SPECTRAL CLUSTER ESTIMATES FOR $C^{1,1}$ METRICS

HART F. SMITH

Abstract. In this paper, we establish L^p norm bounds for spectral clusters on compact manifolds, under the assumption that the metric is $C^{1,1}$. Precisely, we show that the L^p estimates proven by Sogge [12] in the case of smooth metrics hold under this limited regularity assumption. It is known by examples of Smith-Sogge [11] that such estimates fail for $C^{1,\alpha}$ metrics if $\alpha < 1$.

1. Introduction

Let M be a compact manifold without boundary, of dimension $n \geq 2$, and let P be an elliptic differential operator of second order on M which annihilates the constant function. Assume that P is self-adjoint and non-positive with respect to some density $d\mu$, and denote the eigenvalues of P by $\{-\lambda_j^2\}$. Let Π_λ be the projection of $L^2(d\mu)$ onto the subspace spanned by the eigenfunctions for which $\lambda_j \in [\lambda, \lambda+1]$.

In the case that the coefficients of P are smooth, Sogge [12] established the following bounds for the $L^2 \to L^q$ operator norm of Π_λ:

$$\left\|\Pi_\lambda f\right\|_{L^q(M)} \leq C\,\lambda^{\frac{n-1}{2}(\frac{1}{2}-\frac{1}{q})}\,\|f\|_{L^2(M)}\,, \qquad 2 \leq q \leq \tfrac{2(n+1)}{n-1}\,. \tag{1.1}$$

$$\left\|\Pi_\lambda f\right\|_{L^q(M)} \leq C\,\lambda^{n(\frac{1}{2}-\frac{1}{q})-\frac{1}{2}}\,\|f\|_{L^2(M)}\,, \qquad \tfrac{2(n+1)}{n-1} \leq q \leq \infty\,, \tag{1.2}$$

Both exponents can be shown to be best possible for each q.

In this paper we establish these estimates under the assumption that the coefficients of P are of regularity $C^{1,1}$ (the first derivatives of the coefficient functions are Lipschitz.) In Smith-Sogge [11], for each $\alpha < 1$ examples of operators with coefficients of regularity $C^{1,\alpha}$ were constructed for which estimate (1.1) fails for $q \neq 2$, showing that the regularity $C^{1,1}$ is sharp among the Hölder classes.

We begin by making our assumptions on P more precise. We assume that M is covered by a finite collection of coordinate patches, and that in the induced local coordinates

$$\bigl(Pf\bigr)(x) = \rho(x)^{-1}\sum_{i,j=1}^{n}\partial_i\Bigl(\rho(x)\,\mathrm{g}^{ij}(x)\,\partial_j f(x)\Bigr)\,, \qquad d\mu = \rho(x)\,dx\,. \tag{1.3}$$

Our assumption is that $\rho(x)$ and each $\mathrm{g}^{ij}(x)$ is the restriction to the patch of a function in $C^{1,1}(\mathbb{R}^n)$, with $\rho(x)$ and the matrix valued function $\mathrm{g}(x)$ both real and uniformly positive.

The author was supported in part by NSF grant DMS-0140499.

Theorem 1.1. *Let M and P be as above. Then* (1.1) *and* (1.2) *hold for* $n \geq 2$.

Theorem 1.1 will follow as a corollary of the square function estimates for the Cauchy problem stated in Theorem 1.2 below. Since the square function estimates involve Sobolev spaces of index equal to the exponent of λ in (1.2), the limited regularity of the coefficients places a restriction on how large q can be in Theorem 1.2. In particular, we require $\delta(q) \leq 2$, where

$$\delta(q) = n\Big(\,\frac{1}{2} - \frac{1}{q}\,\Big) - \frac{1}{2}\,.$$

However, as noted below the estimates (1.1) and (1.2) for the full range of q follow by establishing them at the critical exponent $q_n = \frac{2(n+1)}{n-1}$, and Theorem 1.2 applies since $\delta(q_n) \leq 1$ in all dimensions.

We define Sobolev spaces on M by the rule

$$\|f\|_{H^\delta(M)} = \sum_{j=0}^{\infty} \big(\,1+\lambda_j^2\big)^{\delta}\,|c_j|^2\,, \qquad f = \sum_{j=0}^{\infty} c_j\,\psi_j\,,$$

where $P\psi_j = -\lambda_j^2\psi_j$ is the orthonormal eigenbasis. By elliptic regularity (see, for example, Theorem 8.10 of [4]), this norm is equivalent to that defined by the standard Sobolev space in local coordinates, provided $0 \leq \delta \leq 3$. (This requires that the coordinate patches on M be of regularity C^3. In our applications we consider only $0 \leq \delta \leq 2$, so that C^2 coordinate patches are sufficient for what follows.)

Theorem 1.2. *Let P and M be as above, and suppose that $\{\psi_j\}$ is the eigenbasis for P, with eigenvalues $\{-\lambda_j^2\}$. Then, with*

$$u(t,x) \;= \sum_{j=0}^{\infty} c_j \cos(t\lambda_j)\,\psi_j(x)\,, \qquad\qquad f(x) \;= \sum_{j=0}^{\infty} c_j\,\psi_j(x)\,,$$

the following estimate holds, provided $\frac{2(n+1)}{n-1} \leq q \leq \infty$ *and* $\delta(q) \leq 2$,

$$\|u\|_{L^q_x L^2_t(M\times[-1,1])} \leq C\,\|f\|_{H^{\delta(q)}(M)} \tag{1.4}$$

To see that Theorem 1.1 follows from Theorem 1.2, we first consider estimate (1.2) for q such that $\delta(q) \leq 2$. Observe that

$$\Pi_\lambda f = \int_{-1}^{1} e^{-it\lambda} \sum_{j:\lambda_j\in[\lambda,\lambda+1]} \tilde{c}_j \cos(t\lambda_j)\,\psi_j\,,$$

where

$$\tilde{c}_j = \left[\,\frac{\sin(\lambda_j-\lambda)}{\lambda_j-\lambda} + \frac{\sin(\lambda_j+\lambda)}{\lambda_j+\lambda}\,\right]^{-1} c_j\,.$$

For $\lambda_j \in [\lambda, \lambda+1]$ the term in brackets lies in the range $[\frac{1}{2}, 2]$, and the $\tilde{c}_j$ are thus coefficients of a function of comparable $H^{\delta(q)}$ norm. The bound (1.2) for $\delta(q) \leq 2$ then follows immediately from (1.4). Since $\delta(q_n) = q_n^{-1} < \frac{1}{2}$, the case $q = q_n$ holds for (1.2), hence for (1.1). Estimate (1.1) follows on its full range by interpolation of $q = q_n$ and $q = 2$.

The condition $\delta(q) \leq 2$ fails for large q if $n \geq 6$. To obtain (1.2) for all n and q we apply heat kernel methods. We first note that by interpolation it suffices to consider the case $q=\infty$. Let $H_\lambda = \exp(\lambda^{-2}P)$ denote the heat kernel at time $\lambda^{-2}\le 1$. By Markov semigroup results of Varopoulos [16], specifically the Theorem of section 7, page 259, it holds for $n>2$ that

$$\|H_\lambda g\|_{L^\infty(M)} \le C\,\lambda^n\,\|g\|_{L^1(M)}\,.$$

For this we need to check the inequality

$$\|g\|_{L^{\frac{2n}{n-2}}(M)} \le C\left((-Pg,g)^{\frac12} + \|g\|_{L^2(M)}\right),$$

which holds since the right hand side is equivalent to $\|f\|_{H^1(M)}$. Interpolation with the bound

$$\|H_\lambda g\|_{L^\infty(M)} \le \|g\|_{L^\infty(M)}$$

yields that

$$\|H_\lambda g\|_{L^\infty(M)} \le C\,\lambda^{\frac{n}{q}}\,\|g\|_{L^q(M)}\,.$$

If $\lambda_j\in[\lambda,\lambda+1]$ then $\exp(-\lambda_j^2/\lambda^2)$ is bounded away from 0, and we may write

$$\Pi_\lambda f = H_\lambda \Pi_\lambda \tilde f\,,\qquad \|\tilde f\|_{L^2(M)}\approx \|f\|_{L^2(M)}\,.$$

The case $q=\infty$ of (1.2) thus follows from the case $q=q_n$, which as noted above follows from (1.4) in all dimensions.

Theorem 1.2 has the advantage of being localizable to a coordinate patch by finite propagation velocity, after observing that u satisfies the Cauchy problem for ∂_t^2-P. By covering $M\times[-1,1]$ by a finite collection of sets, and scaling the coordinate patches, we may assume that in each local coordinate patch P is arbitrarily close to the standard Laplace operator, and thus reduce Theorem 1.2 to the following result on $\mathbb{R}^n$.

Theorem 1.3. *Suppose that the operator P takes the form* (1.3)*, where for $|x|\ge\frac34$ we have $\mathrm{g}^{ij}(x)=\delta^{ij}$ and $\rho(x)=1$. Assume also that*

$$\sum_{i,j=1}^n \|\mathrm{g}^{ij}-\delta^{ij}\|_{C^{1,1}(\mathbb{R}^n)} + \|\rho-1\|_{C^{1,1}(\mathbb{R}^n)} \le c_0\,,$$

where c_0 is a small number to be fixed. Then, if $u(t,x)$ solves the Cauchy problem

$$\partial_t^2 u(t,x) = Pu(t,x)\,,\qquad u(0,x)=f(x)\,,\quad \partial_t u(0,x)=g(x)\,,$$

the following estimate holds, provided $q_n\le q\le\infty$ and $\delta(q)\le 2$,

$$\|u\|_{L^q_xL^2_t(\mathbb{R}^n\times[-1,1])} \le C\left(\|f\|_{H^{\delta(q)}} + \|g\|_{H^{\delta(q)-1}}\right).$$

The version of Theorem 1.3 for smooth coefficients, without the restriction on $\delta(q)$, was established by Mockenhoupt-Seeger-Sogge [8]. The result of Theorem 1.3 is closely related to the Strichartz estimates, which bound $L^p_tL^q_x$ norms of u. The Strichartz estimates were proven in the setting of $C^{1,1}$ coefficients by the author in [10] for dimensions $n=2,3$, and subsequently by Tataru [14] in all dimensions. We remark that the results of this paper, in particular the estimate (3.15), yield an alternate proof of the homogeneous Strichartz estimates, in the more general time-dependent setting considered by Tataru [14]. We provide the details in the last section.

This paper follows the same initial steps as [10] and [14] by first reducing the estimates, via a Littlewood-Paley decomposition, to the case of u localized in frequency to a dyadic region. One then mollifies the coefficients of P at the scale of the square root of the given frequency, thus reducing the problem to proving uniform estimates on dyadic frequency scales for a family of smooth metrics.

The difference between this paper and the above works is in the control of solutions for the mollified equation. The proofs of the Strichartz estimates essentially are established through dispersive estimates; that is, through the time decay of an approximate fundamental solution to the wave equation. Squarefunction estimates, on the other hand, depend on spatial decay of the approximate fundamental solution away from the light cone. The complication is that, for small t, the smoothing procedure applied to the coefficients involves averaging over a spatial scale larger than the decay scale. To get around this problem, we use a parametrix construction that behaves well under time-space dilation. The desired decay properties of the parametrix for small t are then obtained by scaling from the spatial decay estimates on the kernel at $t = 1$.

A rough outline of the proof is as follows. We begin by reducing Theorem 1.3 to establishing estimates for a first-order equation, essentially the half-wave equation, with both the solution and the coefficients of the equation appropriately localized in frequency. We next apply a continuous wave-packet transform in the x-variables to the solution $u(t, x)$ of the first-order equation, lifting u to a function $v(t, x, \xi)$. The wave-packet transform is at the right scale so that the lift of the half-wave operator applied to u equals the derivative of v along the Hamiltonian flow, modulo an error term bounded in energy. This error term can be absorbed into the driving force, and thus $v(t, x, \xi)$ satisfies an inhomogeneous first-order differential equation. By variation of parameters, we are then reduced to establishing estimates for functions whose lifts are constant along the Hamiltonian flow. The Hamiltonian flow generates a unitary transform on $L^2(dx\, d\xi)$, which lets the crucial "$W_t W_s^* = W_{t-s}$" step go through. We then scale $t - s$ to 1, and show that the rescaled kernel has the desired decay estimates away from the light cone.

Our use of a continuous wave-packet transform is inspired by Tataru's proof of the Strichartz estimates [14]. The main difference is that we here take a wave-packet transform in the x-variable only, rather than in (t, x). We also use a Schwartz function of compact support in ξ as the basis for the transform, as opposed to the Gaussian based FBI transform used in [14], which has the benefit that the transform of a function u is compactly supported in ξ if the Fourier transform of u is. This is important for intended future applications, since as a result one has to deal with the Hamiltonian flow only on a localized set in ξ. For more information on the Fourier-Bros-Iagolnitzer transform, and its applications for wave equations, we refer to [3] and the references therein. The particular scale of FBI transform used here coincides with the Córdoba-Fefferman wave-packet transform, [2].

The outline of this paper is as follows. After generalizing the class of operators we consider, section 2 is concerned with reducing Theorem 1.3 to Theorem 2.5, which involves the half wave equation localized to a dyadic frequency shell. In section 3 we introduce the wave-packet transform used to analyze the solution u. The main result is Lemma 3.3, relating the half-wave operator to differentiation along the Hamiltonian flow. We then reduce matters to proving pointwise estimates for an integral kernel arising from the Hamiltonian flow conjugated by the wave-packet transform. In section 4, we establish the estimates on the Hamiltonian flow necessary to establish the desired decay estimates on the kernel. The main result is that, after rescaling $t - s$ to 1, the kernel of interest has the same decay off of the light cone as one would have for the smooth case. In section 5 we show how these techniques also yield a proof of the homogeneous Strichartz estimates, in the setting of Tataru's result [14].

Notation. We write $a \lesssim b$ to mean that $a \le C\, b$ where C is an allowable constant, in that it depends only on universal quantities (such as the dimension n). We also allow C to depend on the integer N in case of decay statements involving arbitrary N, and α and β in case of estimates involving arbitrary multi-indices α and β. We say that $a \approx b$ if $a \lesssim b$ and $b \lesssim a$. We write $a \ll b$ if $a \le c\, b$, where c is a constant that can be taken arbitrarily small by making c_0 in (2.2) below small.

We use d_x to denote the gradient operation that takes scalar functions of x to vector fields, and vector fields to matrix valued functions. We also use $d_x^k f(x)$ to denote a generic derivative of $f(x)$ of order exactly k. We use d to denote the space-time gradient (d_t, d_x).

In summations, we let $\partial_0 = \partial_t$, and $\xi_0 = \tau$ be the dual variable to t. The symbol ξ denotes $(\xi_1, \dots, \xi_n)$.

2. Reduction to the first order case

The goal of this section is to reduce Theorem 1.3 to the case of a hyperbolic operator of first order, with solution localized to a dyadic frequency shell. We start by introducing a more general class of time-dependent operators, of the type considered by Tataru [14].

We let $\mathrm{g}^{ij}(t,x)$ denote a matrix valued function with $0 \le i, j \le n$, such that

$$\mathrm{g}^{00}(t,x) = 1\,, \qquad \mathrm{g}^{ij}(t,x) = \eta^{ij} \quad \text{if} \quad |x| + |t| \ge \frac{3}{4}\,, \tag{2.1}$$

where η is the Minkowski metric

$$\eta = \mathrm{diag}(1, -1, \dots, -1)\,.$$

We assume that the second derivatives of the functions $\mathrm{g}^{ij}(t,x)$ are sufficiently small in $L^1_t L^\infty_x$ norm,

$$\sup_{i,j} \sup_{|\alpha|=2} \|\partial^\alpha_{t,x} \mathrm{g}^{ij}(t,x)\|_{L^1_t L^\infty_x} \le c_0\,, \tag{2.2}$$

where c_0 is a constant that will be chosen sufficiently small, depending only on the dimension n. Together with (2.1) this implies that g is close to the Minkowski metric in the Lipschitz norm,

$$\sup_{i,j} \sup_{|\alpha|\le 1} \|\partial^\alpha_{t,x} (\mathrm{g}^{ij}(t,x) - \eta^{ij})\|_{L^\infty_t L^\infty_x} \lesssim c_0\,. \tag{2.3}$$

Notice that if $\mathrm{g}(t,x)$ is the smooth cutoff of a time-independent metric, as is the case of interest for Theorem 1.2, then we have the stronger estimate

$$\sup_{i,j} \sup_{|\alpha|=2} \|\partial^\alpha_{t,x} \mathrm{g}^{ij}(t,x)\|_{L^2_t L^\infty_x} \le c_0\,. \tag{2.4}$$

In the following, we let $\partial_0 = \partial_t$, and $\partial_i = \partial_{x_i}$ for $1 \le i \le n$.

Theorem 2.1. *Assume that* (2.1) *and* (2.2) *hold. If* $u(t,x)$ *solves the Cauchy problem*

$$\sum_{i,j=0}^{n} \partial_i\Big(\mathrm{g}^{ij}(t,x)\,\partial_j u(t,x)\Big) = 0\,, \qquad u(0,x)=f(x)\,, \quad \partial_t u(0,x) = g(x)\,, \tag{2.5}$$

then the following estimate holds, provided $q_n \le q < \infty$ *and* $\delta(q)\le 2$,

$$\|u\|_{L^q_x L^2_t(\mathbb{R}^n\times[-1,1])} \le C_q\,\big(\,\|f\|_{H^{\delta(q)}} + \|g\|_{H^{\delta(q)-1}}\big)\,.$$

If (2.1) *and* (2.4) *hold, then the estimate holds provided* $q_n \le q \le \infty$ *and* $\delta(q)\le 2$.

We start by reducing the proof of Theorem 2.1 to the case that u, F, and g^{ij} are all appropriately localized in frequency space. We let

$$1 = \beta_0(\xi) + \sum_{k=1}^{\infty}\beta_k(\xi)\,, \qquad \beta_k(\xi) = \beta_1(2^{-k}\xi)$$

denote a Littlewood-Paley partition of unity on $\mathbb{R}^n$,

$$\mathrm{support}(\beta_0) \subseteq \{\,|\xi|\le 1\}\,, \qquad \mathrm{support}(\beta_1)\subseteq\{\,\tfrac34 \le |\xi| \le 2\}\,.$$

We also introduce a cutoff function $\chi(\tau,\xi)$ supported in the unit ball, which equals 1 for $|\tau,\xi|\le \frac12$. Let $\phi(t)$ be a smooth cutoff to $|t|\le 1$, which vanishes on $|t|\ge \frac32$. We then let

$$u_k(t,x) = \phi(t)\,\beta_k(D_x)u(t,x)\,, \qquad \mathrm{g}^{ij}_k(t,x) = \chi(2^{-\frac{k}{2}}D_t, 2^{-\frac{k}{2}}D_x)\mathrm{g}^{ij}(t,x)\,.$$

Lemma 2.2. *If* u *satisfies* (2.5), *then the functions* $u_k(t,x)$ *satisfy*

$$\sum_{i,j=0}^{n}\partial_i\Big(\mathrm{g}^{ij}_k(t,x)\,\partial_j u_k(t,x)\Big) = F_k(t,x)\,, \quad u_k(0,x) = f_k(x)\,, \quad \partial_t u_k(0,x) = g_k(x)\,, \tag{2.6}$$

where we have the bound, for all $0\le\delta\le 2$,

$$\sum_{k=0}^{\infty}\|f_k\|^2_{H^\delta} + \|g_k\|^2_{H^{\delta-1}} + \|F_k\|^2_{L^1_t H^{\delta-1}} \lesssim \|f\|^2_{H^\delta} + \|g\|^2_{H^{\delta-1}}\,.$$

If (2.4) *holds, we have the bound, for all* $0\le\delta\le 2$,

$$\sum_{k=0}^{\infty}\|f_k\|^2_{H^\delta} + \|g_k\|^2_{H^{\delta-1}} + \|F_k\|^2_{L^2_t H^{\delta-1}} \lesssim \|f\|^2_{H^\delta} + \|g\|^2_{H^{\delta-1}}\,.$$

Proof. We begin by observing that the energy inequality

$$\sup_{t\in[-2,2]}\sum_{0\le j\le 2}\|\partial_t^j u(t,\,\cdot\,)\|_{H^{\delta-j}(\mathbb{R}^n)} \lesssim \|f\|_{H^\delta} + \|g\|_{H^{\delta-1}} \tag{2.7}$$

holds for $0\le\delta\le 2$. This is a consequence of [5] Proposition 6.3.2 and commutation arguments, as well as using the equation to control $\partial_t^2 u$.

The estimates for f_k and g_k follow by orthogonality, as $f_k = \beta_k(D_x)f$, $g_k = \beta_k(D_x)g$. We next write

$$\begin{aligned} F_k = \sum_{i,j=0}^{n}\Big[\,&\partial_i\big((\,\mathrm{g}^{ij}_k - \mathrm{g}^{ij})\,\partial_j u_k\big) + \partial_i\big(\,[\,\mathrm{g}^{ij},\beta_k(D_x)]\,\partial_j(\phi u)\,\big) \\ &+ \beta_k(D_x)\Big(\partial_i\big(\mathrm{g}^{ij}(\partial_j\phi)u\big) + (\partial_i\phi)\mathrm{g}^{ij}\partial_j u\Big)\Big]\,. \end{aligned}$$

Observe that, by (2.6), F_k is localised to frequencies in the range $2^{k-1} \le |\xi| \le 2^{k+1}$, since u_k is similarly limited, and g_k^{ij} is limited to $|\xi| \le 2^{-\frac{k}{2}}$. Thus, we need only control the restriction of each term to frequencies of scale 2^k.

The first term in braces is controlled by bounding

$$\begin{aligned}\|\tilde\beta_k(D_x)d((\mathrm{g}_k-\mathrm{g})du_k)\|_{\ell_k^2 L_t^1 H^{\delta-1}} &\le \|\tilde\beta_k(D_x)d((\mathrm{g}_k-\mathrm{g})du_k)\|_{L_t^1\ell_k^2 H^{\delta-1}}\\ &\lesssim \|d\mathrm{g}\|_{L_t^1L_x^\infty}\|2^{k(\delta-1)}du_k\|_{L_t^\infty\ell_k^2L_x^2}+\|d^2\mathrm{g}\|_{L_t^1L_x^\infty}\|2^{k(\delta-2)}d^2u_k\|_{L_t^\infty\ell_k^2L_x^2}\\ &\lesssim c_0\big(\,\|f\|_{H^\delta}+\|g\|_{H^{\delta-1}}\,\big)\,,\end{aligned}$$

where we used (2.2), (2.7), and the bound

$$\|\,\mathrm{g}_k^{ij}-\mathrm{g}^{ij}\|_{L_t^1L_x^\infty} \lesssim 2^{-k}\|d^2\mathrm{g}^{ij}\|_{L_t^1L_x^\infty}\,.$$

If (2.4) is satisfied, the estimate holds with L_t^1 replaced by L_t^2 in each instance.

To handle the second term in braces, it suffices by the Minkowski inequality to bound

$$\|2^{k(\delta-1)}\,[\,d\mathrm{g},\beta_k(D_x)]\,d(\phi u)\|_{L_t^1\ell_k^2L_x^2}+\|2^{k(\delta-1)}\,[\,\mathrm{g},\beta_k(D_x)]\,d^2(\phi u)\|_{L_t^1\ell_k^2L_x^2}\,.$$

We consider the order $\delta-1$ multiplier $P=\sum_k\varepsilon_k 2^{k(\delta-1)}\beta_k(D_x)$, for an arbitrary choice of $\varepsilon_k=\pm1$. By considering $\varepsilon_k=r_k(t)$ for $t\in[0,1]$, where $r_k(t)$ is the k-th Rademacher function, a standard argument (see e.g. [13] page 464) shows that it suffices to prove that, if $a(x)\in C^{0,1}(\mathbb{R}^n)$, and $b(x)\in C^{1,1}(\mathbb{R}^n)$, then

$$\|\,[a,P]\,f\|_{L^2}\lesssim\|a\|_{C^{0,1}}\|f\|_{H^{\delta-1}}\,,\qquad \|\,[b,P]\,g\|_{L^2}\lesssim\|b\|_{C^{1,1}}\|g\|_{H^{\delta-2}}\,, \tag{2.8}$$

for $0\le\delta\le2$. The first estimate in (2.8) is trivial; one can consider the terms aP and Pa separately. For $\delta=2$, the second estimate in (2.8) follows from the Coifman-Meyer commutator theorem [1], and holds for $b\in C^{0,1}$. See also [15], (3.6.2) and (3.6.35). For $\delta=0,1$, the second estimate holds by commutating this result with d_x, and intermediate values of δ follow by analytic interpolation.

If (2.4) holds, we can similarly replace L_t^1 by L_t^2 in each instance.

Finally, the remaining terms in the braces are controlled by estimating

$$\|\partial_i\big(\mathrm{g}^{ij}(\partial_j\phi)u\big)+(\partial_i\phi)\mathrm{g}^{ij}\partial_ju\|_{L_t^1H^{\delta-1}}\lesssim\|f\|_{H^\delta}+\|g\|_{H^{\delta-1}}\,,$$

and similarly with L_t^1 replaced by L_t^2 if (2.4) holds, which is a simple consequence of (2.7) and (2.2) or (2.4). $\square$

Our goal is to reduce matters to establishing uniform estimates over k for a first order hyperbolic equation. Fix a function $h^+(s)\in C^\infty(\mathbb{R})$, supported in the set $0\le s\le2$, with $h(s)=1$ for $\frac12\le s\le\frac32$. We set $h^-(s)=h^+(-s)$, and let

$$u_k^\pm=h^\pm(2^{-k}D_t)u_k\,.$$

Lemma 2.3. *The function $v_k=u_k-u_k^+-u_k^-$ satisfies, for $2\le q\le\infty$,*

$$\|v_k\|_{L_x^qL_t^2}\lesssim\|f_k\|_{H^{\delta(q)}}+\|g_k\|_{H^{\delta(q)-1}}+\|F_k\|_{L_t^1H^{\delta(q)-1}}\,.$$

The following bound also holds

$$\|v_k\|_{L_x^qL_t^2}\lesssim 2^{-\frac{k}{2}}\big(\,\|f_k\|_{H^{\delta(q)}}+\|g_k\|_{H^{\delta(q)-1}}+\|F_k\|_{L_t^2H^{\delta(q)-1}}\big)\,.$$

Proof. The function v_k is microlocally supported away from the characteristic set of our equation, and the lemma thus reduces to elliptic estimates. Precisely, let

$$b(\tau,\xi)=\bigl(1-h^+(2^{-k}\tau)-h^-(2^{-k}\tau)\bigr)\tilde\beta_k(\xi)\,,$$

where $\tilde\beta_k(\xi)\beta_k(\xi)=\beta_k(\xi)$ and $\tilde\beta_k(\xi)$ is supported in $|\xi|\approx 2^k$, and let $B=b(D_t,D_x)$. We show that there exist pseudodifferential operators Q and R on $\mathbb{R}^{n+1}$ such that

$$B=B\,Q\sum_{i,j=0}^n \partial_i\,\mathrm{g}_k^{ij}\partial_j+BR\,, \tag{2.9}$$

with the property that

$$\|BQF_k\|_{L^q_xL^2_t}\lesssim \min\Bigl(\|F_k\|_{L^1_tH^{\delta(q)-1}}\,,\,2^{-\frac k2}\|F_k\|_{L^2_tH^{\delta(q)-1}}\Bigr) \tag{2.10}$$

$$\|BRu_k\|_{L^q_xL^2_t}\lesssim 2^{-\frac k2}\|u_k\|_{L^\infty_tH^{\delta(q)}}\,. \tag{2.11}$$

The last quantity in (2.11) is controlled using the energy estimates (2.7), and the lemma follows by (2.6), since $Bu_k=v_k$.

We let $a(\tau,\xi)$ be a standard multiplier of order 0, supported in the region $|\tau|+|\xi|\gtrsim 2^k$ and away from the characteristics of the equation, with $a(\tau,\xi)b(\tau,\xi)=b(\tau,\xi)$.

The coefficients $\mathrm{g}_k^{ij}(t,x)$ satisfy

$$|\partial^\beta_{t,x}\mathrm{g}_k^{ij}(t,x)|\lesssim 2^{\frac k2\max(0,|\beta|-1)}\,.$$

It follows that, given N, there exists symbol $Q(t,x,\tau,\xi)\in S^{-2}_{1,\frac12}$, supported in the region $|\tau|+|\xi|\gtrsim 2^k$, such that

$$a(D_t,D_x)=Q\sum_{i,j=0}^n\partial_i\,\mathrm{g}_k^{ij}\partial_j+R\,,$$

with $R\in S^{-N}_{1,\frac12}$. Furthermore, since $\mathrm{g}_k^{ij}(t,x)-\eta^{ij}$ is Schwartz class for $|t|+|x|\ge 1$, the symbol of R will be rapidly decreasing in t and x. The estimate (2.11) then follows easily. To establish (2.10), we may replace Q by $Q\tilde\beta_k(D_x)$, since F_k is supported in the set $|\xi|\approx 2^k$. We then decompose

$$Q(t,x,\tau,\xi)\tilde\beta_k(\xi)=\sum_{j=k}^\infty Q_j(t,x,\tau,\xi)\,,\quad \mathrm{support}(Q_j)\subset\begin{cases}|\tau|\approx 2^j\,, & |\xi|\approx 2^k\,,\quad j>k\\ |\tau|\lesssim 2^k\,, & |\xi|\approx 2^k\,,\quad j=k\end{cases}$$

The integral kernel K_j of Q_j satisfies

$$|K_j(t,x,s,y)|\lesssim 2^{-2j}\,2^{j+nk}\bigl(1+2^j|t-s|+2^k|x-y|\bigr)^{-N}\,,$$

and hence by boundedness of B, the Minkowski inequality, and Young's inequality

$$\|BQ_jF_k\|_{L^q_xL^2_t}\le\|Q_jF_k\|_{L^2_tL^q_x}\lesssim 2^{-\frac32 j}2^{kn(\frac12-\frac1q)}\,\min\Bigl(\|F_k\|_{L^1_tL^2_x}\,,\,2^{-\frac j2}\|F_k\|_{L^2_tL^2_x}\Bigr)\,.$$

The inequality (2.10) follows by summing over $j\ge k$. □

We now set $v=\sum_{k=1}^\infty v_k$. By Littlewood-Paley theory, for $q_n\le q<\infty$ we may bound

$$\|v\|_{L^q_xL^2_t}\le C_q\|v_k\|_{L^q_x\ell^2_kL^2_t}\le C_q\|v_k\|_{\ell^2_kL^q_xL^2_t}\le C_q\,\bigl(\,\|f\|_{H^{\delta(q)}}+\|g\|_{H^{\delta(q)-1}}\bigr)\,,$$

where we use Lemmas 2.2 and 2.3. If (2.4) holds, we control the $L_t^2 H^{\delta(q)-1}$ norm of F and, by the second part of Lemma 2.3, for $q \le \infty$ we may bound

$$\|v\|_{L_x^q L_t^2} \le \|v_k\|_{\ell_k^1 L_x^q L_t^2} \le C\left(\|f\|_{H^{\delta(q)}} + \|g\|_{H^{\delta(q)-1}}\right).$$

Thus, Theorem 2.1 is reduced to establishing the estimate

$$\|u - v\|_{L_x^q L_t^2} \le C\left(\|f\|_{H^{\delta(q)}} + \|g\|_{H^{\delta(q)-1}}\right),$$

which we will show holds for $q_n \le q \le \infty$ under the weaker assumption (2.2). For this, note that the functions $u_k^\pm$ are essentially orthogonal in t, since their Fourier transforms are localized to dyadic intervals $\pm\tau \in [2^k, 2^{k+1}]$. Consequently, for $q_n \le q \le \infty$,

$$\|\textstyle\sum_k u_k^\pm\|_{L_x^q L_t^2} \lesssim \|u_k^\pm\|_{L_x^q \ell_k^2 L_t^2} \lesssim \|u_k^\pm\|_{\ell_k^2 L_x^q L_t^2},$$

where at the last step we may use Minkowski's inequality since $q > 2$. Consequently, by Lemma 2.2, the proof of Theorem 2.1 is reduced to establishing, uniformly over k, the following estimate

$$\|u_k^\pm\|_{L_x^q L_t^2} \lesssim \|f_k\|_{H^{\delta(q)}} + \|g_k\|_{H^{\delta(q)-1}} + \|F_k\|_{L_t^1 H^{\delta(q)-1}}, \tag{2.12}$$

where u_k solves (2.6). Due to the use of orthogonality in t, we need to establish (2.12) where the left hand side norm is taken over $t \in \mathbb{R}$. However, since $u_k^\pm$ are obtained by convolving in t a 2^k-scaled Schwartz function with u_k, which is supported in the interval $|t| \le \frac{3}{2}$, the norm over $|t| \ge 2$ in (2.12) is dominated by $2^{-Nk}\|u_k\|_{L_x^2 L_t^2}$, which is easily controlled by the right hand side. Thus, we may assume that $t \in [-2, 2]$ in (2.12).

We will in fact prove (2.12) in all dimensions, for the full range $\frac{2(n+1)}{n-1} \le q \le \infty$. The restriction $\delta(q) \le 2$ in Theorem 2.1 arises only due to Lemma 2.2.

We now show that the $u_k^\pm$ satisfy an appropriate first order equation. We factor

$$\sum_{i,j=0}^{n} \mathrm{g}_k^{ij}(t,x)\,\xi_i\,\xi_j = \mathrm{g}_k^{00}(t,x)\big(\tau - q_k^+(t,x,\xi)\big)\big(\tau + q_k^-(t,x,\xi)\big),$$

where on the left ξ_j denotes the time dual variable τ in case $j = 0$, and where the functions $q_k^\pm$ are smooth positive functions away from $\xi = 0$, homogeneous of degree 1 in $\xi = (\xi_1, \ldots, \xi_n)$. Since the functions g_k^{ij} are truncations of g^{ij} to frequencies less than $2^{-\frac{k}{2}}$, the assumptions (2.1) and (2.2) imply that for $|\xi| = 1$,

$$\left|\partial_{t,x}^\beta \partial_\xi^\alpha\big(q_k^\pm(t,x,\xi) - |\xi|\big)\right| \lesssim c_0\, 2^{-Nk}\big(1 + |t| + |x|\big)^{-N} \quad \text{if} \quad |t| + |x| \ge 1. \tag{2.13}$$

Also, by (2.2) and (2.3), we have the following derivative estimates on the set $|\xi| = 1$,

$$\left\|\partial_{t,x}^\beta \partial_\xi^\alpha\big(q_k^\pm(t,x,\xi) - |\xi|\big)\right\|_{L_t^1 L_x^\infty} \lesssim c_0\, 2^{\frac{k}{2}\max(0,|\beta|-2)}, \tag{2.14}$$

$$\left\|\partial_{t,x}^\beta \partial_\xi^\alpha\big(q_k^\pm(t,x,\xi) - |\xi|\big)\right\|_{L_t^\infty L_x^\infty} \lesssim c_0\, 2^{\frac{k}{2}\max(0,|\beta|-1)}. \tag{2.15}$$

Finally, we let

$$p_k^\pm(t,x,\xi) = \chi(2^{-\frac{k}{2}} D_t, 2^{-\frac{k}{2}} D_x)\, q_k^\pm(t,x,\xi)$$

be the truncation of the symbols $q_k^\pm$ to frequencies of size at most $2^{\frac{k}{2}}$ in the t and x variables.

Lemma 2.4. *The functions* $u_k^\pm$ *satisfy the Cauchy equations*

$$\partial_t u_k^\pm(t,x) = \pm\, i p_k^\pm(t,x,D_x) u_k^\pm(t,x) + F_k^\pm(t,x)\,, \qquad u_k^\pm(0,x) = f_k^\pm(x)\,,$$

where

$$\|f_k^\pm\|_{H^{\delta(q)}} + \|F_k^\pm\|_{L^1_t H^{\delta(q)}} \lesssim \|f_k\|_{H^{\delta(q)}} + \|g_k\|_{H^{\delta(q)-1}} + \|F_k\|_{L^1_t H^{\delta(q)-1}}\,. \tag{2.16}$$

Proof. We concern ourselves with u_k^-, the case of u_k^+ being identical. We first observe that

$$\|u_k^-\|_{L^\infty_t H^{\delta(q)}} \lesssim \|u_k\|_{L^\infty_t H^{\delta(q)}} \lesssim \|f_k\|_{H^{\delta(q)}} + \|g_k\|_{H^{\delta(q)-1}} + \|F_k\|_{L^1_t H^{\delta(q)-1}}\,, \tag{2.17}$$

which yields the part of (2.16) concerning $f_k^\pm$.

Set $H_k^- = \tilde\beta_k(D_x) h^-(2^{-k} D_t)$, so that $H_k^- u_k = u_k^-$. The estimate (2.15) shows that the symbols $q_k^\pm$ and $\partial_{t,x} q_k^\pm$ are of class $S^1_{1,\frac12}$ on the set $|\xi| \approx 2^k$. We can thus find pseudodifferential operators Q and R such that

$$\bigl(\partial_t + i q_k^-(t,x,D_x)\bigr) H_k^- = Q \sum_{i,j=0}^n \partial_i\, \mathrm{g}_k^{ij} \partial_j\, H_k^- + R\,,$$

where

$$Q(t,x,\tau,\xi) \in S^{-1}_{1,\frac12}\,, \qquad\qquad (t^2+|x|^2)^N R(t,x,\tau,\xi) \in S^0_{1,\frac12} \quad \forall N\,.$$

Then, $\partial_t u_k^- = -i p_k^-(t,x,D_x) u_k^- + F_k^-$, where

$$F_k^- = i(p_k^- - q_k^-) H_k^- u_k + Q H_k^- F_k + Q \sum_{i,j=0}^n \partial_i\,[\,\mathrm{g}_k^{ij}, H_k^-\,]\,\partial_j u_k + R u_k\,. \tag{2.18}$$

We thus need to show that each of the four terms on the right hand side of (2.18) has $L^1_t H^{\delta(q)}$ norm bounded by the right hand side of (2.16).

The operator QH_k^- has symbol of type $S^{-1}_{1,\frac12}$ supported where $|\tau|, |\xi| \approx 2^k$, and thus maps $L^1_t H^{\delta(q)-1} \to L^1_t H^{\delta(q)}$, which handles the second term. The fourth term is easily handled using the rapid decrease of the symbol of R in t and x. To handle the first term, we note by (2.14) that

$$\partial^2_{t,x} q_k^- \cdot \tilde\beta_k \in L^1_t S^1_{1,\frac12}(\mathbb{R}^{2n}_{x,\xi}) \quad \Rightarrow \quad (p_k^- - q_k^-) H_k^- \in L^1_t \Psi^0_{1,\frac12}(\mathbb{R}^n)\,.$$

To handle the third term, we show that

$$\partial_i\,[\,\mathrm{g}_k^{ij}, H_k^-\,] : L^\infty_t H^{\delta(q)-1} \to L^1_t H^{\delta(q)-1}\,. \tag{2.19}$$

Since $\|\partial_j u_k\|_{L^\infty_t H^{\delta(q)-1}}$ is bounded by the right side of (2.16), and Q maps $L^1_t H^{\delta(q)-1} \to L^1_t H^{\delta(q)}$ (note that we are applying Q to a frequency localized term) the bound for the fourth term follows.

The estimate (2.19) follows by using the $S_{1,\frac12}$ calculus on $\mathbb{R}^{n+1}$. The symbol of $[\,\mathrm{g}_k^{ij}, H_k^-\,]$ is of class $S^{-1}_{1,\frac12}$, and localised to $|\tau|, |\xi| \approx 2^k$. Furthermore, it is rapidly decreasing in t and x since g_k^{ij} is. The estimate is a simple consequence. $\square$

We summarize here the reductions we have made. For the rest of this paper, we fix a scale $\lambda = 2^k$, and let $p = p_k^-$. Then $p(t,x,\xi)$ is a real symbol, homogeneous of degree 1 in ξ, which is spectrally limited in t and x to frequencies of size $\lambda^{\frac{1}{2}}$, in that for all ξ,

$$\int e^{i\langle x,\eta\rangle + it\tau} p(t,x,\xi)\,dt\,dx = 0 \qquad \text{if} \quad |\eta| + |\tau| \geq \lambda^{\frac{1}{2}}\,. \tag{2.20}$$

We also have that $p(t,x,\xi)$ is close to $|\xi|$ up to second order, in that for all ξ with $|\xi| = 1$,

$$\sum_{|\alpha|+|\beta|\leq 2} \bigl\|\partial_{t,x}^\beta \partial_\xi^\alpha \bigl(p(t,x,\xi) - |\xi| \bigr) \bigr\|_{L_t^1 L_x^\infty} + \sum_{\substack{|\alpha|+|\beta|\leq 2 \\ |\beta|\leq 1}} \bigl\|\partial_{t,x}^\beta \partial_\xi^\alpha \bigl(p(t,x,\xi) - |\xi| \bigr) \bigr\|_{L_t^\infty L_x^\infty} \leq c_0\,, \tag{2.21}$$

where c_0 is a number to be fixed sufficiently small. Furthermore, for all ξ with $|\xi| = 1$, and all α, β, we have

$$\bigl\|\partial_{t,x}^\beta \partial_\xi^\alpha \bigl(p(t,x,\xi) - |\xi| \bigr) \bigr\|_{L_t^1 L_x^\infty} \lesssim \lambda^{\frac{1}{2}\max(0,|\beta|-2)}\,, \tag{2.22}$$

$$\bigl\|\partial_{t,x}^\beta \partial_\xi^\alpha \bigl(p(t,x,\xi) - |\xi| \bigr) \bigr\|_{L_t^\infty L_x^\infty} \lesssim \lambda^{\frac{1}{2}\max(0,|\beta|-1)}\,. \tag{2.23}$$

Also, (2.21) yields

$$\sup_{|\xi|=1} \sup_{t,x} \bigl|\, d_\xi^2 p(t,x,\xi) - \Pi_\xi^\perp \,\bigr| \leq c_0\,, \tag{2.24}$$

where $\Pi_\xi^\perp$ denotes projection onto the plane normal to ξ, and d_ξ^2 the Hessian with respect to ξ.

Let $p(t,x,D_x)$ denote the corresponding pseudodifferential operator acting on $\mathbb{R}^n$, parametrized by t. Then, by the results of this section, we have reduced Theorem 2.1 to establishing the following, uniformly over $\lambda \geq 1$.

Theorem 2.5. *Suppose that $\partial_t u(t,x) = -i\,p(t,x,D_x)\,u(t,x) + F(t,x)$, where the symbol of p satisfies conditions* (2.20)–(2.24)*. Assume also that, for all t, the partial Fourier transform $\widehat{u}(t,\xi)$ is supported in the region $\frac{1}{4}\lambda < |\xi| < \lambda$. If*

$$\tfrac{2(n+1)}{n-1} \leq q \leq \infty\,, \qquad\qquad \delta(q) = n\Bigl(\tfrac{1}{2} - \tfrac{1}{q}\Bigr) - \tfrac{1}{2}\,,$$

then the following estimate holds

$$\|u\|_{L_x^q L_t^2([-1,1]\times\mathbb{R}^n)} \lesssim \lambda^{\delta(q)} \Bigl(\|u\|_{L_t^\infty L_x^2([-1,1]\times\mathbb{R}^n)} + \|F\|_{L_t^1 L_x^2([-1,1]\times\mathbb{R}^n)} \Bigr)\,.$$

3. The wave transform

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

$$\big(T_\lambda f\big)(x,\xi) = \lambda^{\frac{n}{4}} \int e^{-i\langle \xi, z-x\rangle}\, g\big(\lambda^{\frac12}(z-x)\big)\, f(z)\, dz\,.$$

A simple calculation shows that

$$f(y) = \lambda^{\frac{n}{4}} \int e^{i\langle \xi, y-x\rangle}\, g\big(\lambda^{\frac12}(y-x)\big)\,\big(T_\lambda f\big)(x,\xi)\, dx\, d\xi\,,$$

so that $T_\lambda^* T_\lambda = I$. In particular,

$$\|T_\lambda f\|_{L^2(\mathbb{R}^{2n}_{x,\xi})} = \|f\|_{L^2(\mathbb{R}^n_x)}\,. \tag{3.1}$$

It will be useful to note that this holds in a more general setting.

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

$$\big(T_\lambda f\big)(x,\xi) = \lambda^{\frac{n}{4}} \int e^{-i\langle \xi, z-x\rangle}\, g_{x,\xi}\big(\lambda^{\frac12}(z-x)\big)\, f(z)\, dz$$

satisfies the bound

$$\|T_\lambda f\|_{L^2(\mathbb{R}^{2n}_{x,\xi})} \lesssim \|f\|_{L^2(\mathbb{R}^n_z)}\,.$$

Proof. The operator $T_\lambda T_\lambda^*$ is an integral operator with kernel

$$K(x,\xi;\tilde x,\tilde\xi) = \lambda^{\frac{n}{2}}\, e^{i\langle \xi, x\rangle - i\langle \tilde\xi, \tilde x\rangle} \int e^{i\langle \tilde\xi - \xi, z\rangle} g_{x,\xi}\big(\lambda^{\frac12}(z-x)\big)\, \overline{g_{\tilde x,\tilde\xi}\big(\lambda^{\frac12}(z-\tilde x)\big)}\, dz\,.$$

A simple integration by parts argument shows that

$$\big|K(x,\xi;\tilde x,\tilde\xi)\big| \lesssim \big(\,1 + \lambda^{-\frac12}|\xi - \tilde\xi| + \lambda^{\frac12}|x-\tilde x|\,\big)^{-N}\,,$$

with constants depending only on uniform bounds for a finite collection of seminorms of $g_{x,\xi}$ depending on N. The $L^2(\mathbb{R}^{2n}_{x,\xi})$ boundedness of K then follows by Schur's Lemma. $\square$

Lemma 3.2. *For $\lambda \ge 2^{10}$, and $\frac18\lambda < |\xi| < 2\lambda$, we may write*

$$\Big(p^*(t,y,D_y) - i d_\xi p(t,x,\xi)\cdot d_x + i d_x p(t,x,\xi)\cdot d_\xi\,\Big)\Big[e^{i\langle \xi, y-x\rangle}\, g\big(\lambda^{\frac12}(y-x)\big)\Big] = e^{i\langle \xi, y-x\rangle}\, g_{t,x,\xi}\big(\lambda^{\frac12}(y-x)\big)$$

where $g_{t,x,\xi}(\,\cdot\,)$ denotes a family of Schwartz functions depending on the parameters t, x and ξ, each of which has Fourier transform supported in the ball of radius 2. Furthermore, the Schwartz norms of $\rho(t)^{-1}\, g_{t,x,\xi}(\,\cdot\,)$ are bounded uniformly over t, x and ξ, where

$$\rho(t) = \sup_{|\xi|=1}\,\sup_x \sum_{|\alpha|+|\beta|=2} \big|\,\partial_x^\beta \partial_\xi^\alpha p(t,x,\xi)\,\big|\,.$$

Proof. Letting $\mathfrak{F}$ denote the Fourier transform with respect to y, we write

$$\mathfrak{F}\circ\Big(p^*(t,y,D_y)-id_\xi p(t,x,\xi)\cdot d_x+id_x p(t,x,\xi)\cdot d_\xi\Big)\Big[e^{i\langle\xi,y-x\rangle}\,g\big(\lambda^{\frac12}(y-x)\big)\Big](\eta)$$
$$=e^{-i\langle\eta,x\rangle}\,\lambda^{-\frac n2}\,h_{t,x,\xi}\big(\lambda^{-\frac12}(\eta-\xi)\big)\,,$$

where $h_{t,x,\xi}(\zeta)=\widehat{g}_{t,x,\xi}(\zeta)$ is equal to

$$\int e^{-i\langle\zeta,y\rangle}\Big[p(x+\lambda^{-\frac12}y,\xi+\lambda^{\frac12}\zeta)-p(x,\xi)-\lambda^{\frac12}d_\xi p(x,\xi)\cdot\zeta-\lambda^{-\frac12}d_x p(x,\xi)\cdot y\Big]\,g(y)\,dy$$
$$=\int\int_0^1 e^{-i\langle\zeta,y\rangle}(1-s)\,\partial_s^2\Big(\,\lambda\,p\big(t,x+s\lambda^{-\frac12}y,\lambda^{-1}\xi+s\lambda^{-\frac12}\zeta\big)\Big)\,g(y)\,ds\,dy\,.$$

The spectral restriction on p and g imply that this vanishes for $|\zeta|\geq 2$. Consequently, we are reduced to establishing C^∞ bounds on $\rho(t)^{-1}\,h_{t,x,\xi}(\zeta)$, uniformly over t, x and ξ, for $|\zeta|<2$. Since $\frac18<|\lambda^{-1}\xi|<1$ and $|s\lambda^{-\frac12}\zeta|<\frac1{16}$, the estimates (2.23) imply that any ζ derivative of the integrand is bounded by $C_N\,(1+|y|^2)^{-N}\rho(t)$, with C_N independent of t, y, ξ and ζ for $|\zeta|<2$. The result follows. □

Suppose now that u is as in Theorem 2.5, and let

$$v(t,x,\xi)=\big(T_\lambda u\big)(t,x,\xi)\,,$$

where T_λ acts in the x variable. We assume that $\lambda>2^{10}$, so that $v(t,x,\xi)$ vanishes unless $\frac18\lambda<|\xi|<2\lambda$.

Lemma 3.3. *Under the above conditions, we have*

$$\partial_t v(t,x,\xi)=\Big(d_\xi p(t,x,\xi)\cdot d_x-d_x p(t,x,\xi)\cdot d_\xi\Big)v(t,x,\xi)+\big(T_\lambda F\big)(t,x,\xi)+G(t,x,\xi)\,,$$

where $G(t,x,\xi)=0$ *unless* $\frac18\lambda<|\xi|<2\lambda$, *and*

$$\|G(t,\,\cdot\,)\|_{L^2(\mathbb{R}^{2n}_{x,\xi})}\lesssim\rho(t)\,\|u(t,\,\cdot\,)\|_{L^2(\mathbb{R}^n_x)}\,. \tag{3.2}$$

Proof. Differentiating under the integral sign yields

$$\partial_t v(t,x,\xi)=\big(T_\lambda F\big)(t,x,\xi)-i\,\lambda^{\frac n4}\int\overline{p^*(t,y,D_y)\Big[e^{i\langle\xi,y-x\rangle}\,g\big(\lambda^{\frac12}(y-x)\big)\Big]}\,u(t,y)\,dy\,,$$

By Lemma 3.2, we are done if we can prove the estimate (3.2) for the term

$$G(t,x,\xi)=\lambda^{\frac n4}\int e^{-i\langle\xi,y-x\rangle}\,g_{t,x,\xi}\big(\lambda^{\frac12}(y-x)\big)\,u(t,y)\,dy\,.$$

By Lemma 3.2, each Schwartz norm of $\rho(t)^{-1}g_{t,x,\xi}$ is uniformly bounded over x and ξ. The estimate (3.2) then follows by Lemma 3.1. □

Let $\chi_{s,t}$ denote the canonical transform on $\mathbb{R}^{2n}_{x,\xi}=T^*(\mathbb{R}^n)$ generated by the hamiltonian flow of p. Thus, $\chi_{s,t}(x,\xi)=\gamma(s)$, where γ is the integral curve with $\gamma(t)=(x,\xi)$. Then we have

$$v(t,x,\xi)=\big(T_\lambda f\big)(\chi_{0,t}(x,\xi))+\int_0^t G(r,\chi_{r,t}(x,\xi))\,dr+\int_0^t\big(T_\lambda F\big)(r,\chi_{r,t}(x,\xi))\,dr\,.$$

Writing $u(t,\,\cdot\,) = T_\lambda^* v(t,\,\cdot\,) = T_\lambda^* v_t$ yields

$$|u(t,\,\cdot\,)| \le \bigl|T_\lambda^*\bigl(T_\lambda f \circ \chi_{0,t}\bigr)\bigr| + \int_{-1}^1 \bigl|T_\lambda^*\bigl(G_r \circ \chi_{r,t}\bigr)\bigr|\,dr + \int_{-1}^1 \bigl|T_\lambda^*\bigl(T_\lambda F_r \circ \chi_{r,t}\bigr)\bigr|\,dr\,.$$

For each r, we define a map W_r taking a function $\tilde f$ on $\mathbb{R}^{2n}_{x,\xi}$ to a function on $\mathbb{R}^{1+n}_{s,y}$

$$\begin{aligned}\bigl(W_r \tilde f\,\bigr)(s,y) &= T_\lambda^*\bigl(\tilde f \circ \chi_{r,s}\bigr)(y)\\ &= \lambda^{\frac n4}\int e^{i\langle \xi, y-x\rangle} g\bigl(\lambda^{\frac12}(y-x)\bigr)\,\tilde f(\chi_{r,s}(x,\xi))\,dx\,d\xi\,.\end{aligned}$$

Then, by (3.1) and (3.2), and since the volume form $dx\wedge d\xi$ is invariant under the symplectic map $\chi_{r,t}$, we have reduced Theorem 2.5 to establishing, uniformly over r,

$$\|S_\lambda(D_x) W_r \tilde f\,\|_{L^p_x L^2_t([-1,1]\times\mathbb{R}^n)} \lesssim \lambda^{\delta(p)}\|\tilde f\,\|_{L^2(\mathbb{R}^{2n})}\,,$$

where $S_\lambda(D_x)$ is a Littlewood-Paley cutoff to frequencies at scale λ in x. This is equivalent to proving that

$$\|S_\lambda(D_x) W_r W_r^* S_\lambda(D_x) F\|_{L^p_x L^2_t([-1,1]\times\mathbb{R}^n)} \lesssim \lambda^{2\delta(p)}\|F\|_{L^{p'}_x L^2_t([-1,1]\times\mathbb{R}^n)}\,, \tag{3.3}$$

with p' the dual index to p. The operator $W_r W_r^* S_\lambda(D_x)$ takes the form

$$\lambda^{\frac n2}\int e^{i\langle\xi,y-x\rangle - i\langle \xi_{t,s}, z-x_{t,s}\rangle} g\bigl(\lambda^{\frac12}(y-x)\bigr)\, g\bigl(\lambda^{\frac12}(z-x_{t,s})\bigr)\,\bigl(S_\lambda(D_z)F\bigr)(t,z)\,dt\,dz\,dx\,d\xi\,,$$

where $(x_{t,s},\xi_{t,s}) = \chi_{t,s}(x,\xi)$. In this calculation, we used the fact that $dx\wedge d\xi$ is invariant under $\chi_{r,t}$, and that $\chi_{t,r}\circ\chi_{r,s} = \chi_{t,s}$.

The integral over z vanishes unless $|\xi_{t,s}|\approx\lambda$, hence unless $|\xi|\approx\lambda$, thus we are reduced to establishing $L^{p'}_z L^2_t \to L^p_y L^2_s$ estimates, with norm $\lambda^{2\delta(p)}$, for the integral kernel

$$K(s,y;t,z) = \lambda^{\frac n2}\int e^{i\langle\xi,y-x\rangle - i\langle \xi_{t,s}, z-x_{t,s}\rangle} g\bigl(\lambda^{\frac12}(y-x)\bigr)\, g\bigl(\lambda^{\frac12}(z-x_{t,s})\bigr)\beta(\lambda^{-1}\xi)\,dx\,d\xi\,,$$

with $\beta(\zeta)$ a smooth function supported in the region $|\zeta|\approx 1$.

We next follow the steps of [8], by localizing the operator in the Fourier variables and reducing matters to dispersive type estimates. By decomposing β into a finite number of terms, we may assume that it is supported in a cone of small angle, and without loss of generality we assume that β is supported in a small cone about the ξ_1 axis. We then split $z=(z_1,z')$ and $y=(y_1,y')$, and will prove the following pair of estimates

$$\Bigl\|\int K(s,y;t,z)\,f(t,z')\,dt\,dz'\,\Bigr\|_{L^2_{y'}L^2_s([-1,1]\times\mathbb{R}^{n-1})} \lesssim \|f\|_{L^2_{z'}L^2_t([-1,1]\times\mathbb{R}^{n-1})}\,, \tag{3.4}$$

$$\begin{aligned}\Bigl\|\int K(s,y;t,z)\,f(t,z')\,dt\,dz'\,\Bigr\|_{L^\infty_{y'}L^2_s([-1,1]\times\mathbb{R}^{n-1})}\\ \lesssim \lambda^{n-1}\bigl(1+\lambda\,|y_1-z_1|\,\bigr)^{-\frac{n-1}{2}}\|f\|_{L^1_{z'}L^2_t([-1,1]\times\mathbb{R}^{n-1})}\,.\end{aligned} \tag{3.5}$$

Interpolation yields that, for $\frac{2(n+1)}{n-1} \le p \le \infty$,

$$\Big\| \int K(s,y;t,z)\, f(t,z')\, dt\, dz' \Big\|_{L^p_{y'} L^2_s([-1,1]\times\mathbb{R}^{n-1})} \lesssim \lambda^{2\delta(p)} |y_1 - z_1|^{-1+\frac{1}{p'}-\frac{1}{p}} \|f\|_{L^{p'}_{z'} L^2_t([-1,1]\times\mathbb{R}^{n-1})} \,,$$

and an application of the Hardy-Littlewood inequality yields the desired bound.

Proof of the estimate (3.4). We make the measure preserving change of variables $\chi_{0,s}$ in x and ξ to write

$$K(s,y;t,z) = \lambda^{\frac{n}{2}} \int e^{i\langle \xi_s, y-x_s\rangle - i\langle \xi_t, z-x_t\rangle} g\big(\lambda^{\frac12}(y-x_s)\big)\, g\big(\lambda^{\frac12}(z-x_t)\big) \beta(\lambda^{-1}\xi_s)\, dx\, d\xi \,,$$

where $(x_s,\xi_s) = \chi_{s,0}(x,\xi)$ and $(x_t,\xi_t) = \chi_{t,0}(x,\xi)$.

We will show that, for fixed y_1, the map $T^{y_1}_\lambda$ defined by

$$\big(T^{y_1}_\lambda \tilde f\,\big)(s,y') = \lambda^{\frac{n}{4}} \int e^{i\langle \xi_s, y-x_s\rangle} g\big(\lambda^{\frac12}(y-x_s)\big)\, \beta(\lambda^{-1}\xi_s)\, \tilde f(x,\xi)\, dx\, d\xi$$

is bounded from $L^2(\mathbb{R}^{2n}_{x,\xi}) \to L^2(\mathbb{R}^n_{s,y'})$. The operator K is essentially of the form $T^{y_1}_\lambda (T^{z_1}_\lambda)^*$, since we can insert a harmless cutoff like $\beta(\lambda^{-1}\xi_t)$ as ξ_t is restricted to a small neighborhood if ξ_s is, and the L^2 boundedness of K follows.

Fix y_1. Given (x,ξ), let $\bar s$ denote the unique time s at which x_s lies over y_1, which exists since $(\partial_s x_s)_1 \approx 1$. Then, with $\bar y' = x'_{\bar s}$, $\bar\sigma = -p(\bar s, x_{\bar s}, \xi_{\bar s})$, and $\bar\eta' = \xi'_{\bar s}$, the point $(\bar s, \bar y'; \bar\sigma, \bar\eta')$ is the intersection of the null Hamiltonian curve determined by (x,ξ) with the cotangent bundle $T^*(\mathbb{R}^n_{s,y'})$. We now show that we may write

$$e^{i\langle \xi_s, y-x_s\rangle} g\big(\lambda^{\frac12}(y-x_s)\big)\, \beta(\lambda^{-1}\xi_s) = e^{i\bar\sigma(s-\bar s) + i\langle \bar\eta', y'-\bar y'\rangle} g_{x,\xi}\big(\lambda^{\frac12}(s-\bar s, y'-\bar y')\big)\,, \tag{3.6}$$

where $g_{x,\xi}(\,\cdot\,)$ denotes a family of Schwartz functions parametrized by (x,ξ), with uniform bounds as x and ξ vary.

We use the following facts about the Hamiltonian flow, where $y = (y_1, y')$ with y_1 fixed as above, and (x,ξ) fixed but arbitrary.

$$|y - x_s| \approx |s - \bar s| + |y' - \bar y'|\,, \qquad s \in \mathbb{R}\,,\ \ y' \in \mathbb{R}^{n-1}\,, \tag{3.7}$$

and

$$\big|\partial_s^{j+1} x_s\big| \lesssim \lambda^{\frac{j}{2}}\,, \qquad \big|\partial_s^{j+1}\xi_s\big| \lesssim \lambda^{\frac{j}{2}+1}\,, \quad \text{if } \ j \ge 0 \ \text{ and } \ |\xi| \approx \lambda\,. \tag{3.8}$$

The estimate (3.7) follows since $(\partial_s x_s)_1 \approx 1$ and $|\partial_s x'_s| \lesssim 1$; the estimate (3.8) follows by repeatedly differentiating the Hamilton equations and using (2.23) and (2.22).

We also observe that we may write

$$\langle \xi_s, y - x_s\rangle - \bar\sigma(s-\bar s) - \langle \bar\eta', y'-\bar y'\rangle = q_0(s)(s-\bar s)^2 + \sum_{j=2}^n q_j(s)(s-\bar s)(y'-\bar y')_j\,, \tag{3.9}$$

with

$$\partial_s^m q_i(s) \lesssim \lambda^{\frac{m}{2}}\,. \tag{3.10}$$

To do this, we note that by homogeneity of p we have $\sigma = -\langle \xi_{\bar s}, (\partial_s x_s)|_{s=\bar s}\rangle$, and thus can write the left hand side as

$$\langle \xi'_s - \xi'_{\bar s}, y' - \bar y'\rangle - \langle \xi_s - \xi_{\bar s}, x_s - x_{\bar s}\rangle - \langle \xi_{\bar s}, x_s - x_{\bar s} - (s - \bar s)(\partial_s x_s)|_{s=\bar s}\rangle \,.$$

The result follows by Taylor's theorem and (3.8). Finally, we note that

$$\bigl|\partial_s^j \partial_{y'}^\beta \bigl[\, g\bigl(\lambda^{\frac12}(y - x_s)\bigr)\, \beta(\lambda^{-1}\xi_s)\,\bigr]\bigr| \lesssim \lambda^{\frac12(j+|\beta|)}\bigl(\,1 + \lambda^{\frac12}|y' - \bar y'| + \lambda^{\frac12}|s - \bar s|\,\bigr)^{-N} ,$$

which follows from (3.7) and (3.8). Together with (3.9) and (3.10), this implies (3.6).

We can thus write

$$\bigl(T_\lambda^{y_1} f\bigr)(s, y') = \lambda^{\frac n4} \int e^{i\bar\sigma(s-\bar s) + i\langle \bar\eta', y' - \bar y'\rangle} g_{x,\xi}\bigl(\lambda^{\frac12}(s - \bar s, y' - \bar y')\bigr)\, f(x,\xi)\, dx\, d\xi \,.$$

The map $(x,\xi) \to (\bar s, \bar y', \bar\sigma, \bar\eta')$ is a measure preserving diffeomorphism, since it is symplectic. After this change of variables, the adjoint of $T_\lambda^{y_1}$ is then of the type for which $L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^{2n})$ boundedness was established in Lemma 3.1. □

We establish estimate (3.5) through pointwise bounds on the kernel $K(s,y;t,z)$. We will show in the next section that

$$|K(s,y;t,z)| \lesssim \lambda^n \bigl(\,1 + \lambda\,|y_1 - z_1|\,\bigr)^{-\frac{n-1}{2}} \,\bigl(\,1 + \lambda\,|\,s - \Phi^{\pm}_{t,z}(y)|\,\bigr)^{-N} . \tag{3.11}$$

Here, $\Phi^+_{t,z}$ is the function such that $s = \Phi^+_{t,z}(y)$ defines the forward light cone centered at (t,z) in the region $2 \ge s \ge t$, and $\Phi^+_{t,z}(y) = 2$ if (s,y) is not in the light cone for any $2 \ge s \ge t$. Similarly, $\Phi^-_{t,z}(y)$ defines the backward light cone centered at (t,z) in the region $-2 \le s \le t$, and $\Phi^-_{t,z}(y) = -2$ if y is not in the appropriate domain. The above estimate then holds taking $\Phi^+_{t,z}$ for $s \ge t$ and $\Phi^-_{t,z}$ for $s \le t$.

Symmetry of the form of K shows that (3.11) holds with (t,z) and (s,y) exchanged. The estimate (3.5) then follows by Schur's Lemma.

The next section is devoted to proving these pointwise estimates, which we do through a scaling argument. Without loss of generality we assume that $s = 0$ and $t \in [0,1]$. Let ε denote a scaling factor with $\lambda^{-1} \le \varepsilon \le 1$, and introduce the scaled kernel

$$K_\varepsilon(y;t,z) = \varepsilon^n K(0, \varepsilon y; \varepsilon t, \varepsilon z)\,.$$

Introducing the scaled parameters

$$R = \varepsilon\lambda^{\frac12}\,, \qquad \mu = \varepsilon\lambda\,,$$

we have

$$K_\varepsilon(y;t,z) = R^n \int e^{i\langle \xi, y - x\rangle - i\langle \xi_t, z - x_t\rangle}\, g\bigl(R(y-x)\bigr)\, g\bigl(R(z - x_t)\bigr)\, \beta(\mu^{-1}\xi)\, dx\, d\xi\,,$$

where now (x_t, ξ_t) is an integral curve of the Hamiltonian function

$$H(t,x,\xi) = p(\varepsilon t, \varepsilon x, \xi) \tag{3.12}$$

with $(x_0,\xi_0) = (x,\xi)$. In the next section we prove the following estimate.

Lemma 3.4. *Let Γ_z denote the time $s=0$ slice of the light cone centered at $(1,z)$. Then the following bounds hold uniformly for $\mu \ge 1$,*

$$|K_\varepsilon(y;1,z)| \lesssim \mu^{\frac{n+1}{2}} \bigl(1 + \mu\, d(y,\Gamma_z) \bigr)^{-N} . \tag{3.13}$$

And for $\mu = 1$, the following bounds hold uniformly for $t \in [0,1]$,

$$|K_\varepsilon(y;t,z)| \lesssim \bigl(1 + |y-z| \bigr)^{-N} . \tag{3.14}$$

To deduce the estimate (3.11) (for $s=0$ and $t \in [0,1]$), we consider separately the cases $t \ge \lambda^{-1}$ and $t < \lambda^{-1}$. In case $t \ge \lambda^{-1}$, we take $\varepsilon = t$, and deduce from (3.13) and a scaling argument that

$$|K(0,y;t,z)| \lesssim \lambda^n \bigl(1 + \lambda\, t \bigr)^{-\frac{n-1}{2}} \bigl(1 + \lambda\, d(y,\Gamma_{t,z}) \bigr)^{-N} , \tag{3.15}$$

where $\Gamma_{t,z}$ is the time 0 slice of the light cone centered at (t,z). If $|y-z| \ge \frac{3}{2}$, then $d(y,\Gamma_{t,z}) \gtrsim 1+|y-z|$, and the estimate (3.11) follows easily, since $|\Phi^-_{t,z}| \le 3$. If $|y-z| \le \frac{3}{2}$, then y belongs to the domain of $\Phi_{t,z}$, and by (4.13) below we conclude that

$$|K(0,y;t,z)| \lesssim \lambda^n \bigl(1 + \lambda\, t \bigr)^{-\frac{n-1}{2}} \bigl(1 + \lambda\, |\, \Phi^-_{t,z}(y)| \bigr)^{-N} .$$

On the other hand, by the first part of (4.11) below,

$$\bigl|\, \Phi^-_{t,z}(y) - \bigl(t - |y-z| \bigr) \bigr| \ll |y-z| .$$

Since $|y-z| \ge |y_1 - z_1|$, these estimates together imply (3.11).

In case $t \le \lambda^{-1}$, we take $\varepsilon = \lambda^{-1}$. The estimate (3.14) and a scaling argument yield

$$|K(0,y;t,z)| \lesssim \lambda^n \bigl(1 + \lambda\, |y-z| \bigr)^{-N} . \tag{3.16}$$

Together with the preceeding estimate, and the fact that $\lambda\, t \le 1$, this implies (3.11) for this case.

4. Regularity for the Hamiltonian flow

We work in this section with the Hamiltonian function defined above by (3.12). By (2.23), the following estimates are then satisfied,

$$\int_{-2}^{2} \sup_{|\xi|=1} \sup_x \bigl| \partial^\beta_{t,x} \partial^\alpha_\xi H(t,x,\xi) \bigr| \, dt \le \begin{cases} C_\alpha \,, & |\beta| = 0 \\ C_\alpha\, \varepsilon \,, & |\beta| = 1 \\ C_{\alpha,\beta}\, \varepsilon\, R^{|\beta|-2} \,, & |\beta| \ge 2 \end{cases} \tag{4.1}$$

Lemma 4.1. *Let $\bigl(x_t(x,\xi)\,,\, \xi_t(x,\xi) \bigr)$ denote the solution to the Hamiltonian flow,*

$$\partial_t x_t = \bigl(d_\xi H \bigr)(t, x_t, \xi_t) \,, \quad \partial_t \xi_t = -\bigl(d_x H \bigr)(t, x_t, \xi_t) \,, \quad x_0 = x \,, \quad \xi_0 = \xi \,.$$

Then, uniformly for $-2 \le t \le 2$ and $|\xi| = 1$, the following hold

$$\bigl| \, x_t - t\xi \, \bigr| + \bigl| \, \xi_t - \xi \, \bigr| \ll t \,, \tag{4.2}$$

and

$$\bigl| \, d_x x_t - I \, \bigr| + \bigl| \, d_x \xi_t \bigr| + \bigl| \, d_\xi x_t - \int_0^t d_\xi^2 H(s, x_s, \xi_s) \, ds \, \bigr| + \bigl| \, d_\xi \xi_t - I \, \bigr| \ll 1 \,. \tag{4.3}$$

Furthermore, for derivatives of order $|\alpha|+|\beta|\geq 2$, the following estimates hold, uniformly for $0<t<1$ and $|\xi|=1$,

For $R\leq 1$,

$$\big|\,\partial_x^\beta\partial_\xi^\alpha x_t\,\big|\lesssim\begin{cases}1\,, & |\beta|=0\\ \varepsilon\,, & |\beta|=1\\ \varepsilon\,R^{|\beta|-2}\,, & |\beta|\geq 2\end{cases}\tag{4.4}$$

$$\big|\,\partial_x^\beta\partial_\xi^\alpha \xi_t\,\big|\lesssim\begin{cases}\varepsilon\,, & |\beta|=0\\ \varepsilon\,R^{|\beta|-1}\,, & |\beta|\geq 1\end{cases}\tag{4.5}$$

For $R\geq 1$,

$$\big|\,\partial_x^\beta\partial_\xi^\alpha x_t\,\big|\lesssim\begin{cases}1+\varepsilon\,R^{|\alpha|-1}\,, & |\beta|=0\\ \varepsilon\,R^{|\alpha|+|\beta|-1}\,, & |\beta|\geq 1\end{cases}\tag{4.6}$$

$$\big|\,\partial_x^\beta\partial_\xi^\alpha \xi_t\,\big|\lesssim\;\;\varepsilon\,R^{|\alpha|+|\beta|-1}\tag{4.7}$$

Proof. To begin we note that

$$\big|\partial_t H(t,x_t,\xi_t)\big|=\big|\big(\partial_t H\big)(t,x_t,\xi_t)\big|\lesssim\;H\big(t,x_t,\xi_t\big)\,,$$

and consequently by the Gronwall Lemma that, for $-2\leq t\leq 2$, and $|\xi|=1$,

$$|\xi_t(x,\xi)|\approx 1\,.\tag{4.8}$$

The estimate (4.2) is a simple consequence of (2.21).

We next differentiate the Hamilton equations to write

$$\partial_t\begin{pmatrix}d_x x_t\\ \\ d_x\xi_t\end{pmatrix}=M(t,x_t,\xi_t)\cdot\begin{pmatrix}d_x x_t\\ \\ d_x\xi_t\end{pmatrix}\,,\qquad\qquad \partial_t\begin{pmatrix}d_\xi x_t\\ \\ d_\xi\xi_t\end{pmatrix}=M(t,x_t,\xi_t)\cdot\begin{pmatrix}d_\xi x_t\\ \\ d_\xi\xi_t\end{pmatrix}$$

where

$$M=\begin{pmatrix}\big(d_xd_\xi H\big) & \big(d_\xi d_\xi H\big)\\ \\ -\big(d_xd_x H\big) & -\big(d_\xi d_x H\big)\end{pmatrix}$$

By (4.1) and (4.8), the upper right hand block of M is of norm ≈ 1 in $L^1_t[-2,2]$, and all other blocks in M are of norm $\ll 1$ in $L^1_t[-2,2]$, uniformly over x and ξ. Since at $t=0$ we have $d_x\xi_0=d_\xi x_0=0$, and $d_x x_0=d_\xi\xi_0=I$, the estimate (4.3) follows by the Gronwall Lemma.

To control higher order derivatives we proceed by induction. We write

$$\partial_t\begin{pmatrix}\partial_x^\beta\partial_\xi^\alpha x_t\\ \\ \partial_x^\beta\partial_\xi^\alpha\xi_t\end{pmatrix}=M(t,x_t,\xi_t)\cdot\begin{pmatrix}\partial_x^\beta\partial_\xi^\alpha x_t\\ \\ \partial_x^\beta\partial_\xi^\alpha\xi_t\end{pmatrix}+\begin{pmatrix}E_1\\ \\ E_2\end{pmatrix}$$

where E_1 is a sum of terms of the form

$$\big(d_x^k d_\xi^{j+1}H\big)(t,x_t,\xi_t)\cdot\big(\partial_x^{\beta_1}\partial_\xi^{\alpha_1}x_t\big)\cdots\big(\partial_x^{\beta_k}\partial_\xi^{\alpha_k}x_t\big)\big(\partial_x^{\beta_{k+1}}\partial_\xi^{\alpha_{k+1}}\xi_t\big)\cdots\big(\partial_x^{\beta_{k+j}}\partial_\xi^{\alpha_{k+j}}\xi_t\big)$$

and E_2 is similarly a sum of such terms, but with $d_x^{k+1}d_\xi^j H$. In both cases, $\beta_i+\alpha_i < \beta+\alpha$ for each i, and $\beta_1+\cdots\beta_{j+k}=\beta\,,\ \alpha_1+\cdots\alpha_{j+k}=\alpha\,.$ We thus assume that the estimates (4.3)–(4.7) hold for all terms arising in E_1 and E_2. We use the notation

$$|||E_j||| = \sup_{|\xi|=1}\sup_x \int_{-2}^2 |E_j(t,x,\xi)|\,dt\,.$$

We first consider the case $R<1$. The estimates (4.3) and (4.4) show that all derivatives of x_t and ξ_t are bounded by ≈ 1, and (4.1) implies that

$$|||E_1|||\lesssim 1\,,\qquad |||E_2|||\lesssim \varepsilon\,,$$

which implies the case $\beta=0$ of (4.3) and (4.4). In case $|\beta|=1$, this estimate can be improved to

$$|||E_1|||\lesssim \varepsilon\,,\qquad |||E_2|||\lesssim \varepsilon\,,$$

since at least one term in the product arising in E_1 must be of size ε. For $|\beta|\ge 2$, we must show that

$$|||E_1|||\lesssim \varepsilon\, R^{|\beta|-2}\,,\qquad |||E_2|||\lesssim \varepsilon\, R^{|\beta|-1}\,.$$

It is helpful to note that, since $\varepsilon<R<1$, estimates (4.4) and (4.5) imply

$$|\partial_x^\beta\partial_\xi^\alpha x_t|\lesssim \min\bigl(1,R^{|\beta|-1}\bigr)\,,\qquad |\partial_x^\beta\partial_\xi^\alpha \xi_t|\lesssim \varepsilon\, R^{|\beta|-1}\le R^{|\beta|}\,. \tag{4.9}$$

For E_1, considering separately the cases $k=0$, $k=1$, and $k\ge 2$ leads to

$$|||E_1|||\lesssim 1\cdot\varepsilon\, R^{|\beta|-1}+\varepsilon\cdot R^{|\beta|-1}+\varepsilon\, R^{k-2}\cdot R^{|\beta|-k}\lesssim \varepsilon\, R^{|\beta|-2}\,.$$

The estimate for E_2 follows similarly.

For $R\ge 1$, it is useful to note that (4.6) and (4.7) imply

$$|\partial_x^\beta\partial_\xi^\alpha x_t|\lesssim R^{|\alpha|+|\beta|-1}\,,\qquad |\partial_x^\beta\partial_\xi^\alpha \xi_t|\lesssim R^{|\alpha|+|\beta|-1}\,, \tag{4.10}$$

since $\varepsilon\le 1$. Considering separately the cases $k=0$ and $k=1$ then leads to the estimate

$$|||E_2|||\lesssim \varepsilon\cdot R^{|\alpha|+|\beta|-j}+\varepsilon\, R^{k-1}\cdot R^{|\alpha|+|\beta|-k-j}\lesssim \varepsilon\, R^{|\alpha|+|\beta|-1}\,.$$

This bound also holds for any term arising in the expansion for E_1 for which $k\neq 0$. In case $|\beta|=0$ and $k=0$, we have the bound

$$|||E_1|||\lesssim 1\cdot 1+\varepsilon\, R^{|\alpha|-j}\lesssim 1+\varepsilon\, R^{|\alpha|-1}\,,$$

where the term 1 arises in case $|\alpha_i|=1$ for all i. For $|\beta|\ge 1$ and $k=0$ we have the bound

$$|||E_1|||\lesssim \varepsilon\, R^{|\alpha|+|\beta|-1}\,,$$

since $|\partial_x^\beta\partial_\xi^\alpha\xi_t|\lesssim \varepsilon\, R^{|\alpha|+|\beta|-1}$ whenever $|\beta|\ge 1$. Together with the Gronwall Lemma, these estimates imply (4.6) and (4.7). $\square$

We now study the geometry and regularity of the light cones. By translation invariance of the conditions on the Hamiltonian, it suffices to consider the forward cone centered at the origin. Thus, let $\Gamma\subset[0,2)\times\mathbb{R}^n$ denote the forward light cone centered at $t=0$ and $x=0$,

$$\Gamma=\bigcup_{t\in[0,2)}\ \bigcup_{\xi\in\mathbb{R}^n\setminus 0}\bigl(t,x_t(0,\xi)\bigr)\,,$$

and let $\Gamma^s\subset\mathbb{R}^n$ denote the slice of Γ at time $t=s$.

Theorem 4.2. *The set Γ can be written as the graph $s=\Phi(y)$ of a function $\Phi(y)$ defined on an open subset of $\mathbb{R}^n$. The function Φ satisfies*

$$\bigl|\,\Phi(y)-|y|\,\bigr|\ll|y|\,,\qquad \Bigl|\,d_y\Phi(y)-\frac{y}{|y|}\,\Bigr|\ll 1\,,\qquad \bigl|\,d_y^2\Phi(y)\,\bigr|\lesssim\frac{1}{|y|}\,. \tag{4.11}$$

Furthermore, for each $0<s<2$, the level set $\Gamma^s:\{\Phi(y)=s\}$ is a convex hypersurface in $\mathbb{R}^n$, and for $x,y\in\Gamma^t$ we have

$$\Bigl|\,\frac{d_x\Phi(x)}{|d_x\Phi(x)|}-\frac{d_y\Phi(y)}{|d_y\Phi(y)|}\,\Bigr|\approx s\,|x-y|\,. \tag{4.12}$$

Finally, for each y belonging to the domain of $\Phi(y)$, and $s\in[0,2)$, we have

$$d(y,\Gamma^s)\approx\bigl|\,s-\Phi(y)\,\bigr|\,. \tag{4.13}$$

Proof. We prove the result for the part of Γ corresponding to $1<t<2$; the case of $2^{-j-1}<t<2^{-j}$ for $j\ge 0$ will follow by a scaling argument. We consider the map $z\to y$ defined as follows, for $|z|<2$. We write $z=r\omega$, where $|\omega|=1$ and $r\in[0,2)$, and let

$$y(z)=x_r(0,\omega)\,,\qquad \xi(z)=\xi_r(0,\omega)\,.$$

Then by (4.2)–(4.3), (2.21) and (2.24), we have

$$\bigl|\partial_r y-\omega\bigr|\ll 1\,,\qquad \bigl|\partial_\omega y-r\Pi_\omega^\perp\bigr|\ll 1\,,\qquad \bigl|\partial_r\xi\bigr|\ll 1\,,\qquad \bigl|\partial_\omega\xi-\Pi_\omega^\perp\bigr|\ll 1\,.$$

Consequently, on the set $1<|z|<2$, it follows that

$$\|\,y(z)-z\,\|_{C^1}\ll 1\,,\qquad \|\,\xi(z)-\omega\,\|_{C^1}\ll 1\,.$$

It follows that the map $z\to y$ is a C^1 diffeomorphism on the set $1<|z|<2$.

The set Γ is defined by the condition $s=r$. Letting $\Phi(y)=r$ yields that Γ is the graph of $\Phi(y)$. To see that Φ is C^2 as a function of y, we note that its differential is given by

$$\bigl(d_y\Phi\bigr)(y(z))\;=\;\frac{\xi(z)}{H\bigl(r,y(z),\xi(z)\bigr)}$$

which follows from the fact that ξ is normal to $d_\omega x$, since the flow is symplectic, and the fact that $\xi\cdot\partial_r x=H(r,x,\xi)$. Consequently, $d_y\Phi(y)$ is C^1 close to ω in the coordinates z, hence C^1 close in the coordinates y. The estimates in (4.2) follow for $1<|z|<2$, and then for $0<|z|<2$ by scaling.

We next observe that for $y\in\Gamma^1$,

$$\frac{d_y\Phi(y)}{|d_y\Phi(y)|}=\frac{\xi(\omega)}{|\xi(\omega)|}$$

if $y=y(\omega)$. By the above, the map $\omega\to y(\omega)$ is bilipschitz from $S^{n-1}\to\Gamma^1$. Also,

$$\bigl|\,d_\omega\xi(\omega)-\Pi_\omega^\perp\,\bigr|\ll 1\,,$$

where $\Pi_\omega^\perp$ denotes the differential of the inclusion of $S^{n-1}\to\mathbb{R}^n$. This implies (4.12) for $t=1$, as well as the fact that Γ^1 is convex. The result for $t\in(0,2)$ follows by scaling.

To establish (4.13), we note that, since Γ is contained in the set $.99\,s<|y|<1.01\,s$, we may restrict to the case $.98\,s<|y|<1.02\,s$. After scaling, we may then assume $t=1$

and $.98 < |y| < 1.02$. There is then a unique point $x \in \Gamma^1$ closest to y, and since $y - x$ is normal to Γ^1 at x, hence close in angle to the vector x, then

$$\big|\, d(y, \Gamma^1) - |\langle x, y - x\rangle| \,\big| \ll 1\,.$$

On the other hand,

$$\big|\, \Phi(y) - 1 \,\big| = \Big|\, (y - x) \cdot \int_0^1 \big(d_y \Phi\big)(sx + (1-s)y)\, ds \,\Big|\,,$$

and by the preceeding estimate and (4.11) this is close to $d(y, \Gamma^1)$. $\square$

The first estimate in (4.3) implies that, for each fixed $\xi \neq 0$, and each $t \in [-2, 2]$, the map $x \to z = x_t(x, \xi)$ is a local diffeomorphism of $\mathbb{R}^n$. It is proper since $|z - x| \le 5$, and consequently is both an open and a closed mapping, which by connectivity of $\mathbb{R}^n$ implies that it is onto. By simple connectivity it is one-to-one, and hence a global diffeomorphism of $\mathbb{R}^n$, for each fixed t and $\xi \neq 0$. We may thus take its inverse to define a map $\bar{x}_t(z, \xi)$. The inverse function theorem, together with (4.3), shows that, for $|\xi| = 1$ and $|t| \le 2$,

$$\big|\, d_{z,\xi} \bar{x}_t(z, \xi) \,\big| \lesssim 1\,. \tag{4.14}$$

For higher order derivatives, we have the following.

Corollary 4.3. *The following hold for $|\alpha| + |\beta| \ge 2$, uniformly for $|\xi| = 1$ and $|t| \le 2$. For $R \le 1$,*

$$\big|\, \partial_z^\beta \partial_\xi^\alpha \bar{x}_t(z, \xi) \,\big| \lesssim \begin{cases} 1\,, & |\beta| = 0 \\ \varepsilon\,, & |\beta| = 1 \\ \varepsilon\, R^{|\beta| - 2}\,, & |\beta| \ge 2 \end{cases} \tag{4.15}$$

For $R \ge 1$,

$$\big|\, \partial_z^\beta \partial_\xi^\alpha \bar{x}_t(z, \xi) \,\big| \lesssim \begin{cases} 1 + \varepsilon\, R^{|\alpha| - 1}\,, & |\beta| = 0 \\ \varepsilon\, R^{|\alpha| + |\beta| - 1}\,, & |\beta| \ge 1 \end{cases} \tag{4.16}$$

Proof. The estimates follow from (4.4)–(4.7), by writing $\partial_z^\beta \partial_\xi^\alpha d_{z,\xi} \bar{x}_t(z, \xi)$ as a sum of terms

$$f(d_{x,\xi} x_t)\,\big(\partial_x^{\beta_1} \partial_\xi^{\alpha_1} d_{x,\xi} x_t\big) \cdots \big(\partial_x^{\beta_k} \partial_\xi^{\alpha_k} d_{x,\xi} x_t\big)$$

where f is a rational function, smooth on the range of its argument, where

$$\alpha_1 + \cdots + \alpha_k = \alpha\,, \qquad \beta_1 + \cdots + \beta_k = \beta\,,$$

and where the right hand side is evaluated at $x = \bar{x}_t(z, \xi)$. $\square$

We introduce the phase function

$$\varphi_t(z, \xi) = \xi \cdot \bar{x}_t(z, \xi) = \sum_{i=1}^n \xi_i\, \bar{x}_t^i(z, \xi)\,. \tag{4.17}$$

This is a generating function for the symplectic map $(x, \xi) \to (x_t, \xi_t)$, in that

$$d_z \varphi_t\big(z, \xi\big) = \xi_t\big(\bar{x}_t(z, \xi), \xi\big)\,, \qquad d_\xi \varphi_t(z, \xi) = \bar{x}_t(z, \xi)\,. \tag{4.18}$$

We will need symbol bounds only for the second order in z derivatives of φ_t. To state them succinctly, we use the bracket norm

$$\langle R\rangle = \big(1+R^2\big)^{\frac12} \approx 1+R\,.$$

Lemma 4.4. *The following estimates hold, uniformly for $|\xi|=1$ and $|t|\le 2$.*

$$\big|\partial_z^\beta\partial_\xi^\alpha d_z^2\varphi_t(z,\xi)\big| \lesssim \varepsilon\, R^{|\beta|}\,\langle R\rangle^{|\alpha|}\,.$$

Proof. We use (4.18) to write $\partial_z^\beta\partial_\xi^\alpha d_z^2\varphi_t(z,\xi)$ as a sum of terms of the following form,

$$\big(d_x^k d_\xi^j\xi_t\big)\big(\bar x_t(z,\xi),\xi\big)\cdot\partial_z^{\gamma_1}\partial_\xi^{\theta_1}\bar x_t(z,\xi)\,\cdots\,\partial_z^{\gamma_k}\partial_\xi^{\theta_k}\bar x_t(z,\xi)\,,$$

where $|\gamma_1|+\cdots+|\gamma_k|=|\beta|+1$, and $j+|\theta_1|+\cdots+|\theta_k|=|\alpha|$. By (4.5), (4.15) and the fact that $\varepsilon<R$, this is dominated, for $R\le 1$, by

$$\varepsilon\, R^{k-1}\cdot R^{|\gamma|-k} = \varepsilon\, R^{|\beta|}\,.$$

For $R\ge 1$, we use the bound (4.7) and (4.16), together with $\varepsilon<1$, to dominate this by

$$\varepsilon\, R^{j+k-1}\cdot R^{|\gamma_1|+\cdots+|\gamma_k|+|\theta_1|+\cdots+|\theta_k|-k} = \varepsilon\, R^{|\alpha|+|\beta|}\,.$$

□

We now consider the estimates (3.13) and (3.14), where $t\in[0,1]$ and $\mu\ge 1$. By taking the Fourier transform of the first factor of g, we write

$$\begin{aligned} K_\varepsilon(y;t,z) &= \int e^{i\langle\eta,y-x\rangle - i\langle\xi_t,z-x_t\rangle}\,h\big(R^{-1}(\eta-\xi)\big)\,g\big(R(z-x_t)\big)\,dx\,d\xi\,d\eta \\ &= \int e^{i\langle y,\eta\rangle - i\varphi_t(z,\eta)}\,a_t(z,\eta)\,d\eta\,, \end{aligned}$$

where we define the symbol

$$a_t(z,\eta) = \int e^{i\varphi_t(z,\eta)-i\langle x,\eta\rangle - i\langle\xi_t(x,\xi),z-x_t(x,\xi)\rangle} \\ h\big(R^{-1}(\eta-\xi)\big)\,g\big(R\,(z-x_t(x,\xi))\big)\,\beta\big(\mu^{-1}\xi\big)\,dx\,d\xi\,. \tag{4.19}$$

Recall that h is a smooth function supported in the unit ball of $\mathbb{R}^n$, and β is a smooth function supported in the ball of radius $\frac14$ about $(1,0,\ldots,0)$. The function g is of Schwartz class. By taking λ large we have $R\le\frac18\mu$, so that $a_t(z,\eta)$ is supported in a ball of radius $\frac38\,\mu$ about $(\mu,0,\ldots,0)$.

Recall also that $\mu=\varepsilon\lambda$ and $R=\varepsilon\lambda^{\frac12}$, so that

$$\varepsilon\,\mu\,R^{-2} = 1\,. \tag{4.20}$$

Theorem 4.5. *The symbol $a_t(z,\eta)$ defined by (4.19) satisfies, uniformly for $t\in[0,1]$,*

$$\big|\partial_\eta^\alpha\,\langle\eta,d_\eta\rangle^k a_t(z,\eta)\big| \lesssim \langle R\rangle^{|\alpha|}\,\mu^{-|\alpha|}\,. \tag{4.21}$$

Proof. We first consider the case $k=0$ of (4.21). We will say that a function $p(x,z,\xi,\eta)$ is a symbol of size A if it satisfies bounds

$$\big|\partial_\eta^\theta\partial_\xi^\alpha p(x,z,\xi,\eta)\big| \lesssim A\cdot\langle R\rangle^{|\alpha|+|\theta|}\mu^{-|\alpha|-|\theta|}\,.$$

We then consider a more general integral of the form

$$a_t(z,\eta) = \int e^{i\varphi_t(z,\eta)-i\langle x,\eta\rangle - i\langle \xi_t(x,\xi), z-x_t(x,\xi)\rangle}$$
$$p(x,z,\xi,\eta)\, h\big(R^{-1}(\eta-\xi)\big)\, g\big(R\,(z-x_t(x,\xi))\big)\, \beta\big(\mu^{-1}\xi\big)\, dx\, d\xi\,.$$

Simple absolute bounds on g and h, and the fact that $x \to x_t(x,\xi)$ is a diffeomorphism for fixed t, ξ, show that if $|p| \lesssim 1$ then $|a_t(z,\eta)| \lesssim 1$. We next show that, if $p(x,z,\xi,\eta)$ is a symbol of size one, then $\partial_\eta a_t(z,\eta)$ can be written in the same form (that is, with modified g, h, and β which satisfy the conditions stated following (4.19)) but with p replaced by a symbol of size $\langle R\rangle \mu^{-1}$. The theorem for $k=0$ then follows by induction.

The operator ∂_{η_j} applied to the integrand can act either on the phase or on h (we ignore the effect on p, which is handled trivially.) The effect on h is the same as applying $-\partial_{\xi_j}$ to h, and integrating by parts leads to terms where ∂_{ξ_j} acts on the phase or on g. (The effect of ∂_{ξ_j} on p or β is handled trivially.) Consequently, ignoring trivial terms, the effect of applying ∂_{η_j} to the integrand is the same as multiplying it by the symbol

$$\partial_{\eta_j}\varphi_t(z,\eta) - x_j - \partial_{\xi_j}\langle \xi_t(x,\xi), z-x_t(x,\xi)\rangle - R\,\partial_{\xi_j}x_t(x,\xi)\,,$$

and we need to show that this expression is in effect a symbol of order $\langle R\rangle\,\mu^{-1}$, where "in effect" means that it can be written as such a symbol after replacing $z-x_t(x,\xi)$ by R^{-1}, and $\eta-\xi$ by R, which can be done at the expense of changing the form of g and h.

By (4.9) and (4.10) and homogeneity, the function $\partial_{\xi_j}x_t(x,\xi)$ is a symbol of size μ^{-1}, which handles the last term. By (4.18), we may write

$$\begin{aligned} d_\eta\varphi_t(z,\eta) - d_\xi\varphi_t(z,\xi) = {} & \bar{x}_t(z,\eta) - \bar{x}_t(z,\xi) \\ = {} & \int_0^1 (d_\xi\bar{x}_t)(z, s\eta+(1-s)\xi)\cdot(\eta-\xi)\,. \end{aligned}$$

Replacing $\eta-\xi$ by R, the estimates (4.15) and (4.16) and homogeneity show that this is in effect a symbol of size $R\,\mu^{-1}$.

Since the terms $(x_t,\xi_t;x,\xi)$ lie on the graph of $(z, d_z\varphi_t(z,\xi); d_\xi\varphi_t(z,\xi),\xi)$, we observe that $\xi_t(x,\xi) = \big(d_z\varphi_t\big)(x_t(x,\xi),\xi)$, and $\langle x,\xi\rangle = \varphi_t(x_t(x,\xi),\xi)$. We are then left to show that the expression

$$\partial_{\xi_j}\Big[\,\varphi_t(z,\xi) - \varphi_t(x_t(x,\xi),\xi) - \big(z-x_t(x,\xi)\big)\cdot\big(d_z\varphi_t\big)(x_t(x,\xi),\xi)\,\Big]$$

is in effect a symbol of size $\langle R\rangle\,\mu^{-1}$. We write this as

$$\partial_{\xi_j}\Big[\,\big(z-x_t(x,\xi)\big)^2\cdot\int_0^1 \big(d_z^2\varphi_t\big)(sz+(1-s)x_t(x,\xi),\xi)\,(1-s)\,ds\,\Big]\,. \tag{4.22}$$

Finally, we note that, uniformly for each fixed s, $\big(d_z^2\varphi_t\big)(sz+(1-s)x_t(x,\xi),\xi)$ is a symbol of size $\varepsilon\,\mu$. This is a simple consequence of Lemma 4.4 and the fact that $\partial_{\xi_j}x_t(x,\xi)$ is a symbol of size μ^{-1}. The first order derivatives in ξ of $(d_z^2\varphi_t)(\cdots)$ are then symbols of size $\varepsilon\,\langle R\rangle$. The expression (4.22) is then in effect a symbol of size

$$R^{-1}\cdot\mu^{-1}\cdot\varepsilon\,\mu + R^{-2}\cdot\varepsilon\,\langle R\rangle \lesssim \langle R\rangle\,\mu^{-1}\,,$$

where we use (4.20) for the last inequality.

We now consider the case $k \geq 1$ of (4.21). Note that

$$\langle \eta, d_\eta \rangle^k a(z,\eta) = (r\partial_r)^k a(z, r\eta)\big|_{r=1}$$

and, by changing variables $\xi \to r\xi$, that

$$a(z, r\eta) = r^n \int e^{ir[\,\varphi_t(z,\eta) - i\langle x,\eta\rangle - i\langle \xi_t(x,\xi), z - x_t(x,\xi)\rangle\,]} \\ h\bigl(rR^{-1}(\eta-\xi)\bigr)\, g\bigl(R\,(z - x_t(x,\xi))\bigr)\, \beta\bigl(r\mu^{-1}\xi\bigr)\, dx\, d\xi\,.$$

Acting on the h or β terms with ∂_r simply changes the particular form of h or β. Thus, the theorem will follow by showing that the phase function

$$\varphi_t(z,\eta) - \langle x,\eta\rangle - \langle \xi_t(x,\xi), z - x_t(x,\xi)\rangle$$

is in effect a symbol of size 1. The function

$$\varphi_t(z,\xi) - \langle x,\xi\rangle - \langle \xi_t(x,\xi), z - x_t(x,\xi)\rangle$$

is the same as the expression in braces in (4.22), and the first line following (4.22) shows that this is in effect a symbol of size $R^{-2}\,\varepsilon\,\mu = 1$. It remains to show that

$$\begin{aligned} \varphi_t(z,\eta) - \varphi_t(z,\xi) - \langle x,\eta\rangle + \langle x,\xi\rangle &= \int_0^1 \langle d_\xi\varphi_t(z, s\eta + (1-s)\xi) - x, \eta - \xi\rangle\, ds \\ &= \int_0^1 \langle \bar{x}_t(z, s\eta + (1-s)\xi) - x, \eta - \xi\rangle\, ds \end{aligned}$$

is in effect a symbol of size 1, which will follow if we show that the last integrand is in effect a symbol of size 1, uniformly for each $s \in [0,1]$.

We now write $x = \bar{x}_t(x_t(x,\xi),\xi)$, and note that

$$\langle \bar{x}_t(z,\xi) - \bar{x}_t(x_t(x,\xi),\xi), \eta - \xi\rangle = (z - x_t(x,\xi))\cdot(\eta-\xi)\cdot \int_0^1 d_z\bar{x}_t(sz + (1-s)x_t(x,\xi),\xi)\, ds\,.$$

Corollary 4.3, with the fact that $d_\xi x_t(x,\xi)$ is a symbol of size μ^{-1}, shows that the integral is a symbol of size 1, and hence this term is in effect of size $R^{-1}\cdot R\cdot 1 = 1$.

This leaves the term

$$\langle \bar{x}_t(z, s\eta + (1-s)\xi) - \bar{x}_t(z,\xi), \eta - \xi\rangle = (\eta-\xi)^2 \cdot \int_0^1 d_\xi\bar{x}_t(z, \xi + rs(\eta-\xi))\, dr\,.$$

The last integral is a symbol of size μ^{-1}, so this term is in effect of size $R^2\,\mu^{-1} = \varepsilon\,.$ □

Proof of the estimate (3.13)**.** We use a second-dyadic decomposition partition of unity over angles (see Stein [13]) to write $a_1(z,\eta) = \sum_\nu a^\nu(z,\eta)$, where $a^\nu(z,\eta)$ is supported in a cone of angle $\mu^{-\frac12}$ about the unit vector ω_ν, and satisfies the estimates

$$\bigl|\langle \omega_\nu, d_\eta\rangle^k \partial_\eta^\alpha a^\nu(z,\eta)\bigr| \lesssim \mu^{-k-\frac12|\alpha|}\,. \tag{4.23}$$

These estimates follow by (4.21), and the fact that $R \leq \mu^{\frac12}$. We may take the vectors ω_ν to be separated by distance $\mu^{-\frac12}$.

Next, on the support of $a^\nu(z,\eta)$ we write

$$\varphi_1(z,\eta) = \langle (d_\eta\varphi_1)(z,\omega_\nu), \eta\rangle + R^\nu(z,\eta)\,,$$

where

$$\big|\langle\omega_\nu, d_\eta\rangle^k\partial_\eta^\alpha R^\nu(z,\eta)\big| \lesssim \mu^{-k-\frac12|\alpha|}\,. \tag{4.24}$$

The proof of the estimates (4.24) is essentially from Seeger-Sogge-Stein [9], where it was established for smooth phase functions. To prove them, note that by homogeneity we may consider $k=0$, and observe that $R^\nu(z,\eta)$ is equal to the error for second order Taylor expansion of $\varphi_1(z,\eta)$ in η about any conveniently chosen point on the ray through ω_ν. By fixing the point within distance $\mu^{\frac12}$ of η, the estimates then follow from Corollary 4.3, and the fact that $d_\eta^2\varphi_1(z,\eta)=d_\eta\bar x_1(z,\eta)$.

We now write $K_\varepsilon(y;1,z)=\sum_\nu K_\varepsilon^\nu(y;z)$, with

$$\begin{aligned} K_\varepsilon^\nu(y;z) &= \int e^{i\langle y,\eta\rangle - i\varphi_1(z,\eta)}\,a^\nu(z,\eta)\,d\eta \\ &= \int e^{i\langle y-x_\nu,\eta\rangle}\Big[e^{-iR^\nu(z,\eta)}\,a^\nu(z,\eta)\Big]\,d\eta\,, \end{aligned}$$

where we set $x_\nu=(d_\eta\varphi_1)(z,\omega_\nu)=\bar x_1(z,\omega_\nu)$. The term in braces satisfies the same estimates (4.23) as $a^\nu(z,\eta)$, and thus

$$\big|K_\varepsilon^\nu(y,z)\big| \lesssim \mu^{\frac{n+1}{2}}\,\big(1+\mu\,|\langle\omega_\nu,y-x_\nu\rangle| + \mu\,|y-x_\nu|^2\,\big)^{-N}\,.$$

The point x_ν is the unique point on Γ_z at which the normal equals ω_ν. Consequently,

$$|\,x_\nu - x_{\nu'}\,| \approx |\,\omega_\nu-\omega_{\nu'}\,|$$

We may assume that $d(y,\Gamma_z)\le .01$, since the result is immediate otherwise. Let x_0 denote the unique point on Γ_z closest to y, and ω_0 the normal at x_0, so that

$$d(y,\Gamma_z) = |\,y-x_0| = |\langle\omega_0,y-x_0\rangle|\,.$$

We claim that

$$|\langle\omega_\nu,y-x_\nu\rangle| + |\,y-x_\nu|^2 \approx |\langle\omega_0,y-x_0\rangle| + |\,\omega_0-\omega_\nu|^2\,. \tag{4.25}$$

To show this, we first observe that, since $|y-x_0|\le|y-x_\nu|$, then

$$|\,y-x_\nu|^2 \approx |\,y-x_0|^2 + |\,x_0-x_\nu|^2 \approx |\,y-x_0|^2 + |\,\omega_0-\omega_\nu|^2\,.$$

On the other hand,

$$\begin{aligned} \Big|\,|\langle\omega_0,y-x_0\rangle| - |\langle\omega_\nu,y-x_\nu\rangle|\,\Big| \;&\le\; |\langle\omega_0-\omega_\nu,y-x_0\rangle| + |\langle\omega_0,x_\nu-x_0\rangle| \\ &\lesssim\; |\,y-x_0|^2 + |\,\omega_0-\omega_\nu|^2 \end{aligned}$$

where we use that Γ_z is C^2 close to the unit sphere and ω_0 is normal to Γ_z at x_0. Together, these imply (4.25).

We now have

$$\begin{aligned} \big|K_\varepsilon(y;z)\big| \;&\le\; \sum_\nu \big|K_\varepsilon^\nu(y,z)\big| \\ &\le\; \mu^{\frac{n+1}{2}}\sum_\nu \big(1+\mu\,d(y,\Gamma_z)+\mu\,|\,\omega_0-\omega_\nu|^2\,\big)^{-N} \\ &\le\; \mu^{\frac{n+1}{2}}\,\big(1+\mu\,d(y,\Gamma_z)\,\big)^{-N} \end{aligned}$$

where we use that the points ω_ν are $\mu^{\frac12}$ evenly spaced over the unit sphere. $\square$

Proof of the estimate (3.14)**.** In this case $\mu = 1$, $R \le 1$, and $t \in [0,1]$. It follows by (4.14) and (4.15) that

$$\big|\, \partial_\eta^\alpha \bar{x}_t(z,\eta)\,\big| \lesssim 1\,.$$

We write $\varphi_t(z,\eta) = \langle \bar{x}_t(z,\eta), \eta\rangle$. Since $|\,\bar{x}_t(z,\eta) - z| \le 2$, together with (4.21) this yields

$$\Big|\, \partial_\eta^\alpha \Big(e^{i\langle z,\eta\rangle - \varphi_t(z,\eta)}\, a_t(z,\eta)\Big)\,\Big| \lesssim 1\,.$$

The symbol $a_t(z,\eta)$ is supported in the set $|\eta| \le 2$. Writing

$$K_\varepsilon(y;t,z) = \int e^{i\langle y-z,\eta\rangle}\,\Big(e^{i\langle z,\eta\rangle - \varphi_t(z,\eta)}\, a_t(z,\eta)\Big)\, d\eta$$

easily yields (3.14). □

5. Strichartz estimates

We provide here details on how the results of the previous sections can be adapted to give an alternate proof of Strichartz estimates for the homogeneous wave equation, for time dependent metrics satisfying (2.1)–(2.2). Tataru established $L^r_t L^s_x \to L^p_t L^q_x$ estimates for the inhomogeneous equation for such metrics in [14]. The steps in section 3 which reduce matters to bounds for the operator W_r work only for the homogeneous $H^\delta \to L^p_t L^q_x$ estimates, however.

Theorem 5.1. *Suppose that the metric* $\mathrm{g}^{ij}(t,x)$ *satisfies conditions* (2.1)–(2.2)*. Let* u *satisfy the Cauchy problem*

$$\sum_{i,j=0}^{n} \partial_i\Big(\mathrm{g}^{ij}(t,x)\,\partial_j u(t,x)\Big) = 0\,, \qquad u(0,x) = f(x)\,, \quad \partial_t u(0,x) = g(x)\,. \tag{5.1}$$

Then for $n \ge 2$ *the following estimates hold,*

$$\|u\|_{L^p_t L^q_x(\mathbb{R}^n\times[-1,1])} \le C_{p,q}\,\big(\,\|f\|_{H^{\delta(p,q)}} + \|g\|_{H^{\delta(p,q)-1}}\big)\,,$$

provided that $2 \le p \le \infty$, $2 \le q < \infty$,

$$\delta(p,q) = n\Big(\frac12 - \frac1q\Big) - \frac1p\,, \qquad \frac2p + \frac{n-1}{q} \le \frac{n-1}{2}\,,$$

and $\delta(p,q) \le 2$.

The condition $\delta(p,q) \le 2$ is necessary to apply Lemma 2.2 to localize matters to a fixed dyadic frequency range. This condition may be dropped if one considers instead the estimate

$$\|\langle D_x\rangle^{1-\delta(p,q)} u\|_{L^p_t L^q_x(\mathbb{R}^n\times[-1,1])} \le C_{p,q}\,\big(\,\|f\|_{H^1} + \|g\|_{L^2}\big)\,.$$

For the Minkowski wave operator the above theorem was established under the restriction $p > 2$ by Lindblad and Sogge [7]. It was shown there that such estimates are a consequence of the $L^1 \to L^\infty$ dispersive estimates for the wave group. The endpoint case $p = 2$ was settled by Keel and Tao [6] using a bilinear interpolation scheme. That paper also formulated the general results for $p \ge 2$ in an abstract Hilbert space setting that works well for the operators W_r of section 3.

We start by observing that, for $q < \infty$, the results of Section 2 reduce matters to proving the following analogue of Theorem 2.5.

Theorem 5.2. *Suppose that $\partial_t u(t,x) = -i\, p(t,x,D_x)\, u(t,x) + F(t,x)$, where the symbol of p satisfies conditions (2.20)–(2.24). Assume also that, for all t, the partial Fourier transform $\widehat{u}(t,\xi)$ is supported in the region $\frac{1}{4}\lambda < |\xi| < \lambda$. Then*

$$\|u\|_{L^p_t L^q_x([-1,1]\times\mathbb{R}^n)} \lesssim \lambda^{\delta(p,q)} \Big(\|u\|_{L^\infty_t L^2_x([-1,1]\times\mathbb{R}^n)} + \|F\|_{L^1_t L^2_x([-1,1]\times\mathbb{R}^n)} \Big),$$

for p, q, and $\delta(p,q)$ as above.

The wave-transform methods of section 3 then reduce matters to proving the mapping properties, uniformly over $r \in [-1,1]$,

$$\|S_\lambda(D_x) W_r \tilde{f}\|_{L^p_t L^q_x([-1,1]\times\mathbb{R}^n)} \le C_{p,q}\, \lambda^{\delta(p,q)}\, \|\tilde{f}\|_{L^2(\mathbb{R}^{2n})} \tag{5.2}$$

where $\tilde{f}$ is a function of (x,ξ), and as before

$$(W_r \tilde{f}\,)(s,y) = T^*_\lambda \big[\tilde{f} \circ \chi_{r,s}\,\big](y)\,.$$

To work in the setting of Theorem 1.2 of [6], we fix r and let

$$U(s)\tilde{f} = S_\lambda(D_x)(W_r \tilde{f}\,)(s,\,\cdot\,)\,.$$

Since $\chi_{r,s}$ preserves $dx \wedge d\xi$, the L^2 boundedness of T_λ implies

$$\|U(s)\tilde{f}\,\|_{L^2(\mathbb{R}^n)} \le \|\tilde{f}\,\|_{L^2(\mathbb{R}^{2n})}\,.$$

On the other hand,

$$\|U(s)U(t)^* g\|_{L^\infty(\mathbb{R}^n)} \lesssim \lambda^n \big(\,1 + \lambda\,|s-t|\,\big)^{-\frac{n-1}{2}} \|g\|_{L^1(\mathbb{R}^n)}\,.$$

To verify this, we note that $U(s)U(t)^* = S_\lambda(D_x) K_{s,t} S_\lambda(D_x)$, where $K_{s,t}$ is the integral kernel of section 3 for fixed variables s and t. The preceding bound is then a consequence of (3.15) and (3.16).

The setting of [6] thus applies with $H = L^2(dx\, d\xi)$. Precisely, conditions (1) and (3) of that paper are satisfied by the rescaled operator $\delta_\lambda \circ U(s)$, where δ_λ is space-time dilation by λ, normalized to preserve the $L^2(\mathbb{R}^n)$ norm. Theorem 1.2 of [6] then yields (5.2). $\square$

References

[1] R. Coifman and Y. Meyer, *Commutateurs d'integrales singulieres et operateurs multilineaires*, Ann. Inst. Fourier Grenoble **28** (1978), 177-202.

[2] A. Córdoba and C. Fefferman, *Wave packets and Fourier integral operators*, Comm. Partial Differential Equations **3** (1978), 979–1005.

[3] J-M. Delort, *F.B.I. Transformation*, Second Microlocalization and Semilinear Caustics. Lecture Notes in Mathematics **1522**, Springer-Verlag, Berlin, 1992.

[4] D. Gilbarg and N. Trudinger, Elliptic Partial Differential Equations of Second Order, 2nd edition. Springer, New York, 1983.

[5] L. Hormander, Lectures on Nonlinear Hyperbolic Equations. Springer-Verlag, Berlin, 1997.

[6] M. Keel and T. Tao, *Endpoint Strichartz estimates*, Amer. J. Math. **120** (1998), 955–980.

[7] H. Lindblad and C. Sogge, *On existence and scattering with minimal regularity for semilinear wave equations*, J. Funct. Anal. **130** (1995), 357-426.

[8] G. Mockenhaupt, A. Seeger and C. Sogge, *Local smoothing of Fourier integrals and Carleson-Sjölin estimates*, J. Amer. Math. Soc. **6** (1993), 65–130.

[9] A. Seeger, C. Sogge and E. Stein, *Regularity properties of Fourier integral operators*, Annals of Math. **134** (1991), 231–251.

[10] H. Smith, *A parametrix construction for wave equations with $C^{1,1}$ coefficients*, Annales de l'Institut Fourier **48** (1998), 797–835.

[11] H. Smith and C. Sogge, *On Strichartz and eigenfunction estimates for low regularity metrics*, Math. Res. Lett. **1** (1994), 729–737.

[12] C. Sogge, *Concerning the L^p norm of spectral clusters for second order elliptic operators on compact manifolds*, J. Funct. Anal. **77** (1988), 123–134.

[13] E.M. Stein, Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton University Press, Princeton, 1993.

[14] D. Tataru, *Strichartz estimates for operators with nonsmooth coefficients III*, J. Amer. Math. Soc. **15** (2002), 419–442.

[15] M. Taylor, Pseudodifferential Operators and Nonlinear PDE. Progress in Mathematics, vol. 100, Birkhäuser, Boston, 1991.

[16] N. Varopoulos, *Hardy-Littlewood theory for semigroups*, J. Funct. Anal. **63** (1985), 240–260.

Department of Mathematics, University of Washington, Seattle, WA 98195

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