SHARP LOCAL WELL-POSEDNESS RESULTS FOR THE NONLINEAR WAVE EQUATION

HART SMITH AND DANIEL TATARU

Abstract. This article is concerned with local well-posedness of the Cauchy problem for second order quasilinear hyperbolic equations with rough initial data. The new results obtained here are sharp in low dimension.

1. Introduction

1.1. **The results.** We consider in this paper second order, nonlinear hyperbolic equations of the form

$$\mathrm{g}^{ij}(u)\,\partial_i\partial_j u = q^{ij}(u)\,\partial_i u\,\partial_j u \tag{1.1}$$

on $\mathbb{R}\times\mathbb{R}^n$, with Cauchy data prescribed at time 0,

$$u(0,x) = u_0(x)\,, \qquad \partial_0 u(0,x) = u_1(x)\,. \tag{1.2}$$

The indices i and j run from 0 to n, with the index 0 corresponding to the time variable. The symmetric matrix $\mathrm{g}^{ij}(u)$ and its inverse $\mathrm{g}_{ij}(u)$ are assumed to satisfy the hyperbolicity condition, that is, have signature $(n,1)$. The functions g^{ij}, g_{ij} and q^{ij} are assumed to be smooth, bounded, and have globally bounded derivatives as functions of u. To insure that the level surfaces of t are space-like we assume that $\mathrm{g}^{00}=-1$. We then consider the following question:

> *For which values of s is the problem* (1.1) *and* (1.2) *locally well-posed in $H^s\times H^{s-1}$?*

In general, well-posedness involves existence, uniqueness and continuous dependence on the initial data. Naively, one would hope to have these properties hold for solutions in $C(H^s)\cap C^1(H^{s-1})$, but it appears that there is little chance to establish uniqueness under this condition for the low values of s that we consider in this paper. Our definition of well-posedness thus includes an additional assumption on the solution

Date: February 4, 2002.

Research partially supported by NSF grant DMS-9970407.
Research partially supported by NSF grant DMS-9970297.

u to insure uniqueness, while also providing useful information about the solution.

Definition 1.1. *We say that the Cauchy problem* (1.1) *and* (1.2) *is locally well-posed in* $H^s \times H^{s-1}$ *if, for each* $R > 0$*, there exist constants* $T, M, C > 0$*, so that the following properties are satisfied:*

(WP1) *For each initial data set* (u_0, u_1) *satisfying*

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

there exists a unique solution $u \in C\big([-T,T]; H^s\big) \cap C^1\big([-T,T]; H^{s-1}\big)$ *subject to the condition* $du \in L^2\big([-T,T]; L^\infty\big)$.

(WP2) *The solution* u *depends continuously on the initial data in the above topologies.*

(WP3) *The solution* u *satisfies*

$$\|du\|_{L^2_t L^\infty_x} + \|du\|_{L^\infty_t H^{s-1}_x} \leq M\,.$$

(WP4) *For* $1 \leq r \leq s+1$*, and for each* $t_0 \in [-T,T]$*, the linear equation*

$$\begin{cases} \mathrm{g}^{ij}(u)\, \partial_i \partial_j v = 0\,, & (t,x) \in [-T,T] \times \mathbb{R}^n\,, \\ v(t_0,\cdot) = v_0 \in H^r(\mathbb{R}^n)\,, & \partial_0 v(t_0,\cdot) = v_1 \in H^{r-1}(\mathbb{R}^n)\,, \end{cases} \tag{1.3}$$

admits a solution $v \in C\big([-T,T]; H^r\big) \cap C^1\big([-T,T]; H^{r-1}\big)$, *and the following estimates hold:*

$$\|v\|_{L^\infty_t H^r_x} + \|\partial_0 v\|_{L^\infty_t H^{r-1}_x} \leq C\, \|(v_0, v_1)\|_{H^r \times H^{r-1}}\,. \tag{1.4}$$

Additionally, the following estimates hold, provided $\rho < r - \frac{3}{4}$ *if* $n = 2$*, and* $\rho < r - \frac{n-1}{2}$ *if* $n \geq 3$*,*

$$\begin{aligned} \|\langle D_x\rangle^\rho v\|_{L^4_t L^\infty_x} &\leq C\, \|(v_0, v_1)\|_{H^r \times H^{r-1}}\,, \quad n = 2\,, \\ \|\langle D_x\rangle^\rho v\|_{L^2_t L^\infty_x} &\leq C\, \|(v_0, v_1)\|_{H^r \times H^{r-1}}\,, \quad n \geq 3\,, \end{aligned} \tag{1.5}$$

and the same estimates hold with $\langle D_x\rangle^\rho$ *replaced by* $\langle D_x\rangle^{\rho-1} d$.

We prove the result for a sufficiently small T, depending on R. However, it is a simple matter to see that uniqueness, as well as condition (WP4), holds up to any time T for which there exists a solution $u \in C\big([-T,T]; H^s\big) \cap C^1\big([-T,T]; H^{s-1}\big)$ which satisfies $du \in L^2\big([-T,T]; L^\infty\big)$.

Observe that we do not ask for uniformly continuous dependence on the initial data. This in general is not expected to hold for nonlinear hyperbolic equations. Indeed, even a small perturbation of the solution suffices in order to change the Hamilton flow for the corresponding linear equation, which in turn modifies the propagation of high frequency solutions.

As a consequence of the $L_t^2 L_x^\infty$ bound for du it follows that if the initial data is of higher regularity, then the solution u retains that regularity up to time T. Hence, one can naturally obtain solutions for rough initial data as limits of smooth solutions. This switches the emphasis to establishing *a priori* estimates for smooth solutions. One can think of the $L_t^2 L_x^\infty$ bound for du as a special case of (1.5), which is a statement about Strichartz estimates for the linear wave equation. Establishing this estimate plays a central role in this article.

Our main result is the following:

Theorem 1.2. *The Cauchy problem* (1.1) *and* (1.2) *is locally well-posed in $H^s \times H^{s-1}$ provided that*

$$s > \frac{n}{2} + \frac{3}{4} \qquad \textit{for} \qquad n = 2\,,$$

$$s > \frac{n+1}{2} \qquad \textit{for} \qquad n = 3, 4, 5\,.$$

Remark 1.3. There are precisely two places in this paper at which our argument breaks down for $n \geq 6$, occurring in Lemmas 8.5 and 8.6. Both are related to the orthogonality argument for wave packets. Presumably this could be remedied with a more precise analysis of the geometry of the wave packets, but we do not pursue this question here.

As a byproduct of our result, it also follows that certain Strichartz estimates hold for the corresponding linear equation (1.3). Interpolation of (1.4) with (1.5), combined with Sobolev embedding estimates, yields

$$\|\langle D_x\rangle^\rho v\|_{L_t^p L_x^q} \leq C\,\|(v_0, v_1)\|_{H^r \times H^{r-1}}\,, \qquad \frac{2}{p} + \frac{1}{q} \leq \frac{1}{2}\,, \qquad n = 2,$$

$$\|\langle D_x\rangle^\rho v\|_{L_t^p L_x^q} \leq C\,\|(v_0, v_1)\|_{H^r \times H^{r-1}}\,, \qquad \frac{1}{p} + \frac{1}{q} \leq \frac{1}{2}\,, \qquad n = 3, 4, 5,$$

provided that

$$1 \leq r \leq s + 1\,, \qquad \text{and} \qquad r - \rho > \frac{n}{2} - \frac{1}{p} - \frac{n}{q}\,.$$

Note that in the usual Strichartz estimates (which hold for a smooth metric g) one permits equality in the second condition on ρ. The estimates we prove in this paper have a logarithmic loss in the frequency, so we need the strict inequality above. Also, we do not get the full range of $L_t^p L_x^q$ spaces for $n \geq 4$. This remains an open question for now.

1.2. **Comments.** To gain some intuition into our result it is useful to consider two aspects of the equation. The first aspect is scaling. We note that the equation (1.1) is invariant with respect to the dimensionless scaling $u(t,x) \to u(rt, rx)$. This scaling preserves the Sobolev space of exponent $s_c = \frac{n}{2}$, which is then, heuristically, a lower bound for the range of permissible s.

The second aspect to be considered is that of blow-up. There are two known mechanisms for blow-up; see Alinhac [1]. The simplest blowup mechanism is a space-independent type blow-up, which can occur already in the case of semilinear equations. Roughly, the idea is that if we eliminate the spatial derivatives from the equation, then one obtains an ordinary differential equation, which can have solutions that blow-up as a negative power of $(t-T)$. For a hyperbolic equation, this type of blow-up is countered by the dispersive effect, but only provided that s is sufficiently large. On the other hand, for the quasilinear equation (1.1) one can also have blow-up caused by geometric focusing. This occurs when a family of null geodesics come together tangentially at a point. Both patterns were studied by Lindblad [13],[14]. Surprisingly, they yield blow-up at the same exponent s, namely $s = \frac{n+5}{4}$. Together with scaling, this leads to the restriction

$$s > \max\left\{\frac{n}{2}, \frac{n+5}{4}\right\} .$$

Comparing this with Theorem 1.2, we see that for $n = 2$ and $n = 3$ the exponents match, therefore both our result and the counterexample are sharp. However, if $n \geq 4$ then there is a gap, and it is not clear whether one needs to improve the counterexamples or the positive result. For comparison purposes one should consider the semilinear equation

$$\Box u = |du|^2 .$$

For this equation it is known (see Tataru [21]) that well-posedness holds for s as above, so that the counterexamples are sharp. However, if one restricts the allowed tools to energy and Strichartz estimates, which are the tools used in this paper, then it is only possible to deduce the more restrictive range in Theorem 1.2. Adapting the ideas in [21] to quasilinear equations appears intractable for now.

To describe the ideas used to establish Theorem 1.2, we recall a classical result[1]:

[1]See the footnote following Lemma 2.3

Lemma 1.4. *Let u be a smooth solution to* (1.1) *and* (1.2) *on* $[0, T]$. *Then, for each $s \geq 0$, the following estimate holds*

$$\|du(t)\|_{H^{s-1}} \lesssim \|du(0)\|_{H^{s-1}} e^{c \int_0^t \|du(h)\|_\infty dh} . \tag{1.6}$$

For integer values of s this result is due to Klainerman [10]. For non-integer s, the argument of Klainerman needs to be combined with a more recent commutator estimate of Kato-Ponce [8]. As an immediate consequence, one obtains

Corollary 1.5. *Let u be a smooth solution to* (1.1) *and* (1.2) *on* $[0, T)$ *which satisfies* $\|du\|_{L^1_t L^\infty_x} < \infty$. *Then u is smooth at time T, and can therefore be extended as a smooth solution beyond time T.*

Thus, to establish existence of smooth solutions, one seeks to establish *a priori* bounds on $\|du\|_{L^1_t L^\infty_x}$. In case $s > \frac{n}{2} + 1$, one can obtain such bounds from the Sobolev embedding $H^s \subset L^\infty$. A simple iteration argument then leads to the classical result of Hughes-Kato-Marsden [6] of well-posedness for $s > \frac{n}{2} + 1$. Note that in this case one obtains $L^\infty_t L^\infty_x$ bounds on du instead of $L^1_t L^\infty_x$. The difference in scaling between L^1_t and L^∞_t corresponds to the one derivative difference between the classical existence result and the scaling exponent.

To improve upon the classical existence result one thus seeks to establish bounds on $\|du\|_{L^p_t L^\infty_x}$, for $p < \infty$. This leads naturally to considering the Strichartz estimates for the operator $\Box_{g(u)}$. For solutions u to the constant coefficient wave equation $\Box u = 0$, the following estimates are known to hold:

$$\|du\|_{L^4_t L^\infty_x} \lesssim \|(u_0, u_1)\|_{H^s \times H^{s-1}} , \qquad s > \tfrac{7}{4} , \qquad n = 2 ,$$

$$\|du\|_{L^2_t L^\infty_x} \lesssim \|(u_0, u_1)\|_{H^s \times H^{s-1}} , \qquad s > \tfrac{n+1}{2} , \qquad n \geq 3 .$$

To establish such estimates with $\Box$ replaced by $\Box_{g(u)}$, however, requires dealing with operators with rough coefficients. Indeed, at first glance one is faced with having only bounds on $\|dg\|_{L^2_t L^\infty_x \cap L^\infty_t H^{s-1}_x}$. (Here and below, for simplicity we discuss the case $n \geq 3$.)

The first Strichartz estimates for the wave equation with variable coefficients were obtained in Kapitanskii [7] and Mockenhaupt-Seeger-Sogge [15], in the case of smooth coefficients. The first result for rough coefficients is due to Smith [17], who used wave packet techniques to show that the Strichartz estimates hold under the condition $g \in C^2$, for dimensions $n = 2$ and $n = 3$. At the same time, counterexamples constructed in Smith-Sogge [18] showed that for all $\alpha < 2$ there exist $g \in C^\alpha$ for which the Strichartz estimates fail.

The first result on Strichartz estimates leading to an improvement in the well-posedness problem for the nonlinear wave equation was obtained in Tataru [22]; roughly it says that if $d\mathrm{g} \in L^2_t L^\infty_x$, then the Strichartz estimates hold with a 1/4 derivative loss. This gives well-posedness for the nonlinear problem with $s > \frac{n+1}{2} + \frac{1}{4}$; independently the same result was also obtained by Bahouri-Chemin [3]. Shortly afterward, the Strichartz estimates were established in all dimensions for $\mathrm{g} \in C^2$ in Tataru [23], a condition that was subsequently relaxed in Tataru [20], where the full estimates are established provided that the coefficients satisfy $d^2\mathrm{g} \in L^1_t L^\infty_x$. As a byproduct, this last estimates implies Strichartz estimates with a loss of $\frac{1}{6}$ derivative in the case $d\mathrm{g} \in L^1_t L^\infty_x$, and hence well-posedness for (1.1) and (1.2) for Sobolev indices $s > \frac{n+1}{2} + \frac{1}{6}$. Around the same time, Bahouri-Chemin [2] improved their earlier 1/4 result to slightly better than 1/5. This line of attack for the nonlinear problem, however, reached a dead end when Smith-Tataru [16] showed that the $\frac{1}{6}$ loss is sharp for general metrics of regularity C^1.

Thus, to obtain an improvement over the 1/6 result, one needs to exploit the additional geometric information on the metric g that comes from the fact that g itself is a solution an equation of type (1.1). The first work to do so was that of Klainerman-Rodnianski [11], where for $n = 3$ the well-posedness was established for $s > \frac{n+1}{2} + \frac{2-\sqrt{3}}{2}$. The central idea is that for solutions u to $\Box_{\mathrm{g}} u = 0$, one has better estimates on derivatives of u in directions tangent to null light cones. This in turn leads to a better regularity of tangential components of the curvature tensor than one would expect at first glance, and hence to better regularity of the null cones themselves. A key role in improving the regularity of the tangential curvature components is played by an observation of Klainerman [9] that the Ricci component $\mathrm{Ric}(l, l)$ admits a decomposition which yields improved regularity upon integration over a null geodesic.

The present work follows the same tack, in exploiting the improved regularity of solutions on null surfaces. In this paper, we work with foliations of space-time by null hypersurfaces corresponding to plane waves rather than light cones, but the principle difference appears to be in the machinery used to establish the Strichartz estimates. In this work we are able to establish such estimates without making reference to the variation of the geodesic flow field as one moves from one null surface to another (other than using estimates which follow immediately from the regularity of the individual surfaces themselves.) We note that recently Klainerman and Rodnianski have announced the conclusion of Theorem 1.2 in the case of the three dimensional vacuum Einstein equations, where the condition $\mathrm{Ric} = 0$ allows one to obtain some control over normal derivatives of the geodesic flow field l in terms of tangential derivatives of l.

Although all the results quoted above point in the same direction, the methods used are quite different. The idea of Bahouri and Chemin in [3] and [2] was to push the classical Hadamard parametrix construction as far as possible, on small time intervals, and then to piece together the results measuring the loss in terms of derivatives. The results in Tataru [22],[23] and [20], are based on the use of the FBI transform as a precise tool to localize both in space and in frequency. This leads to parametrices which resemble Fourier integral operators with complex phase, where both the phase and the symbol are smooth precisely on the scale of the localization provided by the imaginary part of the phase. The work of Klainerman-Rodnianski [11] is based on energy estimates obtained after commuting the equation with well-chosen vector fields. Strichartz estimates are then obtained following a vector field approach developed in [9].

A common point of the three approaches above is a paradifferential localization of the solution at a given frequency λ, followed by a truncation of the coefficients at frequency λ^a for some $a < 1$. Interestingly enough, it is precisely this truncation of the coefficients which is absent in the present paper. Our argument here relies instead on a wave-packet parametrix construction for the nontruncated metric $\mathrm{g}(u)$. This involves representing approximate solutions to the linear equation as a square summable superposition of wave packets, which are special approximate solutions to the linear equation, that are highly localized in phase space. The use of wave packets of such localization to represent solutions to the linear equation is inspired by the work of Smith [17], but the ansatz for the development of such packets, as well as the orthogonality arguments for them, is considerably more delicate in this paper due to the decreased regularity of the metric.

1.3. **Overview of the paper.** The next two sections of this paper are concerned with reducing the proof of Theorem 1.2 to establishing an existence result for smooth data of small norm. Precisely, in section 2 we use energy type estimates to obtain uniqueness and stability results, and thus reduce Theorem 1.2 to an existence result for smooth initial data, namely Proposition 2.1. Section 3 contains scaling and localization arguments which further reduce the problem to establishing time $T = 1$ existence for the case of smooth, compactly supported data of small norm, namely Proposition 3.1.

In Section 4 we present the proof of Proposition 3.1 by the continuity method. At the heart of this proof is a recursive estimate on the regularity of the solutions to the nonlinear equation, stated in Proposition 4.1. For the recursion argument to work, in addition to controlling the norm of the solution u in the Sobolev and $L^2_t L^\infty_x$ norms, we also need to control an appropriate norm of the characteristic foliations by plane waves associated to $\mathrm{g}(u)$. This additional information is collected in the nonlinear G functional.

The core of the paper is devoted to the proof of the estimates used in Proposition 4.1. In Section 5 we study the geometry of the plane wave surfaces; Proposition 5.2 contains the recursive estimate for the G functional. A key role is played by a decomposition of the tangential curvature components stated in Lemma 5.8, analogous to the decomposition for $\mathrm{Ric}(l, l)$ in [9]. It then remains to establish certain dispersive type estimates for the linear equation with metric $\mathrm{g}(u)$.

In Section 6 we study the geometry of characteristic light cones, which plays an essential role for the orthogonality and dispersive estimates. Section 7 contains a paradifferential decomposition which allows us to localize in frequency and reduce the dispersive estimates to their dyadic counterparts.

Section 8 contains the construction of a parametrix for the linear equation. We start by using the information we have for the characteristic plane wave surfaces in order to construct a family of highly localized approximate solutions to the linear equation, which we call wave-packets. These are spatially concentrated in thin curved rectangles, which we call slabs. We then produce approximate solutions as square summable superpositions of wave packets. For this we need to establish orthogonality of distinct wave packets, which depends on the geometric information we have established for both the characteristic light cones, as well as for the plane wave hypersurfaces.

Section 9 contains a bound on the number of distinct slabs which pass through two given points in the spacetime. This bound is at the heart of the dispersive estimates contained in Section 10, which complete the circle of estimates behind the proof of Theorem 1.2. Finally, the Appendix contains the proof of the two dimensional stability estimate, which turns out to be considerably more delicate than its higher dimensional counterpart.

1.4. **Notations.** In this paper, we use the notation $X \lesssim Y$ to mean that $X \leq CY$, with a constant C which depends only on the dimension n, and on global pointwise bounds for finitely many derivatives of g^{ij}, g_{ij}

and q^{ij}. Similarly, the notation $X \ll Y$ means $X \leq C^{-1} Y$, for a sufficiently large constant C as above.

We use four small parameters

$$\epsilon_3 \leq \epsilon_2 \leq \epsilon_1 \leq \epsilon_0 \ll 1\,.$$

In order for all our estimates to fit together, we will actually need the stronger condition

$$\epsilon_3 \ll \epsilon_2 \ll \epsilon_1 \ll \epsilon_0\,. \tag{1.7}$$

Without any restriction in generality we assume that $\frac{n+1}{2} < s < \frac{n}{2}+1$ for $n \geq 3$, respectively $\frac{7}{4} < s < 2$ for $n = 2$. Denote $\delta_0 = s - \frac{n+1}{2}$ for $n \geq 3$, respectively $\delta_0 = s - \frac{7}{4}$ for $n = 2$, and let δ denote a number with $0 < \delta < \delta_0$.

We denote by ξ the space Fourier variable, and let

$$\langle \xi \rangle = (1 + |\xi|^2)^{\frac{1}{2}}\,.$$

Denote by $\langle D_x \rangle$ the corresponding Bessel potential multiplier. We introduce a Littlewood-Paley decomposition in the spatial frequency ξ,

$$1 = S_0 + \sum_{\lambda \text{ dyadic}} S_\lambda\,,$$

where the spherically symmetric symbols of S_0 and S_λ are supported respectively in the sets $\{\,|\xi| \leq 1\,\}$ and $\{\,|\xi| \in [\lambda/2, 2\lambda]\,\}$. We set

$$S_{<\lambda} = \sum_{8\mu < \lambda} S_\mu\,.$$

We let du denote the full space time gradient, and $d_x u$ the space gradient of u, so that

$$du = (\partial_0 u, \ldots, \partial_n u)\,, \qquad d_x u = (\partial_1 u, \ldots, \partial_n u)\,.$$

Finally, let

$$\Box_{\mathrm{g}(u)} v = \mathrm{g}^{ij}(u)\,\partial_i \partial_j v\,.$$

We may then symbolically write

$$\Box_{\mathrm{g}(u)} v = -\partial_0^2 v + \mathrm{g}(u)\, d_x dv\,.$$

2. Uniqueness and stability

In this section we reduce our main theorem to the case of smooth initial data. Precisely, we show that Theorem 1.2 is a consequence of the following existence result for smooth initial data.

Proposition 2.1. *For each $R > 0$ there exist $T, M, C > 0$ such that, for each smooth initial data (u_0, u_1) which satisfies*

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

there exists a smooth solution u to (1.1) *and* (1.2) *on $[-T, T] \times \mathbb{R}^n$, which furthermore satisfies the conditions* (WP3) *and* (WP4).

The uniqueness of such a smooth solution is well known.

2.1. **Commutators and energy estimates.** We begin with a slight generalization of Lemma 1.4. The purpose of this is twofold, both to make this article self contained, and to have a setup which is better suited to our purposes. In the process we also record certain commutator estimates which are independently used later on. We consider spherically symmetric elliptic symbols $a(\langle \xi \rangle)$, where the function $a : [0, \infty) \to [1, \infty)$ satisfies

$$r_0 \leq \frac{x\, a'(x)}{a(x)} \leq r_1, \qquad a(1) = 1 \,, \tag{2.1}$$

for some positive r_0, r_1. This implies that

$$\langle \xi \rangle^{r_0} \leq a(\langle \xi \rangle) \leq \langle \xi \rangle^{r_1} \,,$$

and also that a is slowly varying on a dyadic scale. Thus,

$$a(\langle \xi \rangle) \approx \sum_{\lambda \text{ dyadic}} a(\lambda)\, S_\lambda(\xi) \,.$$

Then the following result holds:

Lemma 2.2. *Let a be as above, and $A = a(\langle D_x \rangle)$. Let u be a smooth solution to* (1.1) *and* (1.2) *on $[0, T] \times \mathbb{R}^n$. Set $m = \sup_{t,x} |u(t, x)|$. Then the following estimate holds:*

$$\|dAu(t)\|_{L^2_x} \lesssim \|dAu(0)\|_{L^2_x}\, e^{c(m) \int_0^t \|du(h)\|_\infty dh} \,, \quad t \in [0, T] \,. \tag{2.2}$$

This yields Lemma 1.4 in the special case of $a(\langle \xi \rangle) = \langle \xi \rangle^{s-1}$. On the other hand, it also allows for the use of weights which are almost but not quite polynomial.

Proof. For the linear equation

$$\Box_{\mathrm{g}} v = f \,, \tag{2.3}$$

we have the associated energy functional

$$E(v(t)) = \frac{1}{2} \int \Big(-\mathrm{g}^{00} |\partial_0 v|^2 + \sum_{i,j=1}^n \mathrm{g}^{ij}\, \partial_i v\, \partial_j v \Big)\, dx \,.$$

Then a standard computation leads to

$$\frac{d}{dt}E(v(t)) \lesssim \int \big(|f|\,|\partial_0 v| + |d\mathrm{g}|\,|dv|^2 \big)\, dx\,,$$

and hence to

$$\frac{d}{dt}E(v(t))^{\frac12} \lesssim \|f(t)\|_{L^2_x} + \|d\mathrm{g}(t)\|_{L^\infty_x} E(v(t))^{\frac12}\,. \tag{2.4}$$

Return now to (1.1) and set $v = Au$. Then v solves (2.3) with

$$f = (\,\mathrm{g} - A\mathrm{g}A^{-1}\,)\, d_x dv + A\big(q(u)(du)^2\big)\,.$$

We claim that the two terms in f satisfy the estimate

$$\|(\,\mathrm{g} - A\mathrm{g}A^{-1}\,)\, d_x dv\|_{L^2_x} + \|A\big(q(u)(du)^2\big)\|_{L^2_x} \lesssim \|du\|_{L^\infty_x}\|dAu\|_{L^2_x}\,, \tag{2.5}$$

where the constant may depend on m. Given this, we can apply (2.4) to obtain

$$\frac{d}{dt}E(v(t))^{\frac12} \lesssim c(m)\,\|du\|_{L^\infty_x} E(v(t))^{\frac12}\,,$$

which by Gronwall's inequality implies (2.2).

It remains to prove (2.5). This is a consequence of the next lemma:

Lemma 2.3. *Suppose that a satisfies* (2.1). *Then the following estimates hold:*

$$\|A\big(q(u)(du)^2\big)\|_{L^2_x} \lesssim c(m)\,\|du\|_{L^\infty}\|A\,du\|_{L^2_x} \tag{2.6}$$

$$\|A\,d_x(\mathrm{g}(u))\|_{L^2_x} \lesssim c(m)\,\|A\,d_x u\|_{L^2_x} \tag{2.7}$$

$$\|A(fg)\|_{L^2_x} \lesssim \|f\|_{L^\infty_x}\|Ag\|_{L^2_x} + \|g\|_{L^\infty_x}\|Af\|_{L^2_x} \tag{2.8}$$

$$\|A(f d_x g)\|_{L^2_x} \lesssim \|f\|_{L^\infty_x}\|A\,d_x g\|_{L^2_x} + \|g\|_{L^\infty_x}\|A\,d_x f\|_{L^2_x} \tag{2.9}$$

$$\|(\mathrm{g}A - A\mathrm{g})d_x w\|_{L^2_x} \lesssim \|d_x\mathrm{g}\|_{L^\infty_x}\|Aw\|_{L^2_x} + \|A\,d_x\mathrm{g}\|_{L^2_x}\|w\|_{L^\infty_x} \tag{2.10}$$

The proof of Lemma 2.3 uses paraproduct type arguments. Estimate (2.6) is of Moser type. Its proof involves writing the telescoping series

$$q(S_0u)(dS_0u)^2 + \sum_{\lambda\ \mathrm{dyadic}} q(S_{<\lambda}u)(dS_{<\lambda}u)^2 - q(S_{<\lambda/2}u)(dS_{<\lambda/2}u)^2$$

as a combination of three terms, each of which takes the form of an operator of type $S^0_{1,1}$ acting on du, where any given seminorm of the symbol is bounded by $c(m)\|du\|_{L^\infty_x}$, with $c(m)$ an appropriate power of m.[2] The result is thus reduced to showing that, if P is a pseudodifferential operator of type $S^0_{1,1}$, then

$$\|APu\|_{L^2_x} \lesssim \|Au\|_{L^2_x},$$

which for the case $A = \langle D_x \rangle^s$ with $s > 0$ is due to Stein [19], and for the case of A as above is a simple modification.

Estimates (2.7) through (2.9) are similarly reduced. To establish (2.10), we first write

$$(\mathrm{g}A - A\mathrm{g})\, d_x w = -(d_x \mathrm{g})Aw + A(d_x \mathrm{g})w + d_x(\mathrm{g}A - A\mathrm{g})w\,.$$

The first two terms are treated as above. The bound on the last term is a simple variation on the commutator estimate of Kato-Ponce [8], where the result is established for the case $A = \langle D_x \rangle^s$. For further details, we refer to Chapter 3 of Taylor [24]. $\square$

2.2. **Stability estimates.** The next step in the proof is to obtain stability estimates for lower Sobolev norms. As an immediate consequence of these we obtain the uniqueness result. Later on we also use them in order to show the strong continuous dependence on the initial data.[3]

Lemma 2.4. *Suppose that u is a solution to* (1.1) *and* (1.2) *which satisfies the conditions* (WP3) *and* (WP4). *Let v be another solution to the equation* (1.1) *with initial data $(v_0, v_1) \in H^s \times H^{s-1}$, such that $dv \in L^\infty_t H^{s-1}_x \cap L^2_t L^\infty_x$. Then, for $n = 2$,*

$$\|d(u-v)\|_{L^\infty_t H^{-1/4}_x} \le C_v \,\|(u_0 - v_0, u_1 - v_1)\|_{H^{3/4} \times H^{-1/4}}\,, \tag{2.11}$$

and for $n = 3, 4, 5$,

$$\|d(u-v)\|_{L^\infty_t L^2_x} \le C_v \,\|(u_0 - v_0, u_1 - v_1)\|_{H^1 \times L^2}\,, \tag{2.12}$$

where C_v depends on u, and on $\|dv\|_{L^\infty_t H^{s-1}_x \cap L^2_t L^\infty_x}$.

We note that for the proof it does not suffice to only use the Sobolev regularity of u and v; we also need the dispersive estimates in Proposition 2.1. On the bright side, it suffices to know these only for u, and therefore to have a less restrictive condition for v.

[2] This step requires that the coefficient $q^{00}(u)$ of $(\partial_0 u)^2$ be constant, since for one term it involves transferring a factor of λ from $S_{<\lambda} du$ to $S_\lambda u$. We can avoid this assumption by weakening Lemmas 1.4 and 2.2 to require $L^2 L^\infty$ bounds on du instead of $L^1 L^\infty$ bounds, which suffices for our application.

[3] For the case $n = 2$, which we handle in the Appendix, we strengthen condition (WP4) to include additional estimates which play a crucial role in the $n = 2$ stability of solutions. This has no effect on the rest of the paper.

Proof. We prove the result here for the case $n \geq 3$. The case $n = 2$ is considerably more delicate and is discussed in the Appendix. The first step is to note that the function $w = u - v$ satisfies the equation

$$\Box_{\mathrm{g}(u)} w = a_0\, dw + a_1\, w\,, \tag{2.13}$$

where the functions a_0 and a_1 are of the form

$$a_0 = q(v)\, d(u,v)\,, \quad a_1 = a(u,v)\, d_x dv + b(u,v)\,(du)^2\,,$$

with q, a, b smooth and bounded functions of u, v. By interpolation,

$$dv \in L_t^\infty H_x^{s-1} \cap L_t^2 L_x^\infty \quad \longrightarrow \quad d_x dv \in L_t^{\frac{2(n-1)}{n-3}} L_x^{n-1+\epsilon}\,,$$

for some $\epsilon > 0$. This yields

$$a_0 \in L_t^2 L_x^\infty\,, \quad a_1 \in L_t^{\frac{2(n-1)}{n-3}} L_x^{n-1+\epsilon}\,.$$

On the other hand, the Strichartz estimates implied by (WP4) show that, if $\Box_{\mathbf{g}(u)} w = 0$, then

$$\|w\|_{L_t^{n-1} L_x^{\frac{2(n-1)}{n-3+\epsilon}}} \lesssim \|(w_0, w_1)\|_{H^1 \times L^2}\,,$$

for all $\epsilon > 0$, and consequently

$$\|a_0\, dw + a_1\, w\|_{L_t^2 L_x^2} \lesssim \|(w_0, w_1)\|_{H^1 \times L^2}\,.$$

By the Duhamel principle and a contraction argument, this is sufficient to show that, for T small, solutions to (2.13) satisfy

$$\|dw\|_{L_t^\infty L_x^2} \lesssim \|(w_0, w_1)\|_{H^1 \times L^2}\,.$$

The result may then be easily extended to any interval on which the conditions of the Lemma hold. $\Box$

2.3. **Existence, uniqueness and stability for rough data.** Again we argue in the case $n \geq 3$; obvious changes are required for $n = 2$. Consider arbitrary initial data $(u_0, u_1) \in H^s \times H^{s-1}$ such that

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

Let (u_0^k, u_1^k) be a sequence of smooth data converging to (u_0, u_1), which also satisfy the same bound. Then the conclusion of Proposition 2.1 applies uniformly to the corresponding solutions u^k.

In particular, it follows that the sequence du^k is bounded in the space $C([-T,T]; H^{s-1})$. We can use compactness to improve upon this. More precisely, since (u_0^k, u_1^k) converges to (u_0, u_1) in $H^s \times H^{s-1}$, it follows that there is a multiplier A satisfying (2.1), such that

$$\lim_{\xi \to \infty} \frac{a(\xi)}{|\xi|^{s-1}} = \infty\,,$$

while the sequence $Adu^k(0)$ is still bounded. By Theorem 2.2, it follows that Adu^k is bounded in $C([-T,T];L^2)$. On the other hand, by Lemma 2.4 the sequence du^k is Cauchy in $L^\infty_t L^2_x$. Combining these two properties, it follows that du^k is Cauchy in $C([-T,T];H^{s-1})$, and we let u denote its limit.

As a consequence of (2.5) applied to $A = \langle D_x \rangle^{s-1}$, the right hand sides $q(u^k)(du^k)^2$ of the equations for u^k are uniformly bounded in the space $L^2([-T,T];H^{s-1})$. Then (WP4) combined with Duhamel's formula show that du^k is uniformly bounded in $L^2([-T,T];C^\delta)$. Together with the above this implies that du^k converges to du in $L^2([-T,T];L^\infty)$.

The above information is more that sufficient to allow passage to the limit in the equation (1.1) and show that u is a solution in the sense of distributions, yielding the existence part of (WP1). The conditions (WP3) and (WP4) hold for u since they hold uniformly for u^k. The uniqueness part of (WP1) then follows by Lemma 2.4. Finally, if (u_0^k, u_1^k) is any sequence of initial data converging to (u_0,u_1), it follows as above that u^k converges to u in both the Sobolev and $L^2_t L^\infty_x$ norms.

3. Reduction to existence for small, smooth, compactly supported data

In this section we take advantage of scaling and the finite speed of propagation to further simplify the problem. Denote by c the largest speed of propagation corresponding to all possible values of $\mathrm{g} = \mathrm{g}(u)$. The intermediate result which will be established in subsequent sections is the following:

Proposition 3.1. *Suppose* (1.7) *holds. Assume that the data* (u_0,u_1) *is smooth, supported in* $B(0,c+2)$*, and satisfies*

$$\|u_0\|_{H^s} + \|u_1\|_{H^{s-1}} \le \epsilon_3 \,.$$

Then the equations (1.1) *and* (1.2) *admit a smooth solution* u *defined on* $\mathbb{R}^n \times [-1,1]$*, and the following properties hold:*

(i) (energy estimate)

$$\|du\|_{L^\infty_t H^{s-1}_x} \le \epsilon_2 \,, \tag{3.1}$$

(ii) (dispersive estimate for u)

$$\begin{aligned} &\|du\|_{L^4_t C^\delta_x} \le \epsilon_2 \,, \qquad n = 2, \\ &\|du\|_{L^2_t C^\delta_x} \le \epsilon_2 \,, \qquad n = 3,4,5, \end{aligned} \tag{3.2}$$

(iii) (dispersive estimates for the linear equation) *For* $1 \leq r \leq s+1$ *the equation* (1.3) *with* $\mathrm{g} = \mathrm{g}(u)$ *is well-posed in* $H^r \times H^{r-1}$, *and the following estimate holds:*

$$(3.3)\qquad \begin{aligned} &\|\langle D_x\rangle^\rho v\|_{L_t^4 L_x^\infty} \lesssim \|(v_0, v_1)\|_{H^r \times H^{r-1}}, \quad \rho < r - \tfrac{3}{4}, \qquad n = 2, \\ &\|\langle D_x\rangle^\rho v\|_{L_t^2 L_x^\infty} \lesssim \|(v_0, v_1)\|_{H^r \times H^{r-1}}, \quad \rho < r - \tfrac{n-1}{2}, \quad n = 3, 4, 5, \end{aligned}$$

and the same estimates hold with $\langle D_x\rangle^\rho$ *replaced by* $\langle D_x\rangle^{\rho-1} d$.

In the remainder of this section we show that Proposition 3.1 implies Proposition 2.1.

3.1. **Scaling.** Consider a smooth initial data set (u_0, u_1) which satisfies

$$\|u_0\|_{H^s} + \|u_1\|_{H^{s-1}} \leq R.$$

For this we seek a smooth solution u to (1.1), (1.2) in a time interval $[-T, T]$. We rescale the problem to time scale 1 by setting

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

Then we ask that $\tilde{u}$ be a solution to the equation (1.1), and note that its initial data satisfies

$$\|\tilde{u}(0)\|_{\dot{H}^s} + \|\tilde{u}_t(0)\|_{\dot{H}^{s-1}} \leq RT^{s-\frac{n}{2}},$$

and

$$\|\tilde{u}(0)\|_{H^1} + \|\tilde{u}_t(0)\|_{L^2} \leq RT^{-\frac{n}{2}}.$$

Let ϵ_3 be as in Proposition 3.1, and choose T so that

$$RT^{s-\frac{n}{2}} \ll \epsilon_3.$$

By doing this we have reduced the problem to the case where $T = 1$, and where

$$\|u_0\|_{\dot{H}^s} + \|u_1\|_{\dot{H}^{s-1}} \ll \epsilon_3,$$

while

$$\|u_0\|_{L^\infty} \lesssim R, \qquad \|u_0\|_{H^1} + \|u_1\|_{L^2} \leq M,$$

for some large M.

3.2. **Localization.** In the previous step there is seemingly a loss, because we had to replace homogeneous spaces by inhomogeneous ones. This is remedied here by taking advantage of the finite speed of propagation. Since c is the largest possible speed of propagation, the solution in a unit cylinder $B(y, 1) \times [-1, 1]$ is uniquely determined by the initial data in the ball $B(y, 1 + c)$. Hence it is natural to truncate the initial data in a slightly larger region. Some care is required, however, since we need the truncated data to be small, which means we only want to use the control of the homogeneous norms, which might not see constants, or, more general, polynomials. In our case we are assuming that $s < \frac{n}{2} + 1$, therefore it suffices to account for the constants in u_0.

Let χ be a smooth function supported in $B(0, c+2)$, and which equals 1 in $B(0, c+1)$. Given $y \in \mathbb{R}^n$ we define the localized initial data near y,

$$u_0^y(x) = \chi(x-y)\,(u_0 - u_0(y))\,, \qquad u_1^y = \chi(x-y)\,u_1\,.$$

Since $s < \frac{n}{2} + 1$, it is easy to see that

$$\|(u_0^y, u_1^y)\|_{H^s \times H^{s-1}} \lesssim \|(u_0, u_1)\|_{\dot H^s \times \dot H^{s-1}}\,,$$

so that

$$\|(u_0^y, u_1^y)\|_{H^s \times H^{s-1}} \leq \epsilon_3\,.$$

Hence, by Proposition 3.1 we have a smooth solution u^y on $[-1,1] \times \mathbb{R}^n$ to the equation

$$\begin{cases} \Box_{\mathrm{g}(u^y + u_0(y))} = q^{ij}(u^y + u_0(y))\,\partial_i u^y\,\partial_j u^y\,, \\ u^y(0) = u_0^y\,, \quad u_t^y(0) = u_1^y\,. \end{cases}$$

Then the function $u^y + u_0(y)$ solves (1.1), and its initial data coincides with (u_0, u_1) in $B(y, c{+}1)$. We now consider the restrictions, for $y \in \mathbb{R}^n$,

$$(u^y + u_0(y))|_{K^y}\,, \quad K^y = \{(t,x) \,:\, ct + |x-y| \leq c+1\,,\ |t| < 1\}\,.$$

The restrictions solve (1.1) and (1.2) on K^y, therefore, by finite speed of propagation, any two must coincide on their common domain. Hence we obtain a smooth solution u in $[-1,1] \times \mathbb{R}^n$ by setting

$$u(t,x) = u^y(t,x) + u_0(y), \quad (t,x) \in K^y\,.$$

It remains to show that u satisfies (WP3) and (WP4). We consider the cartesian grid $n^{-\frac{1}{2}}\mathbb{Z}^n$ in $\mathbb{R}^n$, and a corresponding smooth partition of unity

$$1 = \sum_{y \in n^{-\frac{1}{2}}\mathbb{Z}^n} \psi(x-y)\,,$$

such that the function ψ is supported in the unit ball.

For (WP3) we first obtain the corresponding estimates for u^y. Applying the energy estimates in Lemma 1.4 yields

$$\|du^y\|_{L_t^\infty H_x^{s-1}} \lesssim \|(u_0^y, u_1^y)\|_{H^s \times H^{s-1}}\,.$$

On the other hand, (3.3) combined with Duhamel's formula yields

$$\|du^y\|_{L_t^2 L_x^\infty} \lesssim \|(u_0^y, u_1^y)\|_{H^s \times H^{s-1}} + \|q^{ij}(u^y + u_0(y))\partial_i u^y \partial_j u^y\|_{L_t^1 H_x^{s-1}}\,.$$

By (2.5) with $A = \langle D_x \rangle^{s-1}$ we can estimate the last term to conclude that

$$\begin{aligned} \|du^y\|_{L^2_t L^\infty_x} &\lesssim \|(u^y_0, u^y_1)\|_{H^s \times H^{s-1}} + \|du^y\|_{L^\infty_t H^{s-1}_x} \|du^y\|_{L^2_t L^\infty_x} \\ &\lesssim \|(u^y_0, u^y_1)\|_{H^s \times H^{s-1}} + \epsilon_2 \|du^y\|_{L^2_t L^\infty_x} . \end{aligned}$$

Since $\epsilon_2 \ll 1$, this implies

$$\|du^y\|_{L^2_t L^\infty_x} \lesssim \|(u^y_0, u^y_1)\|_{H^s \times H^{s-1}} .$$

It remains to sum up the estimates for u^y in order to obtain the estimates for u. We have

$$u(x,t) = \sum_{y \in n^{-\frac{1}{2}} \mathbb{Z}^n} \psi(x-y)(u^y(x,t) + u_0(y)) ,$$

therefore

$$\begin{aligned} \|du\|^2_{L^2_t L^\infty_x \cap L^\infty_t H^{s-1}_x} &\lesssim \sum_{y \in n^{-\frac{1}{2}} \mathbb{Z}^n} \|d(\psi(x-y)(u^y + u_0(y))\|^2_{L^2_t L^\infty_x \cap L^\infty_t H^{s-1}_x} \\ &\lesssim \sum_{y \in n^{-\frac{1}{2}} \mathbb{Z}^n} \|\chi(x-y)(u_0, u_1)\|^2_{H^s \times H^{s-1}} + |u_0(y)|^2 \\ &\lesssim \|(u_0, u_1)\|_{H^s \times H^{s-1}} . \end{aligned}$$

For (WP4) we consider the solutions v^y for the localized linear equations

$$\begin{cases} \Box_{g(u^y + u_0(y))} v^y = 0 , \\ v^y(0) = \chi(x-y) v_0 , \quad v^y_t(0) = \chi(x-y) v_1 . \end{cases}$$

We again use the finite speed of propagation to conclude that $v_y = v$ in K_y. Then we can represent v as

$$v(x,t) = \sum_{y \in n^{-\frac{1}{2}} \mathbb{Z}^n} \psi(x-y) v^y(x,t) ,$$

and use (3.3) to estimate

$$\begin{aligned} \|\langle D_x \rangle^\rho dv\|^2_{L^2_t L^\infty_x} &\lesssim \sum_{y \in n^{-\frac{1}{2}} \mathbb{Z}^n} \|\psi(x-y) v^y(x,t)\|^2_{L^2_t L^\infty_x} \\ &\lesssim \sum_{y \in n^{-\frac{1}{2}} \mathbb{Z}^n} \|\chi(x-y)(v_0, v_1)\|^2_{H^r \times H^{r-1}} \\ &\lesssim \|(v_0, v_1)\|^2_{H^r \times H^{r-1}} . \end{aligned}$$

4. A recursive argument

We will establish Proposition 3.1 via a continuity argument. More precisely, we consider a one parameter family of smooth initial data (hu_0, hu_1) with $h \in [0,1]$. Since the data (u_0, u_1) is smooth, for small h the equation has a smooth solution u^h. We seek to extend the range of h for which a solution exists to the value $h = 1$. We do this by establishing uniform bounds on the u^h in the norm of $L^2_t C^\delta_x$; this in turn implies uniform bounds on u^h in the Sobolev norm.

Our proof of the bounds on the u^h in $L^2_t C^\delta_x$ relies on a parametrix construction, which in turn depends on the regularity of certain null-foliations of space time. Rather than attempt to obtain the regularity of these foliations directly, we build their regularity into the continuity argument. This works since we need only assume that the appropriate norm $G(u)$ of the foliations is small compared to 1 in order to deduce that it is in fact bounded by a multiple of the norm of the initial data. We set aside for the moment the definition of $G(u)$ and outline the general recursive argument.

Let η^{ij} be the standard Minkowski metric,

$$\eta^{00} = -1\,, \qquad \eta^{jj} = 1\,, \quad 1 \le j \le n\,, \qquad \eta^{ij} = 0 \quad \text{if} \quad i \ne j\,.$$

After making a linear change of coordinates which preserves dt we may assume that $\mathrm{g}^{ij}(0) = \eta^{ij}$.

For technical reasons it is convenient to replace the original metric function g by a truncated one. Let χ be a smooth cutoff function supported in the region $B(0, 3+2c) \times [-\frac{3}{2}, \frac{3}{2}]$, which equals 1 in the region $B(0, 2+2c) \times [-1, 1]$. Set

$$\mathbf{g}(t,x,u) = \chi(t,x)\big(\mathrm{g}(u) - \mathrm{g}(0)\big) + \mathrm{g}(0)\,, \quad \mathbf{q}(t,x,u) = \chi(t,x)\, q(u)\,,$$

and introduce the truncated equation

$$\Box_{\mathbf{g}(t,x,u)} u = \mathbf{q}^{ij}(t,x,u)\partial_i u\, \partial_j u\,. \tag{4.1}$$

Because of the finite speed of propagation, any solution to (4.1) for $t \in [-2,2]$ with initial data supported in $B(0, 2+c)$ is also a solution to (1.1) for $t \in [-1, 1]$.

We denote by $\mathcal{H}$ the family of smooth solutions u to the equation (4.1) for $t \in [-2,2]$, with initial data (u_0, u_1) supported in $B(0, 2+c)$, and for which

$$\|u_0\|_{H^s} + \|u_1\|_{H^{s-1}} \le \epsilon_3\,, \tag{4.2}$$

$$\|du\|_{L^\infty_t H^{s-1}_x} + \|du\|_{L^2_t C^\delta_x} \le 2\epsilon_2\,. \tag{4.3}$$

On $\mathcal{H}$ we use the induced C^∞ topology. Then our bootstrap argument can be stated as follows:

Proposition 4.1. *Assume that* (1.7) *holds. Then there is a continuous functional $G : \mathcal{H} \to \mathbb{R}^+$, satisfying $G(0) = 0$, so that for each $u \in \mathcal{H}$ satisfying $G(u) \le 2\epsilon_1$ the following hold:*

(i) *The function u satisfies $G(u) \le \epsilon_1$.*

(ii) *The following estimate holds,*

$$\|du\|_{L^\infty_t H^{s-1}_x} + \|du\|_{L^2_t C^\delta_x} \le \epsilon_2 \,. \tag{4.4}$$

(iii) *For $1 \le r \le s+1$, the equation* (1.3) *with* $\mathrm{g} = \mathbf{g}(t,x,u)$ *is well-posed in $H^r \times H^{r-1}$, and the Strichartz estimates* (3.3) *hold.*

Proposition 4.1 will follow as a result of Proposition 5.2 and Proposition 7.1. We provide the definition of $G(u)$ shortly; here we show that Proposition 4.1 implies Proposition 3.1. Thus, consider initial data (u_0, u_1) which satisfies

$$\|u_0\|_{H^s} + \|u_1\|_{H^{s-1}} \le \epsilon_3 \,.$$

We denote by A the subset of those $h \in [0,1]$ such that the equation (4.1) admits a smooth solution u^h having initial data

$$u^h(0) = hu_0 \,, \quad u^h_t(0) = hu_1 \,,$$

and such that $G(u^h) \le \epsilon_1$ and (4.4) holds. We trivially have $0 \in A$, since $u^0 = 0$. Proposition 3.1 would follow if we knew that $1 \in A$, and so it suffices to show that A is both open and closed in $[0,1]$.

A is open. Let $k \in A$. Since u^k is smooth, a perturbation argument shows that for h close to k the equation (4.1) has a smooth solution u^h, which depends continuously on h. By the continuity of G, it follows that for h close to k we have $G(u^h) \le 2\epsilon_1$ and also (4.3). Then by Proposition 4.1 we obtain $G(u^h) \le \epsilon_1$ and (4.4), showing that $h \in A$.

A is closed. Let $h_i \in A$, $h_i \to h$. Then (4.4) implies that the sequence du^{h_i} is bounded in $L^2_t C^\delta_x$. Lemma 1.4 then shows that the sequence u^{h_i} is in fact bounded in all Sobolev spaces. We thus can obtain a smooth solution u^h as the limit of some subsequence. The continuity of G then shows that $G(u) \le \epsilon_1$, and similarly (4.4) must also hold for u^h.

4.1. **The Hamilton flow and the G functional.** Let $u \in \mathcal{H}$, and consider the corresponding metric $\mathbf{g} = \mathbf{g}(t,x,u)$, which equals the Minkowski metric for $t \in [-2, -\frac{3}{2}]$. For each $\theta \in S^{n-1}$ we consider a foliation of the slice $t = -2$ by taking level sets of the function $r_\theta(-2,x) = \theta \cdot x + 2$. Then $\theta \cdot dx - dt$ is a null covector field over $t = -2$ which is conormal to the level sets of $r_\theta(-2)$. We let Λ_θ be the flowout of this section under the Hamitonian flow of $\mathbf{g}$.

A crucial step in the proof of the Strichartz estimates is to establish that, for each θ, the null Lagrangian manifold Λ_θ is the graph of a null covector field given by dr_θ, where r_θ is a smooth extension of $\theta \cdot x - t$, and that the level sets of r_θ are small perturbations of the level sets of the function $\theta\cdot x - t$ in a certain norm captured by $G(u)$. In establishing Proposition 4.1 we will actually establish that $u \in \mathcal{H}$ implies Λ_θ is the graph of an appropriate null covector field dr_θ, so we only define $G(u)$ in this situation.

Thus, assume that Λ_θ and r_θ are as above, and let $\Sigma_{\theta,r}$ for $r \in \mathbb{R}$ denote the level sets of r_θ. The characteristic hypersurface $\Sigma_{\theta,r}$ is thus the flowout of the set $\theta \cdot x = r - 2$ along the null geodesic flow in the direction θ at $t = -2$.

We introduce an orthonormal sets of coordinates on $\mathbb{R}^n$ by setting $x_\theta = \theta \cdot x$, and letting x'_θ be given orthonormal coordinates on the hyperplane perpendicular to θ, which then define coordinates on $\mathbb{R}^n$ by projection along θ. Then (t, x'_θ) induce coordinates on $\Sigma_{\theta,r}$, and $\Sigma_{\theta,r}$ is given by

$$\Sigma_{\theta,r} = \{\, (t,x) \,:\, x_\theta - \phi_{\theta,r} = 0\}$$

for a smooth function $\phi_{\theta,r}(t,x'_\theta)$. We now introduce two norms for functions defined on $[-2,2]\times\mathbb{R}^n$,

$$\begin{aligned}
|\!|\!| u |\!|\!|_{s,\infty} &= \sup_{-2\le t\le 2}\ \sup_{0\le j\le 1} \|\partial_t^j u(t,\cdot)\|_{H^{s-j}(\mathbb{R}^n)}\,,\\
|\!|\!| u |\!|\!|_{s,2} &= \left(\int_{-2}^{2} \sup_{0\le j\le 1} \|\partial_t^j u(t,\cdot)\|^2_{H^{s-j}(\mathbb{R}^n)}\,dt\right)^{\frac12}.
\end{aligned}$$

The same notation applies for functions in $[-2,2]\times\mathbb{R}^{n-1}$. We denote

$$|\!|\!| \, f \, |\!|\!|_{s,2,\Sigma_{\theta,r}} = |\!|\!| \, f \,|_{\Sigma_{\theta,r}} |\!|\!|_{s,2}\,,$$

where the right hand side is the norm of the restriction of f to $\Sigma_{\theta,r}$, taken over the (t,x'_θ) variables used to parametrise $\Sigma_{\theta,r}$. Similarly,

$$\|f\|_{H^s(\Sigma^t_{\theta,r})}$$

denotes the $H^s(\mathbb{R}^{n-1})$ norm of f restricted to the time t slice of $\Sigma_{\theta,r}$ using the x'_θ coordinates on $\Sigma^t_{\theta,r}$.

We now set

$$G(u) = \sup_{\theta,r} |\!|\!| d\phi_{\theta,r} - dt |\!|\!|_{s,2,\Sigma_{\theta,r}}\,. \tag{4.5}$$

Note that G is nonlinear, as $\phi_{\theta,r}$ depends in a nonlinear way on u. Since all functions in $\mathcal{H}$ are supported in a fixed compact set, it follows that we can restrict ourselves to a compact set of values for r. Then the continuity of G as a function of u with respect to the C^∞ topology easily follows.

5. Regularity of null surfaces

The goal of this section is to establish the following. The functional $G(u)$ is defined in (4.5).

Proposition 5.1. *Let $u \in \mathcal{H}$ so that $G(u) \le 2\epsilon_1$. Let $\mathbf{g}_\lambda$ denote the localization, in the x-variables, of $\mathbf{g}$ to frequencies less than or comparable to λ. Then*

$$\|\mathbf{g}^{ij} - \eta^{ij}\|_{s,2,\Sigma_{\theta,r}} + \|(\lambda(\mathbf{g}^{ij} - \mathbf{g}^{ij}_\lambda), d\mathbf{g}^{ij}_\lambda, \lambda^{-1}\partial_x d\mathbf{g}^{ij}_\lambda)\|_{s-1,2,\Sigma_{\theta,r}} \lesssim \epsilon_2 \,.$$

Proposition 5.2. *Let $u \in \mathcal{H}$ so that $G(u) \le 2\epsilon_1$. Then $G(u) \lesssim \epsilon_2$. Furthermore, for each t it holds that*

$$\|d\phi_{\theta,r}(t,\cdot) - dt\|_{C^{1,\delta}_{x'}(\mathbb{R}^{n-1})} \lesssim \epsilon_2 + \sup_{i,j} \|d\mathbf{g}^{ij}(t,\cdot)\|_{C^\delta_x(\mathbb{R}^n)} \,. \tag{5.1}$$

Proposition 5.1 is essentially a variation on the theme of characteristic energy estimates for the variable coefficient wave equation. Together with the estimate (5.1) it is used later in the parametrix construction for the linear equation.

The first part of Proposition 5.2 is a much deeper result which lies at the heart of our paper. It gives the recursive estimate, namely part (i) of Proposition 4.1.

5.1. **Setup.** Since the proof of Propositions 5.1 and 5.2 is lengthy, it is useful to summarize at this stage the information we have about the function u and the metric $\mathbf{g}$.

In this section, we deal more generally with equations of the form

$$\mathbf{g}^{ij}(t,x,u)\,\partial_i\,\partial_j u = Q(t,x,u;du)\,, \tag{5.2}$$

where Q takes the form

$$Q(t,x,u;du) = \sum_{ij} q^{ij}(t,x,u)\partial_i u \partial_j u + \sum_j q^j(t,x,u)\partial_j u + q_0(t,x,u)u\,,$$

and $\mathbf{g}^{ij}$, q^{ij}, q^i, and q_0 are smooth functions of the variables t, x, u.

By doing so, we note that we may also write such an equation as

$$\partial_i\,\mathbf{g}^{ij}(t,x,u)\,\partial_j u = Q(t,x,u;du)\,,$$

for a different Q of the same form, and by combining terms we may assume that $\mathbf{g}^{0j} = 0$ for $j \ne 0$. This means that the coefficients of the

Lorentzian form $\langle \cdot, \cdot \rangle_{\mathbf{g}}$ are given by $\frac{1}{2}\big(\mathbf{g}^{ij} + \mathbf{g}^{ji}\big)$, rather than by $\mathbf{g}^{ij}$. Furthermore, for each k, l, we may also write

$$\mathbf{g}^{ij}(t,x,u)\,\partial_i\,\partial_j\,\mathbf{g}^{kl}(t,x,u) = Q(t,x,u;du)\,, \tag{5.3}$$

with Q of the same form. Recall also that $\mathbf{g}^{ij}(0) = \eta^{ij}$, and that

$$\mathbf{g}^{ij} = \eta^{ij} \quad \text{if} \quad |t| \geq \tfrac{3}{2} \quad \text{or} \quad |x| \geq 3 + 2c\,.$$

The function u belongs to $\mathcal{H}$, therefore it satisfies

$$\|du\|_{L^2_t C^\delta_x} + \|u\|_{s,\infty} \lesssim \epsilon_2\,. \tag{5.4}$$

In particular u is pointwise small, $|u| \lesssim \epsilon_2$. Thus $|\mathrm{g}(u) - \eta| \lesssim \epsilon_2$, which in turn yields a similar bound for $\mathbf{g}$,

$$\|d\mathbf{g}^{ij}\|_{L^2_t C^\delta_x} + \|\mathbf{g}^{ij} - \eta^{ij}\|_{s,\infty} \lesssim \epsilon_2\,. \tag{5.5}$$

For the proof of Propositions 5.1 and 5.2 it suffices to consider the case where $\theta = (0, \ldots, 0, 1)$ and $r = 0$. We fix this choice, and suppress θ and r in our notation. Instead of (x_θ, x'_θ) we use (x_n, x'). Then Σ is defined by

$$\Sigma = \{ x_n - \phi(t, x') = 0 \}\,.$$

The hypothesis $G(u) < 2\epsilon_1$ implies that

$$\|d\phi(t,x') - dt\|_{s,2,\Sigma} \leq 2\epsilon_1\,. \tag{5.6}$$

Note that by Sobolev embedding, this implies that

$$\|d\phi(t,x') - dt\|_{L^2_t C^{1,\delta}_{x'}} + \|\partial_t d\phi(t,x')\|_{L^2_t C^\delta_{x'}} \lesssim \epsilon_1\,. \tag{5.7}$$

As a consequence of this it follows that $\phi - t$ is small in C^1.

5.2. **Characteristic energy estimates.** We use a basic fact about Sobolev norms, which is a simple paraproduct result.

Lemma 5.3. *Suppose that $0 \leq r, r' < \frac{n}{2}$ and $r + r' > \frac{n}{2}$. Then*

$$\|fg\|_{H^{r+r'-\frac{n}{2}}(\mathbb{R}^n)} \leq C_{r,r'}\,\|f\|_{H^r(\mathbb{R}^n)}\,\|g\|_{H^{r'}(\mathbb{R}^n)}\,. \tag{5.8}$$

If $-r \leq r' \leq r$ and $r > \frac{n}{2}$ then

$$\|fg\|_{H^{r'}(\mathbb{R}^n)} \leq C_{r,r'}\,\|f\|_{H^r(\mathbb{R}^n)}\,\|g\|_{H^{r'}(\mathbb{R}^n)}\,. \tag{5.9}$$

As a consequence we have the following facts about the triple norm.

Lemma 5.4. *For $r \geq 1$, we have*

$$\begin{aligned} &\sup_{t\in[-2,2]} \|f\|_{H^{r-\frac{1}{2}}(\mathbb{R}^n)} \leq C_r\,|\!|\!| f |\!|\!|_{r,2}\,, \\ &\sup_{t\in[-2,2]} \|f\|_{H^{r-\frac{1}{2}}(\Sigma^t)} \leq C_r\,|\!|\!| f |\!|\!|_{r,2,\Sigma}\,. \end{aligned} \tag{5.10}$$

If $r > (n+1)/2$, then

$$|\!|\!| fg |\!|\!|_{r,2} \leq C_r\,|\!|\!| f |\!|\!|_{r,2}\,|\!|\!| g |\!|\!|_{r,2}\,. \tag{5.11}$$

Similarly, if $r > n/2$, then

$$\|\!|fg|\!\|_{r,2,\Sigma} \le C_r \,\|\!|f|\!\|_{r,2,\Sigma}\,\|\!|g|\!\|_{r,2,\Sigma}\,. \tag{5.12}$$

Proof. The first result follows from the trace theorem:

$$\|f\|_{L^\infty_t H_x^{r-\frac12}} = \|\langle D_x\rangle^{r-1} f\|_{L^\infty_t H_x^{\frac12}} \lesssim \|\langle D_x\rangle^{r-1} f\|_{H^1([-2,2]\times\mathbb{R}^n)} = \|f\|_{r,2}\,.$$

The bound (5.10) follows similarly. To establish (5.11), we use (5.9) and the preceding estimate to bound

$$\begin{aligned}\|\!|fg|\!\|_{r,2} &\le \|fg\|_{L^2_t H^{r-1}_x} + \|d(fg)\|_{L^2_t H^{r-1}_x}\\ &\lesssim \|f\|_{L^\infty_t H_x^{r-\frac12}}\big(\|g\|_{L^2_t H^r_x} + \|dg\|_{L^2_t H^{r-1}_x}\big) + \|df\|_{L^2_t H^{r-1}_x}\|g\|_{L^\infty_t H_x^{r-\frac12}}\\ &\lesssim \|\!|f|\!\|_{r,2}\,\|\!|g|\!\|_{r,2}\,.\end{aligned}$$

The inequality (5.12) follows similarly. $\square$

We now show that the triple norm of u is preserved under the change of coordinates which flattens Σ.

Lemma 5.5. *Let $\tilde w(t,x) = w(t,x',x_n+\phi(t,x'))$. Then*

$$\|\!|\tilde w|\!\|_{s,\infty} \lesssim \|\!|w|\!\|_{s,\infty}\,, \qquad \|d\tilde w\|_{L^2_t L^\infty_x} \lesssim \|dw\|_{L^2_t L^\infty_x}\,.$$

Proof. The second inequality is immediate from the C^1 bounds on ϕ. For the first, recall that $s = m+\sigma$, where $0<\sigma<1$. Since ϕ is C^1, we need to show that, for $|\alpha|\le m$, and with ∂^α involving at most one derivative in t, we have

$$\big(\partial + (\partial\phi)\partial_n\big)^\alpha w \in L^\infty_t H^\sigma_x\,.$$

The product may be expanded as a sum of terms

$$(\partial\phi)^j\,(\partial^{\alpha_1}\partial\phi)\,\cdots\,(\partial^{\alpha_k}\partial\phi)\,\partial^{\alpha_0} w\,,$$

where $\alpha_0+\alpha_1+\cdots+\alpha_k=\alpha$, and $\alpha_0\neq 0$, and each term involves at most one derivative in t. By (5.8) we may bound the $H_x^{s-|\alpha|}$ norm of the product by

$$\|\partial\phi\|^j_{H_x^{s-\frac12}}\|\partial^{\alpha_1}\partial\phi\|_{H_x^{s-\frac12-|\alpha_1|}}\cdots\|\partial^{\alpha_k}\partial\phi\|_{H_x^{s-\frac12-|\alpha_2|}}\|\partial^{\alpha_0}w\|_{H_x^{s-|\alpha_0|}}\,.$$

$\square$

Remark. A similar proof shows that, for $0\le s'\le s$, we have for all t

$$\|\tilde w(t,\cdot)\|_{H_x^{s'}} \lesssim \|w(t,\cdot)\|_{H_x^{s'}}\,. \tag{5.13}$$

We continue with the characteristic energy estimate:

Lemma 5.6. *Assume that w satisfies the linear equation*

$$\partial_i\big(\mathbf{g}^{ij}\,\partial_j w\big) = F\,.$$

Then

$$\|dw\|_{s-1,2,\Sigma} \lesssim \|dw\|_{L^\infty_t H^{s-1}_x} + \|dw\|_{L^2_t L^\infty_x} + \|F\|_{L^2_t H^{s-1}_x}\,.$$

Proof. Let

$$\|dw\|_{L^\infty_t H^{s-1}_x} + \|dw\|_{L^2_t L^\infty_x} + \|F\|_{L^2_t H^{s-1}_x} = \epsilon\,.$$

Under the change of coordinates $x_n \to x_n - \phi(t,x')$, the equation transforms to

$$\sum_{i,j=0}^{n} (\partial_i - (\partial_i\phi)\partial_n)\,\big(\tilde{\mathbf{g}}^{ij}\,(\partial_j - (\partial_j\phi)\partial_n)\tilde{w}\big) = \tilde{F}\,,$$

where $\tilde{}$ denotes the function expressed in the new coordinates. Recall that we have $\mathbf{g}^{00} = -1$, and $\mathbf{g}^{0j} = 0$ for $j \neq 0$, and that ϕ is independent of x_n. For $i \neq 0$ we now define

$$\mathbf{h}^{ij} = \tilde{\mathbf{g}}^{ij} - \delta^{in}(\partial_k\phi)\tilde{\mathbf{g}}^{kj} - \tilde{\mathbf{g}}^{ik}(\partial_k\phi)\delta^{jn} + \delta^{in}\delta^{jn}(\partial_k\phi)(\partial_\ell\phi)\tilde{\mathbf{g}}^{k\ell} - \delta^{in}\delta^{j0}(\partial_0\phi),$$

and set $\mathbf{h}^{00} = -1$ and $\mathbf{h}^{0j} = 0$ for $j \neq 0$. Then the above equation takes the form

$$\sum_{i,j=0}^{n} \partial_i\,\big(\mathbf{h}^{ij}\,\partial_j\tilde{w}\big) = \tilde{F} - (\partial_0^2\phi)\,\partial_n\tilde{w} = G\,.$$

We use the following bounds on $\mathbf{h}^{ij}$.

$$\|\mathbf{h}^{ij}\|_{s,2} + \|\mathbf{h}^{ij}\|_{L^\infty_t H^{s-\frac{1}{2}}_{x'}(\Sigma)} \lesssim 1\,, \tag{5.14}$$

$$\|d\mathbf{h}^{ij}\|_{L^2_t L^\infty_x} + \|\partial_x\mathbf{h}^{ij}\|_{L^\infty_t H^{s-\frac{3}{2}}_{x'}(\Sigma)} \lesssim 1\,. \tag{5.15}$$

The first term in (5.14) is bounded using (5.5), (5.6), and (5.11). The second term in (5.14) is bounded using (5.9), (5.10), and the trace theorem applied to $\mathbf{g}^{ij}$. The first term in (5.15) uses the uniform bounds on $\mathbf{g}^{ij}$ and $d\phi$, as well as the $L^2_t L^\infty_x$ bounds on $d\mathbf{g}^{ij}$ and $d^2\phi$, the latter a consequence of (5.6) and the Sobolev embedding $H^{s-1}(\Sigma^t) \subset L^\infty(\Sigma^t)$. For the second term in (5.15), by the line above we need only consider the case ∂_x replaced by ∂_{x_n}, for which case we use the inequality

$$\|(\partial\phi)^\alpha(\partial\tilde{\mathbf{g}}^{ij})\|_{L^\infty_t H^{s-\frac{3}{2}}_{x'}(\Sigma)} \lesssim \|\partial\phi\|^{|\alpha|}_{L^\infty_t H^{s-\frac{1}{2}}_{x'}(\Sigma)}\,\|\partial\mathbf{g}^{ij}\|_{L^\infty_t H^{s-\frac{3}{2}}_{x'}(\Sigma)}\,.$$

To continue write

$$\begin{aligned}\partial_i\left(\mathbf{h}^{ij}\,\partial_j(\partial_x\langle D_{x'}\rangle^{s-2}\tilde{w})\right) &= \partial_x\langle D_{x'}\rangle^{s-2}G - \left[\partial_i\partial_x\langle D_{x'}\rangle^{s-2},\mathbf{h}^{ij}\right]\partial_j\tilde{w}\\ &\quad + \left(\partial_i\mathbf{h}^{ij}\right)\partial_x\,\langle D_{x'}\rangle^{s-2}\partial_j\tilde{w}\,.\end{aligned}$$

By the Kato-Ponce commutator estimate, noting that $i \neq 0$ in the commutator term, we have for each fixed t the bound

$$\left\| \left[\partial_i \partial_x \langle D_{x'} \rangle^{s-2}, \mathbf{h}^{ij} \right] \partial_j \tilde{w} \right\|_{L^2_x} \lesssim \|\mathbf{h}^{ij}\|_{\mathrm{Lip}_x} \|dw\|_{H^{s-1}_x} + \|\mathbf{h}^{ij}\|_{H^s_x} \|dw\|_{L^\infty_x},$$

where all norms are taken over an arbitrary slice $t = \text{constant}$, and we use (5.13) to bound norms of $\tilde{w}$ by the same norms of w. Also,

$$\|d_x \langle D_{x'} \rangle^{s-2} G\|_{L^2_x} \lesssim \|F\|_{H^{s-1}_x} + \|\partial_0^2 \phi\|_{L^\infty_x} \|\partial_n w\|_{H^{s-1}_x} + \|\partial_0^2 \phi\|_{H^{s-1}_x} \|\partial_n w\|_{L^\infty_x}$$

and

$$\|\left(\partial_i \mathbf{h}^{ij}\right) \partial_x \langle D_{x'} \rangle^{s-2} \partial_j \tilde{w}\|_{L^2_x} \lesssim \|d\mathbf{h}^{ij}\|_{L^\infty_x} \|\partial_j w\|_{H^{s-1}_x}.$$

Consequently,

$$\left\| \partial_i \left(\mathbf{h}^{ij} \, \partial_j (\partial_x \langle D_{x'} \rangle^{s-2} \tilde{w}) \right) \right\|_{L^1_t L^2_x} \lesssim \epsilon.$$

Recall that Σ is a null surface, defined in these coordinates by $x_n = 0$. By the energy inequality, we thus obtain

$$\|\partial_t \partial_x \langle D_{x'} \rangle^{s-2} \tilde{w}\|_{L^2(\Sigma)} + \|\partial_{x'} \partial_x \langle D_{x'} \rangle^{s-2} \tilde{w}\|_{L^2(\Sigma)} \lesssim \epsilon.$$

The trace theorem shows that $\|dw\|_{L^2(\Sigma)} \lesssim \epsilon$, and it therefore remains to show that

$$\|\partial_t^2 \langle D_{x'} \rangle^{s-2} \tilde{w}\|_{L^2(\Sigma)} \lesssim \epsilon.$$

Since $\mathbf{h}^{00} = -1$, we may write

$$\partial_t^2 \tilde{w} = \tilde{F} - (\partial_0^2 \phi) \, \partial_n \tilde{w} + \sum_{i=1}^n \sum_{j=0}^n \partial_i \left(\mathbf{h}^{ij} \, \partial_j \tilde{w} \right).$$

To handle the contribution from the first two terms we apply the trace theorem and the fact that $s - 1 > \frac{n-1}{2}$. to get

$$\|\langle D_{x'} \rangle^{s-\frac{3}{2}} \tilde{F}\|_{L^2(\Sigma)} \lesssim \|F\|_{L^2 H^{s-1}} \lesssim \epsilon$$

and

$$\begin{aligned} \|\langle D_{x'} \rangle^{s-\frac{3}{2}} (\partial_0^2 \phi \, \partial_n \tilde{w})\|_{L^2(\Sigma)} &\lesssim \|\langle D_{x'} \rangle^{s-1} \partial_0^2 \phi\|_{L^2(\Sigma)} \|\langle D_{x'} \rangle^{s-\frac{3}{2}} \partial_n w\|_{L^\infty_t L^2_{x'}(\Sigma)} \\ &\lesssim \|d\phi\|_{s,2,\Sigma} \, \|\partial_n w\|_{L^\infty_t H^{s-1}_x} \\ &\lesssim \epsilon, \end{aligned}$$

For the remaining terms we first note that, since $2s - 3 > (n-1)/2$, we may apply (5.8) and (5.15) to bound

$$\begin{aligned} \|(\partial_i \mathbf{h}^{ij}) \, (\partial_j \tilde{w})\|_{L^\infty_t H^{s-2}_{x'}(\Sigma)} &\lesssim \|(\partial_i \mathbf{h}^{ij})\|_{L^\infty_t H^{s-\frac{3}{2}}_{x'}(\Sigma)} \, (\partial_j \tilde{w})\|_{L^\infty_t H^{s-\frac{3}{2}}_{x'}(\Sigma)} \\ &\lesssim \|\partial_x \mathbf{h}^{ij}\|_{L^\infty_t H^{s-\frac{3}{2}}_{x'}(\Sigma)} \|dw\|_{L^\infty_t H^{s-1}_x} \\ &\lesssim \epsilon. \end{aligned}$$

Next, since Σ is null and is defined by $x_n = 0$, we get $\mathbf{h}^{nn} = 0$ on Σ. Then,

$$\|\mathbf{h}^{ij}\,\partial_i\,\partial_j\tilde{w}\|_{L^2_t H^{s-2}_{x'}(\Sigma)} \lesssim \|\mathbf{h}^{ij}\|_{L^\infty_t H^{s-\frac{1}{2}}_{x'}(\Sigma)}\,\|\langle D_{x'}\rangle^{s-2}\,\partial_i\,\partial_j\tilde{w}\|_{L^2(\Sigma)}$$
$$\lesssim \epsilon\,,$$

where we use (5.14) and the fact that $i \neq 0$ and $(i,j) \neq (n,n)$. $\square$

Corollary 5.7. *Suppose that w satisfies the conditions of Lemma 5.6. Then*

$$\|\!|\big(\lambda(w - w_\lambda), dw_\lambda, \lambda^{-1} d\partial_x w_\lambda\big)\|\!|_{s-1,2,\Sigma} \lesssim \|dw\|_{L^\infty_t H^{s-1}_x \cap L^2_t L^\infty_x} + \|F\|_{L^2_t H^{s-1}_x}$$

Proof. As before, let

$$\|dw\|_{L^\infty_t H^{s-1}_x} + \|dw\|_{L^2_t L^\infty_x} + \|F\|_{L^2_t H^{s-1}_x} = \epsilon\,.$$

Suppose that P is a standard multiplier of order 0 on $\mathbb{R}^n_x$, such that P is additionally bounded on $L^\infty(\mathbb{R}^n_x)$. Then

$$\partial_i\big(\mathbf{g}^{ij}\,\partial_j\,Pw\big) = G = (\partial_i\,\mathbf{g}^{ij})\,P\,\partial_j w + \big[\mathbf{g}^{ij}, \partial_i P\big]\,\partial_j w + PF\,.$$

The Kato-Ponce commutator estimate and the assumptions on $\mathbf{g}^{ij}$ imply

$$\|G\|_{L^2_t H^{s-1}_x} \lesssim \epsilon\,,$$

and Lemma 5.6 then shows that

$$\|\!|dPw\|\!|_{s-1,2,\Sigma} \lesssim \epsilon\,. \tag{5.16}$$

To control the norm of $\lambda\,(w - w_\lambda)$, we write

$$\lambda\,(w - w_\lambda) = \sum_{k=1}^{n} \partial_k\,P_k w\,,$$

where P_k satisfies the above conditions for P. Applying (5.16) yields the desired bound.

Finally, applying (5.16) to $P = S_{<\lambda}$ and $P = \lambda^{-1}\partial_x S_{<\lambda}$ shows that

$$\|\!|dw_\lambda\|\!|_{s-1,2,\Sigma} + \lambda^{-1}\,\|\!|d\,\partial_x w_\lambda\|\!|_{s-1,2,\Sigma} \lesssim \epsilon\,.$$

$\square$

5.3. **Proof of Proposition 5.1.** This is an immediate consequence of Lemma 5.6 and Corollary 5.7, and (5.5), once we verify that, for each k, ℓ,

$$\|\mathbf{g}^{ij}\partial_i\,\partial_j\,\mathbf{g}^{k\ell}\|_{L^2_t H^{s-1}_x} \lesssim \epsilon_2\,.$$

To begin, suppose that $f(t,x,u)$ is a smooth, compactly supported function of its arguments. Then since $s > \frac{n}{2}$, we have the bound

$$\|f(t,x,u)\|_{L^\infty_t H^s_x} \le C\,,$$

where C depends on uniform bounds on a finite number of derivatives of f. Consequently, by (5.4) we have the bound

$$\|f(t,x,u)\,(u,\partial_i u)\|_{L^\infty_t H^{s-1}_x} + \|f(t,x,u)\,\partial_i u\,\partial_j u\|_{L^2_t H^{s-1}_x} \lesssim \epsilon_2\,,$$

where the second term is bounded as a consequence of the inequality

$$\|(du)^2\|_{H^{s-1}_x} \lesssim \|du\|_{L^\infty_x}\|du\|_{H^{s-1}_x}\,.$$

The result now follows as a consequence of (5.3). $\square$

5.4. **The null frame and an elliptic estimate.** We introduce a null frame along Σ as follows. First, we let

$$V = (dr)^*\,,$$

where r is the defining function of the foliation Σ, and where $*$ denotes the identification of covectors and vectors induced by $\mathbf{g}$. Then V is the null geodesic flow field tangent to Σ. Let

$$\sigma = dt(V)\,, \qquad l = \sigma^{-1}\,V\,. \tag{5.17}$$

Thus l is the $\mathbf{g}$-normal field to Σ normalized so $dt(l) = 1$, hence

$$l = \langle dt, dx_n - d\phi\rangle_{\mathbf{g}}^{-1}\,\big(\,dx_n - d\phi\,\big)^*\,, \tag{5.18}$$

so the coefficients l^j are smooth functions of u and $d\phi$. Conversely,

$$dx_n - d\phi = \langle l, \partial_{x_n}\rangle_{\mathbf{g}}^{-1}\,l^*\,, \tag{5.19}$$

so that $d\phi$ is a smooth function of u and the coefficients of l.

Next we introduce vector fields $e_a : 1 \le a \le n-1$ tangent to the fixed-time slices Σ^t of Σ. We do this by applying Grahm-Schmidt orthogonalization in the metric $\mathbf{g}_{ij} : 1 \le i, j \le n$ to the Σ^t-tangent vector fields $\partial_{x_a} + (\partial_{x_a}\phi)\,\partial_{x_n}$.

Finally, we let

$$\underline{l} = l + 2(dt)^*\,.$$

It follows that $\{l\,,\,\underline{l}\,,e_a\}$ form a null frame in the sense that

$$\langle l, \underline{l}\,\rangle_{\mathbf{g}} = 2\,, \qquad \langle e_a, e_b\rangle_{\mathbf{g}} = \delta_{ab}\,,$$

$$\langle l, l\rangle_{\mathbf{g}} = \langle\,\underline{l}\,,\,\underline{l}\,\rangle_{\mathbf{g}} = \langle l, e_a\rangle_{\mathbf{g}} = \langle\,\underline{l}\,, e_a\rangle_{\mathbf{g}} = 0\,.$$

The coefficients of each of the fields is a smooth function of u and $d\phi$, and by assumption it also follows that we have pointwise bounds

$$| \, e_a - \partial_{x_a} \, | + | \, l - (\partial_t + \partial_{x_n}) \, | + | \, \underline{l} - (-\partial_t + \partial_{x_n}) \, | \lesssim \epsilon_1 \, .$$

Lemma 5.8. *Suppose that* $\mathbf{g}^{ij} \, \partial_i \, \partial_j w = F$. *Let* $(t, x', \phi(t, x'))$ *denote the projective parametrisation of* Σ, *and for* $0 \le i, j \le n-1$, *let* $\not{\partial}_i$ *denote differentiation along* Σ *in the induced coordinates. Then, for* $0 \le i, j \le n-1$, *one may write*

$$\not{\partial}_i \, \not{\partial}_j \, (w|_\Sigma) = l(f_2) + f_1 \, ,$$

where

$$\begin{aligned} &\|f_2\|_{L^2_t H^{s-1}_{x'}(\Sigma)} + \|f_1\|_{L^1_t H^{s-1}_{x'}(\Sigma)} \\ &\qquad \lesssim \|dw\|_{L^\infty_t H^{s-1}_x} + \|dw\|_{L^2_t L^\infty_x} + \|F\|_{L^2_t H^{s-1}_x} + \|F\|_{L^1_t H^{s-1}_{x'}(\Sigma)} \, , \end{aligned}$$

and for each value of t,

$$\|f_2(t\,,\cdot\,)\|_{C^\delta_{x'}(\Sigma^t)} \lesssim \|dw(t, \cdot\,)\|_{C^\delta_x(\mathbb{R}^n)} \, .$$

Proof. Let

$$\|dw\|_{L^\infty_t H^{s-1}_x} + \|dw\|_{L^2_t L^\infty_x} + \|F\|_{L^2_t H^{s-1}_x} + \|F\|_{L^1_t H^{s-1}_{x'}(\Sigma)} = \epsilon \, .$$

The conditions of Lemma 5.6 are satisfied since

$$\begin{aligned} \|(\partial_i \mathbf{g}^{ij}) \partial_j w\|_{L^2_t H^{s-1}_x} &\lesssim \|d\mathbf{g}^{ij}\|_{L^2_t L^\infty_x} \|dw\|_{L^\infty_t H^{s-1}_x} + \|d\mathbf{g}^{ij}\|_{L^\infty_t H^{s-1}_x} \|dw\|_{L^2_t L^\infty_x} \\ &\lesssim \;\; \epsilon \, . \end{aligned}$$

Consequently $\|dw\|_{s-1,2,\Sigma} \lesssim \epsilon$. We make the change of coordinates $x_n \to x_n - \phi(t, x')$ as before, which reduces Σ to the set $x_n = 0$. In these coordinates the equation takes the form

$$\mathbf{h}^{ij} \, \partial_i \, \partial_j \tilde{w} = \tilde{F} + \tilde{\mathbf{g}}^{k\ell} \, (\partial_k \, \partial_\ell \phi) \, \partial_n \tilde{w} = F_1 \, ,$$

where now

$$\mathbf{h}^{ij} = \tilde{\mathbf{g}}^{ij} - \delta^{in} \, (\partial_k \phi) \, \tilde{\mathbf{g}}^{kj} - \tilde{\mathbf{g}}^{ik} \, (\partial_k \phi) \, \delta^{jn} + \delta^{in} \, \delta^{jn} \, (\partial_k \phi) \, (\partial_\ell \phi) \, \tilde{\mathbf{g}}^{k\ell} \, .$$

The metric $\mathbf{h}^{ij}$ satisfies the bounds (5.14) and (5.15) as before. Also,

$$\|F_1\|_{L^1_t H^{s-1}_{x'}(\Sigma)} \lesssim \epsilon \, .$$

To see this, we note that $H^{s-1}_{x'}(\Sigma^t)$ is an algebra, and that by the trace theorem, (5.13), and (5.6), and by Lemma 5.6, we have

$$\|\tilde{\mathbf{g}}^{k\ell}\|_{L^\infty_t H^{s-1}_{x'}(\Sigma)} + \|\partial_k \partial_\ell \phi\|_{L^2_t H^{s-1}_{x'}(\Sigma)} \lesssim 1 \, , \qquad \|\partial_n w\|_{L^2_t H^{s-1}_{x'}(\Sigma)} \lesssim \epsilon \, .$$

We let l^i, $\underline{l}^{\,i}$, e^i_a denote the coefficients of the null frame $\{l, \, \underline{l}\,, e_a\}$ in these coordinates. Thus, $l^0 = 1$, and $l^n = e^n_a = 0$. Each coefficient may be written along Σ as a smooth combination of the $\mathbf{h}^{ij}$, and is equal to its constant coefficient version for $|x|$ large. Consequently, (5.6), Proposition 5.1, and (5.12) together imply

$$\|l^i-\delta^{i0}\|_{s,2,\Sigma}+\|\,\underline{l}^{\,i}+\delta^{i0}-2\delta^{in}\|_{s,2,\Sigma}+\|e_a^i-\delta^{ia}\|_{s,2,\Sigma}\lesssim\epsilon_1\,. \tag{5.20}$$

In particular,

$$\begin{aligned}&\|l^i\|_{L^\infty_t H^{s-1}_{x'}(\Sigma)}\lesssim 1\,,\\ &\|\not\partial\, l^i\|_{L^2_tH^{s-1}_{x'}(\Sigma)}+\|\not\partial\,\underline{l}^{\,i}\|_{L^2_tH^{s-1}_{x'}(\Sigma)}+\|\not\partial e_a^i\|_{L^2_tH^{s-1}_{x'}(\Sigma)}\lesssim\epsilon_1\,.\end{aligned} \tag{5.21}$$

Since $\|dw\|_{L^2_tH^{s-1}_{x'}(\Sigma)}\lesssim\epsilon$, we may reduce matters to showing that we can write

$$\partial_k\,\partial_\ell\,w=l(f_2)+f_1\,,\quad\text{where}\quad 1\le k,\ell\le n-1\,.$$

We define

$$\not\Delta w=\sum_{i,j,a=1}^{n-1}e_a^i\,e_a^j\,\partial_i\,\partial_j w\,.$$

Since $\{l\,,\ \underline{l}\,,\ e_a\}$ is a null frame, we have

$$\not\Delta w=-\sum_{i,j=0}^{n}l^i\ \underline{l}^{\,j}\partial_i\,\partial_j\,w+F_1$$

$$=-l(\,\underline{l}\,w)+G\,,$$

where $G=l(\,\underline{l}^{\,j})\,\partial_jw+F_1$ satisfies $\|G\|_{L^1_tH^{s-1}_{x'}(\Sigma)}\lesssim\epsilon$ by (5.21). We thus write

$$\partial_k\,\partial_\ell\,w=l\left(\partial_k\,\partial_\ell\not\Delta^{-1}(\,\underline{l}\,w)\right)+\left[\partial_k\,\partial_\ell\not\Delta^{-1},l\,\right](\,\underline{l}\,w)+\partial_k\,\partial_\ell\not\Delta^{-1}G\,,$$

where, with $\not\Delta_0=\sum_{i=1}^{n-1}\partial_i^2$, we may expand

$$\partial_k\,\partial_\ell\not\Delta^{-1}=\partial_k\,\partial_\ell\not\Delta_0^{-1}\sum_{k=0}^{\infty}\Big((\delta^{ij}-e_a^ie_a^j)\,\partial_i\,\partial_j\not\Delta_0^{-1}\Big)^k\,,$$

which by (5.20) and the algebra property of $H^{s-1}_{x'}(\Sigma^t)$ is for each t a bounded operator on $H^{s-1}_{x'}(\Sigma^t)$ with norm independent of t. It follows that

$$\|\partial_k\,\partial_\ell\not\Delta^{-1}(\,\underline{l}\,w)\|_{L^2_tH^{s-1}_{x'}(\Sigma)}\lesssim\epsilon\,,$$

$$\|\partial_k\,\partial_\ell\not\Delta^{-1}G\|_{L^1_tH^{s-1}_{x'}(\Sigma)}\lesssim\epsilon\,.$$

To handle the commutator term, it suffices to show that

$$\| \, [(\delta^{ij} - e_a^i e_a^j)\partial_i\partial_j \not\Delta_0^{-1}, l\,] f \|_{L_t^1 H_{x'}^{s-1}(\Sigma)} \lesssim \|f\|_{L_t^2 H_{x'}^{s-1}(\Sigma)}\,.$$

To do this, we bound the left hand side by

$$\|l(e_a^i e_a^j)\,\partial_i\partial_j \not\Delta_0^{-1} f\|_{L_t^1 H_{x'}^{s-1}(\Sigma)} + \|(\delta^{ij} - e_a^i e_a^j)\partial_i\partial_j \not\Delta_0^{-1}\big((\partial_k l^k) f\big)\|_{L_t^1 H_{x'}^{s-1}(\Sigma)}$$
$$+ \|(\delta^{ij} - e_a^i e_a^j)\,[\,\partial_i\partial_j\partial_k \not\Delta_0^{-1}, l^k\,] f\|_{L_t^1 H_{x'}^{s-1}(\Sigma)}\,,$$

where we have $1 \le k \le n-1$, since $l^0 = 1$. The first two terms have the desired bound by the algebra property of $H_{x'}^{s-1}$. For the third term, we use the Kato-Ponce estimate

$$\| \, [\partial_i\partial_j\partial_k \not\Delta_0^{-1}, l^k\,]\, f\,\|_{H_{x'}^{s-1}(\Sigma^t)} \lesssim \|\, l^k - \delta^{k0}\|_{\mathrm{Lip}_{x'}(\Sigma^t)}\, \|f\|_{H_{x'}^{s-1}(\Sigma^t)}$$
$$+ \|\, l^k - \delta^{k0}\|_{H_{x'}^{s}(\Sigma^t)}\, \|f\|_{L_{x'}^\infty(\Sigma^t)}$$

$$\lesssim \|\, l^k - \delta^{k0}\|_{H_{x'}^{s}(\Sigma^t)}\, \|f\|_{H_{x'}^{s-1}(\Sigma^t)}\,.$$

To conclude the proof, we need establish the $C_{x'}^\delta$ bounds on $f_2(t,\cdot\,)$. By (5.20), it follows that the coefficients of the null frame belong to $C_{x'}^\delta(\Sigma^t)$, with uniform bounds over t. As above we may thus reduce consideration to $\partial_i\,\partial_j\,w$, in the projective coordinates on Σ. Since (5.20) shows that $\|e_a^i\,e_a^j - \delta^{ij}\|_{C_{x'}^\delta(\Sigma^t)} \lesssim \epsilon_1$, we have

$$\|f_2(t,\cdot\,)\|_{C_{x'}^\delta(\Sigma^t)} = \|\partial_i\,\partial_j\,\not\Delta^{-1}(\,\underline{l}\,w)(t,\cdot\,)\|_{C_{x'}^\delta(\Sigma^t)} \lesssim \|dw(t,\cdot\,)\|_{C_x^\delta(\mathbb{R}^n)}\,.$$

□

Corollary 5.9. *Let R be the Riemann curvature tensor for the metric $\mathbf{g}$, and let $e_0 = l$. Then for any collection $0 \le a,b,c,d \le n-1$, We may write*

$$\langle R(e_a,e_b)\,e_c,e_d\rangle_{\mathbf{g}}\,|_\Sigma = l(f_2) + f_1\,,$$

where

$$\|f_2\|_{L_t^2 H_{x'}^{s-1}(\Sigma)} + \|f_1\|_{L_t^1 H_{x'}^{s-1}(\Sigma)} \lesssim \epsilon_2\,,$$

and for each value of t,

$$\|f_2(t,\cdot\,)\|_{C_{x'}^\delta(\Sigma^t)} \lesssim \sup_{i,j}\,\|d\mathbf{g}^{ij}(t,\cdot\,)\|_{C_x^\delta(\mathbb{R}^n)}\,.$$

Proof. The curvature expression takes the form $R_{ijk\ell}\,e_a^i e_b^j e_c^k e_d^\ell$, where

$$R_{ijk\ell} = \frac{1}{2}\Big[\partial_i\,\partial_k\,\mathbf{g}_{j\ell} + \partial_j\,\partial_\ell\,\mathbf{g}_{ik} - \partial_j\,\partial_k\,\mathbf{g}_{i\ell} - \partial_i\,\partial_\ell\,\mathbf{g}_{jk}\Big] + Q\big(\mathbf{g}^{ij}, d\mathbf{g}_{ij}\big)\,,$$

where Q is a sum of products of coefficients of $\mathbf{g}^{ij}$ with quadratic forms in $d\mathbf{g}_{ij}$. It follows by Proposition 5.1, which applies to $\mathbf{g}_{ij}$ as well as

$\mathbf{g}^{ij}$, that the term Q satisfies the bound required of f_1 . We therefore look at the term $e_a^i e_c^k \partial_i\, \partial_k\, \mathbf{g}_{j\ell}$, which is typical. By (5.20) and Proposition 5.1, the term $e_a(e_c^k)\, \partial_k \mathbf{g}_{j\ell}$ satisfies the bound required of f_1, so we consider $e_a(e_c\, \mathbf{g}_{j\ell})$. Finally, since the coefficients of e_c in the basis $\not\partial_i$ have tangential derivatives bounded in $L^2_t H^{s-1}_{x'}(\Sigma)$, we are reduced by Lemma 5.8 to verifying that

$$\|\mathbf{g}^{ij}\, \partial_i\, \partial_j\, \mathbf{g}_{k\ell}\|_{L^2_t H^{s-1}_x} + \|\mathbf{g}^{ij}\, \partial_i\, \partial_j\, \mathbf{g}_{k\ell}\|_{L^1_t H^{s-1}_{x'}(\Sigma)} \lesssim \epsilon_2\, .$$

The bound on the first term follows by the proof of Proposition 5.1. The same proof, together with the bound $\|\partial_i u\|_{L^2_t H^{s-1}_{x'}(\Sigma)} \lesssim \epsilon_2$, also bounds the second term. □

5.5. **Connection coefficients and the Raychaudhuri equation.** We will work with the following selected subset of the connection coefficients for the null frame with respect to covariant differentiation along Σ,

$$\chi_{ab} = \langle D_{e_a} l, e_b\rangle_{\mathbf{g}} \qquad l(\,\ln\sigma) = \frac{1}{2}\langle D_l\, \underline{l}\,, l\,\rangle_{\mathbf{g}} \qquad \mu_{0ab} = \langle D_l e_a, e_b\rangle_{\mathbf{g}}$$

For σ set the initial data $\sigma = 1$ at time -2. The coefficients of l and e_a are combinations of coefficients of $\mathbf{g}$ and $d\phi$, by (5.18) and the orthogonalization process. Consequently by (5.12), together with Proposition 5.1 and (5.6), it follows that

$$\|\chi_{ab}\|_{L^2_t H^{s-1}_{x'}(\Sigma)} + \|l(\ln\sigma)\|_{L^2_t H^{s-1}_{x'}(\Sigma)} + \|\mu_{0ab}\|_{L^2_t H^{s-1}_{x'}(\Sigma)} \lesssim \epsilon_1\, . \tag{5.22}$$

Similarly, if we expand $l = l^a \not\partial_a$ in the tangent frame $\partial_t, \partial_{x'}$ on Σ, then

$$l^0 = 1\,, \qquad \sup_{1\le a\le n-1} \|\!|l^a|\!\|_{s,2,\Sigma} \lesssim \epsilon_1\, . \tag{5.23}$$

Lemma 5.10. *Let χ_{ab} be defined as above. Then, for $1 \le a, b \le n-1$,*

$$\|\chi_{ab}\|_{L^2_t H^{s-1}_{x'}(\Sigma)} \lesssim \epsilon_2\, .$$

Furthermore, for each value of t,

$$\|\chi_{ab}(t,\cdot)\|_{C^\delta_{x'}(\Sigma^t)} \lesssim \epsilon_2 + \sup_{i,j} \|d\mathbf{g}^{ij}(t,\cdot)\|_{C^\delta_x(\mathbb{R}^n)}\, .$$

Proof. It follows from (5.23) and Sobolev embedding that the tangent field l, expressed in the basis $\partial_t, \partial_{x'}$ of tangent vector fields on Σ in (t,x') coordinates, differs from ∂_t by a field with coefficients of small norm in $L^2_t C^{1,\delta}_{x'}$. Consequently, if we introduce coordinates (t,y') on Σ, such that $l(y') = 0$ and $y' = x'$ at $t = -2$, then the y' are a small C^1 perturbation of x'.

We use the transport equation for χ_{ab},

$$l(\chi_{ab}) = \langle R(l,e_a)l, e_b\rangle_{\mathbf{g}} - \chi_{ac}\chi_{cb} - l(\,\ln\sigma)\,\chi_{ab} + \mu_{0ac}\,\chi_{cb} + \mu_{0bc}\,\chi_{ac}\,.$$

By Corollary 5.9, we may write this in the form

$$l(\chi_{ab} - f_2^{ab}) = f_1^{ab} - \chi_{ac}\chi_{cb} - l(\,\ln\sigma)\,\chi_{ab} + \mu_{0ac}\,\chi_{cb} + \mu_{0bc}\,\chi_{ac}\,.$$

As before, let Λ^{s-1} be the fractional derivative operator in the x' variables. Then, since $H^{s-1}_{x'}(\Sigma^t)$ is an algebra, we may for each t bound the norm of the right hand side in $H^{s-1}_{x'}(\Sigma^t)$ by

$$h_1(t) + h_2(t)\,\sup_{a,b}\,\|\Lambda^{s-1}(\chi_{ab} - f_2^{ab})(t,\cdot)\|_{L^2_{x'}(\Sigma^t)}\,,$$

where by (5.22) and Corollary 5.9 we have

$$\|h_1\|_{L^1([-2,2])} \lesssim \epsilon_2\,, \qquad \|h_2\|_{L^2([-2,2])} \lesssim \epsilon_1\,.$$

We next bound

$$\begin{aligned}\|\,[\Lambda^{s-1}, l](\chi_{ab} - f_2^{ab})(t,\cdot)\,\|_{L^2_{x'}(\Sigma^t)} \le \|(\partial\!\!\!/_c l^c)(\chi_{ab} - f_2^{ab})(t,\cdot)\|_{H^{s-1}_{x'}(\Sigma^t)}\\ + \|\,[\Lambda^{s-1}\partial\!\!\!/_c, l^c](\chi_{ab} - f_2^{ab})(t,\cdot)\,\|_{L^2_{x'}(\Sigma^t)}\,,\end{aligned}$$

which by the Kato-Ponce commutator estimate and the Sobolev embedding theorem is bounded by

$$\sup_{1\le c\le n-1}\|l^c(t,\cdot)\|_{H^s_{x'}(\Sigma^t)}\,\|\Lambda^{s-1}(\chi_{ab} - f_2^{ab})(t,\cdot)\|_{H^{s-1}_{x'}(\Sigma^t)}\,.$$

By (5.23) we thus have the bound

$$\begin{aligned}&\sup_{a,b}\,\|\,l\,\Lambda^{s-1}(\chi_{ab} - f_2^{ab})(t,\cdot)\|_{L^2_{x'}(\Sigma^t)}\\ &\qquad\qquad \le h_1(t) + h_2(t)\,\sup_{a,b}\,\|\Lambda^{s-1}(\chi_{ab} - f_2^{ab})(t,\cdot)\|_{L^2_{x'}(\Sigma^t)}\,,\end{aligned}$$

where $h_1(t)$ and $h_2(t)$ satisfy the bound above. Since the flow of l is C^1 as noted above, together with Gronwall's Lemma this implies that

$$\sup_t\,\|(\chi_{ab} - f_2^{ab})(t,\cdot)\|_{H^{s-1}_{x'}(\Sigma^t)} \lesssim \epsilon_2\,.$$

The conclusion now follows by Corollary 5.9 and Sobolev embedding. $\square$

5.6. **Proof of Proposition 5.2.** Recall that we have fixed $r = 0$ and $\theta = (0, \ldots, 0, 1)$. Note that since $\phi(t, x') = t$ for $t \le -\frac{3}{2}$, it follows by (5.10) and Sobolev embedding that

$$\|\phi(t,x') - t\|_{C^1} \lesssim \|d\phi(t,x') - dt\|_{s,2}\,,$$

so it suffices to dominate the latter quantity by ϵ_2. By (5.19), together with (5.12) and the bounds on $\|\mathbf{g}^{ij} - \eta^{ij}\|_{s,2,\Sigma}$ from Proposition 5.1, this in turn will follow as a consequence of the bound

$$\|\, l - (\partial_t - \partial_{x_n})\|_{s,2,\Sigma} \lesssim \epsilon_2\,,$$

where it is understood that one takes the norm of the coefficients of $l - (\partial_t - \partial_{x_n})$ in the standard frame on $\mathbb{R}^{n+1}$. The geodesic equation, together with the bound for Christoffel symbols $\|\Gamma^i_{jk}\|_{L^2_t L^\infty_x} \lesssim \epsilon_2$, imply that

$$\|\, l - (\partial_t - \partial_{x_n})\|_{L^\infty_{t,x}} \lesssim \epsilon_2\,,$$

so it suffices to bound the tangential derivatives of the coefficients of $l - (\partial_t - \partial_{x_n})$ in the norm $L^2_t H^{s-1}_{x'}(\Sigma)$. Finally, we claim that we can now reduce matters to dominating the coefficients of $D_l l$ and $D_{e_a} l$ in the tangent frame $\{e_a, l\}$. To see this, we note that the coefficients of e_a and l are small perturbations of their constant coefficient analogs in the $L^2_t H^s_x(\Sigma)$ norm. Also, by Proposition 5.1, we have the bounds for the Christoffel symbols

$$\|\Gamma^i_{jk}\|_{L^2_t H^{s-1}_{x'}(\Sigma)} \lesssim \epsilon_2\,,$$

so that, for instance,

$$\|\Gamma^i_{jk} e^j_a l^k\|_{L^2_t H^{s-1}_{x'}(\Sigma)} \lesssim \epsilon_2\,,$$

and in particular the covariant derivatives of $\partial_t - \partial_{x_n}$ are small in $L^2_t H^{s-1}_{x'}(\Sigma)$.

Thus, we need to establish the following bound,

$$\|\langle D_{e_a} l, e_b\rangle_{\mathbf{g}}\|_{L^2_t H^{s-1}_{x'}(\Sigma)} + \|\langle D_{e_a} l, \underline{l}\,\rangle_{\mathbf{g}}\|_{L^2_t H^{s-1}_{x'}(\Sigma)} + \|\langle D_l l, \underline{l}\,\rangle_{\mathbf{g}}\|_{L^2_t H^{s-1}_{x'}(\Sigma)} \lesssim \epsilon_2.$$

The first term is χ_{ab} which is bounded by Lemma 5.10. For the second we note that

$$\langle D_{e_a} l, \underline{l}\,\rangle_{\mathbf{g}} = \langle D_{e_a} l, 2(dt)^*\rangle_{\mathbf{g}} = -2\,\langle D_{e_a}(dt)^*, l\rangle_{\mathbf{g}}\,.$$

Since the coefficients of $(dt)^*$ are combinations of the $\mathbf{g}^{ij}$, bounds for this term, as well as the last, follow from Proposition 5.1.

It remains to show that

$$\|d\phi(t,\cdot) - dt\|_{C^{1,\delta}_{x'}(\mathbb{R}^{n-1})} \lesssim \epsilon_2 + \|d\mathbf{g}^{ij}(t,\cdot)\|_{C^\delta_x(\mathbb{R}^n)}\,,$$

for which it suffices to show that

$$\| \, l(t,\cdot) - (\partial_t - \partial_{x_n})\|_{C^{1,\delta}_{x'}(\Sigma^t)} \lesssim \epsilon_2 + \|d\mathbf{g}^{ij}(t,\cdot)\|_{C^\delta_x(\mathbb{R}^n)} \,.$$

The coefficients of e_a are small $C^\delta_{x'}(\Sigma^t)$ perturbations of their constant coefficient analogs, so it suffices to show that

$$\|\langle D_{e_a} l, e_b\rangle_{\mathbf{g}}(t,\cdot)\|_{C^\delta_{x'}(\Sigma^t)} + \|\langle D_{e_a} l, \underline{l}\,\rangle_{\mathbf{g}}(t,\cdot)\|_{C^\delta_{x'}(\Sigma^t)} \lesssim \epsilon_2 + \|d\mathbf{g}^{ij}(t,\cdot)\|_{C^\delta_x(\mathbb{R}^n)}.$$

The first term is bounded by Lemma 5.10, and the second by noting that

$$\|\langle D_{e_a}(dt)^*, l\rangle_{\mathbf{g}}(t,\cdot)\|_{C^\delta_{x'}(\Sigma^t)} \lesssim \|d\mathbf{g}^{ij}(t,\cdot)\|_{C^\delta_x(\mathbb{R}^n)} \,. \quad \square$$

6. Geometry of cones

The purpose of this section is to show that any two null foliations Σ_ω and Σ_θ, as defined in section 5, intersect at each point at an angle comparable to $|\omega-\theta|$.

Precisely, let l_ω be the $\mathbf{g}$-normal field to the foliation Σ_ω, normalized as before so that $dt(l_\omega)=1$. We use $o(r)$ to denote a quantity that is bounded by $c\,r$, where c is a small constant which can be made arbitrarily close to 0 by taking ϵ_2 of (5.4) and (5.3) small.

Proposition 6.1. *For all unit vectors ω , $\theta \in S^{n-1}$, uniformly at all points in space-time,*

$$l_\omega - l_\theta = \omega - \theta + o(\,|\omega-\theta|\,)\,. \tag{6.1}$$

As an immediate consequence of this and the fact that l_ω and l_θ are null in $\mathbf{g}$, we have that, uniformly at all points in space-time,

$$\langle l_\omega\,,\, l_\theta\rangle_{\mathbf{g}} = -\frac{1}{2}\,|\omega-\theta|^2 + o(\,|\omega-\theta|^2\,)\,. \tag{6.2}$$

We also establish the following fact about the geodesic flow from a point. For a given point x_1, we let γ_θ denote the null geodesic curve, reparametrised by t, such that $\gamma_\theta(t_1)=x_1$, and $\dot\gamma_\theta(t_1)$ lies along the direction θ.

Proposition 6.2. *For all (t_1,x_1), with $t_1\in[-2,2]$, and all $t\in[-2,2]$,*

$$\gamma_\omega(t)-\gamma_\theta(t) = (t-t_1)\,(\omega-\theta) + o(\,|t-t_1|\,|\omega-\theta|\,)\,. \tag{6.3}$$

To establish (6.1) and (6.2) at a given point (t_1,x_1), we study the cone spanned by the null geodesics through that point.

For the rest of this section, we fix a point (t_1,x_1) , with $t_1\in[-2,2]$. Given $\omega\in S^{n-1}$, let $r(\omega)>0$ be defined so that the vector $(1,r(\omega)\,\omega)$ is null at (t_1,x_1) . Then $r(\omega)=1+o(1)$, and $\partial_\omega r(\omega)=o(1)$. Let $\big(t,\gamma_\omega(t)\big)$ for $t\in[-2,2]$ be the null geodesic curve such that

$$\gamma_\omega(t_1)=x_1\,,\qquad \frac{d\gamma_\omega}{dt}(t_1)=r(\omega)\,\omega\,.$$

At $t=-2$, the metric $\mathbf{g}$ is the standard Minkowski metric, and hence we may write

$$\frac{d\gamma_\omega}{dt}(-2)=\theta(\omega)\,,$$

which defines $\omega\to\theta(\omega)$ as a map $S^{n-1}\to S^{n-1}$. Note that $\theta(\omega)$ is the vector such that γ_ω is tangent to the foliation $\Sigma^{\theta(\omega)}$. Our proof establishes that $\theta(\omega)$ is a small C^1 perturbation of the identity map, which yields (6.1). Since $\theta(\omega)$ is the normal map to the $t=-2$ slice of the light cone with vertex at (t_1,x_1), this in effect says that the map $\gamma_\omega(-2)$ is a small C^2 perturbation of multiplication by $-(2+t_1)\,\omega$. We prove this in turn by first establishing (6.3), which implies that $\gamma_\omega(-2)$ is a small C^1 perturbation of $-(2+t_1)\,\omega$, and then showing that the second fundamental form of the cone is a small C^0 perturbation of that of the tangent cone over (t_1,x_1) . We begin by establishing (6.3).

We start by noting that the bounds on the Christoffel symbols,

$$\|\Gamma_{ij}^k\|_{L^2_tL^\infty_x}=o(1)$$

imply that

$$\frac{d\gamma_\omega}{dt}(t)=r(\omega)\,\omega+o(|t-t_1|^{\frac12})\,,$$

hence that

$$\gamma_\omega(t)=x_1+(t-t_1)\,r(\omega)\,\omega+o(\,|t-t_1|^{\frac32}\,)\,. \tag{6.4}$$

Given a tangent vector v to S^{n-1} at ω, we let Z_v denote the purely spatial vector field along $(t,\gamma_\omega(t))$,

$$Z_v(t)=v\cdot\partial_\omega\gamma_\omega(t)\,,$$

so that

$$Z_v(t_1)=0\,,\qquad \frac{DZ_v}{dt}(t_1)=v\cdot\partial_\omega\big(\,1,r(\omega)\,\omega\,\big)=\big(\,0,v+o(|v|)\,\big)\,.$$

As a variation of reparametrised geodesics, $Z_v(t)$ differs from a Jacobi field along $(t,\gamma_\omega(t))$ by a multiple of $\dot\gamma_\omega(t)$. Hence, using $\frac{D}{dt}$ to denote covariant differentiation along γ_ω , we have

$$\frac{D^2Z_v}{dt^2}=R(\,\dot\gamma_\omega,Z_v\,)\,\dot\gamma_\omega\;-\;\frac{d\ln(\sigma)}{dt}\;\frac{DZ_v}{dt}\qquad \mathrm{mod}\;(\dot\gamma_\omega)\,.$$

Here, σ denotes $\frac{dt}{ds}$, where s is any affine parametrisation of the geodesic γ_ω. By taking s to be the parametrisation with $\sigma(-2) = 1$, then $\sigma = \sigma_{\theta(\omega)}$ where σ_θ is defined as in (5.17). In particular, we have that

$$\left\| \frac{d\ln(\sigma)}{dt} \right\|_{L^2(-2,2)} = o(1)\,. \tag{6.5}$$

The above together imply that $\langle Z_v, \dot\gamma_\omega\rangle_{\mathbf{g}} = 0$ for all t. We now fix a set e_a of purely spatial vector fields along $(t, \gamma_\omega(t))$, orthonormal under $\mathbf{g}$, which together with $(1, \dot\gamma_\omega)$ span the ortho-complement of $(1, \dot\gamma_\omega)$ under $\mathbf{g}$. We may choose e_a such that

$$\frac{De_a}{dt} = 0 \qquad \mathrm{mod}\ (\dot\gamma_\omega)\,,$$

for instance by parallel translating an orthonormal frame along γ_ω and subtracting a multiple of $\dot\gamma_\omega$ to make them purely spatial. We set

$$z_v^a(t) = \langle Z_v(t), e_a(t)\rangle_{\mathbf{g}}\,,$$

and derive the formula

$$\frac{d^2 z_v^a}{dt^2} = \langle R(\,\dot\gamma_\omega, e_b\,)\,\dot\gamma_\omega, e_a\rangle_{\mathbf{g}}\, z_v^b \;-\; \frac{d\ln(\sigma)}{dt}\,\frac{dz_v^a}{dt}\,.$$

By the parallel transport equations, the coefficients of e_a relative to the frame ∂_{x_i} have derivative with small L^2 norm. Hence we may apply Corollary 5.9 to rewrite this equation in the form (along γ_ω)

$$\frac{d}{dt}\Big(\frac{dz_v^a}{dt} - f_2^{ab}\, z_v^b\Big) = \tilde f_2^{ab}\,\Big(\frac{dz_v^a}{dt} - f_2^{ab}\, z_v^b\Big) + f_1^{ab}\, z_v^b\,,$$

where

$$\|f_2^{ab}\|_{L^2(-2,2)} + \|\tilde f_2^{ab}\|_{L^2(-2,2)} + \|f_1^{ab}\|_{L^1(-2,2)} = o(1)\,.$$

Since $z_v^a(t_1) = 0$, and $|\dot z_v^a(t_1)| \le 2\,|v|$, this implies that

$$z_v^a(t) = (t - t_1)\,\dot z_v^a(t_1) + o(\,|t-t_1|^{\frac32}\,|v|\,)\,.$$

Since $e_a(t) = e_a(t_1) + o(|t-t_1|^{\frac32})$ (relative to the frame ∂_{x_i}), this yields

$$Z_v(t) = (t-t_1)\frac{DZ_v}{dt}(t_1) + o(\,|t-t_1|^{\frac32}\,|v|\,)\,,$$

again relative to the frame ∂_{x_i}. Consequently,

$$\gamma_\omega(t) - \gamma_\theta(t) = (t-t_1)\,(\,r(\omega)\omega - r(\theta)\theta\,) + o(\,|t-t_1|^{\frac32}\,|\omega-\theta|\,)\,, \tag{6.6}$$

which in particular implies (6.3).

Together, (6.4) and (6.6) imply that the map $\omega \to \gamma_\omega(-2)$ is an embedding of S^{n-1} into $\mathbb{R}^n$, which is a small C^1 perturbation of the mapping $\omega \to -(2+t_1)\,\omega$. It remains to show that the function $\theta(\omega) = \dot\gamma_\omega(-2)$, considered as a function on this manifold, is a small

C^1 perturbation of the function ω. To do this, we show that, uniformly for each ω,

$$-(2+t_1)\,\langle D_{e_a}\theta(\omega), e_b\rangle_{\mathbf{g}} = \delta_{ab} + o(1)\,.$$

Together with (6.4) and (6.6), this implies (6.1).

We fix ω, and along $\gamma_\omega(t)$ we set

$$H_{ab}(t) = \langle D_{e_a(t)}\,\dot\gamma_\omega(t), e_b(t)\rangle_{\mathbf{g}}\,.$$

Then $H_{ab}(t)$ is well defined and smooth in t for $t \neq t_1$, since the above argument and dilation show that $\omega \to \gamma_\omega(t)$ is a C^∞ embedding for $t \neq t_1$, as $\mathbf{g}_{ab}$ is assumed to be C^∞. A dilation argument shows that

$$H_{ab}(t) = (t-t_1)^{-1}\,\delta_{ab} + h_{ab}(t)\,,$$

where

$$\sup_{t\neq t_1} |h_{ab}(t)| < \infty\,. \tag{6.7}$$

Furthermore, for $t \neq t_1$, we have

$$\frac{dH_{ab}}{dt} = \langle\, R(\,\dot\gamma_\omega, e_a\,)\,\dot\gamma_\omega, e_b\,\rangle_{\mathbf{g}} - \frac{d\ln(\sigma)}{dt}\,H_{ab} - H_{ac}\,H_{cb}\,.$$

Applying Corollary 5.9 as before, and setting $\tilde f_2 = -d\ln(\sigma)/dt$, we obtain upon substitution for H_{ab} the differential equation for h_{ab},

$$\frac{dh_{ab}}{dt} = \frac{df_2^{ab}}{dt} + f_1^{ab} + (t-t_1)^{-1}\,\tilde f_2\delta_{ab} + \tilde f_2 h_{ab} - 2(t-t_1)^{-1}\,h_{ab} - h_{ac}h_{cb}\,,$$

which we may rewrite in the form

$$\begin{aligned}\frac{d}{dt}\big(\,(t-t_1)^2\,(h_{ab}-f_2^{ab})\,\big) = -2\,(t-t_1)f_2^{ab} + (t-t_1)^2\,f_1^{ab} + (t-t_1)\,\tilde f_2\,\delta_{ab}\\ + (t-t_1)^2\,\tilde f_2\,h_{ab} - (t-t_1)^2\,h_{ac}\,h_{cb}\,,\end{aligned}$$

where as before

$$\|f_2^{ab}\|_{L^2(-2,2)} + \|\tilde f_2\|_{L^2(-2,2)} + \|f_1^{ab}\|_{L^1(-2,2)} = o(1)\,.$$

Applying (6.7) leads to the inequality

$$\begin{aligned}\big|h_{ab}(t)\big| \le \big|f_2^{ab}(t)\big| + 2\,|t-t_1|^{-1}\int_{t_1}^t \big(\,\big|f_2^{ab}\big| + \big|\tilde f_2\,\delta_{ab}\big|\,\big)\,ds + \int_{t_1}^t \big|f_1^{ab}\big|\,ds\\ + \int_{t_1}^t \big|\tilde f_2\,h_{ab}\big|\,ds + \int_{t_1}^t \big|h_{ac}\,h_{cb}\big|\,ds\,,\end{aligned}$$

with the order of integration reversed for $t < t_1$. We next note that the first integral on the right hand side is dominated pointwise in t by $M(f_2^{ab}) + M(\tilde f_2)$, where M is the Hardy-Littlewood maximal function, hence the second term has small norm in $L^2(dt)$. A continuity argument in r applied to $\|h_{ab}\|_{L^2(|t-t_1|<r)}$ shows that

$$\|h_{ab}\|_{L^2(-2,2)} = o(1) .$$

Furthermore, since $f_2^{ab}(t) = 0$ for $t < -1$, we have

$$|h_{ab}(-2)| = o(1) .$$

7. The paradifferential decomposition

To conclude the proof of Proposition 4.1 we establish the following:

Proposition 7.1. *Suppose that $u \in \mathcal{H}$, and that $G(u) \leq 2\epsilon_1$. Then condition* (WP4) *is satisfied with* g(u) *replaced by* $\mathbf{g}(t,x,u)$. *That is, the linear equation $\Box_{\mathbf{g}(t,x,u)} v = 0$ is well-posed for data in $H^r \times H^{r-1}$ if $1 \leq r \leq s+1$, and the solutions satisfy the Strichartz estimates* (3.3).

First we show that this yields Proposition 4.1. By Proposition 5.2, we need to show that

$$\|du\|_{L^\infty_t(H^{s-1}_x)} \lesssim \epsilon_3 , \qquad \|du\|_{L^2_t C^\delta_x} \lesssim \epsilon_3 .$$

The first of these is a consequence of Lemma 1.4, since by assumption $\|du\|_{L^2_t C^\delta_x} \leq 2\epsilon_2$. It remains to bound du in $L^2_t C^\delta_x$. The bound would follow directly from Proposition 7.1 if the right hand side of (1.1) were zero. In our case, the result follows by the Duhamel variation of parameters formula, upon verifying that

$$\|\mathbf{q}(t,x,u)\,(du)^2\|_{L^1_t(H^{s-1}_x)} \lesssim \epsilon_3 .$$

But this follows from the fixed time multiplicative estimate

$$\|\mathbf{q}(t,x,u)\,(du)^2\|_{H^{s-1}_x} \lesssim \|du\|_{L^\infty_x}\|du\|_{H^{s-1}_x} ,$$

which is in turn a consequence of (2.5) with $a(\xi) = \langle\xi\rangle^{s-1}$. $\square$

We will establish Proposition 7.1 via an appropriate parametrix construction for the equation $\Box_{\mathbf{g}} v = 0$. The first step in the construction is to make a paradifferential decomposition in order to localize the problem in the frequency variable dual to x. Given a frequency scale $\lambda \geq 1$, we consider the regularized coefficients

$$\mathbf{g}_\lambda = S_{<\lambda}\mathbf{g} ,$$

which we use to study the localized problem at frequency λ. We will begin by showing that Proposition 7.1 is a result of the following

Proposition 7.2. *Suppose that $u \in \mathcal{H}$, and that $G(u) \le 2\epsilon_1$. Then for each $(v_0, v_1) \in H^1 \times L^2$ there exists a function v_λ in $C^\infty([-2,2]\times\mathbb{R}^n)$, with*

$$\text{support}\ \widehat{v_\lambda(t,\cdot)}(\xi) \subseteq \{\,\xi\,:\,\lambda/8 \le |\xi| \le 8\lambda\,\}\,,$$

such that

$$(7.1)\qquad \begin{cases} \|\Box_{\mathbf{g}_\lambda} v_\lambda\|_{L^1_t(L^2_x)} \lesssim \epsilon_0\,\big(\,\|v_0\|_{H^1} + \|v_1\|_{L^2}\big)\,, \\[2ex] v_\lambda(-2) = S_\lambda v_0\,, \quad \partial_t v_\lambda(-2) = S_\lambda v_1\,, \end{cases}$$

and such that the following Strichartz estimates hold, provided $r > \frac{3}{4}$ if $n = 2$, and $r > \frac{n-1}{2}$ if $n \ge 3$,

$$(7.2)\qquad \begin{aligned} &\|v_\lambda\|_{L^4_t(L^\infty_x)} \lesssim \epsilon_0^{-\frac14}\,\lambda^{r-1}\,\big(\,\|v_0\|_{H^1} + \|v_1\|_{L^2}\big)\,, \qquad n = 2\,,\\ &\|v_\lambda\|_{L^2_t(L^\infty_x)} \lesssim \epsilon_0^{-\frac12}\,\lambda^{r-1}\,\big(\,\|v_0\|_{H^1} + \|v_1\|_{L^2}\big)\,, \qquad n = 3,4,5\,. \end{aligned}$$

Roughly speaking, this says that we can find a "good" approximate solution v_λ for the equation

$$(7.3)\qquad \Box_{\mathbf{g}_\lambda} v_\lambda = 0, \qquad v_\lambda(-2) = S_\lambda v_0, \qquad \partial_t v_\lambda(-2) = S_\lambda v_1\,.$$

This result is almost trivial if $\epsilon_0\lambda \lesssim 1$. Indeed, in this case we can let $v_\lambda = S_\lambda v$, where v is the exact solution to $\Box_{\mathbf{g}_\lambda} v = 0$ with data (v_0, v_1), in which case

$$\begin{aligned} \|\Box_{\mathbf{g}_\lambda} v_\lambda\|_{L^2_t L^2_x} &\lesssim \|\,[\mathbf{g}^{ij}_\lambda, \partial_i S_\lambda]\,\partial_j v\,\|_{L^2_t L^2_x} + \|S_\lambda(\partial_i \mathbf{g}^{ij}_\lambda)(\partial_j v)\|_{L^2_t L^2_x}\\ &\lesssim \|d\mathbf{g}^{ij}_\lambda\|_{L^2_t L^\infty_x}\|\partial_i v\|_{L^\infty_t L^2_x}\\ &\lesssim \epsilon_0\,\big(\,\|v_0\|_{H^1} + \|v_1\|_{L^2}\big)\,. \end{aligned}$$

The Strichartz estimates then follow from the Sobolev embedding and the energy estimates. Hence, in the next section we will restrict ourselves to establishing Proposition 7.2 in the case that

$$\epsilon_0\lambda \gg 1\,,$$

in which case we will show that (7.2) holds without the factors of ϵ_0 on the right hand side. For the rest of this section we show that Proposition 7.2 implies Proposition 7.1.

7.1. **Replacing $L^1_t L^2_x$ by $L^2_t L^2_x$.** We prove here that v_λ can be chosen so that a stronger version of (7.1) holds; namely so that

$$\|\Box_{\mathbf{g}_\lambda} v_\lambda\|_{L^2_t L^2_x} \lesssim \epsilon_0 \left(\|v_0\|_{H^1} + \|v_1\|_{L^2}\right).$$

We fix a Littlewood-Paley cutoff $\tilde S_\lambda$ so that $\tilde S_\lambda S_\lambda = S_\lambda$, and so that $\tilde S_\lambda$ is supported in the range $|\xi| \in [\lambda/8, 8\lambda]$.

Suppose we are given initial data (v_0, v_1) with frequencies supported in the range $[\lambda/8, 8\lambda]$. Then $S_\lambda(v_0, v_1)$ vanishes except for a fixed number of dyadic values. Applying Proposition 7.2 to each of these pieces and summing up the resulting approximate solutions we produce a function v which is localized at frequency λ and satisfies

$$\|\Box_{\mathbf{g}_\lambda} v\|_{L^1_t L^2_x} \lesssim \epsilon_0 \left(\|v_0\|_{H^1} + \|v_1\|_{L^2}\right), \quad v(-2) = v_0\,, \quad v_t(-2) = v_1\,.$$

(We use the fact that one may replace $\mathbf{g}_\lambda$ by $\mathbf{g}_{\lambda'}$ with $\lambda' \in [\lambda/4, 4\lambda]$ without changing the result of Proposition 7.2.) We set $v_\lambda = S_\lambda v$ and compute

$$\Box_{\mathbf{g}_\lambda} v_\lambda = S_\lambda f_1 + f_2\,, \qquad f_1 = \Box_{\mathbf{g}_\lambda} v\,, \qquad f_2 = [\Box_{\mathbf{g}_\lambda}, S_\lambda]\, v\,.$$

The commutator term can be estimated as above,

$$\|[\Box_{\mathbf{g}_\lambda}, S_\lambda]\, v\|_{L^2_t L^2_x} \lesssim \|d\mathbf{g}\|_{L^2_t L^\infty_x} \|dv\|_{L^\infty_t L^2_x}\,. \tag{7.4}$$

We thus obtain a smooth function v_λ with

$$v_\lambda(-2) = S_\lambda v_0\,, \qquad \partial_t v_\lambda(-2) = S_\lambda v_1\,,$$

and such that $\Box_{\mathbf{g}_\lambda} v_\lambda = S_\lambda f_1 + f_2$, where

$$\|f_1\|_{L^1_t L^2_x} + \|f_2\|_{L^2_t L^2_x} \lesssim \epsilon_0 \left(\|v_0\|_{H^1} + \|v_1\|_{L^2}\right). \tag{7.5}$$

This is already an improvement over (7.1), since $\Box_{\mathbf{g}_\lambda} v_\lambda$ is the sum of a good term f_2 and a bad term which has the special form $S_\lambda f_1$. We want to eliminate the bad term using an iterative argument based on the Duhamel variation of parameters formula. To do so, we need to construct approximate solutions for Cauchy data specified at arbitrary initial time $t_0 \in [-2, 2]$, and not just $t_0 = -2$.

Precisely, given $(w_0, w_1) \in H^1 \times L^2$, we seek w_λ so that

$$\Box_{\mathbf{g}_\lambda} w_\lambda = S_\lambda f_1 + f_2\,, \quad w(t_0) = S_\lambda w_0\,, \quad w_t(t_0) = S_\lambda w_1\,, \tag{7.6}$$

and such that (7.5) and also the Strichartz estimates (7.2) are satisfied with v replaced by w.

For this, we start with the exact solution w to

$$\Box_{\mathbf{g}_\lambda} w = 0\,, \qquad w(t_0) = w_0\,, \qquad w_t(t_0) = w_1\,,$$

and let $v_\lambda = S_\lambda v$ be the approximate solution constructed as above, with $v_0 = \tilde S_\lambda w(-2)$, $v_1 = \tilde S_\lambda w_t(-2)$. Then by energy estimates and (7.5), v_λ satisfies the correct estimate,

$$\Box_{\mathbf{g}_\lambda} v_\lambda = S_\lambda f_1 + f_2, \quad \|f_1\|_{L^1L^2} + \|f_2\|_{L^2} \lesssim \epsilon_0 \left(\|w_0\|_{H^1} + \|w_1\|_{L^2} \right),$$

as well as the Strichartz estimates (7.2), but it does not match exactly the data at time t_0. However, we have

$$S_\lambda w_0 - v_\lambda(t_0) = S_\lambda w_0^1, \qquad w_0^1 = \tilde S_\lambda w_0 - v(t_0),$$

$$S_\lambda w_1 - \partial_t v_\lambda(t_0) = S_\lambda w_1^1, \qquad w_1^1 = \tilde S_\lambda w_1 - v_t(t_0).$$

We can use energy estimates and the commutator estimate (7.4) to bound the size of the error,

$$\begin{aligned} \|w_0^1\|_{H^1} + \|w_1^1\|_{L^2} &\lesssim \|\Box_{\mathbf{g}_\lambda}(\tilde S_\lambda w - v)\|_{L^1_t L^2_x} \\ &\lesssim \|\Box_{\mathbf{g}_\lambda} v\|_{L^1_t L^2_x} + \|[\Box_{\mathbf{g}_\lambda}, \tilde S_\lambda] w\|_{L^2_t L^2_x} \\ &\lesssim \epsilon_0 \left(\|w_0\|_{H^1} + \|w_1\|_{L^2} \right). \end{aligned}$$

Since the norm of the error is much smaller than the initial size of the data, we may repeat this process with data (w_0^1, w_1^1), and sum the resulting series to obtain a smooth function w_λ with data $S_\lambda(w_0, w_1)$ at time t_0, such that the Strichartz estimates (7.2) are satisfied, and such that $\Box_{\mathbf{g}} w = S_\lambda f_1 + f_2$, with

$$\|f_1\|_{L^1_t L^2_x} + \|f_2\|_{L^2_t L^2_x} \lesssim \epsilon_0 \left(\|w_0\|_{H^1} + \|w_1\|_{L^2} \right).$$

An iteration argument now allows us to eliminate the bad term $S_\lambda f_1$ in $\Box_{\mathbf{g}_\lambda} v_\lambda$. Note also that the above argument implies the result of Proposition 7.2 with -2 replaced by arbitrary $t_0 \in [-2, 2]$. Combining these results, we obtain the following strengthening of Proposition 7.2,

Corollary 7.3. *Suppose that $u \in \mathcal{H}$, and that $G(u) \le 2\epsilon_1$. Then for each $(v_0, v_1) \in H^1 \times L^2$, and each $t_0 \in [-2, 2]$, there exists a function v_λ in $C^\infty([-2, 2] \times \mathbb{R}^n)$, with*

$$\text{support } \widehat{v_\lambda(t, \cdot)}(\xi) \subseteq \{\, \xi \,:\, \lambda/8 \le |\xi| \le 8\lambda \,\},$$

such that

$$\begin{cases} \|\Box_{\mathbf{g}_\lambda} v_\lambda\|_{L^2_t L^2_x} \lesssim \epsilon_0 \left(\|v_0\|_{H^1} + \|v_1\|_{L^2} \right), \\ v_\lambda(t_0) = S_\lambda v_0\,, \quad \partial_t v_\lambda(t_0) = S_\lambda v_1\,, \end{cases}$$

and such that the Strichartz estimates (7.2) *hold.*

7.2. **The case** $r = 1$**.** Since $d\mathbf{g}_\lambda \in L^2_t L^\infty_x$, it follows that equation (1.3) is well-posed in $H^1 \times L^2$, and its solution satisfies the energy estimates. It remains to show that its solution v also satisfies the Strichartz estimates (3.3). Without loss of generality we take $t_0 = 0$.

Given arbitrary initial data $(v_0, v_1) \in H^1 \times L^2$, and general $t_0 \in [-2, 2]$, we take the Littlewood-Paley decomposition

$$v_0 = \sum S_\lambda v_0 \,, \quad v_1 = \sum S_\lambda v_1 \,,$$

and for each λ take the corresponding v_λ as in (7.1). Set

$$v = \sum v_\lambda \,.$$

Then v matches the initial data (v_0, v_1) at time t_0, and satisfies the Strichartz estimates (3.3) (with a constant depending on ϵ_0). We claim that v is also an approximate solution for $\Box_{\mathbf{g}}$, in that

$$\|\Box_{\mathbf{g}} v\|_{L^2_t L^2_x} \lesssim \epsilon_0 \left(\|v_0\|_{H^1} + \|v_1\|_{L^2} \right) .$$

Indeed, we have

$$\Box_{\mathbf{g}} v = \sum_{\lambda \text{ dyadic}} \Box_{\mathbf{g}_\lambda} v_\lambda + \sum_{\lambda \text{ dyadic}} \Box_{\mathbf{g} - \mathbf{g}_\lambda} v_\lambda \,.$$

The first sum is controlled by Corollary 7.3 since the terms have finite overlap on the Fourier transform side. For the second, we first observe that it contains no second order time derivatives, since $\mathbf{g}^{00} = 1$. We set $w_\lambda = dv_\lambda \in L^\infty_t L^2_x$, and rewrite the second term as

$$\sum_{\lambda \text{ dyadic}} (\mathbf{g} - \mathbf{g}_\lambda)\, d_x w_\lambda \,.$$

The bound on this term follows from the fixed time estimate

$$\| \sum_{\lambda \text{ dyadic}} (\mathbf{g} - \mathbf{g}_\lambda)\, d_x w_\lambda \|_{L^2_x} \lesssim \|d\mathbf{g}\|_{C^\delta_x} \Big(\sum \|w_\lambda\|^2_{L^2_x} \Big)^{\frac{1}{2}} ,$$

which follows from the bound

$$\|\mathbf{g} - \mathbf{g}_\lambda\|_{L^\infty_x} \lesssim \lambda^{-1-\delta} \, \|\mathbf{g}\|_{C^\delta_x} \,.$$

Given $F \in L^1_t L^2_x$, we now form the function

$$TF(t,x) = \int_0^t v^s(t,x)\, ds \,,$$

where v^s is the approximate solution formed above, with Cauchy condition

$$v^s(s) = 0 \,, \quad (\partial_t v^s)(s) = F(s, \cdot) \,.$$

Then the above shows that

$$\| \Box_{\mathbf{g}} TF - F \|_{L^2_t L^2_x} \lesssim \epsilon_0 \, \|F\|_{L^1_t L^2_x} \, .$$

Hence the contraction principle implies that we may write the solution v in the form

$$v = \tilde{v} + TF \, ,$$

where $\tilde{v}$ is the approximate solution formed above for data (v_0, v_1) specified at time $t = 0$, and

$$\|F\|_{L^2_t L^2_x} \lesssim \epsilon_0 \left(\|v_0\|_{H^1} + \|v_1\|_{L^2} \right) .$$

The Strichartz estimates now follow since they hold for each v^s .

7.3. **The case** $r = 2$**.** Again we consider $t_0 = 0$. Given data $(v_0, v_1) \in H^2 \times H^1$, we seek a solution of the form $v = \langle D_x \rangle^{-1} w$. Then we require that w have Cauchy data $\langle D_x \rangle (v_0, v_1) \in H^1 \times L^2$, and that w solve

$$\begin{aligned} \Box_{\mathbf{g}} w \;&=\; \big(\Box_{\mathbf{g}} - \langle D_x \rangle \Box_{\mathbf{g}} \langle D_x \rangle^{-1} \big) w \\ &=\; [\mathbf{g}^{ij}, \langle D_x \rangle] \, \langle D_x \rangle^{-1} \partial_i \partial_j w \, , \end{aligned}$$

where we may assume that $i \neq 0$.

For $F \in L^2_t L^2_x$, we form TF as above, but with v^s the exact solution to $\Box_{\mathbf{g}} v^s = 0$, so that $\Box_{\mathbf{g}} TF = F$, and TF has vanishing Cauchy data at $t = 0$. Let $\tilde{w}$ be the solution for the homogeneous equation $\Box_{\mathbf{g}} \tilde{w} = 0$, with Cauchy data $\langle D_x \rangle (v_0, v_1) \in H^1 \times L^2$, at time 0. Then we may find a solution w of the form $w = \tilde{w} + TF$, provided that we show

$$\| \, [\mathbf{g}^{ij}, \langle D_x \rangle] \, \langle D_x \rangle^{-1} \partial_i \partial_j TF \|_{L^2_t L^2_x} \lesssim \epsilon_0 \, \|F\|_{L^2_t L^2_x} \, .$$

This, however, follows from the fixed-time commutator estimate of Coifman and Meyer [5]

$$\| \, [\mathbf{g}^{ij}, \langle D_x \rangle] \, f \|_{L^2_x} \lesssim \|d\mathbf{g}\|_{L^\infty_x} \, \|f\|_{L^2_x} \, ,$$

and the bound

$$\|dTF\|_{L^\infty_t L^2_x} \lesssim \|F\|_{L^1_t L^2_x} \, .$$

At this point, we note that the Duhamel principle implies that for v solving the inhomogeneous problem

$$\Box_g v = G, \quad v(0) = v_0, \quad v_t(0) = v_1$$

and for $r = 1$ and $r = 2$, we have the bounds

$$\begin{aligned} &\|\langle D_x \rangle^\rho v\|_{L^4_t L^\infty_x} \lesssim \|(v_0, v_1)\|_{H^r \times H^{r-1}} + \|G\|_{L^1_t H^{r-1}_x} \, , \quad n = 2 \, , \\ &\|\langle D_x \rangle^\rho v\|_{L^2_t L^\infty_x} \lesssim \|(v_0, v_1)\|_{H^r \times H^{r-1}} + \|G\|_{L^1_t H^{r-1}_x} \, , \quad n = 3, 4, 5 \, , \end{aligned} \tag{7.7}$$

provided $r-\rho>\frac{3}{4}$ for $n=2$, and $r-\rho>\frac{n-1}{2}$ for $n\geq 3$. As an easy consequence of (7.7), we will now show that the following bounds hold for $r=1$ and $r=2$,

$$
\begin{aligned}
&\|\langle D_x\rangle^\rho dv\|_{L^4_tL^\infty_x} \lesssim \|(v_0,v_1)\|_{H^r\times H^{r-1}} + \|G\|_{L^1_tH^{r-1}_x}\,, \quad n=2\,,\\
&\|\langle D_x\rangle^\rho dv\|_{L^2_tL^\infty_x} \lesssim \|(v_0,v_1)\|_{H^r\times H^{r-1}} + \|G\|_{L^1_tH^{r-1}_x}\,, \quad n=3,4,5\,,
\end{aligned}
\tag{7.8}
$$

provided $r-\rho>\frac{7}{4}$ for $n=2$, and $r-\rho>\frac{n+1}{2}$ for $n\geq 3$.

To establish (7.8) for $r=2$, we first consider the case $G=0$. Then

$$\Box_{\mathbf{g}} dv = (d\mathbf{g}^{ij})\partial_i\partial_j v \in L^2_tL^2_x\,,$$

and it is seen from the equation $\Box_{\mathbf{g}} v=0$ that the Cauchy data of dv is of regularity $H^1\times L^2$ if the Cauchy data of v is of regularity $H^2\times H^1$. The estimate (7.8) with $r=2$ then follows from (7.7) with $r=1$ applied to dv. To handle the case $G\neq 0$, we use the Duhamel formula for v, and note that

$$\langle D_x\rangle^\rho d\int_0^t v^s(t,x)\,ds = \int_0^t \langle D_x\rangle^\rho dv^s(t,x)\,ds\,.$$

To establish (7.8) for $r=1$, we note that, if v has Cauchy data of regularity $H^1\times L^2$, then $\langle D_x\rangle^{-1}v$ has Cauchy data of regularity $H^2\times H^1$, and

$$
\begin{aligned}
\|\Box_{\mathbf{g}}\langle D_x\rangle^{-1}v\|_{L^1_tH^1_x} &= \|\langle D_x\rangle\Box_{\mathbf{g}}\langle D_x\rangle^{-1}v\|_{L^1_tL^2_x}\\
&\lesssim \|\,[\langle D_x\rangle,\mathbf{g}^{ij}]\langle D_x\rangle^{-1}\partial_i\partial_j v\|_{L^2_tL^2_x} + \|G\|_{L^1_tL^2_x}\,,
\end{aligned}
$$

and the commutator term is bounded by the Coifman-Meyer estimate together with energy estimates on v.

7.4. **The general case $1\leq r\leq s+1$.** To handle the general case, we first show that the following estimate holds for $1\leq r\leq s+1$,

$$\|\,[\Box_{\mathbf{g}},\langle D_x\rangle^{r-1}]\langle D_x\rangle^{1-r}w\|_{L^2_tL^2_x} \lesssim \epsilon_0\left(\|dw\|_{L^\infty_tL^2_x} + \|\langle D_x\rangle^m dw\|_{L^2_tL^\infty_x}\right),$$

provided $m>1-s$. To see this, we apply analytic interpolation to the family

$$w\to[\Box_{\mathbf{g}},\langle D_x\rangle^z]\langle D_x\rangle^{-z}w\,.$$

For $\operatorname{Re} z=0$, we let $u=\partial_j w$, and use the fixed time commutator estimate

$$\|\,[\mathbf{g}^{ij},\langle D_x\rangle^z]\partial_i u\|_{L^2_x} \lesssim \|d\mathbf{g}^{ij}\|_{L^\infty_x}\,\|u\|_{L^2_x}\,,$$

which follows by (3.6.35) of [24] , where we recall that $i\neq 0$.

For $\operatorname{Re} z = s$, we use the fixed time commutator bound

$$\| [\mathbf{g}^{ij}, \langle D_x\rangle^z]\langle D_x\rangle^{-z}\partial_i u\|_{L^2_x} \lesssim \|d\mathbf{g}^{ij}\|_{L^\infty_x} \|u\|_{L^2_x} + \|\mathbf{g}^{ij} - \eta^{ij}\|_{H^s_x} \|\langle D_x\rangle^{-z}\partial_i u\|_{L^\infty} ,$$

which is a consequence of the Kato-Ponce commutator estimate, see (3.6.14) of [24]. The estimate now follows by analytic interpolation and the fact that

$$\|d\mathbf{g}^{ij}\|_{L^2_t L^\infty_x} + \|\mathbf{g}^{ij} - \eta^{ij}\|_{L^\infty_t H^s_x} \lesssim \epsilon_0 .$$

For Cauchy data (v_0, v_1) of regularity $H^r \times H^{r-1}$, we seek a solution of the form $v = \langle D_x\rangle^{1-r} w$, where w solves

$$\Box_{\mathbf{g}} w = \big(\Box_{\mathbf{g}} - \langle D_x\rangle^{r-1}\Box_{\mathbf{g}}\langle D_x\rangle^{1-r}\big) w .$$

We may obtain a solution of the form $w = w^0 + TF$, with $F \in L^2_t L^2_x$, and where w solves $\Box_{\mathbf{g}} w = 0$, with Cauchy data $\langle D_x\rangle^{r-1}(v_0, v_1)$ of regularity $H^1 \times L^2$, provided we show that

$$\|\big(\Box_{\mathbf{g}} - \langle D_x\rangle^{r-1}\Box_{\mathbf{g}}\langle D_x\rangle^{1-r}\big) TF\|_{L^2_t L^2_x} \lesssim \epsilon_0 \|F\|_{L^2_t L^2_x} .$$

This, however, is a consequence of the above commutator estimate, provided we show that, for some $m > 1 - s$, we have

$$\|dTF\|_{L^\infty_t L^2_x} + \|\langle D_x\rangle^m dTF\|_{L^2_t L^\infty_x} \lesssim \|F\|_{L^1_t L^2_x} .$$

This in turn follows from the case $r = 1$ of (7.8), since we can take $1 - s < m < \frac{1-n}{2}$ if $n \geq 3$, and $1 - s < m < -\frac{3}{4}$ if $n = 2$.

8. The parametrix construction

8.1. **The construction of wave packets.** We introduce in this section the notion of a wave packet, which is central to our parametrix construction. Roughly, a wave packet is an approximate solution to the equation $\Box_{\mathbf{g}} u = 0$, which has a finer spatial localization than a traveling (plane) wave solutions. More precisely, a frequency λ wave packet is localized within λ^{-1} of a characteristic surface $\Sigma_{\omega,r}$, but also roughly within $\lambda^{-\frac{1}{2}}$ of a bicharacteristic ray on $\Sigma_{\omega,r}$. Thus, the support of a wave packet is contained in a curved rectangle which is roughly of size $1 \times \lambda^{-1} \times (\lambda^{-\frac{1}{2}})^{n-1}$. In the sequel we shall call such a region a slab.

While the natural idea might be to start with initial data which is spatially localized in a $\lambda^{-1} \times (\lambda^{-\frac{1}{2}})^{n-1}$ rectangle, as well as frequency localized in the dual rectangle at frequency λ, and transport it along the geodesic flow of $\mathbf{g}$, such a construction does not seem to work, essentially because applying $\Box_{\mathbf{g}}$ to such a wave packet yields an expression which involves (badly behaved) derivatives of the null frame in the direction $\underline{l}$ transversal to the characteristic surfaces.

To avoid having to deal with the behavior of the null foliations Σ_ω in transversal directions, we construct wave packets by starting with a flow-invariant measure on some $\Sigma_{\omega,r}$ (essentially surface measure multiplied by a $\lambda^{-\frac{1}{2}}$ bump function on $\Sigma_{\omega,r}$), then mollifying it on the λ^{-1} scale. The advantage of this approach is that derivative estimates for a wave packet involve only the tangential behavior of restrictions of various functions to the characteristic surfaces $\Sigma_{\omega,r}$, as opposed to their regularity within the support of the wave packet.

Another aspect worth noting in our construction stems from the fact that one cannot localize sharply in both space and frequency. For most of our arguments a sharp spatial localization is more convenient, but the sharp localization in frequency is exploited in order to gain the orthogonality of the wave packets. Consequently, our definition of a wave packet u involves a sharp spatial localization, but the approximate solutions at frequency λ to $\Box_{\mathbf{g}} u = 0$ are constructed as superpositions of $S_\lambda u$. At all instances where we need to take advantage of spatial localization we are able to discard the harmless factor S_λ.

We introduce a spatially localized mollifier T_λ by setting

$$T_\lambda f = \psi_\lambda * f\,, \qquad \psi_\lambda(y) = \lambda^n\, \psi(\lambda^{-1} y)\,,$$

where $\psi \in C_c^\infty(\mathbb{R}^n)$ is supported in the ball $|x| \le \frac{1}{32}$, and has integral 1. By choosing ψ appropriately, any function u with frequency support contained in $|\xi| \le 4\lambda$ can be factored $u = T_\lambda \tilde{u}$, where $\|\tilde{u}\|_{L^2_x} \approx \|u\|_{L^2_x}$.

Finally, we note that our definition of a wave packet involves the small parameter ϵ_0, which is introduced in order to assure that $\Box_{\mathbf{g}} u$ is of small norm, so that we may later obtain exact solutions by an iteration argument.

Definition 8.1. *Let $\gamma = \gamma(t)$ be a geodesic for $\mathbf{g}$, and let $\Sigma_{\omega,r}$ be the null surface introduced in section 4 that contains γ, defined by*

$$\Sigma_{\omega,r} = \{(t,x)\,:\, x_\omega - \phi_{\omega,r}(t, x'_\omega) = 0\,\}\,,$$

where $x_\omega = \langle \omega, x\rangle$, and $x'_\omega \in \mathbb{R}^n$ are projective coordinates along ω. A normalized wave packet around γ is a function u of the form

$$u = \epsilon_0^{\frac{n-1}{4}}\, \lambda^{\frac{n-7}{4}} T_\lambda(vw)\,,$$

where

(i) The function v is simple surface measure on $\Sigma_{\omega,r}$,

$$v(t,x) = \delta(x_\omega - \phi_{\omega,r}(t, x'_\omega))\,,$$

(ii) The function w has the form

$$w = w_0\left((\epsilon_0\lambda)^{\frac{1}{2}}(x'_\omega - \gamma'_\omega(t))\right),$$

where $w_0(x')$ is a smooth function, supported in the set $|x'| \leq 1$, with uniform bounds on its derivatives

$$|\partial^\alpha_{x'} w_0(x')| \leq c_\alpha .$$

As mentioned above, the small parameter ϵ_0 will play an essential role in insuring that $\Box_{\mathbf{g}} u$ not only belongs to $L^1_t L^2_x$, but that it is also of small norm in this space. It would be possible to replace this by relying instead on a rescaling argument in (t, x), but that moves the burden to a different part of the proof. Given ϵ_0, the construction which follows is of interest only if the frequency λ is large enough, namely if

$$\lambda \geq \epsilon_0^{-1} .$$

This is assumed throughout the rest of the paper.

8.2. **An estimate for single wave packets.** Our first goal is to verify that a normalized wave packet is an approximate solution to the wave equation, normalized with respect to the $H^1 \times L^2$ energy. For a single wave packet u, this means establishing an $L^1_t L^2_x$ estimate on $\Box_{\mathbf{g}} u$, which is fairly straightforward. However, we will later need similar estimates for square summable superpositions of wave packets, so it is useful to be more precise at this stage.

We introduce two notations. We use $L(u, v)$ to denote a translation invariant bilinear operator of the form

$$L(u, v)(x) = \int K(y, z)\, u(x + y)\, v(x + z)\, dy\, dz ,$$

where $K(y, z)$ is a finite measure. The particular operator L that arises in Proposition 8.2 below is fixed, and not dependent on either $\mathbf{g}$ or u.

If X is a Sobolev of Hölder space, then we use X_a to denote the same space but with the norm obtained by dimensionless rescaling by a,

$$\|u\|_{X_a} = \|u(a\,\cdot)\|_X .$$

We note that, since $2(s-1) > n-1$, then for $a < 1$ we have

$$\|u\|_{H^{s-1}_a(\mathbb{R}^{n-1})} \lesssim \|u\|_{H^{s-1}(\mathbb{R}^{n-1})} .$$

Proposition 8.2. *Let $u = T_\lambda(vw)$ be a normalized wave packet. Then there is another normalized wave packet $\tilde{u}$, and functions $\psi_m(t, x'_\omega)$,*

$j = 0, 1, 2$, *so that*
(8.1)
$$\Box_{\mathbf{g}_\lambda} S_\lambda u = L(d\mathbf{g}, d\tilde{S}_\lambda \tilde{u}) + \epsilon_0^{\frac{n-1}{4}} \lambda^{\frac{n-7}{4}} S_\lambda T_\lambda \sum_{m=0,1,2} \psi_m \, \delta^{(m)}(x_\omega - \phi_{\omega,r}) \,,$$

where the functions $\psi_m = \psi_m(t, x'_\omega)$ *satisfy the scaled Sobolev estimates*

$$\|\psi_m\|_{L^2_t H^{s-1}_{a,x'_\omega}} \lesssim \epsilon_0 \, \lambda^{1-m} \,, \quad m = 0, 1, 2, \quad a = (\epsilon_0 \lambda)^{-\frac{1}{2}} \,. \tag{8.2}$$

As an immediate consequence, we obtain

Corollary 8.3. *Let u be a normalized wave packet. Then*

$$\|dS_\lambda u\|_{L^\infty_t(L^2_x)} \lesssim 1, \quad \|\Box_{\mathbf{g}_\lambda} S_\lambda u\|_{L^2_t L^2_x} \lesssim \epsilon_0 \,. \tag{8.3}$$

Proof of Proposition 8.2. For the purpose of this proof we consider the case $\omega = (0, 0, \ldots, 1)$, and dispense with the indices ω and r. Then $x_\omega = x_n$, and $x'_\omega = x'$. We write

$$\Box_{\mathbf{g}_\lambda} S_\lambda u = \lambda^{\frac{n-7}{4}} \left([\Box_{\mathbf{g}_\lambda}, S_\lambda T_\lambda] + S_\lambda T_\lambda \Box_{\mathbf{g}_\lambda} \right)(vw) \,. \tag{8.4}$$

For the first term, we use the fact that $\mathbf{g}_\lambda$ is supported at frequency $\leq \lambda/8$ to conclude that only the frequency λ part of vw is contributing. Then we can write

$$[\Box_{\mathbf{g}_\lambda}, S_\lambda T_\lambda] = [\Box_{\mathbf{g}_\lambda}, S_\lambda T_\lambda] \, \tilde{S}_\lambda \tilde{T}_\lambda$$

for some multipliers $\tilde{S}_\lambda \, , \tilde{T}_\lambda$ which have the same properties as $S_\lambda \, , T_\lambda$. Hence it remains to show that

$$[\Box_{\mathbf{g}_\lambda}, S_\lambda T_\lambda] \, u = L(d\mathbf{g}, du) \,.$$

This, however, is a straightforward consequence of the kernel bounds for $S_\lambda T_\lambda$.

For the second term in (8.4), we use the Leibniz rule

$$\Box_{\mathbf{g}_\lambda}(vw) = w \, \Box_{\mathbf{g}_\lambda} v + (\mathbf{g}_\lambda^{ij} + \mathbf{g}_\lambda^{ji}) \, \partial_i v \, \partial_j w + v \, \Box_{\mathbf{g}_\lambda} w \,. \tag{8.5}$$

We consider the three terms separately. In the following computations, the greek indices take values $0 \leq \alpha, \beta \leq n-1$. We let ν denote the conormal vector field along Σ , $\nu = dx_n - d\phi(t, x')$.

The first term in (8.5)**.** We expand $\Box_{\mathbf{g}_l} v$ as a sum of terms

$$\begin{aligned}
\mathbf{g}_\lambda^{ij}\partial_i\partial_j v &= \mathbf{g}_\lambda^{ij}\,\nu_i\,\nu_j\,\delta^{(2)}_{x_n-\phi} - \mathbf{g}_\lambda^{\alpha\beta}\partial_\alpha\partial_\beta\phi\,\delta^{(1)}_{x_n-\phi}\\
&= \mathbf{g}_\lambda^{ij}(t,x',\phi)\,\nu_i\,\nu_j\,\delta^{(2)}_{x_n-\phi} + 2\,(\partial_n\mathbf{g}_\lambda^{ij})(t,x',\phi)\,\nu_i\,\nu_j\,\delta^{(1)}_{x_n-\phi}\\
&\quad + (\partial_n^2\mathbf{g}_\lambda^{ij})(t,x',\phi)\,\nu_i\,\nu_j\,\delta^{(0)}_{x_n-\phi} - \mathbf{g}_\lambda^{\alpha\beta}(t,x',\phi)\,\partial_\alpha\partial_\beta\phi\,\delta^{(1)}_{x_n-\phi}\\
&\quad + (\partial_n\mathbf{g}_\lambda^{\alpha\beta})(t,x',\phi)\,\partial_\alpha\partial_\beta\phi\,\delta^{(0)}_{x_n-\phi}\,.
\end{aligned}$$

Since Σ is characteristic, it follows that $\mathbf{g}^{ij}(t,x',\phi)\,\nu_i\,\nu_j = 0$. We thus take

$$\begin{aligned}
\psi_0 &= w\,[(\partial_n^2\mathbf{g}_\lambda^{ij})(t,x',\phi)\,\nu_i\,\nu_j + (\partial_n\mathbf{g}_\lambda^{\alpha\beta})(t,x',\phi)\,\partial_\alpha\partial_\beta\phi\,]\,,\\
\psi_1 &= w\,[2(\partial_n\mathbf{g}_\lambda^{ij})(t,x',\phi)\,\nu_i\,\nu_j - \mathbf{g}_\lambda^{\alpha\beta}(t,x',\phi)\,\partial_\alpha\partial_\beta\phi\,]\,,\\
\psi_2 &= w\,(\mathbf{g}_\lambda^{ij} - \mathbf{g}^{ij})(t,x',\phi)\,\nu_i\,\nu_j\,.
\end{aligned}$$

It remains to verify that the ψ_m have the appropriate regularity. The function w is a smooth bump on the $(\epsilon_0\lambda)^{-\frac{1}{2}}$ scale, and therefore harmless. Also

$$(\,\mathbf{g}_\lambda\,,\,\nu\,,\,\lambda^{-1}d\mathbf{g}_\lambda\,) \in L_t^\infty H_{x'}^{s-\frac{1}{2}}(\Sigma)\,,$$

so these factors can also be neglected. The conclusion then follows from Proposition 5.1 and (5.6).

The second term in (8.5)**.** Let $\bar{\mathbf{g}}^{ij} = \frac{1}{2}(\mathbf{g}^{ij} + \mathbf{g}^{ji})$. We have

$$\begin{aligned}
\bar{\mathbf{g}}_\lambda^{ij}\,\partial_i v\,\partial_j w &= \nu_i\,\bar{\mathbf{g}}_\lambda^{i\beta}\,\partial_\beta w\,\delta^{(1)}_{x_n-\phi}\\
&= \nu_i\,\bar{\mathbf{g}}_\lambda^{i\beta}(t,x',\phi)\,\partial_\beta w\,\delta^{(1)}_{x_n-\phi} - \nu_i\,(\partial_n\bar{\mathbf{g}}_\lambda^{i\beta})(t,x',\phi)\,\partial_\beta w\,\delta^{(0)}_{x_n-\phi}\,.
\end{aligned}$$

Then we take

$$\psi_0 = \nu_i\,(\partial_n\bar{\mathbf{g}}_\lambda^{i\beta})(t,x',\phi)\,\partial_\beta w\,,\qquad \psi_1 = \nu_i\,\bar{\mathbf{g}}_\lambda^{i\beta}(t,x',\phi)\,\partial_\beta w\,.$$

For ψ_0 we argue as before; differentiating w yields an $(\epsilon_0\lambda)^{\frac{1}{2}}$ factor which is less than the $\epsilon_0\lambda$ we are allowed to lose. The analysis of ψ_1 is more delicate; a rough argument yields the same $(\epsilon_0\lambda)^{\frac{1}{2}}$ loss, but we are not allowed to lose anything. The first useful observation is that we can replace $\bar{\mathbf{g}}_\lambda$ by $\bar{\mathbf{g}}$ in ψ_1, as the error can be controlled as above. The second is the fact that

$$(\nu_i\,\bar{\mathbf{g}}^{i\beta})(t,\gamma(t))\,\partial_\beta w = 0\,.$$

This follows since $(1,\dot\gamma(t))$ is proportional to $(\nu_i\,\bar{\mathbf{g}}^{ij})(t,\gamma(t))$, by (5.18). Consequently, we can write

$$\psi_1 = \sum_{\beta=0}^{n-1}\big[\,(\nu_i\bar{\mathbf{g}}^{i\beta})(t,x',\phi) - (\nu_i\bar{\mathbf{g}}^{i\beta})(t,\gamma(t))\,\big]\partial_\beta w\,.$$

Again, the function $\partial_\beta w$ equals $(\epsilon_0\lambda)^{\frac12}$ times a unit bump on the $(\epsilon_0\lambda)^{-\frac12}$ scale. Also, the function $d(\nu_i \bar{\mathbf{g}}^{i\beta})$ has norm $\lesssim \epsilon_1$ in $L^2_t H^{s-1}_{x'}(\Sigma)$. Then within the support of w, the above difference has size $(\epsilon_0\lambda)^{-\frac12}\epsilon_1$ in $L^2_t H^{s-1}_{x'}(\Sigma)$, which suffices to obtain the desired bound.

The third term in (8.5)**.** This is the easiest one. It only contributes to ψ_0 by

$$\psi_0 = \mathbf{g}^{\alpha\beta}_\lambda(t,x,\phi)\,\partial_\alpha\partial_\beta w\,.$$

The factor $\mathbf{g}^{\alpha\beta}_\lambda(t,x,\phi(t,x'))$ belongs to $L^\infty_t H^{s-\frac12}_{x'}(\Sigma)$, and is therefore negligible. Two spatial derivatives of w yield an $\epsilon_0\lambda$ loss, which is precisely what we are allowed to lose. When differentiating in time we get smooth unit bumps multiplied by either $\epsilon_0\lambda(\dot\gamma)^2$, or by $(\epsilon_0\lambda)^{\frac12}\ddot\gamma$. Both are acceptable since $\|\ddot\gamma\|_{L^2_t} \lesssim \epsilon_1$. $\square$

8.3. **Superpositions of wave packets.** Given an arbitrary initial data set (u_0,u_1) in $H^1\times L^2$, we will construct in the next section a square summable superposition of wave packets,

$$u = \sum_{\omega,j} a_{\omega,j} u^{\omega,j} = \epsilon_0^{\frac{n-1}{4}}\lambda^{\frac{n-7}{4}} T_\lambda \sum_{\omega,j} a_{\omega,j} v^{\omega,j} w^{\omega,j}\,,$$

such that the Cauchy data of $S_\lambda u$ at $t=-2$ equals $S_\lambda(u_0,u_1)$. The purpose of this section is to obtain estimates on $S_\lambda u$ and $\Box_{\mathbf{g}} S_\lambda u$, and so we outline here the important details about the decomposition.

The index ω, which stands for the initial orientation of the wave packet at $t=-2$, varies over a maximal collection of approximately $\epsilon_0^{-\frac{n-1}{2}}\lambda^{\frac{n-1}{2}}$ unit vectors separated by at least $\epsilon_0^{\frac12}\lambda^{-\frac12}$. For each ω we have the orthonormal coordinate system (x_ω, x'_ω) of $\mathbb{R}^n$, where $x_\omega = x\cdot\omega$, and x'_ω are projective along ω.

Next, we decompose $\mathbb{R}^n$ by a parallel tiling of rectangles, with length $(8\lambda)^{-1}$ in the x_ω direction, and $(4\epsilon_0\lambda)^{-\frac12}$ in the remaining directions x'_ω. The index j corresponds to a counting of the rectangles in this decomposition. We let $R_{\omega,j}$ denote the collection of the doubles of these rectangles, and $\Sigma_{\omega,j}$ will denote the element of the Σ_ω foliation upon which $R_{\omega,j}$ is centered. Distinct $\Sigma_{\omega,j}$ are thus separated by at least $(8\lambda)^{-1}$ at $t=-2$, and thus by $(9\lambda)^{-1}$ at all values of t, as shown in (8.6) below. Let $\gamma_{\omega,j}$ denote the null geodesic contained in $\Sigma_{\omega,j}$ which passes through the center of $R_{\omega,j}$ at time $t=-2$.

We let $T_{\omega,j}$ denote the $(32\lambda)^{-1}$ neighborhood of the set

$$\Sigma_{\omega,j}\cap\{\,|x'_\omega - \gamma_{\omega,j}(t)| \le (\epsilon_0\lambda)^{-\frac12}\,\}\,.$$

For each ω the slabs $T_{\omega,j}$ satisfy a finite-overlap condition; indeed, slabs associated to different elements of Σ_ω are disjoint, and those associated to the same Σ_ω have finite overlap in the x'_ω variable, since the flow restricted to any $\Sigma_{\omega,r}$ is C^1 close to translation. The fixed-time cross sections $T^t_{\omega,j}$ of a slab are thus C^1 close to the translates of the rectangle $R_{\omega,j}$, but their C^2 regularity can be much worse. In particular, the time sections $T^t_{\omega,j}$ are not necessarily comparable to rectangles.

The wave packets $u^{\omega,j}$ that arise in the superposition are normalized wave packets associated to $\Sigma_{\omega,j}$ and $\gamma_{\omega,j}$ as in Definition 8.1, with $u^{\omega,j}$ supported in $T_{\omega,j}$.

We record here some useful facts about the geometry of slabs. We first observe that the results of Section 5 imply a crucial result about the separation of the surfaces $\Sigma_{\theta,r}$ as r varies. Precisely, it follows as a result of the estimates on the null field l following (5.17), and the estimate (5.22), that

$$|dr_\theta - (\theta \cdot dx - dt)| \lesssim \epsilon_1 \,,$$

pointwise uniformly over $[-2,2]\times\mathbb{R}^n$. This implies that

$$|\phi_{\theta,r}(t,x'_\theta) - \phi_{\theta,r'}(t,x'_\theta) - (r-r')| \lesssim \epsilon_1 |\, r-r'\,| \,, \tag{8.6}$$

or that the surfaces $\Sigma_{\theta,r}$ in the foliation essentially maintain a constant separation.

This in turn implies Hölder-$\frac{1}{2}$ bounds on the variation of $d\phi_{\theta,r}$ as r varies. More precisely, from the estimate (8.6) above and the fact that, for each fixed t,

$$\|d^2_{x'_\omega}\phi_{\omega,r}(t,x'_\omega) - d^2_{x'_\omega}\phi_{\omega,r'}(t,x'_\omega)\|_{L^\infty_{x'_\omega}} \lesssim \epsilon_2 + \rho(t)\,,$$

where $\rho(t) = \|d\mathbf{g}(t,\cdot)\|_{C^\delta_x}$, we obtain

$$\|d_{x'_\omega}\phi_{\omega,r}(t,x'_\omega) - d_{x'_\omega}\phi_{\omega,r'}(t,x'_\omega)\|_{L^\infty_{x'_\omega}} \lesssim (\epsilon_2+\rho(t))^{\frac{1}{2}}\,|\, r-r'\,|^{\frac{1}{2}}\,. \tag{8.7}$$

Since $dx_\omega - d\phi_{\omega,r}$ is null, and since $d\mathbf{g} \le \rho(t)$, this also implies Hölder-$\frac{1}{2}$ bounds on $d\phi_{\omega,r}$. To put these in the form we need, suppose that $(t,x)\in\Sigma_{\omega,r}$ and $(t,y)\in\Sigma_{\omega,r'}$, that $|x'_\omega - y'_\omega| \le 2\,(\epsilon_0\lambda)^{-\frac{1}{2}}$, and that $|\, r-r'\,| \le 2\lambda^{-1}$. Then by (5.18), we have

$$|l_\omega(t,x) - l_\omega(t,y)| \lesssim \epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}} + \epsilon_0^{-\frac{1}{2}}\rho(t)\lambda^{-\frac{1}{2}}\,.$$

Since $\dot\gamma_\omega = l_\omega$, and $\|\rho\|_{L^2_t} \lesssim \epsilon_0$, it follows that any geodesic in Σ_ω which intersects a slab $T_{\omega,j}$ must then be contained in the similar slab of half the scale. Alternatively, if $(t,x)\in T_{\omega,j}$, where $T_{\omega,j}$ is of scale λ, then $T_{\omega,j}$ is contained in the slab of scale $\lambda/4$ centered on the geodesic γ_ω through (t,x). This also shows that the slab $T_{\omega,j}$ is comparable to the image of $R_{\omega,j}$ under the geodesic flow tangent to Σ_ω, up to a change in the scale λ.

We now state the counterpart of Corollary 8.3 for superpositions of wave packets.

Proposition 8.4. *Assume that $n \leq 5$. Let*

$$u = \sum_{\omega,j} a_{\omega,j} u^{\omega,j} ,$$

where $u^{\omega,j}$ are normalized wave packets supported in $T_{\omega,j}$. Then

$$\|dS_\lambda u\|_{L^\infty_t L^2_x} \lesssim \Big(\sum_{\omega,j} a^2_{\omega,j} \Big)^{\frac{1}{2}} , \tag{8.8}$$

$$\|\Box_{\mathbf{g}_\lambda} S_\lambda u\|_{L^1_t L^2_x} \lesssim \epsilon_0 \Big(\sum_{\omega,j} a^2_{\omega,j} \Big)^{\frac{1}{2}} . \tag{8.9}$$

Proof. Instead of (8.8) we prove a weaker estimate, namely

$$\|dS_\lambda u\|_{L^2_t L^2_x} \lesssim \Big(\sum_{\omega,j} a^2_{\omega,j} \Big)^{\frac{1}{2}} . \tag{8.10}$$

This suffices, since (8.8) follows from (8.9) and (8.10) by energy estimates.

The result will follow from certain fixed time orthogonality estimates for expressions of the form

$$v = \epsilon_0^{\frac{n-1}{4}} \lambda^{\frac{n-3}{4}} S_\lambda \sum_{\omega,j} T_\lambda \big(\psi^{\omega,j} \delta_{x_\omega - \phi_{\omega,j}(t,x'_\omega)} \big) .$$

We do this in several steps. The size of $\rho(t) = \|d\mathbf{g}(t,\cdot)\|_{C^\delta_x}$ plays an essential role in our arguments. We begin with "good" time sections, namely for which $\rho(t)$ is small.

Lemma 8.5. *Let v be as above, and t such that $\rho(t) = \|d\mathbf{g}(t)\|_{C^\delta_x} \leq \epsilon_0$. Let $0 < \mu < \delta$. Then*

$$\|v(t)\|^2_{L^2_x} \lesssim \sum_{\omega,j} \|\psi^{\omega,j}(t)\|^2_{H_a^{\frac{n-1}{2}+\mu}} , \qquad a = (\epsilon_0 \lambda)^{-\frac{1}{2}} . \tag{8.11}$$

Proof. We begin by noting that it suffices to prove the result for a collection of wave packets whose time t sections intersect a fixed cube Q of size $(\epsilon_0\lambda)^{-\frac{1}{2}}$, since the following argument is easily modified to include appropriate polynomial weights. Consider one such wave packet $u^{\omega,j}$, which is supported in the slab $T_{\omega,j}$. Since $\|d\mathbf{g}(t)\|_{C^\delta_x} \leq \epsilon_0$, it follows that the characteristic surface $\Sigma^t_{\omega,j}$ has the regularity $\|\phi_{\omega,j}(t)\|_{C^{2+\delta}_x} \lesssim \epsilon_0$. Thus the time t section $T^t_{\omega,j}$ of $T_{\omega,j}$ is contained within a rectangle $Q_{\omega,j}$

of size $\lambda^{-1} \times [(\epsilon_0\lambda)^{-\frac{1}{2}}]^{n-1}$. It also follows that the conormal direction to $\Sigma^t_{\omega,j}$ varies at most by $\epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$ within $T^t_{\omega,j}$.

Within $Q_{\omega,j}$ we will work with orthonormal coordinates y_1, y' so that $|dy_1 - (dx_\omega - d\phi_{\omega,j})| \lesssim \epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$, and hence such that $Q_{\omega,j}$ is contained in a rectangle $\{\, |y_1 - c_1| \lesssim \lambda^{-1},\ |y' - c'| \lesssim (\epsilon_0\lambda)^{-\frac{1}{2}}\,\}$. The choice of these coordinates admits the freedom of a $O(\epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}})$ rotation, which we shall exploit shortly. We claim that in such coordinates the following estimate holds:

$$\|T_\lambda(\psi^{\omega,j}(t)\delta_{x_\omega-\phi_{\omega,j}(t,x'_\omega)})\|_{L^2_{y_1}H^{\frac{n-1}{2}+\mu}_{a,y'}} \lesssim \lambda^{\frac{1}{2}}\,\|\psi^{\omega,j}(t)\|_{H^{\frac{n-1}{2}+\mu}_{a,x'_\omega}} \tag{8.12}$$

for $a = (\epsilon_0\lambda)^{-\frac{1}{2}}$. Assume for the moment that this is true, and let us see how to conclude the argument. Because of Proposition 6.1 we know that the angle between two intersecting rectangles $Q_{\omega,j}$ and $Q_{\omega',j'}$ is comparable to the angle between ω and ω'. Actually, Proposition 6.1 applies to intersecting surfaces, however (8.7) shows that the conormals to different elements of Σ^t_ω intersecting the same $Q_{\omega,j}$ are comparable.

Hence, for each $Q_{\omega,j}$ the number of the $Q_{\omega',j'}$'s which intersect it and whose conormal direction is at angle less than $10\,\epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$ is bounded from above by an absolute constant. We now relabel the rectangles as follows. We choose a collection Ω of directions which are $\epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$ separated on the unit sphere, and to each $Q_{\omega,j}$ we associate a direction $\theta \in \Omega$ which is angle at most $\epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$ to $dx_\omega - d\phi_{\omega,j}$. For a fixed $\theta \in \Omega$ we label the associated $Q_{\omega,j}$'s intersecting Q based on their position with respect to the θ direction, in increments of λ^{-1} (the thickness of our rectangles). Thus we may write

$$\{\, Q_{\omega,j} \;:\; Q_{\omega,j}\cap Q \neq \emptyset \,\} = \{\, Q_{\theta,k};\ \theta\in\Omega\,,\ 1 \le k \le \epsilon_0^{-\frac{1}{2}}\lambda^{\frac{1}{2}}\,\}\,.$$

This is somewhat imprecise in that more than one rectangle may have the same label θ, k. However, the above argument shows that the number of such repetitions is bounded from above by an absolute constant, so we shall neglect it.

We also use the same association to relabel the functions as follows,

$$u^{\theta,k} = \epsilon_0^{\frac{n-1}{4}}\lambda^{\frac{n-3}{4}}T_\lambda(\psi^{\omega,j}\delta_{x_\omega-\phi_{\omega,j}})\,,$$

and set

$$v_\theta = S_\lambda \sum_k u^{\theta,k}\,.$$

For each θ we fix orthonormal coordinates y_θ, y'_θ with $y_\theta = \theta \cdot x$. By combining (8.12) with the λ^{-1} separation in k, we conclude that

$$\|v_\theta\|^2_{L^2_{y_\theta} H^{\frac{n-1}{2}+\mu}_{a,y'_\theta}} \lesssim (\epsilon_0\lambda)^{\frac{n-1}{2}} \sum_k \|\psi^{\theta,k}\|^2_{H_a^{\frac{n-1}{2}+\mu}}, \quad a = (\epsilon_0\lambda)^{-\frac{1}{2}}.$$

To sum up the v_θ's we use an orthogonality argument in the frequency variable. We have, with $a = (\epsilon_0\lambda)^{-\frac{1}{2}}$,

$$\begin{aligned}
\|v\|^2_{L^2_x} &= \|S_\lambda(\xi) \sum_{\theta\in\Omega} \widehat{v}_\theta\|^2_{L^2_\xi} \\
&= \int S^2_\lambda(\xi) \Big| \sum_{\theta\in\Omega} \widehat{v}_\theta \big(1 + a|\xi'_\theta|\big)^{\frac{n-1}{2}+\mu} \big(1 + a|\xi'_\theta|\big)^{-\frac{n-1}{2}-\mu} \Big|^2 d\xi \\
&\lesssim \int \Big(\sum_{\theta\in\Omega} \big|\widehat{v}_\theta \big(1 + a|\xi'_\theta|\big)^{\frac{n-1}{2}+\mu}\big|^2 \Big) S^2_\lambda(\xi) \sum_{\theta\in\Omega} \big(1 + a|\xi'_\theta|\big)^{-(n-1)-2\mu} d\xi .
\end{aligned}$$

However,

$$\big\|\widehat{v}_\theta \big(1 + a|\xi'_\theta|\big)^{\frac{n-1}{2}+\mu}\|^2_{L^2_\xi} = a^{\frac{n-1}{2}} \|v_\theta\|^2_{L^2_{y_\theta} H^{\frac{n-1}{2}+\mu}_{a,y'_\theta}} \lesssim \sum_k \|\psi^{\theta,k}\|^2_{H_a^{\frac{n-1}{2}+\mu}},$$

therefore we have

$$\|v\|^2_{L^2_x} \lesssim \sum_{\theta,k} \|\psi^{\theta,k}\|^2_{H_a^{\frac{n-1}{2}+\mu}} \sup_\xi \Big[S^2_\lambda(\xi) \sum_{\theta\in\Omega} \big(1 + a|\xi'_\theta|\big)^{-(n-1)-2\mu} \Big].$$

To conclude the argument it suffices to verify that the above supremum is bounded by some absolute constant. This is true because each term in the sum is essentially concentrated within an $(\epsilon_0\lambda)^{\frac{1}{2}}$ neighborhood of the line with direction θ. At frequency λ these regions are disjoint due to the $\epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$ angular separation between different directions. Precisely,

$$S^2_\lambda(\xi) \sum_{\theta\in\Omega} \big(1 + a|\xi'_\theta|\big)^{-(n-1)-2\mu} \lesssim \sum_{k\in\mathbb{Z}^{n-1}} \big(1 + |k|\big)^{-(n-1)-2\mu} \lesssim 1 .$$

It remains to prove (8.12) in the coordinates $(y_1, y') = (y_\theta, y'_\theta)$. We begin by noting that it suffices to prove the bound using the coordinates

$$y_1 = y_\theta = \theta \cdot x, \quad y' = x'_\omega,$$

since these have the same level sets $y_\theta = c$. We let α' be the vector perpendicular to ω such that $\omega - \alpha' = |\omega - \alpha'|\,\theta$. Then, since the angle of $\omega - d_x\phi_{\omega,j}$ to θ is bounded by $\epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$, and since both $d_x\phi_{\omega,j}$ and α' are perpendicular to ω, it follows that

$$|\,\alpha' - d_x\phi_{\omega,j}| \lesssim \epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}},$$

uniformly over $Q_{\omega,j}$. We now write

$$\begin{aligned} x_\omega - \phi_{\omega,j}(t,x'_\omega) &= |\,\omega-\alpha'|\,y_\theta - \big(\phi_{\omega,j}(t,x'_\omega) - \alpha'\cdot x\big) \\ &\equiv |\omega-\alpha'|\,\big(\,y_1 - \phi(t,y')\,\big)\,, \end{aligned}$$

to see that we are reduced to establishing the bound

$$\|T_\lambda(\,\psi(y')\delta_{y_1-\phi(y')}\,)\|_{L^2_{y_1}H^{\frac{n-1}{2}+\mu}_{y',a}} \lesssim \lambda^{\frac12}\|\psi\|_{H_a^{\frac{n-1}{2}+\mu}}\,, \quad a = (\epsilon_0\lambda)^{-\frac12}\,,$$

where within the support of ψ we have $\phi\in C^{2+\delta}$, $|d\phi|\lesssim \epsilon_0^{\frac12}\lambda^{-\frac12}$. We can also subtract a constant from ϕ to insure that $|\phi|\lesssim\lambda^{-1}$, which implies that $\lambda\phi\in C^{2+\delta}_a$, with a as above.

The Fourier transform of $\psi(y')\,\delta_{y_1-\phi(y')}$ in the y_1 direction equals $\psi\,e^{i\eta\phi}$, therefore it suffices to show that

$$\|\psi\,e^{i\eta\phi}\|_{H^{\frac{n-1}{2}+\mu}_{y',a}} \lesssim \big(1+\lambda^{-1}|\eta|\big)^N\,\|\psi\|_{H_a^{\frac{n-1}{2}+\mu}}\,.$$

After rescaling η by λ and y' by a, we are reduced to verifying that

$$\|\psi e^{i\eta\phi}\|_{H^{\frac{n-1}{2}+\mu}} \lesssim \big(1+|\eta|\big)^N\,\|\psi\|_{H^{\frac{n-1}{2}+\mu}}\,, \quad \|\phi\|_{C^{2+\delta}}\lesssim 1\,.$$

But

$$\|e^{i\eta\phi}\|_{C^{2+\delta}} \lesssim \big(1+|\eta|\big)^{2+\delta}\,,$$

therefore the conclusion follows from the multiplicative estimate

$$H^{\frac{n-1}{2}+\mu}\cdot C^{2+\delta}\subset H^{\frac{n-1}{2}+\mu}\,, \quad 0<\mu<\delta\,.$$

Note that this requires $n\le 5$. □

Our next step is to obtain a fixed time estimate for values of t at which $d\mathbf{g}$ is large. The following estimate is a simple variation of the preceding argument, which unfortunately is useful only when $\frac{n-1}{2}\le 2$, that is, for dimensions $n\le 5$.

Lemma 8.6. *Let v be as above, and t such that $\rho(t)=\|d\mathbf{g}(t)\|_{C^\delta_x}\ge\epsilon_0$. Let $0<\mu<\delta$. Then*

$$\|v(t)\|^2_{L^2_x} \lesssim \epsilon_0^{-\frac{n-1}{2}}\rho(t)^{\frac{n-1}{2}}\sum_{\omega,j}\|\psi^{\omega,j}(t)\|^2_{H_a^{\frac{n-1}{2}+\mu}}\,, \quad a=(\epsilon_0\lambda)^{-\frac12}\,. \tag{8.13}$$

Proof. If $\rho(t)\ge\lambda$, then the above bound follows from the Schwartz inequality by noting that, for each fixed ω,

$$\Big\|S_\lambda\sum_j T_\lambda(\,\psi^{\omega,j}\delta_{x_\omega-\phi_{\omega,j}}\,)\Big\|^2_{L^2_x} \le \lambda\sum_j\|\psi^{\omega,j}\|^2_{L^2_{x'_\omega}}\,.$$

This estimate in turn is a simple consequence of the fact that wave packets on the same $\Sigma_{\omega,j}$ have finite overlap, together with the fact that the $\Sigma_{\omega,j}$ are small C^1 perturbations of flat surfaces, with uniform separation of order λ^{-1}. We thus subsequently assume that $\rho(t) \leq \lambda$.

By Proposition 5.2, we have $\|d_x^2\phi_{\omega,j}\|_{C_x^\delta} \lesssim \rho(t)$. It follows that, within a cube of size r, the conormal direction to $\Sigma_{\omega,j}^t$ varies by at most $\rho(t)r$. At frequency scale λ this leads to a frequency spread of $\lambda\rho(t)r$, which is consistent with a decomposition into cubes of size r provided that

$$r = \rho(t)^{-\frac{1}{2}}\lambda^{-\frac{1}{2}}.$$

We thus take a partition of unity on each $\Sigma_{\omega,j}^t$ over cubes of sidelength r in x'_ω to split

$$\psi^{\omega,j} = \sum_m \chi_m \psi^{\omega,j}.$$

Then each term $T_\lambda(\chi_m\psi^{\omega,j}\delta_{x_\omega - \phi_{\omega,j}})$ is supported in a rectangle $Q_{\omega,j,m}$ of dimensions $\lambda^{-1}\times r^{n-1}$.

The analogue of (8.12) is the estimate, for $a = (\epsilon_0\lambda)^{-\frac{1}{2}}$ and r as above,

$$\|T_\lambda(\chi_m\psi^{\omega,j}\delta_{x_\omega-\phi_{\omega,j}})\|_{L^2_{y_\theta} H^{\frac{n-1}{2}+\mu}_{r,y'_\theta}} \lesssim \lambda^{\frac{1}{2}}\|\psi^{\omega,j}\|_{H_a^{\frac{n-1}{2}+\mu}}, \tag{8.14}$$

where θ is such that, uniformly over the support of χ_m,

$$|\,\theta - (dx_\omega - d_x\phi_{\omega,j})\,| \lesssim \rho(t)r = (\lambda r)^{-1}.$$

The proof of (8.14) is similar to that of (8.12). Indeed, on the one hand

$$\|\chi_m\psi^{\omega,j}\|_{H_r^{\frac{n-1}{2}+\mu}} \lesssim \|\psi^{\omega,j}\|_{H_a^{\frac{n-1}{2}+\mu}},$$

since $r \leq a$. On the other hand, we may work in the coordinates $y_1 = y_\theta$, $y' = x'_\omega$, in which case we may assume that $|d_x\phi_{\omega,j}| \lesssim (\lambda r)^{-1}$ within the r cube Q_r which contains the support of $\chi_m\psi^{\omega,j}$, and after subtracting a constant that $|\phi_{\omega,j}| \lesssim \lambda^{-1}$. Combined with the relation $|\phi_{\omega,j}|_{C_x^{2+\delta}} \lesssim \rho(t) = r^{-2}\lambda^{-1}$ this implies

$$\lambda\|\phi\|_{C_r^{2+\delta}(Q_r)} \lesssim 1,$$

and the proof of (8.14) proceeds as before.

At this point we repeat the argument in the proof of Lemma 8.5 to obtain a square summability result within a cube Q_r of size r. For the most part this requires simply replacing ϵ_0 by $\rho(t)$ in the previous argument. We take the collection $\Omega_{\rho(t)}$ of directions to be $\rho(t)^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$ separated, and consider the collection

$$\{Q_{\theta,k} \;:\; \theta\in\Omega_{\rho(t)}, \quad 1 \leq k \leq \rho(t)^{-\frac{1}{2}}\lambda^{\frac{1}{2}}\}$$

of rectangles of size $\lambda^{-1} \times r^{n-1}$ contained in Q_r and with dimension λ^{-1} in the direction θ. To each $Q_{\theta,k}$ we want to associate the truncated wave packet sections which are supported within it. However, the angles of the $Q_{\theta,k}$'s are separated on the $\rho(t)^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$ scale, while the angles of wave packets are only separated on the $\epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$. The estimate (8.7) shows that if two truncated wave packet sections are associated to the same $Q_{\theta,k}$, then their initial angles ω and ω' must satisfy

$$|\,\omega - \omega'| \lesssim \rho(t)^{\frac{1}{2}}\, \lambda^{-\frac{1}{2}}\,,$$

which implies that each $Q_{\theta,k}$ supports approximately $\rho(t)^{\frac{n-1}{2}}\epsilon_0^{-\frac{n-1}{2}}$ truncated wave packet sections. Hence, we can relabel

$$\begin{aligned}\{\,\epsilon_0^{\frac{n-1}{4}}\lambda^{\frac{n-3}{4}}T_\lambda(\chi_m\psi^{\omega,j}\delta_{x_\omega-\phi_{\omega,j}})\,:\,Q_{\omega,j,m}\cap Q_r\neq\emptyset\,\} \;=\\ \{u^{\theta,k,l}\,:\,\theta\in\Omega_\rho(t)\,,\ 1\le k\le\rho(t)^{-\frac{1}{2}}\lambda^{\frac{1}{2}}\,,\ 1\le l\le\rho(t)^{\frac{n-1}{2}}\epsilon_0^{-\frac{n-1}{2}}\,\}\,.\end{aligned}$$

Denote

$$u^{\theta,k} = \sum_l u^{\theta,k,l}\,,\quad v_\theta = S_\lambda \sum_k u^{\theta,k}\,,\quad v_{Q_r} = \sum_\theta v_\theta\,.$$

By (8.14) we have

$$\|u^{\theta,k,l}\|_{L^2_{y_\theta}H^{\frac{n-1}{2}+\mu}_{r,y'_\theta}} \lesssim \epsilon_0^{\frac{n-1}{4}}\lambda^{\frac{n-1}{4}}\|\psi^{\omega,j}\|_{H_a^{\frac{n-1}{2}+\mu}}\,,\quad a=(\epsilon_0\lambda)^{-\frac{1}{2}}\,.$$

Summing over l and using the Cauchy-Schwartz inequality yields

$$\|u^{\theta,k}\|^2_{L^2_{y_\theta}H^{\frac{n-1}{2}+\mu}_{r,y'_\theta}} \lesssim \rho(t)^{\frac{n-1}{2}}\epsilon_0^{-\frac{n-1}{2}}\sum_l \|u^{\theta,k,l}\|^2_{L^2_{y_\theta}H^{\frac{n-1}{2}+\mu}_{r,y'_\theta}}\,.$$

The λ^{-1} separation in the y_θ direction yields

$$\|v_\theta\|^2_{L^2_{y_\theta}H^{\frac{n-1}{2}+\mu}_{r,y'_\theta}} \lesssim \sum_k \|u^{\theta,k}\|^2_{L^2_{y_\theta}H^{\frac{n-1}{2}+\mu}_{r,y'_\theta}}\,.$$

Finally, repeating the orthogonality argument in frequency with respect to θ in the proof of Lemma 8.5 yields

$$\|v_{Q_r}\|^2_{L^2_x} \lesssim \rho(t)^{-\frac{n-1}{2}}\lambda^{-\frac{n-1}{2}}\sum_\theta \|v_\theta\|^2_{L^2_{y_\theta}H^{\frac{n-1}{2}+\mu}_{r,y'_\theta}}\,.$$

We now combine the last four relations to obtain

$$\|v_{Q_r}\|^2_{L^2_x} \lesssim \sum_{\substack{\omega,j\\ Q_{\omega,j}\cap Q_r\neq\emptyset}} \|\psi^{\omega,j}\|^2_{H_a^{\frac{n-1}{2}+\mu}}\,,\quad a=(\epsilon_0\lambda)^{-\frac{1}{2}}\,.$$

The above bound also holds with appropriate polynomially growing weights, so that the v_{Q_r} may be considered orthogonal for different Q_r.

When summing over the different Q_r, each wave packet gets counted $\rho(t)^{\frac{n-1}{2}}\epsilon_0^{-\frac{n-1}{2}}$ times, therefore we obtain

$$\|v(t)\|_{L_x^2}^2 \lesssim \rho(t)^{\frac{n-1}{2}}\epsilon_0^{-\frac{n-1}{2}} \sum_{\omega,j} \|\psi^{\omega,j}\|^2_{H_a^{\frac{n-1}{2}+\mu}}, \quad a = (\epsilon_0\lambda)^{-\frac{1}{2}}.$$

□

We can now conclude the proof of Proposition 8.4. We begin by establishing (8.10). If we apply (8.11) and (8.13) with S_λ replaced by $\lambda^{-1}d_x S_\lambda$, and $\psi^{\omega,j} = a_{\omega,j}w^{\omega,j}$, we get

$$\|d_x S_\lambda u(t)\|_{L_x^2}^2 \lesssim \left(1 + \rho(t)^{\frac{n-1}{2}}\epsilon_0^{-\frac{n-1}{2}}\right) \sum_{\omega,j} a_{\omega,j}^2 .$$

Since $\|\rho\|_{L_t^2} \lesssim \epsilon_0$ and $n \leq 5$, this gives (8.10) for d replaced by d_x . To handle the time derivative of u, we note that we may write

$$\partial_t w = \dot\gamma(t)\,(e_0\lambda)^{\frac{1}{2}}\,\tilde w\,, \quad \partial_t\delta(x_\omega - \phi_{\omega,j}) = \nu_0\,\delta^{(1)}(x_\omega - \phi_{\omega,j})\,.$$

The first term is handled as above since $\dot\gamma(t) \in L_t^\infty$ and $(\epsilon_0\lambda)^{\frac{1}{2}} \leq \lambda$. The second term will be handled below, noting that the term ν_0 is harmless.

To prove (8.8) we use the representation in (8.1). On the one hand, by (8.10) we have

$$\|L(d\mathbf{g}, d\tilde S_\lambda \tilde u)\|_{L_t^1L_x^2} \lesssim \|d\mathbf{g}\|_{L_t^2L_x^\infty}\|d\tilde S_\lambda\tilde u\|_{L_t^2L_x^2} \lesssim \epsilon_1\,.$$

On the other hand, we can apply (8.11) and (8.13) for the remaining three right hand side terms in (8.1). If we set

$$f = \epsilon_0^{\frac{n-1}{2}}\lambda^{\frac{n-7}{4}} S_\lambda T_\lambda \sum_{\omega,j} a_{\omega,j} \sum_{m=0,1,2} \psi_m^{\omega,j}\,\delta^{(m)}_{x_\omega-\phi_{\omega,j}}$$

then

$$\|f(t)\|_{L_x^2}^2 \lesssim \left(1 + \rho(t)^{\frac{n-1}{2}}\epsilon_0^{-\frac{n-1}{2}}\right) \sum_{\omega,j} a_{\omega,j}^2 \sum_{m=0,1,2} \lambda^{m-1}\|\psi_m^{\omega,j}(t)\|^2_{H_a^{\frac{n-1}{2}+\mu}},$$

which yields

$$\|f\|_{L_t^1L_x^2}^2 \lesssim \left(\int 1 + \rho(t)^{\frac{n-1}{2}}\epsilon_0^{-\frac{n-1}{2}}\,dt\right) \times \left(\int \sum_{\omega,j} a_{\omega,j}^2 \sum_{m=0,1,2} \lambda^{2(m-1)}\|\psi_m^{\omega,j}(t)\|^2_{H_a^{\frac{n-1}{2}+\mu}}\,dt\right).$$

By hypothesis, $\|\rho(t)\|_{L^2_t} \lesssim \epsilon_0$. Combining this with (8.2) we obtain

$$\|f\|_{L^1_t L^2_x} \lesssim \epsilon_0 \Big(\sum_{\omega,j} a^2_{\omega,j} \Big)^{\frac{1}{2}},$$

which concludes the proof of (8.9). $\square$

8.4. **Matching the initial data.** In order to complete the construction of an approximate solution for the initial value problem, it remains to verify that the approximate solutions which are superpositions of wave packets can be chosen so that they match the initial data at time $t = -2$. Since the metric $\mathbf{g}$ equals the Minkowski metric for times t in a neighborhood of -2, it actually suffices to work with wave packets near $t = -2$ for the Minkowski wave operator, since the definition of a wave packet, together with the regularity of the $\Sigma_{\omega,r}$, show that these may be continued to wave packets for $\mathbf{g}$ up to time $t = 2$.

Proposition 8.7. *Given any initial data $(u_0, u_1) \in H^1 \times L^2$, there exists a function of the form*

$$u = \sum_{\omega,j} a_{\omega,j} u^{\omega,j},$$

where the functions $u_{\omega,j}$ are normalized wave packets, such that

$$S_\lambda u(-2) = S_\lambda u_0, \qquad \partial_t S_\lambda u(-2) = S_\lambda u_1.$$

Furthermore

$$\sum_{\omega,j} a^2_{\omega,j} \lesssim \big(\|d_x u_0\|^2_{L^2_x} + \|u_1\|^2_{L^2_x} \big).$$

Proof. We consider a maximal collection of unit vectors Ω in $\mathbb{R}^n$ with spacing $\epsilon_0^{\frac{1}{2}} \lambda^{-\frac{1}{2}}$. Then Ω contains about $\epsilon_0^{-\frac{n-1}{2}} \lambda^{\frac{n-1}{2}}$ elements. Without loss of generality, we may assume that (u_0, u_1) have Fourier transform supported in the range $\lambda/4 \le |\xi| \le 4\lambda$. Via a partition of unity, we may decompose

$$u_0 = \sum_{\omega \in \Omega} u_0^\omega, \qquad u_1 = \sum_{\omega \in \Omega} u_1^\omega,$$

where the Fourier transform of u_j^ω is supported in a $(4\lambda) \times (4\epsilon_0^{\frac{1}{2}} \lambda^{\frac{1}{2}})^{n-1}$ rectangle, with the long direction parallel to ω. The approximate solution to the Minkowski equation with Cauchy data (u_0^ω, u_1^ω) is the function

$$u^\omega = \frac{1}{2} \big[u_0^\omega(x + t\omega) + u_0^\omega(x - t\omega) + \\ \big((\omega \cdot D_x)^{-1} u_1^\omega\big)(x + t\omega) - \big((\omega \cdot D_x)^{-1} u_1^\omega\big)(x - t\omega) \big],$$

which involves wave packets in both the ω and $-\omega$ directions. We will show that we may write the function $u_0^\omega(x - t\omega)$ as an appropriate sum of normalized wave packets in the ω direction, which implies the desired result.

Because of the support condition on $\widehat{u_0^\omega}$, we may write $u_0^\omega = T_\lambda \tilde{u}_0^\omega$, where $\tilde{u}_0^\omega$ is of comparable L^2 norm. We extend the Fourier transform of $\tilde{u}_0^\omega$ as a periodic function of period $16\pi\lambda\,\omega$, which we also denote by $\tilde{u}_0^\omega$, so that $S_\lambda u_0^\omega = S_\lambda T_\lambda \tilde{u}_0^\omega$. Then $\tilde{u}_0^\omega$ has the form

$$\tilde{u}_0^\omega = \sum_{k\in\mathbb{Z}} w^{\omega,k}(x'_\omega)\,\delta_{x_\omega - \frac{k}{8\lambda}}\,,$$

where the functions $w^{\omega,k}(x'_\omega)$ have Fourier transform supported in the region $|\xi'_\omega| \leq 4(\epsilon_0\lambda)^{\frac{1}{2}}$, and satisfy the Plancherel identity

$$\sum_k \|w^{\omega,k}\|^2_{L^2_{x'_\omega}} \approx \lambda^{-1}\|u_0^\omega\|^2_{L^2_x}\,.$$

We now take a partition of unity on the $(\epsilon_0\lambda)^{-\frac{1}{2}}$ scale with respect to the transversal variables x'_ω,

$$1 = \sum_l w_l(x'_\omega)\,,$$

and set

$$w^{\omega,k,l}(x'_\omega) = w^{\omega,k}(x'_\omega)\,w_l(x'_\omega)\,.$$

Then

$$\sum_l \|w^{\omega,k,l}\|^2_{L^2_{x'_\omega}} \lesssim \|w^{\omega,k}\|^2_{L^2_{x'_\omega}}\,.$$

As a result of the support property of the Fourier transform of $w^{\omega,k}$, we also obtain

$$\sum_l \|\partial^\alpha_{x'_\omega} w^{\omega,k,l}\|^2_{L^2_{x'_\omega}} \leq c_\alpha(\epsilon_0\lambda)^{|\alpha|}\|w^{\omega,k}\|^2_{L^2_{x'_\omega}}\,.$$

Consolidating the indices k, l into a single index j, we obtain the desired decomposition

$$S_\lambda u_0^\omega = S_\lambda T_\lambda \sum_j w^{\omega,j}(x'_\omega)\,\delta_{x_\omega - \frac{k(j)}{8\lambda}}\,.$$

$\square$

9. Overlap estimates

An essential role in the proof of Proposition 10.1 is played by an upper bound estimate on the number of λ-slabs that contain two given points in space-time. Given two points P_1, P_2 we denote by $N_\lambda(P_1, P_2)$ the number of λ slabs containing both P_1 and P_2. In order to obtain sharp bounds on $N_\lambda(P_1, P_2)$ we introduce some additional notation.

We set $P_1 = (t_1, x_1)$, $P_2 = (t_2, x_2)$. Without any restriction in generality we assume that $t_1 < t_2$. We denote by K_{P_1} the forward light cone starting at P_1. The analysis in Section 6 shows that K_{P_1} is smooth; in addition, by Proposition 6.2, its time t_2 section $K_{P_1}^{t_2}$ is $o(|t_2 - t_1|)$-close in the C^1 topology to the sphere of radius $t_2 - t_1$ centered at x_1. Changing notation slightly from Section 6, for each θ we let γ_θ be the null geodesic contained in Σ_θ such that $\gamma_\theta(t_1) = x_1$, and set $Q_\theta = \gamma_\theta(t_2) \in K_{P_1}^{t_2}$.

Let d denote the distance of x_2 to the t_2 slice of the light cone K_{P_1} centered at P_1,

$$d = \inf_{\omega \in S^{n-1}} \left| x_2 - \gamma_\omega(t_2) \right| .$$

We use the functions $r_\theta(t, x)$ defined in Section 4.1, whose level sets yield the Σ_θ foliation,

$$\Sigma_{\theta,r} = \{ \, (t,x) \, : \, r_\theta(t,x) = r \, \} .$$

If we work in the null frame $\{l_\theta, \underline{l}_\theta, e_{a,\theta}\}$ associated to $\Sigma_{\theta,r}$, then $l_\theta r_\theta = e_{a,\theta} r_\theta = 0$, while

$$l_\theta(\underline{l}_\theta r_\theta) = [l_\theta, \underline{l}_\theta] r_\theta = \frac{1}{2} \langle [l_\theta, \underline{l}_\theta], l_\theta \rangle \, \underline{l}_\theta r_\theta = (l_\theta \ln \sigma_\theta) \, \underline{l}_\theta r_\theta \, ,$$

where the connection coefficient σ_θ, introduced in Section 5.5, measures the infinitesimal separation between neighboring surfaces in the Σ_θ foliation. As the functions $\underline{l}_\theta r_\theta$ and σ_θ agree at time -2, it follows that

$$dr_\theta(t,x) \cdot v = \sigma_\theta(t,x) \, \langle l_\theta(t,x), v \rangle_{\mathbf{g}} \tag{9.1}$$

where d denotes the differential in the (t, x) variables. By (6.5), we have

$$\sigma_\theta(t,x) = 1 + o(1) \, . \tag{9.2}$$

We next introduce the parameter m defined by

$$m = \max_{\omega \in S^{n-1}} r_\omega(P_2) - r_\omega(P_1) \, ,$$

and fix some ω_0 at which the maximum occurs. As we prove next, m plays the role of a signed distance to the cone. Since $K_{P_1}^{t_2}$ is suitably close to a sphere, its interior and exterior are well defined.

Lemma 9.1. *The parameter m is negative if P_2 is inside the light cone K_{P_1}, and positive if P_2 is exterior to K_{P_1}. Furthermore, $|m| \approx d$.*

Proof. (a.) If P_2 is on the cone K_{P_1}, then $x_2 = Q_\theta$ for some $\theta \in S^{n-1}$, which shows that $r_\theta(P_2) = r_\theta(P_1)$. For any other $\omega \in S^{n-1}$, we use (9.1) to write

$$r_\omega(P_2) - r_\omega(P_1) = \int_{t_1}^{t_2} \sigma_\omega(s, \gamma_\theta(s)) \, \langle l_\omega, l_\theta \rangle_{\mathbf{g}} \, ds \leq 0 \, .$$

This shows that $m = 0$. In addition, the functions r_θ are Lipschitz continuous in x with Lipschitz norm $1 + o(1)$. This implies that in general

$$|m| \leq (1 + o(1))\, d \, .$$

(b.) Suppose that P_2 is outside K_{P_1}. Choose θ which minimizes the euclidean distance $|\overrightarrow{Q_\theta x_2}|$ of x_2 to points on $K_{P_1}^{t_2}$. The outer normal direction to $K_{P_1}^{t_2}$ at Q_θ is $o(1)$-close to θ, therefore $\overrightarrow{Q_\theta x_2} = (\theta + o(1))|\overrightarrow{Q_\theta x_2}|$. Set

$$\mu(s) = s\, x_2 + (1-s)\gamma_\theta(t_2) \, .$$

Since $r_\theta(P_1) = r_\theta(t_2, Q_\theta)$, by (9.1) we compute

$$\begin{aligned} r_\theta(P_2) - r_\theta(P_1) &= \int_0^1 \sigma_\omega(t_2, \mu(s)) \, \langle \dot\mu(s), l_\theta \rangle ds \\ &= (1 + o(1)) \, |\overrightarrow{Q_\theta x_2}| \approx d \, . \end{aligned}$$

This implies the missing inequality $m \gtrsim d$.

(c.) Suppose that P_2 is inside K_{P_1}. Then $d \leq (t_2 - t_1) + o(t_2 - t_1)$. We choose θ as before, but now we have $\overrightarrow{Q_\theta x_2} = -(\theta + o(1)) \, |Q_\theta x_2|$. Given $\omega \in S^{n-1}$, and μ as before, we write

$$\begin{aligned} r_\omega(P_2) - r_\omega(P_1) &= \int_{t_1}^{t_2} \sigma_\omega(s, \gamma_\theta(s)) \, \langle l_\omega, l_\theta \rangle_{\mathbf{g}} \, ds \\ &\qquad + \int_0^1 \sigma_\omega(t_2, \mu(s)) \, \langle \dot\mu(s), l_\omega \rangle \, ds \\ &= -\Big(\frac{1}{2} + o(1)\Big)(t_2 - t_1)|\theta - \omega|^2 - d\,(\theta \cdot \omega + o(1)) \\ &\leq -d + o(d) \, , \end{aligned}$$

where we estimated the first integral using (6.1). ∎

We are now ready to state our main result:

Proposition 9.2. *For all points $P_1 = (t_1, x_1)$ and $P_2 = (t_2, x_2)$ in space-time, and $\epsilon_0 \lambda \geq 1$, the number $N_\lambda(P_1, P_2)$ of slabs of scale λ that contain both P_1 and P_2 satisfies the bound*

$$N_\lambda(P_1,P_2) \lesssim \begin{cases} \epsilon_0^{-\frac{n-1}{2}} \lambda^{\frac{n-1}{2}} (\,1+\lambda d\,)^{\frac{n-3}{2}}\,(1+\lambda\,|t_2-t_1|\,)^{-\frac{n-1}{2}}\,, & m \in I_1 \\ \epsilon_0^{-\frac{n-1}{2}} \lambda^{\frac{n-1}{2}} (\,1+\lambda d\,)^{-1}\,, & m \in I_2 \\ 0\,, & m \notin I_1 \cup I_2 \end{cases}$$

where

$$I_1 \;=\; \big\{ -4\lambda^{-1} \leq m \leq \min\big(\,2\,|t_2-t_1|\,,\, C\,\epsilon_0^{-1}\lambda^{-1}|t_2-t_1|^{-1}\,\big)\big\}\,,$$

$$I_2 \;=\; \big\{\, 2\,|t_2-t_1| \leq m \leq C\,\epsilon_0^{-\frac{1}{2}}\lambda^{-\frac{1}{2}}\big\}\,,$$

and C is a large constant.

Proof. The above result coincides with the estimate that holds in the constant coefficient case. The challenge in the proof is that the surfaces we work with are not C^2 close to their constant coefficient analogues, so we need to work only with the aspects of the geometry which we control.

By the comments following (8.7), if a slab in direction ω contains both P_1 and P_2, then the slab centered on γ_ω of scale $\lambda/4$ must also contain P_2. Thus we seek to bound the number of $\epsilon_0^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$ balls needed to cover the set $A_\lambda \subseteq S^{n-1}$ defined by

$$A_\lambda = \big\{\omega \,:\, |r_\omega(t_2,x_2) - r_\omega(t_1,x_1)| \leq \lambda^{-1},\; |\gamma_\omega(t_2) - x_2| \leq (\epsilon_0\lambda)^{-\frac{1}{2}}\big\}\,.$$

For each $\theta \in A_\lambda$, we may choose $Q_2 \in \mathbb{R}^n$ with $|Q_2 - x_2| \leq 2\,\lambda^{-1}$, such that $r_\theta(t_2,Q_2) - r_\theta(t_1,x_1) = 0$, and observe that

$$A_\lambda \subset \big\{\omega : |r_\omega(t_2,Q_2) - r_\omega(t_1,x_1)| \leq 3\lambda^{-1}\,, |\gamma_\omega(t_2) - Q_2| \leq 3(\epsilon_0\lambda)^{-\frac{1}{2}}\big\}\,.$$

Continuing, we let $\mu_\theta(s)$ denote the C^1 path for $s \in [0,1]$ contained in $\Sigma^t_{\theta,r}$, $r = r_\theta(t_1,x_1)$, which goes from $Q_\theta = \gamma_\theta(t_2)$ to Q_2, and which is obtained by projecting the straight line segment $\overrightarrow{Q_\theta Q_2}$ onto the surface $\Sigma^t_{\theta,r}$ along the direction θ. We note that

$$\text{(9.3)} \qquad \dot\mu_\theta - \overrightarrow{Q_\theta Q_2} = o\big(\,|Q_\theta - Q_2|\,\big)\,, \qquad \theta \cdot \overrightarrow{Q_\theta Q_2} = o\big(\,|Q_\theta - Q_2|\,\big)\,.$$

We now write

$$r_\omega(t_2,Q_2)-r_\omega(t_1,x_1)=$$
$$\int_{t_1}^{t_2}\sigma_\omega(s,\gamma_\theta(s))\,\langle l_\omega,l_\theta\rangle_{\mathbf g}\,ds+\int_0^1\sigma_\omega(t_2,\mu_\theta(s))\,\langle l_\omega,\dot\mu_\theta\rangle_{\mathbf g}\,ds\,.$$

By (6.2) and (9.2) applied to the first integral, and by (6.1),(9.3), and $\langle l_\theta,\dot\mu_\theta\rangle_{\mathbf g}=0$ applied to the second integral, we then have

$$\text{(9.4)}\quad r_\omega(t_2,Q_2)-r_\omega(t_1,x_1)=-\frac{1}{2}(t_2-t_1)\,|\omega-\theta|^2+(\omega-\theta)\cdot\overrightarrow{Q_\theta Q_2}$$
$$+\,o\big(\,|t_2-t_1|\,|\omega-\theta|^2\,\big)+o\big(\,|\omega-\theta|\,|Q_\theta-Q_2|\,\big)\,.$$

We consider several cases with respect to the values of m and t_2-t_1.

Case 1: $|m|<-4\lambda^{-1}$**.** In this case $N_\lambda(P_1,P_2)=0$.

Case 2: $|m|<4\lambda^{-1}$**,** $(t_2-t_1)<2\lambda^{-1}$**.** Here we use the trivial bound

$$N_\lambda(P_1,P_2)\lesssim\epsilon_0^{-\frac{n-1}{2}}\lambda^{\frac{n-1}{2}}\,.$$

Case 3: $|m|<4\lambda^{-1}$**,** $(t_2-t_1)\ge 2\lambda^{-1}$**.** Then by (9.4) it follows that

$$A_\lambda\subset\big\{\,\omega\,:\,|\omega-\omega_0|\le C\,\lambda^{-\frac12}\,(t_2-t_1)^{-\frac12}\,\big\}\,,$$

which is covered by $\approx C^{n-1}\epsilon_0^{-\frac{n-1}{2}}(t_2-t_1)^{-\frac{n-1}{2}}$ balls of radius $\epsilon_0^{\frac12}\lambda^{-\frac12}$.

Case 4: $4\lambda^{-1}\le m\le 2\,(t_2-t_1)\,$. Then for $\theta\in A_\lambda$, if Q_2 is chosen as above depending on θ, it follows that

$$\max_\omega\, r_\omega(t_2,Q_2)-r_\omega(t_1,x_1)\approx m\,,$$

which by (9.4) implies that

$$m\approx|Q_\theta-Q_2|^2\,(t_2-t_1)^{-1}\lesssim\epsilon_0^{-1}\lambda^{-1}(t_2-t_1)^{-1}\,,$$

where we are assuming $N_\lambda(P_1,P_2)\ne 0$ to conclude that $|Q_\theta-Q_2|\le 2\,\epsilon_0^{-\frac12}\,\lambda^{-\frac12}$. Since the maximum, up to λ^{-1}, occurs at $\omega=\omega_0$, it follows that, for all $\theta\in A_\lambda$,

$$|\theta-\omega_0|\approx|Q_2-Q_\theta|\,(t_2-t_1)^{-1}\approx m^{\frac12}\,(t_2-t_1)^{-\frac12}\,.$$

We can thus cover A_λ by $\approx c^{1-n}$ of balls of radius $\rho=c\,m^{\frac12}\,(t_2-t_1)^{-\frac12}$. Consequently, we are reduced to showing that the intersection of A_λ with a ball $B_\rho(\theta)$ of radius ρ, centered on $\theta\in A_\lambda$, can be covered by the indicated number of balls of radius $\epsilon_0^{\frac12}\lambda^{-\frac12}$, for sufficiently small c.

If now θ' is any point in $B_\rho(\theta)$ such that $r_{\theta'}(t_1,x_1) - r_{\theta'}(t_2,Q_2) = 0$, then applying (9.4) to $\theta = \theta'$ yields that, for $\omega \in B_\rho(\theta)$,

$$\begin{aligned} r_\omega(t_2,Q_2) - r_\omega(t_1,x_1) &= (\omega-\theta')\cdot\overrightarrow{Q_{\theta'}Q_2} + o\big(\,|\omega-\theta'|\,|Q_{\theta'}-Q_2|\,\big) \\ &= (\omega-\theta')\cdot\overrightarrow{Q_\theta Q_2} + o\big(\,|\omega-\theta'|\,|Q_\theta-Q_2|\,\big)\,, \end{aligned}$$

where we use (6.3) in the second step. By the second part of (9.3), it follows that the set of $\theta' \in B_\rho(\theta)$ for which $r_{\theta'}(t_1,x_1) - r_{\theta'}(t_2,Q_2) = 0$ forms a graph in the direction $\overrightarrow{Q_\theta Q_2}$, with small Lipschitz constant, and that the set $A_\lambda \cap B_\rho(\theta)$ is contained in a neighborhood of this graph of thickness $4\,\lambda^{-1}\,|Q_\theta - Q_2|^{-1}$.

We have $|Q_\theta - Q_2| \le 4\,(\epsilon_0\lambda)^{-\frac12}$, since Q_θ and x_2 are in the same slab, and hence the thickness is bounded below by $\epsilon_0^{\frac12}\lambda^{-\frac12}$. We may therefore control the number of $\epsilon_0^{\frac12}\lambda^{-\frac12}$ balls needed to cover $A_\lambda \cap B_\rho(\theta)$ by

$$(9.5)\quad \epsilon_0^{-\frac{n-1}{2}}\lambda^{\frac{n-3}{2}}\,\rho^{n-2}\,|Q_\theta - Q_2|^{-1} \approx \epsilon_0^{-\frac{n-1}{2}}\lambda^{\frac{n-3}{2}}\,\rho^{n-2}\,m^{-\frac12}\,(t_2-t_1)^{-\frac12}\,.$$

The result follows since $d \approx m \ge \lambda^{-1}$, and $\rho \lesssim m^{\frac12}\,(t_2-t_1)^{-\frac12}$.

Case 5: $2\,(t_2-t_1) \le m$, $4\lambda^{-1} \le m$. In this case, $|Q_\theta - Q_2| \approx m$ for all θ, which by (9.4) and the second part of (9.3) implies that, for $\theta \in A_\lambda$, we have $|\theta - \omega_0| \approx 1$. We may thus bound A_λ by $\approx c^{1-n}$ balls of radius c, and the proof proceeds as above. $\square$

For the proof of the dispersive estimates in the next section we do not need the full strength of Proposition 9.2. Instead, it suffices to consider the worst case for fixed t_1, t_2. This happens if $|m| \le 4\lambda^{-1}$ for $n=2$, if $-4\lambda^{-1} \le m \le 2(t_2-t_1) \lesssim (\epsilon_0\lambda)^{-\frac12}$ for $n=3$, and if $m \approx 2(t_2-t_1) \lesssim (\epsilon_0\lambda)^{-\frac12}$ for $n \ge 4$. We obtain

Corollary 9.3. *For all points $P_1 = (t_1,x_1)$ and $P_2 = (t_2,x_2)$ in space-time, and $\epsilon_0\lambda \ge 1$, the number $N_\lambda(P_1,P_2)$ of slabs of scale λ that contain both P_1 and P_2 satisfies the bound*

$$N_\lambda(P_1,P_2) \lesssim \epsilon_0^{-\frac{n-1}{2}}\lambda^{\frac{n-3}{2}}\,|t_1-t_2|^{-1}\,, \qquad n \ge 3\,,$$

$$N_\lambda(P_1,P_2) \lesssim \epsilon_0^{-\frac12}\,|t_1-t_2|^{-\frac12}\,, \qquad n = 2\,.$$

Another variation on the same theme is required for the proof of the two dimensional stability estimates in the Appendix.

Corollary 9.4. *Set $n=2$. Let $P_1 = (t_1,x_1)$, and let $Q_R^{t_2}$ be a square at time t_2, of sidelength R, with $R \ll (\epsilon_0\lambda)^{-\frac12}$. Then*

$$N_\lambda(P_1,Q_R^{t_2}) \lesssim \epsilon_0^{-\frac12}\,|t_1-t_2|^{-\frac12}(1+\lambda R)^{\frac12}\,.$$

Proof. The result is trivial if $|t_1 - t_2| \lesssim \lambda^{-1}$, or if $|t_1 - t_2| \lesssim R$. Hence assume that $R + \lambda^{-1} \ll |t_1 - t_2|$.

Case 1. If the distance d between $Q_R^{t_2}$ and $K_{P_1}^{t_2}$ is at least R, then we consider a set of points P_2^j spaced by $(4\lambda)^{-1}$ on the boundary of R. The slabs have thickness λ, so each slab intersecting $Q_R^{t_2}$ must also contain at least one of the points P_2^j. Since there are $\approx \lambda R$ such points, using Proposition 9.2 we obtain

$$N_\lambda(P_1, Q_R^{t_2}) \leq \sum_j N_\lambda(P_1, P_2^j) \lesssim \lambda\, R\, \epsilon_0^{-\frac{1}{2}} \lambda^{\frac{1}{2}} (1 + \lambda d)^{-\frac{1}{2}} \lesssim \epsilon_0^{-\frac{1}{2}} \lambda^{\frac{1}{2}} (\lambda R)^{\frac{1}{2}},$$

which is stronger than we need.

Case 2. If the distance between $Q_R^{t_2}$ and $K_{P_1}^{t_2}$ is at most R then $4Q_R^{t_2}$ and $K_{P_1}^{t_2}$ intersect. Fix Q_θ in the intersection. Then any slab through P_1 which intersects $Q_R^{t_2}$ must have direction ω with $|\omega - \theta| \leq c$, for some small c. We take the line L through Q_θ and of direction θ, and a $(4\lambda)^{-1}$ spaced set $\{P_2^j\}_{|j| \leq 32\lambda R}$ on L extending $8R$ on both sides of Q_θ. If a slab through P_1 intersects $Q_R^{t_2}$, then the slab with ϵ_0 replaced by $\epsilon_0/2$ must contain at least one of the points P_2^j. Since P_2^j is at distance $d \approx j\lambda^{-1}$ from the cone section $K_{P_1}^{t_2}$, we can again use Proposition 9.2 to compute

$$\begin{aligned} N_\lambda(P_1, Q_R^{t_2}) &\leq \sum_{|j| \leq 32\lambda R} N_\lambda(P_1, P_2^j) \\ &\lesssim \sum_{|j| \leq 32\lambda R} \epsilon_0^{-\frac{1}{2}} \lambda^{\frac{1}{2}} (1 + \lambda |t_1 - t_2|)^{-\frac{1}{2}} (1 + j)^{-\frac{1}{2}} \\ &\lesssim \epsilon_0^{-\frac{1}{2}} \lambda^{\frac{1}{2}} (1 + \lambda |t_1 - t_2|)^{-\frac{1}{2}} (1 + \lambda R)^{\frac{1}{2}}. \end{aligned}$$

□

10. Dispersive estimates

Our proof of the dispersive estimates for the parametrix (7.2) uses only pointwise bounds on the wave packets, not their oscillation. Since normalized wave packets have size at most $O\big(\epsilon_0^{\frac{n-1}{4}} \lambda^{\frac{n-3}{4}}\big)$, the estimate (7.2) is a consequence of the following result.

Proposition 10.1. *Let*

$$u = \sum_{T \in \mathcal{T}} a_T \chi_T,$$

where $\sum_T a_T^2 \le 1$. *Then*

$$\|u\|_{L^4_t(L^\infty_x)} \lesssim \epsilon_0^{-\frac{1}{4}} (\ln \lambda)^2\,, \qquad n = 2\,,$$

$$\|u\|_{L^2_t(L^\infty_x)} \lesssim \epsilon_0^{-\frac{n-1}{4}} \lambda^{\frac{n-3}{4}} (\ln \lambda)^3\,, \qquad n \ge 3\,.$$

Proof. **Case 1:** $n = 2$**.** We consider λ points $P_j = (t_j, x_j)$, where the t_j are separated by λ^{-1}, but with x_j arbitrarily chosen. Then we need to show that

$$\sum_j |u(t_j, x_j)|^4 \lesssim \epsilon_0^{-1} \lambda (\log \lambda)^2\,.$$

We may assume that $|a_T| \ge \epsilon^{\frac{1}{2}} \lambda^{-\frac{1}{2}}$, since each point lies in at most $\approx \epsilon^{-\frac{1}{2}} \lambda^{\frac{1}{2}}$ slabs. We then decompose the sum $u = \sum a_T \chi_T$ dyadically with respect to the size of a_T. It thus suffices to prove the result, with $(\log \lambda)^2$ replaced by $\log \lambda$, in the case that we have a sum over exactly N slabs $T \in T_N$ for which $|a_T| \approx N^{-\frac{1}{2}}$.

We next decompose the sum over j via a dyadic decomposition in the number of slabs containing (t_j, x_j). We may assume that we are summing over M points (t_j, x_j), each of which is contained in approximately L slabs. Then $|u(t_j, x_j)| \lesssim N^{-\frac{1}{2}} L$ and

$$\sum_j |u(t_j, x_j)|^4 \lesssim L^4 M N^{-2}\,.$$

Hence, to conclude, we need to prove that

$$L^4 M \lesssim \epsilon_0^{-1} \lambda N^2\,. \tag{10.1}$$

This is a counting problem, which we will solve by evaluating in two different ways the number K of pairs (i, j) for which P_i and P_j are contained in a common slab, counted with multiplicity. For $T \in \mathcal{T}_N$, we denote by n_T the number of points P_j contained in T. Then

$$K \approx \sum_{n_T \ge 2} n_T^2 \gtrsim N^{-1} \Big(\sum_{n_T \ge 2} n_T \Big)^2\,.$$

Note that $\sum_{T \in \mathcal{T}_N} n_T \approx ML$. We consider two cases. If

$$\sum_{n_T \ge 2} n_T \le \sum_{n_T = 1} n_T\,,$$

then $N \approx ML$. Combined with the trivial bound $L \lesssim \epsilon_0^{-\frac{1}{2}}\lambda^{\frac{1}{2}}$, this directly gives (10.1). Otherwise, we have the following lower bound,

$$K \gtrsim N^{-1}M^2L^2\,. \tag{10.2}$$

On the other hand, by Corollary 9.3, the number of slabs which contain both (t_i, x_i) and (t_j, x_j) is dominated by $\epsilon_0^{-\frac{1}{2}}|t_i - t_j|^{-\frac{1}{2}}$. Hence

$$K \lesssim \epsilon_0^{-\frac{1}{2}} \sum_{\substack{1\le i,j\le M\\ i\ne j}} |t_i - t_j|^{-\frac{1}{2}}\,.$$

The sum is maximized in case the t_j are as close as possible, that is, if the t_j are consecutive multiples of λ^{-1}. Thus,

$$K \lesssim \epsilon_0^{-\frac{1}{2}}\lambda^{\frac{1}{2}} \sum_{\substack{1\le i,j\le M\\ i\ne j}} |i-j|^{-\frac{1}{2}} \lesssim \epsilon_0^{-\frac{1}{2}}\lambda^{\frac{1}{2}}M^{\frac{3}{2}}\,. \tag{10.3}$$

Combining (10.2) and (10.3) yields (10.1).

Case 2: $n \ge 3$. We proceed in a similar way, with the same notation. We need to prove that

$$\sum_j |u(t_j, x_j)|^2 \lesssim \epsilon_0^{-\frac{n-1}{2}}\lambda^{\frac{n-1}{2}}(\log\lambda)^3\,.$$

We make similar decompositions of the sums as for $n = 2$. Then, since

$$\sum_j |u(t_j, x_j)|^2 \approx L^2MN^{-1}\,,$$

it remains to prove that

$$L^2M \lesssim \epsilon_0^{-\frac{n-1}{2}}\lambda^{\frac{n-1}{2}}N\log\lambda\,. \tag{10.4}$$

As above, there are two cases to consider: either $N \approx ML$, or the estimate (10.2) holds. In the first case, (10.4) follows from the trivial bound $L \lesssim \epsilon_0^{-\frac{n-1}{2}}\lambda^{\frac{n-1}{2}}$. In the second case, we have the following substitute for (10.3),

$$K \lesssim \epsilon_0^{-\frac{n-1}{2}}\lambda^{\frac{n-1}{2}} \sum_{\substack{1\le i,j\le M\\ i\ne j}} |t_i - t_j|^{-1} \lesssim M\epsilon_0^{-\frac{n-1}{2}}\lambda^{\frac{n-1}{2}}\log\lambda\,.$$

Together with (10.2) this yields (10.4). $\square$

Appendix A. The stability estimates in $2+1$ dimensions

The aim of this section is to establish the $n = 2$ part of Lemma 2.4, namely the estimate (2.11). This turns out to be considerably more difficult than the higher dimensional estimate. On the positive side, the proof uses a type of estimate which has not previously played a role in the study of quasi-linear hyperbolic equations.

The crux of the problem is the need for improved low frequency estimates for the product of two high frequency waves. Such estimates are known to be true in the constant coefficient case, see [12]. In our case, the line of argument in [12] appears hopeless. Furthermore, we need such estimates in a context where the two factors solve different wave equations. This motivates us to think of the bilinear estimates as a byproduct of certain multiscale linear estimates.

To describe our results, we first introduce the appropriate function spaces. Given $\mu \geq 1$, we decompose the space-time into cubes of sidelength μ^{-1}, and introduce the notation

$$l^p l^q (L^r L^s)_\mu = l^p_t l^q_x (L^r_t L^s_x)_\mu \,,$$

where the $L^r_t L^s_x$ norm is evaluated within each μ^{-1} cube, and the $l^p_t l^q_x$ norm is then taken with respect to the collection of such cubes. Our main estimate has roughly the form

$$\| \langle D_x \rangle^{\frac{1}{2}} v \|_{l^{\frac{16}{5}} l^\infty (L^{16} L^2)_\mu} \lesssim \mu^{-\frac{1}{2}} \big(\, \|dv(0)\|_{L^2_x} + \|\Box_{\mathrm{g}} v\|_{L^1_t L^2_x} \big) \,.$$

The indices can be improved in the constant coefficient case, but the above suffices for our purposes. Similar estimates also hold in higher dimensions, and are in fact easier to prove. The $2+1$ dimensional case has certain unique features which make it more delicate.

The plan of this appendix is as follows. In A.1 we obtain localized energy estimates on intermediate scales between λ^{-1} and $\lambda^{-\frac{1}{2}}$. This is combined in section A.2 with the modified overlap result in Corollary 9.4 to show that our frequency localized parametrix satisfies the new estimates. These are then extended to the exact solutions in A.3, which also contains the crucial $H^{\frac{3}{4}} \times H^{-\frac{1}{4}}$ well-posedness result for the linear equation. Finally, in A.4, we use these estimates to prove the stability result (2.11). To keep the notation simple we neglect the parameter ϵ_0, which is irrelevant for the arguments here.

A.1. **Localized energy estimates.** In this section we prove an energy estimate for superpositions of wave packets restricted to smaller cubes. Let $\lambda^{-1} \leq R \leq \lambda^{-\frac{1}{2}}$. We seek an L^2_x estimate for a superposition of λ-wave packets in a cube Q_R of size R. We denote by $\mathcal{T}_\lambda$ the family of λ-slabs introduced in section 8.4, and by $\mathcal{T}_\lambda(Q_R)$ the subset of λ-slabs which intersect Q_R. By Proposition 6.1, if two characteristic surfaces Σ_{ω_1,r_1} and Σ_{ω_2,r_2} pass within distance λ^{-1} at some point in Q_R, then their respective intersections with Q_R remain within distance λ^{-1} of each other, provided that $|\omega_1-\omega_2| \leq (\lambda R)^{-1}$. Consequently, two λ-slabs which intersect within Q_R have essentially the same intersection with Q_R provided that their initial angles differ by less then $(\lambda R)^{-1}$. This motivates a decomposition

$$\mathcal{T}_\lambda(Q_R) = \cup_{j\in J}\mathcal{T}_\lambda^j(Q_R)$$

where for each $j \in J$, $\mathcal{T}_\lambda^j(Q_R)$ contains a family of slabs which intersect within Q_R and which have initial angle less than $(\lambda R)^{-1}$ from each other. The number of distinct groups needed is $|J| \approx \lambda^2 R^2$, and each $\mathcal{T}_\lambda^j(Q_R)$ contains about $\lambda^{-\frac{1}{2}}R^{-1}$ slabs. Furthermore, for any given angle, the subcollection of $\mathcal{T}_\lambda^j(Q_R)$ with that initial angle has the finite overlap property.

Lemma A.1. *Let χ_{Q_R} be a smooth cutoff to Q_R, and let u_T be a family of normalized wave packets associated to $T \in \mathcal{T}_\lambda$. Then*

$$\| \, dS_\lambda \sum_{T\in\mathcal{T}_\lambda} \chi_{Q_R} a_T u_T \, \|_{L^2_x}^2 \lesssim \rho(t)^{\frac{1}{2}}\, R^2\, \lambda \sum_j \Big(\sum_{T\in\mathcal{T}_\lambda^j(Q_R)} |a_T| \Big)^2 , \tag{A.1}$$

where

$$\rho(t) = \max\{ \, \|d\mathrm{g}(t)\|_{L^\infty_x} \, , R^{-2}\lambda^{-1} \, \} \, .$$

Proof. We use the same computation as in the proof of Lemma 8.6 but with $\rho(t)$ chosen as above. This guarantees that the decomposition scale r satisfies

$$r = \rho(t)^{-\frac{1}{2}}\lambda^{-\frac{1}{2}} \lesssim R \, .$$

Within a sub-cube Q_r of size r, the argument in the proof of Lemma 8.6 shows that we have almost orthogonality for angles which are at least $(\lambda r)^{-1} = \rho(t)^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$ separated. More precisely, we have

$$\| \, dS_\lambda \sum_{T\in\mathcal{T}_\lambda} \chi_{Q_r} a_T u_T \, \|_{L^2_x}^2 \lesssim r\, \lambda^{\frac{1}{2}} \sum_j \Big(\sum_{T\in\mathcal{T}_\lambda^j(Q_r)} |a_T| \Big)^2 ,$$

where the factor $r\lambda^{\frac{1}{2}}$ represents the square L^2_x norm of the gradient of a normalized wave packet in an r cube. Each set $\mathcal{T}_\lambda^j(Q_r)$ has elements T in common with at most Rr^{-1} sets $\mathcal{T}_\lambda^j(Q_R)$. Hence, an application of the Schwarz inequality yields

$$\| \, dS_\lambda \sum_{T\in\mathcal{T}_\lambda} \chi_{Q_r} a_T u_T \, \|_{L^2_x}^2 \lesssim R\, \lambda^{\frac{1}{2}} \sum_j \Big(\sum_{\substack{T\in\mathcal{T}_\lambda^j(Q_R) \\ T\cap Q_r \neq\emptyset}} |a_T| \Big)^2 .$$

To conclude, we sum over a grid of r-cubes in Q_R. Each $T \in \mathcal{T}_\lambda^j(Q_R)$ intersects at most $R r^{-1}$ cubes Q_r, hence

$$\| dS_\lambda \sum_{T\in\mathcal{T}_\lambda} \chi_{Q_R} a_T u_T \|_{L^2_x}^2 \lesssim R^2\, r^{-1}\, \lambda^{\frac{1}{2}} \sum_j \Big(\sum_{T\in\mathcal{T}_\lambda^j(Q_R)} |a_T| \Big)^2 .$$

The conclusion follows now since

$$R^2\, r^{-1}\, \lambda^{\frac{1}{2}} = \rho(t)^{\frac{1}{2}}\, R^2\, \lambda .$$

□

A.2. **Dispersive estimates.** The dispersive/energy estimate we establish here is

Lemma A.2. *For $T \in \mathcal{T}_\lambda$ we consider a family u_T of normalized wave packets. Then*

$$\| dS_\lambda \sum_{T\in\mathcal{T}_\lambda} a_T u_T \|_{l^{\frac{16}{5}} l^\infty (L^{16} L^2)_\mu} \lesssim (\log\lambda)^3 \lambda^{\frac{1}{2}} \mu^{-\frac{1}{2}} \Big(\sum_{T\in\mathcal{T}_\lambda} |a_T|^2 \Big)^{\frac{1}{2}} . \tag{A.2}$$

Proof. The result is a straightforward consequence of the energy estimates if $\mu \le \lambda^{\frac{1}{2}}$, and of the dispersive estimates in Proposition 10.1 if $\lambda \le \mu$. Hence in what follows we assume that $\lambda^{\frac{1}{2}} \le \mu \le \lambda$.

We select a family of μ^{-1} space-time cubes $\{Q_\mu^k\}_{k=1,\mu}$ which are μ^{-1} equidistant in time but arbitrarily chosen spatially. We denote by I_k the corresponding time sections. Within each Q_μ^k we choose an arbitrary time section Q_μ^{k,t_k} at time t_k. We claim that the following estimate holds:

$$\Bigg[\sum_k \Big(\sum_j \Big(\sum_{T\in\mathcal{T}_\lambda^j(Q_\mu^{k,t_k})} |a_T| \Big)^2 \Big)^2 \Bigg]^{\frac{1}{4}} \lesssim (\log\lambda)^3\, \lambda^{\frac{1}{4}} \Big(\sum_{T\in\mathcal{T}_\lambda} |a_T|^2 \Big)^{\frac{1}{2}} . \tag{A.3}$$

We postpone for now the proof of (A.3) and instead show that, together with (A.1), it implies (A.2).

Within each section $Q_\mu^{k,t}$ we apply (A.1) with $R = \mu^{-1}$ to obtain

$$\| dS_\lambda \sum_{T\in\mathcal{T}_\lambda} \chi_{Q_\mu^k} a_T u_T(t) \|_{L^2_x}^2 \lesssim \rho(t)^{\frac{1}{2}} \mu^{-2} \lambda \sum_j \Big(\sum_{T\in\mathcal{T}_\lambda^j(Q_\mu^{k,t})} |a_T| \Big)^2 ,$$

where

$$\rho(t) = \max\{ \|dg(t)\|_{L^\infty}, \mu^2\lambda^{-1} \} .$$

Integrating over t, this yields

$$\| dS_\lambda \sum_{T\in\mathcal{T}_\lambda} \chi_{Q^k_\mu} a_T u_T \|_{L^{16}_t L^2_x} \lesssim \\ \|\rho\|^{\frac14}_{L^4_t(I_k)} \, \mu^{-1} \, \lambda^{\frac12} \Big(\sup_{t\in I_k} \sum_j \Big(\sum_{T\in\mathcal{T}^j_\lambda(Q^{k,t}_\mu)} |a_T| \Big)^2 \Big)^{\frac12} .$$

We sum over k and apply Hölder's inequality to yield

$$\Big(\sum_k \| dS_\lambda \sum_{T\in\mathcal{T}_\lambda} \chi_{Q^k_\mu} a_T u_T \|^{\frac{16}{5}}_{L^{16}_t L^2_x} \Big)^{\frac{5}{16}} \lesssim \\ \|\rho\|^{\frac14}_{L^4_t} \, \mu^{-1} \, \lambda^{\frac12} \Big[\sum_k \Big(\sup_{t\in I_k} \sum_j \Big(\sum_{T\in\mathcal{T}^j_\lambda(Q^{k,t}_\mu)} |a_T| \Big)^2 \Big)^2 \Big]^{\frac14} .$$

Since $dg \in L^4_t L^\infty_x$, it follows that $\|\rho(t)\|_{L^4_t} \lesssim \mu^2 \lambda^{-1}$. To conclude, we observe that we may bound the quantity

$$\sup_{Q_\mu} \| dS_\lambda \sum_{T\in\mathcal{T}_\lambda} a_T u_T \|_{L^{16}_t L^2_x(Q_\mu)}$$

taken over cubes Q_μ belonging to the same time slice I_k, by an expression of the form

$$\sup_{Q_\mu} \| dS_\lambda \sum_{T\in\mathcal{T}_\lambda} \chi_{Q_\mu} a_T u_T \|_{L^{16}_t L^2_x} ,$$

taken over the same cubes, since inserting an appropriate polynomially growing weight leaves such an expression unchanged. Thus, the last inequality combined with (A.3) implies (A.2).

It remains to prove (A.3). We begin by partitioning the initial angles into sectors of size $\lambda^{-1}\mu$, and take a corresponding partition of $\mathcal{T}_\lambda$,

$$\mathcal{T}_\lambda = \bigcup_\omega \mathcal{T}^\omega_\lambda .$$

For each j, k, t the set $\mathcal{T}^j_\lambda(Q^{k,t}_\mu)$ consists of packets with initial angles within $\lambda^{-1}\mu$ of each other, therefore it is contained in finitely many

$\mathcal{T}_\lambda^\omega$'s. Then

$$\left[\sum_k\Big(\sum_j\Big(\sum_{T\in\mathcal{T}_\lambda^j(Q_\mu^{k,t_k})}|a_T|\Big)^2\Big)^2\right]^{\frac12}\lesssim\left[\sum_k\Big(\sum_\omega\sum_j\Big(\sum_{T\in\mathcal{T}_\lambda^\omega\cap\mathcal{T}_\lambda^j(Q_\mu^{k,t_k})}|a_T|\Big)^2\Big)^2\right]^{\frac12}$$
$$\lesssim\sum_\omega\left[\sum_k\Big(\sum_j\Big(\sum_{T\in\mathcal{T}_\lambda^\omega\cap\mathcal{T}_\lambda^j(Q_\mu^{k,t_k})}|a_T|\Big)^2\Big)^2\right]^{\frac12}$$

which implies that (A.3) can be reduced to the case where all slabs are contained in a single $\mathcal{T}_\lambda^\omega$. The advantage of this reduction is that the intersections of such slabs with the sections Q_μ^{k,t_k} are easier to describe. To describe their intersections, we use a λ^{-1} spaced subset of the foliation Σ_ω to decompose each set Q_μ^{k,t_k} into approximately $\mu^{-1}\lambda$ curved rectangles, which we call *leafs.* Any slab in $\mathcal{T}_\lambda^\omega$ can intersect at most a bounded number of leafs, and so there is essentially a 1-1 correspondence between the sets $\mathcal{T}_\lambda^\omega\cap\mathcal{T}_\lambda^j(Q_\mu^{k,t_k})$ which are nonempty, and the leafs of Q_μ^{k,t_k}.

We are now in a position to use an argument similar to that in the proof of Proposition 10.1. By taking successive decompositions of the index sets for T, j, and k, with at most $(\log\lambda)^3$ terms in all, we may reduce to the case that in the sum there are

- N slabs in $\mathcal{T}_\lambda^\omega$, for which $a_T=1$,
- each leaf in the sum is intersected by about L slabs,
- each section Q_μ^{k,t_k} contains roughly P leafs,
- there are a total of M sections Q_μ^{k,t_k}.

Then (A.3) reduces to the estimate

$$MP^2L^4\lesssim\lambda N^2\,. \tag{A.4}$$

We index the chosen slabs by s, the chosen leafs by ℓ, and set $\chi(s,\ell)=1$ if s intersects ℓ, and $\chi(s,l)=0$ otherwise. Set

$$K(\ell)=\sum_s\chi(s,l)\sum_{\ell'}\chi(s,\ell')\,,\qquad K=\sum_\ell K(\ell)\,.$$

Thus, $K(\ell)$ is the number of leafs, including ℓ, which can be reached from ℓ by the selected slabs, counted with multiplicity.

Since there are MP indices ℓ and N indices s, an application of the Schwarz inequality yields the lower bound

$$K\gtrsim N^{-1}(MPL)^2\,.$$

Hence it remains to establish the upper bound

$$K \lesssim M^{\frac{3}{2}} P \lambda^{\frac{1}{2}} ,$$

for which it suffices to show that, for each ℓ,

$$K(\ell) \lesssim M^{\frac{1}{2}} \lambda^{\frac{1}{2}} .$$

Given a leaf ℓ, we fix a point $P_0 = (t_0, x_0) \in \ell$. Then if a chosen slab s intersects ℓ, since we are restricting ourselves to slabs within angle $\mu\lambda^{-1}$ of ω, it follows that the slab of half the frequency must contain P_0. Thus, we need show that

$$K(P_0) \lesssim M^{\frac{1}{2}} \lambda^{\frac{1}{2}} ,$$

where $K(P_0)$ has the obvious definition. At this point, we recall Corollary 9.4, which says that the number of slabs which intersect both P_0 and the μ^{-1} square time section Q_μ^{k,t_k} at time t_k is at most

$$K(P_0, Q_\mu^{k,t_k}) \lesssim \min \left[\lambda^{\frac{1}{2}} , \frac{\mu^{\frac{1}{2}} \lambda^{\frac{1}{2}}}{(t_0 - t_k)^{\frac{1}{2}}} \right] .$$

Since each slab passes through at most a fixed number of leafs in any given Q_μ^{k,t_k}, and since the time slices are spaced by μ^{-1}, it follows that

$$K(P_0) \lesssim \sum_{k=1}^{M} \frac{\mu^{\frac{1}{2}} \lambda^{\frac{1}{2}}}{(t_0 - t_k)^{\frac{1}{2}}} \lesssim \lambda^{\frac{1}{2}} \sum_{k=1}^{M} k^{-\frac{1}{2}} \approx \lambda^{\frac{1}{2}} M^{\frac{1}{2}} .$$

□

A.3. **Linear local well-posedness.** Here we improve the Sobolev well-posedness range in Proposition 7.1 for the linear equation

$$\left\{ \begin{array}{l} \Box_{\mathrm{g}} v = 0 \ \text{ in } \ [-T, T] \times \mathbb{R}^2 \\ \\ v(t_0) = v_0, \quad \partial_t v(t_0) = v_1 , \end{array} \right. \tag{A.5}$$

and complement it with the new dispersive/energy estimates.

We assume that u is the smooth solution on $[-T, T] \times \mathbb{R}^2$ to the equation

$$\Box_{\mathrm{g}(u)} u = q^{ij}(u)\, \partial_i u\, \partial_j u$$

produced in the proof of Proposition 2.1. We first show that for this solution we may strengthen the condition (WP4) to the following result.

Lemma A.3. *For $\frac{3}{4} \leq r \leq s + 1$, the equation* (A.5) *with* $\mathrm{g} = \mathrm{g}(u)$ *is well-posed in $H^r \times H^{r-1}$, and the following estimates hold:*

$$\|v\|_{L_t^\infty H_x^r} + \|\partial_t v\|_{L_t^\infty H_x^{r-1}} \lesssim \|(v_0, v_1)\|_{H^r \times H^{r-1}} ,$$

$$\|\langle D_x\rangle^\rho v\|_{L_t^4 L_x^\infty} + \|\langle D_x\rangle^{\rho-1} dv\|_{L_t^4 L_x^\infty} \lesssim \|(v_0, v_1)\|_{H^r \times H^{r-1}}\,, \quad \rho < r - \frac{3}{4}\,,$$
$$\|\langle D_x\rangle^\rho dv\|_{l^{\frac{16}{5}} l^\infty (L^{16} L^2)_\mu} \lesssim \mu^{-\frac12}\, \|(v_0, v_1)\|_{H^r \times H^{r-1}}\,, \quad \rho < r - \frac{3}{2}\,.$$

Proof. We will prove the result for $T = 1$ with g replaced by $\mathbf{g}$. The arguments of section 3 then yield the result as stated.

By (A.2), we know that the desired estimates hold for our frequency localized parametrix, so it remains to show that they hold for the actual solution v. In the range $1 \le r \le s+1$, the arguments in the proof of Proposition 7.1 apply with virtually no change. Hence we only need to consider the case $\frac34 \le r < 1$, which we illustrate in the case $r = \frac34$.

We set $w = \langle D_x\rangle^{-\frac14} v$. Then w must solve

$$\Box_{\mathbf{g}} w = [\,\mathbf{g} - \langle D_x\rangle^{-\frac14} \mathbf{g} \langle D_x\rangle^{\frac14}]\, d_x dw\,. \tag{A.6}$$

We solve this equation with Cauchy data $\langle D_x\rangle^{-\frac14}(v_0, v_1)$ by using the $r = 1$ local well-posedness result combined with a fixed point argument. Given $F \in L_t^1 L_x^2$, we let w_f denote the unique solution to the equation

$$\Box_{\mathbf{g}} w_F = F$$

with Cauchy data $\langle D_x\rangle^{-\frac14}(v_0, v_1)$, and we set

$$LF = B(\mathbf{g}, dw_F) = [\,\mathbf{g} - \langle D_x\rangle^{-\frac14} \mathbf{g} \langle D_x\rangle^{\frac14}]\, d_x dw_F\,.$$

If F is a fixed point for L, then w_F solves (A.6). We find F using the contraction principle. For this it suffices to prove the estimate

$$\|LF\|_{L_t^2 L_x^2} \lesssim \epsilon_0 \left(\|(v_0, v_1)\|_{H^{\frac34} \times H^{\frac14}} + \|F\|_{L_t^1 L_x^2} \right). \tag{A.7}$$

To prove (A.7), we use the bounds for w which follow from the $r = 1$ case of Lemma A.3 combined with Duhamel's formula, namely that the quantity

$$\|dw\|_{L_t^\infty L_x^2} + \|\langle D_x\rangle^{-\frac34 - \delta} dw\|_{L_t^4 L_x^\infty} + \mu^{\frac12} \|\langle D_x\rangle^{-\frac12 - \delta} dw\|_{l^{\frac{16}{5}} l^\infty (L^{16} L^2)_\mu}$$

for any $\delta > 0$ is bounded by the quantity

$$\|(v_0, v_1)\|_{H^{\frac34} \times H^{-\frac14}} + \|F\|_{L_t^1 L_x^2}\,.$$

For $\mathbf{g}$, on the other hand, we use the bounds for $\lambda > 1$,

$$\|d\mathbf{g}\|_{L_t^4 C_x^\delta \cap L_t^\infty H_x^{s-1}} \lesssim \epsilon_0\,, \quad \|S_\lambda \mathbf{g}\|_{l^\infty l^2 (L^\infty L^2)_\mu} \lesssim \epsilon_0\, \lambda^{-s}\,.$$

The second bound is a consequence of energy estimates and finite propagation velocity arguments applied to $S_\lambda \mathbf{g}$.

We take a paradifferential decomposition of LF,

$$LF = \sum_{\lambda} B(S_{<\lambda}\mathbf{g}, S_\lambda dw_F) + B(S_\lambda \mathbf{g}, S_{<\lambda} dw_F) + B(S_\lambda \mathbf{g}, S_\lambda dw_F)\,.$$

The first term is localized at frequency λ, and we can bound the sum in $L^2_t L^2_x$ using orthogonality over λ together with the fixed time commutator estimate,

$$\|B(S_{<\lambda}\mathbf{g}, S_\lambda dw_F)(t)\|_{L^2_x} \lesssim \|d\mathbf{g}(t)\|_{L^\infty_x} \|S_\lambda dw_F(t)\|_{L^2_x}\,.$$

The second term is also localized at frequency λ, but this time there is no gain from the commutation. Instead, the two components of B are estimated separately at fixed time, to obtain

$$\|B(S_\lambda \mathbf{g}, S_{<\lambda} dw_F)(t)\|_{L^2_x} \lesssim \|S_\lambda d\mathbf{g}(t)\|_{H^{s-1}} \|\langle D_x\rangle^{-\frac{3}{4}-\delta} w_F(t)\|_{L^\infty_x}\,.$$

For the third term, one handles the case $\lambda = 1$ as for the first term. For $\lambda > 1$, there is again no gain from the commutation, and we handle the two components of B separately. The first component is easy to estimate,

$$\sum_{\lambda} \|(S_\lambda \mathbf{g})(S_\lambda d_x dw_F)(t)\|_{L^2_x} \lesssim \|d\mathbf{g}(t)\|_{C^\delta_x} \|dw_F(t)\|_{L^2_x}\,.$$

The second component, $\langle D_x\rangle^{-\frac{1}{4}}(S_\lambda \mathbf{g})\langle D_x\rangle^{\frac{1}{4}} d_x S_\lambda dw_F$, is the term which causes the most difficulties. Because the product of two frequency λ functions contributes to all lower frequencies, we need a better estimate due to the $\langle D_x\rangle^{-\frac{1}{4}}$ operator. We first bound

$$\begin{aligned}
&\|\langle D_x\rangle^{-\frac{1}{4}}(S_\lambda \mathbf{g})\langle D_x\rangle^{\frac{1}{4}} d_x S_\lambda dw_F\|_{L^2_t L^2_x} \\
&\qquad \lesssim \sum_{\mu \lesssim \lambda} |\mu|^{-\frac{1}{4}} \|S_\mu [(S_\lambda \mathbf{g})\langle D_x\rangle^{\frac{1}{4}} d_x S_\lambda dw_F]\|_{L^2_t L^2_x} \\
&\qquad\qquad \approx \sum_{\mu \lesssim \lambda} |\mu|^{-\frac{1}{4}} \|S_\mu [(S_\lambda \mathbf{g})\langle D_x\rangle^{\frac{1}{4}} d_x S_\lambda dw_F]\|_{l^2 l^2 (L^2 L^2)_\mu}\,.
\end{aligned}$$

Since S_μ is mollification on the μ^{-1} scale, we may use the Young and Hölder inequalities to bound this by

$$\sum_{\mu \lesssim \lambda} |\mu|^{\frac{1}{2}} \|(S_\lambda \mathbf{g})\langle D_x\rangle^{\frac{1}{4}} d_x S_\lambda dw_F\|_{l^{\frac{16}{5}} l^2 (L^{16} L^1)_\mu}\,.$$

The Hölder inequality is now used to bound this by

$$\sum_{\mu \lesssim \lambda} |\mu|^{\frac{1}{2}} \|S_\lambda \mathbf{g}\|_{l^\infty l^2 (L^\infty L^2)_\mu} \lambda^{\frac{5}{4}} \|S_\lambda dw_F\|_{l^{\frac{16}{5}} l^\infty (L^{16} L^2)_\mu}$$

$$\lesssim \sum_{\mu \lesssim \lambda} \lambda^{\frac{7}{4}+\delta} \|S_\lambda \mathbf{g}\|_{l^\infty l^2 (L^\infty L^2)_\mu} |\mu|^{\frac{1}{2}} \|\langle D_x \rangle^{-\frac{1}{2}-\delta} dw_F\|_{l^{\frac{16}{5}} l^\infty (L^{16} L^2)_\mu} .$$

The desired bound now follows. □

A.4. **Stability estimates.** Here we prove the two dimensional stability estimate in Lemma 2.4 which for convenience we restate below. We prove stability under stronger conditions than those stated in Lemma 2.4, but the existence of the stronger solution shows that Lemma 2.4 can be used as stated.

Lemma A.4. *Let u be a solution to* (1.1) *and* (1.2) *on $[-T, T] \times \mathbb{R}^2$ which satisfies the conditions* (WP3) *and* (WP4)*, as well as the conclusion of Lemma A.3. Let v be another solution to the equation* (1.1) *with initial data $(v_0, v_1) \in H^s \times H^{s-1}$, such that $dv \in L_t^\infty H_x^{s-1} \cap L_t^2 L_x^\infty$. Then*

$$\|d(u-v)\|_{L_t^\infty H_x^{-\frac{1}{4}}} \leq C_v \|(u_0 - v_0, u_1 - v_1)\|_{H^{\frac{3}{4}} \times H^{-\frac{1}{4}}} , \tag{A.8}$$

where C_v depends on u and on $\|dv\|_{L_t^\infty H_x^{s-1} \cap L_t^2 L_x^\infty}$.

Proof. The function $w = u - v$ solves the equation

$$\Box_{g(u)} w = a_0\, dw + a_1\, w , \tag{A.9}$$

where the functions a_0 and a_1 are of the form

$$a_0 = q(v)\, d(u,v) , \quad a_1 = a(u,v)\, d_x dv + b(u,v)\, (du)^2 .$$

To show that (A.9) is well-posed in $H^{\frac{3}{4}} \times H^{-\frac{1}{4}}$, we use the conclusion of Lemma A.3, together with a fixed point argument based on the Duhamel principle. Given $F \in L_t^1 H_x^{-\frac{1}{4}}$, we denote by w_F the unique solution to the equation

$$\Box_{g(u)} w_F = F ,$$

with Cauchy data (w_0, w_1), and we set

$$LF = a_0\, dw_F + a_1 w_F .$$

If F is a fixed point for L then w_F solves (A.9). The uniqueness of this fixed point, together with the easily verified condition that, for u and v as given, we have

$$a_0\, d(u-v) + a_1\, (u-v) \in L_t^\infty H_x^{-\frac{1}{4}} ,$$

will show that w must coincide with w_F.

To show that L is a contraction on $L^1_t H_x^{-\frac{1}{4}}$, it suffices to show that

$$\|LF\|_{L^2_t H_x^{-\frac{1}{4}}} \lesssim \|(w_0, w_1)\|_{H^{-\frac{3}{4}} \times H^{-\frac{1}{4}}} + \|F\|_{L^1_t H_x^{-\frac{1}{4}}} . \tag{A.10}$$

We denote

$$M = \|(w_0, w_1)\|_{H^{-\frac{3}{4}} \times H^{-\frac{1}{4}}} + \|F\|_{L^1_t H_x^{-\frac{1}{4}}} .$$

By the conditions of Lemma A.3 and the Duhamel principle, we have

$$\|dw_F\|_{L^\infty_t H_x^{-\frac{1}{4}}} + \|\langle D_x\rangle^{-1-\delta'} dw_F\|_{L^4_t L^\infty_x} \lesssim M ,$$

$$\sup_\mu \mu^{\frac{1}{2}} \|\langle D_x\rangle^{-\frac{3}{4}-\delta'} dw_F\|_{l^{\frac{16}{5}} l^\infty (L^{16} L^2)_\mu} \lesssim M ,$$

for each $\delta' > 0$. On the other hand, for a_0 and a_1 we have the estimates

$$a_0 \in L^2_t L^\infty_x \cap L^\infty_t H_x^{\frac{3}{4}+\delta} , \qquad a_1 \in L^\infty_t H_x^{-\frac{1}{4}+\delta} \cap d_x L^2_t L^\infty_x .$$

In addition, we have the localized energy bounds

$$\lambda^{\frac{3}{4}+\delta} \|S_\lambda a_0\|_{l^\infty l^2 (L^\infty L^2)_\mu} + \lambda^{-\frac{1}{4}+\delta} \|S_\lambda a_1\|_{l^\infty l^2 (L^\infty L^2)_\mu} \lesssim C_v$$

as a consequence of the Sobolev bounds for u and v and finite propagation velocity.

We estimate LF using a paraproduct decomposition on each fixed time slices. We consider the term $a_0\, dw_F$; the estimate for the second term is similar. Taking the paraproduct decomposition (see Bony [4], or also Taylor [24], chapter 3), we write

$$a_0\, dw_F = T_{a_0} dw_F + T_{dw_F} a_0 + R(a_0, dw_F) .$$

The first two terms are easy to estimate:

$$\|T_{a_0} dw_F(t)\|_{H^{-\frac{1}{4}}} \lesssim \|dw_F(t)\|_{H^{-\frac{1}{4}}} \|a_0(t)\|_{L^\infty} ,$$

respectively

$$\|T_{dw_F} a_0(t)\|_{H^{-\frac{1}{4}}} \lesssim \|a_0(t)\|_{H^{\frac{3}{4}+\delta}} \|\langle D_x\rangle^{-1-\delta} dw_F(t)\|_{L^\infty} .$$

It is more difficult to estimate the remainder

$$R(a_0, dw_F) \approx \sum_\lambda (S_\lambda a_0)(S_\lambda dw_F) .$$

As in the proof of Lemma A.3, we used the improved bound on μ-cubes for $\mu \lesssim \lambda$. We first bound

$$\begin{aligned}
&\|\langle D_x\rangle^{-\frac{1}{4}}(S_\lambda a_0)(S_\lambda dw_F)\|_{L^2_t L^2_x} \\
&\qquad\qquad\lesssim \sum_{\mu\lesssim\lambda} |\mu|^{-\frac{1}{4}} \|S_\mu\left[(S_\lambda a_0)(S_\lambda dw_F)\right]\|_{L^2_t L^2_x} \\
&\qquad\qquad\qquad\approx \sum_{\mu\lesssim\lambda} |\mu|^{-\frac{1}{4}} \|S_\mu\left[(S_\lambda a_0)(S_\lambda dw_F)\right]\|_{l^2 l^2 (L^2 L^2)_\mu} .
\end{aligned}$$

Since S_μ is mollification on the μ^{-1} scale, we may as before use the Young and Hölder inequalities to bound this by

$$\begin{aligned}
&\sum_{\mu\lesssim\lambda} |\mu|^{\frac{1}{2}} \|(S_\lambda a_0)(S_\lambda dw_F)\|_{l^{\frac{16}{5}} l^2 (L^{16} L^1)_\mu} \\
&\quad\lesssim \sum_{\mu\lesssim\lambda} |\mu|^{\frac{1}{2}} \|S_\lambda a_0\|_{l^\infty l^2 (L^\infty L^2)_\mu} \|S_\lambda dw_F\|_{l^{\frac{16}{5}} l^\infty (L^{16} L^2)_\mu} \\
&\quad\lesssim \sum_{\mu\lesssim\lambda} \lambda^{\frac{3}{4}+\delta'} \|S_\lambda a_0\|_{l^\infty l^2 (L^\infty L^2)_\mu} |\mu|^{\frac{1}{2}} \|\langle D_x\rangle^{-\frac{3}{4}-\delta'} dw_F\|_{l^{\frac{16}{5}} l^\infty (L^{16} L^2)_\mu} \\
&\quad\lesssim \lambda^{\delta'-\delta} (\log\lambda)\, M .
\end{aligned}$$

□

References

[1] Serge Alinhac. *Blowup for nonlinear hyperbolic equations.* Boston: Birkhäuser, 1995. Progress in Nonlinear Differential Equations and their Applications, 17.

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

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

[4] J. Bony. Calcul symbolique et propogation des singularites pour les equations aux derivees nonlineaires. *Ann. Sci. Ecole Norm.*, 14:209–246, 1981.

[5] R. Coifman and Y. Meyer. Commutateurs d'integrales singulieres et operateurs multilineaires. *Ann. Inst. Fourier Grenobles*, 28:177–202, 1978.

[6] Thomas J. R. Hughes, Tosio Kato, and Jerrold E. Marsden. Well-posed quasi-linear second-order hyperbolic systems with applications to nonlinear elastodynamics and general relativity. *Arch. Rational Mech. Anal.*, 63(3):273–294 (1977), 1976.

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

[8] Tosio Kato and Gustavo Ponce. Commutator estimates and the Euler and Navier-Stokes equations. *Comm. Pure Appl. Math.*, 41(7):891–907, 1988.

[9] Sergiu Klainerman. A commuting vector fields approach to strichartz type inequalities and applications to quasilinear wave equations. IMRN, to appear.

[10] Sergiu Klainerman. Long-time behavior of solutions to nonlinear evolution equations. *Arch. Rational Mech. Anal.*, 78(1):73–98, 1982.

[11] Sergiu Klainerman and Igor Rodnianski. Improved local well posedness for quasilinear wave equations in dimension three. preprint.

[12] Sergiu Klainerman and Daniel Tataru. On the optimal local regularity for Yang-Mills equations in R^{4+1}. *J. Amer. Math. Soc.*, 12(1):93–116, 1999.

[13] Hans Lindblad. Counterexamples to local existence for semi-linear wave equations. *Amer. J. Math.*, 118(1):1–16, 1996.

[14] Hans Lindblad. Counterexamples to local existence for quasilinear wave equations. *Math. Res. Lett.*, 5(5):605–622, 1998.

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

[16] Hart Smith and Daniel Tataru. Counterexamples to Strichartz estimates for the wave equation with nonsmooth coefficients. preprint.

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

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

[19] E.M. Stein. *Singular Integrals and Pseudodifferential Operators.* Princeton University, 1972. Graduate Lecture Notes.

[20] Daniel Tataru. Strichartz estimates for operators with nonsmooth coefficients III. to appear, J. Amer. Math. Soc.

[21] Daniel Tataru. On the equation $\Box u = |\nabla u|^2$ in $5+1$ dimensions. *Math. Res. Lett.*, 6(5-6):469–485, 1999.

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

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

[24] Michael E. Taylor. *Pseudodifferential operators and nonlinear PDE.* Birkhäuser, Boston, 1991.

Department of Mathematics, University of Washington, Seattle

Department of Mathematics, University of California, Berkeley