde{P} u\|^2_{L^1 \dot{H}^s + Y^s} + \epsilon \|\nabla u\|^2_{X^s}$$

$$\lesssim \|\nabla u(0)\|^2_{\dot{H}^s} + \|P_a u\|^2_{L^1 \dot{H}^s + Y^s} + \epsilon \|\nabla u\|^2_{X^s} \qquad \text{(by (35))}$$

$$\lesssim \|\nabla u(0)\|^2_{\dot{H}^s} + \|P u\|^2_{L^1 \dot{H}^s + Y^s} + \epsilon \|\nabla u\|^2_{X^s} \qquad \text{(by Lemma 3).}$$

For small ϵ we can neglect the last right hand side term to obtain

$$(51) \qquad \|\nabla u\|^2_{L^\infty \dot{H}^s \cap X^s} \lesssim \|\nabla u(0)\|^2_{\dot{H}^s} + \|P u\|^2_{L^1 \dot{H}^s + Y^s}$$

which holds in any time interval containing 0.

Reverting the transformation in Section 3, we see that without any restriction in generality we can write P in its self-adjoint divergence form. Assuming that $b = 0$, we may then use a duality argument to show that for any $f \in L^1 \dot{H}^s \cap Y^s$, there is a v solving

$$P v = f, \qquad v(0) = v_0, \qquad v_t(0) = v_1$$

with

$$\|\nabla v\|_{L^\infty \dot{H}^s \cap X^s} \lesssim \|\nabla v(0)\|_{\dot{H}^s} + \|f\|_{L^1 \dot{H}^s + Y^s}.$$

Due to (19) this extends perturbatively to the case of nonzero b.

By (51), this solution is unique, and the proof of Theorem 4 is concluded.

23

6. The half wave decomposition

In this section we reduce the study of the wave equation (7) to the study of two half-wave equations. We first factor the principal symbol as

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

where $a^\pm$ are 1-homogeneous in ξ satisfying the symmetry property

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

and are chosen so that $a^+ > a^-$. The symbols $a^\pm$ can be written down explicitly as

$$a^\pm(t,x,\xi) = -a^{0j}(t,x)\xi_j \pm \sqrt{(a^{0j}(t,x)\xi_j)^2 + a^{ij}(t,x)\xi_i\xi_j}.$$

In the sequel, we shall, however, only need the properties listed above. This will permit us, in Section 8 and beyond, to free up the a, b, c notation. There the focus will only be on the half-wave operators and the symbols $a^\pm$. The notations a, b, c will no longer be reserved for the coefficients of P but will be used for abstract terms which play the analogous roles.

Mollifying the symbols $a^\pm$ with respect to x as in Section 4 we obtain the symbols $a^\pm_{(k)}(t,x,\xi)$ which we use at frequency 2^k. We note that $a^\pm_{(k)}$ are **not** the symbols obtained from the factorization of the principal symbol of $P_{(k)}$; also one cannot define them in this way since algebraic operations (such as square roots) do not preserve the frequency localization.

We also denote

$$l(t,x,\xi) = (a^+(t,x,\xi) - a^-(t,x,\xi))^{-1}$$

and let $l_{(k)}(t,x,\xi)$ be the corresponding regularizations. We note that $l_{(k)}(t,x,\xi)$ is obtained by regularizing $l(t,x,\xi)$ and not by algebraically combining the symbols $a^\pm_{(k)}(t,x,\xi)$.

We are interested in operator properties matching the above algebraic properties. We work at frequency 1, but by rescaling the results extend to all dyadic frequencies.

Proposition 13. *Define the error operators*

(52) $$R^+ = P_{(0)} + (D_t + A^-_{(0)})(D_t + A^+_{(0)}), \qquad R^- = P_{(0)} + (D_t + A^+_{(0)})(D_t + A^-_{(0)}).$$

Then for all functions u and f localized at frequency 1, we have

(53) $$\|R^\pm u\|_{X'_0} \lesssim \|\nabla u\|_{X_0}$$

(54) $$\|\langle x\rangle(L_{(0)}(A^+_{(0)} - A^-_{(0)}) - I)f\|_{L^2_x} \lesssim \|f\|_{L^2_x}$$

(55) $$\|[L_{(0)}, P_{(0)}]u\|_{X'_0} \lesssim \|\nabla u\|_{X_0}.$$

24

Proof. We write R^+ in the form

$$R^+ = -i\partial_t A^+_{(0)}(t,x,D) - 2i(\partial_j a^{j0}_{(0)})D_t - i(\partial_j a^{jk}_{(0)})D_k$$
$$- (A^- A^+)_{(0)}(t,x,D) + A^-_{(0)}(t,x,D)A^+_{(0)}(t,x,D).$$

The first three terms are easily estimated by (32). Consider the remaining two terms. For $|\xi| \approx 1$ the symbols $a^\pm(t,x,\xi)$ are smooth and homogeneous in ξ. Expanding them into spherical harmonics we can assume without any restriction in generality that both $a^\pm$ have the form

$$a^\pm(t,x,\xi) = a^\pm(t,x)h^\pm(\xi)$$

where $a^\pm(t,x)$ satisfy bounds similar to the bounds for a^{ij}, namely

(56) $$|a^\pm(t,x) - a^\pm_\infty| + |x||\nabla a^\pm(t,x)| + |x|^2|\nabla^2 a^\pm(t,x)| \lesssim \epsilon(|x|).$$

Then the last two terms in R^+ have the form

$$a^-_{(0)}(t,x)h^-(D)a^+_{(0)}(t,x)h^+(D) - (a^- a^+)_{(0)}(t,x)h^-(D)h^+(D) =$$
$$(a^-_{(0)}(t,x)a^+_{(0)}(t,x) - (a^- a^+)_{(0)}(t,x))h^-(D)h^+(D) + a^-_{(0)}(t,x)[h^-(D), a^+_{(0)}(t,x)]h^+(D).$$

The operators $h^\pm$ are bounded in X_0 on functions of frequency 1. The commutator estimate

(57) $$[h^-(D), a^-_{(0)}(t,x)] : X_0 \to X'_0$$

on frequency 1 functions follows due to the bound

$$|\nabla a^-_{(0)}(t,x)| \lesssim \epsilon_0(|x|)\langle x\rangle^{-1}.$$

Hence the estimate (53) is proved if we can show that

Lemma 14. *Let $a^\pm$ be functions satisfying* (56). *Then*

(58) $$|a^-_{(0)}(t,x)a^+_{(0)}(t,x) - (a^- a^+)_{(0)}(t,x)| \lesssim \epsilon_0(|x|)\langle x\rangle^{-1}.$$

Proof. Without any restriction in generality we can assume that $a^\pm_\infty = 0$. As in the case of the coefficients a^{ij}, the regularized functions have size

$$|a^\pm_{(0)}(t,x)| \lesssim \epsilon_0(|x|).$$

We separate the contributions coming from small x and from large x. The contribution from small x decays rapidly at infinity,

$$|(\chi_{\leq 0}a^\pm)_{(0)}(t,x)| \lesssim \epsilon_0(|x|)\langle x\rangle^{-N},$$

and the corresponding part in (58) will satisfy a similar bound. Hence without any restriction in generality we assume that $a^\pm$ are both supported in $A_{\geq 0}$. This allows us to replace (56)

with a better bound
$$|a^\pm(t,x)| + \langle x\rangle |\nabla a^\pm(t,x)| + \langle x\rangle^2 |\nabla^2 a^\pm(t,x)| \lesssim \epsilon_0(|x|).$$

Using the analogues of (31), this allows us to estimate the differences
$$|a_{(0)}^\pm(t,x) - a^\pm(t,x)| \lesssim \epsilon_0(|x|)\langle x\rangle^{-1}$$

and similarly for their product a^+a^-. The conclusion of the lemma follows. $\qquad \square$

The proof of (54) is virtually identical, the roles of $a^\pm(t,x,\xi)$ are played by $l(t,x,\xi)$ and $a^+(t,x,\xi) - a^-(t,x,\xi)$.

For (55) we expand $l(t,x,\xi)$ in spherical harmonics and reduce the problem to the case when
$$l(t,x,\xi) = l(t,x)h(\xi)$$

with $l(t,x)$ satisfying (56). Then the proof of (55) reduces to commutator estimates similar to (57). $\qquad \square$

7. PARAMETRICES AND STRICHARTZ ESTIMATES

Here we reduce the proof of Theorem 6 to the construction of a suitable parametrix for $D_t + A_{(0)}^\pm$. Our main result concerning parametrices is

Proposition 15. *Assume that ϵ is sufficiently small. Then there are parametrices $K_0^\pm$ for $D_t + A_{(0)}^\pm$ which are localized at frequency 1 and have the following properties:*
  (i) L^2 bound:

$$\tag{59} \|K_0^\pm(t,s)\|_{L_x^2 \to L_x^2} \lesssim 1,$$

  (ii) Error estimate:

$$\tag{60} \begin{aligned} \|(1+|x|)^N (D_t + A_{(0)}^\pm)K_0^\pm(t,s)\|_{L_x^2 \to L_x^2} &\lesssim (1+|t-s|)^{-N}, & t \neq s, \\ \|(1+|x|)^N D_t(D_t + A_{(0)}^\pm)K_0^\pm(t,s)\|_{L_x^2 \to L_x^2} &\lesssim (1+|t-s|)^{-N}, & t \neq s, \end{aligned}$$

  (iii) Jump condition: $K_0^\pm(s+0,s)$ and $K_0^\pm(s-0,s)$ are $S_{1,0}^0$ type pseudodifferential operators satisfying

$$(K_0^\pm(s+0,s) - K_0^\pm(s-0,s))S_0 = S_0,$$

  (iv) Outgoing parametrix:

$$\tag{61} \|1_{\{|x|<2^{-10}|t-s|\}}K_0^\pm(t,s)\|_{L_x^2 \to L_x^2} \lesssim (1+|t-s|)^{-N},$$

  (v) Pointwise decay:

$$\tag{62} \|K_0^\pm(t,s)\|_{L_x^1 \to L_x^\infty} \lesssim (1+|t-s|)^{-\frac{n-1}{2}}.$$

Here $K_0^\pm$ is defined by

$$K_0^\pm f(t) = \int_{-\infty}^\infty K_0^\pm(t,s)f(s)ds.$$

We leave the proof of this result for later sections, and we show that it implies Theorems 6,7. As an intermediate step we have the following localized Strichartz estimates for the parametrix:

Proposition 16. *The parametrix $K_0^\pm$ given by Proposition 15 has the following properties:*

(i) (regularity) For any Strichartz pairs (p_1, q_1) respectively (p_2, q_2) with $q_1 \le q_2$ we have

(63) $$\|K_0^\pm f\|_{L^{p_1}L^{q_1}\cap X_0} \lesssim \|f\|_{L^{p_2'}L^{q_2'}}.$$

(ii) (error estimate) For any Strichartz pair (p,q) we have

(64) $$\|[(D_t + A_{(0)}^\pm)K_0^\pm - 1]f\|_{X_0'} \lesssim \|f\|_{L^{p'}L^{q'}}.$$

In both (63) *and* (64) *the function f is assumed to be localized at frequency 1.*

The proof is identical to the proof of the similar result in [37, Proposition 12] and is omitted. The proof of (63) follows that of the Strichartz estimates in the constant coefficient case as it consists of interpolating between (59) and (62), using a TT^* argument, and applying the Hardy-Littlewood-Sobolev inequality. The error estimate (64) follows somewhat directly from (60).

We can use the half-wave parametrices to construct a full wave parametrix. Precisely we have

Proposition 17. *Assume that ϵ is sufficiently small. Then there is a parametrix K_0 for $P_{(0)}$ which has the following properties:*

(i) (regularity) For any Strichartz pairs (p_1, q_1) respectively (p_2, q_2) with $q_1 \le q_2$, we have

(65) $$\|\nabla K_0 f\|_{L^{p_1}L^{q_1}\cap X_0} \lesssim \|f\|_{L^{p_2'}L^{q_2'}}.$$

(ii) (error estimate) For any Strichartz pair (p,q) we have

(66) $$\|(P_{(0)}K_0 - 1)f\|_{X_0'} \lesssim \|f\|_{L^{p'}L^{q'}}.$$

In all of the above the function f is assumed to be localized at frequency 1.

Proof. Our first approximation for K_0 is the operator K_{00} defined by

$$K_{00} = L_{(0)}(K_0^+ - K_0^-).$$

The operator $L_{(0)}$ is bounded in both $L^{p_1}L^{q_1}$ and X_0; therefore from (63) we obtain part of (65), namely

$$\|D_x K_{00} f\|_{L^{p_1}L^{q_1} \cap X_0} \lesssim \|f\|_{L^{p_2'}L^{q_2'}}.$$

We can also bound $D_t K_{00} f$ in X_0. We have

$$D_t K_{00} f = [D_t, L_{(0)}](K_0^+ - K_0^-) + L_{(0)} D_t(K_0^+ - K_0^-),$$

and the first commutator is bounded in X_0. For the second term we use the X_0 bound for $L_{(0)}$ and write

$$D_t(K_0^+ - K_0^-)f = (D_t + A_{(0)}^+)K_0^+ f - (D_t + A_{(0)}^-)K_0^- f - A_{(0)}^+ K_0^+ f + A_{(0)}^- K_0^- f.$$

Now we use (64) and the embedding $X_0' \subset X_0$ for the first two terms and the X_0 boundedness of $A_{(0)}^+$ and $A_{(0)}^-$. Summing up we have proved that

$$\|D_t K_{00} f\|_{X_0} \lesssim \|f\|_{L^{p_2'}L^{q_2'}}.$$

We still have to estimate $D_t K_{00} f$ in $L^{p_1}L^{q_1}$, but we postpone this for later.

Next we estimate the error

$$P_{(0)} K_{00} - 1.$$

The kernel $K_{00}(s,t)$ of K_{00} is smooth in s,t away from the diagonal. However, we need to compute its singularity on the diagonal. Due to the property (iii) in Proposition 15 we see that the jump of K_{00} on the diagonal vanishes, namely

$$[K_{00}(t,t)] := K_{00}(t+0,t) - K_{00}(t-0,t) = 0.$$

However, the jump of the t derivative of $K_{00}(t,s)$ on the diagonal is nontrivial. Precisely, we have

$$\begin{aligned}
[D_t K_{00}(t,t)] &= L_{(0)}[D_t(K_0^+ - K_0^-)(t,t)] \\
&= L_{(0)}(-A_{(0)}^+[K_0^+(t,t)] + A_{(0)}^-[K_0^-(t,t)]) \\
&\quad + L_{(0)}([(D_t + A_{(0)}^+)K_0^+(t,t)] - [(D_t + A_{(0)}^-)K_0^-(t,t)]) \\
&= L_{(0)}(A_{(0)}^- - A_{(0)}^+) + L_{(0)}([(D_t + A_{(0)}^+)K_0^+(t,t)] - [(D_t + A_{(0)}^-)K_0^-(t,t)]).
\end{aligned}$$

By (54) the first term is close to the identity, while the second can be estimated by (60). For f localized at frequency 1 we obtain

(67) $$\|(1 + |x|)([D_t K_{00}(t,t)] - I)f\|_{L^{p_2'}L^2} \lesssim \|f\|_{L^{p_2'}L^2} \lesssim \|f\|_{L^{p_2'}L^{q_2'}}.$$

Next we compute

$$(P_{(0)} K_{00} - 1)f = R_0 f + ([D_t K_{00}(t,t)] - I)f$$

28

where the first term represents the off-diagonal contribution and the last term represents the contribution due to the jump of $D_t K_{00}(t, s)$ on the diagonal.

We use the factorization (52) for $P_{(0)}$ to compute the kernel of R_0,

$$
\begin{aligned}
R_0(t, s) &= L_{(0)} P_{(0)} (K_0^+ - K_0^-)(t, s) - [L_{(0)}, P_{(0)}](K_0^+ - K_0^-)(t, s) \\
&= - L_{(0)} \left((D_t + A_{(0)}^-)(D_t + A_{(0)}^+) K_0^+(t, s) - (D_t + A_{(0)}^+)(D_t + A_{(0)}^-) K_0^-(t, s) \right) \\
&\quad + L_{(0)} \left(R^+ K_0^+(t, s) - R^- K_0^-(t, s) \right) - [L_{(0)}, P_{(0)}](K_0^+ - K_0^-)(t, s).
\end{aligned}
$$

For the expression on the first line we use the X_0' boundedness of $L_{(0)}$ and $A_{(0)}^\pm$, together with the error estimates in (60). For the $R^\pm$ terms we use (53) together with (63) and (60); the latter is needed to bound the time derivative $D_t K(t, s)$. Finally, for the last terms we use (55). Summing up, we obtain

(68) $$\|R_0 f\|_{X_0'} \lesssim \|f\|_{L^{p'} L^{q'}}.$$

This is an acceptable error.

The expression

$$
f_1 = -([D_t K_{00}(t, t)] - I)f,
$$

however, is not an acceptable error because it does not yield to a similar bound of its X_0' norm. It has better decay at infinity; therefore we can account for it by setting

$$
K_0 f = K_{00} f + K_{01} f_1
$$

where

$$
K_{01} f_1(t) = -\frac{1}{2} \int_{-\infty}^{\infty} e^{-|t-s|} f_1(s) ds,
$$

which solves

$$
\partial_t^2 K_{01} f_1 = f_1 + K_{01} f_1.
$$

Using the bound (67) for f_1, it is easy to see that $K_{01} f_1$ satisfies

$$
\|\nabla K_{01} f_1\|_{L^{p_1} L^{q_1} \cap X_0} \lesssim \|f\|_{L^{p_2'} L^{q_2'}}.
$$

On the other hand the f_1 component of the error is replaced by

$$
f_2 = P_{(0)} K_{01} f_1 - f_1 = (2 D_i a_{(0)}^{i0} D_t + D_i a_{(0)}^{ij} D_j + I) K_{01} f_1,
$$

which we can estimate by

$$
\|f_2\|_{X_0'} \lesssim \|D_t K_{01} f_1\|_{X_0'} + \|K_{01} f_1\|_{X_0'} \lesssim \|\langle x \rangle f_1\|_{L^{p_2'} L^2}
$$

and then apply (67).

The last step of the argument is to prove the $L^{p_1}L^{q_1}$ bound for $D_t K_0 f$. We will show that for u at frequency 1 we have

$$\|D_t u\|_{L^{p_1}L^{q_1}} \lesssim \|u\|_{L^{p_1}L^{q_1}} + \|P_{(0)}u\|_{L^{p'_2}L^{q'_2}+X'_0}$$

from which the desired bound follows after an application of (66). This would follow from

$$\|D_t u\|_{L^{p_1}L^{q_1}} \lesssim \|u\|_{L^{p_1}L^{q_1}} + \|P_{(0)}u\|_{(L^1+L^{p_1})L^{q_1}}$$

or equivalently,

$$\|D_t u\|_{L^{p_1}L^{q_1}} \lesssim \|u\|_{L^{p_1}L^{q_1}} + \|g_1\|_{L^1 L^{q_1}} + \|g_2\|_{L^{p_1}L^{q_1}}, \qquad P_{(0)}u = g_1 + g_2.$$

The above is easily reduced to the case $g_1 = 0$ by substituting

$$u := u - v, \qquad v = -\frac{1}{2}\int_{-\infty}^{\infty} e^{-|t-s|}g_1(s)ds$$

since

$$\|v\|_{(L^1 \cap L^\infty)L^{q_1}} + \|D_t v\|_{(L^1 \cap L^\infty)L^{q_1}} \lesssim \|g_1\|_{L^1 L^{q_1}}.$$

We are left with proving

$$\|D_t u\|_{L^{p_1}L^{q_1}} \lesssim \|u\|_{L^{p_1}L^{q_1}} + \|P_{(0)}u\|_{L^{p_1}L^{q_1}}$$

which follows from the interpolation inequality

$$\|D_t u\|^2_{L^{p_1}L^{q_1}} \lesssim \|u\|_{L^{p_1}L^{q_1}} \|D_t^2 u\|_{L^{p_1}L^{q_1}}.$$

$\square$

Proposition 17 is useful only if ϵ is small. However, a similar result holds even if ϵ is not small:

Proposition 18. *Assume that the coefficients $a^{i\beta}$ satisfy (8). Then there is a parametrix K_0 for $P_{(0)}$ localized at frequency 1 and which satisfies*

(i) (regularity) For any Strichartz pairs (p_1, q_1) respectively (p_2, q_2) with $q_1 \leq q_2$, we have

(69)
$$\|\nabla K_0 f\|_{L^{p_1}L^{q_1} \cap X_0} \lesssim \|f\|_{L^{p'_2}L^{q'_2}}.$$

(ii) (error estimate) For any Strichartz pair (p, q), we have

(70)
$$\|[P_{(0)}K_0 - 1]f\|_{X'_0} \lesssim \|f\|_{L^{p'}L^{q'}}.$$

The proof is identical to the proof of the similar result in [37, Proposition 15] and is omitted. The idea is that the smallness condition is violated only on finitely many dyadic spatial regions. In [37] it is argued that a fixed dyadic spatial region can be partitioned into finitely many cubes on which the smallness holds with respect to a different coordinate frame. The local parametrices are then assembled together using a partition of unity. Alternatively, in a fixed dyadic region the problem of constructing a parametrix as above can be localized to a similar time scale and then rescaled into a local problem.

Proof of Theorems 6, 7. In what follows we work in a time interval $[T^-, T^+]$, possibly infinite. By (33) we can replace the operator $P_{(0)}$ by $\tilde{P}$ in Propositions 17, 18. Rescaling this result we obtain similar parametrices K_j at any dyadic frequency 2^j. We first assemble these dyadic parametrices and set

$$K = \sum_{j=-\infty}^{\infty} K_j S_j.$$

The properties of K are summarized in the next lemma.

Lemma 19. *The parametrix K for P_a has the following properties:*

(i) (regularity) For any Strichartz pairs (ρ_1, p_1, q_1) respectively (ρ_2, p_2, q_2) with $q_1 \leq q_2$, we have

(71) $$\|\nabla K f\|_{|D_x|^{\rho_1-s}L^{p_1}L^{q_1}\cap X^s} \lesssim \|f\|_{|D_x|^{-\rho_2-s}L^{p_2'}L^{q_2'}}.$$

(ii) (error estimate) For any Strichartz pair (ρ, p, q), we have

(72) $$\|(P_a K - I)f\|_{Y^s} \lesssim \|f\|_{|D|^{-\rho-s}L^{p'}L^{q'}}.$$

Part (i) follows directly from the Littlewood-Paley theory[3]. Similarly we get part (ii) but with $\tilde{P}$ instead of P_a, since we can write

$$\tilde{P}K - I = \sum_{j\in\mathbb{Z}} (\tilde{P} - P_{(j)})K_j S_j + (P_{(j)}K_j - I)S_j$$

However, by (35) we can freely interchange P_a and $\tilde{P}$. Since (19) allows us to further pass from P_a to P with $c = 0$, this establishes the bounds for the first and last terms in the left side of (25).

A second step is to use duality to establish an $L^2 \to L^p L^q$ bound. This establishes (24), i.e. the first part of Theorem 7 (a).

---

[3] As mentioned in the Introduction, the Littlewood-Paley theory with respect to the spatial variables cannot be used in dimension $n = 2$ for the $L^4 L^\infty$, respectively the $L^{4/3}L^1$ norms. Here we instead only obtain estimates in appropriate l^2 Besov spaces.

Lemma 20. *If there is a parametrix K for P_a as in Lemma 19 and (ρ, p, q) is a Strichartz pair, then*

(73)
$$\|\nabla u\|_{|D_x|^{\rho-s} L^p L^q} \lesssim \|\nabla u\|_{L^\infty \dot{H}^s \cap X^s} + \|P_a u\|_{Y^s}.$$

Proof. Without any restriction in generality we assume that T^- and T^+ are finite but prove the bound with constants which are independent of T^+ and T^-. For $g^\alpha \in |D_x|^{s-\rho} L^{p'} L^{q'}$ we use integration by parts

$$\int_{T^-}^{T^+} \langle \nabla u, g \rangle \, dt = \int_{T^-}^{T^+} \langle \nabla u, P_a K g \rangle \, dt - \int_{T^-}^{T^+} \langle \nabla u, [P_a K - 1]g \rangle \, dt$$

$$= \int_{T^-}^{T^+} \Big[-\langle P_a u, \nabla \cdot K g \rangle - \langle \nabla u, [P_a K - 1]g \rangle - 2\langle \partial_i u, (\nabla a^{i0}) \cdot \partial_t K g \rangle$$

$$- \langle \partial_j u, (\nabla a^{ij}) \cdot \partial_i K g \rangle + 2\langle \partial_i u, (\partial_t a^{i0}) \nabla \cdot K g \rangle - 2\langle \partial_t u, (\partial_i a^{i0}) \nabla \cdot K g \rangle \Big] dt$$

$$+ \langle \nabla u, \partial_t K g \rangle|_{T^-}^{T^+} + \langle \partial_t u, \nabla \cdot K g \rangle|_{T^-}^{T^+} - 2\langle a^{i0} \partial_i u, \nabla \cdot K g \rangle|_{T^-}^{T^+}$$

$$- \langle \partial_t u, \partial_t K g^0 \rangle|_{T^-}^{T^+} + 2\langle a^{i0} \partial_i u, \partial_t K g^0 \rangle|_{T^-}^{T^+} + \langle a^{ij} \partial_j u, \partial_i K g^0 \rangle|_{T^-}^{T^+}.$$

Then by (71) and (72) we obtain

$$\left| \int_{T^-}^{T^+} \langle \nabla u, g \rangle \, dt \right| \lesssim \|g\|_{|D|^{s-\rho} L^{p'} L^{q'}} \left(\|\nabla u\|_{L^\infty \dot{H}^s \cap X^s} + \|P_a u\|_{Y^s} \right).$$

Here we have also used (19) with b replaced by ∇a, which according to (8) satisfies (9). The conclusion follows. □

Next we prove that the conclusion of Lemma 19 is also valid for $q_1 > q_2$:

Lemma 21. *The parametrix K in Lemma 19 also satisfies (71) when $q_1 > q_2$.*

Proof. We repeat the computation in the previous lemma with

$$u = Kf, \qquad g^\alpha \in |D|^{s-\rho_1} L^{p_1'} L^{q_1'}.$$

All the terms are estimated in the same way except for

$$\int_{T^-}^{T^+} \langle P_a u, \nabla \cdot K g \rangle \, dt = \int_{T^-}^{T^+} \langle (P_a K - I) f, \nabla \cdot K g \rangle \, dt + \int_{T^-}^{T^+} \langle f, \nabla \cdot K g \rangle \, dt$$

for which we use (71) and (72) to estimate

$$\left| \int_{T^-}^{T^+} \langle P_a u, \nabla \cdot K g \rangle \, dt \right| \lesssim \|(P_a K - I) f\|_{Y^s} \|\nabla K g\|_{X^{-s}} + \|f\|_{|D|^{-s-\rho_2} L_2^{p'} L_2^{q'}} \|\nabla K g\|_{|D|^{s+\rho_2} L^{p_2} L^{q_2}}$$

$$\lesssim \|f\|_{|D|^{-s-\rho_2} L_2^{p'} L_2^{q'}} \|g\|_{|D|^{s-\rho_1} L_1^{p'} L_1^{q'}}.$$

32

Then as in the previous lemma we obtain

$$\left| \int_{T^-}^{T^+} \langle \nabla u, g \rangle \, dt \right| \lesssim \|g\|_{|D|^{s-\rho_1} L_1^{p'} L^{q_1'}} \left(\|\nabla u\|_{L^\infty \dot{H}^s \cap X^s} + \|f\|_{|D|^{-s-\rho_2} L_2^{p'} L^{q_2'}} \right)$$

$$\lesssim \|g\|_{|D|^{s-\rho_1} L_1^{p'} L^{q_1'}} \|f\|_{|D|^{-s-\rho_2} L_2^{p'} L^{q_2'}}$$

which concludes the proof. $\qquad\square$

The bound (71) on Kf allows us to estimate $\|Kf\|_{X^{s+1}}$. However, if $s + 1 \geq \frac{n-1}{2}$ then in order to conclude the proof of (25), i.e. the remainder of Theorem 7 (a), we need to have a bound for the stronger norm $\|Kf\|_{\tilde{X}^{s+1}}$. This is achieved in the next lemma.

Lemma 22. *There is a parametrix $\tilde{K}$ for P_a which satisfies:*

(i) (regularity) For any[4] Strichartz pairs (ρ_1, p_1, q_1) respectively (ρ_2, p_2, q_2), we have

$$(74) \qquad \|\nabla \tilde{K} f\|_{|D_x|^{\rho_1 - s} L^{p_1} L^{q_1} \cap X^s} + \|\tilde{K} f\|_{\tilde{X}^{s+1}} \lesssim \|f\|_{|D_x|^{-\rho_2 - s} L_2^{p'} L^{q_2'}}.$$

(ii) (error estimate) For any Strichartz pair (ρ, p, q), we have

$$(75) \qquad \|(P_a \tilde{K} - I) f\|_{Y^s} \lesssim \|f\|_{|D|^{-\rho - s} L^{p'} L^{q'}}.$$

Proof. Let K be as in Lemma 19. If we think of Kf as the sum of its dyadic pieces which are measured in X_k, then for $s + 1 \geq \frac{n-1}{2}$ we fail to obtain a $\tilde{X}^{s+1}$ bound for Kf due to the accumulation near the origin of the contributions below the uncertainty principle scale $\{|x| \lesssim |\xi|^{-1}\}$. To remedy this we attempt to remove these contributions.

We consider a Schwartz function ϕ with

$$\phi(0) = 1, \qquad \operatorname{supp} \hat{\phi} \subset \{|\xi| \in [1/2, 2]\}$$

and set $\phi_k(x) = \phi(2^k x)$. In a first approximation we replace the parametrix K with $(1-T)K$, with T defined by

$$Tu = \sum_{k=-\infty}^{\infty} T_k S_k u, \qquad T_k u = u(t, 0) \phi_k.$$

This substitution improves the left hand side of (74). We shall show that

$$(76) \qquad \|(1 - T)u\|_{\tilde{X}^{s+1}} \lesssim \|u\|_{X^{s+1}}, \qquad \frac{n-1}{2} \leq s + 1 < \frac{n+1}{2}$$

$$(77) \qquad \|\nabla T K f\|_{|D_x|^{\rho_1 - s} L^{p_1} L^{q_1} \cap X^s} \lesssim \|f\|_{|D_x|^{\rho_2 - s} L_2^{p'} L^{q_2'}}.$$

For functions u localized at frequency 2^k, we have the fixed time pointwise bound

$$|u(t, 0)| \lesssim 2^{\frac{n-1}{2}k} \|u\|_{X_k^0}$$

---

[4] We are again largely ignoring the $L^4 L^\infty$, respectively $L^{4/3} L^1$, estimates in $n = 2$.

which implies that

$$\|T_k u\|_{X_k} \lesssim \|u\|_{X_k}.$$

Here, X_k^0 is the spatial part of the X_k norm, i.e., $X_k = L^2 X_k^0$. Hence we easily obtain

$$\|Tu\|_{X^{s+1}} \lesssim \|u\|_{X^{s+1}}, \qquad \|\nabla Tu\|_{X^s} \lesssim \|\nabla u\|_{X^s}.$$

The X^s bound of (76) follows immediately, and in order to obtain the X^s bound of (77), we then apply (71). The $L^{p_1} L^{q_1}$ estimate uses a similar argument involving a Bernstein estimate, Littlewood-Paley estimates, and the bound (71).

To prove the L^2 part of the bound (76), we take advantage of the fact that $((1 - T_k)S_k u)(t, 0) = 0$ to obtain the better bound

$$\sup_j \| |x|^{-1-\frac{n}{2}} (|x| + 2^{-k})^{\frac{1}{2}+\frac{n}{2}} (1 - T_k)S_k u \|_{L^2(A_j)} \lesssim \|S_k u\|_{X_k},$$

which after summation yields

$$\| |x|^{-s-\frac{3}{2}} (1 - T)u \|_{L^2} \lesssim \|u\|_{X^{s+1}}, \qquad \frac{n-1}{2} \leq s + 1 < \frac{n+1}{2}.$$

Consider now the error estimate for $(1 - T)K$. We claim that

(78)
$$\|\tilde{P}TKf - Tf\|_{Y^s} \lesssim \|(\tilde{P}K - 1)f\|_{Y^s} + \|\nabla Kf\|_{X^s}.$$

It is easily seen that T_k is bounded in X'_k; therefore T is bounded in Y^s. It remains to show that

$$\|\tilde{P}Tu - T\tilde{P}u\|_{Y^s} \lesssim \|\nabla u\|_{X^s}$$

which reduces to

$$\|P_{(k)}T_k S_k u - T_k P_{(k)} S_k u\|_{X'_k} \lesssim \|\nabla S_k u\|_{X_k}.$$

After rescaling to $k = 0$ this is straightforward. What is important is that the second order time derivatives cancel. All the remaining terms can be estimated separately.

It remains to consider separately the outstanding error estimate for TKf. This cannot be placed in Y^s because it does not have enough time integrability. Hence we need to add a correction to the parametrix $(1 - T)K$ which accounts for this. Our final parametrix $\tilde{K}$ has the form

$$\tilde{K} = (1 - T)K + R_T f$$

where the operator R_T verifies the following properties:

(79)
$$\|\nabla R_T f\|_{|D_x|^{\rho_1 - s} L^{p_1} L^{q_1} \cap X^s} + \|R_T f\|_{\tilde{X}^{s+1}} \lesssim \|f\|_{|D_x|^{-\rho_2 - s} L^{p'_2} L^{q'_2}}$$

and

(80)
$$\|(\tilde{P}R_T - T)f\|_{Y^s} \lesssim \|f\|_{|D_x|^{-\rho - s} L^{p'} L^{q'}}.$$

34

For Tf we have the representation

$$Tf = \sum_{k \in \mathbb{Z}} \phi_k(x) f_k(t), \qquad f_k(t) = S_k f(t, 0).$$

Then we define

$$R_T f = \sum_{k \in \mathbb{Z}} \sum_{j \geq k} (\phi_{j+1}(x) - \phi_j(x)) D_t^{-2} S_{>j}^t f_k(t).$$

Here S_j^t is a Littlewood-Paley decomposition in the time-frequency variable. That is,

$$1 = \sum_{j=-\infty}^{\infty} S_j^t(D_t)$$

with

$$\operatorname{supp} s_j^t \subset \{2^{j-1} < |\tau| < 2^{j+1}\}.$$

The notions of $S_{>j}^t$, $S_{\leq j}^t$, etc. are then analogous to those defined in Section 2. D_t^{-2} denotes the operator with Fourier multiplier τ^{-2}, where τ is the frequency variable dual to t.

For f_k we estimate

$$\|f_k\|_{L^{p_2'}} \lesssim 2^{\frac{nk}{q_2'}} \|S_k f\|_{L^{p_2'} L^{q_2'}} \lesssim 2^{(-s-\rho_2+\frac{n}{q_2'})k} \|S_k f\|_{|D_x|^{-\rho_2-s} L^{p_2'} L^{q_2'}}.$$

Since

$$-\rho_2 + \frac{n}{q_2'} = \frac{n}{2} + \frac{1}{p_2},$$

we obtain

(81)
$$\|f_k\|_{L^{p_2'}} \lesssim 2^{(-s+\frac{n}{2}+\frac{1}{p_2})k} \|S_k f\|_{|D_x|^{-\rho_2-s} L^{p_2'} L^{q_2'}}.$$

We now proceed to estimate $R_T f$. By Bernstein's inequality in time we have

$$\|\nabla (\phi_{j+1}(x) - \phi_j(x)) D_t^{-2} S_{>j}^t f_k(t)\|_{|D_x|^{\rho_1-s} L^{p_1} L^{q_1}} \lesssim 2^{-\frac{n}{q_1} j} 2^{-j} 2^{(s-\rho_1)j} 2^{(\frac{1}{p_2'}-\frac{1}{p_1})j} \|f_k\|_{L^{p_2'}}$$

$$= 2^{-(-s+\frac{n}{2}+\frac{1}{p_2})j} \|f_k\|_{L^{p_2'}}$$

$$\lesssim 2^{(-s+\frac{n}{2}+\frac{1}{p_2})(k-j)} \|S_k f\|_{|D_x|^{-\rho_2-s} L^{p_2'} L^{q_2'}}.$$

Since

$$s < \frac{n-1}{2} \leq \frac{n}{2} + \frac{1}{p_2},$$

it follows that we have off-diagonal decay, while the diagonal summation is controlled by the Littlewood-Paley theory. This works if $q_1 \neq \infty$. In the special case $q_1 = \infty$ we also need to observe that the bump functions $\phi_{j+1}(x) - \phi_j(x)$ concentrate in different spatial regions; therefore cannot produce pointwise accumulation.

35

We continue with the X^s norm:

$$\|\nabla(\phi_{j+1}(x) - \phi_j(x))D_t^{-2}S_{>j}^t f_k(t)\|_{X^s} \lesssim 2^{-\frac{n}{2}j}2^{(s-\frac{1}{2})j}2^{(\frac{1}{p_2'}-\frac{1}{2})j}\|f_k\|_{L^{p_2'}}$$

$$= 2^{-(-s+\frac{n}{2}+\frac{1}{p_2})j}\|f_k\|_{L^{p_2'}}$$

$$\lesssim 2^{(-s+\frac{n}{2}+\frac{1}{p_2})(k-j)}\|S_k f\|_{|D_x|^{-\rho_2-s}L^{p_2'}L^{q_2'}},$$

and the summation works out as before.

The $\tilde{X}^{s+1}$ norm is next. Taking advantage of the fact that $(\phi_{j+1} - \phi_j)(0) = 0$ we compute

$$\||x|^{-s-\frac{3}{2}}(\phi_{j+1}(x) - \phi_j(x))D_t^{-2}S_{>j}^t f_k(t)\|_{L^2} \lesssim 2^{-\frac{n}{2}j}2^{(s-\frac{1}{2})j}2^{(\frac{1}{p_2'}-\frac{1}{2})j}\|f_k\|_{L^{p_2'}}$$

$$\lesssim 2^{(-s+\frac{n}{2}+\frac{1}{p_2})(k-j)}\|S_k f\|_{|D_x|^{-\rho_2-s}L^{p_2'}L^{q_2'}}$$

where the restriction $s < \frac{n-1}{2}$ insures that the norm on the left is finite. We still have off-diagonal decay, and for the diagonal summation we can use spatial orthogonality. This concludes the proof of (79).

For the error estimate (80) we split

$$\tilde{P} = -D_t^2 + \tilde{P}_1.$$

The expression $\tilde{P}_1 R_T f$ is bounded in the same manner as above. On the other hand we have

$$-D_t^2 R_T f - Tf = \sum_k \sum_{j \geq k}(\phi_{j+1}(x) - \phi_j(x))S_{\leq j}^t f_k(t),$$

and for the summand on the right we can use again Bernstein's inequality with respect to t. $\qquad\square$

Now we prove (23). If

$$P_a u = f + g, \qquad f \in |D_x|^{-\rho_2-s}L^{p_2'}L^{q_2'}, \ g \in Y^s,$$

then we write

$$u = Kf + v.$$

We use (71) to bound ∇Kf in $|D_x|^{\rho_1-s}L^{p_1}L^{q_1} \cap X^s$. It remains to bound v, which solves

$$P_a v = (1 - P_a K)f + g.$$

In the case of Theorem 6 we use successively (73), Theorem 4, (71), and (72). We obtain

$$\|\nabla v\|_{|D_x|^{\rho_1-s}L^{p_1}L^{q_1}} \lesssim \|\nabla v\|_{L^\infty \dot{H}^s \cap X^s} + \|P_a v\|_{Y^s}$$

$$\lesssim \|\nabla v(0)\|_{\dot{H}^s} + \|P_a v\|_{Y^s}$$

$$\lesssim \|\nabla u(0)\|_{\dot{H}^s} + \|\nabla Kf\|_{L^\infty \dot{H}^s} + \|(1 - P_a K)f\|_{Y^s} + \|g\|_{Y^s}$$

$$\lesssim \|\nabla u(0)\|_{\dot{H}^s} + \|f\|_{|D_x|^{-\rho_2-s}L^{p_2'}L^{q_2'}} + \|g\|_{Y^s}.$$

36

This establishes (23) with P replaced by P_a. Using (19), (23) then follows.

In the case of Theorem 7 the argument is similar, but instead of using Theorem 4 we assume that the localized energy estimates hold. $\qquad\square$

8. Pseudodifferential operators and phase space transforms

Here we tersely introduce the microlocal setup which will be required in the sequel. A more detailed exposition can be found in [36], [37], and the references therein.

Precisely, our initial goal is to provide a phase-space description of the flow for a pseudo-differential evolution of the form

$$(82) \qquad (D_t + a^w(t, x, D))u = 0, \quad u(0) = u_0$$

with a real symbol a. We begin by introducing a simpler set-up, which suffices in order to obtain a short time description of the flow. In terms of symbol classes, we begin with the standard class S_{00}^0 of symbols a satisfying

$$|\partial_x^\alpha \partial_\xi^\beta a(x, \xi)| \le c_{\alpha\beta}, \qquad |\alpha| + |\beta| \ge 0.$$

We also need the following generalizations $S^{(k)} = S_{00}^{0,(k)}$ of the above class, defined by

$$|\partial_x^\alpha \partial_\xi^\beta a(x, \xi)| \le c_{\alpha\beta}, \qquad |\alpha| + |\beta| \ge k.$$

For a phase space transform we use the Bargman transform T defined by

$$Tu(x, \xi) = c_n \int e^{-\frac{(x-y)^2}{2}} e^{i\xi(x-y)} u(y)\, dy.$$

This is an isometry from $L_x^2(\mathbb{R}^n)$ to $L_{x,\xi}^2(\mathbb{R}^{2n})$ and thus satisfies $T^*T = I$. However, T is not an isomorphism; instead, its range consists of functions which satisfy the Cauchy-Riemann type equation

$$(83) \qquad i\partial_\xi T = (\partial_x - i\xi)T.$$

To each pseudodifferential operator $a^w(x, D)$ we associate its phase space kernel, i.e. the kernel of the conjugated operator $Ta^w(x, D)T^*$. A simple example of the correspondence between the symbol class and the phase space kernel is the relation (see [36, Theorem 1])

$$a \in S^{(0)} \Leftrightarrow |K((x, \xi), (y, \eta))| \le c_N (1 + |(x, \xi) - (y, \eta)|)^{-N} \quad \forall\, N \in \mathbb{N}.$$

This leads to an easy proof of the Calderón-Vaillancourt theorem, which asserts that the operator a^w is L^2 bounded if $a \in S^{(0)}$.

We now turn our attention to the equation (82) where we assume a is a real symbol, is in $S^{(2)}$ uniformly in $t \in [0, 1]$, and is continuous in $t \in [0, 1]$. This suffices in order to guarantee that (82) is well-posed in L^2. We let $S(t, s)$ denote the evolution operators corresponding to

(82); these are all L^2 isometries. We denote by $K(t, s)$ the phase space kernels of $S(t, s)$, i.e. the kernels of $TS(t, s)T^*$. It is natural to try to characterize the kernels $K(t, s)$ in terms of the Hamilton flow associated to (82):

$$(84) \qquad \begin{cases} \dot{x} = a_\xi(t, x, \xi) \\ \dot{\xi} = -a_x(t, x, \xi). \end{cases}$$

The corresponding phase space evolution is denoted by $\chi(t, s)$. These are canonical transformations in $\mathbb{R}^{2n}$. Furthermore, the condition $a \in S^{(2)}$ guarantees that $\chi(t, s)$ are bilipschitz uniformly with respect to $(t, s) \in [0, 1]$. As it turns out, the phase space kernel $K(t, s)$ of $S(t, s)$ can indeed be easily characterized as follows:

Proposition 23. [36, Corollary 7.4] *Assume that a is a real symbol in $S^{(2)}$ uniformly in $t \in [0, 1]$. Then the phase space kernels $K(t, s)$ of $S(t, s)$ satisfy*

$$|K(t, x, \xi, s, y, \eta)| \leq c_N (1 + |(x, \xi) - \chi(t, s)(y, \eta)|)^{-N}.$$

We also have a corresponding Egorov theorem. For a pdo $q^w(0)$, we define its conjugate with respect to the flow by

$$q^w(t) = S(t, 0)q^w(0)S(0, t).$$

Then the counterpart of Egorov's theorem in this setting is

Proposition 24. [36, Proposition 7.6, Proposition 7.7] *Assume that a is a real symbol in $S^{(2)}$ uniformly in $t \in [0, 1]$.*

(a.) *If $q(0) \in S^{(0)}$, then $q(t) \in S^{(0)}$ uniformly in t.*
(b.) *If $q(0) \in S^{(1)}$, then $q(t) \in S^{(1)}$ uniformly in t, and*

$$q(t, x, \xi) - q(0) \circ \chi(0, t) \in S^{(0)}.$$

The counterpart of this result for $q(0) \in S^{(2)}$ is not valid in general. However, we can prove it in a special case, which will be useful later.

Proposition 25. *Let $\lambda \geq 1$. Assume that $a(t, x, \xi) = \lambda|\xi|$, and let $q(0) \in S^{(2)}$ be an operator which is localized at frequency λ. Then, $q(t) \in S^{(2)}$ uniformly in $t \in [0, 1]$ and*

$$q(t, x, \xi) - q(0) \circ \chi(0, t) \in S^{(0)}.$$

Here, analogous to the definition in Section 2, we say that an operator K is localized at frequency λ if both $\widehat{Kf}$ and $\widehat{K^*f}$ are supported in $\{2^{-10}\lambda < |\xi| < 2^{10}\lambda\}$ for all $f \in \mathcal{S}(\mathbb{R}^n)$.

We remark that, in the context of the Schrödinger equation, a similar result was proved in [37] for $a(\xi) = \xi^2$.

Proof of Proposition 25. We explicitly compute

$$\chi(t,s)(x,\xi) = (x + \lambda(t-s)\xi|\xi|^{-1}, \xi).$$

Then we want to show that

$$r^w(t,x,D) = e^{-it\lambda|D|}q^w(0,x,D)e^{it\lambda|D|} - q^w(x - \lambda t D|D|^{-1}, D) \in OPS^{(0)}$$

uniformly in $t \in [0,1]$. Compute

$$\frac{d}{dt}e^{it\lambda|D|}r^w(t,x,D)e^{-it\lambda|D|} = e^{it\lambda|D|}r_1^w(t,x,D)e^{-it\lambda|D|}$$

where

$$r_1^w(s,x,D) = -i\lambda[|D|, q^w(x - s\lambda D|D|^{-1}, D)] - \frac{d}{ds}q^w(x - s\lambda D|D|^{-1}, D).$$

Using the Weyl calculus, as $|\xi| \approx \lambda$ we get

$$r_1(s,x,\xi) \in S^{(0)}.$$

By Proposition 24, conjugation by $e^{\pm i\lambda t|D|}$ leaves the $S^{(0)}$ class unchanged from which the conclusion follows. $\qquad\qquad\square$

From the perspective of the present work, the main disadvantage of Proposition 23 is that it can only be used on a fixed time-scale. Of course, appropriate versions can be obtained for other time scales simply by rescaling. For instance, in order to obtain results which are valid up to time s we need to replace the Bargman transform T with its rescaled versions

$$T_{\frac{1}{s}}u(t,x,\xi) = c_n s^{-\frac{n}{4}} \int e^{-\frac{(x-y)^2}{2s}} e^{i\xi(x-y)} u(t,y)\, dy.$$

This is often called the FBI transform. It is still an L^2 isometry, and its range consists of functions satisfying the rescaled Cauchy-Riemann type equation

(85)
$$\frac{i}{s}\partial_\xi T_{\frac{1}{s}} = (\partial_x - i\xi)T_{\frac{1}{s}}.$$

Correspondingly, the symbol classes $S^{(k)}$ are replaced by $S_s^{(k)}$ defined by

$$a \in S_s^{(k)} := \{|\partial_x^\alpha \partial_\xi^\beta a(s,x,\xi)| \leq c_{\alpha\beta} s^{\frac{|\beta|-|\alpha|}{2}}, \quad |\alpha| + |\beta| \geq k\}.$$

These are rescaled versions of the $S^{(k)}$ spaces, and thus, results on $S^{(k)}$ can easily be transferred to these classes. In this context, the decay of phase space kernels would be measured with a rescaled distance function

$$d_s((x,\xi),(y,\eta))^2 = s^{-1}|x-y|^2 + s|\xi - \eta|^2.$$

39

Still, rescaling does not bring us closer to our goal, which is to work on an infinite time scale. This difficulty was resolved in [37] by using a time dependent scale to study the evolution (82).

9. A long time phase space parametrix

In this section, following [37], we consider global in time evolutions of the form

$$(86) \qquad (D_t + a^w(t, x, D) - ib^w(t, x, D) + c^w(t, x, D))u = 0, \quad t > 0$$

with time dependent scales for the symbols a, b, c. Precisely, we introduce the classes $l^1 S^{(k)}$ of symbols in $\mathbb{R} \times T^* \mathbb{R}^n$ whose seminorms are given by

$$\sum_j 2^{j\left(1 + \frac{|\alpha| - |\beta|}{2}\right)} \|\partial_x^\alpha \partial_\xi^\beta a(t, x, \xi)\|_{L^\infty(\{t \approx 2^j\})}, \quad |\alpha| + |\beta| \geq k.$$

When $k = 2$ we also need to better track the second derivatives of the symbols using the function $\epsilon(t)$ introduced in Section 2. We denote by $l^1 S_\epsilon^{(2)}$ the subset of $l^1 S^{(k)}$ whose seminorms are $O(\epsilon)$ when $|\alpha| + |\beta| = 2$. This additional condition can be rewritten as

$$(87) \qquad |\partial_x^\alpha \partial_\xi^\beta a(t, x, \xi)| \lesssim \frac{\epsilon(t)}{t} t^{\frac{|\beta| - |\alpha|}{2}}, \quad |\alpha| + |\beta| = 2.$$

We consider the equation (86) with a real symbol $a \in l^1 S_\epsilon^{(2)}$, which drives the evolution, $b \in l^1 S^{(1)}$ with $b \geq 0$, which is a damping term, and a possibly complex symbol $c \in l^1 S^{(0)}$, which can be regarded as a bounded error.

We remark that the symbols a, b, c above are not related to the coefficients a, b, c of P, though they play somewhat similar roles. This slight abuse of notation is harmless since at this stage our arguments no longer involve the coefficients a, b, c of P; instead all the analysis in the parametrix construction is done on the half-wave evolutions at frequency 1, using the symbols $a^\pm$.

We let $S(t, s)$ now denote the evolution operator corresponding to (86). The following result on the L^2 evolution was shown in [37] and follows from fairly standard energy estimate techniques.

Proposition 26. [37, Proposition 28] *Assume that $a \in l^1 S_\epsilon^{(2)}$ and $b \in l^1 S^{(1)}$ are real symbols with $b \geq 0$, while $c \in l^1 S^{(0)}$. Then the equation (86) is forward well-posed in $L^2(\mathbb{R}^n)$, and the corresponding evolution operators satisfy*

$$\|S(t, s)\|_{L^2 \to L^2} \lesssim 1, \qquad 0 < s < t.$$

The evolution (86) is considered in [37] using a time-dependent phase space transform. Precisely, at time t one uses the FBI transform $T_{\frac{1}{t}}$. Thus, the phase space kernels $K(t,s)$ of $S(t,s)$ are defined to be the kernels of the conjugated operators

$$\tilde{S}(t,s) = T_{\frac{1}{t}} S(t,s) T^*_{\frac{1}{s}}.$$

A main result in [37] is to establish precise bounds on the phase space kernels $K(t,s)$. These bounds are described in terms of the Hamilton flow dictated by a and the decay dictated by the damping b.

The Hamilton flow of $D_t + a^w$ is given by (84) and as above, we use $\chi(t,s)$ to denote the evolution operators. We shall use

$$t \to (x_t, \xi_t)$$

to denote the trajectories of the flow. Using the linearized equations, one can compute the Lipschitz regularity of this flow. See [37, Proposition 29]. It turns out, however, to be more convenient to parametrize $\chi(t,s)$ using the variables (x_s, ξ_t). In this context, one obtains the following regularity.

Proposition 27. [37, Equation (73)] If $a \in l^1 S_\epsilon^{(2)}$ with ϵ small and $s < t$ then

$$(88) \qquad \frac{\partial(x_t, \xi_s)}{\partial(x_s, \xi_t)} = \begin{pmatrix} I_n + \epsilon O(1) & \epsilon O(t) \\ \epsilon O\left(\frac{1}{s}\right) & I_n + \epsilon O(1) \end{pmatrix}.$$

In order to describe the decay caused by the damping, we define

$$\psi(t, x_t, \xi_t) = \int_1^t b(s, x_s, \xi_s)\, ds.$$

We expect b to cause the energy to decay like $e^{-\psi(t, x_t, \xi_t)}$ along the flow. Using the linearized flow, it can be shown that

Proposition 28. [37, Proposition 30] If $a \in l^1 S_\epsilon^{(2)}$ with ϵ small, $b \in l^1 S^{(1)}$ and $t > s$ then

$$(89) \qquad \frac{\partial(\psi(x_t, \xi_t) - \psi(x_s, \xi_s))}{\partial(x_s, \xi_t)} = (O(s^{-\frac{1}{2}}), O(t^{\frac{1}{2}})).$$

In terms of the above quantities, we can now state the pointwise bound on the kernel of the phase space operator $\tilde{S}(t,s)$. This is one of the principal results of [37].

41

Theorem 29. [37, Theorem 31] *Let $a \in l^1 S_\epsilon^{(2)}$, $b \in l^1 S^{(1)}$ be real symbols with $b \geq 0$ and $c \in l^1 S^{(0)}$. Then for $s < t$ the kernels $K(t,s)$ of the operators $\tilde{S}(t,s)$ satisfy the bound*

$$(90) \quad |K(t,x,\xi_t,s,x_s,\xi)| \lesssim t^{-\frac{n}{4}} s^{\frac{n}{4}} \left(1 + (\psi(x_s,\xi_s) - \psi(x_t,\xi_t))^2 + \frac{|x-x_t|^2}{t} + s|\xi - \xi_s|^2 \right)^{-N}.$$

To prove this result one considers the phase space evolution of $v(t) = T_{\frac{1}{t}} u(t)$ where u solves (86). As it turns out, modulo negligible errors this evolution is governed by a degenerate parabolic equation with the following components:

(a) A transport term along the Hamilton flow of a

(b) A damping term produced by b

(c) A degenerate parabolic term which is due to the change of scale in the FBI transform.

Pointwise bounds for the kernel of this degenerate diffusion are obtained in [37] using the maximum principle.

10. A PERTURBATION OF THE HALF WAVE EQUATION

The results in Theorem 29 apply for symbols a which satisfy the smallness condition $a \in l^1 S_\epsilon^{(2)}$. Instead, the symbols $a^\pm$ are a small perturbation of $\pm|\xi|$, precisely

$$a^\pm \in \pm|\xi| + l^1 S_\epsilon^{(2)}, \qquad |\xi| \approx 1, \ |x| \approx t.$$

To remedy this, in this section we consider the evolution equation

$$(D_t + |D_x| + a_0^w(t,x,D) - ib_0^w(t,x,D))u = 0$$

where $a_0 \in l^1 S_\epsilon^{(2)}$, $b_0 \in l^1 S^{(1)}$ are real symbols with $b_0 \geq 0$. Since we are interested in this evolution only at frequency 1, we will also make the symplifying assumption that a_0^w is localized at frequency 1 and that $b_0^w - \tilde{b}_0(t)$ is also localized at frequency 1. Here $\tilde{b}_0$ is simply a function of t. These assumptions guarantee that if the initial data $u(t_0)$ is localized at frequencies $\{|\xi| \in [2^{-10}, 2^{10}]\}$ then the solution u inherits this localization. The above evolution will serve as the model for our outgoing parametrix.

We denote by $S_0(t,s)$ the L^2 evolution generated by the above equation. Due to the above frequency localization of a_0^w and $b_0^w - \tilde{b}$ we have

$$S_0(t,s) S_{-10<\cdot<10} = S_{-10<\cdot<10} S_0(t,s) = S_{-10<\cdot<10} S_0(t,s) S_{-10<\cdot<10}.$$

We denote by $\tilde{S}_0(t,s)$ its (frequency localized) phase space image

$$\tilde{S}_0(t,s) = T_{\frac{1}{t}} S_0(t,s) S_{-10<\cdot<10} T_{\frac{1}{s}}^*.$$

42

We want to obtain bounds on the kernel of $\tilde{S}_0(t, s)$ which are similar to the ones in Theorem 29. As a preliminary step we need to study the regularity of the associated Hamilton flow which we denote by $\chi_0(t, s)$. This can be done directly, but for our purposes it is more convenient to reduce it to the case considered in the previous section.

At each time t we consider the symplectic map μ defined by

$$\mu_t(x, \xi) = (x + t\xi|\xi|^{-1}, \xi).$$

which corresponds to the Hamilton flow for the $D_t + |D|$ evolution. This extends to a space-time symplectic map

$$\mu(t, \tau, x, \xi) = (t, \tau - |\xi|, x + t\xi|\xi|^{-1}, \xi).$$

If p_0 is the symbol

$$p_0(t, \tau, x, \xi) = \tau + |\xi| + a_0(t, x, \xi),$$

then its image through μ is

$$p_0(\mu(t, \tau, x, \xi)) = \tau + a(t, \tau, x, \xi), \qquad a(t, x, \xi) = a_0(t, x + t\xi|\xi|^{-1}, \xi).$$

Hence the conjugate of the Hamilton flow $\chi_0(t, s)$ for $\tau + |\xi| + a_0$ with respect to μ_t is the Hamilton flow $\chi(t, s)$ for $\tau + a(t, x, \xi)$,

$$\chi_0(t, s) = \mu_t \circ \chi(t, s) \circ \mu_s^{-1}.$$

We note that $a \in l^1 S_\epsilon^{(2)}$ iff $a_0 \in l^1 S_\epsilon^{(2)}$. Hence from (88) we obtain its counterpart for the χ_0 flow,

Proposition 30. *If $a_0 \in l^1 S_\epsilon^{(2)}$ with ϵ sufficiently small and $t > s$ then the Hamilton flow $\chi_0(t, s)$ has the Lipschitz regularity*

$$(91) \qquad \frac{\partial(x_t, \xi_s)}{\partial(x_s, \xi_t)} = \begin{pmatrix} I_n + \epsilon O(1) & 2(t - s)|\xi|^{-3}(|\xi|^2 I_n - \xi \otimes \xi) + \epsilon O(t) \\ \epsilon O(\frac{1}{s}) & I_n + \epsilon O(1) \end{pmatrix}.$$

We proceed in a similar manner with b_0 and set

$$b(t, x, \xi) = b_0(t, x + t\xi|\xi|^{-1}, \xi).$$

Then the integral ψ_0 of b_0 along the χ_0 flow is the μ conjugate of the integral ψ of b along the χ flow. Hence we also trivially obtain the analog of Proposition 28, namely

Proposition 31. *If $a_0 \in l^1 S_\epsilon^{(2)}$ with ϵ sufficiently small and $b_0 \in l^1 S^{(1)}$ then for $t > s$ we have*

$$(92) \qquad \frac{\partial(\psi_0(x_t, \xi_t) - \psi_0(x_s, \xi_s))}{\partial(x_s, \xi_t)} = (O(s^{-\frac{1}{2}}), O(t^{\frac{1}{2}})).$$

43

Now we can state our main result:

Theorem 32. *Let $a_0 \in l^1 S_\epsilon^{(2)}$, $b_0 \in l^1 S^{(1)}$ be real symbols with $b_0 \geq 0$ with ϵ sufficiently small, so that a_0 and $b_0 - \tilde{b}_0(t)$ are localized at frequency 1. Then for $s < t$ the kernel K_0 of the operator $\tilde{S}_0(t, s)$ satisfies the bound*

(93)
$$|K_0(t, x, \xi_t, s, x_s, \xi)| \lesssim t^{-\frac{n}{4}} s^{\frac{n}{4}} \left(1 + (\psi_0(x_s, \xi_s) - \psi_0(x_t, \xi_t))^2 + \frac{|x - x_t|^2}{t} + s|\xi - \xi_s|^2 \right)^{-N}$$
$$\times (1 + t \ d(|\xi_s|, [2^{-10}, 2^{10}]))^{-N}.$$

Proof. We use Theorem 29 via a conjugation with respect to the flat half-wave flow, which corresponds to the canonical transformations μ_t. Denote

$$S(t, s) = e^{it|D|} S_0(t, s) e^{-is|D|}.$$

Then we compute

$$\frac{d}{dt} S(t, s) = -i e^{it|D|} (-a_0^w(t, x, D) + i b_0^w(t, x, D)) e^{-it|D|} S(t, s).$$

Hence $S(t, s)$ is the evolution associated to the pseudodifferential operator

$$e^{it|D|} (-a_0^w(t, x, D) + i b_0^w(t, x, D)) e^{-it|D|}.$$

Using rescaled versions of Propositions 24,25, this operator can be expressed in the form

$$a^w(t, x, D) - i b^w(t, x, D) + c^w(t, x, D)$$

where the remainder term satisfies $c \in l^1 S^{(0)}$. Hence the phase space kernel of $S(t, s)$ satisfies the bounds given by Theorem 29.

Returning to the original equation, for the phase space evolution $\tilde{S}_0(t, s)$ we can write

$$\tilde{S}_0(t, s) = T_{\frac{1}{t}} e^{-it|D|} S_{-10<\cdot<10} S(t, s) S_{-10<\cdot<10} e^{is|D|} T_{\frac{1}{s}}^*$$
$$= T_{\frac{1}{t}} e^{-it|D|} S_{-10<\cdot<10} T_{\frac{1}{t}}^* T_{\frac{1}{t}} S(t, s) T_{\frac{1}{s}}^* T_{\frac{1}{s}} S_{-10<\cdot<10} e^{is|D|} T_{\frac{1}{s}}^*$$
$$= (T_{\frac{1}{t}} e^{-it|D|} S_{-10<\cdot<10} T_{\frac{1}{t}}^*) \tilde{S}(t, s) (T_{\frac{1}{s}} S_{-10<\cdot<10} e^{is|D|} T_{\frac{1}{s}}^*).$$

By a rescaled version of Proposition 23 the kernel of the first factor $T_{\frac{1}{t}} e^{-it|D|} S_{-10<\cdot<10} T_{\frac{1}{t}}^*$ is rapidly decreasing on the $t^{\frac{1}{2}} \times t^{-\frac{1}{2}}$ scale away from the graph of μ_t as well as away from the support of the symbol $S_{-10<\cdot<10}$, while the kernel of the last factor $T_{\frac{1}{s}} e^{is|D|} S_{-10<\cdot<10} T_{\frac{1}{s}}^*$ is rapidly decreasing on the $s^{\frac{1}{2}} \times s^{-\frac{1}{2}}$ scale away from the graph of μ_s^{-1} as well as away from

the support of the symbol $S_{-10<\cdot<10}$. Hence the composition simply replaces the Hamilton flow associated to a by the Hamilton flow associated to a_0 and the function ψ with ψ_0 in the kernel bounds. Thus (90) implies (93), and the proof is concluded. $\qquad\square$

11. THE PARAMETRIX CONSTRUCTION

We end with a proof of Proposition 15. That is, we construct a parametrix K_0^+ for $D_t + A_{(0)}^+$, and easy modifications yield also a parametrix for $D_t + A_{(0)}^-$. In the sequel, we shall drop the $+$ signs and denote these by K_0 and $D_t + A_{(0)}$ respectively. The $\pm$ signs will be reserved to distinguish waves which are outgoing forward, respectively backward, in time.

We partition the annulus $|\xi| \approx 1$ in phase space

$$s_{-1}(\xi) + s_0(\xi) + s_1(\xi) = \sum_{\pm} \sum_{j \geq 0} p_j^{\pm}(x, \xi),$$

with

$$\operatorname{supp} p_j^{\pm} \subset \{2^{j-1} < |x| < 2^{j+1}, \ \pm x\xi \geq -2^{-5}|x|\}, \quad j \geq 1,$$

$$\operatorname{supp} p_0^{\pm} \subset \{|x| < 2, \ \pm x\xi \geq -2^{-5}|x|\}.$$

At the expense of Schwartz tails which play no role in the sequel, we may replace $p_j^{\pm}$ by $S_{<-10}(D_x)p_j^{\pm}$. As such, we shall do so without changing the notation. This allows us to assume that the operators $P_j^{\pm}$ are frequency localized to frequency 1.

In the proposition which follows, we construct evolution operators $S_j^{\pm}(t, s)$ as the evolutions associated to a certain damped half-wave equation. We then form K_0 by setting

$$K_0(t, s) = \begin{cases} \sum_{j=1}^{\infty} S_j^-(t, s)(P_j^-)^w(x, D), & t < s \\ \sum_{j=1}^{\infty} S_j^+(t, s)(P_j^+)^w(s, D), & t > s. \end{cases}$$

The properties of K_0 listed in Proposition 15 follow easily, after summing, from the given properties of $S_j^{\pm}$.

Proposition 33. *Assume that ϵ is sufficiently small. Then for each $s \in \mathbb{R}$, there is an outgoing parametrix S_j^+ for $D_t + A_{(0)}$ in $\{t > s\}$ which is localized at frequency 1 and satisfies the following:*

(i) *L^2 bound:*

$$\|S_j^+(t, s)\|_{L^2 \to L^2} \lesssim 1$$

(ii) *Error estimate:*

(94)
$$\|x^{\alpha}(D_t + A_{(0)})S_j^+(t, s)P_j^+\|_{L^2 \to L^2} \lesssim (2^j + |t - s|)^{-N}$$
$$\|x^{\alpha}D_t(D_t + A_{(0)})S_j^+(t, s)P_j^+\|_{L^2 \to L^2} \lesssim (2^j + |t - s|)^{-N}$$

(iii) *Initial data:*
$$S_j^+(s+0, s) = I$$

(iv) *Outgoing parametrix:*

(95)
$$\|\mathbf{1}_{\{|x|<2^{-10}(|t-s|+2^j)\}} S_j^+(t,s)P_j^+\|_{L^2 \to L^2} \lesssim (|t-s|+2^j)^{-N}$$

(v) *Finite speed:*

(96)
$$\|x^\alpha \mathbf{1}_{\{|x|>2^{10}(|t-s|+2^j)\}} S_j^+(t,s)P_j^+\|_{L^2 \to L^2} \lesssim (|t-s|+2^j)^{-N}$$

(vi) *Frequency localization:*

(97)
$$\|(1-P_{[-4,4]})S_j^+(t,s)P_j^+\|_{L^2 \to L^2} \lesssim (|t-s|+2^j)^{-N}$$

(vii) *Pointwise decay:*

(98)
$$\|S_j^+(t,s)P_j^+\|_{L^1 \to L^\infty} \lesssim (1+|t-s|)^{-\frac{n-1}{2}}.$$

With obvious modifications, the same hold for S_j^-.

By translation invariance, without any loss of generality we may assume that $s = 2^j$. We first reduce the problem to the study of an evolution of a perturbed half-wave equation as in Section 10. Heuristically we observe that in the support of the symbol of P_j^+ we have

$$|\xi| \in [2^{-2}, 2^2], \qquad |x| \approx s, \qquad x \cdot \xi \geq -\frac{1}{5}|x||\xi|.$$

An easy computation shows that along the forward Hamilton flow starting here we have

$$|\xi| \in [2^{-3}, 2^3], \qquad |x| \approx t.$$

But in this region we have

$$a_{(0)}(t, x, \xi) - |\xi| \in l^1 S_\epsilon^{(2)}$$

which follows from the analog of (32) which holds for $a_{(0)}$. Thanks to (95), (96) and (97), we can freely modify the symbol of $a_{(0)}$ in the regions $\{|x| \ll t\}$ and $\{|\xi| \notin [2^{-5}, 2^5]\}$ at the expense of producing a negligible error in (94). It thus suffices to study the evolution governed by a symbol

$$|\xi| + a_0(t, x, \xi), \quad a_0 \in l^1 S_\epsilon^{(2)}$$

so that a_0 vanishes if $\{|\xi| \notin [2^{-6}, 2^6]\}$ and a_0^w is localized at frequency 1.

In essence, $a_0 = a_{(0)} - |\xi|$, and thus by (32), we may assume the better regularity

(99)
$$|\partial_x^\alpha \partial_\xi^\beta a_0(t, x, \xi)| \lesssim \epsilon(t)t^{-|\alpha|}, \quad |\alpha| \leq 2$$
$$|\partial_x^\alpha \partial_\xi^\beta a_0(t, x, \xi)| \lesssim \epsilon(t)t^{-1-\frac{|\alpha|}{2}}, \quad |\alpha| \geq 2.$$

This additional decay shall be used on time scales which are too small to allow $s^{\frac{1}{2}} \times s^{-\frac{1}{2}}$ packets at time s to separate in time t.

Unfortunately, simply defining the parametrix S_j^+ by the evolution associated to the operator $|D| + a_{(0)}$ does not seem to work. Precisely, the bounds (95), (96) and (97) appear to fail. This is because at each time t, there is leakage caused by the uncertainty principle to the regions appearing in (95), (96) and (97), which are outside the propagation region indicated by the Hamilton flow. While this leakage does have rapid spatial decay, its time evolution yields output which does not have the rapid decay in time as needed in (95), (96) and (97).

Thus, in order to be able to prove the rapid t-decay in, e.g., (95) (96) and (97), we shall introduce an artificial damping term $b_0 \in l^1 S^{(1)}$, $b_0(t, x, \xi) \geq 0$. The role of b_0 is precisely to put a damping on the time evolution of the above mentioned leakeage. At the same time, b_0 is taken to be 0 in the main propagation region. We would like to define $S_j^+(t, s)$ to be the forward evolution operator of the equation

$$(D_t + |D| + a_0^w(t, x, D))u = ib_0^w(t, x, D)u.$$

However, in order to insure the frequency localization of our parametrix we replace $S_j^+(t, s)$ by the truncated operator

$$S_{[-7,7]}(D_x) \cdot S_j^+.$$

We shall show that

(100) $$\|x^\alpha S_{<-5}(D_x)S_j^+(t, s)P_j^+\|_{L^2 \to L^2} \lesssim (|t - s| + 2^j)^{-N},$$

(101) $$\|x^\alpha \partial^\beta S_{>5}(D_x)S_j^+(t, s)P_j^+\|_{L^2 \to L^2} \lesssim (|t - s| + 2^j)^{-N},$$

and thus, the errors in (94) which result from this truncation are negligible. We shall further prove the following bound on the damping term

(102)
$$\|x^\alpha b_0^w(t, x, D)S_j^+(t, s)P_j^+\|_{L^2 \to L^2} \lesssim (|t - s| + 2^j)^{-N},$$
$$\|x^\alpha D_t b_0^w(t, x, D)S_j^+(t, s)P_j^+\|_{L^2 \to L^2} \lesssim (|t - s| + 2^j)^{-N},$$

which shall yield (94).

With $S_j^+(t, s)$ now fixed, property (iii) is trivial, and (i) follows from Proposition 26. We proceed to the argument which yields our main pointwise bound (98). Here, we examine three cases separately.

Case 1: $|t - s| \geq s$. In this regime, we may neglect the damping. For initial data $u(s) = \delta_0$, we have

$$T_{\frac{1}{s}}u(s, x_s, \xi) = s^{-\frac{n}{4}}e^{-\frac{x_s^2}{2s}}e^{ix_s\xi}.$$

Using Theorem 32, we see that

$$|T_{\frac{1}{t}}u(t,x,\xi_t)| \lesssim t^{-\frac{n}{4}} \int (1+t^{-1}|x-x_t(\xi_t,x_s)|^2)^{-N}(1+s|\xi-\xi_s(x_s,\xi_t)|^2)^{-N}e^{-\frac{x_s^2}{2s}} \, dx_s \, d\xi$$

$$\lesssim t^{-\frac{n}{4}}s^{-\frac{n}{2}} \int (1+t^{-1}|x-x_t(\xi_t,x_s)|^2)^{-N}e^{-\frac{x_s^2}{2s}} \, dx_s.$$

For the remaining integral, we use that $x_s \to x_t(\xi_t,x_s)$ is Lipschitz. See (91). Integrating in x_s then yields

$$|T_{\frac{1}{t}}u(t,x,\xi_t)| \lesssim t^{-\frac{n}{4}}(1+t^{-1}|x-x_t(\xi_t,0)|^2)^{-N},$$

and by applying $T_{\frac{1}{t}}^*$, we have

$$|u(t,y)| \lesssim t^{-\frac{n}{2}} \int (1+t^{-1}|x-x_t(\xi_t,0)|^2)^{-N}e^{-\frac{|y-x|^2}{2t}} \, dx \, d\xi_t$$

$$\lesssim \int (1+t^{-1}|y-x_t(\xi_t,0)|^2)^{-N} \, d\xi_t.$$

If $|t-s| \geq s$, then the map $\xi_t \to x_t(\xi_t,0)$ is zero homogeneous, by (91) has Lipschitz constant which is bounded by t, and has maximal rank $n-1$. Hence, integration with respect to ξ_t yields

$$|u(t,y)| \lesssim t^{-\frac{n-1}{2}}.$$

Case 2: $1 \leq |t-s| \leq s$. Here, we reinitialize the time scale to prevent difficulties which result from $s^{\frac{1}{2}} \times s^{-\frac{1}{2}}$ packets at time s not separating before time t. In addition to (99), we similarly require

(103)
$$|\partial_x^\alpha \partial_\xi^\beta b_0(t,x,\xi)| \lesssim t^{-\frac{1}{2}-|\alpha|}, \quad |\alpha| \leq 1,$$
$$|\partial_x^\alpha \partial_\xi^\beta b_0(t,x,\xi)| \lesssim t^{-1-\frac{|\alpha|}{2}}, \quad |\alpha| \geq 1$$

for $|t-s| < 2^j$. The additional regularity (99) and (103) is sufficient to show that a_0, b_0 remain in the appropriate symbol classes after the time translation which sets the initial time to $t-s$. Theorem 32 thus remains valid, and the bound follows from the computation above in the translated coordinates.

Case 3: $0 \leq |t-s| \leq 1$. Here, since our initial data is localized at frequency 1, we may simply use Sobolev embeddings combined with the L^2 bounds from Proposition 26:

$$\|S_j^+(t,s)P_j^+u_0\|_{L^\infty} \lesssim \|S_j^+(t,s)P_j^+u_0\|_{L^2} \lesssim \|P_j^+u_0\|_{L^2} \lesssim \|u_0\|_{L^1}.$$

The rest of the proof is based on properties of b_0. In particular, we use a construction which is quite similar to that of [37] to build a b_0 which allows us to prove the remaining required estimates: (95), (96), (100), (101), and (102). In particular, we have

Lemma 34. *There exists a symbol* $b \in l^1 S^{(1)}$ *which satisfies, in addition to* (103),

48

(b1) $t^{\frac{3}{4}}b$ is nonincreasing along the Hamiltonian flow for $D_t + |D_x| + a_0^w$, and

$$0 < t^{\frac{3}{4}}b(t, x_t, \xi_t) < 1 \implies b(2t, x_{2t}, \xi_{2t}) = 0.$$

(b2) *At the initial time, we have*

$$b(2^j, x, \xi) = 0, \quad in \quad \{2^{-3} < |\xi| < 2^3, \quad 2^{j-2} < |x| < 2^{j+2}, \quad x\xi > -2^{-4}|x|\}.$$

(b3) *At any time $t \geq 2^j$, we have*

$$b(t, x, \xi) = t^{-\frac{3}{4}}, \quad outside \quad \{2^{-4} < |\xi| < 2^4\} \cap \{2^{-6}t < |x| < 2^6 t\}.$$

Before proving this lemma, let us explain how such a damping term can be used to complete the proof of Proposition 15. Indeed, we have the following lemma which is essentially from [37]:

Lemma 35. *Assume that the symbol $b_0 \in l^1 S^{(1)}$ satisfies the properties (b1), (b2), and (b3) above. Then, the bounds (95), (96), (97), (100), (101), and (102) hold.*

Indeed, once Theorem 32 has been established, the necessary modifications to the arguments from [37] are quite simple. It only remains to establish the second estimate from (102) which did not appear in [37]. Here, however, we note that, modulo negligible errors due to the frequency truncation of S^+,

$$D_t b_0^w(t, x, D)S^+(t, s) = -i(\partial_t b_0)^w(t, x, D)S^+(t, s) - b_0^w(t, x, D)|D_x|S^+(t, s)$$
$$- b_0^w(t, x, D)a_0^w(t, x, D)S^+(t, s) + ib_0^w(t, x, D)b_0^w(t, x, D)S^+(t, s).$$

Since the symbols of $(\partial_t b_0)^w$, $b_0^w|D_x|$, $b_0^w a_0^w$, and $b_0^w b_0^w$ are all in $S_t^{(1)}$ and have supports which are contained in the support of b_0, we may similarly apply Proposition 17 of [37] to obtain the estimate.

It now only remains to complete the construction of said damping terms b.

Proof of Lemma 34. We define the increasing bounded function $e(s)$ by

$$e(s) = \epsilon^{-1} \int_0^s \frac{\epsilon(\sigma)}{\sigma} d\sigma.$$

Letting ϕ be a smooth, nondecreasing cutoff function which equals 0 in $(-\infty, 0)$ and 1 in $(1, \infty)$, we set

$$b(t, x, \xi) = t^{-\frac{3}{4}}(1 - \phi(b_1)\phi(b_2)\phi(b_3)\phi(b_4)\phi(b_5))$$

with

- *Cutoff frequencies which are too large*

$$b_1(t, \xi) = \frac{2^{7/2} + e(t) - |\xi|}{\epsilon(t)},$$

- *Cutoff frequencies which are too small*

$$b_2(t, \xi) = \frac{|\xi| - 2^{-7/2} + ce(t)}{\epsilon(t)}$$

  where c is a fixed small constant,

- *Select outgoing waves*

$$b_3(t, x, \xi) = \frac{2^{-\frac{1}{2}}|x||\xi| + x\xi}{2^{-12}|x|},$$

- *Cutoff values of $|x|$ which are too large*

$$b_4(t, x) = \frac{2^6 t - |x|}{t},$$

- *Cutoff values of $|x|$ which are too small*

$$b_5(t, x, \xi) = \frac{|x||\xi| - 2^{-5} t|\xi| + x\xi}{2^{-10} t}.$$

We note that

$$\{2^{-3} < |\xi| < 2^3\} \cap \{2^{-2} t < |x| < 2^2 t\} \cap \{x\xi > -2^{-4}|x|\} = D_t \subset \{b = 0\}$$

if ϵ is sufficiently small, while

$$\{t^{\frac{3}{4}} b < 1\} \subset E_t = \{2^{-4} < |\xi| < 2^4\} \cap \{2^{-6} t < |x| < 2^6 t\} \cap \{x\xi > -2^{-1/2}|x||\xi|\}.$$

So, the conditions (b2) and (b3) are easily satisfied.

To prove (b1), it suffices to study the behavior of b along the Hamilton flow within E_t and show that for each b_j, we have

(104) $$\frac{d}{dt} b_j(t, x_t, \xi_t) \geq \frac{2}{t}, \quad \text{in } E_t \cap \{0 \leq b_j \leq 1\}.$$

Here $t \to (x_t, \xi_t)$ now denotes a trajectory of the flow for $D_t + |D_x| + a_0^w$. For $(x_t, \xi_t) \in E_t$, we have

$$\frac{d}{dt} \xi_t = O\left(\frac{\epsilon(t)}{t}\right), \quad \frac{d}{dt} x_t = \frac{\xi_t}{|\xi_t|} + O(\epsilon(t)).$$

We simply calculate

$$\frac{d}{dt} b_1(t, \xi_t) \geq \frac{1}{\epsilon t} - \frac{1}{t} - \frac{\epsilon'(t)}{\epsilon^2(t)} (2^{7/2} + e(t) - |\xi_t|) \geq \frac{2}{t}, \quad \text{in } \{0 \leq b_1 \leq 1\}$$

for ϵ sufficiently small. The computation for b_2 is identical. For b_3, we have

$$\frac{d}{dt} b_3(t, x_t, \xi_t) = \frac{|\xi_t|^2 |x_t|^2 - (x_t \xi_t)^2}{2^{-12}|x_t|^3 |\xi_t|} + \frac{O(\epsilon(t))}{t} \geq \frac{2}{t}, \quad \text{in } E_t \cap \{0 \leq b_3 \leq 1\}.$$

50

For b_4, we compute

$$\frac{d}{dt}b_4(t, x_t) = \frac{|x_t|^2|\xi_t| - tx_t\xi_t}{t^2|x_t||\xi_t|} + \frac{O(\epsilon(t))}{t} \geq \frac{2^5}{t}, \quad \text{in } 0 \leq b_4 \leq 1.$$

Finally, for b_5 we also compute

$$\frac{d}{dt}b_5(t, x_t, \xi_t) = \frac{|x_t|^{-1}x_t\xi_t + |\xi_t|}{2^{-10}t} - \frac{|x_t||\xi_t| + x_t\xi_t}{2^{-10}t^2} + \frac{O(\epsilon(t))}{t}$$

$$\geq \frac{2^5|\xi_t|}{t} \geq \frac{2}{t}, \quad \text{in } E_t \cap \{0 \leq b_5 \leq 1\}.$$

It remains to verify (103), and hence that $b \in l^1 S^{(1)}$, but this is straightforward. $\qquad\square$

References

[1] Serge Alinhac. On the Morawetz–Keel-Smith-Sogge inequality for the wave equation on a curved background. *Publ. Res. Inst. Math. Sci.*, 42(3):705–720, 2006.

[2] Jean-Marc Bouclet and Nikolay Tzvetkov. On global Strichartz estimates for non-trapping metrics. *J. Funct. Anal.*, 254(6):1661–1682, 2008.

[3] Philip Brenner. On $L_p - L_{p'}$ estimates for the wave-equation. *Math. Z.*, 145(3):251–254, 1975.

[4] Peter Constantin and Jean-Claude Saut. Effets régularisants locaux pour des équations dispersives générales. *C. R. Acad. Sci. Paris Sér. I Math.*, 304(14):407–410, 1987.

[5] Walter Craig, Thomas Kappeler, and Walter Strauss. Microlocal dispersive smoothing for the Schrödinger equation. *Comm. Pure Appl. Math.*, 48(8):769–860, 1995.

[6] Jean-Marc Delort. *F.B.I. transformation.* Springer-Verlag, Berlin, 1992. Second microlocalization and semilinear caustics.

[7] Shin-ichi Doi. Smoothing effects for Schrödinger evolution equation and global behavior of geodesic flow. *Math. Ann.*, 318(2):355–389, 2000.

[8] Charles L. Fefferman. The uncertainty principle. *Bull. Amer. Math. Soc. (N.S.)*, 9(2):129–206, 1983.

[9] Gerald B. Folland. *Harmonic analysis in phase space.* Princeton University Press, Princeton, NJ, 1989.

[10] J. Ginibre and G. Velo. Smoothing properties and retarded estimates for some dispersive evolution equations. *Comm. Math. Phys.*, 144(1):163–188, 1992.

[11] Andrew Hassell, Terence Tao, and Jared Wunsch. A Strichartz inequality for the Schrödinger equation on nontrapping asymptotically conic manifolds. *Comm. Partial Differential Equations*, 30(1-3):157–205, 2005.

[12] Markus Keel, Hart F. Smith, and Christopher D. Sogge. Almost global existence for some semilinear wave equations. *J. Anal. Math.*, 87:265–279, 2002. Dedicated to the memory of Thomas H. Wolff.

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

[14] Carlos E. Kenig, Gustavo Ponce, and Luis Vega. On the Zakharov and Zakharov-Schulman systems. *J. Funct. Anal.*, 127(1):204–234, 1995.

[15] Herbert Koch and Daniel Tataru. Dispersive estimates for principally normal pseudodifferential operators. *Comm. Pure Appl. Math.*, 58(2):217–284, 2005.

[16] Peter D. Lax and Ralph S. Phillips. *Scattering theory*, volume 26 of *Pure and Applied Mathematics.* Academic Press Inc., Boston, MA, second edition, 1989. With appendices by Cathleen S. Morawetz and Georg Schmidt.

[17] Jeremy Marzuola, Jason Metcalfe, and Daniel Tataru. Strichartz estimates and local smoothing estimates for asymptotically flat Schrödinger equations. *J. Funct. Anal.*, 255(6):1497–1553, 2008.

[18] Jason Metcalfe and Christopher D. Sogge. Long-time existence of quasilinear wave equations exterior to star-shaped obstacles via energy methods. *SIAM J. Math. Anal.*, 38(1):188–209 (electronic), 2006.

[19] 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.

[20] Cathleen S. Morawetz. Time decay for the nonlinear Klein-Gordon equations. *Proc. Roy. Soc. Ser. A*, 306:291–296, 1968.

[21] Luc Robbiano and Claude Zuily. Strichartz estimates for Schrödinger equations with variable coefficients. *Mém. Soc. Math. Fr. (N.S.)*, (101-102):vi+208, 2005.

[22] I. Rodnianski and T. Tao. Longtime decay estimates for the Schrödinger equation on manifolds. In *Mathematical aspects of nonlinear dispersive equations*, volume 163 of *Ann. of Math. Stud.*, pages 223–253. Princeton Univ. Press, Princeton, NJ, 2007.

[23] Per Sjölin. Regularity of solutions to the Schrödinger equation. *Duke Math. J.*, 55(3):699–715, 1987.

[24] Johannes Sjöstrand. Singularités analytiques microlocales. In *Astérisque, 95*, pages 1–166. Soc. Math. France, Paris, 1982.

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

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

[27] Hart F. Smith and Christopher D. Sogge. Global Strichartz estimates for nontrapping perturbations of the Laplacian. *Comm. Partial Differential Equations*, 25(11-12):2171–2183, 2000.

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

[29] Gigliola Staffilani and Daniel Tataru. Strichartz estimates for a Schrödinger operator with nonsmooth coefficients. *Comm. Partial Differential Equations*, 27(7-8):1337–1372, 2002.

[30] Walter A. Strauss. Dispersal of waves vanishing on the boundary of an exterior domain. *Comm. Pure Appl. Math.*, 28:265–278, 1975.

[31] Robert S. Strichartz. Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations. *Duke Math. J.*, 44(3):705–714, 1977.

[32] Daniel Tataru. Strichartz estimates for operators with nonsmooth coefficients and the nonlinear wave equation. *Amer. J. Math.*, 122(2):349–376, 2000.

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

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

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

[36] Daniel Tataru. Phase space transforms and microlocal analysis. In *Phase space analysis of partial differential equations. Vol. II*, Pubbl. Cent. Ric. Mat. Ennio Giorgi, pages 505–524. Scuola Norm. Sup., Pisa, 2004.

[37] Daniel Tataru. Parametrices and dispersive estimates for Schrödinger operators with variable coefficients. *Amer. J. Math.*, 130(3):571–634, 2008.

[38] Michael E. Taylor. *Tools for PDE*, volume 81 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000. Pseudodifferential operators, paradifferential operators, and layer potentials.

[39] Luis Vega. Schrödinger equations: pointwise convergence to the initial data. *Proc. Amer. Math. Soc.*, 102(4):874–878, 1988.

Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599-3250, USA

Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA