

PROPAGATION OF SINGULARITIES FOR ROUGH METRICS

HART F. SMITH

ABSTRACT. We use a wave packet transform and weighted norm estimates in phase space to establish propagation of singularities for solutions to time-dependent scalar hyperbolic equations that have coefficients of limited regularity. It is assumed that the second order derivatives of the principal coefficients belong to $L_t^1 L_x^\infty$, and that u is a solution to the homogeneous equation, of global Sobolev regularity $s_0 = 0$ or 1 . It is then proven that the H^{s_0+1} wavefront set of u is a union of maximally extended null bicharacteristic curves.

1. INTRODUCTION

In this paper we establish a propagation of singularities theorem for second-order, scalar hyperbolic operators of $(t, x) \in (-T, T) \times \mathbb{R}^n$ of the form

$$L = D_t^2 - 2b^j(t, x)D_j D_t - c^{ij}(t, x)D_i D_j + d^0(t, x)D_t + d^j(t, x)D_j, \quad 1 \leq i, j \leq n,$$

where summation notation is used, and $D_t = -i\partial_t$, $D_j = -i\partial_{x_j}$ for $1 \leq j \leq n$. Under the assumption that the second derivatives of the principal coefficients belong to $L_t^1 L_x^\infty$, we establish the following.

Theorem 1.1. *Suppose that $s_0 \in \{0, 1\}$. Suppose that $Lu = 0$, and that*

$$(1.1) \quad u \in C^0((-T, T), H^{s_0}(\mathbb{R}^n)), \quad D_t u \in C^0((-T, T), H^{s_0-1}(\mathbb{R}^n)).$$

If $\gamma(t)$ is a null bicharacteristic curve of L , and $\gamma(t_0) \notin WF_{s_0+1}(u)$ for some $t_0 \in (-T, T)$, then $\gamma \cap WF_{s_0+1}(u) = \emptyset$.

The improvement of this paper over prior results for twice-differentiable coefficients is the gain of 1 derivative over the background regularity, which we show by example is the best possible in the setting we consider. Also, we assume integrability in t of the second order derivatives, as opposed to uniform bounds, which by a limiting argument will show that the theorem holds for piecewise regular coefficients. By Theorem 8.2, the assumptions on u imply that $WF_{s_0+1}(u)$ is contained in the characteristic set of L , so the restriction to null bicharacteristics is natural. Theorem 1.1 can be localized in x , see Remark 8.7. Also, for $s_0 = 1$ the regularity assumption on u may be reduced to $u \in H^1((-T, T) \times \mathbb{R}^n)$ by Theorem 8.6. For $s_0 = 0$ it is not clear how to interpret Lu in case $u \in L^2((-T, T) \times \mathbb{R}^n)$. However, if L is of divergence form, or if the regularity assumption on the coefficients of L is increased to $b^j, c^{ij} \in C^{1,1}((-T, T) \times \mathbb{R}^n)$, and $d^0, d^j \in C^{0,1}((-T, T) \times \mathbb{R}^n)$, then Remark 8.8 shows that Theorem 1.1 holds for L^2 solutions.

This material is based upon work supported by the National Science Foundation under Grant DMS-1161283. This work was partially supported by a grant from the Simons Foundation (#266371 to Hart Smith).

Hörmander's theorem [9] on propagation of singularities for operators of real principal type shows that if the coefficients of L are C^∞ , then the conclusion of Theorem 1.1 holds for all s_0 , with no global regularity condition required of u . Propagation of singularities theorems in the setting of nonsmooth nonlinear equations were obtained by Bony [2], where the method of paradifferential operators was introduced. In that case a local regularity assumption is required of u . Related work on nonlinear equations includes Rauch and Reed [10], and Beals and Reed [1]. More closely related to this paper, Taylor used the positive commutator method and paradifferential theory in [15] to establish propagation of singularities for linear differential equations, including results for coefficients of Hölder regularity less than 2. In the case of $C^{1,1}$ coefficients, Proposition 11.4 of [15, Chapter 3] implies invariance of the H^{s_0+1} wavefront set if $u \in H^{s_0+\epsilon}$, any $\epsilon > 0$, for $s_0 \in [-1, 1]$. In [7], de Hoop et. al. studied reflection of the H^{s_0} wavefront set off conormal singularities of metrics with singularities of Hölder regularity $C^{1,\alpha}$, where $0 < \alpha \leq 1$. The limiting result in [7] for $\alpha = 1$ would be a gain of $1/2$ derivative relative to the assumed background regularity of u . For C^2 metrics in domains with C^3 boundary, Burq [3] established the propagation result for microlocal defect measures. In the setting of [3], as well as in that of [7], there may be multiple generalized bicharacteristics passing through a given initial point in phase space.

We now make more precise the regularity conditions that we place on the coefficients. We assume that the coefficient functions b^j and c^{ij} are real, and that the equation is uniformly hyperbolic in t ,

$$(1.2) \quad \sum_{i,j=1}^n c^{ij}(t,x) \xi_i \xi_j \geq c_0 |\xi|^2, \quad c_0 > 0.$$

The b^j and c^{ij} are assumed continuously differentiable, with uniform bounds

$$(1.3) \quad \sup_{|t|<T, x \in \mathbb{R}^n} \sum_{|\gamma| \leq 1} \left(|\partial_{t,x}^\gamma b^j(t,x)| + |\partial_{t,x}^\gamma c^{ij}(t,x)| \right) \leq C_0.$$

In addition, we assume that the second-order derivatives of b^j and c^{ij} belong to $L^1 L^\infty$. Precisely, we assume that their distributional derivatives of second-order are locally integrable functions of (t,x) , and that there exists a function $\alpha(t) \in L^1((-T, T))$ such that

$$(1.4) \quad \sup_{x \in \mathbb{R}^n} \sum_{|\gamma|=2} \left(|\partial_{t,x}^\gamma b^j(t,x)| + |\partial_{t,x}^\gamma c^{ij}(t,x)| \right) \leq \alpha(t).$$

This condition in fact implies that the coefficients are continuously differentiable functions of (t,x) , so that the assumption of C^1 coefficients (as opposed to Lipschitz) is redundant. It also follows from (1.4) that

$$(1.5) \quad \|c^{ij}(t, \cdot) - c^{ij}(s, \cdot)\|_{C^1(\mathbb{R}^n)} \leq \int_s^t \alpha(r) dr,$$

so the map $s \rightarrow c^{ij}(s, \cdot)$ is continuous from $(-T, T)$ into $C^1(\mathbb{R}^n)$, similarly for b^j .

The coefficients d^0 and d^j are assumed to have the same regularity as the first order derivatives of b^j and c^{ij} ; that is, d^0 and d^j are assumed to be continuous functions of (t,x) with uniform upper bounds, with first order derivatives in $L^1 L^\infty$.

Precisely, for d denoting either d^0 or d^j , we assume bounds with $\alpha(t)$ as above,

$$(1.6) \quad \sup_{|t| < T, x \in \mathbb{R}^n} |d(t, x)| \leq C_0, \quad \sup_{x \in \mathbb{R}^n} \sum_{|\gamma|=1} |\partial_{t,x}^\gamma d(t, x)| \leq \alpha(t).$$

The coefficients of L all admit extensions to $\mathbb{R} \times \mathbb{R}^n$ with the same regularity. For example, consider $c(t, x)$ defined on $t > 0$ with second order derivatives belonging to $L^1 L^\infty(\mathbb{R}_+ \times \mathbb{R}^n)$. By (1.5), $c(t, x)$ extends to a C^1 function on $t \geq 0$. For $t < 0$ define

$$(1.7) \quad c(t, x) = 3c(-t, x) - 2c(-2t, x).$$

It is then easily verified that all second order distributional derivatives of c belong to $L^1 L^\infty(\mathbb{R}^{1+n})$, and that the same extension preserves the first order regularity of d^0, d^j . For convenience we will thus assume when needed that all coefficients of L have been extended to $\mathbb{R} \times \mathbb{R}^n$, and in addition that L equals the standard wave operator for $|t| \geq T + 1$.

We note that the product of functions satisfying (1.3) and (1.4) is of the same type, hence there is no loss of generality in our assumption that the coefficient of D_t^2 is 1. Such an L can also be written in divergence form

$$L = D_t^2 - 2D_j b^j(t, x) D_t - D_i c^{ij}(t, x) D_j + \tilde{d}^0(t, x) D_t + \tilde{d}^j(t, x) D_j,$$

for \tilde{d} satisfying (1.6). This form will be more convenient for certain proofs.

Consider the principal symbol of L , where (τ, ξ) are the phase space coordinates dual to (t, x) ,

$$H(t, x, \tau, \xi) = \tau^2 - 2 \sum_{j=1}^n b^j(t, x) \xi_j \tau - \sum_{i,j=1}^n c^{ij} \xi_i \xi_j.$$

This factors as

$$H(t, x, \tau, \xi) = (\tau - p_+(t, x, \xi))(\tau + p_-(t, x, \xi))$$

where

$$(1.8) \quad \begin{aligned} p_\pm(t, x, \xi) &= p(t, x, \xi) \pm b^j(t, x) \xi_j, \\ p(t, x, \xi) &= \left(c^{ij}(t, x) \xi_i \xi_j + \left(b^j(t, x) \xi_j \right)^2 \right)^{\frac{1}{2}}. \end{aligned}$$

We modify $p(t, x, \xi)$ near $\xi = 0$ so that it is smooth in ξ , and homogeneous of degree 1 for $|\xi| > 1$. The symbols p and p_\pm are continuously differentiable in (t, x) , and satisfy

$$(1.9) \quad \begin{aligned} \sup_{|\xi|=1} \sup_{|\beta| \leq 1} |\partial_\xi^\gamma \partial_{t,x}^\beta p(t, x, \xi)| &\leq C_\gamma, \\ \sup_{|\xi|=1} \sup_{|\beta|=2} |\partial_\xi^\gamma \partial_{t,x}^\beta p(t, x, \xi)| &\leq C_\gamma \alpha(t), \end{aligned}$$

and similarly for p_\pm . As a consequence, the Hamiltonian flow of $\pm p_\pm$,

$$\frac{dx_t}{dt} = \pm d_\xi p_\pm(t, x_t, \xi_t), \quad \frac{d\xi_t}{dt} = \mp d_x p_\pm(t, x_t, \xi_t),$$

is well-posed and induces a bilipschitz homeomorphism on $\mathbb{R}_{x,\xi}^{2n}$, since the Lipschitz norm of p_{\pm} with respect to (x, ξ) is bounded by $\alpha(t) \in L^1((-T, T))$. The null bicharacteristics of $H(t, x, \tau, \xi)$ are, after reparametrization, curves of the form

$$\gamma(t) = (t, x_t, \pm p_{\pm}(t, x_t, \xi_t), \xi_t),$$

where (x_t, ξ_t) is respectively an integral curve of $\pm p_{\pm}$. We will refer to such curves $\gamma(t)$ as the null bicharacteristic curves of L .

The outline of this paper is as follows. In Section 2 we reduce the proof of Theorem 1.1 to an analogous result for a first order pseudodifferential equation, which requires a careful factorization of L . In Section 3 we construct the evolution groups for the first order factors of L as one-parameter families of operators on the appropriate range of Sobolev spaces, through the use of wave packet transform methods. In Section 4 we establish spatial-wavefront mapping properties (pseudolocality) for the evolution operators at fixed time. This is the heart of the paper, where pseudolocality is established via weighted-norm estimates on the fixed-time evolution operators expressed in the wave-packet frame. In Section 5 we deduce the space-time wavefront propagation of Theorem 1.1 from the fixed time result. In Section 6 we show that Theorem 1.1 applies, through a limiting process, to coefficients that satisfy the above regularity assumptions on the elements of a partition of $(-T, T) \times \mathbb{R}^n$ into time slices, with matching assumptions at the endpoints. We then produce an example of such a metric showing that the assumption of H^{s_0} regularity on u cannot be relaxed when establishing propagation of H^{s_0+1} singularities. The first appendix, Section 7, contains the various commutator and paraproduct estimates that are used throughout the paper. Some of these results are standard in paraproduct theory, but we collect them here for reference. The second appendix, Section 8, contains energy estimates and well-posedness results for the operators considered in this paper.

Notation. We use the following notation for function spaces. For $1 \leq p \leq \infty$, and $s \in \mathbb{R}$, $L^p H^s$ denotes functions for which $\|u(t)\|_{H^s(\mathbb{R}^n)}$ belongs to $L^p((-T, T))$, with norm

$$\|u\|_{L^p H^s} = \left(\int_{-T}^T \|u(t)\|_{H^s}^p dt \right)^{\frac{1}{p}},$$

with the obvious modification if $p = \infty$, and where we write L^2 instead of H^0 . Here and throughout this paper, $u(t)$ denotes the function $x \rightarrow u(t, x)$. The $L^p L^q$ norm is similarly defined as $\|u\|_{L^p((-T, T), L^q(\mathbb{R}^n))}$.

The space $C^{k,1}$, for nonnegative integer k , consists of functions whose k -th derivatives satisfy a Lipschitz condition,

$$\|f\|_{C^{0,1}} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|},$$

$$\|f\|_{C^{k,1}} = \sup_{|\alpha| \leq k} \|\partial_x^\alpha f\|_{C^{0,1}}.$$

For k a nonnegative integer, $C^k H^s$ denotes the space of u such that $t \rightarrow u(t)$ is a C^k map of $(-T, T) \rightarrow H^s(\mathbb{R}^n)$, with the norm

$$\|u\|_{C^k H^s} = \sup_{t \in (-T, T)} \sup_{j \leq k} \|\partial_t^j u(t)\|_{H^s}.$$

The notation $\|f\|_{H^s}$ denotes the norm in the Sobolev space $H^s(\mathbb{R}^n)$. In case that we use the norm in $H^s((-T, T) \times \mathbb{R}^n)$ or $H^s(\mathbb{R}^{1+n})$ we write the domain explicitly unless it is obvious from the context; in the first case s will be a nonnegative integer.

For a sequence of functions $f = \{f_k\}_{k=0}^\infty$,

$$\|f\|_{\ell^2 L^2} = \left(\sum_{k=0}^\infty \|f_k\|_{L^2}^2 \right)^{\frac{1}{2}}, \quad \|f\|_{2^{k\sigma} \ell^2 L^2} = \left(\sum_{k=0}^\infty 2^{2k\sigma} \|f_k\|_{L^2}^2 \right)^{\frac{1}{2}}.$$

The space $S^m \subset C^\infty(\mathbb{R}^n)$ denotes smooth symbols satisfying the standard multiplier condition; that is, for all multi-indices α ,

$$|\partial_\xi^\alpha q(\xi)| \leq C_\alpha (1 + |\xi|)^{m-|\alpha|}.$$

The space $S_{cl}^m \subset S^m$ denotes symbols that are homogeneous of degree m on $|\xi| \geq 1$,

$$q(r\xi) = r^m q(\xi), \quad r \geq 1, \quad |\xi| \geq 1.$$

Given two positive functions f and g we say that $f \lesssim g$, respectively $f \approx g$, if there is a constant $C < \infty$ such that

$$f \leq Cg, \quad \text{respectively} \quad C^{-1}g \leq f \leq Cg.$$

2. REDUCTION TO A FIRST ORDER OPERATOR

In this section we reduce Theorem 1.1 to results for a first order pseudodifferential equation, through a factorization of the operator L . We introduce the notation

$$(2.1) \quad P = p(t, x, D), \quad P^\pm = p_\pm(t, x, D) = P \pm b^j D_j,$$

with p and p_\pm defined by (1.8).

Throughout, $D = (D_1, \dots, D_n) = -i\partial_x$, and always $1 \leq i, j \leq n$. The operator $P(t)$, respectively $P^\pm(t)$, will denote the corresponding pseudodifferential operator acting on functions of x , obtained by freezing the t variable.

We start with a factorization of L of the form

$$(2.2) \quad L = (D_t + P^- + d^0)(D_t - P^+) + R_1^+,$$

where $R_1^+(t)$ is a one-parameter family of first-order operators acting on functions of x , with the precise form of $R_1^+(t)$ stated below.

Since $P^\pm = P \pm b^j D_j$, the product of parentheses on the right hand side expands to

$$D_t^2 - 2b^j D_j D_t + b^i b^j D_i D_j + d^0 D_t - P^2 + R,$$

where

$$R = -((D_t b^j) - b^i (D_i b^j)) D_j - [D_t, P] + [b^j D_j, P] - d^0 P^+.$$

Using a symbol expansion of the homogeneous symbol $p(t, x, \xi)$ as in (7.1), we see that R is a convergent sum of terms of the form

$$(2.3) \quad d(t, x) q_0(D) \quad \text{and} \quad a_1(t, x) [a_2(t, x), q_1(D)] q_2(D),$$

where d satisfies (1.6), each a_j satisfies the regularity conditions (1.3) and (1.4), and each $q_j(\xi) \in S_{cl}^1(\mathbb{R}^n)$.

Next, observe that

$$p(t, x, D)^2 = p^2(t, x, D) + R = c^{ij} D_i D_j + b^i b^j D_i D_j + R,$$

with R is again of the form (2.3), as seen by using the symbol expansion (7.1) of p . Thus, (2.2) holds with R_1^+ a convergent sum of terms of the form (2.3). By (1.3) and (1.4), and Theorem 7.1, we have the following bounds for each $t \in (-T, T)$,

$$(2.4) \quad \|R_1^+(t)f\|_{L^2} \leq C \|f\|_{H^1},$$

$$(2.5) \quad \|R_1^+(t)f\|_{H^s} \leq C \alpha(t) \|f\|_{H^{s+1}}, \quad -1 \leq s \leq 1,$$

$$(2.6) \quad \|D_t R_1^+(t)f\|_{L^2} + \|[q_1(D), R_1^+(t)]f\|_{L^2} \leq C \alpha(t) \|f\|_{H^1},$$

whenever $q_1(D)$ is an order 0 multiplier in the x -variable. Additionally, by (2.6) or (1.5), we have the following norm-continuity of $R_1^+(t)$ with respect to t ,

$$(2.7) \quad \|R_1^+(t)f - R_1^+(s)f\|_{L^2} \leq C \left(\int_s^t \alpha(r) dr \right) \|f\|_{H^1}.$$

We now fix $s_0 \in \{0, 1\}$, and produce a factorization of L modulo order 0 terms,

$$(2.8) \quad L = \left(D_t + P^- + d^0 + Q^+ \right) \left(D_t - P^+ - Q^+ \right) + R_0^+$$

where $Q^+ = Q^+(t)$ will be a uniformly bounded family of operators on $H^{s_0}(\mathbb{R}^n)$, depending on the parameter t , and where the form of Q^+ will depend on the choice of $s_0 \in \{0, 1\}$. Here, $R_0^+(t)$ is a one-parameter family of operators on $H^{s_0}(\mathbb{R}^n)$, and we construct $Q^+(t)$ such that

$$(2.9) \quad \|R_0^+(t)f\|_{H^{s_0}} \leq C \alpha(t) \|f\|_{H^{s_0}},$$

and such that

$$(2.10) \quad \begin{aligned} \|Q^+(t)f\|_{H^s} &\leq C \alpha(t) \|f\|_{H^s}, \quad s_0 - 1 \leq s \leq s_0 + 1, \\ \|Q^+(t)f\|_{H^{s_0}} &\leq C \|f\|_{H^{s_0}}, \\ \|Q^+(t)f - Q^+(s)f\|_{H^{s_0}} &\leq C \left(\int_s^t \alpha(r) dr \right) \|f\|_{H^{s_0}}. \end{aligned}$$

In particular, $Q^+(t)$ is a continuous function of t in the H^{s_0} operator norm.

Expanding the product of parentheses in (2.8) leads to

$$L - R_1^+(t) - P(t)Q^+(t) - Q^+(t)P(t) - R(t),$$

where

$$R = [D_t, Q^+] - [b^j, Q^+]D_j - b^j [D_j, Q^+] + (Q^+)^2 + d^0 Q^+.$$

Assuming (2.10), and since $\|d^0(t, \cdot)\|_{C^{0,1}} \leq C \alpha(t)$, the last two terms satisfy the bound in (2.9), hence can be absorbed into the error $R_0^+(t)$. The estimates (2.12) below will imply that the first three terms also satisfy the bound in (2.9).

So, given R_1^+ of the form (2.3), it suffices to construct $Q^+(t)$ solving

$$(2.11) \quad P(t)Q^+(t) + Q^+(t)P(t) - R_1^+(t) = R_0^+(t),$$

with $R_0^+(t)$ satisfying (2.9), and $Q^+(t)$ satisfying the following two conditions:

$$(2.12) \quad \begin{aligned} \|D_t Q^+(t)f\|_{H^{s_0}} + \|[Q^+(t), q(D)]f\|_{H^{s_0}} &\leq C \alpha(t) \|f\|_{H^{s_0}}, \\ \|[b, Q^+(t)]f\|_{H^{s_0}} &\leq C \alpha(t) \|b\|_{C^{0,1}} \|f\|_{H^{s_0-1}}, \end{aligned}$$

where $q(D)$ denotes a general $S_{cl}^1(\mathbb{R}^n)$ multiplier in the x -variable, and $b(x)$ a general Lipschitz function of x . An immediate corollary of (2.12) is that

$$(2.13) \quad \|[Q^+(t), q_0(D)]f\|_{H^{s+1}} \leq C \alpha(t) \|f\|_{H^s}, \quad s_0 - 1 \leq s \leq s_0,$$

whenever $q_0 \in S_{cl}^0(\mathbb{R}^n)$, as is seen by interpolation and writing

$$[Q^+(t), q_0(D)]D = [Q^+(t), q_0(D)D] - q_0(D)[Q^+(t), D],$$

$$D[Q^+(t), q_0(D)] = [Q^+(t), q_0(D)D] + [Q^+(t), D]q_0(D).$$

After adding a harmless constant to $p(t, x, \xi)$, by Lemma 7.10 the operator $P(t)$ is invertible for every t , and with uniform bounds over $t \in (-T, T)$,

$$\|P(t)^{-1}f\|_{H^s} \leq C \|f\|_{H^{s-1}}, \quad 0 \leq s \leq 2.$$

For the case $s_0 = 1$, we define

$$(2.14) \quad Q^+(t) = \frac{1}{2}P(t)^{-1}R_1^+(t).$$

Then (2.11) holds with

$$R_0^+(t) = \frac{1}{2}P(t)^{-1}[R_1^+(t), P(t)].$$

Lemma 2.1 below will show that

$$\|R_0^+(t)f\|_{H^1} \leq C \alpha(t) \|f\|_{H^1}.$$

Thus (2.9) holds with $s_0 = 1$. Furthermore, the operator

$$(D_t Q^+)(t) = \frac{1}{2}P(t)^{-1}(D_t P)(t)P(t)^{-1}R_1^+(t) + \frac{1}{2}P(t)^{-1}(D_t R_1^+)(t)$$

has the same mapping properties by (2.6). This also holds for $[D_j, Q^+(t)]$. Finally, if $b \in C^{0,1}$, then

$$2[b, Q^+(t)] = P(t)^{-1}[b, P(t)]P(t)^{-1}R_1^+(t) + P(t)^{-1}[b, R_1^+(t)].$$

The first term on the right maps L^2 to H^1 with norm $\lesssim \alpha(t)\|b\|_{C^{0,1}}$, which follows by Theorem 7.1 together with (2.5) for $s = -1$. For the second term we apply Lemma 2.1 below, to see that it satisfies similar bounds on H^1 . Thus, (2.12) holds with $s_0 = 1$.

For the case $s_0 = 0$, we set

$$(2.15) \quad Q^+(t) = \frac{1}{2}R_1^+(t)P(t)^{-1}.$$

Then (2.11) holds with

$$R_0^+(t) = \frac{1}{2}[P(t), R_1^+(t)]P(t)^{-1},$$

hence by Lemma 2.1

$$\|R_0^+(t)f\|_{L^2} \leq C \alpha(t) \|f\|_{L^2}.$$

Furthermore, (2.10) and (2.12) hold with $s_0 = 0$, so that the other terms in $R_0^+(t)$ have the same mapping property, hence (2.9) holds with $s_0 = 0$ for this choice of $Q^+(t)$.

Lemma 2.1. *Assuming $R_1^+(t)$ is a convergent sum of terms of the form (2.3), then*

$$\| [P(t), R_1^+(t)] f \|_{L^2} \leq C \alpha(t) \| f \|_{H^1},$$

and for $b \in C^{0,1}(\mathbb{R}^n)$,

$$\| [R_1^+(t), b] f \|_{L^2} \leq C \alpha(t) \| b \|_{C^{0,1}} \| f \|_{L^2}.$$

Proof. In these estimates, the type of terms in $R_1^+(t)$ of the form $d(t, x)q(D)$ lead to commutators that are easily handled, so we replace $R_1^+(t)$ in the statement by an operator of the form $a_1 [a_2, q_1(D)] q_2(D)$.

For the first estimate, we take the symbol expansion (7.1) of $p(t, x, \xi)$, and consider a term of the form

$$\begin{aligned} & [a_0 q_0(D), a_1 [a_2, q_1(D)] q_2(D)] = \\ & a_0 [q_0(D), a_1] [a_2, q_1(D)] q_2(D) + a_0 a_1 [q_0(D), [a_2, q_1(D)]] q_2(D) \\ & + a_1 [a_0, [a_2, q_1(D)]] q_2(D) q_0(D) + a_1 [a_2, q_1(D)] [a_0, q_2(D)] q_0(D), \end{aligned}$$

where each $q_j \in S_{cl}^1(\mathbb{R}^n)$. Each term on the right satisfies the desired bound by Theorem 7.1 and Lemma 7.7.

For the second estimate, we need consider

$$[[a_2, q_1(D)] q_2(D), b] = [a_2, q_1(D)] [q_2(D), b] + [b, [a_2, q_1(D)]] q_2(D),$$

which is handled similarly. \square

The same calculation also constructs one-parameter families of operators $R_1^-(t)$, $Q^-(t)$, and $R_0^-(t)$ satisfying the above conditions, such that

$$(2.16) \quad L = (D_t - P^+ + d^0)(D_t + P^-) + R_1^-,$$

and

$$(2.17) \quad L = (D_t - P^+ + d^0 - Q^-)(D_t + P^- + Q^-) + R_0^-.$$

Suppose now that we are given s_0 from Theorem 1.1, and construct the corresponding $Q^\pm(t)$ as above. In the next section we construct evolution groups $E_\pm(t, t_0)$, for $t, t_0 \in (-T, T)$, satisfying

$$(2.18) \quad D_t E_\pm(t, t_0) = \pm (P^\pm(t) + Q^\pm(t)) E_\pm(t, t_0), \quad E_\pm(t_0, t_0) = I.$$

Precisely, $E_\pm(t, t_0)$ is a bounded family of maps on $H^s(\mathbb{R}^n)$ for $s_0 \leq s \leq s_0 + 1$, which is strongly continuous in t and t_0 , such that if $f \in H^s(\mathbb{R}^n)$ and $t_0 \in (-T, T)$ then

$$E_\pm(t, t_0) f \in C^0 H^s \cap C^1 H^{s-1}, \quad s_0 \leq s \leq s_0 + 1,$$

and such that

$$D_t E_\pm(t, t_0) f = \pm (P^\pm(t) + Q^\pm(t)) E_\pm(t, t_0) f, \quad E(t_0, t_0) f = f.$$

Then the above factorizations show that

$$L E_\pm(t, t_0) f = R_0^\pm(t) E_\pm(t, t_0) f.$$

Given the E_\pm , and a $u \in C^0 H^{s_0} \cap C^1 H^{s_0-1}$ that solves the Cauchy problem

$$Lu = 0, \quad u(t_0) = u_0, \quad D_t u(t_0) = u_1, \quad (u_0, u_1) \in H^{s_0} \times H^{s_0-1}$$

for some given $t_0 \in (-T, T)$, we can write u in the form

$$u = v + \sum_{\pm} E_{\pm}(t, t_0) f_{\pm}, \quad \text{where } f_{\pm} \in H^{s_0}, \quad \text{WF}_{s_0+1}(v) \cap \text{char}(L) = \emptyset.$$

To see this, we impose the conditions

$$f_+ + f_- = u_0, \quad (P^+(t_0) + Q^+(t_0))f_+ - (P^-(t_0) + Q^-(t_0))f_- = u_1,$$

which is solved by

$$f_{\pm} = (2P(t_0) + Q^-(t_0) + Q^+(t_0))^{-1} (P^{\mp}(t_0)u_0 + Q^{\mp}(t_0)u_0 \pm u_1) \in H^{s_0}(\mathbb{R}^n),$$

where the inverse exists by Lemma 7.10, after adding a harmless constant to p . We then write

$$L\left(u - \sum_{\pm} E_{\pm} f_{\pm}\right) = - \sum_{\pm} R_0^{\pm} E_{\pm} f_{\pm} \in L^1 H^{s_0}.$$

Also,

$$\left(u - \sum_{\pm} E_{\pm} f_{\pm}\right)\Big|_{t=t_0} = 0, \quad D_t\left(u - \sum_{\pm} E_{\pm} f_{\pm}\right)\Big|_{t=t_0} = 0.$$

Thus, by Theorem 8.6

$$v = u - \sum_{\pm} E_{\pm} f_{\pm} \in C^0 H^{s_0+1} \cap C^1 H^{s_0},$$

and in particular $\text{WF}_{s_0+1}(v) \cap \text{char}(L) = \emptyset$, since $\text{char}(L) \subset \{\xi \neq 0\}$.

We thus can reduce Theorem 1.1 to a result about the functions $E_{\pm}(t, t_0) f_{\pm}$. Suppose for simplicity that the γ in the statement of Theorem 1.1 is contained in the forward cone $\tau = p_+(t, x, \xi)$. By Theorem 8.4 of the appendix,

$$\gamma \cap \text{WF}_{s_0+1}(E_-(t, t_0) f_-) = \emptyset.$$

Since $\gamma \cap \text{WF}_{s_0+1}(v) = \emptyset$, Theorem 1.1 is reduced to the following.

Theorem 2.2. *Suppose that $s_0 \in \{0, 1\}$, and that $E_+(t, t_0)$ is the wave group constructed in Section 3 for*

$$D_t E_+(t, t_0) = (P^+(t) + Q^+(t)) E_+(t, t_0), \quad E(t_0, t_0) = I,$$

where $Q^+(t)$ is given by (2.14) or (2.15) respectively if $s_0 = 1$ or $s_0 = 0$.

Let $\gamma(t) = (t, x_t, p_+(t, x_t, \xi_t), \xi_t)$ be a null bicharacteristic curve of L . If $f \in H^{s_0}$, and for some $t_0 \in (-T, T)$ we have $\gamma(t_0) \notin \text{WF}_{s_0+1}(E_+(t, t_0) f)$, it follows that

$$\gamma \cap \text{WF}_{s_0+1}(E_+(t, t_0) f) = \emptyset.$$

The analogous result holds for the wave group $E_-(t, t_0)$.

3. THE WAVE PACKET TRANSFORM AND CONSTRUCTION OF THE WAVE GROUP

In this section we construct the wave groups $E_{\pm}(t, t_0)$. For simplicity we drop the superscripts $+$ and $-$, and let $P(t)$ be either $P^{\pm}(t)$. Given $s_0 \in \{0, 1\}$, we let $Q(t)$ denote either $Q^{\pm}(t)$, given by (2.14) or (2.15) respectively if $s_0 = 1$ or $s_0 = 0$, where $R_1(t)$ is a convergent sum of expressions of the form (2.3). There is a minor inconsistency in that the $P(t)$ in (2.14) and (2.15) refers to the original $p(t, x, D)$ as in (2.1), but this is unimportant as all three symbols p and p_{\pm} have the same regularity.

We construct $E(t, t_0) : H^{s_0} \rightarrow C^0 H^{s_0}$, such that

$$D_t E(t, t_0) f = (P(t) + Q(t)) E(t, t_0) f, \quad E(t_0, t_0) f = f, \quad f \in H^{s_0}.$$

By Theorem 8.5 the evolution group $E(t, t_0)$ is uniquely determined, although in the proof of Theorem 1.1 the existence of $E(t, t_0)$ with the desired properties is all that is used. Our construction will show that $E(t, t_0)$ is also uniformly bounded on H^s for $s_0 \leq s \leq s_0 + 1$, and is strongly continuous in both t and t_0 on each such H^s . It follows from (2.10) that if $f \in H^s$ for some $s_0 \leq s \leq s_0 + 1$, then

$$Q(t) E(t, t_0) f \in C^0 H^{s_0} \subset C^0 H^{s-1}.$$

The same holds for $P(t) E(t, t_0) f$. Thus, for $s_0 \leq s \leq s_0 + 1$,

$$(3.1) \quad E(t, t_0) f \in C^0 H^s \cap C^1 H^{s-1}, \quad \text{if } f \in H^s.$$

Since the proof below works equally well if $Q \equiv 0$, it will also construct the evolution groups $E_{0,\pm}(t, t_0)$ for the equation

$$(3.2) \quad D_t E_{0,\pm}(t, t_0) f = \pm P^\pm(t) E_{0,\pm}(t, t_0) f, \quad E_{0,\pm}(t_0, t_0) f = f,$$

and (3.1) holds for $f \in H^s$ if $0 \leq s \leq 2$.

Following [12], we work with a scaled wave-packet transform, similar to the FBI transform used in Tataru [14], but based on a Schwartz function with Fourier transform of compact support instead of a Gaussian.

We fix a real, even Schwartz function $h(x) \in \mathcal{S}(\mathbb{R}^n)$, with $\|h\|_{L^2} = (2\pi)^{-\frac{n}{2}}$, and assume that its Fourier transform $\widehat{h}(\xi)$ is supported in the unit ball $\{|\xi| \leq 1\}$. For $k \geq 0$, we define $T_k : \mathcal{S}'(\mathbb{R}^n) \rightarrow C^\infty(\mathbb{R}^{2n})$ by the rule

$$(T_k g)(x, \xi) = 2^{\frac{nk}{4}} \int e^{-i\langle \xi, z-x \rangle} h(2^{\frac{k}{2}}(z-x)) g(z) dz.$$

A simple calculation shows that

$$g(y) = 2^{\frac{nk}{4}} \int e^{i\langle \xi, y-x \rangle} h(2^{\frac{k}{2}}(y-x)) (T_k g)(x, \xi) dx d\xi,$$

so that $T_k^* T_k = I$. In particular,

$$\|T_k g\|_{L^2(\mathbb{R}^{2n})} = \|g\|_{L^2}.$$

The following, which is Lemma 3.1 of [12], shows that the L^2 -continuity of T_k holds under relaxed conditions.

Lemma 3.1. *Suppose that $h_{x,\xi}(z)$ is a family of Schwartz functions on \mathbb{R}^n , depending on the parameters x and ξ , with uniform bounds over x and ξ on each Schwartz semi-norm of h . Then the operator*

$$(R_k g)(x, \xi) = 2^{\frac{nk}{4}} \int e^{-i\langle \xi, z-x \rangle} h_{x,\xi}(2^{\frac{k}{2}}(z-x)) g(z) dz$$

satisfies the bound

$$\|R_k g\|_{L^2(\mathbb{R}^{2n})} \leq C \|g\|_{L^2}.$$

We will apply T_k to the localization of u at frequency k . We introduce a nonnegative function $\beta(s) \in C_c^\infty(\mathbb{R})$, supported in the interval $[2^{-\delta}, 2^{1+\delta}]$, where $\delta > 0$ will

be taken sufficiently small. With $\beta_k(\xi) = \beta(2^{-k}|\xi|)$ if $k \geq 1$, and β_0 an appropriate compactly supported function on \mathbb{R}^n , we assume that

$$(3.3) \quad \sum_{k=0}^{\infty} \beta_k(\xi)^2 = 1.$$

Now define $T : L^2(\mathbb{R}^n) \rightarrow \ell^2(\mathbb{N}, L^2(\mathbb{R}^{2n}))$ by

$$Tg \equiv \tilde{g} \equiv \{\tilde{g}_k\}_{k=0}^{\infty}, \quad \tilde{g}_k = T_k \beta_k(D)g.$$

Then T is a norm isomorphism, hence $T^*T = I$. Furthermore, for k large enough so that $2^{-k/2} \leq 2^{-\delta}(1 - 2^{-\delta})$, \tilde{g}_k is supported in the set $\{2^{k-2\delta} \leq |\xi| \leq 2^{k+1+2\delta}\}$. It follows that, for $\sigma \in \mathbb{R}$,

$$(3.4) \quad \|g\|_{H^\sigma} \approx \left(\sum_{k=0}^{\infty} 2^{2k\sigma} \|\tilde{g}_k\|_{L^2(\mathbb{R}^{2n})}^2 \right)^{\frac{1}{2}}.$$

We will obtain $E(t, t_0)$ by constructing its lift $\tilde{E}(t, t_0)$ to $\ell^2(\mathbb{N}, L^2(\mathbb{R}^{2n}))$ via the wave packet transform T ,

$$\tilde{E}(t, t_0)f = TE(t, t_0)T^*f.$$

The group $\tilde{E}(t, t_0)$ will be constructed in a manner similar to that used in [12], approximating the lifted equation by the Hamiltonian flow of an appropriately mollified p and obtaining $\tilde{E}(t, t_0)$ by a convergent iteration from the Hamiltonian flow group.

For $k \in \mathbb{N}$, we introduce the spatial regularization of p ,

$$p_k(t, x, \xi) = \phi(2^{-\frac{k}{2}}D)p(t, x, \xi),$$

which regularizes the symbol in x to frequencies of magnitude $\leq c2^{\frac{k}{2}}$, some small fixed $c > 0$. We let $P_k(t) = p_k(t, x, D)$. We remark that in [12] the symbol regularization was over both t and x variables, but that is unimportant for the results from [12] that we use in this paper, specifically [12, Lemma 3.2] and [12, Lemma 3.3].

Let $V_k = V_k(t, x, \xi, \partial_x, \partial_\xi)$ denote the real, linear first-order differential operator

$$V_k f = d_\xi p_k(t, x, \xi) \cdot d_x f - d_x p_k(t, x, \xi) \cdot d_\xi f,$$

This vector field is Lipschitz regular in (x, ξ) provided $|\xi|$ is bounded above, with Lipschitz constant $\alpha(t) \in L^1((-T, T))$. Hence the associated flow group is well-posed.

Let $\Theta_{s,t}^k$ denote the associated time $t \rightarrow s$ flow map on \mathbb{R}^{2n} ,

$$\partial_t f(\Theta_{s,t}^k(x, \xi)) = V_k f(\Theta_{s,t}^k(x, \xi)),$$

which is the Hamiltonian flow induced by p_k . Also let $\Theta_{s,t}$ denote the $t \rightarrow s$ Hamiltonian flow map for p . By a simple extension of [11, Lemma 3.6], if (x_t, ξ_t) is the flowout of (x_0, ξ_0) through p , and (x_t^k, ξ_t^k) is the flowout of (x_0, ξ_0) through p_k , and $|\xi_0| \approx 1$, then

$$|x_t^k - x_t| + |\xi_t^k - \xi_t| \lesssim 2^{-\frac{k}{2}}.$$

Also $\Theta_{s,t}$, and each $\Theta_{s,t}^k$, are bi-Lipschitz, measure preserving maps on \mathbb{R}^{2n} , homogeneous of degree 1 in ξ , and by homogeneity it holds that $|\xi_t| \approx |\xi_0|$, similarly $|\xi_t^k| \approx |\xi_0|$.

We define a unitary evolution group $W(t, s)$ on $\ell^2(\mathbb{N}, L^2(\mathbb{R}^{2n}))$ by evolving each f_k along V_k . Thus, for $f = \{f_k(x, \xi)\}_{k=0}^\infty \in \ell^2(\mathbb{N}, L^2(\mathbb{R}^{2n}))$, we set

$$(W(t, s)f)_k = f_k \circ \Theta_{s,t}^k.$$

Suppose that $\tilde{u}(t) = T(u(t))$. Then the equation $D_t u - P(t)u = Q(t)u$ is equivalent to the collection of equations for $k \in \mathbb{N}$,

$$(3.5) \quad -i(\partial_t - V_k)\tilde{u}_k = (T_k P_k + iV_k T_k)\beta_k(D)u \\ + T_k[\beta_k(D), P_k]u + T_k\beta_k(D)(P - P_k)u + T_k\beta_k(D)Qu.$$

Inserting $u = T^*\tilde{u}$, we can write this as a series of equations

$$(3.6) \quad (\partial_t - V_k)\tilde{u}_k(t) = (B(t)\tilde{u}(t))_k,$$

where $(B\tilde{u})_k$ is the right hand side of (3.5) applied to $u = T^*\tilde{u}$. Note that $(B\tilde{u})_k$ is supported where $|\xi| \in [2^{k-3\delta}, 2^{k+1+3\delta}]$, by the frequency localization of P_k and T_k .

We will show that

$$(3.7) \quad \|B(t)f\|_{2^{k\sigma}\ell^2L^2} \leq C\alpha(t)\|f\|_{2^{k\sigma}\ell^2L^2}, \quad s_0 - 1 \leq \sigma \leq s_0 + 1,$$

where the norm denotes the one on the right hand side in (3.4). We can then obtain the solution to (3.6) with given initial condition $\tilde{u}(t_0)$ by solving the integral equation

$$(3.8) \quad \tilde{u}(t) = W(t, t_0)\tilde{u}(t_0) + \int_{t_0}^t W(t, s)B(s)\tilde{u}(s) ds.$$

Indeed, for $u(t_0) \in H^\sigma(\mathbb{R}^n)$, $s_0 - 1 \leq \sigma \leq s_0 + 1$, the integral equation (3.8) admits a series solution $\tilde{u} = \sum_{j=0}^\infty \tilde{u}^{(j)}$, convergent in $C^0((-T, T), 2^{k\sigma}\ell^2L^2)$, where

$$\tilde{u}^{(0)}(t) = W(t, t_0)\tilde{u}(t_0), \quad \tilde{u}^{(j+1)}(t) = \int_{t_0}^t W(t, s)B(s)\tilde{u}^{(j)}(s) ds.$$

We express the solution as $\tilde{u}(t) = \tilde{E}(t, t_0)u(t_0)$, which by uniqueness determines an evolution group $\tilde{E}(t, t_0)$. Note that each $\tilde{u}_k^{(j+1)}$ is supported where $C^{-1}2^k \leq |\xi| \leq C2^k$, for some fixed C , by the localization of $(B\tilde{u})_k$ and homogeneity of $W(t, s)$.

It is easily seen from its construction that the group is strongly continuous in the $2^{k\sigma}\ell^2L^2$ norm, as a function of the parameters $(t, t_0) \in (-T, T)^2$. Since (3.6) is obtained by lifting the equation $D_t u - P(t)u = Q(t)u$, it follows that $\tilde{E}(t, t_0)$ preserves the range of T , and thus is of the form $TE(t, t_0)T^*$, where $E(t, t_0) = T^*\tilde{E}(t, t_0)T$ is consequently strongly continuous on H^σ in both parameters. It follows from (3.8) that $E(t, t_0)u(t_0)$ is a distribution solution of the equation $D_t u - P(t)u = Q(t)u$, which as noted before belongs to $C^0H^\sigma \cap C^1H^{\sigma-1}$ provided that $s_0 \leq \sigma \leq s_0 + 1$.

It remains to establish (3.7). Let $B_{kj}(t)$ denote the kj component of $B(t)$, so $(B\tilde{u})_k(t) = \sum_j B_{kj}(t)\tilde{u}_j(t)$. By the above, B_{kj} is the sum of 4 terms,

$$B_{kj} = (T_k P_k + iV_k T_k)\beta_k(D)\beta_j(D)T_j^* + T_k[\beta_k(D), P_k]\beta_j(D)T_j^* \\ + T_k\beta_k(D)(P - P_k)\beta_j(D)T_j^* + T_k\beta_k(D)Q\beta_j(D)T_j^* \\ \equiv {}_1B_{kj} + {}_2B_{kj} + {}_3B_{kj} + {}_4B_{kj}.$$

The bounds in (3.7) are satisfied by the operator ${}_4B = TQ(t)T^*$ by (3.4) and (2.10), so we focus on the first three components of $B(t)$. The terms ${}_1B_{kj}$ and ${}_2B_{kj}$ vanish unless $|j-k| \leq 1$. Thus, it suffices to prove that each is bounded on $L^2(\mathbb{R}^{2n})$ with norm $\lesssim \alpha(t)$, uniformly over j and k . For ${}_2B_{jk}$ this follows by Theorem 7.1 (or indeed the $S_{1, \frac{1}{2}}$ pseudodifferential calculus). For ${}_1B_{jk}$, it follows by Lemmas 3.1 and 3.2 of [12]. In the next section we will prove even stronger estimates for these terms.

To handle the term ${}_3B$, we take the symbol expansion (7.1) of $p(t, x, \xi)$ to reduce matters to considering $p(t, x, D) = a(t, x)q(D)$. For $|j-k| \leq 1$, uniform boundedness of ${}_3B_{jk}$ follows since $\|a - a_k\|_{L^\infty} \lesssim 2^{-k}\alpha(t)$.

If $|j-k| \geq 2$, then after this substitution

$${}_3B_{kj} = T_k \beta_k(D) a(t, x) q(D) \beta_j(D) T_j^*, \quad |j-k| \geq 2.$$

These off-diagonal terms give an operator which is in fact smoothing of order 1, as we now show.

Set $P(t) = a(t, x)q(D)$. If $2 \leq |j-k| \leq 3$, then since $\|a(t, \cdot)\|_{C^{1,1}} \leq \alpha(t)$, and $q(D)$ is of order 1, we have

$$(3.9) \quad \|T_k \beta_k(D) a(t, x) q(D) \beta_j(D) T_j^*\|_{L^2 \rightarrow L^2} \lesssim C \alpha(t) 2^{-k}.$$

If $|j-k| \geq 4$, then using the Littlewood-Paley partition of unity given by $\psi_j = \beta_j^2$,

$${}_3B_{kj} = \sum_{|l-m| \geq 2} T_k \beta_k(D) \psi_l(D) a(t, x) q(D) \psi_m(D) \beta_j(D) T_j^*,$$

and hence

$$TR_{a(t)}q(D)T^* - \sum_{|j-k| \geq 4} {}_3B_{kj}(t) = \sum_{\substack{|j-k| \leq 3 \\ |l-m| \geq 2}} T_k \beta_k(D) \psi_l(D) a(t, x) q(D) \psi_m(D) \beta_j(D) T_j^*$$

where R_a is defined as in Lemma 7.4. In the latter sum, j, k, l, m differ by at most 5, and each term satisfies the bound in (3.9). Combined with Lemma 7.4, we see that

$$(3.10) \quad \left\| \sum_{|j-k| \geq 2} {}_3B_{kj}(t) f_j \right\|_{2^{k(\sigma+1)} \ell^2 L^2} \leq C \alpha(t) \|f\|_{2^{k\sigma} \ell^2 L^2}, \quad -1 \leq \sigma \leq 1.$$

As a consequence, the operator ${}_3B$ satisfies the bound in (3.7) on the range $-1 \leq s \leq 2$, which contains $s_0 - 1 \leq s \leq s_0 + 1$ for $s_0 \in \{0, 1\}$. This concludes the proof of (3.7), and hence the existence of $E(t, t_0)$. If $Q \equiv 0$, then (3.7) holds on the union of the ranges, $-1 \leq \sigma \leq 2$, hence the wave group $E_0(t, t_0)$ exists on the range $-1 \leq s \leq 2$.

We summarize the results of this section.

Theorem 3.2. *Suppose that $s_0 \in \{0, 1\}$, and that $Q^\pm(t)$ is respectively given by (2.15) or (2.14). Then an evolution group $E_\pm(t, t_0)$ for equation (2.18) exists as a family of bounded maps on $H^s(\mathbb{R}^n)$ for $s_0 - 1 \leq s \leq s_0 + 1$, and is strongly continuous in both t and t_0 . Additionally, for $s_0 \leq s \leq s_0 + 1$,*

$$E_\pm(t, t_0) f \in C^0 H^s \cap C^1 H^{s-1} \quad \text{when } f \in H^s.$$

The evolution group $E_{0, \pm}(t, t_0)$ for the equation

$$D_t E_{0, \pm}(t, t_0) = \pm P^\pm(t) E_{0, \pm}(t, t_0), \quad E_{0, \pm}(t_0, t_0) = I,$$

similarly exists, is strongly continuous in both variables on H^s for $-1 \leq s \leq 2$, and the following holds if $0 \leq s \leq 2$,

$$E_{0,\pm}(t, t_0)f \in C^0 H^s \cap C^1 H^{s-1} \quad \text{when } f \in H^s.$$

4. WEIGHTED ESTIMATES FOR THE WAVE GROUP

The null bicharacteristics of $\tau \mp p_{\pm}(t, x, \xi)$ are in one-one correspondence with the Hamiltonian curves (x_t, ξ_t) for $\pm p_{\pm}(t, x, \xi)$. In this section we prove the following result about wavefront mapping properties on \mathbb{R}^n for the fixed time wave groups $E_{\pm}(t, t_0)$ constructed in the previous section, and in Section 5 derive Theorem 2.2 as a corollary.

Theorem 4.1. *Given $s_0 \in \{0, 1\}$, let $E_{\pm}(t, t_0)$ be the wave group constructed in Section 3. Let (x_t, ξ_t) be the Hamiltonian curve of the corresponding $\pm p_{\pm}(t, x, \xi)$ that passes through (x_0, ξ_0) at $t = t_0$.*

Then, given $f \in H^{s_0}$, if $(x_0, \xi_0) \notin \text{WF}_{s_0+1}(f)$ it follows that

$$(x_t, \xi_t) \notin \text{WF}_{s_0+1}(E_{\pm}(t, t_0)f), \quad t \in (-T, T).$$

Furthermore, if $T < \infty$ there is a constant $c > 0$ such that if $\chi_t(x)$ is a $C_c^\infty(\mathbb{R}^n)$ -bounded family of cutoffs, respectively supported in the ball of radius c about x_t , and $\Gamma_t(\xi)$ is a $S^0(\mathbb{R}^n)$ -bounded family of conic cutoffs, respectively supported in the cone of angle c about ξ_t , then with uniform bounds over $t \in (-T, T)$

$$\Gamma_t(D)\chi_t(x)(E_{\pm}(t, t_0)f) \in H^{s_0+1}.$$

We consider the case of $E_+(t, t_0)$, and denote the wavegroup simply by $E(t, t_0)$. We prove Theorem 4.1 through weighted-norm estimates on the lifted evolution group $\tilde{E}(t, s) = TE(t, s)T^*$, where the weights are time-dependent functions of (x, ξ) . It suffices to consider the case $t \geq t_0$ in Theorem 4.1, which we will assume in the rest of this section. Also, by making a smooth, t -dependent change of variables in x , we will from now on assume that ξ_t remains within a small cone about the positive ξ_1 axis.

Suppose that $M(t, x, \xi)$ is a family of strictly positive functions on $(-T, T) \times \mathbb{R}^{2n}$, continuous in all parameters, such that for some $C < \infty$

$$C^{-1}\langle \xi \rangle^{s_0} \leq M(t, x, \xi) \leq C \langle \xi \rangle^{s_0+1}.$$

Assume that the following holds, where $B(t)$ and $W(t, s)$ are as in Section 3,

$$(4.1) \quad \|M(s, x, \xi)B(s)f\|_{\ell^2 L^2} \leq C \alpha(s) \|M(s, x, \xi)f\|_{\ell^2 L^2},$$

and, in addition, for $t_0 \leq s \leq t \leq T$ assume that

$$(4.2) \quad \|M(t, x, \xi)W(t, s)f\|_{\ell^2 L^2} \leq C \|M(s, x, \xi)f\|_{\ell^2 L^2}.$$

It follows from (3.8) that

$$\begin{aligned} \|M(t, x, \xi)\tilde{u}^{(j+1)}(t)\|_{\ell^2 L^2} &\leq C \|M(t_0, x, \xi)\tilde{u}(t_0)\|_{\ell^2 L^2} \\ &\quad + C \int_0^t \alpha(s) \|M(s, x, \xi)\tilde{u}^{(j)}(s)\|_{\ell^2 L^2} ds. \end{aligned}$$

Since $\alpha \in L^1((-T, T))$, the sum of the $\tilde{u}^{(j)}$ converges to \tilde{u} in the weighted norms, and we conclude that

(4.3)

$$\sup_{t \in (-T, T)} \|M(t, x, \xi) \tilde{u}(t)\|_{\ell^2 L^2} \leq C \exp\left(C \int_{t_0}^t \alpha(s) ds\right) \|M(t_0, x, \xi) \tilde{u}(t_0)\|_{\ell^2 L^2}.$$

With data $u(t_0) \in H^{s_0}(\mathbb{R}^n)$, we thus need to construct $M(t, x, \xi)$ such that the right hand side is finite if $(x_0, \xi_0) \notin WF_{s_0+1}(u(t_0))$, and such that finiteness of the left hand side implies $(x_t, \xi_t) \notin WF_{s_0+1}(u(t))$. The weight $M(t, x, \xi)$ we construct will be of size $\langle \xi \rangle^{s_0+1}$ on some locally uniform conic set about (x_t, ξ_t) , so the statement about uniformity of the neighborhoods in Theorem 4.1 will be a consequence of the following arguments, and we thus focus on the fixed time estimates.

We start by equating weighted $L^2(\mathbb{R}^{2n})$ estimates on $(Tg)(x, \xi)$ to multiplier estimates on $g(x)$.

Lemma 4.2. *Suppose that $w(\xi) \in C(\mathbb{R}^n)$ is a strictly positive function for which there is a constant $m < \infty$ such that, if $k \geq 0$ and $2^{k-1} \leq |\xi|, |\eta| \leq 2^{k+2}$, the following holds*

$$w(\eta) \leq C w(\xi) (1 + 2^{-\frac{k}{2}} |\xi - \eta|)^m.$$

Assume also that $C^{-1} \langle \xi \rangle^{-N} \leq w(\xi) \leq C \langle \xi \rangle^N$ for some N and $C < \infty$. If $g \in H^s(\mathbb{R}^n)$ for some $s \in \mathbb{R}$, then

$$\|w(D)^{\pm 1} g\|_{L^2} \approx \|w(\xi)^{\pm 1} Tg\|_{\ell^2 L^2},$$

and consequently

$$\|w(D)T^*f\|_{L^2} \lesssim \|w(\xi)f\|_{\ell^2 L^2}.$$

Proof. If $\chi \in C_c^\infty$ is supported in the cube $[-.6, .6]^n$, such that $\sum_{j \in \mathbb{Z}^n} \chi(\xi - j) = 1$, then on the set $2^{k-1} \leq |\xi| \leq 2^{k+2}$

$$w(\xi) \approx \sum_{j \in \mathbb{Z}^n} w(2^{\frac{k}{2}} j) \chi(2^{-\frac{k}{2}} \xi - j).$$

If we replace w by the right hand side, then for $|\xi|, |\eta| \approx 2^k$,

$$(4.4) \quad |\partial_\eta^\alpha w(\eta)^{\pm 1}| \leq C_\alpha 2^{-k \frac{|\alpha|}{2}} w(\xi)^{\pm 1} (1 + 2^{-\frac{k}{2}} |\xi - \eta|)^m.$$

Smoothing out w in this way on each component of w with respect to a Littlewood-Paley decomposition, we may assume that (4.4) is satisfied whenever $2^{k-1} \leq |\xi|, |\eta| \leq 2^{k+2}$.

Since the conditions on w are symmetric in w and w^{-1} , it suffices to show that

$$\|w(\xi)T w(D)^{-1} g\|_{\ell^2 L^2}^2 \lesssim \|g\|_{L^2}^2,$$

as writing $g = T^*Tg$ and using the adjoint bound with w replaced by w^{-1} implies the reverse inequality. Let $g_k = \beta_k(D)g$, and write

$$\begin{aligned} w(\xi)T_k w(D)^{-1} g_k &= 2^{-\frac{nk}{4}} \int e^{i\langle \zeta, x \rangle} w(\xi) w(\zeta)^{-1} \widehat{h}(2^{-\frac{k}{2}}(\zeta - \xi)) \widehat{g}_k(\zeta) d\zeta \\ &= 2^{-\frac{nk}{4}} \int e^{i\langle \zeta, x \rangle} \widehat{h}_\xi(2^{-\frac{k}{2}}(\zeta - \xi)) \widehat{g}_k(\zeta) d\zeta, \end{aligned}$$

where

$$\widehat{h}_\xi(\eta) = w(\xi) w(\xi - 2^{\frac{k}{2}} \eta)^{-1} \widehat{h}(\eta).$$

Here, $|\xi| \approx 2^k$ and $|\eta| \leq 1$, so by (4.4) it follows that the function $h_\xi(z)$ is a smooth function of z , with Schwartz seminorms bounded uniformly over ξ . By Lemma 3.1,

$$\|w(\xi)T_k w(D)^{-1}g_k\|_{L^2} \lesssim \|g_k\|_{L^2},$$

and the result follows. \square

For weights in x the analogue is the following result, which holds by a similar proof.

Lemma 4.3. *Suppose that $w_k(x) \in C^0(\mathbb{R}^n)$ is a strictly positive function, such that for some $m < \infty$ the following holds*

$$w_k(x) \leq C w_k(y) (1 + 2^{\frac{k}{2}}|x - y|)^m.$$

Then

$$\|w_k(x)^{\pm 1}T_k g\|_{L^2(\mathbb{R}^{2n})} \approx \|w_k(x)^{\pm 1}g\|_{L^2},$$

and consequently

$$\|w_k(x)T_k^* f\|_{L^2} \lesssim \|w_k(x)f\|_{L^2(\mathbb{R}^{2n})},$$

Furthermore, the constants in the bounds are independent of k .

We can now use weighted estimates to characterize the H^σ -wavefront set of g .

Lemma 4.4. *Suppose $g \in H^s(\mathbb{R}^n)$ for some s . Then $(x_0, \xi_0) \notin \text{WF}_\sigma(g)$ if and only if there exists an open ball Ω centered on x_0 , and an open conic set $\Gamma \subset \mathbb{R}^n$ centered on ξ_0 , such that*

$$(4.5) \quad \sum_{k=0}^{\infty} \int_{\Omega \times \Gamma} \langle \xi \rangle^{2\sigma} |\tilde{g}_k(x, \xi)|^2 dx d\xi < \infty.$$

Proof. Suppose that $(x_0, \xi_0) \notin \text{WF}_\sigma(g)$. For $\chi(x) \in C_c^\infty(\mathbb{R}^n)$, and $q(\xi) \in S_{cl}^\sigma$ real and homogeneous of degree σ for $|\xi| \geq 1$, we consider

$$\int \chi(x)q(\xi)^2 |\tilde{g}_k(x, \xi)|^2 dx d\xi = \int e^{i\langle x, \eta - \zeta \rangle} \chi(x) b_k(\zeta, \eta) \widehat{g}(\eta) \overline{\widehat{g}(\zeta)} d\eta dx d\zeta,$$

where

$$b_k(\zeta, \eta) = 2^{-\frac{n_k}{2}} \int q(\xi)^2 \widehat{h}(2^{-\frac{k}{2}}(\eta - \xi)) \widehat{h}(2^{-\frac{k}{2}}(\zeta - \xi)) \beta_k(\eta) \beta_k(\zeta) d\xi.$$

Since \widehat{h} is supported in the unit ball, $b_k(\zeta, \eta)$ vanishes unless

$$2^{k-\delta} \leq |\eta| \leq 2^{k+1+\delta}, \quad \text{dist}(\eta, \text{supp}(q)) \leq 2^{-\frac{k}{2}},$$

and the same condition holds with η replaced by ζ . In particular, if Γ' is an open cone containing the support of q , then $b_k(\zeta, \eta)$ is supported in $\Gamma' \times \Gamma'$ for k sufficiently large. Additionally, a simple calculation shows that

$$|\partial_\zeta^\alpha \partial_\eta^\beta b_k(\zeta, \eta)| \leq C_{\alpha, \beta} 2^{2k\sigma - \frac{k}{2}(|\alpha| + |\beta|)}.$$

Hence, the compound symbol $a(\zeta, x, \eta) = \chi(x) \sum_{k=0}^{\infty} b_k(\zeta, \eta)$ is of type $S_{\frac{1}{2}, \frac{1}{2}, 0}^{\sigma, \sigma}$. If the support of $\chi(x)q(\xi)$ is contained in a small conic neighborhood of (x_0, ξ_0) , then standard pseudodifferential calculus arguments show that

$$\int \overline{g(x)} a(D, x, D) g(x) dx < \infty.$$

The bound (4.5) follows by taking a sufficiently small conic neighborhood $\Omega \times \Gamma$ of (x_0, ξ_0) with $\chi(x)q(\xi)$ equal to one on $\Omega \times \Gamma$.

Conversely, suppose (4.5) holds and $g \in H^s(\mathbb{R}^n)$. Let $q(\xi) \in S_{cl}^\sigma$, and write

$$\chi(y)(q(D)g)(y) = \sum_{k=0}^{\infty} 2^{-\frac{nk}{4}} e^{i\langle y-x, \eta \rangle} \chi(y) q(\eta) \beta_k(\eta) \widehat{h}(2^{-\frac{k}{2}}(\eta - \xi)) \tilde{g}_k(x, \xi) dx d\xi d\eta.$$

Let

$$\begin{aligned} K_{j,k}(x', \xi'; x, \xi) &= 2^{-\frac{n(j+k)}{4}} \int e^{i\langle y-x, \eta \rangle - i\langle y-x', \zeta \rangle} \chi^2(y) \\ &\quad \times q(\zeta) \beta_j(\zeta) \widehat{h}(2^{-\frac{j}{2}}(\zeta - \xi')) q(\eta) \beta_k(\eta) \widehat{h}(2^{-\frac{k}{2}}(\eta - \xi)) d\eta dy d\zeta. \end{aligned}$$

Then $K_{j,k}$ vanishes unless ξ and ξ' both lie in a small conic neighborhood of the support of q , and $|\xi| \approx 2^k$, $|\xi'| \approx 2^j$. Additionally, for all N ,

$$\begin{aligned} |K_{j,k}(x', \xi'; x, \xi)| &\leq C_N 2^{\sigma(j+k)} 2^{-N|j-k|} (1+2^{\frac{\min(j,k)}{2}}|x-x'|)^{-N} (1+2^{-\frac{\max(j,k)}{2}}|\xi-\xi'|)^{-N} \\ &\quad \times (1+2^{\frac{k}{2}}\text{dist}(x, \text{supp}(\chi)))^{-N} (1+2^{\frac{j}{2}}\text{dist}(x', \text{supp}(\chi)))^{-N}. \end{aligned}$$

An application of the Schur test and the Schwarz inequality then show that, if χ is supported inside Ω , and q supported inside the open cone Γ , the following holds,

$$\|\chi(x)q(D)g\|_{L^2}^2 \lesssim \sum_{k=0}^{\infty} \int_{\Omega \times \Gamma} \langle \xi \rangle^{2\sigma} |\tilde{g}_k(x, \xi)|^2 dx d\xi + \sum_{k=0}^{\infty} \int_{\mathbb{R}^{2n}} \langle \xi \rangle^{2s} |\tilde{g}_k(x, \xi)|^2 dx d\xi,$$

hence $(x_0, \xi_0) \notin WF_\sigma(g)$ by elliptic regularity. \square

Suppose that $(x_0, \xi_0) \notin WF_{s_0+1}(u(t_0))$. Given $\Omega \times \Gamma$ as in Lemma 4.4, we will produce a family of t -dependent weight functions $M(t, x, \xi)$ for $t \geq t_0$, such that

$$(4.6) \quad \begin{aligned} C^{-1} \langle \xi \rangle^{s_0} &\leq M(t, x, \xi) \leq C \langle \xi \rangle^{s_0+1}, \\ M(t_0, x, \xi) &\leq C \langle \xi \rangle^{s_0} \quad \text{for } (x, \xi) \notin \Omega \times \Gamma. \end{aligned}$$

Also, for some $c_t > 0$, if

$$\Omega_t = \{x : |x - x_t| < c_t\}, \quad \Gamma_t = \left\{ \xi : \left| \frac{\xi}{|\xi|} - \frac{\xi_t}{|\xi_t|} \right| < c_t \right\},$$

then the following will hold

$$(4.7) \quad M(t, x, \xi) \geq C^{-1} \langle \xi \rangle^{s_0+1} \quad \text{for } (x, \xi) \in \Omega_t \times \Gamma_t.$$

In addition, we will show that (4.1) and (4.2) hold. Theorem 4.1 then follows immediately from Lemma 4.4 and (4.3).

4.1. The weight function. For $c_t > 0$, and ξ_t close to the positive ξ_1 axis, we take $M(t, x, \xi)$ to be the weight function

$$\langle \xi \rangle^{s_0+1} \left(1 + |\xi| \min \left(1, \text{dist}^2(x, \Omega_{c_t}(x_t)) \right) + |\xi| \text{dist}^2 \left(\frac{\xi}{|\xi|}, K_{c_t}(\xi_t) \right) \right)^{-1},$$

where $\Omega_{c_t}(x_t)$ is the ball of radius c_t centered on x_t , and $K_{c_t}(\xi_t)$ is the closed conic set contained in the half-space $\xi_1 > 0$ whose intersection with the set $\xi_1 = 1$ is the cube of sidelength $2c_t$ centered on $\xi_t/(\xi_t)_1$, with sides parallel to the ξ_j axes. The time-dependent number c_t is given in Lemma 4.5 below, where c_{t_0} is chosen as follows.

Provided that $\Omega_{2c_{t_0}}(x_0) \subset \Omega$ and $K_{2c_{t_0}}(\xi_0) \subset \Gamma$, then condition (4.6) is seen to hold. Thus, if $(x_0, \xi_0) \notin WF_{s_0+1}(u(t_0))$, we can choose c_{t_0} small so that

$$\|M(t_0, x, \xi)\tilde{u}(t_0)\|_{\ell^2 L^2} < \infty.$$

Also, (4.7) holds (with the same c_t) since $\Gamma_t \subset K_{c_t}(\xi_t)$. It remains thus to verify the mapping bounds (4.1) and (4.2) for $t_0 \leq s \leq t \leq T$.

We start with the proof of (4.2), which reduces to showing that, uniformly in k, t, s ,

$$\int M(t, x, \xi)^2 |f \circ \Theta_{s,t}^k|^2(x, \xi) dx d\xi \leq C \int M(s, x, \xi)^2 |f|^2(x, \xi) dx d\xi, \quad s \leq t.$$

Since each $\Theta_{s,t}^k$ is a volume preserving diffeomorphism, this is equivalent to the bound

$$(4.8) \quad M(t, \Theta_{t,s}(x, \xi)) \leq M(s, x, \xi), \quad s \leq t.$$

The map $\Theta_{s,t}^k$ is homogeneous of degree 1 in ξ , and preserves $|\xi|$ up to a uniform multiple, so the factor $\langle \xi \rangle^{s_0+1}$ can be ignored. Furthermore, the projective map induced by $\Theta_{s,t}^k$ on the cosphere bundle is a bilipschitz map, with uniform bounds over k, s, t . Thus, (4.8) holds as a consequence of the following.

Lemma 4.5. *For $c_0 > 0$, let*

$$c_t = c_0 \exp\left(-C \int_{t_0}^t \alpha(r) dr\right).$$

Then for c_0 sufficiently small, and C given below,

$$\Theta_{t,s}\left(\Omega_{c_s}(x_s) \times K_{c_s}(\xi_s)\right) \supset \Omega_{c_t}(x_t) \times K_{c_t}(\xi_t), \quad s \leq t.$$

Proof. Write $\xi = (\xi_1, \xi')$, and consider the projection $(x_t, \xi_t) \rightarrow (x_t, 1, (\xi_t)_1^{-1} \xi'_t)$ of a Hamiltonian curve onto the set $\xi_1 = 1$. Let $\zeta_t = (\xi_t)_1^{-1} \xi'_t$. Then by homogeneity of p_k ,

$$\dot{x}_t = d_\xi p_k(x_t, 1, \zeta_t), \quad \dot{\zeta}_t = -d_{x'} p_k(x_t, 1, \zeta_t) + d_{x_1} p_k(x_t, 1, \zeta_t) \zeta_t.$$

On the set $|\zeta| \leq 10$, the right hand side is Lipschitz in (x, ζ) , with Lipschitz constant $C\alpha(t)$. Hence, if we let

$$Q_c(x^0, \zeta^0) = \{(x, \zeta) : |x - x^0| + \sup_{2 \leq i \leq n} |\zeta_i - \zeta_i^0| \leq c\},$$

then for $t > s$ the image of $Q_{c_t}(x_t, \zeta_t)$ under the reverse-time projected flow is contained in $Q_{c_s}(x_s, \zeta_s)$, where c_s is as in the statement. Since $K_{c_t}(\xi_t)$ is the conic subset of $\mathbb{R}^n \cap \{\xi_1 > 0\}$ whose intersection with $\{\xi_1 = 1\}$ equals $Q_{c_t}(x_t, \zeta_t)$, then by homogeneity of the Hamiltonian flow, for $t > s$

$$\Theta_{t,s}\left(\Omega_{c_s}(x_s) \times K_{c_s}(\xi_s)\right) \supset \Omega_{c_t}(x_t) \times K_{c_t}(\xi_t),$$

provided we choose c_{t_0} small enough so that $Q_{c_t}(x_t, \zeta_t)$ remains within the set $|\zeta| < 10$. Here we use that ζ_t remains in the set $|\zeta_t| < 1$, by the assumption that ξ_t lies in a cone of small angle about the ξ_1 axis. \square

4.2. Fixed time weight bounds for $B(t)$. We now turn to the proof of (4.1). The operator $B(t) : \ell^2 L^2 \rightarrow \ell^2 L^2$ is a sum of four terms, $B = {}_1B + {}_2B + {}_3B + {}_4B$. As before, we let $B_{kj}(t) : L^2(\mathbb{R}^{2n}) \rightarrow L^2(\mathbb{R}^{2n})$ denote the $j \rightarrow k$ component of $B(t)$, which we recall is localized dyadically in ξ on each side.

We start by considering the terms ${}_mB$ for $m = 1, 2, 3$. Recall that ${}_1B_{kj}$ and ${}_2B_{kj}$ vanish unless $|j - k| \leq 1$. For ${}_3B_{kj}$ we may also restrict attention to $|j - k| \leq 1$, since (4.1) holds for the sum over $|j - k| \geq 2$ by (3.10) and (4.6).

Consider then the terms for $|j - k| \leq 1$. By finite overlap in ξ these are almost orthogonal in k , hence we are reduced to establishing, for $m = 1, 2, 3$, that uniformly for j, k with $|j - k| \leq 1$, and $f \in L^2(\mathbb{R}^{2n})$,

$$(4.9) \quad \|M(s, x, \xi) {}_mB_{kj}(s) f\|_{L^2(\mathbb{R}^{2n})} \leq C \alpha(s) \|M(s, x, \xi) f\|_{L^2(\mathbb{R}^{2n})}.$$

We will consider the case $j = k$, as the terms with $j = k \pm 1$ are handled the same way. We ignore the factor $\langle \xi \rangle^{s_0+1}$ in the definition of M , since it introduces the same factor of $2^{k(s_0+1)}$ on both sides.

The terms ${}_1B_{kk}$ and ${}_2B_{kk}$ are the simplest to handle. The operator ${}_1B_{kk}$ is, by Lemmas 3.1 and 3.2 of [12], represented by an integral kernel operator K satisfying

$$|K(x, \xi; y, \eta)| \leq C_N \alpha(s) \left(1 + 2^{\frac{k}{2}} |x - y| + 2^{-\frac{k}{2}} |\xi - \eta|\right)^{-N}.$$

On the other hand, if $|\xi| \approx |\eta| \approx 2^k$, then

$$\left[\frac{M(s, x, \xi)}{M(s, y, \eta)} \right]^{\pm 1} \leq C \left(1 + 2^{\frac{k}{2}} |x - y| + 2^{-\frac{k}{2}} |\xi - \eta|\right)^2,$$

so the term ${}_1B_{kk}$ is seen by the Schur test to satisfy the desired weighted L^2 bound (4.9). The operator ${}_2B_{kk}$ is represented by a similar kernel; this follows from the fact that $\alpha(s)^{-1} [\beta_k(D), a_k(s, x)] q(D)$ is a $S_{1, \frac{1}{2}}^0$ pseudo-differential operator in x , dyadically localized to $|\xi| \approx 2^k$.

For the term ${}_3B_{kk}$, after substituting $p(t, x, D) = a(t, x) q(D)$, freezing t , and replacing $q(D) \beta_k(D)$ by $2^k \beta_k(D)$ (since the exact form of β is unimportant) we can assume that

$${}_3B_{kk} = T_k \beta_k(D) 2^k (a(x) - a_k(x)) \beta_k(D) T_k^*,$$

and need to show that (4.9) holds with $\alpha(s) = 1$ if $\|D^2 a\|_{L^\infty} \leq 1$. The adjoint operator ${}_3B_{kk}^*$ then has the same form as ${}_3B_{kk}$, so that in the estimate (4.9) we may replace $M(s, x, \xi)$ by $M(s, x, \xi)^{-1}$. Letting $\Omega = \Omega_{c_s}(x_s)$, $K = K_{c_s}(\xi_s)$, then since the estimate is over the region $|\xi| \approx 2^k$, we may thus work with the weight

$$M(x, \xi) = 1 + 2^k \min(1, \text{dist}^2(x, \Omega)) + 2^{-k} \text{dist}^2(\xi, K),$$

and show that the analogue of (4.9) holds for ${}_3B_{kk}$.

The conic set K is obtained by intersecting $2n - 2$ distinct half-spaces. Let $\{\omega_j\}_{j=1}^{2n-1}$ be the collection of their outer normals, together with the vector $-e_1$ pointing on the negative ξ_1 axis. We let

$$\langle \omega_j, \xi \rangle_+ = \max(\langle \omega_j, \xi \rangle, 0),$$

and claim that

$$(4.10) \quad \text{dist}^2(\xi, K) \approx \sum_{j=1}^{2n-1} \langle \omega_j, \xi \rangle_+^2.$$

To see this, we first note that each term on the right vanishes on K , so the right side is dominated by the left. To prove the converse, we make an affine transformation preserving ξ_1 so that K is centered on the ξ_1 axis. The collection of $\langle \omega_j, \xi \rangle$ are then equivalent to the collection of $\pm \xi_j - c \xi_1$ and $-\xi_1$. In case $\xi_1 \leq 0$, then $\langle -e_1, \xi \rangle_+ = |\xi_1|$. Since $\xi_1 \leq 0$, then

$$|\xi_j| \leq \sum_{\pm} (\pm \xi_j - c \xi_1)_+,$$

so the right hand side of (4.10) dominates $|\xi|^2 \geq \text{dist}^2(\xi, K)$. If $\xi_1 > 0$, let η be the point in K closest to ξ . If $|\xi_j| \leq c \xi_1$, then $\eta_j = \xi_j$, so by reducing dimension and multiplying ξ_j by -1 if needed, we may assume that $\xi_j > c \xi_1$ for each j . Then

$$\sum_{j=1}^{2n-1} \langle \omega_j, \xi \rangle_+^2 = \sum_{j=2}^n |\xi_j - c \xi_1|^2 = \text{dist}^2(\xi, (\xi_1, c \xi_1, \dots, c \xi_1))^2 \geq \text{dist}^2(\xi, K).$$

Including the spatial weight, we can thus replace $M(x, \xi)$ by a sum of $2n$ weights, and it suffices to establish the analogue of (4.9) separately for each. Precisely, by Lemmas 4.2 and 4.3, it suffices to show that multiplication by $2^k(a(x) - a_k(x))$ preserves the spaces with norms

$$\|(1 + 2^k \min(1, \text{dist}^2(x, \Omega)))g(x)\|_{L^2(dx)}, \quad \|(1 + 2^{-k} \langle \omega, \xi \rangle_+^2)\widehat{g}(\xi)\|_{L^2(d\xi)},$$

for a general unit vector ω .

Boundedness in the first norm is immediate since $\|a - a_k\|_{L^\infty} \leq 2^{-k}\|D^2 a\|_{L^\infty}$. For the second norm, we make a rotation to reduce to the case $\omega = (1, 0, \dots, 0)$. Let $b(x) = 2^k(a - a_k)(2^{-\frac{k}{2}}x)$. Then

$$\|b\|_{L^\infty} + \|Db\|_{L^\infty} + \|D^2 b\|_{L^\infty} \lesssim 1.$$

Thus, after scaling $x \rightarrow 2^{\frac{k}{2}}x$, we need to show that

$$\|(1 + (\xi_1)_+^2)\widehat{bf}\|_{L^2(d\xi)} \lesssim \|b\|_{C^{1,1}}\|(1 + (\xi_1)_+^2)\widehat{f}\|_{L^2(d\xi)}.$$

Since the weight is a function of ξ_1 only, and $C^{1,1}(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^{n-1}, C^{1,1}(\mathbb{R}))$, we may assume $n = 1$, that $b \in C^{1,1}(\mathbb{R})$, and need show that

$$\|(1 + \xi_+^2)\widehat{bf}\|_{L^2(\mathbb{R})} \lesssim \|b\|_{C^{1,1}}\|(1 + \xi_+^2)\widehat{f}\|_{L^2(\mathbb{R})}.$$

If \widehat{f} is supported in $[0, \infty)$, then the bound follows from the fact that

$$\|\langle D \rangle^2(bf)\|_{L^2} \lesssim \|b\|_{C^{1,1}}\|\langle D \rangle^2 f\|_{L^2}.$$

Hence we may assume \widehat{f} is supported in $-\infty, 0]$. Since

$$\|\chi_{(-\infty, 2]}\widehat{bf}\|_{L^2} \leq \|bf\|_{L^2} \leq \|b\|_{L^\infty}\|f\|_{L^2} \leq \|b\|_{C^{1,1}}\|(1 + \xi_+^2)\widehat{f}\|_{L^2},$$

it suffices to then bound

$$\|\langle \xi \rangle^2 \widehat{bf}\|_{L^2([2, \infty))} \lesssim \|b\|_{C^{1,1}}\|f\|_{L^2}, \quad \text{supp}(\widehat{f}) \subset (-\infty, 0].$$

Given $h \in L^2(\mathbb{R})$ with \widehat{h} supported in $[2, \infty)$, using the functions ϕ_j and ψ_j from (7.2), we may write

$$\begin{aligned} \left| \int h \langle D \rangle^2 (bf) dx \right| &= \left| \sum_{k \leq j+2} \int (\psi_k(D) \langle D \rangle^2 h) (\psi_j(D)b) (\phi_{j+4}(D)f) dx \right| \\ &\leq \sum_{k \leq j+2} 4^{k-j} \int |2^{-2k} \psi_k(D) \langle D \rangle^2 h| |2^{2j} \psi_j(D)b| |\phi_{j+4}(D)f| dx \\ &\lesssim \|h\|_{L^2} \left(\sum_{j=0}^{\infty} \int |2^{2j} \psi_j(D)b|^2 |\phi_{j+4}(D)f|^2 dx \right)^{\frac{1}{2}} \\ &\lesssim \|h\|_{L^2} \|b\|_{C^{1,1}} \|f\|_{L^2} \end{aligned}$$

where at the last step we use Theorem 7.3. This completes the proof for ${}_3B_{jk}$.

We now establish (4.1) for the term ${}_4B(t) = TQ(t)T^*$. Recall that

$$Q(t) = \begin{cases} P(t)^{-1}R_1(t), & s_0 = 1, \\ R_1(t)P(t)^{-1}, & s_0 = 0, \end{cases}$$

where $R_1(t)$ is a convergent sum of terms of the form (2.3). We observe that if $M_{s_0} = M_{s_0}(t, x, \xi)$ denotes the weight for s_0 , then

$$\|M_1 Tg\|_{\ell^2 L^2} \approx \|M_0 T \langle D \rangle g\|_{\ell^2 L^2}.$$

Also, if $q(\xi) \in S^0(\mathbb{R}^n)$, then

$$\|M_{s_0} Tq(D)T^*f\|_{\ell^2 L^2} \leq C \|M_{s_0}f\|_{\ell^2 L^2},$$

since the operator $T_k \beta_k(D)q(D)\beta_j(D)T_j^*$ vanishes unless $|j-k| \leq 1$, and for $|j-k| \leq 1$ is given by an integral kernel with bound

$$(4.11) \quad |K(x, \xi; y, \eta)| \leq C_N \left(1 + 2^{\frac{k}{2}}|x-y| + 2^{-\frac{k}{2}}|\xi-\eta|\right)^{-N}.$$

Since $T^*T = I$, it therefore suffices to show the following bounds:

$$(4.12) \quad \|M_0 T a T^*f\|_{\ell^2 L^2} \leq C \|a\|_{C^{0,1}} \|M_0 f\|_{\ell^2 L^2}$$

$$(4.13) \quad \|M_0 T[a, q(D)]T^*f\|_{\ell^2 L^2} \leq C \|a\|_{C^{1,1}} \|M_0 f\|_{\ell^2 L^2}$$

$$(4.14) \quad \|M_0 T \langle D \rangle P(t)^{-1}T^*f\|_{\ell^2 L^2} \leq C \|M_0 f\|_{\ell^2 L^2}$$

where in (4.13) the multiplier $q(\xi)$ belongs to $S_{cl}^1(\mathbb{R}^n)$.

To establish (4.12) it suffices to prove

$$(4.15) \quad \|M_0 T_j \beta_j(D) a \beta_k(D) T_k^* f\|_{L^2} \leq C \|a\|_{C^{0,1}} \|M_0 f\|_{L^2}, \quad |j-k| \leq 1,$$

since the terms for $|j-k| \geq 2$ are handled by the arguments leading to (3.10), together with Lemma 7.4 and the fact that $c \leq M_0 \leq \langle \xi \rangle$.

By taking adjoints, we may replace M_0 by M_0^{-1} in (4.15), and ignore the factor $\langle \xi \rangle$ in M_0 since $|j-k| \leq 1$. Furthermore, we may replace a by the operator $(\phi_k(D)a)\phi_k(D)$ for a compactly supported $\phi(\xi)$. As with the handling of the term ${}_3B_{kk}$, it suffices to prove that if $\|a\|_{C^{0,1}(\mathbb{R}^n)} \leq 1$, then $(\phi_k(D)a)\phi_k(D)$ preserves the spaces with norms

$$(4.16) \quad \|(1 + 2^k \min(1, \text{dist}^2(x, \Omega)))g(x)\|_{L^2}, \quad \|(1 + 2^{-k} \langle \omega, \xi \rangle_+^2)\widehat{g}(\xi)\|_{L^2}.$$

Boundedness in the first norm is simple, since $\phi_k(D)$ is a convolution kernel that is rapidly decreasing on scale 2^{-k} . For the second norm, we can reduce to the one-dimensional case, and need to prove that

$$\|(1 + 2^{-k}D_+^2)(\phi_k(D)a)(\phi_k(D)g)\|_{L^2(\mathbb{R})} \leq C \|a\|_{C^{0,1}} \|(1 + 2^{-k}D_+^2)\widehat{g}\|_{L^2(\mathbb{R})},$$

where D_+ is the operator with multiplier $\xi_+ = \max(\xi, 0)$.

Consider first the case that \widehat{g} is supported in $\xi \leq 0$. Then since $|\xi| \lesssim 2^k$ on the frequency support of $(\phi_k(D)a)(\phi_k(D)g)$, this follows from the bound

$$\|(1 + D_+)(ag)\|_{L^2} \leq C \|a\|_{C^{0,1}} \|(1 + D_+)g\|_{L^2},$$

which holds by Theorem 7.1 since ξ_+ is a classical first order multiplier. If \widehat{g} is supported in $\xi \geq 0$, it suffices to prove the bound

$$\|2^{-k}D^2(\phi_k(D)a)(\phi_k(D)g)\|_{L^2} \leq C \|a\|_{C^{0,1}} \left(\|g\|_{L^2} + 2^{-k}\|D^2g\|_{L^2} \right).$$

This holds by distributing derivatives, and using that $\|D^2\phi_k(D)a\|_{L^\infty} \lesssim 2^k\|a\|_{C^{0,1}}$, in addition to $\|D\phi_k(D)g\|_{L^2} \lesssim 2^k\|g\|_{L^2}$.

The estimate (4.13) is similarly reduced by Lemma 7.4 to handling $|j - k| \leq 1$. We then need to show that the commutator $[(\phi_k(D)a), \rho_k(D)q(D)]$ is bounded in the norms (4.16) with operator norm $\lesssim \|a\|_{C^{1,1}}$. Here $\rho_k(\xi)q(\xi)$ is an order 1 classical symbol dyadically localized to $|\xi| \approx 2^k$. Thus, the kernel $K(x, y)$ of the commutator has bounds

$$|K(x, y)| \leq C_N \|a\|_{C^{0,1}} 2^{kn} (1 + 2^k|x - y|)^{-N},$$

so boundedness in the first norm in (4.16) follows by the Schur test as the weight is slowly varying over distance 2^{-k} . For boundedness in the second norm, we assume $\langle \omega, \xi \rangle = \xi_1$, let $q_k(D) = \rho_k(D)q(D)$, and need show that, if \widehat{g} vanishes for $|\xi| \geq 2^k$, then

$$(4.17) \quad \|(1 + D_{1,+})[(\phi_k(D)a), q_k(D)]g\|_{L^2} \leq C \|a\|_{C^{1,1}} \|(1 + D_{1,+})g\|_{L^2},$$

$$(4.18) \quad \|2^{-k}D_1^2[(\phi_k(D)a), q_k(D)]g\|_{L^2} \leq C \|a\|_{C^{1,1}} \left(\|g\|_{L^2} + 2^{-k}\|D_1^2g\|_{L^2} \right).$$

The estimate (4.17) follows from Corollary 7.9, since we may replace the multiplier $1 + \xi_{1,+}$ by its truncation to $|\xi| \lesssim 2^k$. The estimate (4.18) follows by distributing derivatives similar to above, and using that $\|D_1a\|_{C^{0,1}} + 2^{-k}\|D_1^2\phi_k(D)a\|_{C^{0,1}} \lesssim \|a\|_{C^{1,1}}$, together with Theorem 7.1.

We now turn to the proof of (4.14). By Lemma 7.10,

$$\langle D \rangle P(t)^{-1} = \langle D \rangle (p^\sharp(t, x, D) + c)^{-1} \sum_{n=0}^{\infty} \left(p^\flat(t, x, D) (p^\sharp(t, x, D) + c)^{-1} \right)^n,$$

where the sum converges as a map on $L^2(\mathbb{R}^n)$, uniformly over t . Since $(p^\sharp(t, x, D) + c)^{-1}$ is a pseudodifferential operator of class $S_{1, \frac{1}{2}}^{-1}$, it follows from (7.6) that

$$\left\| \left(p^\flat(t, x, D) (p^\sharp(t, x, D) + c)^{-1} \right)^2 g \right\|_{H^1} \leq C \|g\|_{L^2}$$

uniformly in t . Thus, the sum of terms over $n \geq 2$ gives a bounded map from L^2 to H^1 , and the bound (4.14) holds for these terms since $c \leq M_0 \leq \langle \xi \rangle$.

It thus suffices to show that

$$(4.19) \quad \|M_0 T p^\flat(t, x, D) \langle D \rangle^{-1} T^* f\|_{\ell^2 L^2} \leq C \|M_0 f\|_{\ell^2 L^2}$$

$$(4.20) \quad \|M_0 T \langle D \rangle (p^\sharp(t, x, D) + c)^{-1} T^* f\|_{\ell^2 L^2} \leq C \|M_0 f\|_{\ell^2 L^2}$$

In proving (4.19), we take the symbol expansion (7.1) of p , and use that order 0 multipliers are bounded in the M_0 norm, to replace $p^\flat(x, D) \langle D \rangle^{-1}$ by a^\flat , and thus need show

$$\|M_0 T a^\flat T^* f\|_{\ell^2 L^2} \leq C_0 \|a\|_{C^{0,1}} \|M_0 f\|_{\ell^2 L^2}.$$

This is proven as for (4.12). Indeed the off-diagonal terms of a^\flat are the same as for multiplication by a , and if $|j - k| \leq 1$ the bound holds for both a and $\phi_{\lfloor k/2 \rfloor}(D)a$.

The bound for (4.20) is simpler. The operator $\langle D \rangle (p^\sharp(t, x, D) + c)^{-1}$ is a pseudo-differential operator of type $S_{1, \frac{1}{2}}^0$, so the off-diagonal part is a smoothing operator; in particular it maps $L^2(\mathbb{R}^n)$ to $H^1(\mathbb{R}^n)$. And for $|j - k| \leq 1$, the operator

$$T_j \beta_j(D) \langle D \rangle (p^\sharp(t, x, D) + c)^{-1} \beta_k(D) T_k^*$$

is an integral kernel operator with kernel satisfying (4.11). \square

5. THE SPACE-TIME VERSION: PROOF OF THEOREM 2.2.

In this section we deduce the space-time wavefront estimate in Theorem 2.2 from the fixed-time wavefront estimate established in Theorem 4.1. We use the notation of Section 3, with $P = p(t, x, D)$ denoting a choice of $p_\pm(t, x, D)$, and $Q(t)$ constructed according to the choice of $s_0 \in \{0, 1\}$.

Lemma 5.1. *If $u \in C^0 H^{s_0}$ satisfies $D_t u - P(t)u - Q(t)u = 0$, then*

$$(t_0, x_0, p(t_0, x_0, \xi_0), \xi_0) \notin WF_{s_0+1}(u) \implies (x_0, \xi_0) \notin WF_{s_0+1}(u(t_0)).$$

Proof. Let $\chi(t, x)$ and $\tilde{\chi}(t, x)$ denote cutoff functions, with $\tilde{\chi} = 1$ on a neighborhood of the support of χ , and $\chi = 1$ on a neighborhood of (t_0, x_0) . Also, let $\Gamma(\xi)$ and $\tilde{\Gamma}(\xi)$ denote conic cutoffs, equal to one on a neighborhood of ξ_0 , with $\tilde{\Gamma} = 1$ on a neighborhood of the support of Γ .

Let $\phi \in C_c^\infty(\mathbb{R})$ equal 1 near 0. If the support of $\phi \in C_c^\infty(\mathbb{R})$ is suitably close to 0, and $\tilde{\chi}$ and $\tilde{\Gamma}$ also have suitably small support, then $\gamma(t_0) \notin WF_{s_0+1}(u)$ implies that

$$\tilde{\Gamma}(D) \phi(1 - p(t_0, x_0, D) D_t^{-1})(\tilde{\chi}u) \in H^{s_0+1}(\mathbb{R}^{1+n}).$$

On the other hand, (2.10) implies that $Q(t)u(t) \in C^0 H^{s_0}$, so Theorem 8.4 implies

$$\left(1 - \phi(1 - p(t_0, x_0, D) D_t^{-1})\right)(\tilde{\chi}u) \in L^2 H^{s_0+1},$$

hence

$$(5.1) \quad \tilde{\Gamma}(D)(\tilde{\chi}u) \in L^2 H^{s_0+1}.$$

We next show that

$$(5.2) \quad (D_t - P(t) - Q(t))\Gamma(D)(\chi u) \in L^1 H^{s_0+1}.$$

Together with (5.1) and Theorem 8.5 this implies $\Gamma(D)(\chi u) \in C^0 H^{s_0+1}$, hence $(x_0, \xi_0) \notin WF_{s_0+1}(u(t_0))$.

To establish (5.2), we write

$$(5.3) \quad \begin{aligned} (D_t - P(t) - Q(t))\Gamma(D)\chi u &= [\Gamma(D)\chi, Q(t)]u \\ &\quad - [P(t), \Gamma(D)\chi]\tilde{\Gamma}(D)\tilde{\chi}u + \Gamma(D)\chi P(t)(1 - \tilde{\Gamma}(D)\tilde{\chi})u \\ &\quad + P(t)\Gamma(D)\chi(1 - \tilde{\Gamma}(D)\tilde{\chi})u + \Gamma(D)(D_t\chi)u. \end{aligned}$$

The next to last term on the right belongs to L^1H^2 , since $u \in C^0L^2$ and the cutoffs give a smoothing operator, and the last term belongs to $L^2H^{s_0+1}$ by (5.1), hence both are in $L^1H^{s_0+1}$.

Since $u \in C^0H^{s_0}$, the first term on the right in (5.3) belongs to $L^1H^{s_0+1}$, by (2.13) and the fact that χ is C_c^∞ and the components of Q are smooth symbols in the ξ variable.

For the second and third terms in (5.3), by the symbol expansion (7.1) we may substitute $p(t, x, D) = a(t, x)q(D)$, with $q(\xi)$ a symbol of order 1 and $a \in C^{0,1} \cap L^1C^{1,1}$. We note the following consequence of Lemma 7.4,

$$\|\Gamma(D)\chi a(t, \cdot)q(D)(1 - \tilde{\Gamma}(D))u(t, \cdot)\|_{H^{s_0+1}} \leq C \|\chi a(t, \cdot)\|_{C^{1,1}} \|u(t)\|_{H^{s_0}},$$

and also the following consequence of pseudo-locality of $q(D)\tilde{\Gamma}(D)$, where $s_0+1 \leq 2$,

$$\|\Gamma(D)\chi a(t, \cdot)q(D)\tilde{\Gamma}(D)(1 - \tilde{\chi})u(t, \cdot)\|_{H^{s_0+1}} \leq C \|a(t, \cdot)\|_{C^{1,1}} \|u(t)\|_{H^{s_0}}.$$

Since $u \in C^0H^{s_0}$, this handles the third term on the right in (5.3).

Now consider the second term on the right in (5.3), and write

$$[a q(D), \Gamma(D)\chi] = a\Gamma(D)[q(D), \chi] + [a, \Gamma(D)]\chi q(D).$$

Consider the case $s_0 = 0$. By Corollary 7.2 we have

$$\|[a q(D), \Gamma(D)\chi]v\|_{L^2H^1} \lesssim \|a\|_{L^\infty C^{0,1}} \|v\|_{L^2H^1},$$

which we apply to $v = \tilde{\Gamma}(D)\tilde{\chi}u \in L^2H^1$.

In case $s_0 = 1$, we use that $v = \tilde{\Gamma}(D)\tilde{\chi}u \in L^2H^2 \cap C^0H^1$, by (5.1) and since $u \in C^0H^1$. We again apply Corollary 7.2 to obtain

$$\begin{aligned} \|D[a, \Gamma(D)]\chi q(D)v\|_{L^1H^1} &\leq \|[(Da), \Gamma(D)]\chi q(D)v\|_{L^1H^1} + \|[a, \Gamma(D)]D\chi q(D)v\|_{L^1H^1} \\ &\lesssim \|Da\|_{L^1C^{0,1}} \|v\|_{L^\infty H^1} + \|a\|_{L^\infty C^{0,1}} \|v\|_{L^1H^2}. \end{aligned}$$

Similarly, we use that $\Gamma(D)[q(D), \chi]v \in L^2H^2 \cap C^0H^1$, and the following consequence of the Leibniz rule

$$\|aw\|_{L^1H^2} \leq \|Da\|_{L^1C^{0,1}} \|w\|_{L^\infty H^1} + \|a\|_{L^\infty C^{0,1}} \|w\|_{L^1H^2},$$

to handle the remaining term. \square

We now observe that if the conditions of Theorem 2.2 hold, and $s \in (-T, T)$, then by Lemma 5.1 and Theorem 4.1 there is a function $\chi \in C_c^\infty(\mathbb{R}^{1+n})$ equal to 1 on a neighborhood of (s, x_s) , and conic cutoff $\Gamma(\xi)$ equal to 1 on a neighborhood of ξ_s , so that

$$\Gamma(D)\chi u \in L^2H^{s_0+1},$$

where $\gamma(s) = (s, x_s, p(s, x_s, \xi_s), \xi_s)$. Since $p \approx |\xi|$, it follows that $\gamma(s) \notin WF_{s_0+1}(u)$, which completes the proof of Theorem 2.2. \square

6. PIECEWISE REGULAR COEFFICIENTS

We work in this section with L of the form

$$(6.1) \quad L = D_t^2 - 2D_j b^j(t, x)D_t - D_i c^{ij}(t, x)D_j + d^0(t, x)D_t + d^j(t, x)D_j,$$

where the coefficients satisfy certain piecewise regularity conditions with respect to a decomposition of $(-T, T) \times \mathbb{R}^n$ into disjoint time slabs $-T = t_1 < t_2 < \dots < t_{n-1} < t_n = T$. We assume given such a partition, and assume that the coefficients c^{ij} and b^j satisfy the conditions (1.3)-(1.4) separately on each time slab $(t_j, t_{j+1}) \times \mathbb{R}^n$. In addition, we assume that c^{ij} and b^j are continuous on $[-T, T] \times \mathbb{R}^n$. This implies in particular that c^{ij} and b^j belong to $C^{0,1}([-T, T] \times \mathbb{R}^n)$, and that the map $t \rightarrow c^{ij}(t, \cdot)$, respectively $t \rightarrow b^j(t, \cdot)$, is continuous from $[-T, T]$ into $C^1(\mathbb{R}^n)$.

Similarly, we assume d^0 and d^j satisfy (1.6) separately on each time slab, hence on each slab they admit a continuous extension to $[t_j, t_{j+1}] \times \mathbb{R}^n$. We allow d_j to have jumps at t_j for $1 \leq j \leq n$. It is unimportant how d^j is defined at $t = t_j$, but for definiteness we assume it is right continuous.

We assume the coefficient d^0 belongs to $C^0([-T, T] \times \mathbb{R}^n)$, which with the above is equivalent to assuming $\partial_{t,x} d^0 \in L^1 L^\infty((-T, T) \times \mathbb{R}^n)$. The continuity assumption on d^0 is needed for weak solutions of $Lu = 0$ to agree with solutions defined separately on each slab with matching Cauchy data at each t_j . At the end of this section we indicate how to handle jumps in d^0 .

For $s_0 \in \{0, 1, 2\}$, and Cauchy data of regularity $H^{s_0} \times H^{s_0-1}$ at some t_0 , one obtains a solution to $Lu = 0$ of regularity $C^0 H^{s_0} \cap C^1 H^{s_0-1}$ by piecing together solutions on $[t_j, t_{j+1}]$, and imposing continuity of u and $D_t u$ at t_j . Such a solution is easily verified to satisfy $\int u L^t \phi = 0$ for $\phi \in C_c^\infty((-T, T) \times \mathbb{R}^n)$, with L^t the formal transpose of L .

That this u is the unique weak solution of regularity $C^0 H^{s_0} \cap C^1 H^{s_0-1}$ follows immediately from uniqueness for the Cauchy problem on each time slab, by the assumed continuity condition in t of $(u(t), D_t u(t))$.

Since the first-order derivatives in x of c^{ij} and b^j satisfy the regularity conditions of d^0 , one can convert between the standard form of L in the introduction and one of the form (6.1) and preserve the regularity assumptions. Since the first order derivatives in t of b^j satisfy the conditions on d^j for $1 \leq j \leq n$, one could also express a term of the form $D_t b^j D_j$ in the form (6.1). Indeed, up to addition of an $L^1 L^\infty$ function the class (6.1) is closed under transpose.

If we factor the principal symbol of L as before as

$$H(t, x, \tau, \xi) = (\tau - p_+(t, x, \xi))(\tau + p_-(t, x, \xi)),$$

then p_\pm are C^1 , and $\partial_{x,\xi}^2 p_\pm$ belongs to $L^1 L^\infty$, hence the null bicharacteristics of L are well defined, C^1 curves.

Theorem 6.1. *Assume that the coefficients of L are as above. Suppose that $s_0 \in \{0, 1\}$, that $Lu = 0$, and that $u \in C^0 H^{s_0} \cap C^1 H^{s_0-1}$.*

Then if $\gamma(t)$ is a null bicharacteristic curve of L , and $\gamma(t_0) \notin \text{WF}_{s_0+1}(u)$ for some $t_0 \in (-T, T)$, then $\gamma \cap \text{WF}_{s_0+1}(u) = \emptyset$.

Proof. For simplicity, we assume that the partition consists of $[-T, 0] \cup [0, T]$. The general case follows easily. By openness of the wavefront set we may then assume $t_0 \neq 0$, and without loss of generality take $t_0 < 0$.

We derive the result as a limiting case of Theorem 1.1, using uniformity of the wavefront set estimates over bounded sets of coefficients. Precisely, we use the fact that all of the bounds on wavefront sets involve only uniform control over appropriate norms of the coefficients. To define uniform cutoffs, we fix a smooth radial cutoff to the half-unit ball $\chi(t, x)$, supported in the unit ball, and let $\chi_{c, t_0, x_0}(t, x) = \chi(c^{-1}(t - t_0), c^{-1}(x - x_0))$. We also fix a conic cutoff Γ , rotationally symmetric about the ξ_1 axis and supported in the cone of angle $c\pi$,

$$\Gamma_c(\tau, \xi) = \chi(c^{-1}\xi_1^{-1}\tau, c^{-1}\xi_1^{-1}\xi', 0),$$

and define $\Gamma_{c, \tau_0, \xi_0}(\tau, \xi)$ by composing Γ_c with a rotation that centers it on the ray through (τ_0, ξ_0) . The following result is then a consequence of the fact that the bounds and support of the cutoffs in the wavefront estimates in proof of Theorem 1.1 depend only on bounds for the cited quantities in L .

Corollary 6.2. *Suppose that for some $0 < c_0, C_0 < \infty$, the coefficients of L satisfy the bounds (1.2), (1.3), (1.4), and (1.6), where $\|\alpha\|_{L^1} \leq C_0$.*

Suppose that $u \in C^0 H^{s_0} \cap C^1 H^{s_0-1}$ satisfies $Lu = 0$, and that

$$\sup_{t \in (-T, T)} \left(\|u(t)\|_{H^{s_0}} + \|D_t u(t)\|_{H^{s_0-1}} \right) \leq C_0.$$

Let $\gamma(t) = (t, x_t, \tau_t, \xi_t)$ be a null bicharacteristic for L , and suppose that the following holds for some $0 < c_1, C_1 < \infty$, and some t_0 ,

$$\|\Gamma_{c_1, \tau_{t_0}, \xi_{t_0}}(\mathbf{D}) \chi_{c_1, t_0, x_{t_0}} u\|_{H^{s_0+1}} \leq C_1.$$

Then if $T' < T$ there are $0 < c_2, C_2 < \infty$, depending only on c_0, C_0, c_1, C_1 , and T' , so that

$$\|\Gamma_{c_2, \tau_t, \xi_t}(\mathbf{D}) \chi_{c_2, t, x_t} u\|_{H^{s_0+1}} \leq C_2,$$

for all $|t| \leq T'$.

We will consider a family of operators L_n of the form

$$L_n = D_t^2 - 2D_j b_n^j D_t - D_i c_n^{ij} D_j + d_n^0 D_t + d_n^j D_j,$$

which converge appropriately to L , and such that the coefficients of L_n satisfy (1.2), (1.3), (1.4), and (1.6), with constants c_0, C_0 uniform over n . Since the class L of Theorem 1.1 can be expressed in the form (6.1) with comparable c_0, C_0 , the corollary applies to L_n .

To construct L_n , we fix an increasing function $h \in C^\infty(\mathbb{R})$ which vanishes for $s < -1$, equals 1 for $s > 1$, and so that $h(s) + h(-s) = 1$. For $c^{ij}(t, x)$ as above, we let c_-^{ij} denote its restriction to $t \in [-T, 0]$, which using (1.7) we assume extended to a function on $[-T, T]$ satisfying (1.3)-(1.4) there. Similarly, let c_+^{ij} denote its restriction to $[0, T]$, appropriately extended to $[-T, T]$. Define

$$\begin{aligned} c_n^{ij}(t, x) &= h(-nt) c_-^{ij}(t, x) + h(nt) c_+^{ij}(t, x) \\ &= c_-^{ij}(t, x) + h(nt) (c_+^{ij}(t, x) - c_-^{ij}(t, x)). \end{aligned}$$

Since $c_+^{ij}(0, x) = c_-^{ij}(0, x)$, it is seen that the estimates (1.3) and (1.4) are satisfied by c_n^{ij} on $(-T, T)$ with uniform bounds for $\|\alpha_n\|_{L^1}$. Furthermore

$$c_n^{ij}(t, x) = c^{ij}(t, x) \quad \text{if } |t| > \frac{1}{n}.$$

We apply this smoothing technique to the coefficients c^{ij} and b^j of L . Since d^0 is already globally regular we set $d_n^0 = d^0$. We also define

$$d_n^j(t, x) = h(-nt) d_-^j(t, x) + h(nt) d_+^j(t, x), \quad 1 \leq j \leq n,$$

which satisfies (1.6) with uniform bounds on $\|\alpha_n\|_{L^1}$. We then define L_n to be the operator of form (6.1) with modified coefficients, and note that $L = L_n$ for $|t| > \frac{1}{n}$.

If we factor the Hamiltonian of the principal part of L_n as

$$H_n(t, x, \tau, \xi) = (\tau - p_{n,+}(t, x, \xi))(\tau + p_{n,-}(t, x, \xi)),$$

then $p_{n,\pm} = p_{\pm}$ for $|t| > \frac{1}{n}$, and $D_{x,\xi} p_{n,\pm}$ converges uniformly to $D_{x,\xi} p_{\pm}$ on compact sets. It follows that the null bicharacteristic of L_n through a given initial point converges uniformly on $[-T, T]$ to the null bicharacteristic of L through that point.

We henceforth assume n large so that $t_0 < -\frac{1}{n}$. Then the solution to $Lu = 0$ with given Cauchy data at t_0 satisfies $L_n u = 0$ for $-T < t < \frac{1}{n}$, in particular for t near t_0 . Thus, if we let u_n be the solution to $L_n u_n = 0$, with the same Cauchy data as u at t_0 , then $u_n = u$ for $-T < t < -\frac{1}{n}$. In particular, for all n , and c_1 small,

$$\|\Gamma_{c_1, \tau_{t_0}, \xi_{t_0}}(\mathbf{D}) \chi_{c_1, t_0, x_{t_0}} u_n\|_{H^{s_0+1}} \leq C_1.$$

Thus, since the null bicharacteristic of L_n through $(t_0, x_0, \tau_0, \xi_0)$ converges uniformly to the null bicharacteristic of L through $(t_0, x_0, \tau_0, \xi_0)$, Corollary 6.2 shows that, for n large and some small $c_2 > 0$, with C_2 independent of n ,

$$\|\Gamma_{c_2, \tau_t, \xi_t}(\mathbf{D}) \chi_{c_2, t, x_t} u_n\|_{H^{s_0+1}} \leq C_2.$$

We will prove that some subsequence u_{n_j} converges weakly as distributions to u , from which we obtain the desired result

$$\|\Gamma_{c_2, \tau_t, \xi_t}(\mathbf{D}) \chi_{c_2, t, x_t} u\|_{H^{s_0+1}} \leq C_2.$$

To show the convergence, we observe that by weak compactness some subsequence $(u_{n_j}, D_t u_{n_j})$ converges weakly in $L^\infty H^{s_0} \times L^\infty H^{s_0-1}$ to $(v, D_t v) \in L^\infty H^{s_0} \times L^\infty H^{s_0-1}$. We next verify that v is in fact of regularity $C^0 H^{s_0} \cap C^1 H^{s_0-1}$ separately on $[-T, 0] \times \mathbb{R}^n$ and $[0, T] \times \mathbb{R}^n$. For $t < 0$ this is trivial since $u_n = u$ for $t < -\frac{1}{n}$, hence $v = u$ for $t < 0$. For $t > 0$, if $s_0 = 1$ then it follows from Theorem 8.6 since $Lv = 0$ and $v \in H^1((-T, T) \times \mathbb{R}^n)$. If $s_0 = 0$, then Theorem 8.1 yields that $L(\langle D \rangle^{-1} v) \in L^1 L^2$. Since $(\langle D \rangle^{-1} v, D_t \langle D \rangle^{-1} v) \in L^\infty H^1 \times L^\infty L^2$, the result again follows from Theorem 8.6.

Thus, v consists of regular solutions on $(-T, 0)$ and $(0, T)$ to $Lv = 0$, and it remains only to show that the Cauchy data match at $t = 0$, since $v = u$ for $t < 0$. To see that the data match, we note that, for $\psi(t, x) \in C_c^\infty((-T, T) \times \mathbb{R}^n)$,

$$0 = \int v(L^t \psi) dt dx.$$

Integration by parts separately on $t > 0$ and $t < 0$ leads to the following condition:

$$\begin{aligned} & \int_{\mathbb{R}^n} \left(v(0^+, x) - v(0^-, x) \right) \left(D_t \psi(0, x) - b^j(0, x) D_j \psi(0, x) \right) dx \\ &= \int_{\mathbb{R}^n} \left(D_t v(0^+, x) - D_t v(0^-, x) + d^0(0^+, x) v(0^+, x) - d^0(0^-, x) v(0^-, x) \right) \psi(0, x) dx. \end{aligned}$$

Since this vanishes for all ψ , we must have $v(0^+, x) = v(0^-, x)$, and if $d^0(0^+, x) = d^0(0^-, x)$ as we assume, then also $D_t v(0^+, x) = D_t v(0^-, x)$. \square

We remark that if d^0 is piecewise regular with jumps at t_j , that is, of the same regularity as d^j , then the result still holds, but the solution u must be defined by piecing together $C^0 H^{s_0} \cap C^1 H^{s_0-1}$ solutions on $[t_j, t_{j+1}]$ with the following matching conditions on u at each t_j ,

$$(6.2) \quad u(t_j^+) = u(t_j^-), \quad D_t u(t_j^+) - D_t u(t_j^-) + d^0(t_j^+) u(t_j^+) - d^0(t_j^-) u(t_j^-) = 0.$$

The proof shows that the limiting solution v agrees with this solution u , hence the result of Theorem 6.2 holds for u satisfying (6.2). The one modification to the proof is to define d_n^0 similar to d_n^j , so that it meets the regularity conditions (1.6).

6.1. An example showing sharpness. We now show that the assumption of global H^{s_0} regularity on u cannot be lowered in Theorem 6.1, hence it is necessary for Theorem 1.1 to hold with bounds depending only on the appropriate norms of the coefficients. Precisely, we construct a piecewise smooth operator L and a corresponding null bicharacteristic γ , and for each $\sigma \leq 2$ a solution $Lu = 0$ with u of H^σ regularity, such that u is microlocally smooth on $\gamma(t)$ for $t < -1$, but for all $\epsilon > 1$ we have $\gamma(t) \in WF_{\sigma+\epsilon}(u)$ for $t \geq -1$.

Consider the following hyperbolic equation on $\mathbb{R}_t \times \mathbb{R}_x$

$$(6.3) \quad \begin{cases} (\partial_t^2 + t\partial_x^2)u(t, x) = 0, & t \leq -1, \\ (\partial_t^2 - \partial_x^2)u(t, x) = 0, & t \geq -1. \end{cases}$$

Let $A(s) = Ai(\omega s)$ be the solution to the Airy equation $A''(s) + sA(s) = 0$, where $\omega = e^{2\pi i/3}$. Then, for $s < 0$,

$$A(s) = e^{i\frac{2}{3}(-s)^{3/2}} a(-s), \quad a(s) \sim \sum_{k=0}^{\infty} a_k s^{-\frac{1}{4} - \frac{3}{2}k}.$$

Furthermore, $A(s) \neq 0$ for $s < 0$, and each $a_k \neq 0$.

If $\xi \geq 1$, we consider the following solutions to (6.3),

$$u_\xi(t, x) = \begin{cases} e^{i\xi x} \frac{A(\xi^{2/3}t)}{A(-\xi^{2/3})}, & t \leq -1, \\ e^{i\xi x} \left(c_0(\xi) e^{-i\xi(t+1)} + c_1(\xi) e^{i\xi(t+1)} \right), & t \geq -1, \end{cases}$$

where the following matching conditions are met to yield $u \in C^{1,1}(\mathbb{R}^2)$,

$$c_0(\xi) + c_1(\xi) = 1, \quad -i\xi \left(c_0(\xi) - c_1(\xi) \right) = \xi^{2/3} \frac{A'(-\xi^{2/3})}{A(-\xi^{2/3})} = -i\xi \left(1 - i\xi^{-1/3} \frac{a'(\xi^{2/3})}{a(\xi^{2/3})} \right),$$

Then $c_0(\xi)$ and $c_1(\xi)$ are smooth on $\xi \geq 1$, and admit an asymptotic expansion

$$c_0(\xi) = 1 - c_1(\xi), \quad c_1(\xi) \sim \frac{i}{8}\xi^{-1} \left(1 + \sum_{k=1}^{\infty} d_k \xi^{-k}\right).$$

If $\xi \leq -1$, we set

$$(6.4) \quad u_\xi(t, x) = \overline{u_{-\xi}(t, x)}.$$

For $-1 \leq \xi \leq 1$ we take a combination of the solutions A and \bar{A} in the definition of u_ξ so that $u_\xi(-1, x) = e^{i\xi x}$, and (6.4) holds. The following then holds,

$$u_\xi(t, x) = \begin{cases} e^{i\xi x + i\frac{2}{3}\xi(-t)^{3/2} - i\frac{2}{3}\xi} a(t, \xi), & t \leq -1, \\ e^{i\xi x} \left(c_0(\xi) e^{-i\xi(t+1)} + c_1(\xi) e^{i\xi(t+1)} \right), & t \geq -1, \end{cases}$$

where $a(t, \xi)$ and $c_0(\xi)$ are elliptic symbols in ξ of order 0, and $c_1(\xi)$ is elliptic of order -1 .

Let $b(\xi) = \langle \xi \rangle^{-\frac{1}{2}-\sigma} \log(2 + |\xi|)^{-1}$, and set

$$u(t, x) = \int_{-\infty}^{\infty} b(\xi) u_\xi(t, x) d\xi.$$

Then $u \in C^0 H^\sigma \cap C^1 H^{\sigma-1} \cap C^2 H^{\sigma-2}$. Consequently $WF_\sigma(u) = \emptyset$ if $\sigma \leq 2$.

For each $\sigma \in \mathbb{R}$ and $\epsilon > 0$, on the set $t < -1$ we see by stationary phase that

$$WF_{\sigma+\epsilon} u = \left\{ (t, x, \tau, \xi) : x = -\frac{2}{3}(-t)^{3/2} + \frac{2}{3}, \tau + (-t)^{1/2}\xi = 0 \right\}.$$

For $t > -1$ and $0 < \epsilon \leq 1$, since c_1 is of order -1 we have

$$WF_{\sigma+\epsilon} u = \left\{ (t, x, \tau, \xi) : x = 1 + t, \tau + \xi = 0 \right\},$$

On the other hand, for $t > -1$ and $\epsilon > 1$, we have

$$WF_{\sigma+\epsilon} u = \cup_{\pm} \left\{ (t, x, \tau, \xi) : x = \pm(1+t), \tau \pm \xi = 0 \right\}.$$

Thus, if $\epsilon > 1$ then $WF_{\sigma+\epsilon}$ contains the null bicharacteristic $\gamma(t) = (t, -(1+t), 1, 1)$ for $t \geq -1$, but is microlocally smooth on its continuation for $t < -1$, $\gamma(t) = (t, \frac{2}{3}(-t)^{3/2} - \frac{2}{3}, (-t)^{1/2}, 1)$. \square

7. APPENDIX 1: PARAPRODUCT ESTIMATES

In this section we collect the paraproduct and commutator estimates used throughout this paper. By a standard multiplier of order m , we understand a function $q(\xi) \in C^\infty(\mathbb{R}^n)$ such that

$$|\partial_\xi^\alpha q(\xi)| \leq C_\alpha (1 + |\xi|)^{m-|\alpha|},$$

and denote the class of such multipliers by $S^m(\mathbb{R}^n)$. The best constants C_α form the seminorms of q . In the statements of this section, it is implicit that the constant C in any given operator bound for a multiplier depends on a finite number of the C_α . We say that $q \in S^m$ is a classical multiplier if in addition

$$q(r\xi) = r^m q(\xi), \quad r \geq 1, \quad |\xi| \geq 1,$$

and denote this subspace by S_{cl}^m .

The homogeneous symbol $p(t, x, \xi)$ admits a convergent expansion on the set $|\xi| \geq 1$ of the form

$$(7.1) \quad p(t, x, \xi) = \sum_{l=1}^{\infty} a_l(t, x) q_l(\xi), \quad q_l(\xi) = |\xi| \omega_l(\xi/|\xi|) \in S_{cl}^1(\mathbb{R}^n),$$

where ω_l are spherical harmonics, and $a_l(t, x)$ satisfies the regularity conditions in (1.3) and (1.4), with constants C_l that decrease rapidly in l . We may smoothly extend the $q_l(\xi)$ near 0 so that this expansion is valid for all ξ . The seminorms of q_l grow at most polynomially in l , so the bounds in prior sections on R_1^\pm , etc., are convergent.

The Coifman-Meyer commutator theorem [6], which generalizes the Calderón commutator theorem [4] for homogeneous multipliers, is the following.

Theorem 7.1 (Coifman-Meyer commutator theorem). *Suppose that $a \in C^{0,1}(\mathbb{R}^n)$, and $q \in S^1(\mathbb{R}^n)$. Then*

$$\| [a, q(D)] f \|_{L^2} \leq C \|a\|_{C^{0,1}} \|f\|_{L^2}.$$

An immediate corollary, as seen by commuting or composing with D , is the following

Corollary 7.2. *If $q \in S^1(\mathbb{R}^n)$, and $a \in C^{1,1}(\mathbb{R}^n)$, then*

$$\| [a, q(D)] f \|_{H^s} \leq C \|a\|_{C^{1,1}} \|f\|_{H^s}, \quad -1 \leq s \leq 1.$$

If $q \in S^0(\mathbb{R}^n)$, and $a \in C^{0,1}(\mathbb{R}^n)$, respectively $a \in C^{1,1}(\mathbb{R}^n)$, then

$$\| [a, q(D)] f \|_{H^{s+1}} \leq C \|a\|_{C^{0,1}} \|f\|_{H^s}, \quad -1 \leq s \leq 0,$$

$$\| [a, q(D)] f \|_{H^{s+1}} \leq C \|a\|_{C^{1,1}} \|f\|_{H^s}, \quad -2 \leq s \leq 1.$$

A key ingredient in the proof of the commutator theorem is the following estimate, due to Carleson [5] and Fefferman-Stein [8]; for a proof, see [13, II.2.4, IV.4.3].

Theorem 7.3. *Suppose that $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$, and that $\psi(0) = 0$. Let $\psi_j(D) = \psi(2^{-j}D)$, $\phi_j(D) = \phi(2^{-j}D)$. Then*

$$\left(\sum_{j=0}^{\infty} \int_{\mathbb{R}^n} |\psi_j(D)a|^2 |\phi_j(D)f|^2 dx \right)^{\frac{1}{2}} \leq C \|a\|_{BMO} \|f\|_{L^2}.$$

Theorem 7.3 yields smoothing estimates for the off-diagonal terms in paraproducts. To state these, form a Littlewood-Paley partition of unity $\{\psi_k(\xi)\}_{k=1}^{\infty}$ by taking $\psi_k(\xi) = \beta_k^2(\xi)$, with β_k as in (3.3). Then let

$$(7.2) \quad \phi_k(D) = \sum_{j=0}^{k-2} \psi_j(D), \quad \rho_k(D) = \sum_{j=k-1}^{k+1} \psi_j(D).$$

If $g \in L^2(\mathbb{R}^n)$ and $a \in L^\infty(\mathbb{R}^n)$, we decompose $ag = R_a g + T_a g$, where

$$\begin{aligned} R_a g &= \sum_{|j-k| \geq 2} \psi_j(D) (a \psi_k(D) g) \\ &= \sum_{j=2}^{\infty} \psi_j(D) (a \phi_j(D) g) + \sum_{k=2}^{\infty} \phi_k(D) (a \psi_k(D) g), \end{aligned}$$

and

$$\begin{aligned} T_a g &= \sum_{|j-k| \leq 1} \psi_j(D) (a \psi_k(D) g) \\ &= \sum_{j=0}^{\infty} \psi_j(D) \left((\phi_{j+4}(D) a) (\rho_j(D) g) \right). \end{aligned}$$

With the exception of the last identity, in the above a may be replaced by a general bounded linear operator on L^2 .

Lemma 7.4. *Suppose $a \in C^{1,1}(\mathbb{R}^n)$. If $-1 \leq \sigma \leq 1$, then*

$$\| \langle D \rangle^{1+\sigma} R_a (\langle D \rangle^{1-\sigma} g) \|_{L^2} \lesssim \|a\|_{C^{1,1}} \|g\|_{L^2}.$$

Suppose $a \in C^{0,1}(\mathbb{R}^n)$. If $0 \leq \sigma \leq 1$, then

$$\| \langle D \rangle^\sigma R_a (\langle D \rangle^{1-\sigma} g) \|_{L^2} \lesssim \|a\|_{C^{0,1}} \|g\|_{L^2}.$$

Proof. We prove the first estimate; the second follows by similar steps. By interpolation we may restrict attentions to $\sigma = \pm 1$, and by considering adjoints we can assume that $\sigma = -1$. We may then replace $\langle D \rangle^2$ by D^2 , which denotes an arbitrary second-order derivative. First consider

$$\sum_{k=2}^{\infty} \phi_k(D) (a \psi_k(D) D^2 g) = \sum_{k=2}^{\infty} \phi_k(D) ((\rho_k(D) a) \psi_k(D) D^2 g).$$

We take the inner product with $h \in L^2$; by the Cauchy-Schwarz inequality and almost orthogonality over k of $\psi_k(D)g$, we can dominate the result by

$$\left(\sum_{k=2}^{\infty} \int |2^{2k} \rho_k(D) a|^2 \cdot |\phi_k(D) h|^2 dx \right)^{\frac{1}{2}} \|g\|_{L^2} \lesssim \|D^2 a\|_{BMO} \|h\|_{L^2} \|g\|_{L^2},$$

where we use Theorem 7.3, and write $2^{2j} \rho_j(D) a = \bar{\rho}_j(D) D^2 a$.

Now consider the remaining term,

$$\sum_{j=2}^{\infty} \psi_j(D) (a \phi_j(D) D^2 g) = \sum_{j=2}^{\infty} \psi_j(D) ((\rho_j(D) a) \phi_j(D) D^2 g).$$

By almost orthogonality over j we can dominate the L^2 norm of this sum by

$$\left(\sum_{j=2}^{\infty} \int |2^{2j} \rho_j(D) a|^2 \cdot |2^{-2j} \phi_j(D) D^2 g|^2 dx \right)^{\frac{1}{2}} \lesssim \|D^2 a\|_{BMO} \|g\|_{L^2},$$

where we use Theorem 7.3, and write $2^{-2j} \phi_j(D) D^2 g = \bar{\phi}_j(D) g$. \square

Corollary 7.5. For $a \in C^{0,1}(\mathbb{R}^n)$, define the operator a^\flat by

$$a^\flat g = \sum_{j=0}^{\infty} \left(a - (\phi_{\lfloor j/2 \rfloor}(D) a) \right) \psi_j(D) g.$$

Then

$$\|a^\flat g\|_{H^{s+\frac{1}{2}}} \leq C \|a\|_{C^{0,1}} \|g\|_{H^s}, \quad -1 \leq s \leq \frac{1}{2}.$$

If $a \in C^{1,1}(\mathbb{R}^n)$, then

$$\|a^\flat g\|_{H^{s+1}} \leq C \|a\|_{C^{1,1}} \|g\|_{H^s}, \quad -2 \leq s \leq 1.$$

Proof. We write

$$a^\flat g = R_a g + \sum_{j=0}^{\infty} \rho_j(D) \left(a - (\phi_{\lfloor j/2 \rfloor}(D) a) \right) \psi_j(D) g.$$

The desired bound for $R_a g$ follows from Lemma 7.4, and for the second term it follows by orthogonality and the bound

$$\|a - \phi_{\lfloor j/2 \rfloor}(D) a\|_{L^\infty} \leq C \min \left(2^{-\frac{j}{2}} \|a\|_{C^{0,1}}, 2^{-j} \|a\|_{C^{1,1}} \right).$$

□

Corollary 7.6. Suppose $a \in C^{1,1}(\mathbb{R}^n)$, and $q \in S^1(\mathbb{R}^n)$. If $0 \leq \sigma \leq 1$, then

$$\|\langle D \rangle^\sigma R_{[a,q(D)]} (\langle D \rangle^{1-\sigma} g)\|_{L^2} \lesssim \|a\|_{C^{1,1}} \|g\|_{L^2},$$

Proof. We note that $R_{[a,q(D)]} = [R_a, q(D)]$. The estimate then follows by Lemma 7.4, since it yields that, for $0 \leq \sigma \leq 1$,

$$\|q(D) \langle D \rangle^\sigma R_a (\langle D \rangle^{1-\sigma} g)\|_{L^2} + \|\langle D \rangle^\sigma R_a q(D) (\langle D \rangle^{1-\sigma} g)\|_{L^2} \lesssim \|a\|_{C^{1,1}} \|g\|_{L^2}.$$

□

We will need an extension of these results involving double commutators.

Lemma 7.7. Suppose that $a \in C^{1,1}(\mathbb{R}^n)$ and $b \in C^{0,1}(\mathbb{R}^n)$, and that $q_0, q_1 \in S^1(\mathbb{R}^n)$ are Fourier multipliers on \mathbb{R}^n . Then the following hold

$$(7.3) \quad \|[a, q_0(D)], q_1(D)g\|_{L^2} \leq C \|a\|_{C^{1,1}} \|g\|_{L^2},$$

$$(7.4) \quad \|[b, [a, q_0(D)]] q_1(D)g\|_{L^2} \leq C \|a\|_{C^{1,1}} \|b\|_{C^{0,1}} \|g\|_{L^2}.$$

Proof. We start with the proof of (7.3). We decompose the multiplication operator a into $T_a + R_a$. By Corollary 7.6,

$$\|[R_a, q_0(D)] q_1(D)g\|_{L^2} + \|q_1(D) [R_a, q_0(D)]g\|_{L^2} \leq C \|a\|_{C^{1,1}} \|g\|_{L^2}.$$

For the term $[T_a, q_0(D)]$, since $\psi_k(D)$ and $\rho_j(D)$ commute with the $q(D)$, and have finite overlap of support, it suffices to prove that, uniformly over j ,

$$\|\psi_j(D) [[a, q_0(D)], q_1(D)] \rho_j(D)g\|_{L^2} \leq C \|g\|_{L^2}.$$

We may then replace $q_0(D)$ and $q_1(D)$ by their dyadic localization to $|\xi| \approx 2^j$, in which case they are represented by convolution kernels $K_{0,j}$ and $K_{1,j}$ for which

$$|K_j(x-y)| \leq C_N 2^{j(n+1)} (1+2^j|x-y|)^{-N}, \quad \forall N.$$

After this substitution we may ignore the factors $\psi_j(D)$ and $\rho_j(D)$. We next expand

$$(7.5) \quad a(x) - a(y) = a'(x)(x - y) + r(x, y)(x - y)^2, \quad \|r(x, y)\|_{L^\infty} \leq C \|a\|_{C^{1,1}}.$$

The integral kernel $r(x, y)(x - y)^2 K_{0,j}(x - y)$ has operator norm $\lesssim 2^{-j} \|a\|_{C^{1,1}}$, whereas $K_{1,j}$ has operator norm $\lesssim 2^j$, hence this contribution to the double commutator is bounded on L^2 . Letting $q'_0(D)$ denote the L^2 -bounded operator with kernel $(x - y)K_{0,j}(x - y)$, the other term yields $[a', q_1(D)] q'_0(D)$, which is bounded on L^2 with norm $\|a\|_{C^{1,1}}$ by a similar argument, or using Theorem 7.1.

To establish (7.4), we first use Corollary 7.6 to see that

$$\|R_a q_0(D)g\|_{L^2} + \|q_0(D) R_a g\|_{L^2} \leq C \|a\|_{C^{1,1}} \|g\|_{H^{-1}}.$$

We may thus replace a by T_a . In the j -th term for $[T_a, q_0(D)]$, we may replace $q_0(D)$ and $q_1(D)$ by their j -th dyadic localization as above. Expanding a as in (7.5), the second-order remainder term leads to a bounded operator. It thus suffices to show that

$$\left\| \left[b, \sum_j \psi_j(D) a' \rho_j(D) q'_0(D) \right] q_1(D) g \right\|_{L^2} \leq C \|a'\|_{C^{0,1}} \|b\|_{C^{0,1}} \|g\|_{L^2},$$

where $q'_0(D)$ is a multiplier of order 0. We may write the commutator on the left hand side as

$$a' [b, q'_0(D)] q_1(D) - [b, R_{a'} q'_0(D)] q_1(D).$$

The first term has the desired bound on L^2 by Corollary 7.2, and the second term has the desired bound by Lemma 7.4. \square

Remark 7.8. The estimate (7.4) can be established with $\|a\|_{C^{0,1}} \|b\|_{C^{0,1}} \|g\|_{L^2}$ on the right hand side. This is the second commutator estimate; see for example [13]. The simpler estimate in (7.4) suffices for our purposes, however.

Corollary 7.9. *Suppose that $q_1(\xi_1) \in S^1(\mathbb{R})$, and $q_0(\xi) \in S^1(\mathbb{R}^n)$. Then, uniformly over k ,*

$$\left\| \left[[a, \rho_k(D) q_0(D)], \phi_k(D_1) q_1(D_1) \right] g \right\|_{L^2} \leq C \|a\|_{C^{1,1}} \|g\|_{L^2}.$$

Proof. Let $q_{0,k} = \rho_k q_0$. As in the proof of (7.3), we write

$$[a, q_{0,k}(D)] = a'(x) q'_{0,k}(D) + r_k, \quad \|r_k g\|_{L^2} \leq C 2^{-k} \|a\|_{C^{1,1}} \|g\|_{L^2}.$$

The operator $\phi_k(D_1) q_1(D_1)$ has norm $\lesssim 2^k$, so $[r_k, \phi_k(D_1) q_1(D_1)]$ is suitably bounded. This leaves the term $[a'(x), \phi_k(D_1) q_1(D_1)] q'_{0,k}(D)$, which is bounded uniformly over k by Theorem 7.1, since $a'(x)$ is a $C^{0,1}$ function of x_1 , uniformly over (x_2, \dots, x_n) , with norm less than $\|a\|_{C^{1,1}}$. \square

Lemma 7.10. *Let $s_0 \in \{0, 1\}$, and Q^\pm be constructed as in Section 2. Then for $c > 0$ sufficiently large, the operator $2P(t_0) + Q^+(t_0) + Q^-(t_0) + c : H^{s_0} \rightarrow H^{s_0-1}$ has a bounded right inverse for each $t_0 \in (-T, T)$. Furthermore, with uniform bounds over $t \in (-T, T)$,*

$$(2P(t_0) + Q^+(t_0) + Q^-(t_0) + c)^{-1} : H^{s-1} \rightarrow H^s, \quad s_0 \leq s \leq s_0 + 1,$$

and the inverse is a continuous function of t_0 into the operator norm topology. Also, $P(t_0) + c : L^2 \rightarrow H^{-1}$ is invertible, and

$$(P(t_0) + c)^{-1} : H^{s-1} \rightarrow H^s, \quad 0 \leq s \leq 2,$$

with norm-continuity of the inverse over $t \in (-T, T)$.

Proof. Consider a fixed value of t_0 , and let $p(x, \xi) = p(t_0, x, \xi)$. We use only $C^{0,1}$ bounds on the symbol $p(x, \xi)$, so all estimates on p in the following proof will be uniform over t and norm-continuous in t . We write $p(x, \xi) = p^\sharp(x, \xi) + p^\flat(x, \xi)$, with

$$p^\sharp(x, \xi) = \sum_{j=0}^{\infty} (\phi_{\lfloor j/2 \rfloor}(D)p)(x, \xi) \psi_j(\xi) \in S_{1, \frac{1}{2}}^1,$$

where the frequency truncation is in the x variable. By Corollary 7.5 and the symbol expansion (7.1),

$$(7.6) \quad \|p^\flat(x, D)g\|_{H^s} \leq C \|g\|_{H^{s+\frac{1}{2}}}, \quad -\frac{1}{2} \leq s \leq 1.$$

Since $|p^\flat(x, \xi)| \leq C(1 + |\xi|)^{\frac{1}{2}}$, it follows that $2p^\sharp(x, \xi) \geq c_1|\xi| - c_0$. For $c > c_0 + 1$, the symbol $(2p^\sharp(x, \xi) + c)^{-1}$ is a bounded family in $S_{1, \frac{1}{2}}^{-1}$, and $c^{\frac{1}{2}}(2p^\sharp(x, \xi) + c)^{-1}$ is a bounded family in $S_{1, \frac{1}{2}}^{-\frac{1}{2}}$. By the pseudodifferential calculus, the composition of the corresponding operator with $2p^\sharp(x, D) + c$ differs from the identity by a pseudodifferential operator that has seminorm bounds in $S_{1, \frac{1}{2}}^0$ of size $c^{-\frac{1}{2}}$. Hence, for c large the operator $2p^\sharp(x, D) + c$ is left and right invertible on each given H^s , and in particular for some c , and all $|s| \leq 2$,

$$\begin{aligned} \|(2p^\sharp(x, D) + c)^{-1}g\|_{H^{s+1}} &\leq C_s \|g\|_{H^s}, \\ \|(2p^\sharp(x, D) + c)^{-1}g\|_{H^{s+\frac{1}{2}}} &\leq C_s c^{-\frac{1}{2}} \|g\|_{H^s}. \end{aligned}$$

Furthermore, the inverse is a pseudodifferential operator of class $S_{1, \frac{1}{2}}^{-1}$.

It follows that, if $s_0 - \frac{1}{2} \leq s \leq s_0$, where $s_0 \in \{0, 1\}$,

$$\begin{aligned} \|(2p^\flat(x, D) + Q^+(t_0) + Q^-(t_0))(2p^\sharp(x, D) + c)^{-1}g\|_{H^s} &\leq C_s \|g\|_{H^{s-\frac{1}{2}}}, \\ \|(2p^\flat(x, D) + Q^+(t_0) + Q^-(t_0))(2p^\sharp(x, D) + c)^{-1}g\|_{H^s} &\leq C_s c^{-\frac{1}{2}} \|g\|_{H^s}. \end{aligned}$$

The operator $2P(t_0) + Q^-(t_0) + Q^+(t_0)$ as a map from $H^s \rightarrow H^{s-1}$ is invertible provided that

$$\sum_{n=0}^{\infty} (2p^\sharp(x, D) + c)^{-1} \left((2p^\flat(x, D) + Q^+(t_0) + Q^-(t_0))(2p^\sharp(x, D) + c)^{-1} \right)^n$$

converges as a map $H^{s-1} \rightarrow H^s$, which by the above is true for $s_0 \leq s \leq s_0 + 1$. Continuity in t of the inverse follows by norm-continuity of $P(t)$ and $Q^\pm(t)$ as functions of t . The same proof works for $0 \leq s \leq 2$ with $Q_\pm(t_0)$ replaced by 0. \square

8. APPENDIX 2: ENERGY ESTIMATES

In this section we establish the energy bounds and well-posedness results we use for L and its factors. Throughout this section, we assume L satisfies the conditions in the introduction, and make use of the equivalent form (6.1). Since we work with space-time mollification of L in this section, we assume that the coefficients of the operator L have been extended to \mathbb{R}^{1+n} as in (1.7), with the same regularity conditions, and that the coefficients of L are constant for $|t| \geq T + 1$. The solution

u is defined only on $(-T, T) \times \mathbb{R}^n$, however, and function space norms of u are with respect to that domain. Recall that $\mathbf{D} = (D_t, D_1, \dots, D_n) = (D_t, D)$.

The following result will be used when obtaining bounds for solutions of L^2 regularity.

Lemma 8.1. *The commutator $[L, \langle D \rangle^{-1}] = \langle D \rangle^{-1} [L, \langle D \rangle] \langle D \rangle^{-1}$ admits an expansion of the form*

$$(8.1) \quad [L, \langle D \rangle^{-1}] u(t) = B_1(t) (\mathbf{D}u)(t) + B_2(t) (\mathbf{D}u)(t),$$

where

$$\|B_1(t)g\|_{L^2} \leq C \|g\|_{H^{-1}}, \quad \|B_2(t)g\|_{H^1} \leq C \alpha(t) \|g\|_{H^{-1}}.$$

Proof. We write $L = D_t^2 - 2D_j b^j D_t - D_i c^{ij} D_j + d^0 D_t + d^j D_j$, after absorbing derivatives of b^j and c^{ij} into d . We use the following commutator bound for functions $c \in C^{0,1}(\mathbb{R}^n)$,

$$\|[c, \langle D \rangle^{-1}]g\|_{H^1} \leq C \|c\|_{C^{0,1}} \|g\|_{H^{-1}},$$

as seen by writing $[c, \langle D \rangle^{-1}] = \langle D \rangle^{-1} [c, \langle D \rangle] \langle D \rangle^{-1}$ and applying Theorem 7.1.

The terms $D_i [c^{ij}, \langle D \rangle^{-1}] D_j$ and $D_j [b^j, \langle D \rangle^{-1}] D_t$ can thus be written as $B_1(t) \mathbf{D}$, and the terms $[d^j, \langle D \rangle^{-1}] D_j$ and $[d^0, \langle D \rangle^{-1}] D_t$ as $B_2(t) \mathbf{D}$. \square

Theorem 8.2. *If $u \in H_{loc}^1$, and $Lu \in L_{loc}^2$, then $WF_2(u) \subseteq \text{char}(L)$. If $u \in C^0 L^2 \cap C^1 H^{-1}$, and $Lu \in L^1 L^2$, then $WF_1(u) \subseteq \text{char}(L)$.*

Proof. The first result relies on only the Lipschitz nature of L . As in Lemma 8.1, we write $L = \mathbf{D}A\mathbf{D}^T + d^0 D_t + d^j D_j$, where A is a $(n+1) \times (n+1)$ matrix function consisting of 1 , b^j and c^{ij} . The terms $d^0 D_t u + d^j D_j u$ belong to L_{loc}^2 by the assumed regularity of u , so we absorb them into F . We decompose multiplication by A into $A = A^\sharp + A^\flat$ as in Corollary 7.5, but where the regularization $\phi_{[j/2]}(\mathbf{D})$ takes place over both the t and x variables.

Let $\Gamma(\tau, \xi) \chi(t, x)$ be supported away from the characteristic set of L , where $\Gamma \in S_c^0$ is a conic cutoff, and $\chi \in C_c^\infty((-T, T) \times \mathbb{R}^n)$. Suppose $u \in H_{loc}^1$, and write

$$\mathbf{D}A^\sharp \mathbf{D}^T \Gamma(\mathbf{D})(\chi u) = [\mathbf{D}A\mathbf{D}^T, \Gamma(\mathbf{D})\chi] u - \mathbf{D}A^\flat \mathbf{D}^T \Gamma(\mathbf{D})(\chi u).$$

The first term on the right belongs to L^2 by Corollary 7.2, and the second to $H^{-\frac{1}{2}}$ by Corollary 7.5. The operator $\mathbf{D}A^\sharp \mathbf{D}^T$ has symbol in $S_{1, \frac{1}{2}}^2$, and is elliptic away from the characteristic set of L , hence $\Gamma(\mathbf{D})\chi u \in H^{\frac{3}{2}}$. Corollary 7.5 now yields that the second term on the right belongs to L^2 , and we conclude $\Gamma(\mathbf{D})(\chi u) \in H^2$.

For the second result we let $v = \langle D \rangle^{-1} u$. By Lemma 8.1, $Lv \in L^2 + L^1 H^1$. By Theorem 8.6 below, there exists $w \in C^0 H^2 \cap C^1 H^1$ so that $L(v - w) \in L^2$. Since $v - w \in H_{loc}^1$, we may apply the preceding result to see that $WF_2(v - w) \subset \text{char}(L)$. On the other hand $\langle D \rangle w \in H_{loc}^1$, so $WF_1(\langle D \rangle w) = \emptyset$. \square

Under a strengthened regularity assumption we can obtain results for L^2 solutions.

Corollary 8.3. *Suppose that $L = \mathbf{D}A\mathbf{D}^T + d^0 D_t + d^j D_j$, where A , d^0 , and d^j belong to $C^{0,1}(\mathbb{R}^{1+n})$. If $u \in L_{loc}^2$ and $Lu \in H_{loc}^{-1}$, then $WF_1(u) \subset \text{char}(L)$.*

Proof. Under the assumptions on d^0 and d^j we have $\mathbf{D}A\mathbf{D}^T u \in H_{loc}^{-1}$. The proof then follows the same steps as for the first part of Theorem 8.2. \square

Theorem 8.4. *Let $s_0 \in \{0, 1\}$. Suppose that $D_t u - p(t, x, D)u \in L^2 H^{s_0}$, and $u \in L^2 H^{s_0}$. If $\Gamma(\tau, \xi)\chi(t, x)$ vanishes on a neighborhood of the characteristic set $\tau = p(t, x, \xi)$, where $\chi \in C_c^\infty((-T, T) \times \mathbb{R}^n)$ and $\Gamma \in S_{cl}^0$, then $\Gamma(\mathbf{D})(\chi u) \in L^2 H^{s_0+1}$. In particular,*

$$WF_{s_0+1}(u) \subset \{\tau = p(t, x, \xi)\} \cup \{\xi = 0\}.$$

Proof. Consider first the case that $\Gamma(\tau, \xi)$ vanishes near the set $\xi = 0$. We write

$$\begin{aligned} (D_t - p^\sharp(t, x, D))\Gamma(\mathbf{D})(\chi u) &= \Gamma(\mathbf{D})\chi(D_t - p(t, x, D))u \\ &\quad + \Gamma(\mathbf{D})(D_t \chi)u - [p, \Gamma(\mathbf{D})\chi]u + p^\flat(t, x, D)\Gamma(\mathbf{D})(\chi u), \end{aligned}$$

where the frequency regularization defining p^\sharp takes place over both t and x variables. The first two terms on the right belong to $H^{s_0}(\mathbb{R}^{1+n})$, where since Γ vanishes near $\xi = 0$ we have $\Gamma(\mathbf{D}) : L^2 H^{s_0}(\mathbb{R}^{1+n}) \rightarrow H^{s_0}(\mathbb{R}^{1+n})$, and similarly the last term belongs to $H^{s_0-\frac{1}{2}}(\mathbb{R}^{1+n})$ by Corollary 7.5. To see that the third term also belongs to H^{s_0} , we take the symbol expansion (7.1) to replace p by $a(t, x)q(D)$. The commutator of $q(D)$ and χ is bounded on H^{s_0} , so we check that

$$\|[a, \Gamma(\mathbf{D})]\langle D \rangle g\|_{H^{s_0}} \leq C \|a\|_{C^{0,1}} \|g\|_{L^2 H^{s_0}}.$$

This follows from Corollary 7.2, since $\|\langle D \rangle g\|_{H^{s_0-1}} \leq \|g\|_{L^2 H^{s_0}}$ for $s_0 \leq 1$.

The symbol $\tau - p^\sharp(t, x, \xi)$ has a microlocal parametrix of class $S_{1, \frac{1}{2}}^{-1}$ away from the set $\{\xi = 0\} \cup \{\tau = p(t, x, \xi)\}$, and the result follows as in Theorem 8.2.

Suppose then that Γ and $\tilde{\Gamma}$ are supported in a small cone about the τ axis, vanish near $\tau = 0$, with $\tilde{\Gamma}\Gamma = \Gamma$. We write

$$\begin{aligned} (I - p(t, x, D)D_t^{-1}\tilde{\Gamma}(\mathbf{D}))D_t\Gamma(\mathbf{D})\chi u &= \Gamma(\mathbf{D})\chi(D_t - p(t, x, D))u \\ &\quad + \Gamma(\mathbf{D})(D_t \chi)u - [p, \Gamma(\mathbf{D})\chi]u. \end{aligned}$$

The right hand side belongs to $L^2 H^{s_0}$ by steps similar to above.

The operator $p(t, x, D)D_t^{-1}\tilde{\Gamma}(\mathbf{D})$ is of small norm on $L^2 H^s$ for $-1 \leq s \leq 1$, as seen by the symbol expansion (7.1), since $|\xi| \ll \tau$ on the support of $\tilde{\Gamma}$. We conclude that $D_t\Gamma(\mathbf{D})\chi u \in L^2 H^{s_0}$, and hence that $\Gamma(\mathbf{D})\chi u \in L^2 H^{s_0+1}$. \square

Theorem 8.5. *Let $s_0 \in \{0, 1\}$, assume that $p(t, x, \xi)$ satisfies (1.9), and that $Q(t)$ satisfies (2.10). Let $E(t, t_0)$ be the wave group of Theorem 3.2. Suppose that, in the sense of distributions,*

$$(8.2) \quad D_t u - p(t, x, D)u - Q(t)u = F,$$

and $u \in L^2 H^{s_0}$, $F \in L^1 H^{s_0}$. Then $u \in C^0 H^{s_0}$, and for each $t_0 \in (-T, T)$

$$(8.3) \quad u(t) = E(t, t_0)u(t_0) + \int_{t_0}^t E(t, s)F(s)ds.$$

In particular if $u \in L^2 H^{s_0+1}$ and $F \in L^1 H^{s_0+1}$, then $u \in C^0 H^{s_0+1}$.

Proof. We start by proving uniqueness for the equation

$$(8.4) \quad D_t u - p(t, x, D)u = G, \quad \text{supp}(u) \subset \{t > -T + \delta\}, \quad \delta > 0,$$

under the condition $u \in L^2$ and $G \in L^1L^2$, and $T < \infty$. This will show that, for such u ,

$$(8.5) \quad u(t) = \int_{-T}^t E_0(t, s) G(s) ds,$$

where $E_0(t, t_0)$ is as in Theorem 3.2.

Suppose first that $u \in C^\infty((-T, T) \times \mathbb{R}^n)$, and is supported where $t > -T + \delta$ and $|x| \leq R$, for some $\delta > 0$ and $R < \infty$. Let u satisfy (8.4). We calculate

$$\partial_t \int |u(t, x)|^2 dx = -2 \operatorname{Im} \int \overline{u(t)} p(t, x, D) u(t) dx - 2 \operatorname{Im} \int \overline{u(t)} G(t) dx.$$

Since $p(t, x, \xi)$ is real, it follows by Theorem 7.1 and the symbol expansion (7.1) that, uniformly over t ,

$$\|p(t, x, D)^* u(t) - p(t, x, D) u(t)\|_{L^2} \leq C \|u(t)\|_{L^2},$$

hence

$$\partial_t \|u(t)\|_{L^2}^2 \leq 2C \|u(t)\|_{L^2}^2 + 2 \|G(t)\|_{L^2} \|u(t)\|_{L^2}.$$

By the Gronwall inequality

$$\|u(t)\|_{L^2} \leq e^{C(t+T)} \int_{-T}^t \|G(s)\|_{L^2} ds,$$

and in particular

$$(8.6) \quad \|u\|_{L^2((-T, T) \times \mathbb{R}^n)} \leq C_T \|G\|_{L^1L^2((-T, T) \times \mathbb{R}^n)}.$$

Suppose now that $u \in L^2((-T, T) \times \mathbb{R}^n)$ satisfies (8.4). We choose $\chi \in C_c^\infty(\mathbb{R}^{1+n})$ supported in $t > 0$ satisfying $\widehat{\chi}(0) = 1$, and $\phi \in C_c^\infty(\mathbb{R}^{1+n})$ satisfying $\phi(0) = 1$. Let J_ϵ denote the family of causal, compactly supported mollifiers

$$J_\epsilon u = \phi(\epsilon^{-1}(t, x)) \widehat{\chi}(\epsilon \mathbf{D}) u.$$

Note that J_ϵ is a uniformly bounded family of pseudodifferential operators of class $S_{1,0}^0$. By Theorem 7.1, the following holds uniformly over $\epsilon > 0$

$$\|[D_t - p(t, x, D), J_\epsilon] u\|_{L^2} \leq C \|u\|_{L^2}.$$

Since $J_\epsilon \rightarrow I$ strongly on both H^1 and L^2 , as well as L^1L^2 , then by density of $H^1 \subset L^2$ it follows that

$$\lim_{\epsilon \rightarrow 0} \|[D_t - p(t, x, D), J_\epsilon] u\|_{L^2((-T, T) \times \mathbb{R}^n)} = 0.$$

It follows that (8.6) holds for general $u \in L^2$ and $G \in L^1L^2$ under the condition (8.4), yielding uniqueness of the solution, and thus the identity (8.5).

Now suppose that $u \in L^2H^{s_0}$ is supported in $t > -T + \delta$, and satisfies (8.2) with $F = 0$. Since $Q(t)$ and $E_0(t, s)$ are uniformly bounded on H^{s_0} , taking $G(t) = Q(t)u(t)$ in (8.5) we see that

$$\|u\|_{L^2((-T, -T+c), H^{s_0})} \leq C c^{\frac{1}{2}} \|u\|_{L^2((-T, -T+c), H^{s_0})},$$

and by a continuation argument we must have $u \equiv 0$. Hence we have uniqueness for (8.2) for solutions supported in $t > -T + \delta$.

Thus, if $\psi(t) \in C^\infty(\mathbb{R})$ is supported in $t > -T$, and $u \in L^2 H^{s_0}$ satisfies (8.2), then

$$(8.7) \quad \psi(t)u(t) = \int_{-T}^t E(t,s) \left((D_s \psi)(s)u(s) + \psi(s)F(s) \right) ds,$$

since the right hand side is a solution belonging to $C((-T, T), H^{s_0})$. Reversing time, we obtain the following bound for solutions to (8.2) without restrictions on the time-support of u ,

$$\|u\|_{C((-T, T), H^{s_0})} \leq C_T \left(\|u\|_{L^2 H^{s_0}} + \|F\|_{L^1 H^{s_0}} \right).$$

In particular $u(t_0)$ is well defined in H^{s_0} for each $t_0 \in (-T, T)$. Now let ψ be an increasing function in $C^\infty(\mathbb{R})$, which vanishes for $t < t_0 - \epsilon$ and equals 1 for $t > t_0 + \epsilon$. Letting $\epsilon \rightarrow 0$, the formula (8.7) shows that, for $t > t_0$,

$$(8.8) \quad u(t) = E(t, t_0)u(t_0) + \int_{t_0}^t E(t, s) F(s) ds,$$

and by time reversal this holds for all t , which establishes (8.3).

Finally, if $u \in L^2 H^{s_0+1}$ and $F \in L^1 H^{s_0+1}$, then (8.7) necessarily holds, and since $E(t, s)$ is a strongly-continuous evolution group on H^{s_0+1} , the same steps above show that $u \in C^0 H^{s_0+1}$, and that (8.8) holds. \square

Theorem 8.6. *Given $t_0 \in (-T, T)$, and $u_0 \in L^2$, $u_1 \in H^{-1}$, $F \in L^1 H^{-1}$, there exists a unique solution $u \in C^0 L^2 \cap C^1 H^{-1}$ to the Cauchy problem*

$$Lu = F, \quad u(t_0) = u_0, \quad D_t u(t_0) = u_1.$$

If $0 \leq s \leq 2$, and if $u_0 \in H^s$, $u_1 \in H^{s-1}$, $F \in L^1 H^{s-1}$, then the solution satisfies $u \in C^0 H^s \cap C^1 H^{s-1}$.

Also, if $u \in H^1((-T, T) \times \mathbb{R}^n)$ satisfies $Lu \in L^1 L^2$, then $u \in C^0 H^1 \cap C^1 L^2$.

Proof. We start by proving the existence of such a solution to the Cauchy problem. Assume $0 \leq s \leq 2$ and $(u_0, u_1) \in H^s \times H^{s-1}$. We seek a solution u of the form

$$(8.9) \quad u(t) = \sum_{\pm} E_{0,\pm}(t, t_0) f_{\pm} + \int_{t_0}^t (E_{0,+} - E_{0,-})(t, s) (2P(s))^{-1} G(s) ds.$$

Here $E_{0,\pm}$ is the wave group (3.2) for $D_t \mp P^\pm$, we take $G \in L^1 H^{s-1}$, and set

$$f_{\pm} = (2P(t_0))^{-1} (P^\mp(t_0)u_0 \pm u_1) \in H^s,$$

that last inclusion holding by Lemma 7.10. Recall that $P^+ + P^- = 2P$.

Applying L and using (2.2) and (2.16), the equation $Lu = F$ reduces to

$$\begin{aligned} G(t) + \int_{t_0}^t \left(R_1^+(t)E_{0,+}(t, s) - R_1^-(t)E_{0,-}(t, s) \right) (2P(s))^{-1} G(s) ds \\ = F(t) - \sum_{\pm} R_1^\pm(t)E_{0,\pm}(t, t_0) f_{\pm}. \end{aligned}$$

By Theorem 3.2 and (2.5) the right hand side belongs to $L^1 H^{s-1}$. Also, by Lemma 7.10,

$$\| (R_1^+(t)E_{0,+}(t, s) - R_1^-(t)E_{0,-}(t, s)) (2P(s))^{-1} G(s) \|_{H^{s-1}} \leq \alpha(t) \|G(s)\|_{H^{s-1}},$$

so that the Volterra equation for G is uniquely solvable on $L^1 H^{s-1}$, with solution given by a convergent expansion. Note that if $\|F(t)\|_{H^{s-1}} \leq C \alpha(t)$ then the same holds for G . That $u \in C^0 H^s \cap C^1 H^{s-1}$ follows by (8.9) and Theorem 3.2.

We now consider uniqueness. Suppose first that $u \in C^1((-T, T), C^2(\mathbb{R}^n))$ satisfies $Lu = F \in L^1 C^0$, and assume that u is supported in $|x| \leq R$, some $R < \infty$. It follows from $Lu = F$ that $D_t^2 u \in L^1 C^0$. Using integration by parts we calculate

$$\begin{aligned} & \partial_t \int \left(|D_t u(t, x)|^2 + \sum_{i,j=1}^n c^{ij}(t, x) D_i u(t, x) \overline{D_j u(t, x)} + |u(t, x)|^2 \right) dx \\ &= 2i \operatorname{Im} \int F(t, x) \overline{D_t u(t, x)} dx + \int (B(t, x) (u, \mathbf{D}u)(t, x)) \cdot \overline{(u, \mathbf{D}u)(t, x)} dx, \end{aligned}$$

where $B(t, x)$ is a $(n+2) \times (n+2)$ matrix whose coefficients consist of first-order space-time derivatives of the coefficients b^j, c^{ij} , as well as d^0 and d^j , hence

$$\|B^{ij}(t)\|_{C^0} \leq C, \quad \|\mathbf{D}B^{ij}(t)\|_{L^\infty} \leq C \alpha(t).$$

By the positive definite condition on c^{ij} and the Gronwall inequality, we conclude

$$(8.10) \quad \|(u, \mathbf{D}u)(t)\|_{L^2} \leq C e^{\int_{t_0}^t \alpha} \|(u, \mathbf{D}u)(t_0)\|_{L^2} + \int_{t_0}^t e^{\int_s^t \alpha} \|F(s)\|_{L^2} ds.$$

By mollification and truncation with respect to the x -variable, the bound (8.10) holds under the assumption $u \in C^0 H^1 \cap C^1 L^2$ and $F \in L^1 L^2$.

Suppose that $u \in C^0 L^2 \cap C^1 H^{-1}$ satisfies $Lu = F \in L^1 H^{-1}$. By Lemma 8.1, $v = \langle D \rangle^{-1} u$ satisfies

$$Lv = B_0(t)(\mathbf{D}v)(t) + \langle D \rangle^{-1} F(t), \quad \|B_0(t)g\|_{L^2} \leq C \alpha(t) \|g\|_{L^2}.$$

The Gronwall inequality and (8.10) applied to v then imply

$$(8.11) \quad \|(u, \mathbf{D}u)(t)\|_{H^{-1}} \leq C e^{\int_{t_0}^t \alpha} \|(u, \mathbf{D}u)(t_0)\|_{H^{-1}} + \int_{t_0}^t e^{\int_s^t \alpha} \|F(s)\|_{H^{-1}} ds,$$

from which the desired uniqueness follows.

To complete the proof of Theorem 8.6, suppose now that $u \in H^1((-T, T) \times \mathbb{R}^n)$ satisfies $Lu = 0$. We want to show that

$$(8.12) \quad u \in C^0 H^1 \cap C^1 L^2,$$

since by the above the inhomogeneous problems admits a solution of this regularity.

We consider $\psi(t)u$ as in the proof of Theorem 8.5, where $\psi = 0$ near either $\pm T$, and easily verify that $L(\psi(t)u(t)) \in L^2((-T, T) \times \mathbb{R}^n)$. By the above there exists a solution of regularity (8.12) with inhomogeneity $L(\psi(t)u)$, which also vanishes for t near the chosen $\pm T$, hence it suffices to prove that if $u \in H^1((-T, T) \times \mathbb{R}^n)$ satisfies $Lu = 0$, with $u = 0$ for t near either $\pm T$, then $u \equiv 0$. For this we note that (8.10) implies

$$\|u\|_{H^1((-T, T) \times \mathbb{R}^n)} \leq C \|F\|_{L^2((-T, T) \times \mathbb{R}^n)},$$

if u satisfies (8.12) and vanishes near either $\pm T$. The same inequality holds if $u \in H^1$ and $F \in L^2$, as seen by using the space-time mollifiers J_ϵ from Theorem 8.5, and noting that $[L, J_\epsilon]$ maps H^1 to L^2 , uniformly in ϵ . \square

Remark 8.7. Theorem 8.6 together with finite propagation velocity shows that Theorem 1.1 holds for solutions on an open set; that is, it holds for solutions to $Lu = 0$ with $u \in C^0((-T, T), H_{loc}^{s_0}(\Omega))$, $D_t u \in C^0((-T, T), H_{loc}^{s_0-1}(\Omega))$, for an open set $\Omega \subset \mathbb{R}^n$, as long as γ remains above Ω . To see this, assume that $|D_\xi p_\pm| \leq C$. If $\chi \in C_c^\infty(\Omega)$ equals 1 on a ball of radius $2r$ about the spatial projection of $\gamma(t_0)$, then by finite propagation velocity u agrees on the ball of radius r and $|t - t_0| \leq C^{-1}r$ with the solution \tilde{u} on $(-T, T) \times \mathbb{R}^n$ to $L\tilde{u} = 0$, where \tilde{u} has Cauchy data $(\chi u(t_0), \chi D_t u(t_0))$ at $t = t_0$. By Theorem 8.6, $\tilde{u} \in C^0 H^{s_0} \cap C^1 H^{s_0-1}$, and Theorem 1.1 implies that $\gamma(t) \notin WF_{s_0+1}(u)$ for $|t - t_0| \leq C^{-1}r$. The argument may then be repeated starting at $t_0 \pm C^{-1}r$. \square

Remark 8.8. Under increased regularity of the coefficients, solutions $u \in L^2$ to $Lu = 0$ satisfy (1.1). Suppose $L = D_t^2 - 2D_j b^j D_t - D_i c^{ij} D_j + d^0 D_t + d^j D_j$, where b^j and c^{ij} satisfy (1.3) and (1.4) as before, but we make the stronger assumption that $d^0, d^j \in C^{0,1}((-T, T) \times \mathbb{R}^n)$. Suppose that $Lu \in L^2 H^{-1}$. By Corollary 8.3 we then have $\mathbf{D}u \in L_{loc}^2 H^{-1}$, hence $L(\psi(t)u) \in L^2 H^{-1}$ for $\psi \in C_c^\infty((-T, T))$.

Suppose now that $u \in L^2((-T, T) \times \mathbb{R}^n)$ satisfies $Lu = 0$. By Theorem 8.6, there is a solution $v \in C^0 L^2 \cap C^1 H^{-1}$ satisfying $Lv = L(\psi(t)u)$, hence to prove $u \in C^0 L^2 \cap C^1 H^{-1}$ it suffices to prove uniqueness of L^2 solutions to $Lu = F \in L^2 H^{-1}$ when u is supported in $|t| \leq T - \delta$.

Fix $\chi \in C_c^\infty(\mathbb{R}^{1+n})$ with $\hat{\chi}(0) = 1$. Then the commutator $[L, \hat{\chi}(\epsilon \mathbf{D})]$ maps L^2 to $L^2 H^{-1}$, uniformly in ϵ . This follows from Corollary 7.2 by the $C^{0,1}$ regularity of the coefficients, and the form of L ; in particular the D_t^2 term commutes with $\hat{\chi}(\epsilon \mathbf{D})$. By a density argument,

$$\lim_{\epsilon \rightarrow 0} \|[L, \hat{\chi}(\epsilon \mathbf{D})]u\|_{L^2 H^{-1}} = 0, \quad u \in L^2.$$

It follows from (8.11) that

$$\|u\|_{L^2((-T, T) \times \mathbb{R}^n)} \leq C \|F\|_{L^1 H^{-1}},$$

from which the uniqueness of solutions follows. \square

REFERENCES

- [1] M. Beals and M. Reed, *Microlocal regularity theorems for non-smooth pseudodifferential operators and applications to nonlinear problems*, Trans. AMS **285** (1984), 159–184.
- [2] J. M. Bony, *Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles nonlinéaires*, Ann. Sci. Ecole Norm. Sup. **14** (1981), 209–246.
- [3] N. Burq, *Contrôlabilité exacte des ondes dans des ouverts peu réguliers*, Asymptot. Anal. **14** (1997), no. 2, 157–191.
- [4] A. Calderón, *Commutators of singular integral operators*, Proc. Nat. Acad. Sci. U.S.A. **53** (1965), 1092–1099.
- [5] L. Carleson, *Interpolation by bounded analytic functions and the corona problem*, Ann. Math. **76** (1962), 547–559.
- [6] R. Coifman and Y. Meyer, *Au delà des opérateurs pseudo-différentiels*, Astérisque #57, Société Math. de France, 1978.
- [7] M. de Hoop, G. Uhlmann, and A. Vasy, *Diffraction from conormal singularities*, Ann. Sci. Ecole Norm. Sup., to appear.
- [8] C. Fefferman and E. Stein, *H^p spaces of several variables*, Acta Math. **129** (1972), 137–193.
- [9] L. Hörmander, *On the existence and the regularity of solutions of linear pseudo-differential equations*, Enseignement Math. (2), **17** (1971), 99–163.
- [10] M. Rauch and M. Reed, *Propagation of singularities for semilinear hyperbolic equations in one space variable*, Ann. Math. **111** (1980), 531–552.

- [11] H. Smith, *A parametrix construction for wave equations with $C^{1,1}$ coefficients*, Ann. Inst. Fourier, Grenoble **48** 3 (1998), 797–835.
- [12] ———, *Spectral cluster estimates for $C^{1,1}$ metrics*, Amer. Jour. Math. **128** (2006), 1069–1103.
- [13] E. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*. With the assistance of Timothy S. Murphy. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993.
- [14] D. Tataru, *Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. III*, J. Amer. Math. Soc. **15** (2002), no. 2, 419–442.
- [15] M. Taylor, *Tools for PDE: Pseudodifferential Operators, Paradifferential Operators, and Layer Potentials*. Mathematical Surveys and Monographs, no. 81, American Mathematical Society, Providence, 2000.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195

E-mail address: hart@math.washington.edu