Title

The time-like minimal surface equation in Minkowski space: low regularity solutions

Permalink

https://escholarship.org/uc/item/11b8n309

Journal

Inventiones Mathematicae, 235(3)

ISSN

0020-9910

Authors

Ai, Albert

Ifrim, Mihaela

Tataru, Daniel

Publication Date

2024-03-01

DOI

10.1007/s00222-023-01231-3

Copyright Information

Peer reviewed

THE TIME-LIKE MINIMAL SURFACE EQUATION IN MINKOWSKI SPACE: LOW REGULARITY SOLUTIONS

ALBERT AI, MIHAELA IFRIM, AND DANIEL TATARU

ABSTRACT. It has long been conjectured that for nonlinear wave equations that satisfy a nonlinear form of the null condition, the low regularity well-posedness theory can be significantly improved compared to the sharp results of Smith-Tataru for the generic case. The aim of this article is to prove the first result in this direction, namely for the time-like minimal surface equation in the Minkowski space-time. Further, our improvement is substantial, namely by 3/8 derivatives in two space dimensions and by 1/4 derivatives in higher dimensions.

CONTENTS

1. INTRODUCTION

The question of local well-posedness for nonlinear wave equations with rough initial data is a fundamental question in the the study of nonlinear waves, and which has received a lot of attention over the years. The result of Smith and Tataru [38], proved almost 20 years ago, provides the sharp regularity threshold for generic nonlinear wave equations in view of Lindblad's counterexample [30]. On the other hand, it has also been conjectured [45] that for nonlinear wave equations that satisfy a suitable nonlinear null condition, the result of [38] can be improved, and the well-posedness threshold can be lowered. In this paper we provide the first result that proves the validity of this conjecture, for a representative equation in this class, namely the hyperbolic minimal surface equation. Further, our improvement turns

1

out to be substantial; precisely, we gain 3/8 derivatives in two space dimensions and 1/4 derivatives in higher dimension. At this regularity level, the Lorentzian metric g in our problem is no better that $C_{x,t}^{\frac{1}{4}+} \cap L_t^2 C_x^{\frac{1}{2}+}$, $(C_{x,t}^{\frac{3}{8}+} \cap L_t^4 C_x^{\frac{1}{2}+}$ in 2d) far below anything studied before.

Most of the ideas introduced in this paper will likely extend to other nonlinear wave models, and open the way toward further progress in the study of low regularity solutions.

1.1. The Minimal Surface Equation in Minkowski Space.

Let $n \geq 2$, and $\mathfrak{M}^{n+2}$ be the $n + 2$ dimensional Minkowski space-time. A codimension one time-like submanifold $\Sigma \subset \mathfrak{M}^{n+2}$ is called a minimal surface if it is locally a critical point for the area functional

$$\mathcal{L} = \int_\Sigma dA,$$

where the area element is measured relative to the Minkowski metric. A standard way to think of this equation is by representing Σ as a graph over $\mathfrak{M}^{n+1}$,

$$\Sigma = \{x_{n+1} = u(t, x_1, \cdots, x_n)\},$$

where u is a real valued function

$$u : D \subset \mathfrak{M}^{n+1} \to \mathbb{R},$$

which satisfies the constraint

time-like (1.1)
$$u_t^2 < 1 + |\nabla_x u|^2,$$

expressing the condition that its graph is a time-like surface in $\mathfrak{M}^{n+2}$.

Then the surface area functional takes the form

L (1.2)
$$\mathcal{L}(u) = \int \sqrt{1 - u_t^2 + |\nabla_x u|^2} \, dx.$$

Interpreting this as a Lagrangian, the minimal surface equation can be thought of as the associated Euler-Lagrange equation, which takes the form

surface eq (1.3)
$$-\frac{\partial}{\partial t}\left(\frac{u_t}{\sqrt{1 - u_t^2 + |\nabla_x u|^2}}\right) + \sum_{i=1}^n \frac{\partial}{\partial x_i}\left(\frac{u_{x_i}}{\sqrt{1 - u_t^2 + |\nabla_x u|^2}}\right) = 0.$$

Under the condition (1.1), the above equation is a quasilinear wave equation.

The left hand side of the last equation can be also interpreted as the mean curvature of the hypersurface Σ, and as such the minimal surface equation is alternatively described as the *zero mean curvature flow*.

In addition to the above geometric interpretation, the minimal surface equation for time-like surfaces in the Minkowski space is also known as the Born-Infeld model in nonlinear electromagnetism [50], as well as a model for evolution of branes in string theory [15].

On the mathematical side, the question of global existence for small, smooth and localized initial data was considered in work of Lindblad [31], Brendle [8], Stefanov [40] and Wong [48]. The stability of a nonflat steady solution, called the catenoid, was studied in [29, 10]. Some blow-up scenarios due to failure of immersivity were investigated by Wong [49]. Minimal surfaces have also been studied as singular limits of certain semilinear wave equations by Jerrard [22]. The local well-posedness question fits into the similar theory for the broader

2

class of quasilinear wave equations, but there is also one result that is specific to minimal surfaces, due to Ettinger [11]; this is discussed later in the paper.

In our study of the minimal surface equation, the above way of representing it is less useful, and instead it is better to think of it in geometric terms. In particular the fact that the above Lagrangian (1.2) and the equation (1.3) are formulated relative to a background Minkowski metric is absolutely non-essential; one may instead use any flat Lorentzian metric. This is no surprise since any two such metrics are equivalent via a linear transformation. Perhaps less obvious is the fact that the equations may be actually written in an identical fashion, independent of the background metric; see Remark 3.1 in Section 3.

For full details on the structure of the equation we refer the reader to Section 3 of the paper, but here we review the most important facts.

The main geometric object is the metric g that is the trace of the Minkowski metric in $\mathfrak{M}^{n+2}$ on Σ, and which, expressed in the $(t = x_0, x_1, \cdots, x_n)$ coordinates, has the form

$$\boxed{\texttt{metric-g}} \qquad (1.4) \qquad g_{\alpha\beta} := m_{\alpha\beta} + \partial_\alpha u \partial_\beta u,$$

where $m_{\alpha\beta}$ denotes the Minkowski metric with signature $(-1, 1, ..., 1)$ in $\mathfrak{M}^{n+1}$. Since Σ is time-like, this is also a Lorentzian metric. This has determinant

$$\boxed{\texttt{def-g}} \qquad (1.5) \qquad g := |\det(g^{\alpha\beta})| = 1 + m^{\alpha\beta} \partial_\alpha u \, \partial_\beta u,$$

and the dual metric is

$$\boxed{\texttt{tric-intro}} \qquad (1.6) \qquad g^{\alpha\beta} := m^{\alpha\beta} - \frac{m^{\alpha\gamma} m^{\beta\delta} \partial_\gamma u \, \partial_\delta u}{1 + m^{\mu\nu} \partial_\mu u \, \partial_\nu u}.$$

Here, and later in the paper, we carefully avoid raising indices with respect to the Minkowski metric. Instead, all raised indices in this paper will be with respect to the metric g.

Relative to this metric, the equation (1.3) can be expressed in the form

$$\boxed{\texttt{msf}} \qquad (1.7) \qquad \Box_g u = 0,$$

where $\Box_g$ is the covariant d'Alembertian, and which in this problem will be shown to have the simple expression

$$(1.8) \qquad \Box_g = g^{\alpha\beta} \partial_\alpha \partial_\beta.$$

An important role will also be played by the associated linearized equation, which, as it turns out, may be easily expressed in divergence form as

$$(1.9) \qquad \partial_\alpha \hat{g}^{\alpha\beta} \partial_\beta v = 0, \qquad \hat{g}^{\alpha\beta} := g^{-\frac{1}{2}} g^{\alpha\beta}.$$

Our objective in this paper will be to study the local well-posedness of the associated Cauchy problem with initial data at $t = 0$,

$$\boxed{\texttt{cp-msf}} \qquad (1.10) \qquad \begin{cases} \Box_g u = 0, \\ u(t = 0) = u_0, \\ u_t(t = 0) = u_1, \end{cases}$$

where the initial data (u_0, u_1) is taken in classical Sobolev spaces,

$$\boxed{\texttt{data}} \qquad (1.11) \qquad u[0] := (u_0, u_1) \in \mathcal{H}^s := H^s \times H^{s-1},$$

and is subject to the constraint

$$\boxed{\texttt{-like-data}} \qquad (1.12) \qquad u_1^2 - |\nabla_x u_0|^2 < 1.$$

Here we use the following notation for the Cauchy data in (1.3) at time t,

$$u[t] := (u(t, \cdot), u_t(t, \cdot)).$$

We aim to investigate the range of exponents s for which local well-posedness holds, and significantly improve the lower bound for this range.

1.2. **Nonlinear wave equations.** The hyperbolic minimal surface equation (1.3) can be seen as a special case of more general quasilinear wave equations, which have the form

(1.13)
$$g^{\alpha\beta}(\partial u)\partial_\alpha\partial_\beta u = N(u, \partial u),$$

where, again, $g^{\alpha\beta}$ is assumed to be Lorentzian, but without any further structural properties. The simplest case is when u is a scalar, real valued function. But one may equally allow u to be a vector-valued function, in which case we think of the left hand side of the equation as being in diagonal form, with the coupling occurring only via g and N. This generic equation will serve as a reference.

As a starting point, we note that the equation (1.3) (and also (1.13) if $N = 0$) admits the scaling law

$$u(t, x) \to \lambda^{-1} u(\lambda t, \lambda x).$$

This allows us to identify the critical Sobolev exponent as

$$s_c = \frac{n+2}{2}.$$

Heuristically, s_c serves as a universal threshold for local well-posedness, i.e. we have to have $s > s_c$. Taking a naive view, one might think of trying to reach the scaling exponent s_c. However, this is a quasilinear wave equation, and getting to s_c has so far proved impossible in any problem of this type.

As a good threshold from above, one might start with the classical well-posedness result, due to Hughes, Kato, and Marsden [18], and which asserts that local well-posedness holds for $s > s_c + 1$. This applies to all equations of the form (1.13), and can be proved solely by using energy estimates. These have the form

(1.14)
$$\|u[t]\|_{\mathcal{H}^s} \lesssim e^{\int_0^t \|\partial^2 u(s)\|_{L^\infty} ds} \|u[0]\|_{\mathcal{H}^s}.$$

They may also be restated in terms of quasilinear energy functionals E^s that have the following two properties:

(a) Coercivity,
$$E^s(u[t]) \approx \|u[t]\|_{\mathcal{H}^s}^2.$$

(b) Energy growth,

(1.15)
$$\frac{d}{dt} E^s(u) \lesssim \|\partial^2 u\|_{L^\infty} \cdot E^s(u).$$

To close the energy estimates, it then suffices to use Sobolev embeddings, which allow one to bound the above L^∞ norm, which we will refer to as a *control parameter*, in terms of the $\mathcal{H}^s$ Sobolev norm provided that $s > \frac{n}{2} + 2$, which is one derivative above scaling.

The reason a derivative is lost in the above analysis is that one would only need to bound $\|\partial^2 u\|_{L^1 L^\infty}$, whereas the norm that is actually controlled is $\|\partial^2 u\|_{L^\infty L^\infty}$; this exactly accounts for the one derivative difference in scaling. It also suggests that the natural way to improve

4

the classical result is to control the $L^p L^\infty$ norm directly. This is indeed possible in the context of the Strichartz estimates, which in dimension three and higher give the bound

$$\|\partial^2 u\|_{L^2 L^\infty} \lesssim \|u[0]\|_{\mathcal{H}^{\frac{n}{2}+\frac{3}{2}}},$$

with another ϵ derivatives loss in three space dimensions. When true, such a bound yields well-posedness for $s > \frac{n+3}{2}$, which is $1/2$ derivatives above scaling. The numerology changes slightly in two space dimensions, where the best possible Strichartz estimate has the form

$$\|\partial^2 u\|_{L^4 L^\infty} \lesssim \|u[0]\|_{\mathcal{H}^{\frac{n}{2}+\frac{7}{4}}}, \quad \text{where } n = 2,$$

which is $3/4$ derivatives above scaling.

The difficulty in using Strichartz estimates is that, while these are well known in the constant coefficient case [12, 24] and even for smooth variable coefficients [23, 33], that is not as simple in the case of rough coefficients. Indeed, as it turned out, the full Strichartz estimates are true for C^2 metrics, see [35] ($n = 2, 3$), [44] (all n), but not, in general, for C^σ metrics when $\sigma < 2$, see the counterexamples of [36, 37]. This difficulty was resolved in two stages:

(i) *Semiclassical time scales and Strichartz estimates with loss of derivatives.* The idea here, which applies even for C^σ metrics with $\sigma < 2$, is that, associated to each dyadic frequency scale 2^k, there is a corresponding " semiclassical " time scale $T_k = 2^{-\alpha k}$, with α dependent on σ, so that full Strichartz estimates hold at frequency 2^k on the scale T_k. Strichartz estimates with loss of derivatives are then obtained by summing up the short time estimates with respect to the time intervals, separately at each frequency. This idea was independently introduced in [5] and [43], and further refined in [4] and [46].

(ii) *Wave packet coherence and parametrices.* The observation here is that in the study of nonlinear wave equations such as (1.13), in addition to Sobolev-type regularity for the metric, we have an additional piece of information, namely that the metric itself can be seen as a solution to a nonlinear wave equation. This idea was first introduced and partially exploited in [26], but was brought to full fruition in [38], where it was shown that almost loss-less Strichartz estimates hold for the solutions to (1.13) at exactly the correct regularity level.

The result in [38] represents the starting point of the present work, and is concisely stated as follows[1]:

Theorem 1.1 (Smith-Tataru [38]). *(1.13) is locally well-posed in $\mathcal{H}^s$ provided that*

$$(1.16) \qquad\qquad s > s_c + \frac{3}{4}, \qquad n = 2,$$

respectively

$$(1.17) \qquad\qquad s > s_c + \frac{1}{2}, \qquad n = 3, 4, 5.$$

---

[1] The primary result in [38] is for the case when $g = g(u)$, but it directly carries over to equations of the form (1.13). The result as stated below applies equally to both cases, but if $g = g(u)$ then s_c is one unit lower.

As part of this result, almost loss-less Strichartz estimates were obtained both directly for the solution u, and more generally for the associated linearized evolution. We will return to these estimates in Section 10 for a more detailed statement and an in-depth discussion.

The optimality of this result, at least in dimension three, follows from work of Lindblad [30], see also the more recent two dimensional result in [34]. However, this counterexample should only apply to "generic" models, and the local well-posedness threshold might possibly be improved in problems with additional structure, i.e. some form of null condition.

Moving forward, we recall that in [45], a null condition was formulated for quasilinear equations of the form (1.13).

Definition 1.2 ([45])**.** *The nonlinear wave equation (1.13) satisfies the* nonlinear null condition *if*

$$\frac{\partial g^{\alpha\beta}(u,p)}{\partial p_\gamma} \xi_\alpha \xi_\beta \xi_\gamma = 0 \qquad in \qquad g^{\alpha\beta}(u,p)\xi_\alpha\xi_\beta = 0. \tag{1.18}$$

In this definition the vector p is a placeholder for the ∂u variable in (1.13); for added generality, we also allow for the dependence of g on the undifferentiated u.

Here we use the terminology "nonlinear null condition" in order to distinguish it from the classical null condition, which is relative to the Minkowski metric, and was heavily used in the study of global well-posedness for problems with small localized data, see [25] as well as the books [39, 17]. In geometric terms, this null condition may be seen as a cancellation condition for the self interactions of wave packets traveling along null geodesics. In Section 3 we verify that the minimal surface equation indeed satisfies the nonlinear null condition.

Further, it was conjectured in [45] that, for problems satisfying (1.18), the local well-posedness threshold can be lowered below the one in [38]. This conjecture has remained fully open until now, though one should mention two results in [27] and [11] for the Einstein equation, respectively the minimal surface equation, where the endpoint in Theorem 1.1 is reached but not crossed.

The present work provides the first positive result in this direction, specifically for the minimal surface equation. Indeed, not only are we able to lower the local well-posedness threshold in Theorem 1.1, but in effect we obtain a substantial improvement, namely by 3/8 derivatives in two space dimensions and by 1/4 derivatives in higher dimension.

1.3. **The main result.** Our main result, stated in a succinct form, is as follows:

Theorem 1.3. *The Cauchy problem for the minimal surface equation (1.10) is locally well-posed for initial data $u[0]$ in $\mathcal{H}^s$ that satisfy the constraint (1.12), where*

$$s > s_c + \frac{3}{8}, \qquad n = 2, \tag{1.19}$$

respectively

$$s > s_c + \frac{1}{4}, \qquad n = 3, 4, 5. \tag{1.20}$$

Remark 1.4. The constraint $n \leq 5$ in this result is inherited from Theorem 1.1, or more precisely its full formulation provided in Theorem 10.1, which also includes the Strichartz estimates. This result is used as a black box in the present paper, so knowledge of Theorem 10.1 in higher dimensions would directly imply that the above result also holds in higher dimensions.

6

The result is valid regardless of the $\mathcal{H}^s$ size of the initial data. Here we interpret local well-posedness in a strong Hadamard sense, including:

- *existence of solutions* in the class $u[\,\cdot\,] \in C([0,T];\mathcal{H}^s)$, with T depending only on the $\mathcal{H}^s$ size of the initial data.
- *uniqueness of solutions*, in the sense that they are the unique limits of smooth solutions.
- *higher regularity*, i.e. if in addition the initial data $u[0] \in \mathcal{H}^m$ with $m > s$, then the solution satisfies $u[\,\cdot\,] \in C([0,T];\mathcal{H}^m)$, with a bound depending only on the $\mathcal{H}^m$ size of the data,

$$\|u[\,\cdot\,]\|_{C([0,T];\mathcal{H}^m)} \lesssim \|u[0]\|_{\mathcal{H}^m}.$$

- *continuous dependence* in $\mathcal{H}^s$, i.e. continuity of the the the data to solution map

$$\mathcal{H}^s \ni u[0] \to u[\,\cdot\,] \in C([0,T];\mathcal{H}^s).$$

- *weak Lipschitz dependence*, i.e. for two $\mathcal{H}^s$ solutions u and v we have the difference bound

$$\|u[\,\cdot\,] - v[\,\cdot\,]\|_{C([0,T];\mathcal{H}^{\frac{1}{2}})} \lesssim \|u[0] - v[0]\|_{\mathcal{H}^{\frac{1}{2}}}$$

where the exponent $\frac{1}{2}$ is replaced by $\frac{5}{8}$ in two space dimensions.

We remark on the weak Lipschitz dependence, which in more classical results is proved for a much larger range of Sobolev exponents. Here the need for balanced estimates, together with a loss of symmetry in the linearized equation, have the effect of limiting this range, namely to a smaller neighbourhood of the exponent $\frac{1}{2}$. For the present results a single exponent suffices.

In addition to the above components of the local well-posedness result, a key intermediate role in the proof of the above theorem is played by the Strichartz estimates, not only for the solution u, but also, more importantly, for the linearized problem

(1.21)
$$\begin{cases} \partial_\alpha \hat{g}^{\alpha\beta} \partial_\beta v = 0, \\ v[0] = (v_0, v_1), \end{cases}$$

as well as its paradifferential counterpart

(1.22)
$$\begin{cases} \partial_\alpha T_{\hat{g}^{\alpha\beta}} \partial_\beta v = 0, \\ v[0] = (v_0, v_1). \end{cases}$$

Here the paraproducts are defined using the Weyl quantization, see Section 2.2 for more details. For later reference, we state the Strichartz estimates in a separate theorem:

Theorem 1.5. *The following properties hold for every solution u to the minimal surface equation as in Theorem 1.3, in the corresponding time interval $[0,T]$:*

a) There exists some $\delta_0 > 0$, depending on s in (1.19), (1.20) so that the solution u satisfies the Strichartz estimates

(1.23)
$$\|\langle D_x\rangle^{\frac{1}{2}+\delta_0}\partial u\|_{L^4 L^\infty} \lesssim 1, \qquad n = 2,$$
$$\|\langle D_x\rangle^{\frac{1}{2}+\delta_0}\partial u\|_{L^2 L^\infty} \lesssim 1, \qquad n = 3, 4, 5.$$

b) Both the linearized equation (1.21) and its paradifferential version (1.22) are well-posed in $\mathcal{H}^{\frac{5}{8}}$ for $n = 2$ respectively $\mathcal{H}^{\frac{1}{2}}$ for $n = 3, 4, 5$, and the following Strichartz estimates hold

7

for each[2] $\delta > 0$:

ong2-intro | (1.24)
$$\|v[\,\cdot\,]\|_{L^\infty \mathcal{H}^{\frac{5}{8}}} + \|\langle D_x \rangle^{-\frac{n}{2} - \frac{1}{4} - \delta} \partial v\|_{L^4 L^\infty} \lesssim \|v[0]\|_{\mathcal{H}^{\frac{5}{8}}} \qquad n = 2,$$

respectively

ong3-intro | (1.25)
$$\|v[\,\cdot\,]\|_{L^\infty \mathcal{H}^{\frac{1}{2}}} + \|\langle D_x \rangle^{-\frac{n}{2} - \frac{1}{4} - \delta} \partial v\|_{L^2 L^\infty} \lesssim \|v[0]\|_{\mathcal{H}^{\frac{1}{2}}}, \qquad n = 3, 4, 5.$$

We note that the Strichartz estimates in both parts (a) and (b) have derivative losses, namely $1/8$ derivatives in the $L^4 L^\infty$ bound in two dimensions, respectively $1/4$ derivatives in higher dimensions. These estimates only represent the tip of the iceberg. One may also consider the inhomogeneous problem, allow source terms in dual Strichartz spaces, etc. These and other variations that play a role in this paper are discussed in Section 4.

To understand the new ideas in the proof of our main theorem, we recall the two key elements of the proof of the result in [38], namely (i) the classical energy estimates (1.15) and (ii) the nearly lossless Strichartz estimates; at the time, the chief difficulty was to prove the Strichartz estimates.

In this paper we completely turn the tables, taking part (ii) above for granted, and instead working to improve the energy estimates. Let us begin with a simple observation, which is that the minimal surface equation (1.7) has a cubic nonlinearity, which allows one to replace (1.15) with

asic-cubic | (1.26)
$$\frac{d}{dt} E^s(u) \lesssim \|\partial u\|_{L^\infty} \|\partial^2 u\|_{L^\infty} \cdot E^s(u).$$

This is what one calls a *cubic energy estimate*, which is useful in the study of long time solutions but does not yet help with the low regularity well-posedness question. The key to progress lies in developing a much stronger form of this bound, which roughly has the form[3]

c-balanced | (1.27)
$$\frac{d}{dt} E^s(u) \lesssim \|\partial u\|^2_{B^{\frac{1}{2}}_{\infty,2}} \cdot E^s(u),$$

where the two control norms on the right are now balanced, and only require $1/2$ derivative less than (1.26). This is what we call a *balanced energy estimate*, which may only hold for a very carefully chosen energy functional E^s.

This is an idea that originates in our recent work on 2D water waves (see [1]), where balanced energy estimates are also used in order to substantially lower the low regularity well-posedness threshold. Going back further, this has its roots in earlier work of the last two authors and their collaborators [20], [19], in the context of trying to apply normal form methods in order to obtain long time well-posedness results in quasilinear problems. There we have introduced what we called the *modified energy method*, which in a nutshell asserts that in quasilinear problems it is far better to modify the energies in a normal form fashion, rather than to transform the equation. It was the cubic energy estimates of [19] that were later refined in [1] to balanced energy estimates. Along the way, we have also borrowed and adapted another idea from Alazard and Delort [3, 2], which is to prepare the problem with a partial normal form transformation, and is a part of their broader concept of paradiagonalization; that same idea is also used here.

---

[2]Of course with an implicit constant that may depend on δ.
[3]See Section 2 for our Besov norm notations.

8

There are several major difficulties in the way of proving energy estimates such as (1.27):

- The normal form structure is somewhat weaker in the case of the minimal surface equation, compared to water waves. As a consequence, we have to carefully understand which components of the equation can be improved with a normal form analysis and which cannot, and thus have to be estimated directly.
- Not only are the energy functionals E^s not explicit, they have to be constructed in a very delicate way, following a procedure that is reminiscent of Tao's renormalization idea in the context of wave-maps [42], as well as the subsequent work [47] of the third author on the same problem.
- Keeping track of symbol regularities in our energy functionals and in the proof of the energy estimates is also a difficult task. To succeed, here we adapt and refine a suitable notion of paracontrolled distributions, an idea that has already been used successfully in the realm of stochastic pde's [13, 14].
- The balanced energy estimates need to be proved not only for the full equation, but also for the associated linear paradifferential equation, as a key intermediate step, as well as for the full linearized flow. In particular, when linearizing, some of the favourable normal form structure (or null structure, to use the nonlinear wave equations language) is lost, and the proofs become considerably more complex.

Finally, the Strichartz estimates of [38] cannot be used directly here. Instead, we are able to reformulate them in a paradifferential fashion, and to apply them on appropriate semiclassical time scales. After interval summation, this leads to Strichartz estimates on the unit time scale but with derivative losses. Precisely, in our main Strichartz estimates, whose aim is to bound the control parameters in (1.27), we end up losing essentially $1/8$ derivatives in two space dimensions, and $1/4$ derivatives in higher dimension. These losses eventually determine the regularity thresholds in our main result in Theorem 1.3.

One consequence of these energy estimates is the following continuation result for the solutions:

Theorem 1.6. *The $\mathcal{H}^s$ solution u given by Theorem 1.3 can be continued for as long as the following integral remains finite:*

$$(1.28) \qquad \int_0^T \|\partial u(t)\|_{B_{\infty,2}^{\frac{1}{2}}}^2 \, dt < \infty.$$

1.4. **An outline of the paper.**

Paraproducts and paradifferential calculus. The bulk of the paper is written in the language of paradifferential calculus. The notations and some of the basic product and paracommutator bounds are introduced in Section 2. Importantly, we use the Weyl quantization throughout; this plays a substantial role as differences between quantizations are not always perturbative in our analysis. Also of note, we emphasize the difference between balanced and unbalanced bounds, so some of our ΨDO product or commutator expansions have the form

$$\text{commutator} = \text{principal part} + \text{unbalanced lower order} + \text{balanced error}.$$

The geometric form of the minimal surface equation. While the flat d'Alembertian may naively appear to play a role in the expansion (1.3) of the minimal surface equation, this is not at all useful, and instead we need to adopt a geometric viewpoint. As a starting point, in

9

Section 3 we consider several equivalent formulations of the minimal surface equation, leading to its geometric form in (1.7). This is based on the metric g associated to the solution u by (1.4), whose dual we also compute. Two other conformally equivalent metrics will also play a role. In the same section we derive the linearized equation, and also introduce the associated linear paradifferential flow.

Strichartz estimates. As explained earlier, Strichartz estimates play a major role in our analysis. These are applied to several equations, namely the full evolution, the linear paradifferential evolution and finally the linearized equation; in the present paper, we view the bounds for the paradifferential equation as the core ones, and the other bounds as derived bounds, though not necessarily in a directly perturbative fashion. The Strichartz estimates admit a number of formulations: in direct form for the homogeneous flow, in dual form for the inhomogeneous one, or in the full form. The aim of Section 4 is to introduce all of these forms of the Strichartz estimates, as well as to describe the relations between them, in the context of this paper. A new idea here is to allow source terms that are time derivatives of distributions in appropriate spaces; this is achieved by reinterpreting the wave equation as a system.

Control parameters in energy estimates. We begin Section 5 by defining the control parameters $\mathcal{A}$, $\mathcal{A}^\sharp$ and $\mathcal{B}$, which will play a fundamental role in our energy estimate. Here $\mathcal{A}$ and $\mathcal{A}^\sharp$ are scale invariant norms, at the level of $\|\partial u\|_{L^\infty}$, which will remain small uniformly in time. $\mathcal{B}$, on the other hand, is time dependent and at the level of $\||D_x|^{\frac{1}{2}}\partial u\|_{L^\infty}$, and will control the energy growth. Typically, our *balanced cubic energy estimates* will have the form

$$\frac{\partial E}{\partial t} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2 E.$$

To propagate energy bounds we will need to know that $\mathcal{B} \in L^2_t$. Also in the same section we prove a number of core bounds for our solutions in terms of the control parameters.

The multiplier method and paracontrolled distributions. Both the construction of our energies and the proof of the energy estimates are based on a paradifferential implementation of the multiplier method, which leads to space-time identities of the form

$$\int_0^T\!\!\!\int_{\mathbb{R}^n} \Box_g u \cdot Xu \, dxdt = E_X(u)|_0^T + \int_0^T\!\!\!\int_{\mathbb{R}^n} R(u) \, dxdt$$

in a paradifferential format, where the vector field X is our multiplier and E_X is its associated energy, while $R(u)$ is the energy flux term which will have to be estimated perturbatively. A fundamental difficulty is that the multiplier X, which should heuristically be at the regularity level of ∂u, cannot be chosen algebraically, and instead has to be constructed in an inductive manner relative to the dyadic frequency scales. In order to accurately quantify the regularity of X, in Section 6 we use and refine the notion of paracontrolled distributions; in a nutshell, while X may not be chosen to be a function of ∂u, it will still have to be paracontrolled by ∂u, which we denote by $X \ll \partial u$.

Energy estimates for the paradifferential equation. The construction of the energy functionals is carried out in Section 7, primarily at the level of the linear paradifferential equation, first in $\mathcal{H}^1$ and then in $\mathcal{H}^s$. In both cases there are two steps: first the construction of the symbol of the multiplier X, as a paracontrolled distribution, and then the proof of the energy estimates. The difference between the two cases is that X is a vector field in the first case, but a full pseudodifferential operator in the second case; because of this, we prefer to present the two arguments separately.

Energy estimates for the full equation. The aim of Section 8 is to prove that balanced cubic energy estimates hold for the full equation in all $\mathcal{H}^s$ spaces with $s \geq 1$. We do this by thinking about the full equation in a paradifferential form, i.e. as a linear paradifferential equation with a nonlinear source term, and then by applying a normal form transformation to the unbalanced part of the source term.

Well-posedness for the linearized equation. The goal of Section 9 is to establish both energy and Strichartz estimates for $\mathcal{H}^{\frac{1}{2}}$ solutions ($\mathcal{H}^{\frac{5}{8}}$ in dimension two) to the linearized equation. This is achieved under the assumption that both energy and Strichartz estimates for $\mathcal{H}^{\frac{1}{2}}$ solutions ($\mathcal{H}^{\frac{5}{8}}$ in dimension two) for the linear paradifferential equation hold. We remark that, while the energy estimates for the linear paradifferential equation have already been established by this point in the paper, the corresponding Strichartz estimates have yet to be proved.

Short time Strichartz estimates for the full equation. The local well-posedness result of Smith and Tataru [38] yields well-posedness and nearly sharp Strichartz estimates on the unit time scale for initial data that is small in the appropriate Sobolev space. Our objective in Section 10 is to recast this result as a short time result for a corresponding large data problem. This is a somewhat standard argument combining scaling and finite speed of propagation, though with an interesting twist due to the need to use homogeneous Sobolev norms.

Small vs. large $\mathcal{H}^s$ data. In our main well-posedness proof, in order to avoid more cumbersome notations and estimates, it is convenient to work with initial data that is small in $\mathcal{H}^s$. This is not a major problem, as this is a nonlinear wave equation which exhibits finite speed of propagation. This allows us to reduce the large data problem to the small data problem by appropriate localizations. This argument is carried out at the beginning of Section 11.

Rough solutions as limits of smooth solutions. Our sequence of modules discussed so far comes together in Section 11, where we finally obtain our rough solutions u as a limit of smooth solutions u^h with initial data frequency localized below frequency 2^h. The bulk of the proof is organized as a bootstrap argument, where the bootstrap quantities are uniform energy type bounds for both u^h and for their increments $v^h = \dfrac{d}{dh} u^h$, which solve the corresponding linearized equation. The main steps are as follows:

- we use the short time Strichartz estimates derived from [38] for u^h and v^h in order to obtain long time Strichartz estimates for u^h, which in turn implies energy estimates for both the full equation and the paradifferential equation, and closes one half of the bootstrap.

- we combine the short time Strichartz estimates and the long time energy estimates for the paradifferential equation in $\mathcal{H}^{\frac{1}{2}}$ ($\mathcal{H}^{\frac{5}{8}}$ if $n = 2$) to obtain long time Strichartz estimates for the same paradifferential equation.
- we use the energy and Strichartz estimates for the paradifferential equation to obtain similar bounds for the linearized equation. This in turn implies long time energy estimates for v^h, closing the second half of the bootstrap loop.

The well-posedness argument. Once we have a complete collection of energy estimates and Strichartz estimates for both the full equation and the linearized equation, we are able to use frequency envelopes in order to prove the remaining part of the well-posedness results, namely the strong convergence of the smooth solutions, the continuous dependence, and the associated uniqueness property. In this we follow the strategy outlined in the last two authors' expository paper [21].

2. Notations, paraproducts and some commutator type bounds

We begin with some standard notations and conventions:

- The greek indices $\alpha, \beta, \gamma, \delta$ etc. in expressions range from 0 to n, where 0 stands for time. Roman indices i, j are limited to the range from 1 to n, and are associated only to spatial coordinates.
- The differentiation operators with respect to all coordinates are ∂_α, $\alpha = 0, ..., n$. By ∂ without any index we denote the full space-time gradient. To separate only spatial derivatives we use the notation ∂_x.
- We consistently use the Einstein summation convention, where repeated indices are summed over, unless explicitly stated otherwise.
- The inequality sign $x \lesssim y$ means $x \leq Cy$ with a universal implicit constant C. If instead the implicit constant C depends on some parameter A then we write $x \lesssim_A y$.

2.1. **Littlewood-Paley decompositions and Sobolev spaces.** We denote the Fourier variables by ξ_α with $\alpha = 0, ..., n$. To separate the spatial Fourier variables we use the notation ξ'.

2.1.1. *Littlewood-Paley decompositions.* For distributions in $\mathbb{R}^n$ we will use the standard inhomogeneous Littlewood-Paley decomposition

$$u = \sum_{k=0}^{\infty} P_k u,$$

where $P_k = P_k(D_x)$ are multipliers with smooth symbols $p_k(\xi')$, localized in the dyadic frequency region $\{|\xi| \approx 2^k\}$ (unless $k = 0$, where we capture the entire unit ball). We emphasize that no such decompositions are used in the paper with respect to the time variable. We will also use the notations $P_{<k}$, $P_{>k}$ with the standard meaning. Often we will use shorthand for the Littlewood-Paley pieces of u, such as $u_k := P_k u$ or $u_{<k} := P_{<k} u$. On occasion we will need multipliers with slightly larger support, e.g. $\tilde{P}_k$ will denote a multiplier with similar dyadic frequency localization as P_k, but so that $\tilde{P}_k P_k = P_k$.

2.1.2. *Function spaces.* For our main evolution we will use inhomogeneous Sobolev space H^s, often combined as product spaces $\mathcal{H}^s = H^s \times H^{s-1}$ for the position/velocity components of our evolution. In the next to last section of the paper only we will have an auxiliary use for the corresponding homogeneous spaces $\dot{H}^s$, in connection with scaling analysis.

For our estimates we will use L^∞ based control norms. In addition to the standard L^∞ norms, in many estimates we will use the standard inhomogeneous BMO norm, as well as its close relatives BMO^s, with norm defined as

$$\|f\|_{BMO^s} = \|\langle D_x \rangle^s f\|_{BMO}.$$

We will also need several related L^p based Besov norms $B^s_{p,q}$, defined as

$$\|u\|^q_{B^s_{p,q}} = \sum_k 2^{qks} \|P_k u\|^q_{L^p}$$

with the obvious changes if $q = \infty$. In particular we will be often using these norms with $p = \infty$ or $p = 2n$, which correspond to the spaces $B^0_{\infty,1}$, $B^{\frac{1}{2}}_{2n,1}$ and $B^{\frac{1}{2}}_{\infty,2}$; these will be used for our control norms $\mathcal{A}$, $\mathcal{A}^\sharp$ and $\mathcal{B}$.

2.1.3. *Frequency envelopes.* Throughout the paper we will use the notion of *frequency envelopes*, introduced by Tao (see for example [42]), which is a very useful device that tracks the evolution of the energy of solutions between dyadic energy shells.

Definition 2.1. *We say that $\{c_k\}_{k \geq 0} \in \ell^2$ is a frequency envelope for a function u in H^s if we have the following two properties:*

a) Energy bound:

(2.1) $$\|P_k u\|_{H^s} \leq c_k,$$

b) Slowly varying

(2.2) $$\frac{c_k}{c_j} \lesssim 2^{\delta|j-k|}, \quad j, k \in \mathbb{N}.$$

Here δ is a positive constant, which is taken small enough in order to account for energy leakage between nearby frequencies.

One can also limit from above the size of a frequency envelope, by requiring that

$$\|u\|^2_{H^s} \approx \sum c_k^2.$$

13

Such frequency envelopes always exist, for instance one can define

$$c_k = \sup_j 2^{-\delta|j-k|} \|P_j u\|_{H^s}.$$

The same notion can be applied to any Besov norms. In particular we will use it jointly for the Besov norms that define our control parameters $\mathcal{A}$, $\mathcal{A}^\sharp$ and $\mathcal{B}$.

2.2. Paraproducts and paradifferential operators.

For multilinear analysis, we will consistently use paradifferential calculus, for which we refer the reader to [6, 32].

We begin with the simplest bilinear expressions, namely products, for which we will use the Littlewood-Paley trichotomy

$$f \cdot g = T_f g + \Pi(f, g) + T_g f,$$

where the three terms capture the *low×high* frequency interactions, the *high×high* frequency interactions and the *low×high* frequency interactions. The paraproduct $T_f g$ might be heuristically thought of as the dyadic sum

$$T_f g = \sum_k f_{<k-\kappa} g_k,$$

where the frequency gap κ can be simply chosen as a universal parameter, say $\kappa = 4$, or on occasion may be increased and used as a smallness parameter in a large data context. To avoid bulky notations, in this paper we will make a harmless abuse of notation and neglect κ altogether. In other words, our notation $P_{<k}$ stands in effect for $P_{<k-\kappa}$ with a fixed universal constant κ.

However, in our context a definition such as the above one is too imprecise, and the difference between usually equivalent choices may have nonperturbative effects when considering adjoints in our proof of balanced energy estimates later on.

In particular, the symmetry properties of T_f as an operator in L^2 are critical in our energy estimates. For this reason, we choose to work with the Weyl quantization, and we define

$$\mathcal{F}(T_f g)(\zeta) = \int_{\xi+\eta=\zeta} \hat{f}(\eta) \chi \left(\frac{|\eta|}{\langle \xi + \frac{1}{2}\eta \rangle} \right) \hat{g}(\xi) \, d\xi.$$

Here χ is a smooth function supported in a small ball and that equals 1 near the origin. With this convention, if f is real then T_f is an L^2 self-adjoint operator.

For paraproducts we have a number of standard bounds which we list below, and we will refer to as Coifman-Meyer estimates:

$$(2.3) \qquad \qquad \|T_f g\|_{L^p} \lesssim \|f\|_{L^\infty} \|g\|_{L^p},$$

$$(2.4) \qquad \qquad \|T_f g\|_{L^p} \lesssim \|f\|_{L^p} \|g\|_{BMO},$$

$$(2.5) \qquad \qquad \|\Pi(f, g)\|_{L^p} \lesssim \|f\|_{L^p} \|g\|_{BMO}.$$

These hold for $1 < p < \infty$, but there are also endpoint results available roughly corresponding to $p = 1$ and $p = \infty$.

Paraproducts may also be thought of as belonging to the larger class of translation invariant bilinear operators. Such operators

$$f, g \to B(f, g)$$

may be described by their symbols $b(\eta, \xi)$ in the Fourier space, by

$$\mathcal{F}B(f,g)(\zeta) = \int_{\xi+\eta=\zeta} b(\eta, \xi) \hat{f}(\eta) \hat{g}(\xi) \, d\xi.$$

A special class of such operators, which we denote by L_{lh}, will play an important role later in the paper:

Definition 2.2. *By L_{lh} we denote translation invariant bilinear forms whose symbol $\ell_{lh}(\eta, \xi)$ is supported in $\{|\eta| \ll |\xi| + 1\}$ and satisfies bounds of the form*

$$|\partial_\eta^i \partial_\xi^j \ell_{lh}(\eta, \xi)| \lesssim \langle \xi \rangle^{-|i|-|j|}.$$

We remark that in particular the bilinear form $B(f, g) = T_f g$ is an operator of type L_{lh}, with symbol

$$b(\eta, \xi) = \chi\left(\frac{|\eta|}{\langle \xi + \frac{1}{2}\eta \rangle}\right).$$

Here the factor in the denominator $\xi + \eta/2$ is the average of the g input frequency and the output frequency, and corresponds exactly to our use of the Weyl calculus. The L^p bounds and the commutator estimates for such bilinear form mirror exactly the similar bounds for paraproducts.

2.3. Commutator and other paraproduct bounds.

Here we collect a number of general paraproduct estimates, which are relatively standard. See for instance Appendix B of [19] and Sections 2 and 3 of [1] for proofs of the following estimates as well as further references.

We begin with the following standard commutator estimate:

Lemma 2.3 (P_k commutators). *We have*

$$(2.6) \qquad \|[T_f, P_k]\|_{\dot{H}^s \to \dot{H}^s} \lesssim 2^{-k} \|\partial f\|_{L^\infty}.$$

A similar bound holds also in L^p for $1 \le p \le \infty$.

The following commutator-type estimates are exact reproductions of statements from Lemmas 2.4 and 2.6 in Section 2 of [1], respectively:

Lemma 2.4 (Para-commutators). *Assume that $\gamma_1, \gamma_2 < 1$. Then we have*

$$(2.7) \qquad \|T_f T_g - T_g T_f\|_{\dot{H}^s \to \dot{H}^{s+\gamma_1+\gamma_2}} \lesssim \||D|^{\gamma_1} f\|_{BMO} \||D|^{\gamma_2} g\|_{BMO},$$

$$(2.8) \qquad \|T_f T_g - T_g T_f\|_{\dot{B}^s_{\infty,\infty} \to \dot{H}^{s+\gamma_1+\gamma_2}} \lesssim \||D|^{\gamma_1} f\|_{L^2} \||D|^{\gamma_2} g\|_{BMO}.$$

A bound similar to (2.7) holds in the Besov scale of spaces, namely from $\dot{B}^s_{p,q}$ to $\dot{B}^{s+\gamma_1+\gamma_2}_{p,q}$ for real s and $1 \le p, q \le \infty$.

Lemma 2.5 (Para-associativity). *For $s + \gamma_2 \ge 0, s + \gamma_1 + \gamma_2 \ge 0$, and $\gamma_1 < 1$ we have*

$$(2.9) \qquad \|T_f \Pi(v, u) - \Pi(v, T_f u)\|_{\dot{H}^{s+\gamma_1+\gamma_2}} \lesssim \||D|^{\gamma_1} f\|_{BMO} \||D|^{\gamma_2} v\|_{BMO} \|u\|_{\dot{H}^s}.$$

15

We also have a Leibniz-type rule with paraproducts, which closely follows Lemma 3.6 of [1]. Here, our setting is slightly cleaner as we have only $T_f \partial_\alpha$ in place of $\partial_t + T_b \partial_\alpha$, and the dependence on f is captured by the control norm $A_{\frac{1}{4}}$ in [1].

Lemma 2.6 (Para-Leibniz rule). *For the balanced Leibniz rule error*

$$E_L^\pi(u,v) = T_f \partial_\alpha \Pi(u,v) - \Pi(T_f \partial_\alpha u, v) - \Pi(u, T_f \partial_\alpha v)$$

we have the bound

$$(2.10) \qquad \|E_L^\pi(u,v)\|_{H^s} \lesssim \|f\|_{BMO^{\frac{1}{2}}} \|u\|_{BMO^{-\frac{1}{2}-\sigma}} \|v\|_{H^{s+\sigma}}, \qquad \sigma \in \mathbb{R}, \quad s \geq 0.$$

The next paraproduct estimate, see Lemma 2.5 in [1], directly relates multiplication and paramultiplication:

Lemma 2.7 (Para-products). *Assume that $\gamma_1, \gamma_2 < 1$, $\gamma_1 + \gamma_2 \geq 0$. Then*

$$(2.11) \qquad \|T_f T_g - T_{fg}\|_{\dot{H}^s \to \dot{H}^{s+\gamma_1+\gamma_2}} \lesssim \||D|^{\gamma_1} f\|_{BMO} \||D|^{\gamma_2} g\|_{BMO}.$$

A similar bound holds in the Besov scale of spaces, namely from $\dot{B}^s_{p,q}$ to $\dot{B}^{s+\gamma_1+\gamma_2}_{p,q}$ for real s and $1 \leq p, q \leq \infty$.

We will also need the following variant, which applies for a different range of indices:

Lemma 2.8 (Low-high para-products). *Assume that $\gamma_1 > 0$, $\gamma_1 + \gamma_2 \leq 0$. Then*

$$(2.12) \qquad \|T_f T_g - T_{T_f g}\|_{\dot{H}^s \to \dot{H}^{s+\gamma_1+\gamma_2}} \lesssim \||D|^{\gamma_1} f\|_{BMO} \||D|^{\gamma_2} g\|_{BMO}, \qquad \gamma_1 + \gamma_2 \neq 0,$$

$$(2.13) \qquad \|T_f T_g - T_{T_f g}\|_{BMO^s \to \dot{H}^{s+\gamma_1+\gamma_2}} \lesssim \||D|^{\gamma_1} f\|_{BMO} \||D|^{\gamma_2} g\|_{L^2}.$$

The proof of this Lemma only requires a straightforward Littlewood-Paley decomposition for both f and g, where the difference $T_f T_g - T_{T_f g}$ selects the range where the f frequency is at least comparable to the g frequency. The details are left for the reader.

These are stated here in the more elegant homogeneous setting, but there are also obvious modifications that apply in the inhomogeneous case. We end with the following Moser-type result:

Lemma 2.9. *Let F be smooth with $F(0) = 0$, and $w \in H^s$. Set*

$$R(w) = F(w) - T_{F'(w)} w.$$

Then we have the estimate

$$(2.14) \qquad \|R(w)\|_{H^{s+\frac{1}{2}}} \lesssim C(\|w\|_{L^\infty}) \|D^{\frac{1}{2}} w\|_{BMO} \|w\|_{H^s}.$$

This should be a classical result, though we were not able to find a sufficiently accurate reference. Instead of providing a proof here, we refer to similar Moser estimates in Lemmas 5.2 and 5.7, which are more relevant to the current paper and which we prove in Section 5. For further reference, the reader may also view this as a variation of Lemma 2.3 in [1].

16

2.4. **Paradifferential operators.** As a generalization of paraproducts, we will also work with paradifferential operators. Precisely, given a symbol $a(x,\xi)$ in $\mathbb{R}^n$, we define its paradifferential Weyl quantization T_a as the operator

$$\mathcal{F}(T_a g)(\zeta) = \int_{\xi+\eta=\zeta} \chi\left(\frac{|\eta|}{\langle \xi + \frac{1}{2}\eta\rangle}\right) \hat{a}(\eta,\xi)\hat{g}(\xi)\,d\xi,$$

where

$$\hat{a}(\eta,\xi) = \mathcal{F}_x a(x,\xi).$$

The simplest class of symbols one can work with is $L^\infty S^m$, which contains symbols a for which

(2.15) $$|\partial_\xi^\alpha a(x,\xi)| \leq c_\alpha \langle\xi\rangle^{m-|\alpha|}$$

for all multi-indices α. For such symbols, the Calderón-Vaillancourt theorem ensures appropriate boundedness in Sobolev spaces,

$$T_a : H^s \to H^{s-m}.$$

We remark that this class of symbols in the paradifferential quantization is contained in the class denoted by $\mathcal{B}S_{1,1}^m$, see for instance Hormander [16] but also the earlier work of Bony [7] for further properties.

More generally, given a translation invariant space of distributions X, we can define an associated symbol class XS^m of symbols with the property that

(2.16) $$\|\partial_\xi^\alpha a(x,\xi)\|_X \leq c_\alpha \langle\xi\rangle^{m-|\alpha|}$$

for each $\xi \in \mathbb{R}^n$. Later in the paper, we will use several choices of symbols of this type, using function spaces that we will associate to our problem.

3. A COMPLETE SET OF EQUATIONS

Here we aim to further describe the minimal surface equation and the underlying geometry, and, in particular, its null structure. We also derive the linearized equation, and introduce the paralinearization of both the main equation and its linearization.

3.1. **The Lorentzian geometry of the minimal surface.** Starting from the expression of the metric g in (1.4), the dual metric is easily computed to be

(3.1) $$g^{\alpha\beta} := m^{\alpha\beta} - \frac{m^{\alpha\gamma}m^{\beta\delta}\partial_\gamma u\partial_\delta u}{1 + m^{\mu\nu}\partial_\mu u\partial_\nu u}.$$

Also associated to the metric g is its determinant

$$g = \det(g_{\alpha\beta}) = \det(g^{\alpha\beta})^{-1},$$

and the associated volume form

$$dV = \sqrt{g}\,dx.$$

This can be easily computed (e.g. using Sylvester's determinant theorem) as

$$g = 1 + m^{\mu\nu}\partial_\mu u\,\partial_\nu u.$$

In the sequel, we will always raise indices with respect to the metric g, never with respect to Minkowski. In particular we will use the standard notation

(3.2) $$\partial^\alpha = g^{\alpha\beta}\partial_\beta.$$

17

We remark that, when applied to the function u, this operator has nearly the same effect as the corresponding Minkowski operator,

(3.3)
$$\partial^\alpha u = \frac{1}{g} m^{\alpha\beta} \partial_\beta u.$$

3.2. **The minimal surface equation.** Here we rewrite the minimal surface equation in covariant form. Using the g notation above and the Minkowski metric, we rewrite (1.3) as
$$m^{\alpha\beta} \partial_\alpha (g^{-\frac{1}{2}} \partial_\beta u) = 0,$$
or equivalently
$$m^{\alpha\beta} (\partial_\alpha \partial_\beta u - \frac{1}{2g} \partial_\alpha g \partial_\beta u) = 0.$$

Expanding the g derivative, we have

(3.4)
$$\partial_\alpha g = 2 m^{\mu\nu} \partial_\mu u \, \partial_\alpha \partial_\nu u.$$

Then in the previous equation we recognize the expression for the dual metric, and the minimal surface equation becomes

(3.5)
$$g^{\alpha\beta} \partial_\alpha \partial_\beta u = 0.$$

Using the notation (3.2), this is written in an even shorter form,

(3.6)
$$\partial^\alpha \partial_\alpha u = 0.$$

Similarly, using also (3.3), the relation (3.4) becomes

(3.7)
$$\frac{1}{2g} \partial_\alpha g = \partial^\nu u \, \partial_\alpha \partial_\nu u.$$

3.3. **The covariant d'Alembertian.** The covariant d'Alembertian associated to the metric g has the form
$$\Box_g = \frac{1}{\sqrt{g}} \partial_\alpha \sqrt{g} g^{\alpha\beta} \partial_\beta,$$
which we can rewrite as
$$\Box_g = \partial_\alpha g^{\alpha\beta} \partial_\beta + \frac{1}{2g} (\partial_\alpha g) g^{\alpha\beta} \partial_\beta$$
$$= g^{\alpha\beta} \partial_\alpha \partial_\beta + \left(\partial_\alpha g^{\alpha\beta} \right) \partial_\beta + \frac{1}{2g} (\partial_\alpha g) g^{\alpha\beta} \partial_\beta.$$

Next we need to compute the two coefficients in round brackets. The second one is given by (3.7). For the first one, for later use, we perform a slightly more general computation where we differentiate $g^{\alpha\beta}(\partial_\gamma u)$ as a function of its arguments $p_\gamma := \partial_\gamma u$,

(3.8)
$$\frac{\partial g^{\alpha\beta}}{\partial p_\gamma} = -\partial^\alpha u \, g^{\beta\gamma} - \partial^\beta u \, g^{\alpha\gamma}.$$

This formula follows by directly differentiating (3.1) and from (3.3),
$$\frac{\partial g^{\alpha\beta}}{\partial p_\gamma} = -m^{\alpha\gamma} \partial^\beta u - m^{\beta\gamma} \partial^\alpha u + 2g \partial^\alpha u \partial^\beta u \, \partial^\gamma u.$$

18

We use (3.1) once again to get (3.8)

$$\frac{\partial g^{\alpha\beta}}{\partial p_\gamma} = -[g^{\alpha\gamma} + g\partial^\alpha u \partial^\gamma u]\partial^\beta u - [g^{\beta\gamma} + g\partial^\gamma u \partial^\beta u]\partial^\alpha u + 2\partial^\alpha u \partial^\beta u g \partial^\gamma u,$$

$$= -g^{\alpha\gamma}\partial^\beta u - g^{\beta\gamma}\partial^\alpha u.$$

From (3.8) and chain rule, we arrive at

d-gab (3.9)
$$\partial_\gamma g^{\alpha\beta} = -\partial^\alpha u \, g^{\beta\delta}\partial_\gamma\partial_\delta u - \partial^\beta u \, g^{\alpha\sigma}\partial_\gamma\partial_\sigma u.$$

Setting $\gamma = \alpha$ and using the minimal surface equation in the (3.5) formulation, we get

div-g (3.10)
$$\partial_\alpha g^{\alpha\beta} = -\partial^\alpha u \, g^{\beta\delta}\partial_\alpha\partial_\delta u.$$

Comparing this with (3.7), we see that the last two terms in the $\Box_g$ expression above cancel, and we obtain the following simplified form for the covariant d'Alembertian:

box-g-cov (3.11)
$$\Box_g = g^{\alpha\beta}\partial_\alpha\partial_\beta.$$

In particular, we get the covariant form of the minimal surface equation for u:

u-cov (3.12)
$$\Box_g u = 0.$$

For later use, we introduce the notation

def-A (3.13)
$$A^\alpha = -\partial_\beta g^{\alpha\beta} = \frac{1}{2g}\partial^\alpha g = \partial^\beta u \, g^{\alpha\delta}\partial_\beta\partial_\delta u.$$

An interesting observation is that from here on, the Minkowski metric plays absolutely no role:

ll-Lorentz **Remark 3.1.** In order to introduce the minimal surface equations we have started from the Minkowski metric $m^{\alpha\beta}$. However, the formulation (3.5) of the equations together with the relations (3.8) provide a complete description of the equations without any reference to the Minkowski metric, and which is in effect valid for any other Lorentzian metric. Indeed, the equation (3.5) together with the fact that the metric components $g^{\alpha\beta}$ are smooth functions of ∂u satisfying (3.8) are all that is used for the rest of the paper. Thus, our results apply equally for any other Lorentzian metric in $\mathbb{R}^{n+2}$.

3.4. **The linearized equations.** Our objective now is to derive the linearized minimal surface equations. We will denote by v the linearized variable. Then, by (3.8), the linearization of the dual metric $g^{\alpha\beta} = g^{\alpha\beta}(u)$ takes the form

$$\delta g^{\alpha\beta} = -\partial^\alpha u \, g^{\beta\nu}\partial_\nu v - \partial^\beta u \, g^{\alpha\sigma}\partial_\sigma v.$$

Then the linearized equation is directly computed, using the symmetry in α and β, as

$$g^{\alpha\beta}\partial_\alpha\partial_\beta v - 2\partial^\alpha u \, g^{\beta\gamma}\partial_\alpha\partial_\beta u \, \partial_\gamma v = 0.$$

Using the expression of A in (3.13), the linearized equations take the form

earizedeqn (3.14)
$$(g^{\alpha\beta}\partial_\alpha\partial_\beta - 2A^\gamma\partial_\gamma)v = 0.$$

Alternatively this may also be written in a divergence form,

earizedeqn (3.15)
$$(\partial_\alpha g^{\alpha\beta}\partial_\beta - A^\gamma\partial_\gamma)v = 0.$$

or in covariant form,

v-cov (3.16)
$$\Box_g v = 2A^\beta\partial_\beta v.$$

19

3.5. **Null forms and the nonlinear null condition.** The primary null form that plays a role in this article is Q_0, defined by

<div style="text-align: right;">`def-Q0`</div> (3.17)
$$Q_0(v, w) := g^{\alpha\beta} \partial_\alpha v \partial_\beta w = \partial_\alpha v \partial^\alpha w.$$

Now, we verify that the nonlinear null condition ($\overset{\text{null}}{1.18}$) holds; for this we use ($\overset{\text{dg-dp}}{3.8}$) to compute

$$\frac{\partial g^{\alpha\beta}}{\partial p_\gamma} \xi_\alpha \xi_\beta \xi_\gamma = \left(-g^{\alpha\gamma} \partial^\beta u - g^{\beta\gamma} \partial^\alpha u \right) \xi_\alpha \xi_\beta \xi_\gamma,$$

which vanishes on the null cone $g^{\alpha\beta} \xi_\alpha \xi_\beta = 0$.

In addition we would like the contribution of A to the linearized equation to be a null form. We get

$$A^\beta \partial_\beta v = \partial^\alpha u \, Q_0(\partial_\alpha u, v).$$

3.6. **Two conformally equivalent metrics.** While the metric g is the primary metric used in this paper, for technical reasons we will also introduce two additional, conformally equivalent metrics, as follows:

(i) The metric $\tilde{g}$ is defined by

<div style="text-align: right;">`def-tg`</div> (3.18)
$$\tilde{g}^{\alpha\beta} := (g^{00})^{-1} g^{\alpha\beta}.$$

Then the minimal surface equation can be written as

<div style="text-align: right;">`ms-tg`</div> (3.19)
$$\tilde{g}^{\alpha\beta} \partial_\alpha \partial_\beta u = 0$$

while the linearized equation, written in divergence form is

<div style="text-align: right;">`ms-tg-v`</div> (3.20)
$$(\partial_\alpha \tilde{g}^{\alpha\beta} \partial_\beta - \tilde{A}^\alpha \partial_\alpha) v = 0,$$

where, still raising indices only with respect to g,

<div style="text-align: right;">`ta-def`</div> (3.21)
$$\tilde{A}^\alpha = (g^{00})^{-1} A^\alpha - \tilde{g}^{\alpha\beta} \partial_\beta (\ln g^{00}) = \partial^\beta u \, \tilde{g}^{\alpha\delta} \partial_\beta \partial_\delta u + 2 \partial^0 u \, \tilde{g}^{0\delta} \tilde{g}^{\alpha\beta} \partial_\beta \partial_\delta u.$$

The main feature of $\tilde{g}$ is that $\tilde{g}^{00} = 1$. Because of this, it will be useful in the study of the linear paradifferential flow, in order to prevent a nontrivial paracoefficient in front of $\partial_0^2 v$ in the equations.

(ii) The metric $\hat{g}$ is defined by

<div style="text-align: right;">`def-hg`</div> (3.22)
$$\hat{g}^{\alpha\beta} = g^{-\frac{1}{2}} g^{\alpha\beta}.$$

Then the minimal surface equation can be written as

(3.23)
$$\hat{g}^{\alpha\beta} \partial_\alpha \partial_\beta u = 0,$$

which is not so useful. Instead, the advantage of using this metric is that, using ($\overset{\text{def-A}}{3.13}$), the linearized equation can now be written in divergence form,

<div style="text-align: right;">`hg-lin`</div> (3.24)
$$\partial_\alpha \hat{g}^{\alpha\beta} \partial_\beta v = 0.$$

This will be very useful when we study the linearized equation in $\mathcal{H}^{\frac{1}{2}}$ (respectively $\mathcal{H}^{\frac{5}{8}}$ in two dimensions).

3.7. **Paralinearization and the linear paradifferential flow.** A key element in our study of the minimal surface equation is the associated linear paradifferential flow, which is derived from the linearized flow (3.15). In inhomogeneous form, this is

$$(3.25) \qquad (\partial_\alpha T_{g^{\alpha\beta}} \partial_\beta - T_{A^\gamma} \partial_\gamma)w = f.$$

Similarly we can write the paradifferential equations associated to $\tilde{g}$, namely

$$(3.26) \qquad (\partial_\alpha T_{\tilde{g}^{\alpha\beta}} \partial_\beta - T_{\tilde{A}^\gamma} \partial_\gamma)w = f$$

as well as $\hat{g}$, which can be written in divergence form:

$$(3.27) \qquad \partial_\alpha T_{\hat{g}^{\alpha\beta}} \partial_\beta w = f.$$

These are all equivalent up to perturbative errors. Accordingly, we introduce the notation

$$(3.28) \qquad T_P = \partial_\alpha T_{g^{\alpha\beta}} \partial_\beta$$

for the paradifferential wave operator as well as its counterparts $T_{\tilde{P}}$ and $T_{\hat{P}}$ with the metric g replaced by $\tilde{g}$, respectively $\hat{g}$.

We will first use the paradifferential equation in the study of the minimal surface problem (3.5), which we rewrite in the form

$$(3.29) \qquad (\partial_\alpha T_{g^{\alpha\beta}} \partial_\beta - T_{A^\gamma} \partial_\gamma)u = N(u).$$

Here we carefully base our formula on the linearized flow (3.25), rather on a direct paradifferential expansion in (3.5). This is in order to insure that all nonlinear interactions in $N(u)$ are frequency balanced at leading order.

A key contention of our paper is that the nonlinearity N plays a perturbative role. However, this has to be interpreted in a more subtle way, in the sense that N becomes perturbative only after a well chosen partial, variable coefficient normal form transformation.

Secondly, we will use it in the study of the linearized minimal surface equation, which we can write in the form

$$(3.30) \qquad \partial_\alpha T_{\hat{g}^{\alpha\beta}} \partial_\beta v = N_{lin}(u)v.$$

Here the nonlinearity N_{lin} will also play a perturbative role, in the same fashion as above. We caution the reader that this is *not* the linearization of N.

4. Energy and Strichartz estimates

Both energy and Strichartz estimates play an essential role in this paper, in various forms and combinations. These are primarily applied first to the linear paradifferential flow, and then to the linearized flow associated to solutions to our main equation (1.7). Our goal here is to provide a brief overview of these estimates.

Importantly, in this section we do not prove any energy or Strichartz estimates. Instead, we simply provide definitions and context for what will be proved later in the paper, and prove a good number of equivalences between various well-posedness statements and estimates. We do this under absolutely minimal assumptions (e.g. boundedness) on the metric g, in order to be able to apply these properties easily later on. In particular there are no commutator bounds needed or used in this section. The structure of the minimal surface equations also plays no role here.

4.1. The equations. For context, here we consider a pseudo-Riemannian metric g in $I \times \mathbb{R}^n$, where $I = [0, T]$ is a time interval of unspecified length. We will make some minimal universal assumptions on the metric g:

- both g and its inverse are uniformly bounded,
- the time slices are uniformly space-like.

Associated to this metric g, we will consider several equations:

The linear paradifferential flow in divergence form:

<div style="float:left; border:1px solid; padding:2px">Tbox-g-div</div>

(4.1) $$\partial_\alpha T_{g^{\alpha\beta}} \partial_\beta v = f, \qquad v[0] = (v_0, v_1).$$

The linear paradifferential flow in non-divergence form:

<div style="float:left; border:1px solid; padding:2px">ox-g-nodiv</div>

(4.2) $$T_{g^{\alpha\beta}} \partial_\alpha \partial_\beta v = f, \qquad v[0] = (v_0, v_1).$$

The linear flow in divergence form:

<div style="float:left; border:1px solid; padding:2px">box-g-div</div>

(4.3) $$\partial_\alpha g^{\alpha\beta} \partial_\beta v = f, \qquad v[0] = (v_0, v_1).$$

The linear flow in non-divergence form:

<div style="float:left; border:1px solid; padding:2px">ox-g-nodiv</div>

(4.4) $$g^{\alpha\beta} \partial_\alpha \partial_\beta v = f \qquad v[0] = (v_0, v_1).$$

Several comments are in order:

- As written, the above evolutions are inhomogeneous. If $f = 0$ then we will refer to them as the *homogeneous* flows.
- In the context of this paper, we are primarily interested in the metric $\hat{g}$, in which case the equation (4.3) represents our main linearized flow, and (4.1) represents our main linear paradifferential flow. The metric g and the nondivergence form of the equations will be used in order to connect our results with the result of Smith-Tataru, which will be used in our proofs.
- One may also add a gradient potential in the equations above; with the gradient potential added there is no difference between the divergence and the non-divergence form of the equations. We omit it in this section, as it plays no role.

We will consider these evolutions in the inhomogeneous Sobolev spaces $\mathcal{H}^s$. In order to do this uniformly, we will assume that $|I| \leq 1$; else using homogeneous spaces would be more appropriate. The exponent s will be an arbitrary real number in the case of the paradifferential flows, but will have a restricted range otherwise.

4.2. Energy estimates and well-posedness for the homogeneous problem. Here we review some relatively standard definitions and facts about local well-posedness.

Definition 4.1. *For any of the above flows in the homogeneous form, we say that they are (forward) well-posed in $\mathcal{H}^s$ in the time interval $I = [0, T]$ if for each initial data $u[0] \in \mathcal{H}^s$ there exists a unique solution u with the property that*

$$u[\cdot] \in C(I; \mathcal{H}^s).$$

This corresponds to a linear estimate of the form

(4.5) $$\|v[\cdot]\|_{L^\infty(I;\mathcal{H}^s)} \lesssim \|v[0]\|_{\mathcal{H}^s}.$$

Sometimes one establishes additional bounds for the solution (e.g. Strichartz estimates) and these are then added in to the class of solutions for which uniqueness is established. We will

comment on this where needed. If no such assumption is used, we call this *unconditional uniqueness*.

For completeness and reference, we now state without proof a classical well-posedness result:

`t:easy-wp` **Theorem 4.2.** *Assume that $\partial g \in L^1(I; L^\infty)$. Then*
a) The paradiffererential flows (4.1) and (4.2) are wellposed in $\mathcal{H}^s$ for all real s.
b) The divergence form evolution (4.3) is well-posed in $\mathcal{H}^s$ for $s \in [0,1]$, and the non-divergence form evolution (4.4) is well-posed in $\mathcal{H}^s$ for $s \in [1,2]$.

We remark that the metrics g associated with the solutions of Smith-Tataru satisfy the above hypothesis, but the ones associated to the solutions in our paper do not.

A slightly stronger form of well-posedness is to assert the existence of a suitable (time dependent) energy functional E^s in $\mathcal{H}^s$:

`d:eest` **Definition 4.3.** *An energy functional for either of the above problems in $\mathcal{H}^s$ is a bounded quadratic form in $\mathcal{H}^s$ that has the following two properties:*

a) Coercivity,

`ee-equiv` (4.6)
$$E^s(v[t]) \approx \|v[t]\|_{\mathcal{H}^s}^2.$$

b) Bounded growth for solutions v to the homogeneous equation,

`ee-hom` (4.7)
$$\frac{d}{dt} E^s(v[t]) \lesssim B(t) \|v[t]\|_{\mathcal{H}^s}^2,$$

where $B \in L^1$ depends only on g.

Later we will also interpret E^s as a symmetric bilinear form in $\mathcal{H}^s$. Such an interpretation is unique.

We remark that, in the context of Theorem 4.2, where $\partial g \in L^1 L^\infty$, an energy functional E^1 corresponding to $s = 1$ is classically obtained by multiplying the equation with a suitable smooth time-like vector field and integrating by parts; we refer the reader to Section 7.2.1 where this procedure is described in greater detail. Then for $s \neq 1$ one simply defines

$$E^s(v[0]) = E^1(\langle D_x \rangle^{s-1} v[0]),$$

and the corresponding control parameter B may be taken as

$$B(t) = \|\partial g(t)\|_{L^\infty}.$$

`s:Duhamel`
4.3. **The wave equation as a system and the inhomogeneous problem.** Switching now to the associated inhomogeneous flows, the classical set-up is to take a source term $f \in L^1 H^{s-1}$, and then look for solutions v in $C(I; \mathcal{H}^s)$ as above. This is commonly done using the *Duhamel principle*, which is most readily applied by rewriting the wave equation as a system. We next describe this process.

A common choice is to write the system for the pair of variables $(v, \partial_t v)$. However, for us it will be more convenient to make a slightly different linear transformation, and use instead the pair

`bv` (4.8)
$$\mathbf{v}(t) := \begin{pmatrix} v(t) \\ g^{0\alpha} \partial_\alpha v(t) \end{pmatrix} := Q \begin{pmatrix} v \\ \partial_t v \end{pmatrix}, \qquad Q = \begin{pmatrix} 1 & 0 \\ g^{0j} \partial_j & g^{00} \end{pmatrix}$$

23

for (4.3) and (4.4), with products replaced by paraproducts in the case of the equation (4.1) or (4.2). For later use, we record the inverse of Q; this is either

Q-inverse (4.9)
$$Q^{-1} = \begin{pmatrix} 1 & 0 \\ -(g^{00})^{-1} g^{0j} \partial_j & (g^{00})^{-1} \end{pmatrix},$$

or its version with products replaced by paraproducts, as needed.

The system for $\mathbf{v}$ will have the form

bv-syst (4.10)
$$\frac{d}{dt} \mathbf{v}(t) = \mathcal{L} \mathbf{v}(t),$$

with the appropriate choice for the matrix operator $\mathcal{L}$. For instance in the case of the homogeneous equation (4.3) we have

calA (4.11)
$$\mathcal{L} = \begin{pmatrix} -(g^{00})^{-1} g^{0j} \partial_j & (g^{00})^{-1} \\ -\partial_i g^{ij} \partial_j + \partial_i g^{i0} (g^{00})^{-1} g^{0j} \partial_j & -\partial_j (g^{00})^{-1} g^{0j} \end{pmatrix},$$

which has the antisymmetry property

A-sym (4.12)
$$\mathcal{L}^* = -J \mathcal{L} J^{-1}, \qquad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

A similar property holds in the non-divergence case, but only for the principal part.

We will always work in settings where Q is bounded and invertible in $\mathcal{H}^s$. This is nearly automatic in the paradifferential case; there we only need to make sure that the operator $T_{g^{00}}$ is invertible. In the differential case we will have to ask that multiplication by g and by $(g^{00})^{-1}$ are bounded in H^{s-1}. In such settings, $\mathcal{H}^s$ well-posedness for our original wave equation and for the associated system are equivalent. If a good energy functional E^s exists for the wave equation, then we may define an associated energy functional for the system by setting

bE (4.13)
$$\mathbf{E}^s(\mathbf{v}(t)) = E^s(Q^{-1} \mathbf{v}(t)).$$

Then the properties (4.6) and (4.7) directly transfer to the homogeneous system (4.10).

If our system is (forward) well-posed in $\mathcal{H}^s$, then solving it generates a (forward) evolution operator $S(t, s)$ which is bounded in $\mathcal{H}^s$ and maps the data at time s to the solution at time t,

$$S(t, s) \mathbf{v}(s) = \mathbf{v}(t).$$

For the system it is easy to consider the inhomogeneous version

syst-inhom (4.14)
$$\frac{d}{dt} \mathbf{v}(t) = \mathcal{L} \mathbf{v}(t) + \mathbf{f}(t).$$

If $\mathbf{f} \in L^1 \mathcal{H}^s$ then the solution to (4.14) is given by Duhamel's formula,

(4.15)
$$\mathbf{v}(t) = S(t, 0) \mathbf{v}(0) + \int_0^t S(t, s) \mathbf{f}(s) \, ds,$$

and satisfies the bound

eb-bf (4.16)
$$\|\mathbf{v}\|_{L^\infty \mathcal{H}^s} \lesssim \|\mathbf{v}(0)\|_{\mathcal{H}^s} + \|\mathbf{f}\|_{L^1 \mathcal{H}^s}.$$

24

If we have a good energy $\mathbf{E}^s$ for the homogeneous system, then Duhamel's formula easily allows us to obtain the corresponding energy estimate for the inhomogeneous one, namely

$$\boxed{\texttt{ee-bf}} \quad (4.17) \qquad \frac{d}{dt}\mathbf{E}^s(\mathbf{v}(t)) \lesssim \mathbf{E}^s(\mathbf{v}(t), \mathbf{f}(t)) + B(t)\|\mathbf{v}(t)\|^2_{\mathcal{H}^s},$$

where the first term on the right arises due to the fact that the energy is quadratic in $\mathbf{v}(t)$.

Now we are ready to return to our original set of equations, add the source term f and reinterpret the above consequences of Duhamel's formula there. As in the homogeneous case, we define $\mathbf{v}(t) := Qv[t]$. Then adding the source term f in the original equation is equivalent to adding a source term $\mathbf{f}$ in the above system. Indeed, it is readily seen that for all our four equations, $\mathbf{f}$ is given by

$$\boxed{\texttt{first-ff}} \quad (4.18) \qquad \mathbf{f}(t) = \begin{pmatrix} 0 \\ f(t) \end{pmatrix}.$$

To complete the correspondence, we note that for such $\mathbf{f}$ we have

$$Q^{-1}\mathbf{f} = (g^{00})^{-1}\mathbf{f}(t).$$

Then we immediately arrive at the following result:

$\boxed{\texttt{n-wp-inhom}}$ **Theorem 4.4.** *a) Assume that either of the homogeneous paradifferential flows ($\overset{\texttt{Tbox-g-div}}{4.1}$) or ($\overset{\texttt{Tbox-g-nodiv}}{4.2}$) are well-posed in $\mathcal{H}^s$. Then the associated inhomogeneous flows are well-posed in $\mathcal{H}^s$ for $f \in L^1 H^{s-1}$, and the following estimate holds*

$$(4.19) \qquad \|v[\cdot]\|_{L^\infty(I;\mathcal{H}^s)} \lesssim \|v[0]\|_{\mathcal{H}^s} + \|f\|_{L^1 H^{s-1}}.$$

In addition, if an energy functional E^s in $\mathcal{H}^s$ exists, then

$$\boxed{\texttt{ee-f}} \quad (4.20) \qquad \frac{d}{dt}E^s(v[t]) \lesssim E^s(v[t], (T_{g^{00}})^{-1}\mathbf{f}(t)) + B(t)\|v[t]\|^2_{\mathcal{H}^s}.$$

b) The same holds for the flows ($\overset{\texttt{box-g-div}}{4.3}$) or ($\overset{\texttt{box-g-nodiv}}{4.4}$) under the additional assumption that multiplication by g and $(g^{00})^{-1}$ is bounded in H^{s-1}, with the paraproduct replaced by the corresponding product.

For our purposes in this paper, we will also need to allow for a larger class of source terms of the form

$$\boxed{\texttt{f+}} \quad (4.21) \qquad f = \partial_t f_1 + f_2.$$

To understand why this is natural, it is instructive to start from the inhomogeneous system ($\overset{\texttt{bv-syst-inhom}}{4.14}$) and argue backward.

Above, we have used the inhomogeneous system in the case where the first component of $\mathbf{f}$ was zero. Now we will allow for both terms in $\mathbf{f}$ to be nonzero, and derive the corresponding wave equation. For clarity we do this in the context of the equation ($\overset{\texttt{box-g-div}}{4.3}$), for which we have computed the corresponding operator $\mathcal{L}$ in ($\overset{\texttt{calA}}{4.11}$); however, a similar computation will apply in all four cases.

Starting with a solution $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2)^\top$ to the inhomogeneous problem ($\overset{\texttt{bv-syst-inhom}}{4.14}$), we begin by defining

$$\boxed{\texttt{new-v}} \quad (4.22) \qquad v(t) := \mathbf{v}_1(t)$$

as our candidate for the wave equation solution. Then the first equation of the system reads

(4.23)
$$\partial_t v = -(g^{00})^{-1}g^{0j}\partial_j v + (g^{00})^{-1}\mathbf{v}_2 + \mathbf{f}_1,$$

or equivalently

$$g^{0\alpha}\partial_\alpha v = \mathbf{v}_2 + g^{00}\mathbf{f}_1.$$

Differentiating this with respect to time we obtain

$$\partial_\beta g^{\beta\alpha}\partial_\alpha v = \partial_j g^{j\alpha}\partial_\alpha v + \partial_t \mathbf{v}_2 + \partial_t g^{00}\mathbf{f}_1$$
$$= \partial_j g^{jk}\partial_k v + \partial_j g^{j0}\partial_t v + \partial_t \mathbf{v}_2 + \partial_t g^{00}\mathbf{f}_1.$$

Finally we substitute $\partial_t v$ from (4.23) and $\partial_t \mathbf{v}_2$ from the second equation of the system. We already know the right hand side should vanish if $\mathbf{f} = 0$, so it suffices to track the $\mathbf{f}$ terms. Then we easily obtain the desired equation for v:

(4.24)
$$\partial_\beta g^{\beta\alpha}\partial_\alpha v = \partial_\alpha g^{\alpha 0}\mathbf{f}_1 + \mathbf{f}_2.$$

Comparing this with (4.21), we obtain the correspondence between the source terms for the wave equation and the system:

(4.25)
$$f_1 = g^{00}\mathbf{f}_1, \qquad f_2 = \partial_k g^{k0}\mathbf{f}_1 + \mathbf{f}_2.$$

We also record here the correspondence between the solutions, in the form

(4.26)
$$\mathbf{v}_1 = v, \qquad \mathbf{v}_2 = g^{0\alpha}\partial_\alpha v - g^{00}\mathbf{f}_1,$$

noting that this is no longer homogeneous, as in (4.8).

The last step in our analysis is to reinterpret the bounds (4.16) and (4.17) in terms of v and f. To do this we make the assumption that multiplication by g and $(g^{00})^{-1}$ is bounded in both H^s and H^{s-1}. Then from (4.16) we get

$$\|v[\cdot]\|_{L^\infty\mathcal{H}^s} \lesssim \|\mathbf{v}\|_{L^\infty\mathcal{H}^s} + \|\mathbf{f}_1\|_{L^\infty H^{s-1}}$$
$$\lesssim \|\mathbf{v}(0)\|_{\mathcal{H}^s} + \|\mathbf{f}\|_{L^1\mathcal{H}^s} + \|\mathbf{f}_1\|_{L^\infty H^{s-1}}$$
$$\lesssim \|v[0]\|_{\mathcal{H}^s} + \|f_1\|_{L^1 H^s \cap L^\infty H^{s-1}} + \|f_2\|_{L^1 H^{s-1}}.$$

Similarly, from (4.17) and (4.13) we obtain the energy bound

(4.27)
$$\frac{d}{dt}E^s(Q^{-1}\mathbf{v}(t)) \lesssim E^s(Q^{-1}\mathbf{v}(t), Q^{-1}\mathbf{f}(t)) + B(t)\|\mathbf{v}(t)\|_{\mathcal{H}^s}^2.$$

Here we use (4.9) and (4.26) to compute

$$Q^{-1}\mathbf{v}(t) = v[t] - \begin{pmatrix} 0 \\ \mathbf{f}_1 \end{pmatrix} = v[t] - \begin{pmatrix} 0 \\ (g^{00})^{-1}f_1 \end{pmatrix} := v[t] - \tilde{v}[t],$$

respectively, using also (4.25),

$$\tilde{f}[t] := Q^{-1}\mathbf{f}(t) = Q^{-1}\begin{pmatrix} (g^{00})^{-1}f_1 \\ f_2 - \partial_k g^{k0}(g^{00})^{-1}f_1 \end{pmatrix}$$

(4.28)
$$= \begin{pmatrix} (g^{00})^{-1}f_1 \\ (g^{00})^{-1}(f_2 - \partial_k g^{k0}(g^{00})^{-1}f_1 - g^{0k}\partial_k(g^{00})^{-1}f_1) \end{pmatrix}.$$

Thus we obtain the following natural extension of Theorem 4.4 above:

26

-wp-inhom+ **Theorem 4.5.** *a)* *Assume that the homogeneous evolution* ($\overset{\text{box-g-nodiv}}{4.4}$) *or* ($\overset{\text{box-g-div}}{4.3}$) *is well-posed in* $\mathcal{H}^s$, *and that multiplication by g and $(g^{00})^{-1}$ is bounded in H^s and in H^{s-1}. Consider either of the two evolutions with a source term f of the form*

$$f = \partial_t f_1 + f_2, \qquad f_1 \in L^1 H^s \cap C H^{s-1}, \qquad f_2 \in L^1 H^{s-1}.$$

Then a unique solution $u \in C(I, \mathcal{H}^s)$ exists. If in addition the homogeneous problem admits an energy functional E^s as in Definition $\overset{\text{d:eest}}{4.3}$ then we have the energy estimate

ee-corect (4.29)
$$\frac{d}{dt} E^s(v[t] - \tilde{v}[t]) \lesssim E^s(v[t] - \tilde{v}[t], \tilde{f}[t]) + B(t)\|v[t] - \tilde{v}[t]\|^2_{\mathcal{H}^s}$$

with $\tilde{v}$ and $\tilde{f}$ defined above and B as in ($\overset{\text{ee-hom}}{4.7}$).

    b) The same result applies for the paradifferential equations ($\overset{\text{Tbox-g-div}}{4.1}$), respectively ($\overset{\text{Tbox-g-nodiv}}{4.2}$), where all instances of g above are replaced by the corresponding paraproducts T_g.

We emphasize here the somewhat unusual function space for f_1, in an intersection of two spaces. This reflects the fact that f_1 has a dual role, both as a source term and as a velocity correction.

We remark that in the situations where we apply this result, the mapping properties for g and $(g^{00})^{-1}$ will be fairly straightforward to verify. In the paradifferential case, for instance, the continuity of g will suffice.

4.4. A duality argument.

Duality plays an important role in many estimates for evolution equations. We will also use duality considerations in this paper for several arguments. We restrict the discussion below to the problems written in divergence form, as this is what we will use later in the paper. However, similar versions may be formulated in the nondivergence case.

At heart, this is based on the following identity, which in the context of the operator $\partial_\alpha g^{\alpha\beta}\partial_\beta$ is written as follows:

dual (4.30)
$$\int_0^T \!\!\!\int_{\mathbb{R}^n} \partial_\alpha g^{\alpha\beta}\partial_\beta v \cdot w - v \cdot \partial_\alpha g^{\alpha\beta}\partial_\beta w \, dx dt = \int g^{0\alpha}\partial_\alpha v \cdot w - v \cdot g^{0\alpha}\partial_\alpha w \, dx \Big|_0^T.$$

This holds for any test functions v and w. The integral on the right can be viewed as a duality relation between $u[t]$ and $v[t]$,

$$B(v[t], w[t]) = \int g^{0\alpha}\partial_\alpha v \cdot w - v \cdot g^{0\alpha}\partial_\alpha w \, dx.$$

Precisely, assuming that $g : H^{s-1} \to H^{s-1}$ as a multiplication operator, and that g^{00} is invertible, this expression has the following two properties

    (1) Boundedness,
$$B : \mathcal{H}^s \times \mathcal{H}^{1-s} \to \mathbb{R}.$$

    (2) Coercivity,
$$\sup_{\|v\|_{\mathcal{H}^{1-s}} \leq 1} B(u, v) \approx \|u\|_{\mathcal{H}^s}.$$

A standard consequence of this relation is the following property:

Proposition 4.6. *The evolutions ($\overset{\text{box-g-div}}{4.3}$), respectively ($\overset{\text{Tbox-g-div}}{4.1}$) are forward well-posed in $\mathcal{H}^s$ iff they are backward well-posed in $\mathcal{H}^{1-s}$.*

We remark that in the context of this paper forward and backward well-posedness are almost identical, so for us this property says that well-posedness in $\mathcal{H}^s$ and $\mathcal{H}^{1-s}$ are equivalent.

The above proposition may be equivalently reformulated as the corresponding result for the system (4.10). It will be more convenient to view it in this context. To do this, we reinterpret the above duality, in terms of the associated system (4.14). In view of the symmetry property (4.12), we have the relation

$$\text{(4.31)} \qquad \int_0^T\!\!\int_{\mathbb{R}^n} (\partial_t - \mathcal{L})\mathbf{v} \cdot J\mathbf{w} - J\mathbf{v} \cdot (\partial_t - \mathcal{L})\mathbf{w}\, dxdt = \int \mathbf{v} \cdot J\mathbf{w}\, dx\, \bigg|_0^T,$$

where the corresponding duality relation is

$$\text{(4.32)} \qquad \mathbf{B}(\mathbf{u},\mathbf{v}) = \int \mathbf{v} \cdot J\mathbf{w}\, dx,$$

which provides the duality between $\mathcal{H}^s$ and $\mathcal{H}^{1-s}$. Incidentally, a consequence of (4.31) is the duality relation

$$S(t,s) = S(s,t)^*,$$

where the duality between $\mathcal{H}^s$ and $\mathcal{H}^{1-s}$ is the one given by the bilinear form $\mathbf{B}$ above. This can be used to construct the backward evolution in $\mathcal{H}^{1-s}$ given the forward evolution in $\mathcal{H}^s$, and vice-versa. The full equivalence argument is standard, and is omitted.

4.5. Strichartz estimates.
Here we discuss several versions of Strichartz estimates, as well as the connection between them.

4.5.1. *Estimates for homogeneous equations.* In the context of this paper, these have the form

$$\text{(4.33)} \qquad \|v\|_{S^r} + \|\partial_t v\|_{S^{r-1}} \lesssim \|v[0]\|_{H^r},$$

where for the Strichartz space S we will consider two different choices:

i) Almost lossless estimates, akin to those established in Smith-Tataru [38]. The corresponding Strichartz norms, denoted by $S = S_{ST}$ are defined as

$$\text{(4.34)} \qquad \begin{aligned} \|v\|_{S_{ST}^r} &= \|v\|_{L^\infty H^r} + \|\langle D_x\rangle^{r-\frac{3}{4}-\delta}v\|_{L^4 L^\infty}, & n &= 2,\\ \|v\|_{S_{ST}^r} &= \|v\|_{L^\infty H^r} + \|\langle D_x\rangle^{r-\frac{n-1}{2}-\delta}v\|_{L^2 L^\infty}, & n &\geq 3. \end{aligned}$$

Here the loss of derivatives is measured by $\delta > 0$, which is an arbitrarily small parameter.

ii) Estimates with derivative losses, precisely the type that will be established in this paper. The corresponding Strichartz norms, denoted by $S = S_{AIT}$ are defined as

$$\text{(4.35)} \qquad \begin{aligned} \|v\|_{S_{AIT}^r} &= \|v\|_{L^\infty H^r} + \|\langle D_x\rangle^{r-\frac{3}{4}-\frac{1}{8}-\delta}v\|_{L^4 L^\infty}, & n &= 2,\\ \|v\|_{S_{AIT}^r} &= \|v\|_{L^\infty H^r} + \|\langle D_x\rangle^{r-\frac{n-1}{2}-\frac{1}{4}-\delta}v\|_{L^2 L^\infty}, & n &\geq 3. \end{aligned}$$

Here $\delta > 0$ is again an arbitrarily small parameter, but we allow for an additional loss of derivatives in the endpoint (Pecher) estimate, namely $1/8$ derivatives in two space dimensions and $1/4$ in higher dimensions.

28

These estimates can be applied to any of the four equations discussed in this section. There are also appropriate counterparts for the corresponding system (4.10), which have the form

| `r-hom-syst` | (4.36) | $$\|\mathbf{v}_1\|_{S^r} + \|\mathbf{v}_2\|_{S^{r-1}} \lesssim \|\mathbf{v}[0]\|_{\mathcal{H}^r}, \qquad S \in \{S_{ST}, S_{AIT}\}.$$ |

Under very mild assumptions on g, these are equivalent to the ones for the corresponding wave equation:

`-vs-bv-hom` **Proposition 4.7.** *The Strichartz estimates (4.33) for the homogeneous wave equation are equivalent to the Strichartz estimates (4.36) for the associated system.*

We also remark on a very mild extension of the estimate (4.33) to the inhomogeneous case. Precisely, if (4.33) holds then we also have the inhomogeneous bound

| `tr-hom-ext` | (4.37) | $$\|v\|_{S^r} + \|\partial_t v\|_{S^{r-1}} \lesssim \|v[0]\|_{\mathcal{H}^r} + \|f\|_{L^1 H^{r-1}}.$$ |

This follows in a straightforward manner by the Duhamel formula, see the discussion in Section 4.3.

We conclude the discussion of the Strichartz estimates for the homogeneous equation with a simple but important case, which will be useful for us in the sequel, and applies in particular to the solutions in [38].

`p:Str-allr` **Proposition 4.8.** *Assume that $\partial g \in L^1 L^\infty$ and that the Strichartz estimates for the homogeneous equation (4.4) hold in $\mathcal{H}^1$. Then the Strichartz estimates for the homogeneous equation hold in $\mathcal{H}^r$ for all $r \in \mathbb{R}$ for both paradifferential flows (4.1) and (4.2).*

We remark that the implicit constant in these Strichartz estimates depends on the implicit constant in the Strichartz estimate in the hypothesis and on the bound for $\|\partial g\|_{L^1 L^\infty}$. Later when we apply this result we will have uniform control over both, so we obtain uniform control over the $\mathcal{H}^r$ Strichartz norm.

Proof. It will be easier to work with the inhomogeneous bound (4.37), as it is more stable with respect to perturbations. We divide the proof into several steps, all of which are relatively standard.

Step 1: We start with the case $r = 1$ with the additional assumption $g^{00} = -1$. Then the second equation in (4.2) can be seen as a perturbation of (4.4) with an $L^1 L^2$ source term. Hence the bound (4.37) for (4.4) implies the same bound for (4.2).

Step 2: Next, assuming still that $g^{00} = -1$, we extend the bound (4.37) for (4.2) to all real Sobolev exponents r by conjugating by $\langle D_x \rangle^\sigma$ with $\sigma = r - 1$, where we can estimate perturbatively the commutator

$$(4.38) \qquad \|[T_{g^{\alpha\beta}}, \langle D_x \rangle^\sigma] \partial_\alpha \partial_\beta \langle D_x \rangle^{-\sigma} v\|_{L^1 L^2} \lesssim \|\partial g\|_{L^1 L^\infty} \|\partial v\|_{L^\infty L^2}.$$

This is valid for all real σ and, since it involves paraproducts, can be thought of as a frequency localized bound, which is but a version of Lemma 2.3.

Step 3: Using a multiplication by $T_{g^{00}}$, we reduce the problem with nonconstant g^{00} to the case when $g^{00} = -1$. Here we apply Lemma 2.7 with $\gamma_1 + \gamma_2 = 1$ for the composition of paraproducts, and then interpolation; this applies equally for all real s. At the conclusion of this step, we have the bound (4.37) for (4.2) for all r.

Step 4: We commute the paracoefficients $T_{g^{\alpha\beta}}$ inside ∂_α perturbatively, in order to obtain the bound (4.37) for (4.1) for all r. $\qquad\square$

29

4.5.2. *Dual Strichartz estimates.* Here one considers the corresponding inhomogeneous problems, with source terms in dual Strichartz spaces. The estimates have the form

$$\boxed{\texttt{Str-dual}} \qquad (4.39) \qquad \|v[\cdot]\|_{L^\infty \mathcal{H}^r} \lesssim \|v[0]\|_{\mathcal{H}^r} + \|f\|_{(S^{1-r})'}, \qquad S \in \{S_{ST}, S_{AIT}\}.$$

Classically, these are obtained by duality from the homogeneous estimates, as follows:

$\boxed{\texttt{p:dual}}$ **Proposition 4.9.** *If the homogeneous estimates* $\overset{\texttt{Str-hom}}{(4.33)}$ *hold in* $\mathcal{H}^r$ *for the forward (backward) evolution then the dual estimates* $\overset{\texttt{Str-dual}}{(4.39)}$ *hold in* $\mathcal{H}^{1-r}$ *for the backward (forward) evolution.*

However, one can do better than this by going instead through the system form of the equations $\overset{\texttt{bv-syst-inhom}}{(4.14)}$. The dual estimates for $\overset{\texttt{bv-syst-inhom}}{(4.14)}$ have the form

$$\boxed{\texttt{-dual-syst}} \qquad (4.40) \qquad \|\mathbf{v}\|_{L^\infty \mathcal{H}^r} \lesssim \|\mathbf{v}(0)\|_{\mathcal{H}^r} + \|\mathbf{f}_1\|_{(S^{-r})'} + \|\mathbf{f}_2\|_{(S^{-r+1})'}, \qquad S \in \{S_{ST}, S_{AIT}\}.$$

These are directly obtained from the homogeneous estimates for the system $\overset{\texttt{bv-syst}}{(4.10)}$ via the duality $\overset{\texttt{dual-system}}{(4.31)}$:

Proposition 4.10. *If the homogeneous estimates hold in* $\mathcal{H}^r$ *for the forward (backward) evolution* $\overset{\texttt{bv-syst}}{(4.10)}$ *then the dual estimates hold in* $\mathcal{H}^{1-r}$ *for the backward (forward) evolution* $\overset{\texttt{bv-syst-inhom}}{(4.14)}$.

One can now further return to the original inhomogeneous equation with a source term as in $\overset{\texttt{f+}}{(4.21)}$, and use the correspondence $\overset{\texttt{f-vs-ff}}{(4.25)}$ and $\overset{\texttt{v-vs-bv}}{(4.26)}$, in order to transfer the dual bounds back. These dual estimates, which represent a generalization of $\overset{\texttt{Str-dual}}{(4.39)}$, have the form

$$\boxed{\texttt{Str-dual+}} \qquad (4.41) \qquad \|v[\cdot]\|_{L^\infty \mathcal{H}^r} \lesssim \|v[0]\|_{\mathcal{H}^r} + \|f_1\|_{L^\infty H^{r-1} \cap (S^{-r})'} + \|f_2\|_{(S^{1-r})'}, \qquad S \in \{S_{ST}, S_{AIT}\}.$$

We obtain the following strengthening of Proposition $\overset{\texttt{p:dual}}{4.9}$:

$\boxed{\texttt{p:dual+}}$ **Proposition 4.11.** *If the homogeneous estimates* $\overset{\texttt{Str-hom}}{(4.33)}$ *hold in* $\mathcal{H}^r$ *for the forward (backward) evolution then the dual estimates* $\overset{\texttt{Str-dual+}}{(4.41)}$ *hold in* $\mathcal{H}^{1-r}$ *for the backward (forward) evolution.*

4.5.3. *Full (retarded) Strichartz estimates.* Here we combine the homogeneous and dual Strichartz estimates in a single bound for the inhomogeneous problem. The classical form is

$$\boxed{\texttt{Str-full}} \qquad (4.42) \qquad \|v\|_{S^r} + \|\partial_t v\|_{S^{r-1}} \lesssim \|v[0]\|_{\mathcal{H}^r} + \|f\|_{(S^{1-r})'}, \qquad S \in \{S_{ST}, S_{AIT}\}.$$

However, here we need to take the extra step where we allow source terms of the form $f = \partial_t f_1 + f_2$, and then the estimates have the form

$$\boxed{\texttt{Str-full+}} \qquad (4.43) \quad \|v\|_{S^r} + \|\partial_t v\|_{S^{r-1}} \lesssim \|v[0]\|_{\mathcal{H}^r} + \|f_1\|_{S^{r-1} \cap (S^{-r})'} + \|f_2\|_{(S^{1-r})'}, \qquad S \in \{S_{ST}, S_{AIT}\}.$$

As we will see, this is closely related to the corresponding bound for the associated inhomogeneous system $\overset{\texttt{bv-syst-inhom}}{(4.14)}$:

$$\boxed{\texttt{-full-syst}} \qquad (4.44) \qquad \|\mathbf{v}_1\|_{S^r} + \|\mathbf{v}_2\|_{S^{r-1}} \lesssim \|\mathbf{v}(0)\|_{\mathcal{H}^r} + \|\mathbf{f}_1\|_{(S^{-r})'} + \|\mathbf{f}_2\|_{(S^{-r+1})'}, \qquad S \in \{S_{ST}, S_{AIT}\}.$$

Our main result here is as follows:

$\boxed{\texttt{ove-around}}$ **Theorem 4.12.** *Consider either the equation* $\overset{\texttt{box-g-div}}{(4.3)}$ *or* $\overset{\texttt{Tbox-g-div}}{(4.1)}$. *If the homogeneous problem is well-posed forward in* $\mathcal{H}^r$ *and backward in* $\mathcal{H}^{1-r}$ *and satisfies the homogeneous Strichartz estimates* $\overset{\texttt{Str-full}}{(4.42)}$ *in both cases, then the solutions to the associated forward inhomogeneous problem with source term* $f = \partial_t f_1 + f_2$ *satisfy the bounds* $\overset{\texttt{Str-full+}}{(4.43)}$.

30

Proof. The proof consists of four steps:

Step 1: If the homogeneous problem is well-posed forward in $\mathcal{H}^r$ and satisfies the homogeneous Strichartz estimates (4.36), then so does the corresponding system, see Proposition 4.7.

Step 2: If the homogeneous problem is well-posed backward in $\mathcal{H}^{1-r}$ and satisfies the homogeneous Strichartz estimates, then so does the corresponding system. By duality, the inhomogeneous system is well-posed forward in H^r and satisfies the dual Strichartz bounds (4.40).

Step 3: We represent the forward $\mathcal{H}^r$ solution by the Duhamel formula

$$\mathbf{v}(t) = S(t,0)\mathbf{v}(0) + \int_0^t S(t,s)\mathbf{f}(s)\,ds.$$

The first term represents the solution to the homogeneous equation, and is estimated by (4.36). For the second term we have two bounds at our disposal: the dual bound where we fix t and estimate the output in $\mathcal{H}^s$ in terms of the input in the dual Strichartz space, and the homogeneous bound where we fix s, set $\mathbf{f}(s) \in \mathcal{H}^r$ and estimate the output as a function of t in the Strichartz space. Concatenating the two, we get the restricted bound

$$(4.45) \qquad \|\mathbf{v}_1\|_{S^r(J)} + \|\mathbf{v}_2\|_{S^{r-1}(J)} \lesssim \|\mathbf{f}_1\|_{(S^{-r})'(I)} + \|\mathbf{f}_2\|_{(S^{-r+1})'(I)}, \qquad S \in \{S_{ST}, S_{AIT}\},$$

where the source $\mathbf{f}$ is supported in an interval I and the output $\mathbf{v}$ is measured in an interval J so that I precedes J. In two dimensions we can now apply the Christ-Kiselev lemma [9] (or the U^p-V^p spaces, see [28]) to get the full estimate. In three and higher dimensions we have a slight problem which is that neither method applies for bounds from L_t^2 to L_t^2. However in our case this is not an issue, because our estimates allow for at least a loss of δ derivatives. Then we can afford to interpolate between the two endpoints and use the Christ-Kiselev lemma for bounds from L_t^{2-} to L_t^{2+} and then return to the endpoint setting by Bernstein's inequality in space and Holder's inequality in time, all at the expense of an arbitrarily small increase in the size δ of the loss.

Step 4. We transfer the estimate (4.44) back to the original system via the correspondence (4.25), (4.26), in order to obtain (4.43). $\qquad\square$

We conclude with a corollary of Theorem 4.12, which will be used later in the paper and follows by combining this result with Proposition 4.8:

Corollary 4.13. *Assume that $\partial g \in L^1 L^\infty$ and that the Strichartz estimates for the homogeneous equation (4.4) hold in $\mathcal{H}^1$. Then the full Strichartz estimates (4.43) hold in $\mathcal{H}^r$ for all $r \in \mathbb{R}$ for both paradifferential flows (4.1) and (4.2).*

5. Control parameters and related bounds

5.1. **Control parameters.** Here we introduce our main control parameters associated to a solution u to the minimal surface equation, which serve to bound the growth of energy for both solutions to the minimal surface flow and for its linearization. We will use three such primary quantities, $\mathcal{A}$, $\mathcal{A}^\sharp$ and $\mathcal{B}$, which are defined as Besov norms of the solution u. Our notations here mirror similar notations in our earlier water wave paper [1].

We begin with $\mathcal{A}$, which is L^∞ based,

$$(5.1) \qquad \mathcal{A} = \sup_{t\in[0,T]} \sum_k \|P_k \partial u\|_{L^\infty}.$$

31

We next define the slightly stronger variant $\mathcal{A}^\sharp \gtrsim \mathcal{A}$, still at the same scaling but L^{2n} based,

<div style="float:left">def-AAs</div> (5.2)
$$\mathcal{A}^\sharp = \sup_{t \in [0,T]} \sum_k 2^{\frac{k}{2}} \|P_k \partial u\|_{L^{2n}}.$$

Here the choice of the exponent $2n$ is in no way essential, though it does provide some minor simplifications in one or two places.

Finally we define the time dependent $\mathcal{B}$ control parameter which is again L^∞ based:

<div style="float:left">def-BB</div> (5.3)
$$\mathcal{B}(t) = \left(\sum_k 2^k \|P_k \partial u\|_{L^\infty}^2 \right)^{\frac{1}{2}}.$$

In a nutshell, the energy functionals we construct later in the paper will be shown to satisfy *cubic balanced* bounds of the form

<div style="float:left">ee-model</div> (5.4)
$$\frac{dE}{dt} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2 E,$$

which guarantee that energy bounds can be propagated for as long as $\mathcal{A}^\sharp$ remains finite and $\mathcal{B}$ remains in L_t^2. One should compare these bounds with the classical energy estimates, which have the form

<div style="float:left">-classical</div> (5.5)
$$\frac{dE}{dt} \lesssim_{\mathcal{A}} \|\partial^2 u\|_{L^\infty} E,$$

and which require an extra half derivative in the control parameter.

We continue with a few comments concerning our choice of control parameters:

- Here $\mathcal{A}$ and $\mathcal{A}^\sharp$ are critical norms for u, which may be described using the Besov notation as capturing the uniform norms in time

$$\mathcal{A} = \|\partial u\|_{L_t^\infty B_{\infty,1}^0}, \qquad \mathcal{A}^\sharp = \|\partial u\|_{L_t^\infty B_{2n,1}^{\frac{1}{2}}}.$$

In a first approximation, the reader should think of $\mathcal{A}$ as simply capturing the L^∞ norm of ∂u; the slightly stronger Besov norm above is needed for minor technical reasons, and allows us to work with scale invariant bounds. Often we will simply rely on the simpler L^∞-bound, since

(5.6)
$$\|\partial u\|_{L^\infty} \lesssim \mathcal{A} \lesssim \mathcal{A}^\sharp.$$

- The control norm $\mathcal{B}$, taken at fixed time, is $1/2$ derivative above scaling, and may also be described using the Besov notation as

$$\mathcal{B}(t) = \|\partial u(t)\|_{B_{\infty,2}^{\frac{1}{2}}}.$$

Again, in a first approximation one should simply think of it as $\|\partial u\|_{BMO^{\frac{1}{2}}}$, which in effect suffices for most of the analysis. Indeed, we have

(5.7)
$$\|\partial u\|_{BMO^{\frac{1}{2}}} \lesssim \mathcal{B}.$$

- Given the choice of these control parameters, it is not difficult to see that our energy estimates of the form (5.4) are invariant with respect to scaling. This by itself does not mean much; even the classical energy estimates, of the form (5.5), are scale invariant, but much less useful for low regularity well-posedness. What is important here is that our energy estimates are *cubic* and *balanced*.

32

- The fact that our control norms are based on uniform, rather than L^2-bounds, particularly at the level of $\mathcal{B}$, is also critical. This is what allows us to use Strichartz estimates to further improve the low regularity well-posedness threshold in our results.
- Concerning the dependence of constants in our estimates on $\mathcal{A}$, $\mathcal{A}^\sharp$ and $\mathcal{B}$, we adopt a two track system:
    - The dependence on $\mathcal{B}$ is either linear or quadratic, and will always be explicitly stated in all estimates.
    - The dependence on $\mathcal{A}$ and $\mathcal{A}^\sharp$ is often nonlinear, in which case we use the notations $\lesssim_{\mathcal{A}}$, $\lesssim_{\mathcal{A}^\sharp}$. This dependence is less important, as beginning with Section 7 we will assume that $\mathcal{A}^\sharp \ll 1$, and drop it altogether except where the smallness is essential. But for clarity and also for later use, we do track this dependence in this and the next section.
- In terms of using $\mathcal{A}$ versus $\mathcal{A}^\sharp$, we first note that ideally we would like to avoid $\mathcal{A}^\sharp$ altogether, and just use the weaker control norm $\mathcal{A}$. But this appears not to be possible, which is why $\mathcal{A}^\sharp$ was introduced. To streamline the analysis, in what follows we will simply think of the implicit dependence as being on $\mathcal{A}^\sharp$, which suffices for our final result. One may even take the more radical step of dropping $\mathcal{A}$ altogether; we decided against that, both for historical reasons and for easy reference.

For bookkeeping reasons we will use a joint frequency envelope $\{c_k\}_k$ for the dyadic components of each of $\mathcal{A}$, $\mathcal{A}^\sharp$, and $\mathcal{B}$, so that

(i) $\{c_k\}_k$ is normalized in ℓ^2 and slowly varying,

$$\sum c_k^2 = 1;$$

(ii) We have control of dyadic Littlewood-Paley pieces as follows for ∂u:

$\boxed{\texttt{fe-control}}$ (5.8) $\qquad \|P_k \partial u\|_{L^\infty} \lesssim \mathcal{A} c_k^2, \qquad 2^{\frac{k}{2}} \|P_k \partial u\|_{L^{2n}} \lesssim \mathcal{A}^\sharp c_k^2, \qquad 2^{\frac{k}{2}} \|P_k \partial u\|_{L^\infty} \lesssim \mathcal{B} c_k.$

A-priori, these frequency envelopes depend on time. However, at the conclusion of the paper, we will see that for our rough solutions they can be taken to be independent of time, essentially equal to appropriate L^2-type frequency envelopes for the initial data.

5.2. **Related bounds.** We will frequently need to use bounds that are similar to (5.8) in nonlinear expressions, so it is convenient to have a notation for the corresponding space:

$\boxed{\texttt{d:CC}}$ **Definition 5.1.** *The space $\mathfrak{C}_0$ is the Banach space of all distributions v that satisfy the bounds*

$\boxed{\texttt{-control-C}}$ (5.9) $\qquad \|v\|_{L^\infty} \leq C, \qquad 2^{\frac{k}{2}} \|P_k v\|_{L^{2n}} \leq C \mathcal{A}^\sharp c_k^2, \qquad 2^{\frac{k}{2}} \|P_k v\|_{L^\infty} \leq C \mathcal{B} c_k,$

with the norm given by the best constant C in the above inequalities.

For this space we have the following algebra and Moser-type result:

$\boxed{\texttt{r-control0}}$ **Lemma 5.2.** *a) The space $\mathfrak{C}_0$ is closed with respect to multiplication and para-multiplication. In particular $\mathfrak{C}_0$ is an algebra.*

b) Let F be a smooth function with $F(0) = 0$, and $v \in \mathfrak{C}_0$. Then $F(v) \in \mathfrak{C}_0$. In particular if $\|v\|_{\mathfrak{C}_0} \lesssim 1$ then $F(v)$ satisfies

$\boxed{\texttt{control-v0}}$ (5.10) $$\|F(v)\|_{\mathfrak{C}_0} \lesssim_{\mathcal{A}^\sharp} \|v\|_{\mathfrak{C}_0}.$$

In particular the above result applies to the metrics g, $\tilde{g}$ and $\hat{g}$, all of which are smooth functions of ∂u, and thus belong to $\mathfrak{C}_0$ modulo constants (which are simply the Minkowski metric).

Proof. a) We first estimate the $\mathfrak{C}_0$ norm for the paraproduct $T_f g$ for $f, g \in \mathfrak{C}_0$. This is straightforward, using the L^∞ bound for f, for all but the uniform bound in (5.9). For the uniform bound, we change the summation order in the Littlewood-Paley expansion to obtain

$$\|T_f g\|_{L^\infty} \lesssim \sum_k \|P_k f\|_{L^\infty} \|P_{>k} g\|_{L^\infty} \lesssim \|f\|_{\mathfrak{C}_0} \|g\|_{L^\infty}.$$

It now remains to estimate $\Pi(f, g)$ in $\mathfrak{C}_0$. The uniform bound is almost identical to the one above. For the $\mathcal{A}^\sharp$ norm we use Bernstein's inequality

$$2^{\frac{k}{2}} \|P_k \Pi(f, g)\|_{L^{2n}} \lesssim \sum_{j \geq k} 2^k \|f_j g_j\|_{L^n} \lesssim \sum_{j \geq k} 2^k \|f_j\|_{L^{2n}} \|g_j\|_{L^{2n}} \lesssim \sum_{j \geq k} 2^{k-j} c_j^2 \mathcal{A}^{\sharp 2} \|f\|_{\mathfrak{C}_0} \|g\|_{\mathfrak{C}_0},$$

and now the j summation is straightforward.

For the $\mathcal{B}$ norm, on the other hand, we estimate

$$\|P_k \Pi(f, g)\|_{L^\infty} \lesssim \sum_{j \geq k} \|f_j g_j\|_{L^\infty} \lesssim \sum_{j \geq k} \|f_j\|_{L^\infty} \|g_j\|_{L^\infty} \lesssim \sum_{j \geq k} 2^{-\frac{j}{2}} c_j \mathcal{A}^\sharp \mathcal{B} \|f\|_{\mathfrak{C}_0} \|g\|_{\mathfrak{C}_0},$$

and again the j summation is straightforward.

b) To prove the Moser inequality we use a continuous Littlewood-Paley decomposition, which leads to the expansion

$$F(v) = F(v_0) + \int_0^\infty F'(v_{<j}) v_j \, dj.$$

To estimate $P_k F(v)$ we consider several cases:

i) $j = k + O(1)$. Then $c_j \approx c_k$, $F'(v_{<j})$ is directly bounded in L^∞ and our bounds are straightforward.

ii) $j < k - 4$. Then we can insert an additional localization,

$$P_k(F'(v_{<j}) v_j) = P_k(\tilde{P}_k F'(v_{<j}) v_j),$$

where we gain from the frequency difference

(5.11)
$$\|P_k F'(v_{<j})\|_{L^\infty} \lesssim 2^{-N(k-j)},$$

which more than compensates for the difference (ratio) between c_j and c_k.

iii) $j > k + 4$. In this case we reexpand $F'(v_{<j})$ and write

$$F'(v_{<j}) v_j = F'(v_0) v_j + \int_0^\infty F''(v_{<l}) v_l v_j \, dl.$$

We further separate into two cases:

(iii.1) $l = j + O(1)$. Then we simply bound $F''(v_{<l})$ in L^∞, and estimate first for the $\mathcal{A}^\sharp$ bound using Bernstein's inequality

$$2^{\frac{k}{2}} \|P_k(F''(v_{<l}) v_l v_j)\|_{L^{2n}} \lesssim 2^k \|v_l v_j\|_{L^n} \lesssim 2^k \|v_l\|_{L^{2n}} \|v_j\|_{L^{2n}} \lesssim 2^{k-j} c_j^2 \mathcal{A}^{\sharp 2},$$

34

where the j and l integrations are trivial. Next we estimate for the $\mathcal{B}$-bound

$$\|P_k(F''(v_{<l})v_l v_j)\|_{L^\infty} \lesssim \|v_l\|_{L^\infty}\|v_j\|_{L^\infty} \lesssim 2^{-\frac{j}{2}}\mathcal{A}^\sharp \mathcal{B}c_j,$$

again with easy j and l integrations.

(iii.2) $l < j - 4$. Then we can insert another frequency localization,

$$P_k(F''(v_{<l})v_l v_j) = P_k(\tilde{P}_j F''(v_{<l})v_l v_j),$$

and repeat the computation in (b.ii) but using (5.11) to account for the difference between l and j.

$\square$

In order to avoid tampering with causality, the Littlewood-Paley projections we use in this paper are purely spatial. This is more of a choice between different evils than a necessity; see for instance the alternate choice made in [38]. A substantial but worthwhile price to pay is that on occasion we will need to separately estimate double time derivatives, in a somewhat imperfect but sufficient fashion.

A good starting point in this direction is to think of bounds for second derivatives of our solution u. If at least one of the derivatives is spatial, then this is straightforward:

$$\text{(5.12)} \qquad \|P_{<k}\partial_x \partial u\|_{L^\infty} \lesssim 2^{\frac{k}{2}}\mathcal{B}c_k.$$

However, matters become more complex if instead we look at the second time derivative of u. The natural idea is to use the main equation (3.5) to estimate $\partial_t^2 u$, by writing in terms of spatial derivatives,

$$\partial_t^2 u = -\sum_{(\alpha,\beta)\neq(0,0)} \tilde{g}^{\alpha\beta}\partial_\alpha \partial_\beta u.$$

If one takes this view, the main difficulties we face are with the high-high interactions in this expression. But these high-high interactions have the redeeming feature that they are balanced, so they will often play a perturbative role. This leads us to define a corrected expression as follows:

Definition 5.3 ("Good" second order derivatives). *We denote by $\widehat{\partial_0 \partial_0}u$ or shortly $\hat{\partial}_t^2 u$ the expression*

$$\text{(5.13)} \qquad \hat{\partial}_t^2 u = \partial_t^2 u + \sum_{(\alpha,\beta)\neq(0,0)} \Pi(\tilde{g}^{\alpha\beta}, \partial_\alpha \partial_\beta u).$$

On the other hand if $(\alpha,\beta) \neq (0,0)$ then we define

$$\widehat{\partial_\alpha \partial_\beta}u := \partial_\alpha \partial_\beta u.$$

With this notation, we have

Lemma 5.4. *Assume that u solves the equation (3.5). Then for its second time derivative we have the decomposition*

$$\text{(5.14)} \qquad \partial_t^2 u = \hat{\partial}_t^2 u + \pi_2(u),$$

where the two components satisfy the uniform bounds

$$\text{(5.15)} \qquad \|P_{<k}\hat{\partial}_t^2 u\|_{L^\infty} \lesssim 2^{\frac{k}{2}}\mathcal{B}c_k, \qquad \|P_{<k}\hat{\partial}_t^2 u\|_{L^\infty} \lesssim 2^k \mathcal{A}c_k^2,$$

35

respectively

`e-dt2u-err` (5.16)
$$\|\pi_2(u)\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2, \qquad \|\pi_2(u)\|_{L^n} \lesssim_{\mathcal{A}^\sharp} \mathcal{A}^{\sharp 2}.$$

One should compare this with the easier direct bound (5.12) for spatial derivatives; the good part $\hat{\partial}_t^2 u$ satisfies a similar bound, but the error $\pi_2(u)$ does not. Later, when such expressions are involved, we will systematically peel off perturbatively the error, and always avoid differentiating it further.

Proof. The main ingredient here is the Littlewood-Paley decomposition. For expository simplicity we prove (5.15) at fixed frequency k. Using the notation in (5.13) we can rewrite equation (3.19) as

`ave decomp` (5.17)
$$\hat{\partial}_t^2 u + \sum_{(\alpha,\beta)\neq(0,0)} T_{\tilde{g}^{\alpha\beta}}\partial_\alpha\partial_\beta u + T_{\partial_\alpha\partial_\beta u}\tilde{g}^{\alpha\beta} = 0.$$

To finish the proof we consider the expression above localized at frequency 2^k, and evaluated in the L^∞-norm

`decomp` (5.18)
$$\|P_k\hat{\partial}_t^2 u\|_{L^\infty} \leq \sum_{(\alpha,\beta)\neq(0,0)} \|P_{<k}(\tilde{g}^{\alpha\beta})P_k(\partial_\alpha\partial_\beta u)\|_{L^\infty} + \|P_{<k}(\partial_\alpha\partial_\beta u)P_k(\tilde{g}^{\alpha\beta})\|_{L^\infty}.$$

We bound each of the terms separately. For the second we use the fact that $\tilde{g}$ is bounded in L^∞, together with the third bound in (5.8), in order to get

$$\|P_{<k}(\partial_\alpha\partial_\beta u)P_k(\tilde{g}^{\alpha\beta})\|_{L^\infty} \leq 2^{\frac{k}{2}}\mathcal{B}c_k, \qquad (\alpha,\beta)\neq(0,0).$$

For the first term we rely on the same procedure, and hence finish the proof of the first bound in (5.15). The second bound in (5.15) has as a starting point the same decomposition in (5.18), only that this time we want to bound the RHS terms using the control norm $\mathcal{A}$. Here we use the first bound in (5.8) and the algebra property of L^∞ to obtain

$$\|P_{<k}(\partial_\alpha\partial_\beta u)P_k(\tilde{g}^{\alpha\beta})\|_{L^\infty} \leq \|P_{<k}(\partial_\alpha\partial_\beta u)\|_{L^\infty}\|P_k(\tilde{g}^{\alpha\beta})\|_{L^\infty} \leq 2^k\mathcal{A}c_k^2, \qquad (\alpha,\beta)\neq(0,0).$$

The last bound to prove is (5.16), where because of the balanced frequencies we can easily even out the derivatives balance and estimate each of the factors using the $\mathcal{B}$ norm. Explicitly, $\tilde{g}^{\alpha\beta}$ is in $\mathfrak{C}_0$ by Lemma 5.2 and hence, we get that for $(\alpha,\beta)\neq(0,0)$:

$$\|\Pi(\tilde{g}^{\alpha\beta}, \partial_\alpha\partial_\beta u)\|_{L^\infty} \leq \sum_k \|P_k(\tilde{g}^{\alpha\beta})P_k(\partial_\alpha\partial_\beta u)\|_{L^\infty}$$

$$\leq \sum_k 2^{-\frac{k}{2}}\|2^{\frac{k}{2}}P_k(\tilde{g}^{\alpha\beta})\|_{L^\infty}2^{\frac{k}{2}}\|2^{-\frac{k}{2}}P_k(\partial_\alpha\partial_\beta u)\|_{L^\infty}$$

$$\lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2,$$

which is the first bound in (5.16). The second bound in (5.16) is similar, but replacing the L^∞ norms with L^{2n} norms. $\qquad\square$

The above lemma motivates narrowing the space $\mathfrak{C}_0$, in order to also include information about $\partial_t v$. For later use, we also define two additional closely related spaces.

`def:CC` **Definition 5.5.** *a) The space $\mathfrak{C}$ is the space of distributions v that satisfy (5.9) and, in addition, $\partial_t v$ admits a decomposition $\partial_t v = w_1 + w_2$ so that*

`dt2u-defCC` (5.19)
$$\|P_k w_1\|_{L^\infty} \leq C2^{\frac{k}{2}}\mathcal{B}c_k, \qquad \|w_2\|_{L^\infty} \leq C\mathcal{B}^2,$$

36

endowed with the norm defined as the best possible constant C in (5.9) and in the above inequality relative to all such possible decompositions.

b) The space $\mathfrak{DC}$ consists of all functions f that admit a decomposition $f = f_1 + f_2$ so that

DC-decomp (5.20)
$$\|P_k f_1\|_{L^\infty} \leq C 2^{\frac{k}{2}} \mathcal{B} c_k, \qquad \|f_2\|_{L^\infty} \leq C \mathcal{B}^2,$$

endowed with the norm defined as the best possible constant C in the above inequality relative to all such possible decompositions.

c) The space $\partial_x \mathfrak{DC}$ consists of functions f that admit a decomposition $f = f_1 + f_2$ so that

dxdc-norm (5.21)
$$\|P_k f_1\|_{L^\infty} \leq C 2^{\frac{3k}{2}} \mathcal{B} c_k, \qquad \|P_k f_2\|_{L^\infty} \leq C 2^k \mathcal{B}^2,$$

endowed also with the corresponding norm.

We remark that, by definition, we have the simple inclusions

inclus (5.22)
$$\mathfrak{C} \subset \mathfrak{C}_0, \qquad \partial : \mathfrak{C} \to \mathfrak{DC}, \qquad \partial_x : \mathfrak{DC} \to \partial_x \mathfrak{DC}.$$

Based on what we have so far, we begin by identifying some elements of these spaces:

:du-in-dcc **Lemma 5.6.** *We have*

fe-utt (5.23)
$$\|\partial u\|_{\mathfrak{C}} \lesssim 1, \qquad \|\partial^2 u\|_{\mathfrak{DC}} \lesssim 1.$$

Proof. The bounds in (5.23) are trivial unless both derivatives are time derivatives, in which case it follows directly from the previous Lemma 5.4. $\qquad\square$

The Moser estimates of Lemma 5.2 may be extended to this setting to include all smooth functions of ∂u:

er-control **Lemma 5.7.** *a) We have the bilinear multiplicative relations*

C-DC (5.24)
$$\mathfrak{C}_0 \cdot \mathfrak{DC} \to \mathfrak{DC}, \qquad T_{\mathfrak{C}_0} \cdot \mathfrak{DC} \to \mathfrak{DC}, \qquad T_{\mathfrak{DC}} \mathfrak{C}_0 \to \mathcal{A}^\sharp \mathfrak{DC},$$

as well as

TDCC-CC (5.25)
$$T_{\mathfrak{DC}} \mathfrak{C}_0 \to \mathcal{B}^2 L^\infty, \qquad \Pi(\mathfrak{DC}, \mathfrak{C}_0) \to \mathcal{B}^2 L^\infty.$$

b) The space $\mathfrak{C}$ is closed under multiplication and para-multiplication; in particular it is an algebra.

c) Let F be a smooth function, and $v \in \mathfrak{C}$. Then $F(v) \in \mathfrak{C}$. In particular if $\|v\|_{\mathfrak{C}} \lesssim 1$ and $F(0) = 0$ then $F(v)$ satisfies

-control-v (5.26)
$$\|F(v)\|_{\mathfrak{C}} \lesssim_{\mathcal{A}^\sharp} \|v\|_{\mathfrak{C}}.$$

d) In addition we also have the paralinearization error bounds

Fv-paralin (5.27)
$$\|P_k R(v)\|_{L^n} \lesssim_{\mathcal{A}^\sharp} 2^{-k} c_k^2 \mathcal{A}^{\sharp 2}, \qquad \|\partial R(v)\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2,$$

where R is as in Lemma 2.9, namely $R(v) = F(v) - T_{F'(v)} v$.

Here part (a) is the main part, after which parts (b) and (c) become immediate improvements of Lemma 5.2. But the new interesting bound is the one in part (d), where, notably, we also bound the time derivative of $R(v)$.

Proof. a) Let $z \in \mathfrak{C}_0$ and $w \in \mathfrak{DC}$ with the decomposition $w = w_1 + w_2$ as in (5.20). We skip the first bound in (5.24), as it is a consequence of the rest of the estimates in (5.24) and (5.25), and first consider the paraproduct $T_z w$. We will bound the contributions of w_1 and w_2 in the same norms as w_1, respectively w_2. Precisely, we have

$$\|P_k T_z w_1\|_{L^\infty} \lesssim \|z\|_{L^\infty} \|\tilde{P}_k w_1\|_{L^\infty} \lesssim \|z\|_{\mathfrak{C}_0} 2^{\frac{k}{2}} \mathcal{B} c_k \|w\|_{\mathfrak{DC}},$$

respectively

$$\|T_z w_2\|_{L^\infty} \lesssim \sum_k \|z_k\|_{L^\infty} \|P_{>k} w_2\|_{L^\infty} \lesssim \mathcal{A}^\sharp \|z\|_{\mathfrak{C}_0} \|w_2\|_{L^\infty}.$$

Next we consider $T_w z$, where we have two choices. The first choice is to use only the $\mathcal{A}^\sharp$ component of the $\mathfrak{C}_0$ norm of z, and prove the last bound in (5.24). Precisely, we have

$$\|P_k T_{w_1} z\|_{L^\infty} \lesssim \|w_{1,<k}\|_{L^\infty} \|\tilde{P}_k z\|_{L^\infty} \lesssim 2^{\frac{k}{2}} \mathcal{B} c_k \|w\|_{\mathfrak{DC}} \cdot \mathcal{A}^\sharp \|z\|_{\mathfrak{C}_0},$$

respectively

$$\|T_{w_2} z\|_{L^\infty} \lesssim \sum_k \|w_{2,<k}\|_{L^\infty} \|P_k z\|_{L^\infty} \lesssim \|w_2\|_{L^\infty} \cdot \mathcal{A}^\sharp \|z\|_{\mathfrak{C}_0}.$$

Alternatively, we can use the $\mathcal{B}$ component of the $\mathfrak{C}_0$ norm of z in the bound for the w_1 component,

$$\|T_{w_1} z\|_{L^\infty} \lesssim \sum_k \|w_{1,<k}\|_{L^\infty} \|z_k\|_{L^\infty} \lesssim \sum_k 2^{\frac{k}{2}} \mathcal{B} c_k \|w\|_{\mathfrak{DC}} \cdot 2^{-\frac{k}{2}} \mathcal{B} c_k \|z\|_{\mathfrak{C}_0} \lesssim \mathcal{B}^2 \|w\|_{\mathfrak{DC}} \|z\|_{\mathfrak{C}_0},$$

and the $\mathcal{A}^\sharp$ component for the w_2 term,

$$\|T_{w_2} z\|_{L^\infty} \lesssim \sum_k \|w_{2,<k}\|_{L^\infty} \|z_k\|_{L^\infty} \lesssim \mathcal{B}^2 \sum_k \mathcal{A}^\sharp c_k^2 \|z\|_{\mathfrak{C}_0} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2 \|w\|_{\mathfrak{DC}} \|z\|_{\mathfrak{C}_0},$$

which leads to the first bound in (5.25).

It remains to consider the second bound in (5.25), where we have

$$\|\Pi(w_1, z)\|_{L^\infty} \lesssim \sum_k \|w_{1,k}\|_{L^\infty} \|z_k\|_{L^\infty} \lesssim \sum_k 2^{\frac{k}{2}} \mathcal{B} c_k \|w\|_{\mathfrak{DC}} \cdot 2^{-\frac{k}{2}} \mathcal{B} c_k \|z\|_{\mathfrak{C}_0} \lesssim \mathcal{B}^2 \|w\|_{\mathfrak{DC}} \|z\|_{\mathfrak{C}_0},$$

respectively

$$\|\Pi(w_2, z)\|_{L^\infty} \lesssim \sum_k \|w_{2,k}\|_{L^\infty} \|z_k\|_{L^\infty} \lesssim \|w_2\|_{L^\infty} \cdot \mathcal{A}^\sharp \|z\|_{\mathfrak{C}_0} \lesssim \mathcal{B}^2 \|w\|_{\mathfrak{DC}} \cdot \mathcal{A}^\sharp \|z\|_{\mathfrak{C}_0}.$$

b) Compared with part (a) of Lemma 5.2, it remains to estimate the time derivative of products and paraproducts. Using Leibniz's rule, this reduces directly to the multiplicative bounds in (a).

c) Compared with part (b) of Lemma 5.2 it remains to estimate

$$\partial_0 F(v) = F'(v) \partial_0 v$$

as in (5.19), which is the same as placing it in $\mathfrak{DC}$. By Lemma 5.2 we have $F'(v) \in \mathfrak{C}_0$, while $\partial_0 v \in \mathfrak{DC}$. Then we can bound the product in $\mathfrak{DC}$ by part (a) of this Lemma.

d) Subtracting the harmless linear part of F, without any loss of generality we can assume that $F'(0) = 0$. We have

$$\partial R = (F'(v) - T_{F'(v)}) \partial v - T_{\partial F'(v)} v = \Pi(F'(v), \partial v) + T_{\partial v} F'(v) - T_{\partial F'(v)} v.$$

38

By ($\overset{\texttt{fe-control-v}}{5.26}$) we can place $F'(v) \in \mathfrak{C}$. Then the $\mathcal{B}^2$ bound follows directly from ($\overset{\texttt{TDCC-CC}}{5.25}$). For the $\mathcal{A}^{\sharp 2}$ bound we can replace ∂ by ∂_x above, and use only $\mathfrak{C}_0$ bounds. Then we can repeat the argument in Lemma $\overset{\texttt{l:Moser-control0}}{5.2}$ (a). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Applying the above lemma shows that for smooth functions F with $F(0) = 0$ we have $F(\partial u) \in \mathfrak{C}$, and in particular all components of the metrics g, $\tilde{g}$ and $\hat{g}$ are in $\mathfrak{C}$ modulo constants. We also have $F(\partial u)\partial^2 u \in \mathfrak{DC}$, which in particular shows that the gradient potentials A and $\tilde{A}$ belong to $\mathfrak{DC}$.

We will use part (d) when $w = \partial u$ and $F = g$, in which case ($\overset{\texttt{Fv-paralin}}{5.27}$) reads

$\boxed{\texttt{for-du-inf}}$ (5.28) $\qquad \|\partial R(\partial u)\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2, \qquad R = g^{\alpha\beta} + T_{\partial^\alpha u g^{\beta\gamma}}\partial_\gamma u + T_{\partial^\beta u g^{\alpha\gamma}}\partial_\gamma u.$

We remark that a similar H^s type bound for the same R is provided by ($\overset{\texttt{est-R}}{2.14}$), namely

$\boxed{\texttt{t-R-for-du}}$ (5.29) $\qquad\qquad\qquad\qquad \|R(\partial u)\|_{H^{s-\frac{1}{2}}} \lesssim_{\mathcal{A}} \mathcal{B}\|\partial u\|_{H^{s-1}}.$

The next lemma provides us with the primary example of elements of the space $\partial_x \mathfrak{DC}$:

$\boxed{\texttt{3u-in-d2cc}}$ **Lemma 5.8.** *We have*

$\boxed{\texttt{fe-uttt}}$ (5.30) $\qquad\qquad\qquad\qquad\qquad \|\partial_\alpha\widehat{\partial_\beta\partial_\gamma}u\|_{\partial_x\mathfrak{DC}} \lesssim_{\mathcal{A}^\sharp} 1.$

Proof. The bound in ($\overset{\texttt{fe-uttt}}{5.30}$) is trivial if at least two derivatives are spatial, and follows from ($\overset{\texttt{fe-utt}}{5.23}$) unless all indices are zero. It remains to consider the case $\alpha = \beta = \gamma = 0$. Here we rely on the earlier decomposition ($\overset{\texttt{wave decomp}}{5.17}$) to which we further apply a ∂_t:

$\boxed{\texttt{dhatdd}}$ (5.31) $\qquad \partial_t\hat{\partial}_t^2 u = -\sum_{(\alpha,\beta)\neq(0,0)} \left(T_{\partial_t\tilde{g}^{\alpha\beta}}\partial_\alpha\partial_\beta u + T_{\tilde{g}^{\alpha\beta}}\partial_t\partial_\alpha\partial_\beta u + T_{\partial_t\partial_\alpha\partial_\beta u}\tilde{g}^{\alpha\beta} + T_{\partial_\alpha\partial_\beta u}\partial_t\tilde{g}^{\alpha\beta}\right).$

We now investigate each term separately, for fixed $(\alpha,\beta) \neq (0,0)$. We begin with the first term, which needs to be bounded in the $\partial_x\mathfrak{DC}$ norm given in ($\overset{\texttt{dxdc-norm}}{5.21}$). We have

$$\|P_k T_{\partial_t\tilde{g}^{\alpha\beta}}\partial_\alpha\partial_\beta u\|_{L^\infty} \lesssim \|P_{<k}(\partial_t\tilde{g}^{\alpha\beta})\|_{L^\infty}\|P_k(\partial_\alpha\partial_\beta u)\|_{L^\infty}.$$

The term that contains the time derivative falling onto the metric will be bounded using the Moser estimate Lemma $\overset{\texttt{l:Moser-control}}{5.7}$. Explicitly, we know that $\tilde{g}^{\alpha\beta}$ is in $\mathfrak{C}$ modulo constants, and due to Lemma $\overset{\texttt{l:Moser-control}}{5.7}$, part c), we get $\partial_t\tilde{g}^{\alpha\beta} \in \mathfrak{DC}$ which allows us to decompose it as in ($\overset{\texttt{DC-decomp}}{5.20}$), $\partial_t\tilde{g}^{\alpha\beta} = \tilde{g}_1^{\alpha\beta} + \tilde{g}_2^{\alpha\beta}$ where

$$\|P_{<k}\partial_t\tilde{g}^{\alpha\beta}\|_{L^\infty} \lesssim \|\tilde{g}_1^{\alpha\beta}\|_{L^\infty} + \|P_{<k}\tilde{g}_2^{\alpha\beta}\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} C\mathcal{B}^2 + C2^{\frac{k}{2}}\mathcal{B}c_k.$$

We now turn to the last bound which we can estimate in two ways using the last part of ($\overset{\texttt{fe-control}}{5.8}$),

$$\|P_k(\partial_\alpha\partial_\beta u)\|_{L^\infty} \lesssim 2^{\frac{k}{2}}\|2^{-\frac{k}{2}}P_k(\partial_\alpha\partial_\beta u)\|_{L^\infty} \lesssim 2^{\frac{k}{2}}\mathcal{B}c_k,$$

respectively the first part of ($\overset{\texttt{fe-control}}{5.8}$)

$$\|P_k(\partial_\alpha\partial_\beta u)\|_{L^\infty} \lesssim 2^k\|2^{-k}P_k(\partial_\alpha\partial_\beta u)\|_{L^\infty} \lesssim 2^k\mathcal{A}c_k^2.$$

Putting together the bounds we have leads to

$$\|P_k T_{\partial_t\tilde{g}^{\alpha\beta}}\partial_\alpha\partial_\beta u\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} C2^k\mathcal{A}\mathcal{B}^2 c_k^2 + C2^k\mathcal{B}^2 c_k^2.$$

We now bound the second term in ($\overset{\texttt{dhatdd}}{5.31}$) as follows:

$$\|P_k T_{\tilde{g}^{\alpha\beta}}\partial_t\partial_\alpha\partial_\beta u\|_{L^\infty} \lesssim \|\tilde{g}^{\alpha\beta}\|_{L^\infty}\|\tilde{P}_k\partial_t\partial_\alpha\partial_\beta u\|_{L^\infty}.$$

39

Here we know that $(\alpha, \beta) \neq (0,0)$ hence there are two cases to consider: (i) we have either $\alpha = 0$ or $\beta = 0$, but not both zero, which overalls means we need to bound $\partial_t^2 \partial_x u$, or (ii) we have both $\alpha, \beta \neq 0$, in which case we need a pointwise bound for $\partial_t \partial_x^2 u$. However, both cases can be handled in the same way if we observe that $\partial_x(\partial_x \partial_t u)$ and $\partial_x(\partial_t^2 u)$ are elements in $\partial_x \mathfrak{D}\mathfrak{C}$: this is a direct consequence of $\partial^2 u \in \mathfrak{D}\mathfrak{C}$ as shown in (5.23), followed by the inclusion in (5.22).

Finally, the third and fourth terms in (5.31) can be treated in the same way the first term in (5.31) was shown to be bounded. $\qquad\square$

We continue with another, slightly more subtle balanced bound:

Lemma 5.9. *For $g, h \in \mathfrak{C}$ define*

$$r = (T_g T_h - T_{gh})\partial^2 u.$$

Then we have the balanced bound

$$\|r\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2 \|g\|_{\mathfrak{C}} \|h\|_{\mathfrak{C}}. \tag{5.32}$$

Proof. For $\partial^2 u$ we use the $\mathfrak{D}\mathfrak{C}$ decomposition as in (5.20),

$$\partial^2 u = f_1 + f_2.$$

We begin with the contribution r_1 of f_1, which we expand as

$$r_1 = \sum_k (T_g T_h - T_{gh}) f_{1,k}.$$

This vanishes unless the frequencies k_1, k_2 of g and h are either
(i) $k_1, k_2 \leq k$ and $\max\{k_1, k_2\} = k + O(1)$, or
(ii) $k_1 = k_2 > k + O(1)$.

Then we use the $\mathcal{A}^\sharp$ component of the $\mathfrak{C}_0$ norm for the lower frequency and and the $\mathcal{B}$ component for the higher frequency to estimate

$$\|r_1\|_{L^\infty} \lesssim \sum_k \|f_{1,k}\|_{L^\infty} \left(\sum_{k_1 < k+O(1)} \|g_{k_1}\|_{L^\infty} \sum_{k_2 = k+O(1)} \|h_{k_2}\|_{L^\infty} \right.$$

$$\left. + \sum_{k_1 = k+O(1)} \|g_{k_1}\|_{L^\infty} \sum_{k_2 < k+O(1)} \|h_{k_2}\|_{L^\infty} + \sum_{k_1 = k_2 \geq k+O(1)} \|g_{k_1}\|_{L^\infty} \|h_{k_2}\|_{L^\infty} \right)$$

$$\lesssim \|g\|_{\mathfrak{C}} \|h\|_{\mathfrak{C}} \sum_k 2^{\frac{k}{2}} \mathcal{B} c_k (\mathcal{A}^\sharp \cdot 2^{-\frac{k}{2}} \mathcal{B} c_k + 2^{-\frac{k}{2}} \mathcal{B} c_k \cdot \mathcal{A}^\sharp + \mathcal{A}^\sharp \cdot 2^{-k} \mathcal{B} c_k)$$

$$\lesssim \mathcal{A}^\sharp \mathcal{B}^2 \|g\|_{\mathfrak{C}} \|h\|_{\mathfrak{C}},$$

as needed.

For the contribution r_2 of f_2 we use a similar expansion, and the first two lines of the estimate above are largely unchanged, except for the use of Bernstein's inequality in case (ii). But now we only use the $\mathcal{A}^\sharp$ component of the $\mathfrak{C}_0$ norm for both k_1 and k_2 frequency.

This leads to

$$\|r_2\|_{L^\infty} \lesssim \sum_k \|f_{2,k}\|_{L^\infty} \left(\sum_{k_1 < k+O(1)} \|g_{k_1}\|_{L^\infty} \sum_{k_2 = k+O(1)} \|h_{k_2}\|_{L^\infty} \right.$$

$$\left. + \sum_{k_1 = k+O(1)} \|g_{k_1}\|_{L^\infty} \sum_{k_2 < k+O(1)} \|h_{k_2}\|_{L^\infty} + \sum_{k_1 = k_2 \geq k+O(1)} 2^k \|g_{k_1}\|_{L^{2n}} \|h_{k_2}\|_{L^{2n}} \right)$$

$$\lesssim \mathcal{B}^2 \|g\|_{\mathfrak{C}} \|h\|_{\mathfrak{C}} \sum_k \mathcal{A}^{\sharp^2} \left(c_k^2 + \sum_{j \geq k} \mathcal{A}^{\sharp^2} 2^{k-j} c_j^2 \right)$$

$$\lesssim \mathcal{A}^{\sharp^2} \mathcal{B}^2 \|g\|_{\mathfrak{C}} \|h\|_{\mathfrak{C}},$$

which again suffices. $\qquad\square$

As already discussed in the introduction, the paradifferential wave operator

(5.33)
$$T_P = \partial_\alpha T_{g^{\alpha\beta}} \partial_\beta,$$

as well as its counterparts $T_{\tilde{P}}$ and $T_{\hat{P}}$ with the metric g replaced by $\tilde{g}$, respectively $\hat{g}$, play an important role in our context.

Throughout the paper, we will interpret various objects related to u as approximate solutions for the T_P equation. We provide several results of this type, where we use our control parameters $\mathcal{A}, \mathcal{B}$ in order to estimate the source term in the paradifferential equation for both u and for its derivatives.

Lemma 5.10. *We have*

(5.34)
$$\|T_P u\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2,$$

as well as the similar bounds for $T_{\tilde{P}}$ and $T_{\hat{P}}$.

Proof. We first prove the bound (5.34), and for this we begin with the paradifferential equation associated to the minimal surface equation (3.5)

$$T_{g^{\alpha\beta}} \partial_\alpha \partial_\beta u + T_{\partial_\alpha \partial_\beta u} g^{\alpha\beta} + \Pi(g^{\alpha,\beta}, \partial_\alpha \partial_\beta u) = 0,$$

and further isolate the part we are interested in estimating

$$\partial_\beta T_{g^{\alpha\beta}} \partial_\alpha u - T_{\partial_\beta g^{\alpha\beta}} \partial_\alpha u + T_{\partial_\alpha \partial_\beta u} g^{\alpha\beta} + \Pi(g^{\alpha\beta}, \partial_\alpha \partial_\beta u) = 0.$$

The estimate we want relies on getting bounds for the following terms

$$\|T_P u\|_{L^\infty} = \|\partial_\beta T_{g^{\alpha\beta}} \partial_\alpha u\|_{L^\infty} \lesssim \|T_{\partial_\beta g^{\alpha\beta}} \partial_\alpha u\|_{L^\infty} + \|T_{\partial_\alpha \partial_\beta u} g^{\alpha\beta}\|_{L^\infty} + \|\Pi(g^{\alpha\beta}, \partial_\alpha \partial_\beta u)\|_{L^\infty}.$$

However, the bounds for all of these terms rely on the use of the fact that $g^{\alpha\beta}$ is in $\mathfrak{C}$ modulo constants, $\partial g^{\alpha\beta}, \partial_\alpha \partial_\beta u \in \mathfrak{DC}$ (consequence of Lemma 5.7), as well as on the bound given by Lemma 5.4. Precisely the estimate (5.25) implies that

$$\|T_{\partial_\beta g^{\alpha\beta}} \partial_\alpha u\|_{L^\infty} + \|T_{\partial_\alpha \partial_\beta u} g^{\alpha\beta}\|_{L^\infty} + \|\Pi(g^{\alpha\beta}, \partial_\alpha \partial_\beta u)\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2.$$

Similar bounds will be obtained for $T_{\hat{P}}$ and $T_{\tilde{P}}$ using the same results mentioned in the proof of bound (5.34) above.

$\qquad\square$

41

We next consider similar bounds for derivatives of u. Here we will differentiate between space and time derivatives. We begin with spatial derivatives:

`l:tp-dxu` **Lemma 5.11.** *We have*

(5.35) $$\|P_{<k}T_P\partial_x u\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} 2^k \mathcal{B}^2,$$

as well as the similar bounds for $T_{\tilde P}$ and $T_{\hat P}$.

Proof. For this proof we rely on the previous Lemma 5.10 and on Lemma 5.4 . This becomes obvious after we commute the ∂_x across the T_P operator

$$P_k T_P \partial_x u = \partial_x P_k \partial_\alpha T_{g^{\alpha\beta}} \partial_\beta u - P_k \partial_\alpha T_{\partial_x g^{\alpha\beta}} \partial_\beta u.$$

The first term on the RHS of the identity above is bounded using (5.34) as follows:

$$\|\partial_x P_k \partial_\alpha T_{g^{\alpha\beta}} \partial_\beta u\|_{L^\infty} \lesssim 2^k \|\partial_\alpha T_{g^{\alpha\beta}} \partial_\beta u\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} 2^k \mathcal{B}^2.$$

Here we took advantage of the operator ∂_x accompanied by the frequency projector P_k. A similar advantage will not present itself for the last term, where we need to distribute the α derivative

$$P_k \partial_\alpha T_{\partial_x g^{\alpha\beta}} \partial_\beta u = P_k T_{\partial_\alpha \partial_x g^{\alpha\beta}} \partial_\beta u + P_k T_{\partial_x g^{\alpha\beta}} \partial_\alpha \partial_\beta u := e_1 + e_2.$$

We bound e_1 using Lemma 5.7, by placing $\partial_x \partial g^{\alpha\beta} \in \partial_x \mathfrak{DC}$ which means it will admit a decomposition as follows

$$\partial_x \partial g^{\alpha\beta} = f_1 + f_2,$$

where

$$\|P_k f_1\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} 2^{\frac{3k}{2}} \mathcal{B} c_k, \quad \|P_k f_2\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} 2^k \mathcal{B}^2.$$

Thus, we get

$$\|e_1\|_{L^\infty} \lesssim (\|P_{<k} f_1\|_{L^\infty} + \|P_{<k} f_2\|_{L^\infty}) \|\tilde P_k \partial_\beta u\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} \left(2^k \mathcal{B}^2 + 2^{\frac{3k}{2}} \mathcal{B} c_k\right) \|\tilde P_k \partial_\beta u\|_{L^\infty},$$

which leads to the desired bound once we estimate the last term accordingly. The bounds can be one of the following

$$\|\tilde P_k \partial_\beta u\|_{L^\infty} \lesssim 2^{-\frac{k}{2}} \|2^{\frac{k}{2}} \tilde P_k \partial_\beta u\|_{L^\infty} \lesssim 2^{-\frac{k}{2}} c_k \mathcal{B} \qquad \text{or} \qquad \|\partial_\beta u\|_{L^\infty} \lesssim c_k^2 \mathcal{A}.$$

For the first term in the bracket we use the control norm $\mathcal{A}$, and for the second term we use the $\mathcal{B}$ norm bound.

For e_2 we use the decomposition in Lemma 5.4 for $\partial_\alpha \partial_\beta u$ and for g we use the fact that g is in $\mathfrak{C}_0$ modulo constants, where we can use either the $\mathcal{A}$ bound or the $\mathcal{B}$ bound. The computations are similar to the case of e_1.

The bounds for $T_{\tilde P}$ and $T_{\hat P}$ follow from the exact argument as the one used above in the T_P case.

$\square$

`l:tp-du` **Lemma 5.12.** *a) We have*

(5.36) $$T_P \partial u \in \partial(\mathcal{B}^2 L^\infty),$$

i.e. there exists a representation

(5.37) $$T_P \partial u = \partial_\alpha f^\alpha, \qquad |f| \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2.$$

42

b) We also have

eq:5.12   (5.38) $$\|P_k\partial_\alpha T_{g^{\alpha\beta}}\widehat{\partial_\beta\partial_\gamma}u\|_{L^\infty} + \|P_k\partial_\gamma T_{g^{\alpha\beta}}\widehat{\partial_\beta\partial_\alpha}u\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} 2^k\mathcal{B}^2.$$

Similar results hold with g replaced by $\tilde{g}$ or $\hat{g}$.

Proof. a) We write
$$\partial_\alpha T_{g^{\alpha\beta}}\partial_\beta\partial u = \partial\partial_\alpha T_{g^{\alpha\beta}}\partial_\beta u - \partial_\alpha T_{\partial g^{\alpha\beta}}\partial_\beta u.$$
Here for the first term we use Lemma 5.10, while for the second, by (5.25), we have
$$\|T_{\partial g}\partial u\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2.$$

b) The first step here is to reduce to the case of the metric $\tilde{g}$. Each of the other two metrics may be written in the form $h\tilde{g}$, with $h = h(\partial u)$. Then we can write

expresion   (5.39) $$\partial_\alpha T_{h\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_\gamma}u = T_h\partial_\alpha T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_\gamma}u - T_{\partial_\alpha h}T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_\gamma}u + \partial_\alpha(T_{h\tilde{g}^{\alpha\beta}} - T_hT_{\tilde{g}^{\alpha\beta}})\widehat{\partial_\beta\partial_\gamma}u.$$

The first term corresponds to our reduction, and the remaining terms need to be estimated perturbatively. This is straightforward unless $\alpha = 0$, so we focus now on this case. Discarding constants, we can assume here that $h(0) = 0$.

For the middle term in (5.39) we can use the bound (5.26) in Lemma 5.7 to place h in $\mathfrak{C}$. Then $\partial_0 h$ is in $\mathfrak{D}\mathfrak{C}$. Using a $\mathfrak{D}\mathfrak{C}$ decomposition for it, $\partial_0 h = h_1 + h_2$, we can match the two terms with the two pointwise bounds for $\widehat{\partial_\beta\partial_\gamma}u$, namely
$$\|P_k\widehat{\partial_\beta\partial_\gamma}u\|_{L^\infty} \lesssim 2^k\mathcal{A}c_k^2, \qquad \|P_k\widehat{\partial_\beta\partial_\gamma}u\|_{L^\infty} \lesssim 2^{\frac{k}{2}}\mathcal{B}c_k,$$
which follow from (5.8) and (5.15). This yields
$$\|P_k T_{\partial_\alpha h}T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_\gamma}u\|_{L^\infty} \lesssim \|h_1\|_{L^\infty} \cdot \|\tilde{P}_k\widehat{\partial_\beta\partial_\gamma}u\|_{L^\infty} + \|h_{2,<k}\|_{L^\infty} \cdot \|\tilde{P}_k\widehat{\partial_\beta\partial_\gamma}u\|_{L^\infty}$$
$$\lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2 \cdot 2^k\mathcal{A}c_k^2 + 2^{\frac{k}{2}}\mathcal{B}c_k \cdot 2^{\frac{k}{2}}\mathcal{B}c_k,$$
which suffices.

For the last expression in (5.39) we distribute ∂_0,
$$\partial_0(T_{h\tilde{g}^{0\beta}} - T_hT_{\tilde{g}^{0\beta}})\widehat{\partial_\beta\partial_\gamma}u = (T_{h\tilde{g}^{0\beta}} - T_hT_{\tilde{g}^{0\beta}})\partial_0\widehat{\partial_\beta\partial_\gamma}u + (T_{\partial_0 h\tilde{g}^{0\beta}} - T_{\partial_0 h}T_{\tilde{g}^{0\beta}})\widehat{\partial_\beta\partial_\gamma}u$$
$$+ (T_{h\partial_0\tilde{g}^{0\beta}} - T_hT_{\partial_0\tilde{g}^{0\beta}})\partial_0\widehat{\partial_\beta\partial_\gamma}u.$$
For the first term we can combine the bound (5.30) with the para-composition bound in Lemma 2.4 exactly as in the proof of Lemma 5.9. For the second term we use the same $\mathfrak{D}\mathfrak{C}$ decomposition as above for $\partial_0 h$. For the h_1 contribution we have a direct bound without using any cancellations, while for h_2 we use again Lemma 2.4. The third term is similar to the second, with the roles of h and $\tilde{g}^{0\beta}$ interchanged. This concludes our reduction to the case of the metric $\tilde{g}$.

We continue with the second[4] reduction, which is to switch ∂_α and ∂_γ in the expression $\partial_\alpha T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_\gamma}u$; this allows us to replace the first term in (5.38) with the second. For fixed α and γ, we write

(5.40) $$\partial_\alpha T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_\gamma}u = \partial_\gamma T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_\alpha}u + f,$$

---

[4]Note that these two reductions are interchangeable.

where we claim f satisfies

(5.41)
$$\|P_{<k}f\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} 2^k \mathcal{B}^2.$$

This is trivial if $\alpha = \gamma = 0$. If both are nonzero, or if one of them is zero but $\beta \neq 0$, then there is no hat correction and this is a straightforward commutator bound. It remains to discuss the case when $\beta = 0$ and exactly one of α and γ are zero, say $\gamma = 0$. Then we need to consider the difference

$$f = \partial_\alpha T_{\tilde{g}^{\alpha 0}} \widehat{\partial_0 \partial_0} u - \partial_0 T_{\tilde{g}^{\alpha 0}} \partial_0 \partial_\alpha u$$
$$= \partial_\alpha T_{\tilde{g}^{\alpha 0}} \Pi(u, \partial_x \partial u) + T_{\partial_\alpha \tilde{g}^{\alpha 0}} \partial_0 \partial_0 u - T_{\partial_0 \tilde{g}^{\alpha 0}} \partial_0 \partial_\alpha u,$$

which can be estimated as in (5.41) using the fact that $\alpha \neq 0$ as well as the bound (5.15) for the second time derivative of u, respectively the similar bound (5.26) (third estimate) for $\partial_0 \tilde{g}$. Hence (5.41) follows.

Finally, it remains to examine the expression

$$g_\gamma = \partial_\gamma T_{\tilde{g}^{\alpha\beta}} \widehat{\partial_\beta \partial_\alpha} u,$$

where, unlike above, we take advantage of the summation with respect to α and β. Then, using the u equation, we have

$$g_\gamma = \partial_\gamma T_{\partial_\alpha \partial_\beta u} \tilde{g}^{\alpha\beta}.$$

The term where both α and β are zero vanishes since $\tilde{g}^{00}$ is constant; this was a motivation for the first reduction above. So this can be estimated directly as in (5.41) if $\gamma \neq 0$, and using either (5.15) or (5.26) (third estimate) for $\partial_0 \tilde{g}$, otherwise.

$\square$

6. Paracontrolled distributions

To motivate this section, we start from the classical energy estimates for the wave equation, which are obtained using the multiplier method. Precisely, one multiplies the equation $\Box_g u = f$ by Xu and simply integrates by parts. Here X is any regular time-like vector field. In the next section, we prove energy estimates for the paradifferential equation (3.25), by emulating this strategy at the paradifferential level. The challenge is then to uncover a suitable vector field X. Unlike the classical case, here not every time-like vector field X will suffice. Instead X must be carefully chosen, and in particular it will inherently have a limited regularity.

Since the metric g is a function of ∂u, scaling considerations indicate that the vector field X should be at the same regularity level. Naively, one might hope to have an explicit expression $X = X(\partial u)$ for our vector field. Unfortunately, seeking such an X eventually leads to an overdetermined system. At the other extreme, one might enlarge the class of X to all distributions that satisfy the same H^s and Besov norms as ∂u, which is essentially the class of functions that satisfy (5.26). While this class will turn out to contain the correct choice for X, it is nevertheless too large to allow for a clean implementation of the multiplier method.

Instead, there is a more subtle alternative, namely to have the vector X to be *paracontrolled* by ∂u. This terminology was originally introduced by Gubinelli, Imkeller and Perkowski [13] in connection to Bony's calculus, in order to study stochastic pde problems, see also [14]. However, similar constructions have been carried out earlier in the renormalization arguments

e.g. for wave maps, in work of Tao [42], Tataru [47] and Sterbenz-Tataru [41]; the last reference used the name *renormalizable* for the corresponding class of distributions.

In the standard usage, this is more of a principle than an exact notion, which needs to be properly adapted to one's purposes. For our own objective here, we provide a very precise definition of this notion, which is exactly tailored to the problem at hand.

6.1. Definitions and key properties.

Definition 6.1. *We say that a function z is **paracontrolled** by ∂u in a time interval I if it admits a representation*[5] *of the form*

$$z = T_a \partial u + r,$$ (6.1)

where the vector field a and the error r have the following properties:
(i) bounded para-coefficients a:

$$\|a\|_{\mathfrak{C}} \leq C.$$ (6.2)

(ii) balanced error r:

$$\|P_k r\|_{L^n} \leq C 2^{-k} c_k^2 \mathcal{A}^{\sharp 2}, \qquad \|\partial r\|_{L^\infty} \leq C \mathcal{B}^2.$$ (6.3)

It is convenient to think of the space of distributions z paracontrolled by ∂u as a Banach space, which we denote by $\mathfrak{P}(\partial u)$, or simply $\mathfrak{P}$. The norm in this Banach space is defined to be the largest implicit constant in (6.2) and (6.3), minimized over all representations of the form (6.1). If $\|z\|_{\mathfrak{P}} \lesssim_{\mathcal{A}^\sharp} 1$ then we will simply write

$$z \ll \partial u.$$

While for the most part this definition can be applied separately at each time t, in our context we will think of both u and z as functions of time, and think of these bounds as uniform in t. Precisely, above we think of $\mathcal{A}^\sharp$ as a global, time independent parameter, whereas $\mathcal{B}$ is allowed to be a possibly unbounded function of t.

To better understand the space $\mathfrak{P}$ of paracontrolled distributions, it is useful to relate it to the objects we have already discussed in the previous section:

Lemma 6.2. *a) We have the inclusion $\mathfrak{P} \subset \mathfrak{C}$.*
b) If F is a smooth function with $F(0) = 0$, then $F(\partial u) \in \mathfrak{P}$.

Proof. a) Clearly $\partial u \in \mathfrak{P}$. Then the first term in (6.1) can be placed in $\mathfrak{C}$ by part (b) of Lemma 5.7. The error term r also belongs to $\mathfrak{C}_0$ by Bernstein's inequality and interpolation. This is upgraded to $\mathfrak{C}$ using the $\partial_0 r$ bound in the second inequality in (6.3).
b) This is a direct consequence of parts (c), (d) of Lemma 5.7. $\square$

Thus one may think of the class $\mathfrak{P}$ of paracontrolled distributions as an intermediate stage between the class of smooth functions of ∂u, which is too narrow for our purposes, and the larger class $\mathfrak{C}$, which does not carry sufficient structure.

Next we consider nonlinear properties:

---

[5]Such a representation might not be unique in general, though later in the paper we often identify specific choices for the paracoefficients.

45

$\boxed{\text{l:Moser}}$ **Lemma 6.3.** *a) [Algebra property] The space $\mathfrak{P}(\partial u)$ is an algebra. Further, if $z_1, z_2 \in \mathfrak{P}$ have paracoefficients a_1, respectively a_2, then the paracoefficients of $z_1 z_2$ can be taken to be $z_1 a_2 + z_2 a_1$.*

b) [Moser inequality] If F is a smooth function with $F(0) = 0$ and $z \ll \partial u$, then $F(z) \ll \partial u$ and $F(z)$ satisfies

(6.4)
$$\|F(z)\|_{\mathfrak{P}} \lesssim_{\mathcal{A}^\sharp, \|z\|_{L^\infty}} \|z\|_{\mathfrak{P}}.$$

Further, if $z \in \mathfrak{P}$ has paracoefficients a, then the paracoefficients of $F(z)$ can be taken to be $F'(z)a$.

Proof. a) We consider the algebra property. Let

$$z_1 = T_{a_1} \partial u + r_1, \qquad z_2 = T_{a_2} \partial u + r_2$$

and expand $z_1 z_2$.

We first observe that we can place $\Pi(z_1, z_2)$ into the error term. For this it suffices to use the $\mathfrak{C}$ norm for z_1, z_2 and apply the second bound in (5.25).

We next consider $T_{z_1} z_2$ where for z_1 we again use only the $\mathfrak{C}$ norm. We begin with $T_{z_1} r_2$, which we also estimate as an error term. Here we estimate again the more difficult time derivative. If it falls on the first term then we can bound the output exactly as the balanced case above, see (5.25). Else, it suffices to use the uniform bound on z_1.

Finally, we consider the expression

$$T_{z_1} T_{a_2^\gamma} \partial_\gamma u = T_{z_1 a_2^\gamma} \partial_\gamma u + (T_{z_1} T_{a_2^\gamma} - T_{z_1 a_2^\gamma}) \partial_\gamma u,$$

where the first term has a $\mathfrak{C}$ coefficient by the $\mathfrak{C}$ algebra property, and the second may be estimated perturbatively. Here if the time derivative goes on the first factor then we are back to the previous case and no cancellation is needed. Else for $\partial_t \partial_\gamma u$ we use the decomposition in Definition 5.1(a) (or simply Lemma 5.4), combined with Lemma 2.7.

b) To prove the Moser inequality, our starting point is Lemma 5.7(d), which allows us to reduce the problem to estimating $T_{F'(z)} z$, using only the $\mathfrak{C}$ norm of z. But here we can bound $F'(z)$ in $\mathfrak{C}$ using the Moser bound in $\mathfrak{C}$, which allows us to conclude as in part (a). $\qquad \square$

In addition to the above lemmas, functions in $\mathfrak{P}$ essentially solve a paradifferential $\Box_{\tilde{g}}$ equation. This will be used later to estimate lower order terms in the proof of Theorem 7.1, and for (7.108):

$\boxed{\text{unds-extra}}$ **Lemma 6.4.** *Let $h \in \mathfrak{P}$. Then there exist functions f^α so that we have the representation*

$$\partial_\alpha T_{g^{\alpha\beta}} \partial_\beta h = \partial_\alpha f^\alpha,$$

which satisfy the following bounds:

$\boxed{\text{f-a}}$ (6.5)
$$|f^\alpha| \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2 \|h\|_{\mathfrak{P}},$$

respectively

$\boxed{\text{f-b}}$ (6.6)
$$\|P_{<k}(T_{g^{00}} \partial_0 h - f^0)\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} 2^k \|h\|_{\mathfrak{P}}.$$

The same result holds for the metrics $\tilde{g}$, $\hat{g}$.

Proof. We use the representation (6.1) for h. The property in the lemma holds trivially for the r component of h, with

$$f^\alpha = T_{\tilde{g}^{\alpha\beta}}\partial_\beta r.$$

Precisely, the bound (6.5) holds due to the second part of (6.3), while for the bound (6.6), the $\partial_0 r$ component cancels and then we can use the first part of (6.3).

It remains to consider h of the form $h = T_{a^\gamma}\partial_\gamma u$. We write

$$\partial_\alpha T_{\tilde{g}^{\alpha\beta}}\partial_\beta h := \partial_\alpha h^\alpha,$$

noting that the expression on the left hand side of (6.6) is exactly $h^0 - f^0$. We begin by refining the expression for h^α, noting that corrections of size $\mathcal{B}^2$ may be directly included into f^α without harming (6.6). For this we write

$$h^\alpha = \partial_\gamma T_{a^\gamma}T_{g^{\alpha\beta}}\partial_\beta u + T_{\tilde{g}^{\alpha\beta}}T_{\partial_\beta a^\gamma}\partial_\gamma u + [T_{\tilde{g}^{\alpha\beta}}, T_{a^\gamma}]\partial_\gamma\partial_\beta u - T_{a^\gamma}T_{\partial_\gamma \tilde{g}^{\alpha\beta}}\partial_\beta u - \partial_\alpha T_{\partial_\gamma a^\gamma}T_{\tilde{g}^{\alpha\beta}}\partial_\beta u,$$

where the first term on the right is the leading term, while the remaining terms can be estimated by $\mathcal{B}^2$ as follows:

- The second, fourth and fifth terms are estimated directly using (6.2) for $\partial_\beta a^\gamma$, $\partial_\gamma g^{\alpha\beta}$ respectively $\partial_\gamma a^\gamma$.
- The third term is estimated using the commutator bound in Lemma 2.4, as well as Lemma 5.4 if both β and γ are zero.

We have reduced the problem to the case when

$$h^\alpha = \partial_\gamma T_{a^\gamma}T_{g^{\alpha\beta}}\partial_\beta u.$$

At this point we rewrite

$$\partial_\alpha h^\alpha = \partial_\gamma \tilde{h}^\gamma, \qquad \tilde{h}^\gamma = \partial_\alpha T_{a^\gamma}T_{g^{\alpha\beta}}\partial_\beta u,$$

noting that

$$\|P_{<k}(h^0 - \tilde{h}^0)\|_{L^\infty} \lesssim 2^k,$$

which allows us to switch h and $\tilde{h}$ also in (6.6). Then we are allowed to correct $\tilde{h}^\gamma$, by writing

$$\tilde{h}^\gamma = T_{\partial_\alpha a^\gamma}T_{g^{\alpha\beta}}\partial_\beta u + T_{a^\gamma}\partial_\alpha T_{g^{\alpha\beta}}\partial_\beta u.$$

Now both terms on the right can be estimated by $\mathcal{B}^2$ as follows:
- The first term is estimated directly using (6.2) for $\partial_\alpha a^\gamma$.
- The second term is estimated using Lemma 5.10.

Hence the proof of the lemma is concluded.

$\square$

In addition to the class of paracontrolled distributions $\mathfrak{P}$ we also define a secondary class of distributions, which roughly speaking corresponds to derivatives of $\mathfrak{P}$ functions.

Definition 6.5. *The space $\mathfrak{DP}$ of distributions consists of functions y that admit a representation*

$$(6.7) \qquad\qquad y = \partial_\alpha z^\alpha + r,$$

where

$$(6.8) \qquad\qquad \|z^\alpha\|_{\mathfrak{P}} \le C, \qquad \|r\|_{L^\infty} \le C\mathcal{B}^2,$$

with the natural associated norm.

Due to the inclusion $\mathfrak{P} \subset \mathfrak{C}$, we can directly relate it to the class $\mathfrak{DC}$ introduced earlier.

l:DP-in-Q **Lemma 6.6.** *We have* $\mathfrak{DP} \subset \mathfrak{DC}$.

Next we verify that $\mathfrak{DP}$ is stable under multiplication by $\mathfrak{P}$ functions.

l:ppxDPP **Lemma 6.7.** *We have the bilinear bound*

$$(6.9) \qquad \mathfrak{P} \times \mathfrak{DP} \to \mathfrak{DP}.$$

As a corollary of this lemma, it follows that our gradient potentials A^γ and $\tilde{A}^\gamma$ are in $\mathfrak{DP}$.

Proof. For $h, z \in \mathfrak{P}$ we consider the expansion

$$q = h\partial_\alpha z$$
$$= \partial_\alpha T_h z - T_{\partial_\alpha h} z + \pi(h, \partial_\alpha z) + T_{\partial_\alpha z} h.$$

The first term is in $\mathfrak{DP}$ by Lemma 6.3(a). The three remaining terms can be perturbatively estimated by $\mathcal{B}^2$, using the bounds in (5.25). $\qquad\square$

Finally, we consider decompositions for $\mathfrak{DP}$ functions which are akin to Lemma 5.4. We will do this in two different ways, one which is shorter but loses some structure, and another which is more involved but retains full structure.

:switch-dt **Lemma 6.8.** *Let* $w \in \mathfrak{DP}$. *Then* w *admits a representation of the form*

$$w = \partial_x w_1 + r,$$

where

w1-bd $$(6.10) \qquad \|w_1\|_{\mathfrak{P}} \lesssim_{\mathcal{A}^\sharp} \|w\|_{\mathfrak{DP}}, \qquad \|r\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2 \|w\|_{\mathfrak{P}}.$$

Proof. It suffices to consider w of the form $w = \partial_0 z$ where $z \in \mathfrak{P}$, with a representation as in (6.1),

$$z = T_{a^\gamma} \partial_\gamma u + r,$$

with a^γ, r as in (6.2),(6.3). The bound (6.3) allows us to discard the contribution of r to (6.10). It remains to produce an appropriate modification $\partial_x z_1$, with $z_1 \in \partial_x \mathfrak{P}$, for the expression

$$q = \partial_0 T_{a^\gamma} \partial_\gamma u.$$

We successively peel off perturbative $O(\mathcal{B}^2)$ layers from q. First we use (6.2) to write

$$q = T_{a^\gamma} \partial_0 \partial_\gamma u + T_{\partial_0 a^\gamma} \partial_\gamma u = T_{a^\gamma} \partial_0 \partial_\gamma u + O(\mathcal{B}^2).$$

At this point we have two cases to consider:
(i) $\gamma \neq 0$. Then we write

$$q = \partial_\gamma T_{a^\gamma} \partial_0 u + O(\mathcal{B}^2),$$

and the remaining expression is in $\partial_x \mathfrak{P}$.
(ii) $\gamma = 0$. Here we use the equation for u to write

$$\partial_t^2 u = - \sum_{(\alpha,\beta) \neq (0,0)} T_{\tilde{g}^{\alpha\beta}} \partial_\alpha \partial_\beta u + \Pi(\tilde{g}^{\alpha\beta}, \partial_\alpha \partial_\beta u) + T_{\partial_\alpha \partial_\beta u} \tilde{g}^{\alpha\beta}.$$

Here the first term on the right involves at least one spatial derivative and is treated as before, in the case $\gamma \neq 0$, while the contributions of the last two terms are perturbative, and can be bounded by $\mathcal{B}^2$.

$\qquad\square$

Our second representation provides a more explicit recipe to obtain the corrected version not only of $\mathfrak{DP}$ functions, but also of $\mathfrak{P} \times \mathfrak{DP}$ functions:

Lemma 6.9. *Let $w = z_1^\alpha \partial_\alpha z_2$, where $z_1, z_2 \in \mathfrak{P}$, and z_2 has the $\mathfrak{P}$ representation*

$$z_2 = T_{a^\gamma} \partial_\gamma u + r.$$

Define

$$\mathring{w} = T_{z_1^\alpha a^\gamma} \widehat{\partial_\alpha \partial_\gamma} u.$$

Then we have

(6.11)
$$\|w - \mathring{w}\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} \mathcal{B}^2,$$

while

(6.12)
$$\|P_k \mathring{w}\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} 2^{\frac{k}{2}} \mathcal{B} c_k.$$

Proof. The contribution of r is directly perturbative so we discard it. Furthermore, the bounds in (5.25) allow us to replace perturbatively w by

$$w = T_{z_1^\alpha} \partial_\alpha z_2 + O(\mathcal{B}^2) = T_{z_1^\alpha} T_{a^\gamma} \partial_\alpha \partial_\gamma u + O(\mathcal{B}^2).$$

Using also Lemma 5.4 we obtain

$$w = T_{z_1^\alpha} T_{a^\gamma} \widehat{\partial_\alpha \partial_\gamma} u + O(\mathcal{B}^2).$$

Finally, we use Lemma 2.7 to combine the two paraproducts, arriving at

$$w = \mathring{w} + O(\mathcal{B}^2),$$

as needed. Finally, the bound (6.12) is also a consequence of Lemma 5.4. $\qquad\square$

The last lemma helps us uncover a more subtle, hidden $\Box_g$ structure which appears if we compute the double divergence of the metric $\tilde{g}$.

Lemma 6.10. *We have*

(6.13)
$$\|P_{<k} \partial_\alpha \partial_\beta \tilde{g}^{\alpha\beta}\|_{L^\infty} \lesssim_{\mathcal{A}^\sharp} 2^k \mathcal{B}^2.$$

Proof. For fixed β we expand $\partial_\beta \tilde{g}^{\alpha\beta}$ using the relation (3.9) to obtain

$$\partial_\beta \tilde{g}^{\alpha\beta} = -\partial^\beta u \, \tilde{g}^{\alpha\delta} \partial_\beta \partial_\delta u - \partial^\alpha u \, \tilde{g}^{\beta\delta} \partial_\beta \partial_\delta u + 2\partial^0 u \, \tilde{g}^{0\delta} \tilde{g}^{\alpha\beta} \partial_\beta \partial_\delta u.$$

For this expression we define a corresponding ring correction

$$\partial_\beta \mathring{\tilde{g}}^{\alpha\beta} := -T_{\partial^\beta u} T_{\tilde{g}^{\alpha\delta}} \widehat{\partial_\beta \partial_\delta} u - T_{\partial^\alpha u} T_{\tilde{g}^{\beta\delta}} \widehat{\partial_\beta \partial_\delta} u + 2T_{\partial^0 u} T_{\tilde{g}^{0\gamma}} T_{\tilde{g}^{\alpha\beta}} \widehat{\partial_\gamma \partial_\beta} u,$$

which is also chosen to vanish if $(\alpha, \beta) = (0, 0)$. We claim that the difference is perturbative for fixed α and β,

$$|P_{<k} \partial_\alpha (\partial_\beta \tilde{g}^{\alpha\beta} - \partial_\beta \mathring{\tilde{g}}^{\alpha\beta})| \lesssim_{\mathcal{A}^\sharp} 2^k \mathcal{B}^2.$$

Indeed, if $\alpha \neq 0$ then this follows directly from Lemma 6.9. On the other hand if $\alpha = 0$ then $\beta \neq 0$ in which case the hat correction can be discarded and we may distribute the time derivative, using the fact that $\partial u, \tilde{g} \in \mathfrak{C}$ modulo constants, see Lemma 5.7.

It remains to estimate the expression $\partial_\alpha (\partial_\beta \mathring{\tilde{g}}^{\alpha\beta})$, where we return to the standard summation convention and take the sum with respect to all (α, β). Here we separate the three terms in $\partial_\beta \mathring{\tilde{g}}^{\alpha\beta}$, in particular forfeiting the cancellation when $(\alpha, \beta) = (0, 0)$. By Lemma 5.7

49

all paracoefficients are in $\mathfrak{C}$, which allows us to perturbatively commute ∂_α with them as needed. Then it suffices to estimate the expression

$$-T_{\partial^\beta u}\,\partial_\alpha T_{\tilde{g}^{\alpha\delta}}\widehat{\partial_\beta\partial_\delta}u - T_{\partial^\alpha u}\,\partial_\alpha T_{\tilde{g}^{\beta\delta}}\widehat{\partial_\beta\partial_\delta}u + 2T_{\partial^0 u}\,T_{\tilde{g}^{0\gamma}}\partial_\alpha T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\gamma\partial_\beta}u.$$

For all terms here we may directly use Lemma 5.12(b) directly. Hence the proof of the lemma is concluded.

$\square$

6.2. **Symbol classes and the $\mathfrak{P}$DO calculus.** In a similar fashion to the $L^\infty S^m$ classes of symbols, our analysis will involve paradifferential operators with symbols that on the physical side are at either the $\mathfrak{P}$ or the $\mathfrak{D}\mathfrak{P}$ level. Precisely, we will work with both the symbol classes $\mathfrak{P}S^m$ and with the classes $\mathfrak{D}\mathfrak{P}S^m$.

For comparison purposes, we recall that for just paraproducts with $\mathfrak{P}$ functions $f, g \ll \partial u$ we have the uniform in time product bounds

(6.14)
$$\|T_f T_g - T_{fg}\|_{H^s \to H^s} \lesssim \mathcal{A}^{\sharp 2},$$

as well as the time dependent bounds

(6.15)
$$\|T_f T_g - T_{fg}\|_{H^s \to H^{s-1}} \lesssim \mathcal{B}^2,$$

and the corresponding commutator estimates. We also have, for $h \in \mathfrak{D}\mathfrak{P}$,

(6.16)
$$\|T_f T_h - T_{T_f h}\|_{H^s \to H^s} \lesssim \mathcal{B}^2.$$

Our objective in what follows is to expand these kinds of bounds to the ΨDO setting. We will see that things become more complex there. Fortunately, in the present paper we will only need such results primarily when one of the operators is a paraproduct, so we only prove our results in this case and merely make some comments about the general case.

We begin with the uniform in time bounds, i.e. the counterpart of (6.14), where not much changes:

Lemma 6.11. *Let* $f \in \mathfrak{P}S^j$, $g \in \mathfrak{P}S^k$. *Then*

(6.17)
$$\|T_f T_g - T_{fg}\|_{H^s \to H^{s-j-k}} \lesssim \mathcal{A}^{\sharp 2}.$$

Proof. By definition we have $f = f_1 + f_2$ where f_1 is an S^j multiplier and $f_2 \in \mathcal{A}^\sharp L^\infty S^j$, and similarly for g. Since $T_{f_1} = f_1(D)$ and $T_{g_1} = g_1(D)$, the leading parts cancel and we are left only with $O(\mathcal{A}^\sharp)$ terms, which can be estimated directly without using any cancellation.

$\square$

Our next result is concerned with the counterpart of (6.16), where again the result is similar:

Lemma 6.12. *For* $g \in \mathfrak{P}$ *and* $h \in \mathfrak{D}\mathfrak{P}S^m$ *we have*

$$\|T_g T_h - T_{T_g h}\|_{H^s \to H^{s-m}} \lesssim \mathcal{B}^2.$$

Here, by a slight abuse of notation, by $T_g h$ we mean the symbol paraproduct, where the Fourier variable is viewed as a parameter.

50

Proof. All operators in the lemma preserve dyadic frequency localization, so it suffices to fix a dyadic frequency size k and then show that we have

$$\|(T_g T_h - T_{T_g h})P_k u\|_{L^2} \lesssim 2^{mk}\mathcal{B}^2\|u\|_{L^2}.$$

Here we can include the 2^{mk} factor in h and reduce the problem to the case when $m = 0$.

In the first term we can also harmlessly replace g by $g_{<k}$ and T_g by multiplication by $g_{<k}$, as

trick1   (6.18) $$\|(T_g - g_{<k})P_k u\|_{L^2 \to L^2} \lesssim 2^{-\frac{k}{2}}\mathcal{B},$$

while $h_{<k} \in 2^{\frac{k}{2}}\mathcal{B}L^\infty S^0$ therefore

$$\|T_h P_k\|_{L^2 \to L^2} \lesssim 2^{\frac{k}{2}}\mathcal{B}.$$

Similarly, in the second term we can replace $T_g h$ by $g_{<k}h$, as

trick2   (6.19) $$\|P_{<k}(T_g h - g_{<k}h)\|_{L^\infty S^0} \lesssim \mathcal{B}^2,$$

akin to Lemma 6.7.

Thus it remains to bound in L^2 the simpler operator

$$R = (g_{<k}T_h - T_{g_{<k}h})P_k.$$

Our last simplification here is to separate variables in h, and reduce to the case where h has a product form at frequency 2^k, namely

$$h(x, \xi) = f(x)a(\xi), \qquad |\xi| \approx 2^k,$$

where $f \in \mathfrak{DP}$ and $a \in S^0$. This can be done for instance by thinking of h as a function of ξ in a dyadic frequency cube, smooth on the corresponding scale, and by taking a Fourier series in ξ, with coefficients depending on x. The coefficients will inherit the spatial regularity from h, and will be rapidly decreasing since we are taking the Fourier series of a smooth function of ξ.

After this simplification we may represent the operator T_h in the form

$$T_h P_k u = L_{lh}(f, P_k u),$$

where the symbol of the bilinear form L_{lh} depends linearly (and explicitly) on a. In this case we may rewrite the operator R in the form

$$Ru = g_{<k}L_{lh}(f, P_k u) - L_{lh}(g_{<k}f, P_k u),$$

At this point we can apply one last time the method of separation of variables to the symbol of L_{lh} to reduce the problem to the case when the bilinear form L_{lh} is of product type,

$$L_{lh}(f, P_k u) = b_{<k}(D)fc(D)P_k u,$$

where the symbols for both symbols $b_{<k}$ and cP_k are bounded and smooth on the 2^k scale. After this final reduction the operator R has a commutator structure,

$$Ru = [g_{<k}, b_{<k}(D)]fc(D)P_k u.$$

Here $|P_{<k}f| \lesssim 2^{\frac{k}{2}}\mathcal{B}$, while the commutator can be bounded by

$$\|[g_{<k}, b_{<k}(D)]\|_{L^2 \to L^2} \lesssim 2^{-k}\|\partial_x g_{<k}\|_{L^\infty} \lesssim 2^{-\frac{k}{2}}\mathcal{B}.$$

Hence we obtain

$$\|R\|_{L^2 \to L^2} \lesssim \mathcal{B}^2,$$

and the proof of the lemma is concluded.

□

In very limited circumstances, we will also need a more precise commutator expansion, which arises in the context where we commute one paradifferential operator with symbol $h \in \mathfrak{P}S^m$ with a function $g \in \mathfrak{P}$. This will be applied when $g = \tilde{g}^{\alpha\beta}$, but the result holds more generally. The novelty in the commutator expansion below is that we do not simply expand

$$\text{commutator} = \text{principal part} + \text{error}$$

but instead we seek to better understand the structure of the error,

$$\text{commutator} = \text{principal part} + \text{unbalanced subprincipal part} + \text{balanced error}$$

The principal part corresponds exactly with the Lie bracket of the two symbols, interpreted paradifferentially. For possible use later, we define this more generally for two symbols:

Definition 6.13. *The para-Lie bracket of two symbols $f \in \mathfrak{P}S^j$, $g \in \mathfrak{P}S^k$ is defined as*

(6.20) $$\{f, g\}_p = T_{\partial_\xi f}\partial_x g - T_{\partial_\xi g}\partial_x f.$$

This belongs to $\mathfrak{DP}S^{j+k-1}$.

We remark that if f is merely a function, then the first term on the right drops.

While the principal part of the commutator can be described using a paradifferential operator with an appropriate symbol, the unbalanced subprincipal part has a more complex structure which would be described best using a variable coefficient bilinear form. In order to be able to describe this structure, we need a slight expansion of the class L_{lh} of bilinear operators in Definition 2.2:

Definition 6.14. *By $\mathfrak{P}S^m L_{hl}$ we denote any bilinear operator which is a linear combination of operators of the form*

$$T_h L_{lh}, \qquad h \in \mathfrak{P}S^m,$$

which is either finite, or infinite but rapidly convergent.

With this notation, we have the following commutator result:

Proposition 6.15. *For $g \in \mathfrak{P}$ and $h \in \mathfrak{P}S^m$ we have the commutator expansion*

(6.21) $$[T_g, T_h] = -iT_{\{g,h\}_p} + OP\mathfrak{P}S^{m-2}L_{lh}(\partial_x^2 g, \cdot) + R,$$

where

(6.22) $$\|R\|_{H^s \to H^{s-m+1}} \lesssim \mathcal{B}^2.$$

Proof. As in the proof of Lemma 6.12, we first localize in frequency to a dyadic scale 2^k for the input/output, and reduce to the case $s = 0$ and $m = 2$.

We consider first the special case when h is a multiplier, $h(x, \xi) = h(\xi)$. Then

$$\{h, g\}_p = h_\xi g_x.$$

In this case we claim that we have an exact formula,

(6.23) $$[T_g, T_h]u = -iT_{\{g,h\}_p}u + C, \qquad Cu = L_{lh}(\partial_x^2 g, u).$$

A-priori the last term on the right, C, is a lh type translation invariant bilinear form in g, u; all we need to do is to compute its symbol $R(\eta, \xi)$, and verify that it has symbol type

52

regularity and vanishes of second order when $\eta = 0$. The symbol for $T_g u$ as a bilinear form in g and u is

$$\ell(\eta, \xi) = \chi\left(\frac{|\eta|}{|\xi + \frac{1}{2}\eta|}\right).$$

Then the symbol for the commutator is

$$\chi\left(\frac{|\eta|}{|\xi + \frac{1}{2}\eta|}\right)(h(\xi) - h(\xi + \eta)).$$

We expand the last difference as a Taylor series around the middle as

$$h(\xi) - h(\xi + \eta) = -\eta \nabla h(\xi + \frac{1}{2}\eta) + \eta^2 r(\xi, \eta)$$

with r a smooth symbol in both η and ξ on the 2^k scale for $|\eta| \ll |\xi| \approx 2^k$. The middle term gives the symbol of the Weyl quantization for the Lie bracket $\{h, g\}_p$. The last term yields the error term C, which has the η^2 factor corresponding to the two derivatives of g.

Next we turn our attention to the general case, which we seek to reduce to the special case above. This is achieved by separating variables in the symbol h, which allows us to assume without any restriction in generality that the symbol h has the form

$$h(x, \xi) = a(x)b(\xi).$$

Then we have a corresponding decomposition at the operator level,

separate $\qquad$ (6.24) $\qquad\qquad T_h u = T_a B(D) u + C_0 u, \qquad C_0 u = L_{lh}(a_x, u).$

Here we can estimate the commutator with T_g as an error term,

$$[C_0, T_g] = R.$$

This is most readily seen using another separation of variables, which allows us to reduce the problem to the case when

$$C_0 u = C_0^1(D) a_x C_0^2(D) u,$$

after which we may apply Lemma 2.4. The same lemma also shows that the commutator $[T_a, T_g]$ yields an error term, so we arrive at

$$[T_g, T_h] = T_a[B(D), T_g] + R.$$

For the commutator on the right we apply the formula (6.23), which yields

$$[T_g, T_h]u = -iT_a(T_{b_\xi g_x} + L_{lh}(g_{xx}, \cdot)).$$

It remains to refine the first product,

$$T_a T_{b_\xi g_x} = T_{T_{ab_\xi} g_x} + R$$

for which we use Lemma 6.12. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Our final result here is a product formula where we also need an expansion akin to (6.21). One should contrast this with Lemma 6.12, where such expansion was not necessary.

| p:prod-pdo |

Proposition 6.16. *For $g \in \mathfrak{P}S^m$ and $h \in \mathfrak{DP}$ we have the commutator expansion*

| multi-prod |

$$(6.25) \qquad T_g T_h = T_{T_g h} + OP\mathfrak{P}S^{m-1} L_{lh}(\partial_x h, \cdot) + R,$$

where

| multi-pR |

$$(6.26) \qquad \|R\|_{H^s \to H^{s-m}} \lesssim \mathcal{B}^2.$$

Proof. The proof follows the same outline as the proof of the previous proposition, so we only outline the main points.

We localize first in frequency to a dyadic frequency region at scale 2^k and then separating variables in the first factor. If g is simply a multiplier then then (6.25) is an exact identity akin to (6.23) above. If instead

$$g = a(x)b(\xi), \qquad a \in \mathfrak{P},$$

then we expand T_g as in (6.24), and then replace T_a by multiplication by $a_{<k}$, using (6.18), (6.19). After these simplifications, we are left with estimating the difference

$$R_0 = (g_{<k} T_{bh} - T_{g_{<k}bh}) P_k u = g_{<k} L_{lh}(b, P_k u) - L_{lh}(g_{<k}b, P_k u).$$

This difference is easily turned into another commutator and estimated as in (6.26); this is achieved by separating again variables in the symbol of L_{lh} as in the analysis after (6.24). $\qquad \square$

7. Energy estimates for the paradifferential equation

| s:para-ee |

Our objective in this section is to prove that the linear paradifferential flow

| n-inhom-re |

$$(7.1) \qquad (\partial_\alpha T_{g^{\alpha\beta}} \partial_\beta - T_{A^\gamma} \partial_\gamma)v = f$$

is locally well-posed in a range of Sobolev spaces. Precisely, we will show that

| t:para-wp |

Theorem 7.1. *Let u be a smooth solution for the minimal surface equation (3.5) in a time interval $I = [0,T]$, with associated control parameters $\mathcal{A}^\sharp$ and $\mathcal{B}$ so that*

$$(7.2) \qquad \mathcal{A}^\sharp \ll 1, \qquad \mathcal{B} \in L_t^2.$$

Let $s \in \mathbb{R}$. Then the linear paradifferential flow (7.1) is locally well-posed in $\mathcal{H}^s$ in the time interval I. Furthermore, there exists an energy functional $E^s(v) = E^s(v[t])$, depending on u, which is smooth in $\mathcal{H}^{s+1}$, with the following two properties:

a) Energy equivalence:

$$(7.3) \qquad E^s(v[t]) \approx \|v[t]\|^2_{\mathcal{H}^s}.$$

b) Energy estimate:

$$(7.4) \qquad \frac{d}{dt} E^s(v[t]) \lesssim \mathcal{B}^2 E^s(v[t]) + \|f\|_{H^{s-1}} E^s(v[t])^{\frac{1}{2}}.$$

The same result is also valid for the paradifferential equations (3.26), respectively (3.27) associated to the metrics $\tilde{g}$ and $\hat{g}$.

We remark on the modular structure of our arguments. Precisely, from this section it is only the conclusion of this theorem which is used later in the paper. We also remark on the smallness condition for $\mathcal{A}^\sharp$:

Remark 7.2. The condition that $\mathcal{A}^\sharp \ll 1$ in the theorem is a technical convenience rather than a necessity. It is only used in the reduction in Proposition 7.3 in order to ensure that the operator $T_{g^{00}}$ is invertible, and then in Lemma 7.4 in order to insure that our vector field X is forward time-like. Since $|g^{00}| \gtrsim 1$ this may be alternatively guaranteed by a more careful choice of the quantization, respectively construction of X. Another minor advantage is that with this assumption we no longer need to track the dependence on $\mathcal{A}^\sharp$ of implicit constants in all the estimates.

It will be easier to prove the result for the paradifferential flow associated to the metric $\tilde{g}$. Because of this, our first step will be to reduce the problem to this case. Then we will prove the result for $\tilde{g}$ in two steps. First, we show that the desired result holds for $s = 0$. Then, we use a paraconjugation argument to show that the same result holds for all real s.

7.1. Equivalent metrics. The idea here is that we can replace the metric g with the conformally equivalent metric $\tilde{g}$ given by (3.18) in order to simplify the subsequent analysis. A similar equivalence holds for the metric $\hat{g}$; the argument is completely identical.

Then we have the following equivalence:

Proposition 7.3. *Assume that v solves* (7.1). *Then it also satisfies an equation of the form*

$$(7.5) \qquad (\partial_\alpha T_{\tilde{g}^{\alpha\beta}} \partial_\beta - T_{\tilde{A}^\gamma} \partial_\gamma) v = E f + \tilde{R} v,$$

where E is invertible and elliptic,

$$(7.6) \qquad \|E f\|_{H^s} \approx \|f\|_{H^s},$$

and $\tilde{R}$ is balanced,

$$(7.7) \qquad \|\tilde{R} v\|_{H^s} \lesssim \mathcal{B}^2 \|\partial v\|_{H^s}.$$

Proof. We first observe that, since g^{00} is a small, $O(\mathcal{A})$ perturbation of a nonzero constant, it follows that $T_{(g^{00})^{-1}}$ is invertible elliptic, with elliptic inverse $E = (T_{(g^{00})^{-1}})^{-1}$, which satisfies (7.6) for all real s.

Then v solves (7.5) with R of the form

$$\tilde{R} = E \partial_\alpha (T_{\tilde{g}^{\alpha\beta}} - T_{(g^{00})^{-1}} T_{g^{\alpha\beta}}) \partial_\beta - E(T_{(g^{00})^{-1}} T_{A^\gamma} - T_{\tilde{A}^\gamma} - T_{\partial_\alpha (g^{00})^{-1}} T_{g^{\alpha\gamma}}) \partial_\gamma.$$

Here we have the algebraic relations

$$\tilde{g}^{\alpha\beta} = (g^{00})^{-1} g^{\alpha\beta}, \qquad (g^{00})^{-1} A^\gamma = \tilde{A}^\gamma + \partial_\alpha (g^{00})^{-1} g^{\alpha\gamma}.$$

This allows us to estimate $\tilde{R}$ in a balanced fashion using Lemma 5.7 and Lemma 2.7, as desired. $\qquad\square$

As a consequence of this result, we see that it suffices now to prove the result in Theorem 7.1 but with the equation (7.1) replaced by

$$(7.8) \qquad (\partial_\alpha T_{\tilde{g}^{\alpha\beta}} \partial_\beta - 2 T_{\tilde{A}^\gamma} \partial_\gamma) v = f.$$

7.2. The $H^1 \times L^2$ bound. For expository purposes, we first review the multiplier method for proving energy estimates for the wave equation in a simplified setting. Guided by this, we construct a suitable vector field, to be used as our multiplier. Finally, we reinterpret the energy estimates at the paradifferential level, and prove Theorem 7.1 with $s = 1$.

7.2.1. *Energy estimates via the multiplier method.* Suppose that we have a function v that solves a divergence form wave equation,

(7.9)
$$Pv = f, \qquad P = \partial_\alpha g^{\alpha\beta}\partial_\beta - A^\alpha\partial_\alpha.$$

Given a vector field $X = X^\alpha\partial_\alpha$, the standard strategy is to multiply the equation by Xv and integrate by parts. For expository purposes we will follow this path here, noting that another alternative would be to interpret the vector field in the Weyl calculus, and work instead with the skew-adjoint operator

$$X^w = X^\alpha\partial_\alpha + \frac{1}{2}\partial_\alpha X^\alpha.$$

At this point we only seek to identify the principal part of the energy estimates, which will lead us to the choice of the vector field X, so we do not follow this second path. However, later on, once X is chosen and we have switched to the paradifferential setting, we will need to also carefully track the lower order terms, and we will add lower order corrections to our vector field.

To further place the following computations into context, we remark that vector field energy identities for the wave equation are often employed in their covariant form, which is derived by contracting the divergence free relation for the energy momentum tensor with the vector field Xu, and integrating with respect to the measure associated with the metric g. Such a strategy would work but would be counterproductive in our setting, where we will reinterpret all these identities in a paradifferential fashion.

Assuming at first that the function v is compactly supported, integrating by parts several times, in order to essentially commute the second order part of P with X, one arrives at the identity

(7.10)
$$2\iint Pv \cdot Xv \, dxdt = \iint c_X(v,v) \, dxdt,$$

where c_X is a quadratic expression in ∂v of the form

(7.11)
$$c_X(v,v) = c_X^{\alpha\beta}\partial_\alpha v \, \partial_\beta v$$

with coefficients given by the relation

(7.12)
$$c_X(x,\xi) := c_X^{\alpha\beta}\xi_\alpha\xi_\beta = \{p, X^\gamma\xi_\gamma\}(x,\xi) - \partial_\gamma X^\gamma p(x,\xi) + 2A^\gamma\xi_\gamma X^\delta\xi_\delta,$$

where we recall that $p(x,\xi) = g^{\alpha\beta}\xi_\alpha\xi_\beta$. Removing the compact support assumption on v and introducing boundaries at times $t = 0$ and $t = T$, the identity above with the integral taken over $[0,T] \times \mathbb{R}^n$ still holds but with added contributions at these times,

(7.13)
$$2\int_0^T\!\!\int_{\mathbb{R}^n} Pv \cdot Xv \, dxdt = \int_0^T\!\!\int_{\mathbb{R}^n} c_X(v,v) \, dxdt + \int_{\mathbb{R}^n} e_X(v,v) \, dx\Big|_0^T,$$

where the contributions at the initial and final time can be thought of as energies. Here the energy density e_X is a bilinear expression of the form

(7.14)
$$e_X(v,v) = e_X^{\alpha\beta}\partial_\alpha v \cdot \partial_\beta v.$$

This can be written in terms of the energy momentum tensor associated to the $\Box_g$ operator,

| T-def | (7.15) | $$T_{\alpha\beta}[v] = \partial_\alpha v \cdot \partial_\beta v - \frac{1}{2} g_{\alpha\beta} g^{\gamma\delta} \partial_\delta v \cdot \partial_\gamma v.$$ |

Then we have

| eX | (7.16) | $$e_X^{\alpha\beta} \partial_\alpha v \cdot \partial_\beta v = g^{0\alpha} T_{\alpha\beta} X^\beta = T(\partial_t, X).$$ |

Thus we can define the energy functional associated to the vector field X as

| def-EX | (7.17) | $$E_X[v] = \int_{\mathbb{R}^n} e_X(v, v) \, dx.$$ |

The key property of the energy density e_X is that it is classically known to be positive definite in a pointwise sense,

| X-positive | (7.18) | $$e_X(v, v) = T(\partial_t, X) \gtrsim |\partial v|^2,$$ |

provided that the vector fields ∂_t and X are uniformly forward time-like. Then we obtain the energy coercivity property

$$E_X[v(t)] \approx \|\partial v(t)\|_{L^2}^2.$$

With these notations, we can rewrite the integral identity (7.13) as a differential identity

| nergy-diff | (7.19) | $$\frac{d}{dt} E_X(v) = 2 \int_{\mathbb{R}^n} Pv \cdot Xv \, dx - \int_{\mathbb{R}^n} c_X(v, v) \, dx.$$ |

In a nutshell, this computation, interpreted paradifferentially, is at the heart of our proof of the energy estimates. In this context, the choice of the vector field X should naively be governed by the requirement that the energy flux form c_X is balanced. We note that one cannot ask for c_X to be zero, as this would produce an overdetermined system for X, which in particular implies the condition that X is a conformal Killing field for the metric g. Even the requirement that c_X is balanced turns out to be a bit too much, which is why we will need a second step to the above computation.

Precisely, the second step is based on another interesting observation, namely that the contribution of terms in $c_X^{\alpha\beta}$ of the form

$$I = \int_0^T \int_{\mathbb{R}^n} q g^{\alpha\beta} \partial_\alpha v \cdot \partial_\beta v \, dxdt$$

has a favourable structure and can be eliminated using a suitable Lagrangian type energy correction.

Indeed, for compactly supported v, this contribution can be rewritten, integrating by parts, as

$$I = -\iint q \partial_\alpha g^{\alpha\beta} \partial_\beta v \cdot v \, dxdt + \frac{1}{2} \iint \partial_\alpha g^{\alpha\beta} \partial_\beta q \, v^2 dxdt$$

$$= -\iint Pv \cdot qv \, dxdt + \frac{1}{2} \iint (Pq + q\partial_\gamma A^\gamma) v^2 dxdt + \iint A^\gamma \partial_\gamma q \, v^2 dxdt.$$

The first term can be interpreted as a correction to X in (7.10). Introducing the notation

(7.20)
$$\mathfrak{M} = 2X + q,$$

it now takes the form

(7.21)
$$\iint Pv \cdot \mathfrak{M}v \, dxdt = \iint c_X(v,v) - qg^{\alpha\beta}\partial_\alpha v \cdot \partial_\beta v + dv^2 \, dxdt,$$

where the coefficient d of the additional zero order term is

(7.22)
$$d = \frac{1}{2}(Pq + q\partial_\gamma A^\gamma) + A^\gamma \partial_\gamma q.$$

Finally, adding in boundaries at $t = 0, T$ we obtain the integral relation

(7.23)
$$\int_0^T \int_{\mathbb{R}^n} Pv \cdot \mathfrak{M}v \, dxdt = \int_0^T \int_{\mathbb{R}^n} c_{X,q}(v,v) + dv^2 \, dxdt + \int_{\mathbb{R}^n} e_{X,q}(v,v) \, dx \Big|_0^T,$$

where the leading flux density is now

$$c_{X,q}(v,v) = c_X(v,v) - qg^{\alpha\beta}\partial_\alpha v \cdot \partial_\beta v,$$

while the new energy density $e_{X,q}$ has the form

$$e_{X,q}(v,v) = e_X(v,v) + qg^{0\beta}\partial_\beta v \cdot v - \frac{1}{2}(g^{0\beta}\partial_\beta q - qA^0)v^2.$$

We can also convert this into a differential relation akin to ($\overset{\text{energy-diff}}{7.19}$), namely

(7.24)
$$\frac{d}{dt}E_{X,q}(v) = \int_{\mathbb{R}^n} Pv \cdot \mathfrak{M}v \, dx - \int_{\mathbb{R}^n} c_{X,q}(v,v) + dv^2 \, dx.$$

The identity ($\overset{\text{energy-diff-t}}{7.24}$) will heuristically provide the intuition for the proof of the desired energy estimate. However, to make this rigorous we will have to re-implement the above computation at the paradifferential level. There, the treatment of the lower order terms will differ slightly, in part in order to avoid a need for direct bounds on higher order time derivatives of u.

Based on the relation ($\overset{\text{energy-diff-t}}{7.24}$), the vector field X will have to be chosen so that the symbol for the bilinear form $c_{X,q}$ is balanced, or equivalently so that c_X is balanced modulo a Lagrangian contribution. In turn, the Lagrangian correction weight q will have to be chosen carefully, so that it satisfies multiple requirements:

(1) Comparing the form of $c_{X,q}$ with the earlier expression for c_X, a natural choice would seem to be

(7.25)
$$q = \partial_\gamma X^\gamma.$$

(2) Examining the lower order coefficient d above, we will need to have good control over the function Pq. Here it is the second order part of P that matters, as the effect of the magnetic term will turn out to be directly perturbative.

Reconciling these two requirements will play an important role later on in this section.

To complete our discussion here, we need to carry out an additional step, namely to investigate what happens if we replace g, A by $\tilde{g}, \tilde{A}$. Observing that

$$Pv = g^{00}\tilde{P}v$$

it becomes natural to replace the vector field X, the Lagrangian weight q and the multiplier by

$$\tilde{X} = g^{00}X, \qquad \tilde{q} = g^{00}q, \qquad \tilde{\mathfrak{M}} = 2\tilde{X} + \tilde{q}.$$

Then the relation (7.23) remains essentially unchanged,

$$\boxed{\text{IBP-A-tt}} \quad (7.26) \qquad \int_0^T\!\!\int_{\mathbb{R}^n} \tilde{P}v \cdot \mathfrak{M}v \, dxdt = \int_0^T\!\!\int_{\mathbb{R}^n} c_{X,q}(v,v) + dv^2 \, dxdt + \int_{\mathbb{R}^n} e_{X,q}(v,v) \, dx \bigg|_0^T,$$

and the same applies to the differential form (7.24) of the same relation. Here the principal flux symbol can be equivalently expressed in the form

$$\boxed{\text{tcX}} \quad (7.27) \qquad c_{X,q}(x,\xi) := c_{X,q}^{\alpha\beta}\xi_\alpha\xi_\beta = \{\tilde{p}, \tilde{X}\}(x,\xi) - (\partial_\gamma \tilde{X}^\gamma + \tilde{q})p(x,\xi) + 2\tilde{A}^\gamma \xi_\gamma \tilde{X}^\delta \xi_\delta.$$

Our task is now twofold:

- To identify a suitable time-like vector field X so that the energy flux above satisfies a balanced energy estimate, and
- To recast the above computation in the paradifferential setting without losing the energy balance; this will also require a careful choice for q.

7.2.2. *The construction of the vector field X.* Our objective here is to construct a forward time-like vector field X so that the flux coefficients in $c_{X,q}$ are balanced for q as in (7.25). In essence, at this stage we disregard any paradifferential frequency localizations, and work as if v has infinite frequency. We also do not distinguish between g and $\tilde{g}$, as this does not play a role in the choice of X. Our main result governing the choice of the vector field X is where our notion of paracontrolled distributions is first needed, and reads as follows:

$\boxed{\text{l:X}}$ **Lemma 7.4.** *There exists a forward time-like vector field X that is paracontrolled by ∂u, and so that we have the balanced bound*

$$\boxed{\text{x3}} \quad (7.28) \qquad \|c_X^{\alpha\beta} + \partial_\gamma X^\gamma g^{\alpha\beta}\|_{L^\infty} \lesssim \mathcal{B}^2.$$

We remark that the fact that such a vector field exists is closely connected to the fact that our equation satisfies the nonlinear null condition in a strong sense. One should think of our vector field X as the next best thing to a Killing or conformal Killing vector field. Perhaps a good terminology would a *para-Killing vector field*, i.e. whose deformation tensor is balanced, rather than equal to zero or a multiple of the metric.

Proof. Starting from (7.12), we compute the expression in (7.28) as follows:

$$(c_X^{\alpha\beta} + \partial_\gamma \tilde{X}^\gamma g^{\alpha\beta})\xi_\alpha\xi_\beta = 2\xi_\gamma\xi_\alpha g^{\gamma\beta}\partial_\beta X^\alpha - X^\gamma\xi_\alpha\xi_\beta\partial_\gamma g^{\alpha\beta} + 2A^\delta\xi_\delta X^\gamma\xi_\gamma$$

$$= 2\xi_\gamma\xi_\alpha g^{\gamma\beta}\partial_\beta X^\alpha + 2X^\gamma\partial^\beta u\xi_\beta g^{\alpha\delta}\partial_\delta\partial_\gamma u\xi_\alpha + 2\partial^\alpha u\partial_\alpha\partial_\beta u g^{\beta\delta}\xi_\delta X^\gamma\xi_\gamma$$

$$= 2\xi_\gamma\xi_\alpha g^{\gamma\beta}\partial_\beta X^\alpha + 2\xi_\gamma\xi_\alpha X^\beta\partial^\gamma u g^{\alpha\delta}\partial_\delta\partial_\beta u + 2\xi_\alpha\xi_\gamma\partial^\delta u\partial_\delta\partial_\beta u g^{\beta\alpha} X^\gamma$$

$$= 2\xi_\gamma\xi_\alpha g^{\gamma\beta}(\partial_\beta X^\alpha + X^\delta\partial^\alpha u\partial_\delta\partial_\beta u + X^\alpha\partial^\delta u\partial_\delta\partial_\beta u).$$

Here one could freely symmetrize the coefficients relative to the pair of indices (α, γ). We have chosen to neglect the symmetrization, but, instead, we made favourable choices. The above expression would cancel for instance if

$$\boxed{\text{determined}} \quad (7.29) \qquad \partial_\beta X^\alpha = -X^\delta\partial^\alpha u\partial_\beta\partial_\delta u - \partial^\delta u X^\alpha\partial_\delta\partial_\beta u.$$

This is an overdetermined system, so we cannot hope for an exact cancellation. Even if we symmetrize (raising the β index first) and equate the symmetric part of the two sides, it still remains overdetermined.

But we do not need exact cancellation, we only need the difference of the two sides to be balanced. Assume for the moment that X is at the same regularity level as ∂u. Then, examining the right hand side, the expressions there are unbalanced only in the paraproduct case, where the $\partial^2 u$ term is the high frequency, i.e. for the terms $T_{h(h,\nabla u)}\partial^2 u$. Hence we heuristically arrive at the equivalent requirement

$$\partial_\beta X^\alpha \overset{bal}{\approx} -T_{X^\delta \partial^\alpha u}\partial_\beta \partial_\delta u - T_{\partial^\delta u X^\alpha}\partial_\delta \partial_\beta u,$$

where we introduce the notation "$\overset{bal}{\approx}$" to indicate that the difference between the two expressions is balanced, i.e. can be estimated as in (7.28). Then, at leading order we may cancel the β derivative to obtain a single paradifferential relation at one regularity level higher, namely

$$X^\alpha \overset{bal}{\approx} -T_{X^\delta \partial^\alpha u}\partial_\delta u - T_{\partial^\delta u X^\alpha}\partial_\delta u.$$

Modulo balanced terms we may break the paraproducts above in two. This allows us to devise an inductive scheme to construct X as a dyadic sum of frequency localized pieces, by setting

(7.30) $$X = X_0 + \sum_{k=1}^\infty X_k$$

starting with the forward time-like initialization

$$X_0 = \partial_t,$$

and where the functions X_k, localized at frequency 2^k, are defined inductively by

def-Xk   (7.31) $$X_k^\alpha = -(T_{X^\delta}T_{\partial^\alpha u} + T_{X^\alpha}T_{\partial^\delta u})\partial_\delta u_k.$$

It remains to show that, as defined above, the vector field X has all the properties in the Lemma. We will achieve this in three stages:

- We show that X satisfies the same bounds as ∂u, (see (5.8) and (5.15)),

(7.32) $$\|X - X_0\|_{\mathfrak{C}} \lesssim 1.$$

Since $\mathcal{A}^\sharp \ll 1$, this in particular guarantees that X is forward time-like.
- We show that $X - X_0$ is paracontrolled by ∂u.
- Finally, we establish the balanced bound (7.28).

To simplify the notations, we will write schematically that

$$X_k = T_X T_h \partial u_k,$$

with coefficients h of the form $h = F(\partial u)$, which belong to $\mathfrak{C}$ modulo constants.

I. Dyadic bounds for X. These are proved at each dyadic frequency k by induction on k. We do this in two steps, where we first estimate the $\mathfrak{C}_0$ norm of $X - X_0$. Precisely, the first set of statements to be proved by induction for $k > 0$ is as follows:

induction1   (7.33) $$\|X_k\|_{L^{2n}} \le C2^{-\frac{k}{2}}\mathcal{A}^\sharp c_k^2,$$

induction2   (7.34) $$\|X_k\|_{L^\infty} \le C2^{-\frac{k}{2}}\mathcal{B}c_k,$$

60

with a fixed large universal constant C. This implies that $\|X - X_0\|_{\mathfrak{C}_0} \lesssim 1$. The induction hypothesis combined with Bernstein's inequality yields the bound

$$\|X_{<k} - X_0\|_{L^\infty} \lesssim C\mathcal{A}^\sharp.$$

Then we write

$$X_k = T_{X_0} T_h u_k + T_{X_{<k} - X_0} T_h u_k,$$

which yields

$$\|X_k\|_{L^{2n}} \lesssim (1 + C\mathcal{A}^\sharp) 2^{-\frac{k}{2}} \mathcal{A}^\sharp c_k^2,$$

respectively

$$\|X_k\|_{L^\infty} \lesssim (1 + C\mathcal{A}^\sharp) 2^{-\frac{k}{2}} \mathcal{B} c_k.$$

Thus the induction argument closes if C is a large constant and $\mathcal{A}^\sharp \ll 1$.

The second step is to prove that

$$\|X - X_0\|_{\mathfrak{C}} \lesssim 1.$$

To achieve this we will prove by induction that

(7.35) $$\|\partial_t X_{\leq k}\|_{\mathfrak{DC}} \leq C,$$

i.e. that $\partial_t X_{\leq k}$ admits a decomposition $\partial_t X_{\leq k} = f_{k1} + f_{k2}$, where

$$\|f_{1,\leq k}\|_{L^\infty} \leq C\mathcal{B}^2 c_k^2, \qquad \|P_j f_{2,\leq k}\|_{L^\infty} \leq C 2^{\frac{j}{2}} \mathcal{B} c_j.$$

Here again C is a fixed large constant, unrelated to the earlier C.

For this we write

$$\partial_t X_{\leq k} = \partial_t (T_X T_h \partial u_{\leq k}) = (T_{\partial_t X} T_h \partial u_{\leq k} + T_X T_{\partial_t h} \partial u_{\leq k}) + T_X T_h \partial_t \partial u_{\leq k}.$$

Here the X coefficients involve only frequencies below 2^k, so we may use the induction hypothesis in the first term. For the second and third terms it suffices to use the $\mathfrak{C}_0$ bound for X, which we already have from the first induction. Hence, repeatedly applying the bounds in (5.24) we obtain

$$\|\partial_t X_{\leq k}\|_{\mathfrak{DC}} \lesssim \mathcal{A}^\sharp \|\partial_t X_{<k}\|_{\mathfrak{DC}} \|h\|_{L^\infty} \|\partial u\|_{\mathfrak{C}_0} + \mathcal{A}^\sharp \|X\|_{\mathfrak{C}_0} \|\partial_t h\|_{\mathfrak{DC}} \|\partial u\|_{\mathfrak{C}_0} + \|X\|_{\mathfrak{C}_0} \|h\|_{\mathfrak{C}_0} \|\partial_t \partial u\|_{\mathfrak{DC}}$$
$$\lesssim C\mathcal{A}^\sharp + 1,$$

which closes the inductive proof of (7.35) if $C \gg 1$ and $\mathcal{A}^\sharp \ll 1$.

II. X is paracontrolled by ∂u. To prove this, we will establish the representation

(7.36) $$X^\alpha = -(T_{X^\delta \partial^\alpha u} + T_{X^\alpha \partial^\delta u}) \partial_\delta u + r^\alpha.$$

This will play the role of (6.1). The Moser estimates in Lemma 5.7 show that the paracoefficients above satisfy the bounds required of a in (6.2), so it remains to establish that the errors r_α satisfy the bounds (6.3). For this, we write

$$r^\alpha = \sum_k X_k^\alpha + (T_{X^\delta \partial^\alpha u} + T_{X^\alpha \partial^\delta u}) \partial_\delta u_k$$

$$= \sum_k -[(T_{X^\delta \partial^\alpha u} - T_{X^\delta} T_{\partial^\alpha u}) + (T_{X^\alpha \partial^\delta u} - T_{X^\alpha} T_{\partial^\delta u})] \partial_\delta u_k$$

$$:= \sum_k r_k^\alpha.$$

Now we apply Lemma 2.7 to estimate

bd-ra (7.37) $$\|r_k^\alpha\|_{L^n} \lesssim 2^{-k} c_k^2 \mathcal{A}^{\sharp 2}, \qquad \|r_k^\alpha\|_{L^\infty} \lesssim 2^{-k} c_k^2 \mathcal{B}^2$$

as needed. It remains to bound the time derivative of r^α in L^∞. For this we distribute the time derivative. If it falls on any of the para-coefficients then we can directly use the bound (5.25). Else, we use Lemma 5.9.

III. The bound for $c_X^{\alpha\beta} + \partial_\gamma X^\gamma g^{\alpha\beta}$. Here we recall that

$$c_X^{\alpha\gamma} + \partial_\beta X^\beta g^{\alpha\gamma} = 2g^{\gamma\beta}(\partial_\beta X^\alpha + X^\delta \partial^\alpha u \partial_\delta \partial_\beta u + X^\alpha \partial^\delta u \partial_\delta \partial_\beta u).$$

To estimate this, our starting point is the relation (7.36), together with the bounds (7.37) for r^α. Denoting

$$h^{\alpha\delta} = X^\delta \partial^\alpha u + X^\alpha \partial^\delta u$$

we write

$$
\begin{aligned}
c_X^{\alpha\gamma} + \partial_\beta X^\beta g^{\alpha\gamma} &= g^{\gamma\beta}(\partial_\beta X^\alpha + h^{\alpha\delta} \partial_\delta \partial_\beta u)\\
&= T_{g^{\gamma\beta}} \partial_\beta r^\alpha + (T_{g^{\gamma\beta} h^{\alpha\delta}} - T_{g^{\gamma\beta}} T_{h^{\alpha\delta}}) \partial_\alpha \partial_\beta u\\
&\quad + T_{\partial_\beta X^\alpha} g^{\gamma\beta} + T_{\partial_\delta \partial_\beta u}[g^{\gamma\beta} h^{\alpha\delta}] + \Pi(\partial_\beta X^\alpha, g^{\gamma\beta}) + \Pi(\partial_\delta \partial_\beta u, g^{\gamma\beta} h^{\alpha\delta}).
\end{aligned}
$$

For the r^α term we use (7.37), for the next term we use the earlier bound (5.32) and the terms on the last line are estimated directly using the algebra property for $\mathfrak{C}_0$ and the bilinear estimate (5.25). $\square$

7.2.3. *Paradifferential energy estimates associated to X.* Now we use our vector field X to prove the balanced energy estimates for v. To do this, we repeat the computations leading to the key energy relations (7.23) and (7.24) at the paradifferential level.

To fix the notations, we denote by $T_{\tilde P}$ the operator in (7.8),

$$T_{\tilde P} = \partial_\alpha T_{\tilde g^{\alpha\beta}} \partial_\beta - T_{\tilde A^\gamma} \partial_\gamma.$$

By a slight abuse of notation, this is not exactly the same as the Weyl quantized operator with the corresponding symbol, though the difference between the two can be seen to be balanced and thus perturbative in our analysis.

For our multiplier, inspired by the energy relation (7.26), we will use the paradifferential operator

def-hX (7.38) $$T_{\tilde{\mathfrak{M}}} := 2T_{\tilde X^\alpha} \partial_\alpha + T_{\tilde q}.$$

Here ideally we would like to have

$$\tilde q = -g^{00} \partial_\alpha X^\alpha.$$

However, such a choice causes some technical difficulties due to the lack of sufficient time regularity of $\tilde q$. To avoid this, we will forego the above explicit expression for $\tilde q$, and instead ask for $\tilde q$ to satisfy the following two properties:

• it is close to the ideal setting,

choose-d (7.39) $$|\tilde q - g^{00} \partial_\alpha X^\alpha| \lesssim \mathcal{B}^2.$$

• it has the form $\tilde q = \partial_x q_1$, where $q_1 \in \mathfrak{P}$.

We remark that the obvious choice $\tilde{q}_0 := -g^{00}\partial_\alpha X^\alpha$ for the first criteria does not satisfy the second criteria, as it contains expressions involving $\partial_t^2 u$. However, by definition we have $\tilde{q}_0 \in \mathfrak{DP}$, therefore, a good approximation $\tilde{q}$ for $\tilde{q}_0$ as above exists by Lemma 6.8. Note that for this it suffices to use the fact that $X^\alpha \in \mathfrak{P}$ separately for each α, rather than the more precise representation in (7.36).

Now we implement the multiplier method to prove energy estimates in the paradifferential setting. We recall our objective, which is to establish an integral energy identity of the form

(7.40)
$$\iint T_{\tilde{P}}v \cdot T_{\tilde{\mathfrak{M}}}v \, dxdt = E_X(v(t))|_0^T + \int_0^T O(\mathcal{B}^2)\|v(t)\|_{\mathcal{H}^1}^2 \, dt$$

for a suitable positive definite energy functional E_X in $\mathcal{H}$,

(7.41)
$$E_X(v(t)) \approx \|v[t]\|_{\mathcal{H}^1}^2.$$

This may also be interpreted as a differential energy identity,

(7.42)
$$\frac{d}{dt}E_X(v(t)) = \int T_{\tilde{P}}v \cdot T_{\tilde{\mathfrak{M}}}v \, dx + O(\mathcal{B}^2)\|v(t)\|_{\mathcal{H}^1}^2.$$

Notation for errors: There are two types of error/correction terms that appear in our computations:

- Corrections in the energy functional. Here we will denote by $Err(\mathcal{A}^\sharp)$ any fixed time expressions that have size $O(\mathcal{A}^\sharp)\|v[t]\|_{\mathcal{H}^1}^2$. A typical example here is a lower order term of the form
  $$\int_{\mathbb{R}^n} \partial v \cdot T_q v \, dx, \qquad q \in \partial_x\mathfrak{P},$$
  where
  $$\|P_{<k}q\|_{L^\infty} \lesssim 2^k \mathcal{A}.$$

- Corrections in the energy flux term. These are like the last term on the right in (7.40), respectively (7.42). For brevity we will denote the admissible errors in the two identities by $Err(\mathcal{B}^2)$.

To establish (7.40), we consider the contributions of the two terms in $T_{\tilde{\mathfrak{M}}}$.

I. The contribution of $T_{\tilde{X}^\alpha}\partial_\alpha$. Integrating by parts and commuting, this is given by

$$
\begin{aligned}
I_X &= \int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\tilde{P}_A}v \cdot T_{\tilde{X}^\gamma}\partial_\gamma v \, dxdt \\
&= \int_0^T\!\!\!\int_{\mathbb{R}^n} (\partial_\alpha T_{\tilde{g}^{\alpha\beta}}\partial_\beta v - T_{\tilde{A}^\alpha}\partial_\alpha v) \cdot T_{\tilde{X}^\gamma}\partial_\gamma v \, dxdt \\
&= \int_0^T\!\!\!\int_{\mathbb{R}^n} \partial_\alpha T_{\tilde{X}^\gamma}T_{\tilde{g}^{\alpha\beta}}\partial_\beta v \cdot \partial_\gamma v - T_{\partial_\alpha \tilde{X}^\gamma}T_{\tilde{g}^{\alpha\beta}}\partial_\beta v \cdot \partial_\gamma v - T_{\tilde{A}^\alpha}\partial_\alpha v \cdot T_{\tilde{X}^\gamma}\partial_\gamma v \, dxdt \\
&= \int_0^T\!\!\!\int_{\mathbb{R}^n} \partial_\gamma T_{\tilde{X}^\gamma}T_{\tilde{g}^{\alpha\beta}}\partial_\beta v \cdot \partial_\alpha v - T_{\partial_\alpha \tilde{X}^\gamma}T_{\tilde{g}^{\alpha\beta}}\partial_\beta v \cdot \partial_\gamma v - T_{\tilde{A}^\alpha}\partial_\alpha v \cdot T_{\tilde{X}^\gamma}\partial_\gamma v \, dxdt \\
&\quad + \int T_{\tilde{X}^\gamma}T_{\tilde{g}^{0\beta}}\partial_\beta v \cdot \partial_\gamma v - T_{\tilde{X}^0}T_{\tilde{g}^{\alpha\beta}}\partial_\beta v \cdot \partial_\alpha v \, dx \Big|_0^T
\end{aligned}
$$

63

so we obtain

$$I_X = \int_0^T\!\!\int_{\mathbb{R}^n} \frac{1}{2}(\partial_\gamma T_{\tilde{X}^\gamma} T_{\tilde{g}^{\alpha\beta}} - T_{\tilde{g}^{\alpha\beta}} T_{\tilde{X}^\gamma} \partial_\gamma)\partial_\beta v \cdot \partial_\alpha v - T_{\partial_\alpha \tilde{X}^\gamma} T_{\tilde{g}^{\alpha\beta}} \partial_\beta v \cdot \partial_\gamma v - T_{\tilde{A}^\alpha} \partial_\alpha v \cdot T_{\tilde{X}^\gamma} \partial_\gamma v \, dxdt$$

$$+ \int T_{\tilde{X}^\gamma} T_{\tilde{g}^{0\beta}} \partial_\beta v \cdot \partial_\gamma v - T_{\tilde{X}^0} T_{\tilde{g}^{\alpha\beta}} \partial_\beta v \cdot \partial_\alpha v + \frac{1}{2} T_{\tilde{X}^0} T_{\tilde{g}^{\alpha\beta}} \partial_\beta v \cdot \partial_\alpha v \, dx \Big|_0^T .$$

For the double integral we peel off some perturbative contributions. The first term has a commutator structure, and we distinguish several cases. If $(\alpha, \beta) = (0,0)$, then we simply write

$$\partial_\gamma T_{\tilde{X}^\gamma} T_{\tilde{g}^{00}} - T_{\tilde{g}^{00}} T_{\tilde{X}^\gamma} \partial_\gamma = \tilde{g}^{00} T_{\partial_\gamma \tilde{X}^\gamma}.$$

If $(\alpha, \beta) = (0, j)$ then we commute the derivative first,

$$\partial_\gamma T_{\tilde{X}^\gamma} T_{\tilde{g}^{0j}} - T_{\tilde{g}^{0j}} T_{\tilde{X}^\gamma} \partial_\gamma = T_{\partial_\gamma \tilde{X}^\gamma} T_{\tilde{g}^{0j}} + T_{\tilde{X}^\gamma} T_{\partial_\gamma \tilde{g}^{0j}} + [T_{\tilde{X}^\gamma}, T_{\tilde{g}^{0j}}]\partial_\gamma,$$

where the contribution of the commutator term is estimated using Lemma 2.4,

$$\|[T_{\tilde{X}^\gamma}, T_{\tilde{g}^{0j}}]\partial_j\|_{L^2 \to L^2} \lesssim \mathcal{B}^2.$$

If $(\alpha, \beta) = (j, 0)$ then we commute the paraproducts first,

$$\partial_\gamma T_{\tilde{X}^\gamma} T_{\tilde{g}^{j0}} - T_{\tilde{g}^{j0}} T_{\tilde{X}^\gamma} \partial_\gamma = T_{\partial_\gamma \tilde{g}^{0j}} T_{\tilde{X}^\gamma} + T_{\tilde{g}^{j0}} T_{\partial_\gamma \tilde{X}^\gamma} + \partial_\gamma [T_{\tilde{X}^\gamma}, T_{\tilde{g}^{j0}}],$$

where the contribution of the commutator is again perturbative once we integrate by parts with respect to x^γ. If $\gamma = 0$ then this integration by parts contributes to the energy with the expression

$$\int [T_{\tilde{X}^0}, T_{\tilde{g}^{j0}}]\partial_0 v \cdot \partial_j v \, dx,$$

which also plays a perturbative role. For the double paraproducts we use Lemma 2.7 to compound them, as in

$$\|T_g T_{\partial h} - T_{g\partial h}\|_{L^2 \to L^2} \lesssim \mathcal{B}^2.$$

We arrive at the relation

(7.43) $$I_X = \iint T_{\tilde{P}_A} v \cdot T_{\tilde{X}^\gamma} \partial_\gamma v \, dxdt = \iint T_{c_X^{\alpha\beta}} \partial_\alpha v \cdot \partial_\beta v \, dxdt + E_X(v)\big|_0^T + Err(\mathcal{B}^2),$$

where we recall that $c_X^{\alpha\beta}$ is given by the relation (7.27), and the energy functional E_X is given by

$$E_X(v) = \int T_{\tilde{X}^\gamma} T_{\tilde{g}^{0\beta}} \partial_\beta v \cdot \partial_\gamma v - T_{\tilde{X}^0} T_{\tilde{g}^{\alpha\beta}} \partial_\beta v \cdot \partial_\alpha v + \frac{1}{2} T_{\tilde{X}^0} T_{\tilde{g}^{\alpha\beta}} \partial_\beta v \cdot \partial_\alpha v + [T_{\tilde{X}^0}, T_{\tilde{g}^{j0}}]\partial_0 v \cdot \partial_j v \, dx.$$

Here we may compound all double paraproducts and discard the commutator term, at the expense of $Err(\mathcal{A}^\sharp)$ errors. We arrive at

(7.44) $$E_X(v) = \int T_{e_X^{\alpha\beta}} \partial_\alpha v \partial_\beta v \, dx + Err(\mathcal{A}^\sharp),$$

with $e_X^{\alpha\beta}$ given by (7.16), and which therefore belongs to $\mathfrak{P}$ modulo constants. Since $X = \partial_t + O(\mathcal{A}^\sharp)$ is uniformly time-like, it follows that this matrix is positive definite, which implies the positivity property in (7.41).

64

II. The contribution of T_q. Here we need to consider the integral

$$I_q = \iint (\partial_\alpha T_{\tilde{g}^{\alpha\beta}} \partial_\beta + T_{\tilde{A}^\gamma} \partial_\gamma) v \cdot T_{\tilde{q}} v \, dxdt,$$

where we recall that $\tilde{q} = \partial_x q_1$ with $q_1 \in \mathfrak{P}$. The contribution of $\tilde{A}$ is directly perturbative, as $\tilde{A} \in \mathfrak{DP} \subset \mathfrak{DC}$; then one can use the $\mathfrak{DC}$ decomposition $\tilde{A} = \tilde{A}_1 + \tilde{A}_2$ as in Definition 5.5,(b), pairing each of the two associated bounds in (5.20) with the $\mathcal{B}$, respectively the $\mathcal{A}^\sharp$ bound in the $\mathfrak{C}_0$ norm of q_1:

$$\|T_{\tilde{A}_1^\gamma} T_{\tilde{q}}\|_{H^1 \to L^2} \lesssim \mathcal{B}^2, \qquad \|T_{\tilde{A}_2^\gamma} T_{\tilde{q}}\|_{H^1 \to L^2} \lesssim \mathcal{A}^\sharp \mathcal{B}^2.$$

Integrating by parts and using Lemmas 2.7, 2.8, we compute

$$I_q = - \iint T_{\tilde{g}^{\alpha\beta}} \partial_\beta v \cdot (T_{\tilde{q}} \partial_\alpha + T_{\partial_\alpha \tilde{q}}) v \, dxdt + \int T_{\tilde{g}^{0\beta}} \partial_\beta v \cdot T_{\tilde{q}} v \, dx \Big|_0^T + Err(\mathcal{B}^2)$$

$$= - \iint T_{\tilde{g}^{\alpha\beta} \tilde{q}} \partial_\beta v \cdot \partial_\alpha v \, dxdt - \iint \partial_\beta v \cdot T_{T_{\tilde{g}^{\alpha\beta}} \partial_\alpha \tilde{q}} v \, dxdt + \int T_{\tilde{g}^{0\beta}} \partial_\beta v \cdot T_{\tilde{q}} v \, dx \Big|_0^T + Err(\mathcal{B}^2).$$

Here the first term on the right is the one we want and the last term on the right yields an energy correction which is perturbative, i.e. of size $Err(\mathcal{A}^\sharp)$. It remains to show that the second term, which we shall denote by I_q^2, also yields only perturbative contributions. Heuristically, this should be relatively simple, in that we can integrate once more by parts, to obtain

$$I_q^2 := - \iint \partial_\beta v \cdot T_{T_{\tilde{g}^{\alpha\beta}} \partial_\alpha \tilde{q}} v \, dxdt = \frac{1}{2} \iint v \cdot T_{\partial_\beta T_{\tilde{g}^{\alpha\beta}} \partial_\alpha \tilde{q}} v \, dxdt - \frac{1}{2} \int v \cdot T_{T_{\tilde{g}^{\alpha0}} \partial_\alpha \tilde{q}} v \, dx \Big|_0^T.$$

Here we could estimate both integrals perturbatively and conclude directly if we knew that

$$\|P_{<k} \partial_\beta T_{\tilde{g}^{\alpha\beta}} \partial_\alpha \tilde{q}\|_{L^\infty} \lesssim 2^{2k} \mathcal{B}^2, \qquad \|P_{<k} \partial_\alpha \tilde{q}\|_{L^\infty} \lesssim 2^{2k} \mathcal{A}^\sharp.$$

Both of these bounds would be true if q contained no time derivatives of u in its expression. However, this is too much to hope for, so a more careful argument is needed. The first step in this argument has already been carried out earlier, where we saw that we may take $\tilde{q}$ of the form $\tilde{q} = \partial_x q_1$ with $q_1 \in \mathfrak{P}$. This removes one of the two potential time derivatives in q, but not the second. We can use this property to write

$$I_q^2 = \iint \partial_\beta v \cdot T_{\partial_x T_{\tilde{g}^{\alpha\beta}} \partial_\alpha q_1} v \, dxdt - \iint \partial_\beta v \cdot T_{T_{\partial_x \tilde{g}^{\alpha\beta}} \partial_\alpha q_1} v \, dxdt,$$

where the uniform bound

$$\|P_{<k} T_{\partial_x \tilde{g}^{\alpha\beta}} \partial_\alpha q_1\|_{L^\infty} \lesssim 2^k \mathcal{B}^2$$

shows that we can treat the second term perturbatively, to get

$$I_q^2 = \iint \partial_\beta v \cdot T_{\partial_x T_{\tilde{g}^{\alpha\beta}} \partial_\alpha q_1} v \, dxdt + Err(\mathcal{B}^2).$$

At this point, we can use the fact that $q_1 \in \mathfrak{P}$ implies that q_1 solves an approximate paradifferential wave equation. The precise statement we use is the one in Lemma 6.4, which yields the representation

$$\partial_\alpha T_{\tilde{g}^{\alpha\beta}} \partial_\beta q_1 = \partial_\alpha f^\alpha$$

65

with

(7.45) $$\|f^\alpha\|_{L^\infty} \lesssim \mathcal{B}^2, \qquad \|P_{<k}(\partial_0 q_1 - f^0)\|_{L^\infty} \lesssim \mathcal{A}^\sharp.$$

We use this representation to refine the outcome of the naive integration by parts above,

$$
\begin{aligned}
I_q^2 &= -\frac{1}{2} \iint v \cdot T_{\partial_x \partial_\beta T_{\tilde{g}^{\alpha\beta}} \partial_\alpha q_1} v \, dx dt + \frac{1}{2} \int v \cdot T_{\partial_x T_{\tilde{g}^{\alpha 0}} \partial_\alpha q_1} v \, dx \Big|_0^T + Err(\mathcal{B}^2) \\
&= -\frac{1}{2} \iint v \cdot T_{\partial_x \partial_\alpha f^\alpha} v \, dx dt + \frac{1}{2} \int v \cdot T_{\partial_x T_{\tilde{g}^{\alpha 0}} \partial_\alpha q_1} v \, dx \Big|_0^T + Err(\mathcal{B}^2) \\
&= \iint \partial_\alpha v \cdot T_{\partial_x f^\alpha} v \, dx dt + \frac{1}{2} \int v \cdot T_{\partial_x (T_{\tilde{g}^{\alpha 0}} \partial_\alpha q_1 - f^0)} v \, dx \Big|_0^T + Err(\mathcal{B}^2).
\end{aligned}
$$

By the pointwise bound on f^α in (7.45), the first term is perturbative, i.e. $Err(\mathcal{B}^2)$. By the second bound in (7.45), the second term can be seen as a perturbative $Err(\mathcal{A}^\sharp)$ energy correction. We conclude that for I_q we have

(7.46) $\quad I_q = \iint T_{\tilde{g}^{\alpha\beta} q} \partial_\beta v \cdot \partial_\alpha v \, dx dt + \int T_{\tilde{g}^{0\beta}} \partial_\beta v \cdot T_q v + \frac{1}{2} v \cdot T_{\partial_x (T_{\tilde{g}^{\alpha 0}} \partial_\alpha q_1 - f^0)} v \, dx \Big|_0^T + Err(\mathcal{B}^2).$

III. Conclusion. To finish the proof of (7.40), and thus of Theorem 7.1 for $s = 0$, we combine the relations (7.43) and (7.46) to obtain

(7.47) $\qquad 2 \iint T_{\tilde{P}} v \cdot T_{\tilde{\mathfrak{M}}} v \, dx dt = \iint T_{c_X^{\alpha\beta} + \tilde{g}^{\alpha\beta} \tilde{q}} \partial_\alpha v \cdot \partial_\beta v \, dx dt + E_X(v(t))\big|_0^T + Err(\mathcal{B}^2),$

where E_X is redefined as the sum of the two contributions in (7.43) and (7.46), which still has the leading order term as in (7.44) plus an $Err(\mathcal{A}^\sharp)$ correction.

It remains to examine the paracoefficient in the integral on the right, and show that it has size $O(\mathcal{B}^2)$. At this point, we simply invoke the choice of our para-Killing vector field X in Lemma 7.4 for the first term (which we have not used so far), and the choice of $\tilde{q}$ in (7.39) for the second term, thereby completing the proof of (7.40).

7.3. The $H^{s+1} \times H^s$ bound for the linear paradifferential flow.

Here we prove Theorem 7.1 in the general case, where $s \neq 1$. The argument will be a more complex variation of the argument in the case $s = 1$, where paraproduct based multipliers have to be replaced by paradifferential multipliers.

7.3.1. *The conjugated equation.* For simplicity in notations we will consider the linear paradifferential equation in $\mathcal{H}^{s+1}$ with $s \neq 0$. We begin by setting $w = \langle D_x \rangle^s v$, which solves a perturbed linear paradifferential equation of the form

(7.48) $$(\partial_\alpha T_{\tilde{g}^{\alpha\beta}} \partial_\beta - T_{\tilde{A}^\gamma} \partial_\gamma) w = \langle D_x \rangle^s f + \tilde{B} w,$$

where the conjugation error $\tilde{B}$ in the new source term is given by

(7.49) $$\tilde{B} = \langle D_x \rangle^s [\partial_\alpha T_{\tilde{g}^{\alpha\beta}} \partial_\beta - T_{\tilde{A}^\gamma} \partial_\gamma, \langle D_x \rangle^{-s}].$$

Then we need to construct an $H^1 \times L^2$ balanced energy for the solution w to (7.48).

66

We note that $\tilde{B}$ is a paradifferential operator, whose principal symbol $\tilde{b}_0$ is homogeneous of order one and a first degree polynomial in the time variable ξ_0, and is given by

$$\tilde{b}_0(x,\xi) = -i|\xi'|^s\{\tilde{g}^{\alpha\beta}\xi_\alpha\xi_\beta, |\xi'|^{-s}\}.$$

Using the expression (3.9) for the derivatives of the metric g, this can be further written in the form

(7.50)
$$\tilde{b}_0(x,\xi) = 2is(\partial^\beta u\tilde{g}^{\alpha\nu}\partial_j\partial_\nu u - \partial^0 u\tilde{g}^{0\nu}\partial_j\partial_\nu u\tilde{g}^{\alpha\beta})\xi_\alpha\xi_\beta\xi_j|\xi'|^{-2}$$
$$:= 2is\tilde{b}_0^\alpha\xi_\alpha,$$

where

(7.51)
$$\tilde{b}_0^\alpha = (\partial^\beta u\,\tilde{g}^{\alpha\nu} - \partial^0 u\tilde{g}^{0\nu}\tilde{g}^{\alpha\beta})\partial_j\partial_\nu u\,\xi_\beta\xi_j|\xi'|^{-2}.$$

Here the unbalanced part of the coefficients corresponds to the case when the factor $\partial^2 u$ is higher frequency compared to the $\partial^\beta u$ and $\tilde{g}^{\alpha\nu}$ factors. The important feature is that, at the operator level, $T_{\tilde{b}_0^\gamma}\partial_\gamma w$ presents a null form structure of the type $Q_0(\partial u, w)$, with added more regular paradifferential coefficients in $\mathfrak{P}$.

We switch the leading term $2sT_{\tilde{b}_0^\gamma}\partial_\gamma$ to the left hand side of the equation; there it will play a role similar to the gradient term $\tilde{A}^\gamma\partial_\gamma$. The remainder $\tilde{B} - 2sT_{\tilde{b}_0^\gamma}\partial_\gamma$ will play a secondary role; one should think of it as renormalizable, though we will achieve this at the level of the energy, via an energy correction, rather than through an actual normal form transformation. Our equation (7.48) becomes

(7.52)
$$(\partial_\alpha T_{\tilde{g}^{\alpha\beta}}\partial_\beta - T_{\tilde{A}^\gamma}\partial_\gamma - 2sT_{\tilde{b}_0^\gamma}\partial_\gamma)w = \langle D_x\rangle^s f + (\tilde{B} - 2sT_{\tilde{b}_0^\gamma}\partial_\gamma)w,$$

where the leading operator is denoted by

(7.53)
$$T_{\tilde{P}_B} = \partial_\alpha T_{\tilde{g}^{\alpha\beta}}\partial_\beta - T_{\tilde{A}^\gamma}\partial_\gamma - 2sT_{\tilde{b}_0^\gamma}\partial_\gamma.$$

As in the previous case of the $H^1 \times L^2$ bounds, our strategy will be to construct a suitable vector field, or multiplier, denoted $\tilde{X}_s$, which depends only on the principal symbol $\tilde{b}_0$ above, and which formally generates a balanced energy estimate at the leading order. Then, reinterpreting all the analysis at the paradifferential level, we will rigorously prove that the generated energy satisfies favourable, balanced bounds.

7.3.2. *The multiplier $\tilde{X}_s$.* In the previous section, the multiplier $\tilde{X} \in \mathfrak{P}$ was a well-chosen vector field which belongs to our space $\mathfrak{P}$ of paracontrolled distributions. Here, this can no longer work due to the the presence of the operator $T_{\tilde{b}_0}$, which is a pseudodifferential rather than a differential operator. For this reason we will instead use a pseudodifferential "vector field" $i\tilde{X}_s$, where $\tilde{X}_s$ has a real, odd symbol of the form

$$\tilde{X}_s(x,\xi) = \tilde{X}_{s1}(x,\xi') + \tilde{X}_{s0}(x,\xi')\xi_0, \qquad X_{sj} \in \mathfrak{P}S^j,$$

which will be homogeneous away from frequency zero. We carefully note that we want the symbol $\tilde{X}_s$ to be a first order polynomial in ξ_0; this is important so that we can still do integration by parts in time and have a well defined fixed time energy. The symbol $\tilde{X}_s$ may be interpreted as a pseudodifferential operator using the Weyl paradifferential quantization,

(7.54)
$$T_{i\tilde{X}_s} = iT_{\tilde{X}_{s1}} + T_{\tilde{X}_{s0}}\partial_0 + \frac{1}{2}T_{\partial_0\tilde{X}_{s0}}.$$

67

However, as in the $s = 0$ case, we will allow a more general choice for the zero order component, and work instead with the modified multiplier

(7.55)
$$T_{i\tilde{\mathfrak{m}}_s} := iT_{\tilde{X}_{s1}} + T_{\tilde{X}_{s0}}\partial_0 + \frac{1}{2}T_{\tilde{Y}_{s0}},$$

where the real, even zero order symbol $\tilde{Y}_{s0} \in \partial_x \mathfrak{P}S^0$ will be carefully chosen later on in order to provide an appropriate Lagrangian correction in our energy estimates.

Repeating the heuristic computation in the previous subsection, in the absence of time boundaries we have an identity of the form

(7.56)
$$2\iint T_{\tilde{P}_{A,B}}v \cdot (iT_{\tilde{X}_{s1}} + T_{\tilde{X}_{s0}}\partial_0)v\,dxdt = \iint c_{\tilde{X}_s,B}(v,v)\,dxdt,$$

where $c_{\tilde{X}_s,B}(v,v)$ is a bilinear form whose principal symbol $c_{\tilde{X}_s,B}$ is of order two,

(7.57)
$$c_{\tilde{X}_s,B}(x,\xi) = \{\tilde{p}, \tilde{X}_s\}(x,\xi) + 2\tilde{X}_s(x,\xi)(\tilde{A}^\gamma + 2s\tilde{b}_0^\gamma(x,\xi))\xi_\gamma - \partial_0\tilde{X}_{s0}\tilde{p}(x,\xi).$$

The objective would now be to choose the symbols $\tilde{X}_{sj} \in \mathfrak{P}S^j$ so that we cancel the unbalanced part of the symbol $c_{\tilde{X}_s,B}$. However, it is immediately clear that this may be a bit too much to ask, as it conflicts with the requirement that $\tilde{X}_s$ is a first degree polynomial in ξ_0. Hence, as a substitute, we will seek to achieve this cancellation on the characteristic set $p(x,\xi) = 0$. Then, instead of asking for

$$c_{\tilde{X}_s,B} \overset{bal}{\approx} 0,$$

we will settle for the slightly weaker property

$$c_{\tilde{X}_s,B}(x,\xi) \overset{bal}{\approx} \tilde{Y}_{s0}(x,\xi') \cdot \tilde{p}(x,\xi),$$

where $\tilde{Y}_{s0} \in \mathfrak{DP}S^0$ is a purely spatial zero homogeneous symbol, with the spatial dependence at the level of $\partial^2 u$. This term will be harmless, as we will also be able to remove it in our energy energy estimates with a Lagrangian correction, by making a good choice for $\tilde{Y}_{s0}$.

An additional requirement on our paradifferential "vector field" $\tilde{X}_s$ will be that, in the energy estimate generated by $\tilde{X}_s$, the associated energy functional $E_{\tilde{X}_s}$ should be positive definite at the level of it principal part. Earlier, in the case when X was a vector field, this requirement was identified, via the energy momentum tensor, with the property that X is forward time-like. Here we will generalize this notion to symbols:

Definition 7.5. *We say that the (real) symbol $X = \xi_0 X_0 + X_1 \in C^0 S^1$ is forward time-like if the following two properties hold:*
a) $X_0(x,\xi') > 0$.
b) $X(x,\xi_0^1,\xi')X(x,\xi_0^2,\xi') < 0$, where $\xi_0^1(x,\xi') < \xi_0^2(x,\xi')$ are the two real zeros of $p(x,\xi)$ as a polynomial of ξ_0.

We remark that, using X as a multiplier, relative to the metric g, we will obtain an energy functional which at leading order can be described via the symbol

(7.58)
$$e_X(x,\xi) = g^{\alpha\beta}\xi_\alpha\xi_\beta X_0 - 2g^{0\alpha}\xi_\alpha(X_1 + X_0\xi_0)$$
$$:= e_X^0(x,\xi')\xi_0^2 + e_X^1(x,\xi')\xi_0 + e_X^2(x,\xi'),$$

68

which should be compared with the expression ($\overset{\text{eX}}{7.16}$) defined earlier in terms of the energy momentum tensor in the case when X is a vector field. Correspondingly, we define the energy functional

(7.59)
$$E_X[w] = \int -T_{e^0_X} \partial_t w \cdot \partial_t w + T_{ie^1_X} w \cdot \partial_t w + T_{e^2_X} w \cdot w \, dx.$$

The main property of forward time-like symbols is as follows:

Lemma 7.6. *The symbol e_X is positive definite iff X is forward time-like.*

Proof. Assuming X_0 is nonzero, we represent X in the form

$$X = X_0(a_1(\xi_0 - \xi_0^1) + a_2(\xi_0 - \xi_0^2)),$$

where $a_1 + a_2 = 1$. Then e_X has the form

$$e_X = g^{00}(\xi_0 - \xi_0^1)(\xi_0 - \xi_0^2)X_0 - g^{00}(2\xi_0 - \xi_0^1 - \xi_0^2)X_0(a_1(\xi_0 - \xi_0^1) + a_2(\xi_0 - \xi_0^2))$$
$$= -g^{00}X_0[a_1(\xi_0 - \xi_0^1)^2 + a_2(\xi_0 - \xi_0^2)^2].$$

Here $g^{00} = -1$ and at least one of a_1 and a_2 are positive. Then e_X is positive definite iff $X_0 > 0$ and $a_1, a_2 > 0$. This is easily seen to be equivalent with the forward time-like condition in the above definition.

$\square$

7.3.3. *The construction of $\tilde{X}^s$.* Here we return to the matter of choosing $\tilde{X}_s$, whose properties almost exactly mirror those of the vector field $\tilde{X}$ in the previous subsection:

Proposition 7.7. *There exists a real, odd homogeneous symbol of order one $\tilde{X}_s \in \xi_0 + \mathfrak{P}S^1$, which is a first degree polynomial in ξ_0, so that:*
i) $\tilde{X}_s$ is forward time-like.
ii) The principal symbol $c_{\tilde{X}_s,B}$ of the $\tilde{X}_s$ energy flux admits a representation of the form

(7.60)
$$c_{\tilde{X}_s,B}(x,\xi) = \tilde{q}_2(x,\xi) + \tilde{q}_0(x,\xi')\tilde{p}(x,\xi),$$

where $\tilde{q}_2$ is balanced,

(7.61)
$$\|\tilde{q}_2\|_{L^\infty S^2} \lesssim \mathcal{B}^2,$$

and $\tilde{q}_0$ has $\mathfrak{DP}$ type regularity,

(7.62)
$$\|\tilde{q}_0\|_{\mathfrak{DP}S^0} \lesssim \mathcal{A}^\sharp.$$

iii) The symbol $\tilde{X}_s$ admits the $\mathfrak{P}S^1$ representation

(7.63)
$$\tilde{X}_s = \xi_0 + T_{a^\gamma}\partial_\gamma u + r_s,$$

where the para-coefficients $a^\gamma(x,\xi) = a_1^\gamma(x,\xi') + a_0^\gamma(x,\xi')$ with $a_j^\gamma \in \mathfrak{P}S^j$ have the form

(7.64)
$$a^\gamma = -\xi_\delta \partial^\delta u \partial_{\xi_\gamma} \tilde{X}_s - \partial^\gamma u \tilde{X}_s - 2\tilde{X}^s \partial^0 u \tilde{g}^{0\gamma} - s\xi_\delta \partial_{\xi_\gamma} \log|\xi'|^2 \partial^\delta u \tilde{X}_s + pq_0^\gamma$$

with $q_0^\gamma \in \mathfrak{P}S^0$, independent of ξ_0.

69

From the perspective of energy estimates, it might seem that parts (i) and (ii) are the important ones. However, part (ii) will be seen as an immediate consequence of the representation in part (iii), which thus can be thought of as the more fundamental property. Also, in the proof of the energy estimates it will on occasion be more convenient to directly use (7.64). In the sequel we will refer to a^γ as the para-coefficients of $\tilde{X}_s$. We note that the choice of q_0^γ is uniquely determined by the requirement that a^γ are first degree polynomials in ξ_0.

Proof. It will be somewhat easier to construct the corresponding symbol X_s as associated to p, rather than to $\tilde{p}$; this avoids the slight symmetry breaking in the transition from p to $\tilde{p}$. Precisely, we will choose $\tilde{X}_s$ of the form

$$\tilde{X}_s = g^{00} X_s,$$

and then express $c_{\tilde{X}_s,B}$ in terms of X_s as follows:

$$c_{\tilde{X}_s,B} = \{p, X_s\} + 2X_s(A^\gamma + 2sb_0^\gamma)\xi_\gamma + q_{00}p,$$

where q_{00} is given by

$$q_{00} = \{X_{s1}, \log g^{00}\} + \xi_0 \{X_{s0}, \log g^{00}\} - \partial_0 X_{s0} - 2s X_s \partial^0 u\, \tilde{g}^{0\nu} \partial_j \partial_\nu u \xi_j |\xi'|^{-2},$$

and b_0^γ has the form

(7.65)
$$b_0^\gamma = \partial^\beta u\, g^{\gamma\nu} \partial_j \partial_\nu u\, \xi_\beta \xi_j |\xi'|^{-2}.$$

Here we have separated the two terms in $\tilde{b}_0^\gamma$; the first has contributed to b_0^γ, while the second has contributed the last term in the Lagrangian coefficient q_{00}.

For clarity, we note that the exact expression of q_{00} is not important, we will only use the fact that $q_{00} \in \mathfrak{DPS}^0$. On the other hand, for b_0^γ we will need the fact that it has a null structure.

Now we restate the proposition in terms of the new symbol X_s. Our goal will be to find X_s in the same class as $\tilde{X}_s$, so that the reduced symbol

(7.66)
$$c_{X_s,B}^{red} = \{p, X_s\} + 2X_s(A^\gamma + 2sb_0^\gamma)\xi_\gamma$$

can be represented in the form

(7.67)
$$c_{X_s,B}^{red} = q_2(x,\xi) + q_0(x,\xi)p(x,\xi).$$

Here there is a small twist in the argument. While $c_{X_s,B}$ is a second degree polynomial in ξ_0, this is no longer the case for $c_{X_s,B}^{red}$, which contains the term $\xi_0^3 \{g^{00}, X_{s0}\}$. For this reason, in (7.67) we apriori have to allow for symbols q_2, respectively q_0 which are third, respectively first degree polynomials in ξ_0. However, we can eliminate the ξ_0^3 term in q_2 with a ξ_0 correction in q_0. Then, returning to $c_{X_s,B}$, we obtain the representation (7.60) with $\tilde{q}_2$ of second degree and $\tilde{q}_0$ of first degree. But $c_{X_s,B}$ is a second degree polynomial in ξ_0, so we finally conclude that $\tilde{q}_0$ must be independent of ξ_0.

We now proceed to construct the symbol X_s. As a first step in the proof, we seek to obtain a variant X_s^0 of the symbol X_s where we drop the requirement that X_s^0 is a first order polynomial in ξ_0 but we ask for the stronger property that the associated symbol $c_{X_s^0,B}^{red}$ is fully balanced, which corresponds to $q_0 = 0$. Then, at the end, we choose X_s to be the first degree polynomial in ξ_0 that matches X_s^0 at the two roots of $p(x,\xi) = 0$ viewed as a polynomial in ξ_0.

The relation we seek for X_s^0 to satisfy on the characteristic set of p is

<div style="float:left">tX-a</div> (7.68)
$$c_{X_s^0,B}^{red} \overset{bal}{\approx} 0,$$

where, using the expressions (3.13) and (7.65) for A^γ and b_0^γ,

$$c_{X_s^0,B}^{red} = \{g^{\alpha\beta}\xi_\alpha\xi_\beta, X_s^0\} + 2\partial^\alpha u \xi_\delta g^{\beta\delta}\partial_\alpha\partial_\beta u \cdot X_s^0 + 4s X_s^0 \partial^\beta u g^{\alpha\nu}\partial_j\partial_\nu u\, \xi_\alpha\xi_\beta\xi_j |\xi'|^{-2}.$$

Here we recall the expression for the derivatives of g, see (3.9),

<div style="float:left">dg</div> (7.69)
$$\partial_\gamma g^{\alpha\beta}\xi_\alpha\xi_\beta = -2\partial^\beta u g^{\alpha\delta}\partial_\delta\partial_\gamma u\, \xi_\beta\xi_\alpha.$$

Substituting this in the previous expression for $c_{X_s^0,B}^{red}$, we need the following relation to hold modulo balanced terms:

$$2\xi_\alpha g^{\alpha\beta}\partial_\beta X_s^0 \overset{bal}{\approx} -2\partial^\delta u \xi_\delta \partial_{\xi_\gamma} X_s^0 \cdot \xi_\alpha g^{\alpha\beta}\partial_\beta\partial_\gamma u - 2X_s^0 \partial^\gamma u \cdot \xi_\alpha g^{\beta\alpha}\partial_\gamma\partial_\beta u$$
$$- 4s X_s^0 \partial^\delta u \xi_\delta \xi_j |\xi'|^{-2} \cdot \xi_\alpha g^{\alpha\beta}\partial_j\partial_\beta u.$$

We can rewrite this using the following operator

$$L = \xi_\alpha g^{\alpha\beta}\partial_\beta$$

in the form

<div style="float:left">tX-aa</div> (7.70) $$L X_s^0 \overset{bal}{\approx} -\xi_\delta \partial^\delta u \partial_{\xi_\gamma} X_s^0 \cdot L\partial_\gamma u - \partial^\gamma u X_s^0 \cdot L\partial_\gamma u - s\partial^\delta u X_s^0 \xi_\delta \partial_{\xi_j} \log |\xi'|^2 \cdot L\partial_j u.$$

By Lemma 7.4, we already have a solution X for $s = 0$. Thinking of this multiplicatively, it is then natural to look for X_s^0 of the form

$$X_s^0 = Z_s X,$$

where $Z_s \ll \partial u$ should be zero homogeneous in ξ and must satisfy

<div style="float:left">LY</div> (7.71) $$L Z_s \overset{bal}{\approx} -\xi_\delta \partial^\delta u (\partial_{\xi_\gamma} Z_s L\partial_\gamma u + s Z_s \partial_{\xi_j} \log |\xi'|^2 L\partial_j u).$$

We will also assume that Z_s is a positive symbol; this will help later with the time-like condition. Then we can rewrite the above relation as a condition for $\log Z_s$, namely

<div style="float:left">LlogY</div> (7.72) $$L \log Z_s \overset{bal}{\approx} -\xi_\delta \partial^\delta u (\partial_{\xi_\gamma}(\log Z_s) L\partial_\gamma u + s\partial_{\xi_j} \log |\xi'|^2 L\partial_j u).$$

Here the inhomogeneous term is linear in s, so we will also look for a solution $\log Z_s$ which is linear in s.

There is one last algebraic simplification, which is to replace Z_s by $\tilde{Z}_s = Z_s |\xi'|^{2s}$, which is $2s$-homogeneous, even, and inherits the property that $\log \tilde{Z}_s$ is linear in s. Then $\log \tilde{Z}_s$ must solve

<div style="float:left">LlogtY</div> (7.73) $$L \log \tilde{Z}_s \overset{bal}{\approx} -\xi_\delta \partial^\delta u \partial_{\xi_\gamma} \log \tilde{Z}_s L\partial_\gamma u.$$

Dispensing with the log, we replace this by

<div style="float:left">LlogtYa</div> (7.74) $$L \tilde{Z}_s \overset{bal}{\approx} -\xi_\delta \partial^\delta u \partial_{\xi_\gamma} \tilde{Z}_s L\partial_\gamma u.$$

Now we interpret the last relation paradifferentially, formally cancelling the L's. This suggests the following scheme to construct the dyadic parts of $\tilde{Z}_s$ inductively by setting

<div style="float:left">tZs-def</div> (7.75)
$$\tilde{Z}_{s0} = |\xi|^{2s}$$
$$\tilde{Z}_{sk} = \xi_\delta T_{\partial^\delta u} T_{\partial_{\xi_\gamma} \tilde{Z}_s} \partial_\gamma u_k, \qquad k \geq 1.$$

A-priori these dyadic parts have a nontrivial dependence on ξ, which would have to be tracked when considering the convergence in the k summation. However, since $\log \tilde{Z}_s$ is linear in s, it suffices to solve this for some nonzero s. The advantage here is that, if s is a positive integer (say $s = 1$) then all our iterates are polynomials of degree $2s$ in ξ. Hence the convergence issue disappears, due to our smallness condition for u, $\mathcal{A}^\sharp \ll 1$; this is exactly as in the construction of X in Section 7.2.2. This defines $\tilde{Z}_1$ as a positive definite polynomial in ξ of degree 2, so that $\tilde{Z}_1 = \xi^2(1 + O(\mathcal{A}^\sharp))$. Further, by the same argument as in the proof of Lemma 7.4, it follows that the coefficients of $\tilde{Z}_1 - Z_{10}$ are paracontrolled by ∂u; in other words, $\tilde{Z}_1 - \xi^2 \in \mathfrak{P}S^2$. In addition, by (7.75), it also follows that (a choice for) the para-coefficients of $\tilde{Z}_1$, as in Definition 6.1 is given by

$$\boxed{\texttt{tZ1-rep}} \quad (7.76) \qquad \tilde{Z}_1 = \xi^2 + T_{\xi_\delta \partial^\delta u \partial_{\xi_\gamma} \tilde{Z}_1} \partial_\gamma u + r.$$

Remark 7.8. We remark on the symbol $\tilde{Z}_1$, which is quadratic in ξ and para-commutes with p, in the sense that their Lie bracket is balanced and thus bounded by $\mathcal{B}^2$. This symbol plays a role that is similar to that of the first order symbol X constructed earlier.

Now that we have $\tilde{Z}_1$, for all real s we may define

$$\tilde{Z}_s = (\tilde{Z}_1)^s, \qquad X_s^0 = X(\tilde{Z}_1/|\xi'|^2)^s.$$

By Lemma 7.4 in the previous subsection we have $X - \xi_0 \ll \partial u$. Combining this with the similar property of $\tilde{Z}_1$, by the algebra and Moser properties of the space $\mathfrak{P}$ of paracontrolled distributions it follows that $\tilde{X}_s - \xi_0 \ll \partial u$. Finally, combining the representations of X and of $\tilde{Z}_1$ as paracontrolled distributions, as in (7.36) and (7.76), we obtain the corresponding $\mathfrak{P}$ representation for X_s^0 as in Definition 6.1 (see the relation (7.70))

$$\boxed{\texttt{tXs-PP}} \quad (7.77) \qquad X_s^0 = \xi_0 - \xi_\delta T_{\partial^\delta u \partial_{\xi_\gamma} X_s^0} \partial_\gamma u - T_{\partial^\gamma u X_s^0} \partial_\gamma u - s\xi_\delta \partial_{\xi_j} \log|\xi'|^2 T_{\partial^\delta u X_s^0} \partial_j u + r_s.$$

This in turn yields the desired conclusion that $c_{X_s^0, B}^{red}$ is balanced,

$$\boxed{\texttt{c-for-zX}} \quad (7.78) \qquad \|c_{X_s^0, B}^{red}\|_{L^\infty S^2} \lesssim \mathcal{B}^2.$$

Indeed, the equivalent form (7.70) can be obtained by directly applying the operator L in the relation (7.77); this is because the terms where the paracoefficients get differentiated are balanced, so we are left with the terms where L is applied to the main factors ∂u.

Now we carry out the last step of the proof, and define the symbol X_s as the unique first degree polynomial in ξ_0 with the property that

$$X_s(x, \xi) = X_s^0(x, \xi) \qquad \text{on} \quad \{g^{\alpha\beta}\xi_\alpha\xi_\beta = 0\}.$$

We now show that this choice for X_s has the desired properties.

Recall that $\xi_0^1(x, \xi') < \xi_0^2(x, \xi')$ are the two real zeros of $p(x, \xi)$ as a polynomial of ξ_0, which are 1-homogeneous and smooth in ξ' and are also smooth functions of ∂u. Thus,

$$\xi_0^1, \xi_0^2 \in \mathfrak{P}S^1.$$

The coefficients X_{s0} and X_{s1} in X_s are obtained by solving a linear system,

$$X_{s0} = \frac{X_s^0(x, \xi_0^2, \xi') - X_s^0(x, \xi_0^1, \xi')}{\xi_0^2 - \xi_0^1}, \qquad X_{s1} = \frac{X_s^0(x, \xi_0^2, \xi')\xi_0^1 - X_s^0(x, \xi_0^1, \xi')\xi_0^2}{\xi_0^2 - \xi_0^1}.$$

72

By the algebra and Moser properties of the space $\mathfrak{P}$ of paracontrolled distributions, it immediately follows that we have the symbol regularity properties $X_{s0} \in \mathfrak{P}S^0$ and $X_{s1} - 1 \in \mathfrak{P}S^1$. By construction we also have a smooth division,

(7.79)
$$X_s = X_s^0 + dp,$$

where we easily see that the quotient d has regularity $d \in \mathfrak{P}S^{-1}$ by computing directly

$$d(x,\xi) = \frac{X_s(x,\xi) - X_s^0(x,\xi)}{p(x,\xi)}$$
$$= \frac{1}{g^{00}(\xi_0^1 - \xi_0^2)} \left(\frac{X_s^0(x,\xi) - X_s^0(x,\xi_0^1,\xi')}{\xi_0 - \xi_0^1} - \frac{X_s^0(x,\xi) - X_s^0(x,\xi_0^2,\xi')}{\xi_0 - \xi_0^2} \right).$$

One may also interpret this as a form of the Malgrange preparation theorem in an easier case where the roots are separated.

We can now use (7.79) to relate $c_{X_s,B}^{red}$ with $c_{X_s^0,B}^{red}$:

$$c_{X_s,B}^{red} = c_{X_s^0,B}^{red} + p(\{p,d\} + d(A^\gamma + sb_0^\gamma)\xi_\gamma),$$

which is exactly the desired representation (7.67).

We can also use the relation (7.77) for part (iii) of the proposition. For this we first transition from X_s^0 to X_s. Using (7.79) and peeling off balanced terms, this gives the $\mathfrak{P}$ representation

$$X_s = \xi_0 - \xi_\delta T_{\partial^\delta u \partial_{\xi_\gamma} X_s} \partial_\gamma u - T_{\partial^\gamma u X_s} \partial_\gamma u - s\xi_\delta \partial_{\xi_j} \log |\xi'|^2 T_{\partial^\delta u X_s} \partial_j u$$
$$- T_p \left(\xi_\delta T_{\partial^\delta u \partial_{\xi_\gamma} d} \partial_\gamma u + T_{\partial^\gamma u d} \partial_\gamma u + s\xi_\delta \partial_{\xi_j} \log |\xi'|^2 T_{\partial^\delta u d} \partial_j u + d \right)$$
$$+ \left(T_d p + \xi_\delta T_{\partial^\delta u d \partial_{\xi_\gamma} p} \partial_\gamma u \right) + r_s.$$

In view of the paradifferential expansion (5.28) for $g^{\alpha\beta}$, in the last bracket there is a leading order cancellation,

$$T_d p + \xi_\delta T_{\partial^\delta u d \partial_{\xi_\gamma} p} \partial_\gamma u = r_s.$$

This implies that X_s admits a $\mathfrak{P}S^1$ representation of the form

(7.80) $\quad X_s = \xi_0 - \xi_\delta T_{\partial^\delta u \partial_{\xi_\gamma} X_s} \partial_\gamma u - T_{\partial^\gamma u X_s} \partial_\gamma u - \dfrac{s}{2} \xi_\delta \partial_{\xi_j} \log |\xi'|^2 T_{\partial^\delta u X_s} \partial_j u + T_p z + r_s.$

where $z \in \mathfrak{P}S^{-1}$. At this stage we only know that z and r_s are smooth as functions of ξ_0. On the other hand, the remaining terms are at most second degree polynomials in ξ_0. We claim that, without any restriction in generality, we may take z independent of ξ_0 and then r_s has to be at most second degree polynomial in ξ_0.

Subtracting a multiple of p from all the paracoefficients above and discarding balanced contributions, we may reduce to the case of a first degree polynomial, i.e. to a relation of the form

$$T_p z(x,\xi) + r_s(x,\xi) = Z_1(x,\xi') + Z_0(x,\xi')\xi_0,$$

where $z \in \mathfrak{P}S^{-1}$ and $Z_j \in \mathfrak{P}S^j$, while $\partial r_s = O(\mathcal{B}^2)$, with full symbol regularity in ξ. We will show that in this case we must have $\partial Z_j = O(\mathcal{B}^2)$, again with full symbol regularity. This would imply that we may include $T_p z$ into r_s, and thus take $z = 0$ in the last relation.

We begin by differentiating this relation in x and t, noting that $T_{\partial_p} z$ may be placed in ∂r_s:

$$T_p \partial z(x, \xi) + \partial r_s(x, \xi) = \partial Z_1(x, \xi') + \partial Z_0(x, \xi') \xi_0.$$

We may also perturbatively replace T_p with p, arriving at

(7.81) $$p \partial z(x, \xi) + r_s^1(x, \xi) = \partial Z_1(x, \xi') + \partial Z_0(x, \xi') \xi_0,$$

where r^1 has size $\mathcal{B}^2$ and symbol regularity,

$$|\partial_\xi^\alpha r_1| \lesssim \mathcal{B}^2 |\xi|^{1-|\alpha|}.$$

For fixed x, we examine this relation on the characteristic cone $C = \{p(x, \xi) = 0\}$. There we have

$$|\partial Z_1(x, \xi') + \partial Z_0(x, \xi') \xi_0| \lesssim \mathcal{B}^2 |\xi|,$$

so we may directly conclude that $\partial Z_1(x, \xi'), \partial Z_0(x, \xi') = O(\mathcal{B}^2)$. Next we need a similar bound for their derivatives $\partial_{\xi'}^\alpha \partial Z_1(x, \xi'), \partial_{\xi'}^\alpha \partial Z_0(x, \xi')$ with respect to ξ'. We fix x and argue by induction in $|\alpha|$. Then it suffices to use derivatives which are tangent to the cone at that x, which on one hand kill p but on the other hand give a full range of ξ' derivatives for Z_j. Hence, we may indeed assume that z is independent of ξ_0 and r_s is a second degree polynomial in ξ_0.

Lastly we switch from X_s to $\tilde{X}_s$. Again peeling off r_s type contributions, we have

$$\tilde{X}_s = g^{00} X_s$$

$$= \xi_0 + T_{g^{00}} X_s + T_{X_s} g^{00} + r_s$$

$$= \xi_0 - \xi_\delta T_{\partial^\delta u \partial_{\xi_\gamma} \tilde{X}_s} \partial_\gamma u - T_{\partial^\gamma u \tilde{X}_s} \partial_\gamma u - s \xi_\delta \partial_{\xi_j} \log |\xi'|^2 T_{\partial^\delta u \tilde{X}_s} \partial_j u + T_{\tilde{X}_s} \log g^{00} + T_p T_{g^{00}} z + r_s.$$

It remains to expand the fourth term, using Lemma 8.2:

$$T_{\tilde{X}_s} \log g^{00} = -2 T_{\tilde{X}^s \partial^0 u \tilde{g}^{0\alpha}} \partial_\alpha u + r_s.$$

This finally yields the representation (7.63) with the paracoefficients in (7.64), thereby concluding the proof of part (iii) of Proposition 7.7.

The final property of $\tilde{X}_s$ to be verified is that $\tilde{X}_s$ is time-like. This property is easily seen to depend only on the sign of the symbol $\tilde{X}_s$ on the characteristic set $\{p = 0\}$. But by construction, $\tilde{X}_s$ has the same sign as X_s there, which in turn has the same sign as the vector field X in Section 7.2.2. Then the time-like property for $\tilde{X}_s$ follows from the similar property of X. $\qquad \square$

While it is more streamlined to state Proposition 7.7 and its proof directly in terms of the symbol $c_{\tilde{X}_s, B}$, in order to prove energy estimates it is more efficient to peel off balanced components of $c_{\tilde{X}_s, B}$, so that we are left with less debris to contend with.

To start with, let us assume that $\tilde{X}_s \in \mathfrak{P} S^1$ admits the representation (7.63) with $a^\gamma \in \mathfrak{P}$ but without requiring that a^γ satisfy the relation (7.64). For such X_s, we peel off balanced components of $c_{\tilde{X}_s, B}$ following the two steps in Lemma 6.9. These steps are briefly reviewed in the sequel.

In a first stage, we note that all expressions in $c_{\tilde{X}_s, B}$ can be seen as linear combinations on the form $\mathfrak{P} \cdot \partial \mathfrak{P}$, where the output is balanced unless the second factor has higher frequency, see Lemma 5.7(a). This allows us to replace such products by paraproducts of the form

$T_{\mathfrak{P}}\partial\mathfrak{P}$. Further, using the definition of paracontrolled distributions for the second factor, we can discard the error term as balanced and arrive at more precise paraproducts of the form $T_{\mathfrak{P}}\partial^2 u$, namely

(7.82)
$$c_{\tilde{X}_s, B} \overset{bal}{\approx} T_{a^{\alpha\beta}}\partial_\alpha\partial_\beta u,$$

where the coefficients $a^{\alpha\beta}$ and $\tilde{a}^{\alpha\beta}$ are explicitly computable as algebraic expressions in terms of u, $\tilde{X}_s$ and the paracoefficients a^γ of $\tilde{X}_s$. Precisely,

$$c_{\tilde{X}_s, B}(x, \xi) \overset{bal}{\approx} T_{\partial_{\xi_\gamma}\tilde{p}(x,\xi)}\partial_\gamma\tilde{X}_s - T_{\partial_{\xi_\gamma}\tilde{X}_s}\partial_\gamma\tilde{p}(x,\xi) + 2T_{\tilde{X}_s}(\tilde{A}^\gamma + 2s\tilde{b}_0^\gamma(x,\xi))\xi_\gamma + T_p\partial_0\tilde{X}_{s0}$$

$$\overset{bal}{\approx} 2T_{\xi_\alpha\tilde{g}^{\alpha\beta}a^\gamma}\partial_\beta\partial_\gamma u + 2T_{\xi_\alpha\xi_\beta\partial_{\xi_\gamma}\tilde{X}_s\partial^\beta u\tilde{g}^{\alpha\delta}}\partial_\delta\partial_\gamma u + 2T_{\tilde{X}_s(\partial^\beta u\xi_\gamma\tilde{g}^{\gamma\alpha} + \partial^0 u\tilde{g}^{0\beta}\xi_\gamma\tilde{g}^{\gamma\alpha})}\partial_\alpha\partial_\beta u$$

$$+ 2sT_{\tilde{X}_s\partial^\beta u\xi_\gamma g^{\gamma\alpha}\xi_\beta\partial_{\xi_\delta}}(\log|\xi'|^2)\partial_\alpha\partial_\delta u - 2sT_{\tilde{X}_s\partial^0 u\tilde{g}^{0\nu}\xi_\alpha\tilde{g}^{\alpha\beta}\xi_\beta\partial_{\xi_\delta}}(\log|\xi'|^2)\partial_\nu\partial_\delta u$$

$$+ T_{\tilde{p}a_0^\gamma}\partial_0\partial_\gamma u$$

so we obtain, in unsymmetrized form, the relation (7.82) with

(7.83)
$$a^{\alpha\gamma} = 2\xi_\beta\tilde{g}^{\alpha\beta}(a^\gamma + \xi_\delta\partial^\delta u\partial_{\xi_\gamma}\tilde{X}_s + \tilde{X}_s\partial^\gamma u + \tilde{X}_s\partial^0 u\tilde{g}^{0\beta} + \frac{s}{2}\tilde{X}_s\xi_\delta\partial^\delta u\partial_{\xi_\gamma}(\log|\xi'|^2))$$
$$- \tilde{p}(2s\tilde{X}_s\partial^0 u\tilde{g}^{0\alpha}\partial_{\xi_\gamma}(\log|\xi'|^2) - \delta_0^\alpha a_0^\gamma).$$

Finally, the last difficulty we face is that we do not have good enough estimates for $\partial_t^2 u$. This is rectified by using instead the corrected expression $\hat{\partial}_t^2 u$ introduced in (5.13). This yields a corresponding correction of c, namely

(7.84)
$$\mathring{c}_{\tilde{X}_s, B} = T_{a^{\alpha\beta}}\widehat{\partial_\alpha\partial_\beta}u.$$

With these notations, we can now state a more refined version of Proposition 7.7:

Proposition 7.9. *Let $\tilde{X}_s$ be the symbol constructed in Proposition 7.7. Then the conclusion of Proposition 7.7 holds equally for $\mathring{c}_{\tilde{X}_s, B}$, with the corresponding expressions $\mathring{q}_2$ and $\mathring{q}_0$ satisfying a stronger version of (7.61),*

(7.85)
$$\|\mathring{q}_2\|_{L^\infty S^2} \lesssim \mathcal{B}^2, \qquad \|P_{<k}\partial_0\mathring{q}_2\|_{L^\infty S^2} \lesssim 2^k\mathcal{B}^2,$$

and with $\mathring{q}_0$ having $\partial_x\mathfrak{P}$ type regularity,

(7.86)
$$\|\mathring{q}_0\|_{\partial_x\mathfrak{P}S^0} \lesssim \mathcal{A}.$$

Proof. A direct computation using (7.83) and (7.64) shows that the coefficients $a^{\alpha\beta}$ have the form

$$a^{\alpha\beta} = \tilde{p}q^{\alpha\beta}, \qquad q^{\alpha\beta} \in \mathfrak{P}S^0,$$

and thus

$$\mathring{c}_{\tilde{X}_s, B} = T_{\tilde{p}q^{\alpha\beta}}\widehat{\partial_\alpha\partial_\beta}u,$$

We need to express this in the form $T_{\tilde{p}}\partial_x\mathcal{B}$ plus a balanced component. For this we consider two cases:

a) If $(\alpha, \beta) \neq (0, 0)$ then the above component of $\mathring{c}$ has the form

$$T_{\tilde{p}q}\partial_x\partial u = T_{\tilde{p}}\partial_x(T_q\partial u) - T_{\tilde{p}}T_{\partial_x q}\partial u + (T_{\tilde{p}q} - T_{\tilde{p}}T_q)\partial_x\partial u.$$

Here the first term on the right is as needed, so we set

$$\mathring{q}_2 = -T_{\tilde{p}}T_{\partial_x q}\partial u + (T_{\tilde{p}q} - T_{\tilde{p}}T_q)\partial_x\partial u.$$

The first term is balanced by Lemma 5.7(a) and the second is balanced by Lemma 5.9. We still need to estimate $\partial_0 \mathring{q}_2$ as in (7.85), which is immediate using Lemma 5.4, Lemma 5.7(a) and Lemma 5.9.

b) If $(\alpha, \beta) = (0,0)$ then the above component has the form

$$T_{\tilde{p}q}\hat{\partial}_t^2 u = \sum_{(\alpha,\beta)\neq(0,0)} T_{\tilde{p}q}(T_{\tilde{g}^{\alpha\beta}}\partial_\alpha\partial_\beta u + T_{\partial_\alpha\partial_\beta u}\tilde{g}^{\alpha\beta}).$$

The first term on the right is treated exactly as in case (a), by pulling out one spatial derivative, while the second is directly placed in $\mathring{q}_2$ using again Lemma 5.4 and (a minor variation of) Lemma 5.7(a). $\square$

7.3.4. *Paradifferential energy estimates associated to $\tilde{X}_s$.* We now use the symbol $\tilde{X}_s$ given by Proposition 7.7 in order to construct an $H^1 \times L^2$ balanced energy functional for the conjugated problem (7.48). This in turn gives an $H^{s+1} \times H^s$ balanced energy functional for the original linear paradifferential flow (3.25), thus completing the proof of Theorem 7.1.

Broadly speaking, we will be following the analysis in the $s = 0$ case, but with more care since we are replacing the vector field X with the pseudodifferential multiplier $\tilde{X}_s$. In particular, here, instead of paraproducts we we will have to commute paraproducts with paradifferential operators. The difficulty is that we will no longer be able to estimate the commutator contributions in a direct, perturbative fashion; instead, we will need to take into account unbalanced subprincipal commutator terms, and devise an additional zero order correction to $\tilde{X}_s$ in order to deal with them.

We begin by considering the conjugation operator B, for which we provide a favourable decomposition:

Lemma 7.10. *The operator B given by (7.49) admits a decomposition*

$$(7.87) \qquad\qquad B = B_0 + B_1 + B_2,$$

where the three components are as follows:

(i) $B_0 = T_{b_0^\gamma}\partial_\gamma$ is the leading part, with symbol

$$(7.88) \qquad\qquad b_0(x,\xi) = i|\xi|^s\{\tilde{g}^{\alpha\beta}\xi_\alpha\xi_\beta, |\xi|^{-s}\}.$$

(ii) B_1 is unbalanced but with a favourable null structure,

$$(7.89) \qquad\qquad B_1 w = T_{h(\partial u)}T_{\tilde{g}^{\alpha\beta}}L_{lh}(\partial_\alpha\partial u, \partial_\beta|D_x|^{-1}w)$$

with h depending smoothly on ∂u.

(iii) B_2 is balanced,

$$(7.90) \qquad\qquad \|B_2 w\|_{L^2} \lesssim \mathcal{B}^2\|\partial w\|_{L^2}.$$

This result is a direct consequence of Proposition 6.15; we have stated it here separately only for quick reference in this section.

At this point, we can repeat the multiplier computation in the previous section, using as multiplier the operator $T_{\tilde{\mathfrak{M}}_s}$ defined in (7.55). Here $\tilde{X}_s$ will be the symbol constructed in the previous subsection, so it remains to consider the choice of $\tilde{Y}_{s0}$, which will be chosen as

$$(7.91) \qquad\qquad \tilde{Y}_{s0} = -\mathring{q}_0$$

with $\mathring{q}_0$ as in Proposition 7.9.

76

Using the $T_{\tilde{\mathfrak{M}}_s}$ operator as a multiplier, we seek to derive an associated energy identity. Here, at leading order, we would like the energy functional $E_{\tilde{X}_s}$ to be as in (7.59), described by the symbol $e_{\tilde{X}_s}$ defined as in (7.58). On the other hand the energy flux is to be described at leading order by the symbol $\mathring{c}_{\tilde{X}_s,B}$ in (7.84) where we add the contribution of $\tilde{Y}_{s0}$.

To have a modular argument, at first we simply assume that

- $\tilde{X}_s \in \mathfrak{P}S^1$, with the representation (7.63) with $a^\gamma \in \mathfrak{P}S^1$, but without assuming that a^γ are given by (7.64).
- $\tilde{Y}_{s0} \in \partial_x\mathfrak{P}$, but without assuming that $\tilde{Y}_{s0}$ is as in Proposition 7.9.

Given such $\tilde{X}_s$ and $\tilde{Y}_{s0}$, we will describe the leading part of the energy flux using the symbol

$$\tilde{c}_s = \mathring{c}_{\tilde{X}_s,B} + T_p\tilde{Y}_{s0}. \tag{7.92}$$

This is a second degree polynomial in ξ_0, which we expand as

$$\tilde{c}_s(x,\xi) = \tilde{c}_s^0(x,\xi')\xi_0^2 + \tilde{c}_s^1(x,\xi')\xi_0 + \tilde{c}_s^2(x,\xi'). \tag{7.93}$$

To this expansion we associate the bilinear form

$$\tilde{C}_s(w,w) = \int -T_{\tilde{c}_s^0}\partial_t w \cdot \partial_t w + \frac{1}{2}T_{\partial_t\tilde{c}_s^0}w \cdot \partial_t w + T_{i\tilde{c}_s^1}w \cdot \partial_t w + T_{\tilde{c}_s^2}w \cdot w\, dx, \tag{7.94}$$

which, integrated also over time, would yield exactly the quadratic form generated by the symbol $\tilde{c}_s$ in Weyl calculus.

Now we can state our main multiplier energy identity, which is as follows:

Proposition 7.11. *Let $\tilde{X}_s \in \mathfrak{P}S^1$ and $\tilde{Y}_{s0} \in \partial_x\mathfrak{P}$ be as above, and the multiplier $T_{\tilde{\mathfrak{M}}_s}$ be as in (7.55). Then there exists an energy function $E_{\tilde{X}_s,B}$ with the following properties:*

i) Leading order expression:

$$E_{\tilde{X}_s,B}[w] = E_{\tilde{X}_s}[w] + Err(\mathcal{A}^\sharp). \tag{7.95}$$

ii) Energy identity:

$$\frac{d}{dt}E_{\tilde{X}_s,B}[w] = \int T_{\tilde{P}_B}w \cdot T_{\tilde{\mathfrak{M}}_s}w\, dx + \tilde{C}_s(w,w) + Err(\mathcal{B}^2). \tag{7.96}$$

We recall again that here we assume neither that $\tilde{X}_s$ is the "vector field" constructed in the previous subsection nor that $\tilde{X}_s$ is forward time-like. Instead we will add these two assumptions later on when we apply the Proposition, in order to guarantee that $\tilde{C}_s(w,w)$ is controlled by $Err(\mathcal{B}^2)$, respectively that $E_{\tilde{X}_s}$ is positive definite.

Proof. As stated, the result in the Proposition is linear with respect to both $\tilde{X}_s$ and $\tilde{Y}_{s0}$, and also separately in $\tilde{A}$ and $\tilde{b}_0$. This allows us to divide the proof into several cases, which turn out to be easier to manage separately.

I. The contribution of $\tilde{X}_{s1}$ with $\tilde{A} = 0$ and $\tilde{b}_0 = 0$. Our starting point here is the integral

$$I_X^1 = 2\int_0^T\!\!\int_{\mathbb{R}^n} T_{\tilde{P}}w \cdot T_{i\tilde{X}_{s1}}w\, dxdt.$$

The operator $T_{i\tilde{X}_{s1}}$ is purely spatial and antisymmetric, so we can integrate by parts three times in $[0,T] \times \mathbb{R}^n$ to rewrite I_X^1 in the form

$$I_X^1 = 2 \int_0^T \!\!\int_{\mathbb{R}^n} T_{\tilde{g}^{\alpha\beta}} T_{i\partial_\beta \tilde{X}_{s1}} w \cdot \partial_\alpha w \, dxdt + \int_0^T \!\!\int_{\mathbb{R}^n} [T_{\tilde{g}^{\alpha\beta}}, T_{i\tilde{X}_{s1}}] \partial_\alpha w \cdot \partial_\beta w \, dxdt$$
$$+ \left. 2 \int T_{\tilde{g}^{\alpha 0}} T_{i\tilde{X}_{s1}} w \cdot \partial_\alpha w \, dx \right|_0^T.$$

Here the expression on the second line should be thought of as the energy and the expression on the first line represents the energy flux. We remark that if there were no boundaries at times $t = 0, T$ then this would be akin to computing the commutator of $T_{\tilde{P}}$ and $T_{i\tilde{X}_{s1}}$.

The above expression needs some further processing to put it in the desired form. We begin with the energy component, where we need to compound the paraproducts and separate the cases $\alpha = 0$ and $\alpha \neq 0$. This is done using Lemma 6.11, `l:para-pdo-A`

$$\int T_{\tilde{g}^{\alpha 0}} T_{i\tilde{X}_{s1}} w \cdot \partial_\alpha w \, dx = \int T_{\tilde{g}^{00} \tilde{X}_{s1}} w \cdot \partial_0 w \, dx + \int T_{\tilde{g}^{j0} \tilde{X}_{s1}} w \cdot \partial_j w \, dx + O(\mathcal{A}^\sharp) \|w[t]\|_{\mathcal{H}^1}^2,$$

as needed.

We now successively consider the space-time integrals on the first line in I_X^1. In the first integral, the components where the $\tilde{g}^{\alpha\beta}$ frequency is at least comparable to the $\tilde{X}_{s1}$ frequency are balanced, and we can use Lemma 6.12 `l:para-pdo-BDB` to compose the paraproducts as

$$\int_0^T \!\!\int_{\mathbb{R}^n} T_{\tilde{g}^{\alpha\beta}} T_{i\partial_\beta \tilde{X}_{s1}} w \cdot \partial_\alpha w \, dxdt = \int_0^T \!\!\int_{\mathbb{R}^n} T_{iT_{\tilde{g}^{\alpha\beta}} \partial_\beta \tilde{X}_{s1}} w \cdot \partial_\alpha w \, dxdt + Err(\mathcal{B}^2),$$

where the integral on the right can be freely switched to the Weyl calculus if $\alpha \neq 0$, and represents one of the desired components of our energy flux.

For the second space-time integral in I_X^1 we use the commutator expansion in Proposition 6.15 `p:com-pdo` to get a principal part, an unbalanced subprincipal part and a balanced term,

$$\int_0^T \!\!\int_{\mathbb{R}^n} [T_{\tilde{g}^{\alpha\beta}}, T_{i\tilde{X}_{s1}}] \partial_\alpha w \cdot \partial_\beta w \, dxdt = \int_0^T \!\!\int_{\mathbb{R}^n} T_{\{\tilde{g}^{\alpha\beta}, \tilde{X}_{s1}\}_p} \partial_\alpha w \cdot \partial_\beta w \, dxdt + I_{X,sub}^1 + Err(\mathcal{B}^2),$$

where the unbalanced subprincipal part $I_{X,sub}^1$ has the form

`IX-sub` (7.97)
$$I_{X,sub}^1 = \int_0^T \!\!\int_{\mathbb{R}^n} T_{\mathfrak{P} S^{-1}} T_{\tilde{g}^{\alpha\gamma}} L_{lh}(\partial_\gamma \partial_x^2 u, \ \partial_\alpha w) \cdot \partial_\beta w \, dxdt.$$

We postpone the analysis of $I_{X,sub}^1$ for later, and focus now on the principal part, which has symbol

$$\partial_{\xi_j} \tilde{X}_{s1} \partial_j \tilde{g}^{\alpha\beta}.$$

As in Lemma 6.7, `l:ppxDPP` we may perturbatively (with $O(\mathcal{B}^2 L^\infty S^0)$ errors) replace this by

$$h^{\alpha\beta} = T_{\partial_{\xi_j} \tilde{X}_{s1}} \partial_j \tilde{g}^{\alpha\beta}.$$

This is almost in the desired form, except that we need to switch it to Weyl calculus. We observe that we have no contribution if both α and β are zero. We separate the remaining

cases, where switching to the Weyl calculus yields errors as follows,

$$Err = \int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\partial_j h^{j0}} w \cdot \partial_0 w \, dxdt + \frac{1}{2}\int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\partial_j h^{jm}} w \cdot \partial_m w \, dxdt$$

$$= \frac{1}{4}\int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\partial_\alpha \partial_\beta h^{\alpha\beta}} w \cdot w \, dxdt + \int T_{\partial_j h^{j0}} w \cdot w \, dx \Big|_0^T.$$

The last integral is an acceptable energy correction. For the first integral to be an acceptable energy flux error, it suffices to show that

$$\|P_{<k}\partial_\alpha \partial_\beta h^{\alpha\beta}\|_{L^\infty S^0} \lesssim 2^{2k}\mathcal{B}^2.$$

It is easily seen that this is indeed the case if any of the derivatives apply to $\tilde{X}_{s1}$, by using the time derivative component of the $\mathfrak{C}$ bound for either $\tilde{X}_{s1}$ or $\tilde{g}^{\alpha\beta}$ to bound time derivatives (of which we can have at most one). So we are left with showing that

$$\|P_{<k}\partial_\alpha \partial_\beta \tilde{g}^{\alpha\beta}\|_{L^\infty S^0} \lesssim 2^{k}\mathcal{B}^2.$$

But for this we use Lemma 6.10. `l:ddiv-g`

II. The contribution of $\tilde{X}_{s0}$ with $A = 0$. Here we will follow the same road map as in the case of $\tilde{X}_{s1}$, but additional care will be needed in order to handle the additional time derivatives. The integral we need to consider is

$$I_X^0 = 2\int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\tilde{P}} w \cdot T_{\tilde{X}_{s0}} \partial_0 w \, dxdt.$$

We can integrate by parts once in $[0,T]\times\mathbb{R}^n$ to rewrite I_X^0 in the form

$$I_X^0 = -2\int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\tilde{g}^{\alpha\beta}} T_{\partial_\beta \tilde{X}_{s0}} \partial_0 w \cdot \partial_\alpha w \, dxdt - 2\int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\tilde{g}^{\alpha\beta}} T_{\tilde{X}_{s0}} \partial_0 \partial_\beta w \cdot \partial_\alpha w \, dxdt$$

$$+ 2\int T_{\tilde{g}^{\alpha 0}} T_{\tilde{X}_{s0}} \partial_0 w \cdot \partial_\alpha w \, dx \Big|_0^T.$$

In the middle term we switch the operator $T_{\tilde{g}^{\alpha\beta}} T_{\tilde{X}_{s0}} \partial_0$ to the right, while integrating by parts once in time, in order to put it in the more symmetric form

$$I_X^0 = -2\int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\tilde{g}^{\alpha\beta}} T_{\partial_\beta \tilde{X}_{s0}} \partial_0 w \cdot \partial_\alpha w \, dxdt + \int_0^T\!\!\!\int_{\mathbb{R}^n} (\partial_0 T_{\tilde{X}_{s0}} T_{\tilde{g}^{\alpha\beta}} - T_{\tilde{g}^{\alpha\beta}} T_{\tilde{X}_{s0}} \partial_0)\partial_\beta w \cdot \partial_\alpha w \, dxdt$$

$$+ \int 2T_{\tilde{g}^{\alpha 0}} T_{\tilde{X}_{s0}} \partial_0 w \cdot \partial_\alpha w - T_{\tilde{X}_{s0}} T_{\tilde{g}^{\alpha\beta}} \partial_\alpha w \cdot \partial_\beta w \, dx \Big|_0^T.$$

For the energy term there is nothing new, we use as before paraproduct rules to rewrite it as the desired leading part plus an acceptable error. We now consider the second space-time integral, where more care is needed. The operator

$$C = \partial_0 T_{\tilde{X}_{s0}} T_{\tilde{g}^{\alpha\beta}} - T_{\tilde{g}^{\alpha\beta}} T_{\tilde{X}_{s0}} \partial_0$$

has a commutator structure, which is good. However we have to carefully decide on the order in which we commute, because, depending on whether $\alpha = 0$ or $\beta = 0$, we might carelessly end up with a double time derivative. The positive feature, arising from the fact that we work with the metric $\tilde{g}$ rather than g, is that if $(\alpha, \beta) = (0, 0)$ then there is a single

79

commutator which does not involve time derivatives. For clarity we consider the four cases separately:

i) The case $\alpha \neq 0$, $\beta \neq 0$. This is the simplest case, where, commuting and peeling off operators of size $O_{L^2}(\mathcal{B}^2)$, we write

$$C = T_{\partial_0 \tilde{X}_{s0}} T_{\tilde{g}^{\alpha\beta}} + T_{\tilde{X}_{s0}} T_{\partial_0 \tilde{g}^{\alpha\beta}} - [T_{\tilde{g}^{\alpha\beta}}, T_{\tilde{X}_{s0}}] \partial_0.$$

ii) The case $\alpha = 0$, $\beta \neq 0$. Here we use the same order as before.

iii) The case $\alpha \neq 0$, $\beta = 0$. Here we reverse the order, to write

$$C = T_{\tilde{g}^{\alpha\beta}} T_{\partial_0 \tilde{X}_{s0}} + T_{\partial_0 \tilde{g}^{\alpha\beta}} T_{\tilde{X}_{s0}} - \partial_0 [T_{\tilde{g}^{\alpha\beta}}, T_{\tilde{X}_{s0}}],$$

where the middle term is integrated by parts once more to move ∂_0 together with ∂_α, at the expense of another negligible energy correction

$$-\int [T_{\tilde{g}^{\alpha 0}}, T_{\tilde{X}_{s0}}] \partial_0 w \cdot \partial_\alpha w \, dx \Big|_0^T.$$

iv) The case $\alpha = 0$, $\beta = 0$. Here we simply have

$$C = T_{\partial_0 \tilde{X}_{s0}}.$$

Now we put together the terms in the four cases.

a) In the $\partial_0 \tilde{X}_{s0}$ term the multiplication order does not matter, and we can further replace it by $T_{T_{\tilde{g}^{\alpha\beta}} \partial_0 \tilde{X}_{s0}}$ modulo $O_{L^2}(\mathcal{B}^2)$ errors. Thus we retain the integral

$$\int_0^T \int_{\mathbb{R}^n} T_{T_{\tilde{g}^{\alpha\beta}} \partial_0 \tilde{X}_{s0}} \partial_\beta w \cdot \partial_\alpha w \, dx dt.$$

b) In the $\partial_0 \tilde{g}^{\alpha\beta}$ term, however, the commutator is not negligible, so in addition to $T_{\tilde{X}_{s0}} T_{\partial_0 \tilde{g}^{\alpha\beta}}$ we also need the commutator $[T_{\partial_0 \tilde{g}^{\alpha 0}}, T_{\tilde{X}_{s0}}]$. Hence we get two contributions,

$$\int_0^T \int_{\mathbb{R}^n} T_{\tilde{X}_{s0}} T_{\partial_0 \tilde{g}^{\alpha\beta}} \partial_\beta u \cdot \partial_\alpha u \, dx dt + \int_0^T \int_{\mathbb{R}^n} [T_{\partial_0 \tilde{g}^{\alpha 0}}, T_{\tilde{X}_{s0}}] \partial_0 u \cdot \partial_\alpha u \, dx dt.$$

c) In the $[T_{\tilde{g}^{\alpha\beta}}, T_{\tilde{X}_{s0}}]$ term where, distinguishing between $\beta = j \neq 0$ and $\beta = 0$, we get

$$\int_0^T \int_{\mathbb{R}^n} -[T_{\tilde{g}^{\alpha j}}, T_{\tilde{X}_{s0}}] \partial_0 \partial_j w \cdot \partial_\alpha w \, dx dt - \int_0^T \int_{\mathbb{R}^n} \partial_0 w \cdot [T_{\tilde{g}^{\alpha 0}}, T_{\tilde{X}_{s0}}] \partial_0 \partial_\alpha w \, dx dt.$$

In the first integral we move ∂_j and the commutator term to the right, also commuting them, so the above expression is rewritten as

$$\int_0^T \int_{\mathbb{R}^n} \partial_0 w \cdot [T_{\tilde{X}_{s0}}, T_{\tilde{g}^{\alpha\beta}}] \partial_\beta \partial_\alpha w \, dx dt + \int_0^T \int_{\mathbb{R}^n} [T_{\partial_j \tilde{g}^{\alpha j}}, T_{\tilde{X}_{s0}}] \partial_0 w \cdot \partial_\alpha w + [T_{\tilde{g}^{\alpha j}}, \partial_j T_{\tilde{X}_{s0}}] \partial_0 w \cdot \partial_\alpha w \, dx dt.$$

We retain the first term as it is, combine the second one with the second term in part (b) and discard the last one as perturbative, $Err(\mathcal{B}^2)$.

Putting all terms together, we have rewritten I_X^0, modulo perturbative terms, as

$$I_X^0 = -2\int_0^T\!\!\int_{\mathbb{R}^n} T_{T_{\tilde{g}^{\alpha\beta}\partial_\beta\tilde{X}_{s0}}}\partial_0 w\cdot\partial_\alpha w\,dxdt + \int_0^T\!\!\int_{\mathbb{R}^n} T_{T_{\tilde{g}^{\alpha\beta}\partial_0\tilde{X}_{s0}}}\partial_\beta w\cdot\partial_\alpha w\,dxdt$$

$$+\int_0^T\!\!\int_{\mathbb{R}^n} T_{\tilde{X}_{s0}}T_{\partial_0\tilde{g}^{\alpha\beta}}\partial_\beta u\cdot\partial_\alpha u\,dxdt + \int_0^T\!\!\int_{\mathbb{R}^n}\partial_0 w\cdot\partial_\beta[T_{\tilde{X}_{s0}},T_{\tilde{g}^{\alpha\beta}}]\partial_\alpha w\,dxdt$$

$$-\int_0^T\!\!\int_{\mathbb{R}^n}\partial_0 w\cdot[\partial_\beta T_{\tilde{X}_{s0}},T_{\tilde{g}^{\alpha\beta}}]\partial_\alpha w\,dxdt + Err(\mathcal{B}^2)$$

$$:= I_X^{01} + I_X^{02} + I_X^{03} + I_X^{04} + I_X^{05} + Err(\mathcal{B}^2).$$

This can be simplified further by observing that, in view of Lemma 2.3, the term I_X^{05} is also perturbative. Thus we arrive at

$$I_X^0 = I_X^{01} + I_X^{02} + I_X^{03} + I_X^{04} + Err(\mathcal{B}^2).$$

We successively consider these terms:

IIa. The contribution of I_X^{01}. This corresponds to the symbol

$$2T_{T_{\tilde{g}^{\alpha\beta}\partial_\beta\tilde{X}_{s0}}}\xi_0\xi_\alpha,$$

which is akin to one of the components of $c_{X_s,B}$ in (7.57). We can turn this into the corresponding component of $\mathring{c}_{\tilde{X}_s,B}$. Precisely, given $\tilde{X}_{s0}$ as in the representation (7.63), that component is

$$2T_{T_{\tilde{g}^{\alpha\beta}a_0^\gamma\widehat{\partial_\beta\partial_\gamma}u}}\xi_0\xi_\alpha,$$

where we recall that the hat above is understood as nonexistent unless $\beta = \gamma = 0$, in which case it is interpreted as the corrected expression (5.13). The difference between the two coefficients is easily seen to have size $\mathcal{B}^2$, so it is perturbative, as in Lemma 6.9. It remains to switch this modification of I_X^{01} to the Weyl calculus, which requires estimating the integral

$$\int_0^T\!\!\int_{\mathbb{R}^n} T_{\partial_\alpha T_{\tilde{g}^{\alpha\beta}a^\gamma\widehat{\partial_\beta\partial_\gamma}u}}\partial_0 w\cdot w\,dxdt.$$

This follows from the bound

$$\|P_{<k}\partial_\alpha T_{\tilde{g}^{\alpha\beta}a^\gamma}\widehat{\partial_\beta\partial_\gamma}u\|_{L^\infty}\lesssim 2^k\mathcal{B}^2,$$

which in turn follows from Lemma 5.12(b) after commuting a_0^γ out.

IIb. The contribution of I_X^{02}. Exactly as above, this integral also corresponds to a term in $c_{X_s,B}$. Again, after a perturbative $Err(\mathcal{B}^2)$ correction we can turn this into the corresponding term in $\mathring{c}_{\tilde{X}_s,B}$, which has the para-coefficient

$$T_{\tilde{g}^{\alpha\beta}a^\gamma}\widehat{\partial_0\partial_\gamma}u.$$

The associated integral is

$$\int_0^T\!\!\int_{\mathbb{R}^n} T_{T_{\tilde{g}^{\alpha\beta}a^\gamma}\widehat{\partial_0\partial_\gamma}u}\partial_\beta w\cdot\partial_\alpha w\,dxdt.$$

We would like to switch this to Weyl calculus, but we need to be careful here because the convention for the Weyl form differs depending on whether β is zero or not.

If $\beta \neq 0$ then the error corresponds to switching the operator on the left to Weyl calculus, and has the form

(7.98)
$$\frac{1}{2}\int_0^T\!\!\int_{\mathbb{R}^n} T_{\partial_\beta T_{\tilde{g}^{\alpha\beta}a^\gamma}\widehat{\partial_0\partial_\gamma}u}w\cdot\partial_\alpha w\,dxdt.$$

The same applies if $\alpha = \beta = 0$. But if $\alpha \neq 0$ and $\beta = 0$ then we have to switch the paraproduct to the right, and then the Weyl correction is

$$\frac{1}{2}\int_0^T\!\!\int_{\mathbb{R}^n} \partial_\beta w\cdot T_{\partial_\alpha T_{\tilde{g}^{\alpha\beta}a^\gamma}\widehat{\partial_0\partial_\gamma}u}w\,dxdt.$$

We can rectify this discrepancy and switch this correction to the form in (7.98) by integrating twice by parts, first in x_β and then in x_α. Since we are in the case $\beta = 0$, the first step yields an energy correction, namely

$$\frac{1}{4}\int w\cdot T_{\partial_\alpha T_{\tilde{g}^{\alpha 0}a^\gamma}\widehat{\partial_0\partial_\gamma}u}w\,dx\Big|_0^T.$$

As $\alpha \neq 0$, for this to be an acceptable $Err(\mathcal{A}^\sharp)$ error we need the bound

$$\|P_{<k}\widehat{\partial_0\partial_\gamma}u\|_{L^\infty}\lesssim 2^k\mathcal{A}^\sharp.$$

This is obvious if $\gamma \neq 0$, and follows from (5.15) otherwise.

Thus we are left with considering the correction in (7.98) summed over all α and β, and which we would like to estimate perturbatively.

Here there is no structure in the γ summation, so we can fix γ. The easier case is when $\gamma \neq 0$. Then we can commute ∂_γ out, as well as a^γ, and ∂_β in, writing

$$\partial_\beta T_{\tilde{g}^{\alpha\beta}a^\gamma}\partial_0\partial_\gamma u = T_{a^\gamma}\partial_\gamma T_{\tilde{g}^{\alpha\beta}}\partial_\beta\partial_0 u + f,$$

where the error term f satisfies

(7.99)
$$\|P_{<k}f\|_{L^\infty}\lesssim 2^k\mathcal{B}^2.$$

We may also correct the second order time derivative, arriving at

$$\partial_\beta T_{\tilde{g}^{\alpha\beta}a^\gamma}\partial_0\partial_\gamma u = T_{a^\gamma}\partial_\gamma T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_0}u + f.$$

The remaining term is no longer directly perturbative, but its contribution may be instead estimated integrating by parts,

$$\int_0^T\!\!\int_{\mathbb{R}^n} T_{T_{a^\gamma}\partial_\gamma T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_0}u}w\cdot\partial_\alpha w\,dxdt = -\frac{1}{2}\int_0^T\!\!\int_{\mathbb{R}^n} T_{\partial_\alpha T_{a^\gamma}\partial_\gamma T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_0}u}w\cdot w\,dxdt$$
$$+\int T_{T_{a^\gamma}\partial_\gamma T_{\tilde{g}^{0\beta}}\widehat{\partial_\beta\partial_0}u}w\cdot w\,dx\Big|_0^T.$$

The last term is a bounded energy correction, as

$$\|P_k T_{a^\gamma}\partial_\gamma T_{\tilde{g}^{0\beta}}\widehat{\partial_\beta\partial_0}u\|_{L^\infty}\lesssim 2^{2k}\mathcal{A}^\sharp.$$

It remains to show that the first term is also perturbative,

$$\|P_{<k}T_{a^\gamma}\partial_\gamma T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_0}u\|_{L^\infty}\lesssim 2^{2k}\mathcal{B}^2.$$

Commuting ∂_α inside and discarding $a^\gamma\partial_\gamma$, this reduces to

$$\|P_{<k}\partial_\alpha T_{\tilde{g}^{\alpha\beta}}\widehat{\partial_\beta\partial_0}u\|_{L^\infty}\lesssim 2^k\mathcal{B}^2,$$

82

which is again a consequence of Lemma 5.12(b).

It remains to consider the case $\gamma = 0$, where we take advantage of the hat correction. Precisely, using the u equation, we write

$$\widehat{\partial_0 \partial_0} u = - \sum_{(\mu,\nu) \neq (0,0)} T_{\tilde{g}^{\mu\nu}} \partial_\mu \partial_\nu u + T_{\partial_\mu \partial_\nu u} \tilde{g}^{\mu\nu}.$$

We substitute this into the paracoefficient in (7.98), peeling off perturbative contributions. Fixing μ and ν we may assume $\mu \neq 0$ and arrive at

$$\partial_\beta T_{\tilde{g}^{\alpha\beta}{}_a{}^\gamma} \widehat{\partial_0 \partial_0} u = - \sum_{\mu \neq 0} T_{a^\gamma \tilde{g}^{\mu\nu}} \partial_\mu T_{\tilde{g}^{\alpha\beta}} \partial_\beta \partial_\nu u + f,$$

with f as in (7.99). At this point we can repeat the argument in the case $\gamma \neq 0$.

IIc. The contribution of I_X^{03}. We recall that this is

$$I_X^{03} = \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\tilde{X}_{s0}} T_{\partial_0 \tilde{g}^{\alpha\beta}} \partial_\beta u \cdot \partial_\alpha u \, dx dt.$$

This term is easily seen to be perturbative unless the spatial frequency of $\tilde{X}_{s0}$ is smaller than that of $\partial_0 \tilde{g}^{\alpha\beta}$, see Lemma 6.12. Thus we can think of the principal symbol of the product $T_{\tilde{X}_{s0}} T_{\partial_0 \tilde{g}^{\alpha\beta}}$ as being $T_{\tilde{X}_{s0}} \partial_0 \tilde{g}^{\alpha\beta}$. However some care is needed here with the error, which is lower order but not necessarily balanced. Precisely, using Proposition 6.16, we can expand this product into a leading part, an unbalanced subprincipal part and a perturbative term,

$$\text{Xdg-exp} \quad (7.100) \qquad T_{\tilde{X}_{s0}} T_{\partial_0 \tilde{g}^{\alpha\beta}} = T_{T_{\tilde{X}_{s0}} \partial_0 \tilde{g}^{\alpha\beta}} + OP\mathfrak{P}S^{-2} L_{lh}(\partial_x^2 \partial_0 \tilde{g}^{\alpha\beta}, \cdot) + O_{L^2}(\mathcal{B}^2).$$

This yields a corresponding decomposition of I_X^{03} into

$$I_X^{03} = I_{X,main}^{03} + I_{X,sub}^{03} + Err(\mathcal{B}^2).$$

To better describe the first two terms we take a closer look at the coefficient $\partial_0 \tilde{g}^{\alpha\beta}$, for which we compute

$$\text{dtg} \quad (7.101) \qquad \partial_0 \tilde{g}^{\alpha\beta} = - \left(\partial^\beta u \tilde{g}^{\alpha\delta} \partial_\delta \partial_0 u + \partial^\alpha u \tilde{g}^{\beta\delta} \partial_\delta \partial_0 u \right) + 2 \tilde{g}^{\alpha\beta} \partial^0 u \tilde{g}^{0\delta} \partial_\delta \partial_0 u.$$

Here we have a double time derivative $\partial_t^2 u$ when $\delta = 0$, which we replace as before by $\hat{\partial}_t^2 u$ with perturbative errors. Once this is done, we may also replace all products by paraproducts, arriving at the modified expression

$$\mathring{\partial}_0 \tilde{g}^{\alpha\beta} := - \left(T_{\partial^\beta u \tilde{g}^{\alpha\delta}} \widehat{\partial_\delta \partial_0} u + T_{\partial^\alpha u \tilde{g}^{\beta\delta}} \widehat{\partial_\delta \partial_0} u \right) + 2 T_{\partial^0 u \tilde{g}^{0\delta} \tilde{g}^{\alpha\beta}} \widehat{\partial_\delta \partial_0} u$$

so that the difference is perturbative in the sense that

$$\partial_0 \tilde{g}^{\alpha\beta} = \mathring{\partial}_0 \tilde{g}^{\alpha\beta} + O(\mathcal{B}^2).$$

Finally we return to the operator setting, where we make the above substitution. In the principal part can compound the outer paracoefficients at the expense of more negligible errors, writing it in a paradifferential form

$$T_{T_{\tilde{X}_{s0}} \partial_0 \tilde{g}^{\alpha\beta}} = T_{q^{\alpha\beta}} + O_{L^2}(\mathcal{B}^2), \qquad I_{X,main}^{03} = \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{q^{\alpha\beta}} \partial_\alpha w \cdot \partial_\beta w \, dx dt + Err(\mathcal{B}^2),$$

83

where the order zero symbols $q^{\alpha\beta}$ are given by

$$q^{\alpha\beta} = -\left(T_{\tilde{X}_{s0}\partial^\beta u\tilde{g}^{\alpha\delta}}\widehat{\partial_\delta\partial_0}u + T_{\tilde{X}_{s0}\partial^\alpha u\tilde{g}^{\beta\delta}}\widehat{\partial_\delta\partial_0}u\right) + 2T_{\tilde{X}_{s0}\partial^0 u\tilde{g}^{0\delta}\tilde{g}^{\alpha\beta}}\widehat{\partial_\delta\partial_0}u.$$

Here the symbol $q^{\alpha\beta}\xi_\alpha\xi_\beta$ is a component of $\mathring{c}_{\tilde{X}_s,B}$, as desired. All we need now is to convert the last expression for $I^{03}_{X,main}$ to Weyl form. This conversion yields the additional error

$$\frac{1}{2}\int_0^T\!\!\int_{\mathbb{R}^n} T_{\partial_\alpha q^{\alpha\beta}}w \cdot \partial_\beta w\, dxdt,$$

which we need to estimate. Here we separate the three terms in $q^{\alpha\beta}$. For the first term, after one commutation it remains to show that

$$\|P_{<k}\partial_\alpha T_{\tilde{g}^{\alpha\delta}}\widehat{\partial_\delta\partial_0}u\|_{L^\infty} \lesssim 2^k\mathcal{B}^2,$$

which we get from Lemma 5.12. The second term is similar if we integrate by parts to switch α and β, at the expense of a bounded energy correction. Finally, the third term is exactly as in the case of I^{02}_X.

Similarly, in the subprincipal term in (7.100) we may peel off perturbative errors to write it as a linear combination of expressions of the form

$$T_{\mathfrak{P}S^{-1}\tilde{T}_{T_{\tilde{g}^{\alpha\delta}\partial_x}\widehat{\partial_\delta\partial_0}u}}} + T_{\mathfrak{P}S^{-1}\tilde{T}_{T_{\tilde{g}^{\beta\delta}\partial_x}\widehat{\partial_\delta\partial_0}u}}} + T_{\mathfrak{P}S^{-1}\tilde{T}_{T_{\tilde{g}^{\alpha\beta}\partial_x}\widehat{\partial_\delta\partial_0}u}}}.$$

We postpone their analysis for later, for now we simply list the two types of contributions:

(7.102) $$I^{031}_{X,sub} = \int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{\alpha\delta}}L_{lh}(\partial_x\widehat{\partial_\delta\partial_0}u, \partial_\alpha w) \cdot \partial w\, dxdt.$$

(7.103) $$I^{032}_{X,sub} = \int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{\alpha\beta}}L_{lh}(\partial_x\widehat{\partial_\delta\partial_0}u, \partial_\alpha w) \cdot \partial_\beta w\, dxdt.$$

.

IId. The contribution of I^{04}_X. We recall that this is

$$I^{04}_X = \int_0^T\!\!\int_{\mathbb{R}^n} \partial_0 w \cdot \partial_\alpha[T_{\tilde{X}_{s0}}, T_{\tilde{g}^{\alpha\beta}}]\partial_\beta w\, dxdt.$$

This has a similar treatment to I^{03}_X. For the commutator above we must have again a smaller frequency on $\tilde{X}_{s0}$, else this yields a perturbative contribution. Using Proposition 6.15 we expand the commutator into a leading part, an unbalanced subprincipal part and a perturbative term,

(7.104) $$[T_{\tilde{X}_{s0}}, T_{\tilde{g}^{\alpha\beta}}] = T_{T_{\partial_{\xi_j}\tilde{X}_{s0}}\partial_j\tilde{g}^{\alpha\beta}} + OP\mathfrak{P}S^{-2}L_{lh}(\partial^2_x\tilde{g}^{\alpha\beta}, \cdot) + R^{\alpha\beta},$$

where the remainder R satisfies perturbative bounds of the form

$$\|R\|_{L^2\to H^1} \lesssim \mathcal{B}^2, \qquad \|R\|_{H^{-1}\to L^2} \lesssim \mathcal{B}^2, \qquad \|\partial_0 R\|_{L^2\to L^2} \lesssim \mathcal{B}^2.$$

We first consider the contribution of the leading part $I^{04}_{X,main}$. For $\partial_j\tilde{g}^{\alpha\beta}$ we use the expansion in (7.101) with the subscript 0 replaced by $j \neq 0$, arriving at

$$T_{T_{\partial_{\xi_j}\tilde{X}_{s0}}\partial_j\tilde{g}^{\alpha\beta}} = T_{q^{\alpha\beta}} + R^{\alpha\beta},$$

84

where the order zero symbols $q^{\alpha\beta}$ are given by

$$q^{\alpha\beta} = -\left(T_{\partial_{\xi_j}\tilde{X}_{s0}\partial^\beta u\tilde{g}^{\alpha\delta}}\partial_\delta\partial_j u + T_{\partial_{\xi_j}\tilde{X}_{s0}\partial^\alpha u\tilde{g}^{\beta\delta}}\partial_\delta\partial_j u\right) + 2T_{\partial_{\xi_j}\tilde{X}_{s0}\partial^0 u\tilde{g}^{0\delta}\tilde{g}^{\alpha\beta}}\partial_\delta\partial_j u,$$

and the remainder R is as above. Then the leading part can be written as

$$I_{X,main}^{04} = \int_0^T\!\!\!\int_{\mathbb{R}^n} T_{q^{\alpha\beta}}\partial_\alpha\partial_\beta w \cdot \partial_0 w \, dxdt + Err(\mathcal{B}^2).$$

Now the symbol $q^{\alpha\beta}\xi_0\xi_\alpha\xi_\beta$ is a component of $\mathring{c}_{\tilde{X}_s,B}$, as desired. It remains to convert the last expression for $I_{X,main}^{04}$ to Weyl form. The error in doing that is

$$\frac{1}{4}\int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\partial_\alpha\partial_\beta q^{\alpha\beta}} w \cdot \partial_0 w \, dxdt.$$

Estimating this expression requires the bound

$$\|P_{<k}\partial_\alpha\partial_\beta q^{\alpha\beta}\|_{L^\infty} \lesssim 2^k \mathcal{B}^2.$$

Here $q^{00} = 0$ so we avoid the case of two time derivatives. This allows us to commute $\partial_\alpha\partial_\beta$ inside and take ∂_j outside modulo perturbative terms. Then ∂_j yields the 2^k factor, and we have reduced the problem to proving that

$$\left\|P_k\left(T_{\partial_{\xi_j}\tilde{X}_{s0}\partial^\beta u\tilde{g}^{\alpha\delta}} + T_{\partial_{\xi_j}\tilde{X}_{s0}\partial^\alpha u\tilde{g}^{\beta\delta}} - 2T_{\partial_{\xi_j}\tilde{X}_{s0}\partial^0 u\tilde{g}^{0\delta}\tilde{g}^{\alpha\beta}}\right)\partial_\delta\partial_\alpha\partial_\beta u\right\|_{L^\infty} \lesssim \mathcal{B}^2.$$

Re-labeling this becomes

$$\|P_k T_{\partial_{\xi_j}\tilde{X}_{s0}(\partial^\delta u - \partial^0 u\tilde{g}^{0\delta})\tilde{g}^{\alpha\beta}}\partial_\delta\partial_\alpha\partial_\beta u\|_{L^\infty} \lesssim \mathcal{B}^2.$$

The expression on the left vanishes if $\delta = 0$. This allows us to break the para-coefficient in two using Lemma 2.7 and replace this by

$$\|P_k T_{\partial_{\xi_j}\tilde{X}_{s0}(\partial^\delta u - \partial^0 u\tilde{g}^{0\delta})}T_{\tilde{g}^{\alpha\beta}}\partial_\delta\partial_\alpha\partial_\beta u\|_{L^\infty} \lesssim \mathcal{B}^2,$$

which is finally a consequence of Lemma 5.12.

Next we consider the subprincipal term. Here we use again the expansion in (7.101) and recombine paracoefficients to rewrite it as a linear combination of terms of the form

$$\boxed{\text{IXsub-041}} \quad (7.105) \qquad I_{X,sub}^{041} = \int_0^T\!\!\!\int_{\mathbb{R}^n} \partial_0 w \cdot T_h\partial_\alpha T_{\tilde{g}^{\alpha\delta}}L_{lh}(\partial_x^2\partial_\delta u, \partial_\beta w) \, dxdt,$$

$$\boxed{\text{IXsub-042}} \quad (7.106) \qquad I_{X,sub}^{042} = \int_0^T\!\!\!\int_{\mathbb{R}^n} \partial_0 w \cdot T_h\partial_\alpha T_{\tilde{g}^{\beta\delta}}L_{lh}(\partial_x^2\partial_\delta u, \partial_\beta w) \, dxdt,$$

respectively

$$\boxed{\text{IXsub-043}} \quad (7.107) \qquad I_{X,sub}^{043} = \int_0^T\!\!\!\int_{\mathbb{R}^n} \partial_0 w \cdot T_h\partial_\alpha T_{\tilde{g}^{\alpha\beta}}L_{lh}(\partial_x^2\partial_\delta u, \partial_\beta w) \, dxdt,$$

where $h \in \mathfrak{P}S^{-2}$ roughly corresponds to $\partial_\xi^2\tilde{X}_{s0}$. Here we can freely separate variables and reduce to the case when h is a function, including the multiplier part in L_{lh}.

We remark that until now we were able to exclude the case when $\alpha = \beta = 0$. However, at this point we need to separate the three types of contributions in order to take advantage of their structures. Because of this, from here on we have to also allow for the case $\alpha = \beta = 0$,

forfeiting the cancellation that would otherwise occur in this case between the different terms. We postpone the estimate for the subprincipal terms for the end of the proof.

III. The contribution of $\tilde{Y}_{s0}$ with $\tilde{A} = 0$ and $\tilde{b} = 0$. Here we consider the integral

$$I_Y = \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\tilde{P}} w \cdot T_{\tilde{Y}_{s0}} w \, dxdt,$$

where we recall that $\tilde{Y}_{s0} \in \partial_x \mathfrak{P} S^0$. We integrate once by parts to write

$$I_Y = \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\tilde{g}^{\alpha\beta}} \partial_\alpha w \cdot T_{\tilde{Y}_{s0}} \partial_\beta w \, dxdt - \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\tilde{g}^{\alpha\beta}} \partial_\alpha w \cdot T_{\partial_\beta \tilde{Y}_{s0}} w \, dxdt + \int T_{\tilde{g}^{\alpha 0}} \partial_\alpha w \cdot T_{\tilde{Y}_{s0}} w \, dx \Big|_0^T.$$

The last integral is an admissible energy correction. In both space-time integrals we move $T_{\tilde{Y}_{s0}}$ to the left, and combine the two paraproducts as in Lemma 2.7, peeling off perturbative contributions, to get

$$I_Y = \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{T_{\tilde{g}^{\alpha\beta}} \tilde{Y}_{s0}} \partial_\alpha w \cdot \partial_\beta w \, dxdt - \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\partial_\beta T_{\tilde{g}^{\alpha\beta}} \tilde{Y}_{s0}} \partial_\alpha w \cdot w \, dxdt. + Err(\mathcal{B}^2) + Err(\mathcal{A}^\sharp)\big|_0^T.$$

The symbol of the bilinear form in the first integral is the desired component of $\mathring{c}_s$, but we need to convert it to Weyl calculus. This yields an error which is half of the second integral, which in turn needs to be estimated perturbatively. Commuting ∂_β inside, we are left with

$$\int_0^T \!\!\! \int_{\mathbb{R}^n} T_{T_{\tilde{g}^{\alpha\beta}} \partial_\beta \tilde{Y}_{s0}} w \cdot \partial_\alpha w \, dxdt.$$

Here $\tilde{Y}_{s0}$ is of the form $\tilde{Y}_{s0} = \partial_x h$, with $h \in \mathfrak{P} S^0$. We can harmlessly commute ∂_x out, to arrive at

<div style="text-align:right">def-J-Y</div> (7.108)
$$J = \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\partial_x T_{\tilde{g}^{\alpha\beta}} \partial_\beta h} w \cdot \partial_\alpha w \, dxdt.$$

In the absence of boundaries at $t = 0, T$ here we could integrate by parts once more to rewrite this as

$$-\frac{1}{2} \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\partial_x \partial_\alpha T_{\tilde{g}^{\alpha\beta}} \partial_\beta h} w \cdot \partial_\alpha w \, dxdt,$$

and then use Lemma 6.4. The same argument applies if we add in the boundaries, by carefully tracking the boundary contributions. Precisely, we use the lemma to rewrite the expression J in (7.108) as follows:

$$J = \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\partial_x (T_{\tilde{g}^{\alpha\beta}} \partial_\beta h - \delta_0^\alpha f^0)} w \cdot \partial_\alpha w \, dxdt + \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\partial_x f^0} w \cdot \partial_0 w \, dxdt$$

$$= -\frac{1}{2} \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\partial_x (\partial_\alpha T_{\tilde{g}^{\alpha\beta}} \partial_\beta h - \partial_0 f^0)} w \cdot w \, dxdt + \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\partial_x f^0} w \cdot \partial_0 w \, dxdt$$

$$+ \frac{1}{2} \int T_{\partial_x (T_{\tilde{g}^{0\beta}} \partial_\beta h - f^0)} w \cdot w \, dx \Big|_0^T$$

$$= -\frac{1}{2} \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\partial_x \partial_j f^j} w \cdot w \, dxdt + \int_0^T \!\!\! \int_{\mathbb{R}^n} T_{\partial_x f^0} w \cdot \partial_0 w \, dxdt + \frac{1}{2} \int T_{\partial_x (T_{\tilde{g}^{0\beta}} \partial_\beta h - f^0)} w \cdot w \, dx \Big|_0^T.$$

Now, in view of Lemma 6.4, both the energy and the flux terms are perturbative.

<div style="text-align:center">86</div>

IV. The contribution of the gradient potential $\tilde{A}$ and of $\tilde{b}_0$. We discuss the two together, as their contributions are similar. This has the form

$$I_X^2 = \iint T_{\widetilde{\mathfrak{M}}_s} w \cdot (T_{\tilde{A}^\gamma} + T_{\tilde{b}_0^\gamma})\partial_\gamma w \, dxdt,$$

which we need to shift to Weyl calculus after peeling off a perturbative contribution. For instance the contribution of $\tilde{Y}_{s0}$ is directly perturbative. On the other hand, $\tilde{A}^\gamma$ contains $\partial_0^2 u$ terms which need to be corrected, while $\tilde{b}_0^\gamma$ does not. In any case, the correction can be freely added as its contribution has size $Err(\mathcal{B}^2)$. Below we denote by $\mathring{A}$ the corrected version of $\tilde{A}$.

Next we consider the contribution of $\tilde{X}_{s1}$, where we need to shift the operator product $T_{\tilde{X}_{s1}} T_{\tilde{A}^\gamma}$ to the Weyl calculus via Lemma 6.12:

$$T_{\tilde{X}_{s1}} T_{\tilde{A}^\gamma} = T_{T_{\tilde{X}_{s1}} \mathring{A}^\gamma} + T_{\mathfrak{P}S^0} T_{\tilde{g}^{\alpha\gamma}} L_{lh}(\partial_x \widehat{\partial_\alpha \partial} u, \cdot) + O_{H^1 \to L^2}(\mathcal{B}^2),$$

and similarly for b_0, i.e. the desired term plus a null unbalanced lower order term plus a perturbative contribution. We note here that the contribution of the null unbalanced lower order term has the form

$$I_{X,sub}^{31} = \int_0^T \int_{\mathbb{R}^n} T_{\mathfrak{P}S^0} T_{\tilde{g}^{\alpha\gamma}} L_{lh}(\partial_x \widehat{\partial_\alpha \partial} u, \partial_\gamma w) \cdot w \, dxdt.$$

Finally we consider the contribution of $\tilde{X}_{s0}$,

$$I_X^{30} = \int_0^T \int_{\mathbb{R}^n} T_{\tilde{X}_{s0}} \partial_0 w \cdot (T_{\mathring{A}^\gamma} + T_{b_0^\gamma})\partial_\gamma w \, dxdt = \int_0^T \int_{\mathbb{R}^n} \partial_0 w \cdot T_{\tilde{X}_{s0}}(T_{\mathring{A}^\gamma} + T_{b_0^\gamma})\partial_\gamma w \, dxdt.$$

We use again the product formula for paraproducts to write

$$T_{\tilde{X}_{s0}}(T_{\mathring{A}^\gamma} + T_{b_0^\gamma}) = T_{T_{\tilde{X}_{s0}}(\mathring{A}^\gamma + b_0^\gamma)} + T_{\mathfrak{P}S^{-1}} T_{\tilde{g}^{\alpha\gamma}} L_{lh}(\partial_x \widehat{\partial\partial_\alpha} u, \cdot) + O_{L^2}(\mathcal{B}^2),$$

which generates a leading term and a subprincipal term.

The leading term is

$$I_{X,main}^{30} = \int_0^T \int_{\mathbb{R}^n} \partial_0 w \cdot T_{T_{\tilde{X}_{s0}}(\mathring{A}^\gamma + T_{b_0^\gamma})} \partial_\gamma w \, dxdt.$$

Its symbol is as needed, but we still have to switch it to Weyl calculus. This switch introduces an error

$$\frac{1}{2} \int_0^T \int_{\mathbb{R}^n} T_{\partial_\gamma [T_{\tilde{X}_{s0}}(\mathring{A}^\gamma + b_0^\gamma)]} w \cdot \partial_0 w \, dxdt.$$

To bound its contribution, we would like to have the symbol bound

(7.109)
$$|P_{<k} \partial_\gamma T_{\tilde{X}_{s0}}(\mathring{A}^\gamma + b_0^\gamma)| \lesssim 2^k \mathcal{B}^2.$$

Here we use the expressions for $\tilde{A}^\gamma$ and b_0^γ, take out bounded paracoefficients, and we are left with

$$\|P_{<k} \partial_\gamma T_{\tilde{g}^{\gamma\alpha}} \widehat{\partial_\alpha \partial_\beta} u\|_{L^\infty} \lesssim 2^k \mathcal{B}^2.$$

But this is in turn a consequence of Lemma 5.12.

87

To conclude, we record the form of the subprincipal term,

$$\boxed{\text{IXsub-30}} \quad (7.110) \qquad I^{30}_{X,sub} = \int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}} T_{\tilde{g}^{\alpha\gamma}} L_{lh}(\partial_x \widehat{\partial_\alpha \partial} u, \partial_\gamma w) \cdot \partial_0 w \, dxdt.$$

V. The unbalanced lower order terms. These are the expressions identified earlier, which we recall here:

$$\boxed{\text{IX-sub-re}} \quad (7.111) \qquad I^{1}_{X,sub} = \int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}} T_{\tilde{g}^{\alpha\gamma}} L_{lh}(\partial_\gamma \partial_x^2 u, \ \partial_\alpha w) \cdot \partial_\beta w \, dxdt.$$

$$\boxed{\text{sub-031-re}} \quad (7.112) \qquad I^{031}_{X,sub} = \int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}} T_{\tilde{g}^{\alpha\delta}} L_{lh}(\partial_x \widehat{\partial_\delta \partial_0} u, \partial_\alpha w) \cdot \partial w \, dxdt.$$

$$\boxed{\text{sub-032-re}} \quad (7.113) \qquad I^{032}_{X,sub} = \int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}} T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_x \widehat{\partial_\delta \partial_0} u, \partial_\alpha w) \cdot \partial_\beta w \, dxdt.$$

$$\boxed{\text{sub-041-re}} \quad (7.114) \qquad I^{041}_{X,sub} = \int_0^T\!\!\int_{\mathbb{R}^n} \partial_0 w \cdot T_{\mathfrak{P}S^{-2}} \partial_\alpha T_{\tilde{g}^{\alpha\delta}} L_{lh}(\partial_x^2 \partial_\delta u, \partial_\beta w) \, dxdt.$$

$$\boxed{\text{sub-042-re}} \quad (7.115) \qquad I^{042}_{X,sub} = \int_0^T\!\!\int_{\mathbb{R}^n} \partial_0 w \cdot T_{\mathfrak{P}S^{-2}} \partial_\alpha T_{\tilde{g}^{\beta\delta}} L_{lh}(\partial_x^2 \partial_\delta u, \partial_\beta w) \, dxdt.$$

$$\boxed{\text{sub-043-re}} \quad (7.116) \qquad I^{043}_{X,sub} = \int_0^T\!\!\int_{\mathbb{R}^n} \partial_0 w \cdot T_{\mathfrak{P}S^{-2}} \partial_\alpha T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_x^2 \partial_\delta u, \partial_\beta w) \, dxdt.$$

$$\boxed{\text{Xsub-30-re}} \quad (7.117) \qquad I^{30}_{X,sub} = \int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}} T_{\tilde{g}^{\alpha\gamma}} L_{lh}(\partial_x \widehat{\partial_\alpha \partial} u, \partial_\gamma w) \cdot \partial_0 w \, dxdt.$$

All of these exhibit a null structure.

We directly compress four of these into the expression

$$\boxed{\text{IXsub-main}} \quad (7.118) \qquad I_{sub,\gamma\delta} = \int_0^T\!\!\int_{\mathbb{R}^n} T_h T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_x \widehat{\partial_\alpha \partial_\gamma} u, \ \partial_\beta w) \cdot \partial_\delta w \, dxdt, \qquad h \in \mathfrak{P}S^1,$$

where the analysis will be slightly different depending on whether γ and δ are zero or not.

In $I^{042}_{X,sub}$ the case $\alpha \neq 0$ is included above. If instead $\alpha = 0$ then we integrate by parts ∂_α to the left, so that, after a perturbative energy correction, we arrive at

$$\int_0^T\!\!\int_{\mathbb{R}^n} \partial_0^2 w \cdot T_{\mathfrak{P}S^{-2}} T_{\tilde{g}^{\beta\delta}} L_{lh}(\partial_x^2 \partial_\delta u, \partial_\beta w) \, dxdt.$$

Now we use the paradifferential equation $T_{\tilde{P}}$ equation for w, which after more perturbative errors allows us to replace the leading ∂_0^2 operator by $\partial \partial_x$, with a $\mathfrak{P}$ paracoefficient. Then ∂_x combines with $T_{\mathfrak{P}S^{-2}}$ to give $T_{\mathfrak{P}S^{-2}}$, thereby reducing the problem to the case of (7.118).

Finally in $I_{X,sub}^{041}$ we commute inside and distribute the ∂_α derivative, peeling off perturbative errors. We arrive at

$$I_{X,sub}^{041} = \int_0^T\!\!\!\int_{\mathbb{R}^n} \partial_0 w \cdot T_{\mathfrak{P}S^{-2}} T_{\tilde{g}^{\alpha\delta}} L_{lh}(\partial_x^2 \partial_\alpha \partial_\delta u, \partial_\beta w)\, dxdt$$

$$+ \int_0^T\!\!\!\int_{\mathbb{R}^n} \partial_0 w \cdot T_{\mathfrak{P}S^{-2}} T_{\tilde{g}^{\alpha\delta}} L_{lh}(\partial_x^2 \partial_\delta u, \partial_\alpha \partial_\beta w)\, dxdt.$$

The first term is estimated by commuting $T_{\tilde{g}^{\alpha\delta}}$ inside L_{lh} and onto the first argument, after which we use Lemma 5.12. In the second term we pull ∂_β out, reducing the problem either to $I_{X,sub}^{042}$, which was discussed earlier, or to

$$\int_0^T\!\!\!\int_{\mathbb{R}^n} \partial_0 w \cdot T_{\mathfrak{P}S^{-2}} T_{\tilde{g}^{\alpha\delta}} L_{lh}(\partial_x^2 \partial_\delta \partial_\beta u, \partial_\alpha w)\, dxdt.$$

But here we can pull out one of the ∂_x operators to reduce to the case of (7.118).

After this discussion we have reduced the problem to the estimate for $I_{sub,\gamma\delta}$. Here, from easiest to hardest, we need to consider the case when neither of γ or δ is zero, then when one of them is zero, and finally when none of them is zero. We will first illustrate the principle in the easiest case, and then describe the additional complications for the most difficult case. We leave the intermediate case for the reader.

A. The case $\gamma, \delta \neq 0$. The argument in this case consists of three integrations by parts in a circular manner. Here we have $h \in \mathfrak{P}S^{-1}$. We may include ∂_δ in h in which case $h \in \mathfrak{P}S^0$. Separating variables, the problem can be further reduced to $h \in \mathfrak{P}$. In the computations below we omit h altogether, as it does not play any role. Then it remains to bound the integral

$$(7.119) \qquad I_{sub} = \int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_x^2 \partial_\alpha u,\ \partial_\beta w) \cdot w\, dxdt.$$

Similarly, derivatives applied to g yield perturbative contributions, of size $O(\mathcal{B}^2)$, and will not be explicitly written in order to avoid cluttering the formulas. In the absence of boundary terms, we compute as follows, integrating by parts in order to convert the null form into three $T_{\tilde{P}}$ operators modulo admissible errors:

$$I_{sub} = \int_0^T\!\!\!\int_{\mathbb{R}^n} w \cdot T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_\alpha \partial_x^2 u, \partial_\beta w)\, dxdt$$

$$= -\iint \partial_\beta w \cdot T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_\alpha \partial_x^2 u, w)\, dxdt - \iint w \cdot T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_\beta \partial_\alpha \partial_x^2 u, \partial_0 w)\, dxdt + Err(\mathcal{B}^2)$$

$$= \iint \partial_\alpha \partial_\beta w \cdot T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_x^2 u, w)\, dxdt + \iint \partial_\beta w \cdot T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_x^2 u, \partial_\alpha w)\, dxdt$$

$$- \iint w \cdot T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_\beta \partial_\alpha \partial_x^2 u, w)\, dxdt + Err(\mathcal{B}^2)$$

$$= \iint \partial_\alpha \partial_\beta w \cdot T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_x^2 u, w)\, dxdt - I_{sub} - \iint w \cdot T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_x^2 u, \partial_\alpha \partial_\beta w)\, dxdt$$

$$- \iint w \cdot T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_\beta \partial_\alpha \partial_x^2 u, w)\, dxdt + Err(\mathcal{B}^2).$$

89

We now distribute $T_{\tilde{g}^{\alpha\beta}}$, noting that any commutator errors involve derivatives of $\tilde{g}$ and thus are perturbative. We arrive at

(7.120)
$$2I_{sub} = \int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\tilde{g}^{\alpha\beta}}\partial_\alpha\partial_\beta w \cdot L_{lh}(\partial_x^2 u, w)\, dxdt - \int_0^T\!\!\!\int_{\mathbb{R}^n} w \cdot L_{lh}(\partial_x^2 u, T_{\tilde{g}^{\alpha\beta}}\partial_\alpha\partial_\beta w)\, dxdt$$
$$- \int_0^T\!\!\!\int_{\mathbb{R}^n} w \cdot L_{lh}(T_{\tilde{g}^{\alpha\beta}}\partial_\beta\partial_\alpha\partial_x^2 u, w)\, dxdt + Err(\mathcal{B}^2).$$

It remains to add the boundary terms at times $t = 0, T$ into the above computation. Such boundary terms arise from the integration by parts with respect to x_0. We obtain the following enhanced version of (7.120):

(7.121)
$$2I_{sub} = \int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\tilde{g}^{\alpha\beta}}\partial_\alpha\partial_\beta w \cdot L_{lh}(\partial_x^2 u, w)\, dxdt - \int_0^T\!\!\!\int_{\mathbb{R}^n} w \cdot L_{lh}(\partial_x^2 u, T_{\tilde{g}^{\alpha\beta}}\partial_\alpha\partial_\beta w)\, dxdt$$
$$- \int_0^T\!\!\!\int_{\mathbb{R}^n} w \cdot L_{lh}(T_{\tilde{g}^{\alpha\beta}}\partial_\beta\partial_\alpha\partial_x^2 u, w)\, dxdt + Err(\mathcal{B}^2)$$
$$+ \int w \cdot T_{\tilde{g}^{\alpha 0}} L_{lh}(\partial_\alpha\partial_x^2 u, w) - \partial_\beta w \cdot T_{\tilde{g}^{0\beta}} L_{lh}(\partial_x^2 u, w) + w \cdot T_{\tilde{g}^{\alpha 0}} L_{lh}(\partial_x^2 u, \partial_\alpha w)\, dx \Big|_0^T.$$

The boundary terms are easily seen as lower order energy corrections, so it remains to estimate the interior contributions. For the first one we can use the w equation to get the fixed time bounds

(7.122)
$$\|T_{\tilde{g}^{\alpha\beta}}\partial_\alpha\partial_\beta w\|_{L^2} \lesssim \|\tilde{P}_B w\|_{L^2} + (\mathcal{B}^2 + 2^{\frac{k}{2}}\mathcal{B}c_k)\|\partial w\|_{L^2},$$

which suffices[6] by combining the two components of the last term with either the $\mathcal{A}$ or the $\mathcal{B}$ bound for u in L_{lh}. The other two interior contributions reduce to the bound

(7.123)
$$\|P_{<k}T_{\tilde{g}^{\alpha\beta}}\partial_\alpha\partial_\beta\partial_x^2 u\|_{L^\infty} \lesssim 2^{2k}\mathcal{B}^2,$$

which is a consequence of Lemma 5.11 and which suffices to estimate the expressions in (7.121).

    B. The case $\gamma = \delta = 0$. In this case we seek to estimate the integral

(7.124)
$$I_{sub,00} = \int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_x\widehat{\partial_\alpha\partial_0}u,\ \partial_\beta w) \cdot \partial_0 w\, dxdt.$$

Here the hat correction plays a perturbative role and could be omitted. However, in the computations below we need to keep it in order to be able to estimate energy corrections. Our computations emulate the simpler case considered above, but with some care in order to avoid iterated time derivatives. Integrating by parts the β derivative we get

$$I_{sub,00} = -\int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_x\partial_\beta\widehat{\partial_\alpha\partial_0}u,\ w) \cdot \partial_0 w\, dxdt$$
$$- \int_0^T\!\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{\alpha\beta}} L_{lh}(\partial_x\widehat{\partial_\alpha\partial_0}u,\ w) \cdot \partial_0\partial_\beta w\, dx + Err(\mathcal{B}^2) + Err(\mathcal{A}^\sharp)|_0^T,$$

---

[6]Here the $\tilde{P}_B w$ term may be interpreted as arising from a lower order correction to our multiplier $\tilde{\mathfrak{M}}_s$.

where the last term accounts for the boundary contributions obtained when $\beta = 0$. In the first integral we perturbatively move $T_{\tilde{g}^{\alpha\beta}}$ on the first L_{lh} argument and then use Lemma 5.11; this allows us to move the entire first integral into the error, leading us to

$$(7.125) \quad I_{sub,00} = -\int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{\alpha\beta}}L_{lh}(\partial_x\widehat{\partial_\alpha\partial_0}u,\ \partial_\beta w)\cdot\partial_0\partial_\beta w\,dx + Err(\mathcal{B}^2) + Err(\mathcal{A}^\sharp)|_0^T.$$

On the other hand we can perturbatively drop the hat and integrate by parts the α derivative. This gives

$$\begin{aligned}
I_{sub,00} = &-\int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{\alpha\beta}}L_{lh}(\partial_x\partial_0 u,\ \partial_\alpha\partial_\beta w)\cdot\partial_0 w\,dxdt\\
&-\int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{\alpha\beta}}L_{lh}(\partial_x\partial_0 u,\ w)\cdot\partial_0\partial_\alpha w\,dx + Err(\mathcal{B}^2) + Err(\mathcal{A}^\sharp)|_0^T.
\end{aligned}$$

In the first integral we perturbatively move $T_{\tilde{g}^{\alpha\beta}}$ on the second L_{lh} argument; using the paradifferential w equation where the $\tilde{A}$ and $\tilde{b}$ terms are perturbative, this yields a lower order correction to our multiplier. Switching the α and β indices we obtain

$$\begin{aligned}
(7.126) \quad I_{sub,00} = &-\int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{\alpha\beta}}L_{lh}(\partial_x\partial_0 u,\ \partial_\alpha w)\cdot\partial_0\partial_\beta w\,dx + \int_0^T\!\!\int_{\mathbb{R}^n} T_{\tilde{P}}w\cdot Q\partial_0 w\,dxdt\\
&+ Err(\mathcal{B}^2) + Err(\mathcal{A}^\sharp)|_0^T,
\end{aligned}$$

where $\|Q\|_{L^2\to L^2} \lesssim \mathcal{A}^\sharp$.

The next step is to add the relations (7.125) and (7.126). Here we separate the cases $\alpha = j \neq 0$ and $\alpha = 0$, where in the first case we can drop the hat and pull out the ∂_α in L_{lh},

$$\begin{aligned}
2I_{sub,00} = &-\int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{j\beta}}\partial_j L_{lh}(\partial_x\partial_0 u,\ w)\cdot\partial_0\partial_\beta w\,dx\\
&-\int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{0\beta}}[L_{lh}(\partial_x\widehat{\partial_0\partial_0}u,\ w) + L_{hl}(\partial_x\partial_0 u, \partial_0 w)]\cdot\partial_0\partial_\beta w\,dx\\
&+\int_0^T\!\!\int_{\mathbb{R}^n} T_{\tilde{P}}w\cdot Q\partial_0 w\,dxdt + Err(\mathcal{B}^2) + Err(\mathcal{A}^\sharp)|_0^T.
\end{aligned}$$

In the first integral we integrate by parts to switch ∂_0 to the left and then ∂_j to the right. Then we distribute the ∂_0 on the left. This yields

$$\begin{aligned}
2I_{sub,00} = &-\int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{j\beta}}\partial_j[L_{lh}(\partial_x\partial_0\partial_0 u,\ w) + L_{lh}(\partial_x\partial_0 u,\ \partial_0 w)]\cdot\partial_j\partial_\beta w\,dx\\
&-\int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}}T_{\tilde{g}^{0\beta}}[L_{lh}(\partial_x\widehat{\partial_0\partial_0}u,\ w) + L_{hl}(\partial_x\partial_0 u, \partial_0 w)]\cdot\partial_0\partial_\beta w\,dx\\
&+\int_0^T\!\!\int_{\mathbb{R}^n} T_{\tilde{P}}w\cdot Q\partial_0 w\,dxdt + Err(\mathcal{B}^2) + Err(\mathcal{A}^\sharp)|_0^T.
\end{aligned}$$

91

In the first integral we may perturbatively correct $\partial_0^2 u$; this allows us to put back together the cases when α is zero and nonzero,

$$2I_{sub,00} = -\int_0^T\!\!\int_{\mathbb{R}^n} T_{\mathfrak{P}S^{-1}} T_{\tilde{g}^{\alpha\beta}} \partial_j [L_{lh}(\partial_x \widehat{\partial_0 \partial_0} u, \ w) + L_{lh}(\partial_x \partial_0 u, \ \partial_0 w)] \cdot \partial_\alpha \partial_\beta w \, dx$$
$$+ \int_0^T\!\!\int_{\mathbb{R}^n} T_{\tilde{P}} w \cdot Q\partial_0 w \, dxdt + Err(\mathcal{B}^2) + Err(\mathcal{A}^\sharp)|_0^T.$$

Finally, we commute $T_{\tilde{g}^{\alpha\beta}}$ to the right factor, and use the w equation to add another perturbative factor to our multiplier. This gives

$$2I_{sub,00} = \int_0^T\!\!\int_{\mathbb{R}^n} T_{\tilde{P}} w \cdot Q\partial_0 w \, dxdt + Err(\mathcal{B}^2) + Err(\mathcal{A}^\sharp)|_0^T,$$

with a modified Q, as desired.

The proof of Proposition 7.11 is now concluded. $\qquad\qquad\square$

We now conclude the proof of Theorem 7.1 using Proposition 7.11, with the vector field $\tilde{X}_s$ chosen as in Proposition 7.7 and $\tilde{Y}_{s0}$ defined as in (7.91). For these we have at our disposal not only the conclusion of Proposition 7.7, but also the refined version in Proposition 7.9. This guarantees that the symbol $\tilde{c}_s$ in (7.92) has size $\mathcal{B}^2$, in the sense that its coefficients in (7.93) satisfy

$$\tilde{c}_s^j \in \mathcal{B}^2 L^\infty S^j, \qquad 2^{-k} P_{<k} \partial_0 \tilde{c}_s^j \in \mathcal{B}^2 L^\infty S^j.$$

These conditions, in turn, guarantee that the flux term C_s in our energy estimate (7.96) satisfies

$$|C_s(w,w)| \lesssim \mathcal{B}^2 \|\partial w\|_{L^2}^2,$$

and thus the conclusion of Theorem 7.1 follows.

8. Energy estimates for the full equation

Our objective here is to prove energy estimates for the solution u to the minimal surface equation (1.7) in $\mathcal{H}^s = H^s \times H^{s-1}$, in terms of our control parameters $\mathcal{A}^\sharp$ and $\mathcal{B}$.

Theorem 8.1. *For each $s \geq 1$ there exists an energy functional E_{NL}^s for the minimal surface equation (1.7) in $H^s \times H^{s-1}$ with the property that for all $\mathcal{H}^s$ solutions u to (1.7) with $\mathcal{A}^\sharp \ll 1$ and $\mathcal{B} \in L^2$ we have:*

a) coercivity,

$$(8.1) \qquad\qquad E_{NL}^s(u[t]) \approx \|u[t]\|_{\mathcal{H}^s}^2.$$

b) energy bound,

$$(8.2) \qquad\qquad \frac{d}{dt} E_{NL}^s(u[t]) \lesssim \mathcal{B}^2 E_{NL}^s(u[t]).$$

Because of the assumption $\mathcal{A}^\sharp \ll 1$, in this section we no longer need to track the dependence of implicit constants on $\mathcal{A}^\sharp$. The exception to this is in the proof of Lemma 8.4, where the smallness of $\mathcal{A}^\sharp$ is used in order to guarantee the invertibility of our partial normal form transformation; even there, we only need to use linear and quadratic $\mathcal{A}^\sharp$ factors.

The rest of this section is devoted to the proof of the theorem. This has two main ingredients:

(1) Reduction to the paradifferential equation, using normal form analysis.
(2) Energy estimates for the paradifferential equation, which have already been proved in Theorem 7.1.

Hence, our primary objective here will be to carry out the above reduction. We recall the minimal surface equation,

$$g^{\alpha\beta}\partial_\alpha\partial_\beta u = 0.$$

In order to use the energy estimates obtained in the previous section, we write this in paradifferential form:

(8.3) $$(T_{g^{\alpha\beta}}\partial_\alpha\partial_\beta - 2T_{A^\gamma}\partial_\gamma)u = N(u),$$

where the source term $N(u)$ is given by

(8.4) $$N(u) = -\Pi(\partial_\alpha\partial_\beta u, g^{\alpha\beta}) - T_{\partial_\alpha\partial_\beta u}g^{\alpha\beta} - 2T_{A^\gamma}\partial_\gamma u.$$

Here we cannot treat N perturbatively; precisely, we do not have an estimate of the form

$$\|N(u)(t)\|_{H^{s-1}} \lesssim \mathcal{B}^2\|u[t]\|_{\mathcal{H}^s},$$

even though $N(u)$ is cubic in u, and the above inequality is dimensionally correct. This is because N contains some unbalanced contributions.

To address this issue, our strategy will be to correct u via a well chosen normal form transformation, in order to eliminate the unbalanced part of $N(u)$. But in order to do this, we have to first identify the unbalanced part of $N(u)$, and reveal its null structure. A first step in this direction is to better describe the contributions of the metric coefficients $g^{\alpha\beta}$ in N; explicitly we want to extract the renormalizable terms (i.e. the terms to which we can apply a normal form correction). For this we express $g^{\alpha\beta}$ paradifferentially as follows:

Lemma 8.2. *The metric $g^{\alpha\beta}$ can be expressed paradifferentially as follows*

(8.5) $$g^{\alpha\beta}(\partial u) = g^{\alpha\beta}(0) - T_{g^{\alpha\gamma}\partial^\beta u}\partial_\gamma u - T_{g^{\gamma\beta}\partial^\alpha u}\partial_\gamma u + R(u),$$

where $R(u)$ satisfies the following balanced bounds for $s \geq 1$:

(8.6) $$\|R(u)\|_{H^{s-\frac{1}{2}}} \lesssim \mathcal{B}\|\partial u\|_{H^{s-1}},$$

as well as

(8.7) $$\|R(u)\|_{H^{s-1}} \lesssim \|\partial u\|_{H^{s-1}}.$$

Proof. The representation in (8.5) and the bound (8.6) for R follow from (3.8) and Lemma 2.9. To get (8.7) one estimates each term in R separately, using no cancellations. $\square$

This suggests that the nonlinear contribution $N(u)$ should be seen as the sum of two terms

$$N(u) = N_1(u) + N_2(u),$$

where N_1 has null structure and N_2 is balanced,

$$N_1(u) = -2\Pi(\partial_\alpha\partial_\beta u, T_{g^{\alpha\gamma}\partial^\beta u}\partial_\gamma u),$$

$$N_2(u) = -2\left(T_{\partial_\alpha\partial_\beta u}T_{g^{\alpha\gamma}\partial^\beta u}\partial_\gamma u - T_{\partial_\alpha\partial_\beta u g^{\alpha\gamma}\partial^\beta u}\partial_\gamma u\right) + T_{\partial_\alpha\partial_\beta u}R(u) + \Pi(\partial_\alpha\partial_\beta u, R(u)).$$

We will first prove that $N_2(u)$ is a perturbative term:

93

l:n2 **Lemma 8.3.** *The expression given by $N_2(u)$ satisfies the bound*

(8.8) $$\|N_2(u)\|_{H^{s-1}} \lesssim \mathcal{B}^2 \|\partial u\|_{H^{s-1}}, \qquad s \geq 1.$$

Proof. We begin with the first difference in N_2, and look separately at each α, β and γ. If $(\alpha, \beta) \neq (0,0)$ then we apply Lemma 2.7 (**l:para-prod**) to obtain

$$\|\big(T_{\partial_\alpha \partial_\beta u} T_{g^{\alpha\gamma}\partial^\beta u} - T_{\partial_\alpha \partial_\beta u g^{\alpha\gamma}\partial^\beta u}\big)\partial_\gamma u\|_{H^{s-1}} \lesssim \||D|^{-\frac12}\partial_\alpha \partial_\beta u\|_{BMO}\||D|^{\frac12}\big(g^{\alpha\gamma}\partial^\beta u\big)\|_{BMO}\|\partial u\|_{H^{s-1}}$$
$$\lesssim \mathcal{B}^2 \|\partial u\|_{H^{s-1}}.$$

If $(\alpha, \beta) = (0,0)$ then we use the wave equation for u with the $\tilde{g}$ metric (3.19) (**ms-tg**) to write

dt2u-dec (8.9) $$\partial_t^2 u = (T_{\tilde{g}}\partial_x \partial u + T_{\partial_x \partial u}\tilde{g}) + \Pi(\tilde{g}, \partial_x \partial u) =: \hat{\partial}_0^2 u + \pi_2(u),$$

exactly as in Lemma 5.4 (**l:utt**). Then for the first term we have the estimate

f1 (8.10) $$\|\hat{\partial}_0^2 u\|_{BMO^{-\frac12}} \lesssim \mathcal{B}$$

which suffices in order to apply Lemma 2.7 (**l:para-prod**) as above in order to estimate its contribution.

On the other hand, the bound for the contribution of $\pi_2(u)$ is easier because by Lemma 5.4 (**l:utt**) we have the direct uniform bound

f2 (8.11) $$\|\pi_2(u)\|_{L^\infty} \lesssim \mathcal{B}^2.$$

Now, we turn our attention to the second term in $N_2(u)$, where we again discuss separately the $(\alpha, \beta) \neq (0,0)$ and $(\alpha, \beta) = (0,0)$ cases.

For the $(\alpha, \beta) \neq (0,0)$ case we use the bound in (8.6) (**Rb**) to obtain

$$\|T_{\partial_\alpha \partial_\beta u} R(u)\|_{H^{s-1}} \lesssim \|\partial_\alpha \partial_\beta u\|_{BMO^{-\frac12}}\|R(u)\|_{H^{s-\frac12}} \lesssim \mathcal{B}^2\|\partial u\|_{H^{s-1}}.$$

Next we consider the case $(\alpha, \beta) = (0,0)$, and observe that we again need to use the decomposition (8.9) (**dt2u-dec**). The contribution of $\hat{\partial}_0^2 u$ is estimated using (8.10) (**f1**) and the bound (8.6) (**Rb**) for R, exactly as above:

$$\|T_{\hat{\partial}_0^2 u} R(u)\|_{H^{s-1}} \lesssim \||D|^{-\frac12}\hat{\partial}_0^2 u\|_{BMO}\|R(u)\|_{H^{s-\frac12}} \lesssim \mathcal{B}^2\|\partial u\|_{H^{s-1}}.$$

For the $\pi_2(u)$ contribution we use the pointwise bound (8.11) (**f2**) and the H^{s-1} bound (8.7) (**Reb**) for R,

$$\|T_{\pi_2(u)} R(u)\|_{H^{s-1}} \lesssim \||\pi_2(u)\|_{L^\infty}\|R(u)\|_{H^{s-1}} \lesssim \mathcal{B}^2\|\partial u\|_{H^{s-1}}.$$

Finally, a similar analysis leads to the bound for the balanced term $\Pi(\partial_\alpha \partial_\beta u, R(u))$.

$\square$

To account for the unbalanced part $N_1(u)$ of N we introduce a normal form correction

u-nft (8.12) $$\tilde{u} = u - \Pi(\partial_\beta u, T_{\partial^\beta u}u) := u - u_2.$$

Our goal will be to show that the normal form variable solves a linear inhomogeneous paradifferential equation with a balanced source term.

l:nfu2 **Lemma 8.4.** *The normal form correction above has the following properties:*

a) It is bounded from above and below[7],

$$\|\tilde{u}[t]\|_{\mathcal{H}^s} \approx \|u[t]\|_{\mathcal{H}^s}.$$

---

[7]A slight expansion of the argument shows that it is in effect invertible.

b) It solves the an equation of the form

important (8.13)
$$(\partial_\alpha T_{g^{\alpha\beta}}\partial_\beta - T_{A^\gamma}\partial_\gamma)\tilde{u} = N_2(u) + \partial_t R_1(u) + R_2(u),$$

where

R1R2 (8.14)
$$\|R_1(u)\|_{H^s} \lesssim \mathcal{B}^2\|u[t]\|_{\mathcal{H}^s}, \qquad \|R_2(u)\|_{H^{s-1}} \lesssim \mathcal{B}^2\|u[t]\|_{\mathcal{H}^s},$$

and

R1 (8.15)
$$\|R_1(u)\|_{H^{s-1}} \lesssim \mathcal{A}^{\sharp 2}\|u[t]\|_{\mathcal{H}^s}.$$

We remark that here we expand the meaning of "balanced source terms" to include expressions of the form $\partial_t R_1$ with R_1 as above. This is required due to the fact that time derivatives are often more difficult to estimate in our context, and are allowed in view of the result in Theorem 4.5.

Proof. a) In view of the smallness of $\mathcal{A}^\sharp$, for the boundedness of the normal form it suffices to show that we have the fixed time bound

u2 (8.16)
$$\|u_2\|_{H^s} \lesssim \mathcal{A}^{\sharp 2}\|u[t]\|_{\mathcal{H}^s},$$

as well as

dtu2 (8.17)
$$\|\partial_t u_2\|_{H^{s-1}} \lesssim \mathcal{A}^{\sharp 2}\|u[t]\|_{\mathcal{H}^s}.$$

For the first bound we directly have

(8.18)
$$\|\Pi(\partial_\beta u, T_{\partial^\beta u}u)\|_{H^s} \lesssim \|\partial_\beta u\|_{H^{s-1}}\|T_{\partial^\beta u}u\|_{BMO^1} \lesssim \mathcal{A}^{\sharp 2}\|u[t]\|_{\mathcal{H}^s}.$$

We prove the second bound in a similar manner, but we first apply the time derivative and analyze each term separately:

$$\partial_t\Pi(\partial_\beta u, T_{\partial^\beta u}u) = \Pi(\partial_t\partial_\beta u, T_{\partial^\beta u}u) + \Pi(\partial_\beta u, T_{\partial^\beta u}\partial_t u) + \Pi(\partial_\beta u, T_{\partial_t\partial^\beta u}u) := r_1 + r_2 + r_3.$$

There are multiple cases arising from the strategy we will implement for terms involving two times derivatives, as well as from the particular structure of each of the terms.

We begin with r_1, where we need to separate the $\beta \neq 0$ and $\beta = 0$ cases. The easiest case is when $\beta \neq 0$, where we have

pi1 (8.19)
$$\|\Pi(\partial_t\partial_\beta u, T_{\partial^\beta u}u)\|_{H^{s-1}} \lesssim \|T_{\partial^\beta u}u\|_{H^s}\sup_k 2^{-k}\|P_{<k}\partial_t\partial_\beta u\|_{L^\infty} \lesssim \mathcal{A}^{\sharp 2}\|u[t]\|_{\mathcal{H}^s}.$$

Here we have used the energy control we have for $\partial_t u$, which in turn gives control of all spatial derivatives of $\partial_t u$. For the case $\beta = 0$ we use the decomposition for $\partial_t^2 u$ as in Lemma 5.4. For the first component we use the second bound in (5.15),

$$\|\Pi(\hat{\partial}_t^2 u, T_{\partial^\beta u}u)\|_{H^{s-1}} \lesssim \|T_{\partial^\beta u}u\|_{H^s}\sup_k 2^{-k}\|P_{<k}\hat{\partial}_t^2 u\|_{L^\infty} \lesssim \mathcal{A}^{\sharp 2}\|u[t]\|_{\mathcal{H}^s}.$$

For the second component we argue similarly but using the second bound in (5.16) together with Bernstein's inequality,

$$\|\Pi(\pi_2(u), T_{\partial^\beta u}u)\|_{H^{s-1}} \lesssim \|T_{\partial^\beta u}u\|_{H^s}\sup_k 2^{-k}\|P_{<k}\pi_2(u)\|_{L^\infty} \lesssim \mathcal{A}^{\sharp 2}\|u[t]\|_{\mathcal{H}^s}.$$

We continue with the bound for r_2, where we do not need to distinguish between the time and space derivatives,

$$\|\Pi(\partial_\beta u, T_{\partial^\beta u}\partial_t u)\|_{H^{s-1}} \lesssim \|T_{\partial^\beta u}\partial_t u\|_{H^{s-1}}\|\partial_\beta u\|_{L^\infty} \lesssim \mathcal{A}\|\partial^\beta u\|_{L^\infty}\|\partial_t u\|_{H^{s-1}} \lesssim \mathcal{A}^{\sharp 2}\|u[t]\|_{\mathcal{H}^s}.$$

95

Lastly, we need to bound r_3. Here we argue as in ($\overset{\texttt{pi1}}{8.19}$),

$$\|r_3\|_{H^{s-1}} \lesssim \|\partial_\beta u\|_{L^\infty}\|T_{\partial_t\partial^\beta u}u\|_{H^{s-1}} \lesssim \mathcal{A}\|u\|_{H^s}\sup_k 2^{-k}\|P_{<k}\partial_t\partial^\beta u\|_{L^\infty},$$

so it remains to prove the following bound for $\partial_t\partial^\beta u$,

$\boxed{\texttt{pi3}}$ (8.20)
$$\|P_{<k}\partial_t\partial^\beta u\|_{L^\infty} \lesssim 2^k\mathcal{A}^\sharp.$$

Here we use the chain rule for the paracoefficient to write $\partial_t\partial^\beta u$ as a linear combination of $\partial_t^2 u$ and $\partial_t\partial_x u$,

$$\partial_t\partial^\beta u := \partial_t h(\partial u) = h(\partial u)\partial_t^2 u + h(\partial u)\partial_t\partial_x u$$

where $h := h(\partial u)$. For the ∂_t^2 term we use the minimal surface equation ($\overset{\texttt{msf-short}}{3.5}$), arriving at the representation

$$\partial_t\partial^\beta u = \tilde{h}(\partial u)\partial_x\partial_x u,$$

where $\tilde{h}$ incorporates the corresponding metric coefficients. As before, we need to use a Littlewood-Paley decomposition

$$\tilde{h}(\partial u)\partial_x\partial u = T_{\tilde{h}(\partial u)}\partial_x\partial u + T_{\partial_x\partial u}\tilde{h}(\partial u) + \Pi(\tilde{h}(\partial u), \partial_x\partial u).$$

The first two terms are easy to estimate using only L^∞ bounds for ∂u and $\tilde{h}(\partial u)$,

$$\|P_{<k}T_{\tilde{h}(\partial u)}\partial_x\partial u\|_{L^\infty} + \|P_{<k}T_{\partial_x\partial u}\tilde{h}(\partial u)\|_{L^\infty} \lesssim 2^k\mathcal{A}.$$

Finally for the third we use instead the $\mathcal{A}^\sharp$ component of the $\mathfrak{C}_0$ norm for both ∂u and $\tilde{h}(\partial u)$,

$$\|P_{<k}\Pi(\tilde{h}(\partial u), \partial_x\partial u)\|_{L^\infty} \lesssim 2^k\|\Pi(\tilde{h}(\partial u), \partial_x\partial u)\|_{L^n} \lesssim 2^k\mathcal{A}^{\sharp^2}.$$

Hence ($\overset{\texttt{pi3}}{8.20}$) follows. This finishes the proof of ($\overset{\texttt{dtu2}}{8.17}$), and thus of the boundedness from above and below of the normal form transformation in our desired Sobolev space $\mathcal{H}^s$.

b) We begin with the supposedly easier contribution, meaning with the term $T_{A^\gamma}\partial_\gamma u_2$. To bound this term we would like to commute the ∂_γ and place it in front of the product,

$$T_{A^\gamma}\partial_\gamma u_2 = \partial_\gamma T_{A^\gamma}u_2 - T_{\partial_\gamma A^\gamma}u_2.$$

This would look good for the first term on the RHS. However the last term would be problematic, as it may contain three derivatives with respect to time. To avoid this issue we first substitute A^γ, which by ($\overset{\texttt{def-A}}{3.13}$) is given by

$$A^\gamma := \partial^\alpha u g^{\alpha\gamma}\partial_\alpha\partial_\beta u,$$

with the more manageable leading part $\mathring{A}^\gamma$ given by

$$\mathring{A}^\gamma := T_{\partial^\alpha u}T_{g^{\alpha\gamma}}\widehat{\partial_\alpha\partial_\beta u}.$$

Here the hat correction is from the Definition $\overset{\texttt{d:dt2u}}{5.3}$. Then

$$T_{A^\gamma}\partial_\gamma u_2 = (T_{A^\gamma} - T_{\mathring{A}^\gamma})\partial_\gamma u_2 + T_{\mathring{A}^\gamma}\partial_\gamma u_2$$
$$= (T_{A^\gamma} - T_{\mathring{A}^\gamma})\partial_\gamma u_2 + \partial_\gamma T_{\mathring{A}^\gamma}u_2 - T_{\partial_\gamma\mathring{A}^\gamma}u_2.$$

We will successively place each of these terms in $\partial_t R_1$ or R_2. We place the first term in R_2. To prove ($\overset{\texttt{R1R2}}{8.14}$) for this term, we use the bounds ($\overset{\texttt{u2}}{8.16}$) and ($\overset{\texttt{dtu2}}{8.17}$) for u_2, and Lemma $\overset{\texttt{l:circle-app}}{6.9}$ to bound the coefficient

$\boxed{\texttt{diffA}}$ (8.21)
$$\|A^\gamma - \mathring{A}^\gamma\|_{L^\infty} \lesssim \mathcal{B}^2.$$

We will place the second term in $\partial_t R_1$ if $\gamma = 0$ and in R_2 if $\gamma \neq 0$.

In order to prove (8.14) we measure u_2 in $H^{s+\frac{1}{2}}$,

$$\|u_2\|_{H^{s+\frac{1}{2}}} \lesssim \|\partial u\|_{BMO^{\frac{1}{2}}} \|T_{\partial u} u\|_{H^s} \lesssim \|\partial u\|_{BMO^{\frac{1}{2}}} \|\partial u\|_{L^\infty} \|u\|_{H^s} \lesssim \mathcal{A}\mathcal{B}\|u[t]\|_{\mathcal{H}^s}.$$

Then (8.14) follows if we can bound the coefficient $\mathring{A}$ by

$$\|P_{<k} \mathring{A}^\gamma\|_{L^\infty} \lesssim 2^{\frac{k}{2}} \mathcal{B}.$$

But this is also a consequence of Lemma 6.9, see (6.12). On the other hand for the bound (8.15) we estimate u_2 in H^s as in (8.16), and then it suffices to show that

$$\|P_{<k} \mathring{A}^\gamma\|_{L^\infty} \lesssim 2^k \mathcal{A}^\sharp,$$

which is similar.

The last term is placed in R_2 using again the bounds (8.16) and (8.17) for u_2 on one hand and Lemma 5.12 on the other hand, to obtain

$$\|P_{<k} \partial_\gamma \mathring{A}^\gamma\|_{L^\infty} \lesssim 2^k \mathcal{B}^2.$$

Now we consider the main term $\partial_\alpha T_{g^{\alpha\beta}} \partial_\beta u_2$, which can be expanded as

$$\partial_\alpha T_{g^{\alpha\beta}} \partial_\beta u_2 = \partial_\alpha T_{g^{\alpha\beta}} \partial_\beta \left[\Pi(\partial_\gamma u, T_{\partial^\gamma u} u) \right]$$
$$= \partial_\alpha T_{g^{\alpha\beta}} \left[\Pi(\partial_\beta \partial_\gamma u, T_{\partial^\gamma u} u) + \Pi(\partial_\gamma u, T_{\partial_\beta \partial^\gamma u} u) + \Pi(\partial_\gamma u, T_{\partial^\gamma u} \partial_\beta u) \right].$$

Depending on whether $\alpha = 0$ or $\alpha \neq 0$, we place the middle term into $\partial_t R_1$ or R_2, respectively:

$$\|\Pi(\partial_\gamma u, T_{\partial_\beta \partial^\gamma u} u)\|_{H^s} \lesssim \mathcal{B}^2 \|u[t]\|_{\mathcal{H}^s},$$

$$\|\Pi(\partial_\gamma u, T_{\partial_\beta \partial^\gamma u} u)\|_{H^{s-1}} \lesssim \mathcal{A}^{\sharp^2} \|u[t]\|_{\mathcal{H}^s}.$$

Here we use the property $\partial_\beta \partial^\gamma u \in \mathfrak{DC}$ to handle the case when $\beta = 0$ for the first bound, and (8.20) for the second.

The first term can be rewritten in the form

$$(8.22) \qquad \partial_\alpha T_{g^{\alpha\beta}} \Pi(\partial_\beta \partial_\gamma u, T_{\partial^\gamma u} u) = \partial_\alpha \Pi(T_{g^{\alpha\beta}} \widehat{\partial_\beta \partial_\gamma} u, T_{\partial^\gamma u} u) + \partial_t R_1 + R_2,$$

by using Lemma 2.5, as well as Lemma 5.4 for the case $(\beta, \gamma) = (0, 0)$. Similarly the last term can be rewritten in the analogous form

$$(8.23) \qquad \partial_\alpha T_{g^{\alpha\beta}} \Pi(\partial_\gamma u, T_{\partial^\gamma u} \partial_\beta u) = \partial_\alpha \Pi(\partial_\gamma u, T_{\partial^\gamma u} T_{g^{\alpha\beta}} \partial_\beta u) + \partial_t R_1 + R_2.$$

Finally we distribute the α derivative in both (8.22) and (8.23). For the first term on the right in (8.22) we get

$$\partial_\alpha \Pi(T_{g^{\alpha\beta}} \widehat{\partial_\beta \partial_\gamma} u, T_{\partial^\gamma u} u) = \Pi(\partial_\alpha T_{g^{\alpha\beta}} \widehat{\partial_\beta \partial_\gamma} u, T_{\partial^\gamma u} u) + \Pi(T_{g^{\alpha\beta}} \widehat{\partial_\beta \partial_\gamma} u, T_{\partial_\alpha \partial^\gamma u} u)$$
$$+ \Pi(T_{g^{\alpha\beta}} \widehat{\partial_\beta \partial_\gamma} u, T_{\partial^\gamma u} \partial_\alpha u)$$
$$:= s_1 + s_2 + s_3.$$

We place s_1 in R_2 using Lemma 5.12,

$$\|s_1\|_{H^{s-1}} \lesssim \sup_k 2^{-k} \|P_{<k} \partial_\alpha T_{g^{\alpha\beta}} \widehat{\partial_\beta \partial_\gamma} u\|_{L^\infty} \|u[t]\|_{\mathcal{H}^s} \lesssim \mathcal{B}^2 \|u[t]\|_{\mathcal{H}^s}.$$

97

The term s_2 is also estimated perturbatively using the fact that $\partial_\alpha \partial^\gamma u \in \mathfrak{DC}$, which allows us to decompose it as a sum $f_1 + f_2$ as in (5.20). Then we estimate

$$\|s_2\|_{H^{s-1}} \lesssim \sup_k 2^{-\frac{k}{2}} \|P_{<k} T_{g^{\alpha\beta}} \widehat{\partial_\beta \partial_\gamma} u\|_{L^\infty} \sup_j 2^{-\frac{j}{2}} \|P_{<j} f_1\|_{L^\infty} \|u[t]\|_{\mathcal{H}^s}$$
$$+ \sup_k 2^{-k} \|P_{<k} T_{g^{\alpha\beta}} \widehat{\partial_\beta \partial_\gamma} u\|_{L^\infty} \|f_2\|_{L^\infty} \|u[t]\|_{\mathcal{H}^s}$$
$$\lesssim \mathcal{B}^2 \|u[t]\|_{\mathcal{H}^s},$$

using Lemma 5.4 for $\widehat{\partial_\beta \partial_\gamma} u$. In s_3 we can switch T_g onto the other argument of Π using Lemma 2.5 and remove the hat correction, so that it becomes half of N_1.

The last remaining term to bound is the one on the RHS of (8.23). Here we distribute again the α-derivative

$$\partial_\alpha \Pi(\partial_\gamma u, T_{\partial^\gamma u} T_{g^{\alpha\beta}} \partial_\beta u) = \Pi(\partial_\alpha \partial_\gamma u, T_{\partial^\gamma u} T_{g^{\alpha\beta}} \partial_\beta u) + \Pi(\partial_\gamma u, T_{\partial_\alpha \partial^\gamma u} T_{g^{\alpha\beta}} \partial_\beta u)$$
$$+ \Pi(\partial_\gamma u, T_{\partial^\gamma u} T_{\partial_\alpha g^{\alpha\beta}} \partial_\beta u) + \Pi(\partial_\gamma u, T_{\partial^\gamma u} T_{g^{\alpha\beta}} \partial_\beta \partial_\alpha u).$$

By inspection we observe that the first term in the equality above is the second half of N_1. The remaining three terms can be estimated perturbatively using exactly the same approach as in the case of (8.22). $\qquad\square$

In view of Lemma 8.3 we can include $N_2(u)$ into $R_2(u)$ in (8.13), obtaining the shorter representation of the source term

fimportant

(8.24)
$$(\partial_\alpha T_{g^{\alpha\beta}} \partial_\beta - T_{A^\gamma} \partial_\gamma)\tilde{u} = \partial_t R_1(u) + R_2(u),$$

where R_1 and R_2 satisfy the bounds (8.14) and (8.15).

For the homogeneous paradifferential problem we have the $\mathcal{H}^s$ energy E^s given by Theorem 7.1. We will use this to construct our desired nonlinear energy E_{NL}^s in Theorem 8.1. Because we have the source term $\partial_t R_1$, the associated nonlinear energy will not be simply given by $E^s(\tilde{u}[t])$. Instead, the correct energy is the one provided by Theorem 4.5, namely

(8.25)
$$E_{NL}^s(u[t]) := E^s(\tilde{u}[t] - r[t]),$$

where the correction $r[t]$ is given by

(8.26)
$$r[t] = \begin{pmatrix} 0 \\ (T_{g^{00}})^{-1} R_1(u) \end{pmatrix}.$$

Then by the estimate in (4.29) we obtain

(8.27)
$$\frac{d}{dt} E_{NL}^s(u[t]) \lesssim E^s(\tilde{u}[t] - r[t], r_1) + \mathcal{B}^2 E_{NL}^s(u[t]),$$

where $r_1[t]$ is as in (4.28),

$$r_1[t] := \begin{pmatrix} (T_{g^{00}})^{-1} R_1 \\ (T_{g^{00}})^{-1}(R_2 - \partial_k T_{g^{k0}} (T_{g^{00}})^{-1} R_1 - T_{g^{0k}} \partial_k (T_{g^{00}})^{-1} R_1) \end{pmatrix}.$$

Our nonlinear energy E_{NL}^s is coercive because $r[t]$ is small,

$$\|r[t]\|_{\mathcal{H}^s} \lesssim \mathcal{A}^{\sharp 2} \|u[t]\|_{\mathcal{H}^s},$$

due to the bound (8.15). Finally, we control the time derivative of the energy, because

$$\|r_1[t]\|_{\mathcal{H}^s} \lesssim \mathcal{B}^2 \|u[t]\|_{\mathcal{H}^s}.$$

This is due to the bound in (8.14).

<div style="text-align: right">R1R2</div>

9. Energy and Strichartz estimates for the linearized equation

<div style="float:left">linearized</div>

Our objective here is to prove that the homogeneous linearized equation is well-posed in the space $\mathcal{H}^{\frac{1}{2}}$ (respectively $\mathcal{H}^{\frac{5}{8}}$ in two dimensions) and satisfies Strichartz estimates with an appropriate loss of derivatives, namely (4.33) with $S = S_{AIT}$, under the assumption that the associated linear paradifferential equation has similar properties. The main result of this section is as follows:

<div style="float:left">linearized</div>

Theorem 9.1. *Let s be as in (1.20) (respectively (1.19) in dimension $n = 2$). Let u be a smooth solution for the minimal surface equation in a unit time interval, and which satisfies the energy and Strichartz bounds*

<div style="float:left">u-have-now</div>

$$(9.1) \qquad \|u\|_{S^s_{AIT}} + \|\partial_t u\|_{S^{s-1}_{AIT}} \ll 1.$$

Assume also that the associated linear paradifferential equation

<div style="float:left">-inhom-re+</div>

$$(9.2) \qquad \partial_\alpha T_{\hat{g}^{\alpha\beta}} \partial_\beta v = f$$

is well-posed in $\mathcal{H}^{\frac{1}{2}}$ (respectively $\mathcal{H}^{\frac{5}{8}}$ in dimension $n = 2$) in a unit time interval, and satisfies the full Strichartz estimates (4.43), with $S = S_{AIT}$, in the same interval.

Then the homogeneous linearized equation

<div style="float:left">lin-inhom</div>

$$(9.3) \qquad \partial_\alpha \hat{g}^{\alpha\beta} \partial_\beta v = 0$$

is also well-posed in $\mathcal{H}^{\frac{1}{2}}$ (respectively $\mathcal{H}^{\frac{5}{8}}$ in dimension $n = 2$), and satisfies the homogeneous form of the Strichartz estimates in (4.33) with $S = S_{AIT}$.

We continue with several comments on the result in the theorem, in order to better place it into context.

- Up to this point we only know that both the full equation and the associated linear paradifferential equation satisfy good energy estimates, but we do not yet know that they also satisfy the corresponding Strichartz estimates. This is however not a problem, as the main result of this section, namely Theorem 9.1 above, will only be used as a module within our main bootstrap argument in the last section of the paper, by which time we will have already established the energy and Strichartz estimates for both the full equation and for the linear paradifferential equation.
- The exponent s in the above result needs not be the same as the one in our main result in Theorem 1.3; it can be taken to be smaller, as long as it still satisfies the constraints in (1.19), (1.20).
- While we can no longer control the linearized evolution purely in terms of the control parameters $\mathcal{A}$, $\mathcal{A}^\sharp$, and $\mathcal{B}$, these still play role in the analysis. The hypothesis of the theorem guarantees that

$$\mathcal{A}^\sharp \ll 1, \qquad \|\mathcal{B}\|_{L^2} \ll 1.$$

- The exponent δ in the bound (4.43) with $S = S_{AIT}$ should be thought of as being sufficiently small, compared with the distance between s and the threshold in (1.19), (1.20).

<div style="text-align: center">99</div>

- The smoothness assumption on u is not used in any quantitative way. Its role is only to ensure that we already have solutions for the linearized problem, so we can skip an existence argument. Thus, by a density argument, the result of the theorem may be seen as an a-priori estimate for smooth solutions v to the linearized equation. As our rough solutions will be obtained in the last section as limits of smooth solutions, this assumption may be discarded a posteriori.
- The reason we only consider the homogeneous case in the linearized equation (9.3) is to shorten the proof, as this is all that is used later in the last section. However, the result also extends to the inhomogeneous case. In particular in dimension $n \geq 3$ this is an immediate consequence of Theorem 4.12, but in dimension $n = 2$ an extra argument would be needed.

One major simplification in this section, compared with the previous two sections, is that we no longer have the earlier difficulties in estimating the second order time derivatives and even some third order time derivatives of u. In particular, we have the following relatively straightforward lemma:

Lemma 9.2. *For solutions u to the minimal surface equation as in (9.1) we have the bounds*

$$(9.4) \qquad \|\partial^2 u\|_{S_{AIT}^{s-2}} + \|\partial \hat{g}\|_{S_{AIT}^{s-2}} \ll 1,$$

as well as

$$(9.5) \qquad \|\langle D_x \rangle^{\sigma + \delta_0} \partial_\alpha T_{\hat{g}^{\alpha\beta}} \partial_\beta \partial_\gamma u\|_{L^p L^q} \ll 1,$$

where $\delta_0 > 0$ depends only on s and

$$\frac{1}{p} + \frac{1}{q} = 1, \qquad \frac{2}{n - \frac{1}{2}} \leq \frac{1}{q} \leq \frac{1}{2} + \frac{2}{n - \frac{1}{2}}, \qquad \sigma = \left(n - \frac{1}{2}\right)\left(\frac{1}{q} - \frac{2}{n - \frac{1}{2}}\right) \geq 0, \qquad n \geq 3.$$

respectively

$$\frac{2}{p} + \frac{1}{q} = 1, \qquad \frac{4}{7} \leq \frac{1}{q} \leq \frac{11}{14}, \qquad \sigma = \frac{7}{4}\left(\frac{1}{q} - \frac{4}{7}\right) \geq 0, \qquad n = 2.$$

Proof. For (9.4) we only need to consider the second order time derivatives, which we can write using the minimal surface equation as

$$\partial_t^2 u = h(\partial u)\partial_x \partial u.$$

By Moser inequalities we have $\|h(\partial u)\|_{H^{s-1}} \ll 1$. Since $s - 1 > n/2$, it is easily verified that the space S_{AIT}^{s-2} is left unchanged by multiplication by $h(\partial u)$. The same argument applies to derivatives of the metric $\partial \hat{g}$.

For (9.5) we can use again the minimal surface equation to obtain the representation

$$\partial_\alpha T_{\hat{g}^{\alpha\beta}} \partial_\beta \partial_\gamma u = T_{\partial_x \partial \hat{g}} \partial u + T_{\partial \hat{g}} \partial^2 u + \Pi(\partial \hat{g}, \partial^2 u) + \Pi(\hat{g}, \partial^3 u),$$

where we can further write

$$\partial^3 u = \partial_x(\hat{g}\partial^2 u) + \partial \hat{g}\partial^2 u.$$

Hence we need to multiply two functions in S_{AIT}^{s-2}, which contains a range of mixed L^p norms at varying spatial Sobolev regularities. We can do this optimally if both mixed norms can be chosen to have non-negative Sobolev index. In order to avoid using Sobolev embeddings we further limit the range of exponents to the case when one of the Sobolev indices may be taken to be zero. This gives the range of exponents in the lemma. $\qquad \square$

100

Next we discuss the strategy of the proof. The first potential strategy here would be to try to view the equation (9.3) as a perturbation of (9.2). Unfortunately such a strategy does not seem to work in our context, because this would require a balanced estimate for the difference between the two operators, whereas this difference contains some terms which are clearly unbalanced.

To address the above difficulty, the key observation is that the aforementioned difference exhibits a null structure, at least in its unbalanced part. This opens the door to a partial normal form analysis, in order to develop a better reduction of (9.3) to (9.2). Because of this, the proof of the theorem will be done in two steps:

(1) The normal form analysis, where a suitable normal form transformation is constructed.
(2) Reduction to the paradifferential equation, using the above normal form.

9.1. Preliminary bounds for the linearized variable.

The starting point for the proof of the theorem is to rewrite the divergence form of the linearized equation (9.3) as an inhomogeneous paradifferential evolution (9.2) with a perturbative source f, as follows:

$$\text{(9.6)} \qquad T_{\hat{P}}v = -\partial_\alpha T_{\partial_\beta v}\hat{g}^{\alpha\beta} - \partial_\alpha \Pi(\partial_\beta v, \hat{g}^{\alpha\beta}) =: f.$$

While we cannot directly prove a balanced cubic estimate for f, a useful initial step is to establish a quadratic estimate for it. The expression for f involves v and $\partial_t v$, which we already control, but also $\partial_t^2 v$, which we do not. So we estimate it first:

Lemma 9.3. *For solutions v to (9.3) we have*

$$\text{(9.7)} \qquad \begin{aligned} \|\partial_t^2 v(t)\|_{H^{-\frac{3}{2}}} &\lesssim \|v[t]\|_{\mathcal{H}^{\frac{1}{2}}} & n \geq 3, \\ \|\partial_t^2 v(t)\|_{H^{-\frac{11}{8}}} &\lesssim \|v[t]\|_{\mathcal{H}^{\frac{5}{8}}} & n = 2. \end{aligned}$$

Proof. We consider the case $n \geq 3$, and comment on the case $n = 2$ at the end. Using the equation (9.3) for v, we may write

$$\partial_t^2 v = h(\partial u)\partial_x \partial v + h(\partial u)\partial^2 u \partial v.$$

Here by Moser inequalities we have

$$\|h(\partial u)\|_{H^{s-1}} \ll 1.$$

Then, using also (9.4) the conclusion of the Lemma follows from the straightforward multiplicative estimates

$$H^{s-1} \cdot H^{-\frac{3}{2}} \to H^{-\frac{3}{2}}, \qquad H^{s-1} \cdot H^{s-2} \cdot H^{-\frac{1}{2}} \to H^{-\frac{3}{2}}.$$

where it is important that $s > \frac{n}{2} + 1$ and $s > \frac{5}{2}$. This last condition is not valid in dimension $n = 2$, where we only ask that $s > 2 + \frac{3}{8}$. This is why the Sobolev exponents in this case need to be increased by $\frac{1}{8}$. $\qquad \square$

We now return to the quadratic estimate for the source term in (9.6):

Lemma 9.4. *Let $v \in S_{AIT}^{\frac{1}{2}}$ satisfy (9.3). Then v also solves the inhomogeneous paradifferential equation*

$$\text{(9.8)} \qquad T_{\hat{P}}v = f,$$

with source term f satisfies the following bounds:

a) For $n \geq 3$ we have the uniform bound

easy3 (9.9)
$$\|f\|_{H^{-\frac{5}{4}}} \ll \|v[\cdot]\|_{L^\infty \mathcal{H}^{\frac{1}{2}}},$$

and the space-time bound

ong-F-est3 (9.10)
$$\|f\|_{L^p L^q} \ll \|v[\cdot]\|_{L^\infty \mathcal{H}^{\frac{1}{2}}}$$

with
$$\frac{1}{q} = \frac{1}{n - \frac{1}{2}} + \frac{1}{2}, \qquad \frac{1}{p} + \frac{1}{q} = 1.$$

b) For $n = 2$ we have the uniform bound

easy2 (9.11)
$$\|f\|_{H^{-1}} \ll \|v[\cdot]\|_{L^\infty \mathcal{H}^{\frac{5}{8}}},$$

and the space-time bound

ong-F-est2 (9.12)
$$\|f\|_{L^4 L^2} \ll \|v[\cdot]\|_{L^\infty \mathcal{H}^{\frac{5}{8}}}.$$

Proof. To avoid cluttering the notations, we prove the result in the case $n \geq 3$. The two dimensional case is identical up to appropriate adjustments of L^p exponents.

We write
$$-f = T_{\partial_\alpha \partial_\beta v} \hat{g}^{\alpha\beta} + T_{\partial_\beta v} \partial_\alpha \hat{g}^{\alpha\beta} + \Pi(\partial_\alpha \partial_\beta v, \hat{g}^{\alpha\beta}) + \Pi(\partial_\beta v, \partial_\alpha \hat{g}^{\alpha\beta}).$$

For the two terms where ∂_α has fallen on g, we have
$$\|T_{\partial_\beta v} \partial_\alpha \hat{g}^{\alpha\beta} + \Pi(\partial_\beta v, \partial_\alpha \hat{g}^{\alpha\beta})\|_{L^q} \lesssim \|v[t]\|_{\mathcal{H}^{\frac{1}{2}}} \|\langle D_x \rangle^{\frac{1}{2}} \partial_\alpha \hat{g}^{\alpha\beta}\|_{L^{n-\frac{1}{2}}} \lesssim_{\mathcal{A}} \|v[t]\|_{\mathcal{H}^{\frac{1}{2}}} \|\partial u\|_{W^{\frac{3}{2}, n - \frac{1}{2}}}.$$

Finally, in the cases where the ∂_α has fallen on v, we easily obtain the same estimate due to a good balance of derivatives. Here we use Lemma 9.3 to bound second derivatives of v. $\qquad \square$

9.2. **The normal form analysis.** The estimate in Lemma 9.4 is suboptimal as it does not recognize the cubic structure of the source. This is due to components of the source in which the linearized variable v is the second highest frequency, and which are not efficiently balanced with respect to derivatives. In fact, these cubic terms may heuristically be viewed as quadratic with a low frequency coefficient.

To better understand the source terms, we begin with a better description of the metric coefficients. By applying Lemma 2.9 to $g^{-\frac{1}{2}} g^{\alpha\beta}$ (see also (3.8)) and rearranging, we may write the paradifferential expansion

g-paradiff (9.13)
$$\hat{g}^{\alpha\beta}(\partial u) = \hat{g}^{\alpha\beta}(0) - T_{\hat{g}^{\alpha\beta} \partial^\gamma u + \hat{g}^{\alpha\gamma} \partial^\beta u + \hat{g}^{\beta\gamma} \partial^\alpha u} \partial_\gamma u + R(\partial u),$$

where R satisfies favourable balanced bounds, as in Lemma 5.7,

good-R (9.14)
$$\|\partial R\|_{L^n} \lesssim \mathcal{A}^{\sharp 2}, \qquad \|\partial R\|_{L^\infty} \lesssim \mathcal{B}^2.$$

To obtain a cubic estimate for (9.6), we substitute (9.13) in (9.6) and write

paralin2 (9.15)
$$T_{\hat{P}} v = N_1(u) + N_2(u),$$

where
$$N_1(u) = \partial_\alpha T_{\partial_\beta v} T_{\hat{g}^{\alpha\beta} \partial^\gamma u + \hat{g}^{\alpha\gamma} \partial^\beta u + \hat{g}^{\beta\gamma} \partial^\alpha u} \partial_\gamma u + \partial_\alpha \Pi(\partial_\beta v, T_{\hat{g}^{\alpha\beta} \partial^\gamma u + \hat{g}^{\alpha\gamma} \partial^\beta u + \hat{g}^{\beta\gamma} \partial^\alpha u} \partial_\gamma u)$$

102

consists of the essentially quadratic, nonperturbative components, while

$$N_2(u) = -\partial_\alpha T_{\partial_\beta v} R(\partial u) - \partial_\alpha \Pi(\partial_\beta v, R(\partial u))$$

consists of the balanced, directly perturbative components. We address the essentially quadratic components in $N_1(u)$ by passing to a renormalization $\tilde{v}$ of v,

$\boxed{\texttt{lin-nf}}$ (9.16) $\qquad \tilde{v} = v - T_{T_{\partial^\gamma u}\partial_\gamma v}u - T_{T_{\partial^\gamma u}v}\partial_\gamma u - \Pi(T_{\partial^\gamma u}\partial_\gamma v, u) - \Pi(T_{\partial^\gamma u}v, \partial_\gamma u) := v + v_2.$

This renormalization eliminates the components of the source where the linearized variable v is the second highest frequency. We thus replace $N_1(u)$ with a source consisting of terms with v only at the third highest frequency, and hence may be viewed as authentically cubic.

$\boxed{\texttt{radiff-eqn}}$ **Proposition 9.5.** *Let* $v \in S_{AIT}^{\frac{1}{2}}$ *be a solution for* ($\overset{\texttt{lin-inhom}}{9.3}$). *Then the following two properties hold in dimension* $n \geq 3$:

(i) *Equivalent norms:*

$\boxed{\texttt{bdd-nf}}$ (9.17) $\qquad\qquad\qquad\qquad\qquad \|\tilde{v}[t]\|_{\mathcal{H}^{\frac{1}{2}}} \approx \|v[t]\|_{\mathcal{H}^{\frac{1}{2}}}.$

(ii) $\tilde{v}$ *solves a good linear paradifferential equation of the form*

$\boxed{\texttt{radiff-eqn}}$ (9.18) $\qquad\qquad\qquad\qquad\qquad T_{\hat{P}}\tilde{v} = \partial_t f_1 + f_2,$

where the source terms are perturbative:

$\boxed{\texttt{rong-F-est}}$ (9.19) $\qquad \|f_2\|_{(S_{AIT}^{\frac{1}{2}})'} \ll \|v\|_{S_{AIT}^{\frac{1}{2}}}, \qquad \|f_1\|_{(S_{AIT}^{-\frac{1}{2}})'} \ll \|v[\cdot]\|_{L^\infty \mathcal{H}^{\frac{1}{2}}},$

as well as

$\boxed{\texttt{xtra-F-est}}$ (9.20) $\qquad\qquad\qquad\qquad \|f_1(t)\|_{S_{AIT}^{-\frac{1}{2}}} \lesssim \mathcal{A}^\sharp \|v[t]\|_{\mathcal{H}^{\frac{1}{2}}}.$

The same result holds in two space dimensions at the level of $v \in \mathcal{H}^{\frac{5}{8}}$.

Proof. (i) For the bound ($\overset{\texttt{bdd-nf}}{9.17}$), it suffices to estimate v_2 as follows:

$\boxed{\texttt{v2-nf}}$ (9.21) $\qquad\qquad\qquad\qquad \|v_2(t)\|_{H^{\frac{1}{2}}} \lesssim \mathcal{A}^\sharp \|v[t]\|_{\mathcal{H}^{\frac{1}{2}}},$

$\boxed{\texttt{dtv2-nf}}$ (9.22) $\qquad\qquad\qquad\qquad \|\partial_t v_2(t)\|_{H^{-\frac{1}{2}}} \lesssim \mathcal{A}^\sharp \|v[t]\|_{\mathcal{H}^{\frac{1}{2}}}.$

The first term of v_2 can be directly estimated using scale invariant $\mathcal{A}$ bounds,

$$\|T_{T_{\partial^\gamma u}\partial_\gamma v}u\|_{H^{\frac{1}{2}}} \lesssim \|T_{\partial^\gamma u}\partial_\gamma v\|_{H^{-\frac{1}{2}}}\|\partial u\|_{L^\infty} \lesssim \|\partial u\|_{L^\infty}\|\partial_\gamma v\|_{H^{-\frac{1}{2}}}\|\partial u\|_{L^\infty} \lesssim \mathcal{A}^2\|\partial_\gamma v\|_{H^{-\frac{1}{2}}}.$$

The third and the fourth terms are similar. However, for the second term we need to use the $\mathcal{A}^\sharp$ control norm combined with Bernstein's inequality:

$$\|T_{T_{\partial^\gamma u}v}\partial_\gamma u\|_{H^{\frac{1}{2}}} \lesssim \|T_{\partial^\gamma u}v\|_{H^{\frac{1}{2}}}\|\partial u\|_{W^{\frac{1}{2},2n}} \lesssim \|\partial u\|_{L^\infty}\|v\|_{H^{\frac{1}{2}}}\|\partial u\|_{W^{\frac{1}{2},2n}} \lesssim \mathcal{A}\mathcal{A}^\sharp\|v\|_{H^{\frac{1}{2}}}.$$

We next consider ($\overset{\texttt{dtv2-nf}}{9.22}$), where we distribute the time derivative, obtaining several types of terms:

a) Terms with distributed derivatives, namely $T_{T_{\partial u}\partial v}\partial u$ and $\Pi(T_{\partial u}\partial v, \partial u)$. We estimate the first by

$$\|T_{T_{\partial u}\partial v}\partial u\|_{H^{-\frac{1}{2}}} \lesssim \|T_{\partial u}\partial v\|_{H^{-\frac{1}{2}}}\|\partial u\|_{L^\infty} \lesssim \mathcal{A}^2\|\partial v\|_{H^{-\frac{1}{2}}},$$

and the second, using Sobolev embeddings, by

$$\|\Pi(T_{\partial u}\partial v, \partial u)\|_{H^{-\frac{1}{2}}} \lesssim \|\Pi(T_{\partial u}\partial v, \partial u)\|_{L^{\frac{2n}{n+1}}} \lesssim \|T_{\partial u}\partial v\|_{H^{-\frac{1}{2}}}\|\partial u\|_{W^{\frac{1}{2},2n}} \lesssim \mathcal{A}\mathcal{A}^\sharp\|\partial v\|_{H^{-\frac{1}{2}}}.$$

b) Terms with two derivatives on the high frequency u, namely $T_{T_{\partial_u v}}\partial^2 u$ and $\Pi(T_{\partial_u v}, \partial^2 u)$. In view of the bound (9.4), the corresponding estimate is nearly identical to case (a) above.

c) Terms with $\partial_t \partial^\gamma u$. Here we know that $\partial_t \partial^\gamma u \in H^{s-2}$, so we arrive at estimates which are also similar to case (a).

d) Terms with two derivatives on v. If one of them is spatial (i.e. $\gamma \neq 0$) then this is similar to or better than case (a). So we are left with the expressions $T_{T_{\partial_u \partial_t^2 v}}u$ and $\Pi(T_{\partial_u \partial_t^2 v}, u)$. But there we can use the bound (9.7) and complete the analysis again as in case (a).

(ii) The proof of (9.18) along with the estimates (9.19), (9.20) will be accomplished in four steps.

1) We first estimate the balanced source term component $N_2(u)$ from (9.15). We consider below the paraproduct T term, but the Π term is similar. We first consider the cases where the outer derivative $\partial_\alpha = \partial_i$ is a spatial derivative, which we will place in f_2. We have by Lemma 5.7 (see (9.14) above)

$$\|\partial_i T_{\partial_\beta v} R(\partial u)\|_{H^{-1/2}} \lesssim \|v[t]\|_{\mathcal{H}^{1/2}}\|\partial R(\partial u)\|_{L^\infty} \lesssim_{\mathcal{A}} \mathcal{B}^2\|v[t]\|_{\mathcal{H}^{1/2}},$$

where the $\mathcal{B}^2$ factor is integrable in time. We place the case where $\partial_\alpha = \partial_0$ in $\partial_t f_1$, estimating

$$\|T_{\partial_\beta v} R(\partial u)\|_{H^{1/2}} \lesssim \|v[t]\|_{\mathcal{H}^{1/2}}\|\partial R(\partial u)\|_{L^\infty} \lesssim_{\mathcal{A}} \mathcal{B}^2\|v[t]\|_{\mathcal{H}^{1/2}},$$

and

$$\|T_{\partial_\beta v} R(\partial u)\|_{H^{-1/2}} \lesssim \|T_{\partial_\beta v} R(\partial u)\|_{L^{\frac{2n}{1+n}}} \lesssim \|v[t]\|_{\mathcal{H}^{1/2}}\|\partial R(\partial u)\|_{L^n} \lesssim \mathcal{A}^{\sharp^2}\|v[t]\|_{\mathcal{H}^{1/2}},$$

as well as

$$\|\langle D_x \rangle^{-\frac{n}{2}-\frac{1}{4}} T_{\partial_\beta v} R(\partial u)\|_{L^\infty} \lesssim \|T_{\partial_\beta v} R(\partial u)\|_{L^2} \lesssim \|\partial v[t]\|_{\mathcal{H}^{-1/2}}\|\partial R(\partial u)\|_{L^{2n}} \lesssim \mathcal{A}^\sharp \mathcal{B}\|v[t]\|_{\mathcal{H}^{1/2}}.$$

Here we have a single $\mathcal{B}$ factor, which is L^2 in time, as needed for the $L^2 L^\infty$ Strichartz norm in (9.20).

2) Next, we apply product and commutator lemmas to exchange $N_1(u)$ for an equivalent expression up to perturbative errors, in preparation for comparison with the contribution from the normal form corrections. Here, we discuss the first term of $N_1(u)$,

(9.23)
$$\partial_\alpha T_{\partial_\beta v} T_{\hat{g}^{\alpha\beta}\partial^\gamma u}\partial_\gamma u,$$

but the remaining terms, including the balanced Π terms, are similar, using the analogous product and commutator lemmas. We first consider the cases where the outer derivative $\partial_\alpha = \partial_i$ is a spatial derivative, and place all perturbative errors in f_2. By an application of product and commutator Lemmas 2.7 and 2.4, we may replace (9.23) with

$$\partial_\alpha T_{\partial^\gamma u} T_{\partial_\beta v} T_{\hat{g}^{\alpha\beta}}\partial_\gamma u.$$

Then applying Lemma 2.4 and the estimate (2.13) in Lemma 2.8, it suffices to consider

$$\partial_\alpha T_{\hat{g}^{\alpha\beta}} T_{T_{\partial^\gamma u}\partial_\beta v}\partial_\gamma u.$$

In the case where $\partial_\alpha = \partial_0$, we place all perturbative errors in $\partial_t f_1$. The bound for f_1 in (9.19) is similar to the one for f_2, but there is a price to pay, namely that we also need to prove (9.20). Fortunately for (9.20) we may disregard all commutator structure and discard

104

all the para-coefficients, as they are bounded and gain an $\mathcal{A}^\sharp$ factor, so we are left with proving a bound of the form

$$\|T_{\partial v}\partial u\|_{S_{AIT}^{-\frac{1}{2}}} \lesssim \|\partial v\|_{L^\infty H^{-\frac{1}{2}}}.$$

Here for the uniform bound we simply write at fixed time

$$\|T_{\partial v}\partial u\|_{H^{-\frac{1}{2}}} \lesssim \|\partial v\|_{H^{-\frac{1}{2}}}\|\partial u\|_{L^\infty} \lesssim \mathcal{A}\|\partial v\|_{H^{-\frac{1}{2}}},$$

and for the $L^2 L^\infty$ bound we have

$$\|T_{\partial v}\partial u\|_{L^2} \lesssim \mathcal{B}\|\partial v\|_{H^{-\frac{1}{2}}}$$

using $\mathcal{B}$ for the square integrability in time and then applying Bernstein's inequality in space to convert the L^2 bound into L^∞.

Applying the same analysis to the other terms of $N_1(u)$, we have reduced the problem to

$$N_1'(u) = \partial_\alpha(T_{\hat{g}^{\alpha\beta}}T_{T_{\partial^\gamma u}\partial_\beta v}\partial_\gamma u + T_{\hat{g}^{\alpha\gamma}}T_{T_{\partial^\beta u}\partial_\beta v}\partial_\gamma u + T_{\hat{g}^{\beta\gamma}}T_{T_{\partial^\alpha u}\partial_\beta v}\partial_\gamma u)$$
$$+ \partial_\alpha(\Pi(T_{\partial^\gamma u}\partial_\beta v, T_{\hat{g}^{\alpha\beta}}\partial_\gamma u) + \Pi(T_{\partial^\beta u}\partial_\beta v, T_{\hat{g}^{\alpha\gamma}}\partial_\gamma u) + \Pi(T_{\partial^\alpha u}\partial_\beta v, T_{\hat{g}^{\beta\gamma}}\partial_\gamma u)).$$

3) We next establish the cancellation between the normal form correction and $N_1'(u)$. In this step, we discuss only the low-high T paraproduct contributions, and return to the Π contributions in Step 4. Applying $T_{\hat{P}}$ to the T term of v_2 in (9.16), we have the contribution

$$-\partial_\alpha T_{\hat{g}^{\alpha\beta}}\partial_\beta(T_{T_{\partial^\gamma u}\partial_\gamma v}u + T_{T_{\partial^\gamma u}v}\partial_\gamma u). \tag{9.24}$$

a) We first would like to observe that the cases where the derivatives ∂_β and ∂_γ are split, between v and the high frequency u, cancel with the first two terms of $N_1'(u)$. The main task to verify before doing so is that the cases where the ∂_β falls on the lowest frequency para-coefficient $\partial^\gamma u$ are perturbative due to an efficient balance of derivatives, and may be absorbed into f_2 or f_1. To see this, we analyze separately cases involving spatial versus time derivatives. In the case of spatial derivatives $\partial_\alpha = \partial_i$ and $\partial_\beta = \partial_j$, we directly estimate

$$\|\partial_i T_{\hat{g}^{ij}}T_{T_{\partial_j\partial^\gamma u}\partial_\gamma v}u\|_{H^{-1/2}} \lesssim \mathcal{B}^2\|v[t]\|_{\mathcal{H}^{1/2}}.$$

In the case where $\partial_\beta = \partial_0$, we obtain the same estimate in the same manner, except when $\partial_\gamma = \partial_0$. In this case, we may use Lemma 5.4 to estimate the lowest frequency $\partial_0^2 u$.

It remains to consider the case $\partial_\alpha = \partial_0$, which we place in $\partial_t f_1$. We have

$$\|T_{\hat{g}^{0\beta}}T_{T_{\partial_\beta\partial^\gamma u}\partial_\gamma v}u\|_{H^{1/2}} \lesssim \mathcal{B}^2\|v[t]\|_{\mathcal{H}^{1/2}}$$

as before. For f_1 however, we also require an estimate for the full Strichartz norm in (9.20). We separate ∂_β again into spatial and time derivatives. For the spatial case, we have by Sobolev embeddings,

$$\|T_{\hat{g}^{0j}}T_{T_{\partial_j\partial^\gamma u}\partial_\gamma v}u\|_{H^{-1/2}} \lesssim \|T_{T_{\partial_j\partial^\gamma u}\partial_\gamma v}u\|_{L^{\frac{2n}{1+n}}} \lesssim \|\langle D_x\rangle^{1/2}\partial^\gamma u\|_{L^{2n}}\|v[t]\|_{\mathcal{H}^{1/2}}\|\partial u\|_{L^\infty}$$
$$\lesssim \mathcal{A}\mathcal{A}^\sharp\|v[t]\|_{\mathcal{H}^{1/2}}$$

for the uniform bound, as well as

$$\|T_{\hat{g}^{0j}}T_{T_{\partial_j\partial^\gamma u}\partial_\gamma v}u\|_{L^2} \lesssim \|\langle D_x\rangle^{1/2}\partial^\gamma u\|_{L^{2n}}\|v[t]\|_{\mathcal{H}^{1/2}}\|\partial u\|_{BMO^{\frac{1}{2}}} \lesssim \mathcal{B}\mathcal{A}^\sharp\|v[t]\|_{\mathcal{H}^{1/2}}$$

for the $L^2 L^\infty$ bound.

For the case $\partial_\beta = \partial_0$, the lowest frequency includes an instance of $\partial_0^2 u$, where we apply Lemma 5.4. This contributes a spatial component $\hat\partial_t^2 u$ which is estimated as before, as well as a balanced Π interaction, namely $\pi_2(u)$. This case is estimated by

$$\|T_{\hat g^{00}} T_{T_{\pi_2(u)}\partial_0 v} u\|_{H^{-1/2}} \lesssim \|T_{T_{\pi_2(u)}\partial_0 v} u\|_{L^{\frac{2n}{1+n}}} \lesssim \|\pi_2(u)\|_{L^n}\|v[t]\|_{\mathcal{H}^{1/2}}\|\partial u\|_{L^\infty} \lesssim \mathcal{A}\mathcal{A}^{\sharp 2}\|v[t]\|_{\mathcal{H}^{1/2}}$$

for the energy norm respectively

$$\|T_{\hat g^{00}} T_{T_{\pi_2(u)}\partial_0 v} u\|_{L^2} \lesssim \|T_{T_{\pi_2(u)}\partial_0 v} u\|_{L^2} \lesssim \|\pi_2(u)\|_{L^{2n}}\|v[t]\|_{\mathcal{H}^{1/2}}\|\partial u\|_{L^\infty} \lesssim \mathcal{A}\mathcal{A}^{\sharp}\mathcal{B}\|v[t]\|_{\mathcal{H}^{1/2}}$$

for the $L^2 L^\infty$ bound.

Having dismissed the perturbative cases via this analysis, we observe an exact cancellation with the first two terms of $N_1'(u)$. Collecting the remaining paraproduct terms from $N_1'(u)$ and (9.24), we are left with the expression

Tsubstep1 (9.25) $$\partial_\alpha T_{\hat g^{\beta\gamma}} T_{T_{\partial^\alpha u}\partial_\beta v}\partial_\gamma u - \partial_\alpha T_{\hat g^{\alpha\beta}} T_{T_{\partial^\gamma u}\partial_\beta\partial_\gamma v}u - \partial_\alpha T_{\hat g^{\alpha\beta}} T_{T_{\partial^\gamma u}v}\partial_\beta\partial_\gamma u.$$

b) Before proceeding, we further process the first term in (9.25), with the key step being an integration by parts which reveals an instance of $T_{\hat P}v$. Reindexing, we rewrite this term as

$$\partial_\gamma T_{\hat g^{\beta\alpha}} T_{T_{\partial^\gamma u}\partial_\beta v}\partial_\alpha u.$$

Then applying Lemma 2.4 and the estimate (2.13) in Lemma 2.8 to commute $T_{\hat g^{\alpha\beta}}$, similar to step 2), we replace this by

$$\partial_\gamma T_{T_{\partial^\gamma u}T_{\hat g^{\beta\alpha}}\partial_\beta v}\partial_\alpha u.$$

Simulating an integration by parts with respect to ∂_α, we write this as

$$\partial_\alpha\partial_\gamma T_{T_{\partial^\gamma u}T_{\hat g^{\beta\alpha}}\partial_\beta v}u - \partial_\gamma T_{\partial_\alpha T_{\partial^\gamma u}T_{\hat g^{\beta\alpha}}\partial_\beta v}u.$$

We will carry the first of these terms forward to 3c), while the latter term is perturbative. To see this, we observe that ∂_α may commute through $T_{\partial^\gamma u}$, similar to the analysis in 3a). Thus we arrive at the expression

$$\partial_\gamma T_{T_{\partial^\gamma u}\partial_\alpha T_{\hat g^{\beta\alpha}}\partial_\beta v}u = \partial_\gamma T_{T_{\partial^\gamma u}T_{\hat P}v}u.$$

We consider separately via f_2 and $\partial_t f_1$ the contributions corresponding to $\partial_\gamma = \partial_i$ and $\partial_\gamma = \partial_0$ respectively. For the bound (9.19), using Lemma 9.4 and the Strichartz exponents (p_1, q_1) given by

p1q1 (9.26) $$\frac{1}{p_1} = \frac{1}{n - \frac{1}{2}}, \qquad \frac{1}{p_1} + \frac{1}{q_1} = \frac{1}{2},$$

we estimate, in both cases,

$$\|T_{T_{\partial^\gamma u}T_{\hat P}v}u\|_{(S_{AIT}^{-\frac{1}{2}})'} \lesssim \|\langle D_x\rangle^{\frac{3}{2}+\delta}T_{T_{\partial^\gamma u}T_{\hat P}v}u\|_{L^{p_1'}L^{q_1'}}$$

$$\lesssim \mathcal{A}\|T_{\hat P}v\|_{L^p L^q}\|u\|_{L^2 W^{\frac{3}{2}+\delta,\infty}}.$$

It remains to prove the bound (9.20), but this is again a simpler bound where we have a considerable gain. Indeed, using only H^s Sobolev bounds but including (9.4) and (9.7) we obtain at fixed time

$$\|T_{T_{\partial^\gamma u}T_{\hat P}v}u\|_{L^2} \lesssim \mathcal{A}\|T_{\hat P}v\|_{H^{-\frac{5}{4}}}\|u\|_{H^s},$$

which suffices for all the Strichartz bounds.

c) Returning to (9.25) and replacing the first term via the analysis in 3b), we are now left with

$$\partial_\alpha(\partial_\gamma T_{T_{\partial^\gamma u}T_{\hat{g}^{\beta\alpha}}\partial_\beta v}u - T_{\hat{g}^{\alpha\beta}}T_{T_{\partial^\gamma u}\partial_\beta\partial_\gamma v}u - T_{\hat{g}^{\alpha\beta}}T_{T_{\partial^\gamma u}v}\partial_\beta\partial_\gamma u).$$

We observe a cancellation between the first two terms. To see this, we apply the Leibniz rule for the ∂_γ derivative on the first term. Similar to 3a), cases where the derivative falls on the lowest frequency $\partial^\gamma u$ or $\hat{g}^{\beta\alpha}$ are perturbative. We also have a term which cancels the second term, leaving us with

$$\partial_\alpha(T_{T_{\partial^\gamma u}T_{\hat{g}^{\beta\alpha}}\partial_\beta v}\partial_\gamma u - T_{\hat{g}^{\alpha\beta}}T_{T_{\partial^\gamma u}v}\partial_\beta\partial_\gamma u).$$

Applying also the commutator Lemma 2.4 and the bound (2.13) in Lemma 2.8 as in 2), we rewrite this as

$$\partial_\alpha T_{\partial^\gamma u}T_{T_{\hat{g}^{\beta\alpha}}\partial_\beta v}\partial_\gamma u - \partial_\alpha T_{T_{\partial^\gamma u}v}T_{\hat{g}^{\alpha\beta}}\partial_\beta\partial_\gamma u.$$

d) We apply the Leibniz rule with respect to ∂_α. Here we observe that cases where ∂_α falls on lower frequency instances of u or g are perturbative. Note that in contrast to the previous substeps, we no longer have the ∂_α divergence and so we must put all terms in f_2.

We consider for instance the term

$$T_{T_{\partial_\alpha\partial^\gamma u}v}T_{\hat{g}^{\alpha\beta}}\partial_\beta\partial_\gamma u.$$

Excluding the case of two time derivatives in $\partial_\alpha\partial^\gamma u$, this is easily estimated due to a favorable balance of derivatives. In the case of two time derivatives, we have $\partial_\alpha\partial^\gamma u \in \mathfrak{DC}$ so we can use the decomposition in Definition 5.1, say $\partial_\alpha\partial^\gamma u = h_1 + h_2$. The first component can be thought of as a spatial derivative and is again easily estimated. It remains to consider the contribution of the second term $h_2 \in \mathcal{B}^2 L^\infty$:

$$\|T_{T_{h_2}v}T_{\hat{g}^{\alpha\beta}}\partial_\beta\partial_\gamma u\|_{H^{-1/2}} \lesssim \mathcal{A}\|h_2\|_{L^\infty}\|v\|_{H^{\frac{1}{2}}}\|\partial^2 u\|_{H^{s-2}}$$

$$\lesssim \mathcal{B}^2\|v[t]\|_{\mathcal{H}^{1/2}}.$$

A similar analysis applies in the cases where ∂^α falls on a low frequency metric coefficient g.

e) We record the remaining terms after applying the Leibniz rule, and will observe instances of $T_{\hat{P}}$ for which we use the equation (9.6), as well as a cancellation. We arrive at

$$T_{\partial^\gamma u}(T_{\partial_\alpha T_{\hat{g}^{\beta\alpha}}\partial_\beta v}\partial_\gamma u + T_{T_{\hat{g}^{\beta\alpha}}\partial_\beta v}\partial_\alpha\partial_\gamma u) - T_{T_{\partial^\gamma u}\partial_\alpha v}T_{\hat{g}^{\alpha\beta}}\partial_\beta\partial_\gamma u - T_{T_{\partial^\gamma u}v}\partial_\alpha T_{\hat{g}^{\alpha\beta}}\partial_\beta\partial_\gamma u.$$

Reindexing the second term, and applying the bound (2.13) in Lemma 2.8 in the second and the third term, the above expression may be reduced to the form

(9.27) $\qquad T_{\partial^\gamma u}T_{T_{\hat{P}}v}\partial_\gamma u + T_{\partial^\gamma u}T_{\hat{g}^{\alpha\beta}}T_{\partial_\alpha v}\partial_\beta\partial_\gamma u - T_{\partial^\gamma u}T_{\partial_\alpha v}T_{\hat{g}^{\alpha\beta}}\partial_\beta\partial_\gamma u - T_{T_{\partial^\gamma u}v}\partial_\alpha T_{\hat{g}^{\alpha\beta}}\partial_\beta\partial_\gamma u.$

Now in the two middle terms we have a commutator structure, which can be estimated directly by Lemma 2.4. It remains to consider the first and the last term. We apply (9.10) to the first term, and estimate in a dual Strichartz norm with (p_1, q_1) as in (9.26),

$$\|T_{T_{\hat{P}}v}\partial_\gamma u\|_{(S_{AIT}^{\frac{1}{2}})'} \lesssim \|\langle D_x\rangle^{\frac{1}{2}+\delta}T_{T_{\hat{P}}v}\partial_\gamma u\|_{L^{p_1'}L^{q_1'}} \lesssim \mathcal{A}\|T_{\hat{P}}v\|_{L^p L^q}\|\partial u\|_{L^2 W^{\frac{1}{2}+\delta,\infty}}.$$

For the last term, on the other hand, we use a Strichartz bound for v and match it with the bound (9.5) in Lemma 9.2,

$$\|\langle D_x\rangle^\delta T_{T_{\partial^\gamma u}v}\partial_\alpha T_{\hat{g}^{\alpha\beta}}\partial_\beta\partial_\gamma u\|_{L^{p_3'}L^{q_3'}} \lesssim \|\langle D_x\rangle^{-\delta}v\|_{L^{p_3'}L^{q_3'}}\|\langle D_x\rangle^{\delta_0}\partial_\alpha T_{\hat{g}^{\alpha\beta}}\partial_\beta\partial_\gamma u\|_{L^{p_4}L^{q_4}} \ll \|v\|_{S_{AIT}^{\frac{1}{2}}},$$

107

where

$$\frac{1}{p_3} = \frac{1}{n - \frac{1}{2}}, \qquad \frac{1}{p_4} = 1 - \frac{2}{p_3}, \qquad \frac{1}{p_3} + \frac{1}{q_3} = \frac{1}{2}, \qquad \frac{1}{p_4} + \frac{1}{q_4} = 1.$$

Here the Strichartz exponents p_3 and q_3 are chosen so that the first factor on the right is controlled by $\|v\|_{S_{AIT}^{\frac{1}{2}}}$ and δ is arbitrarily small. On the other hand δ_0 is a fixed positive parameter which depends only on the distance between s and its lower bound.

4) It remains to consider the cancellation between the balanced Π terms in the normal form correction and in $N_1'(u)$. Here the analysis is identical to the analysis for the low-high T contributions in Step 3, due to the analogous structure for the T and Π terms in both v_2 and $N_1'(u)$. The main care that is needed is to observe that all negative Sobolev exponent norms have been addressed in Step 3 by either using a divergence structure, or by Sobolev embeddings, which apply equally well to the balanced Π case.

$\square$

9.3. Reduction to the paradifferential equation.
Here we first use the well-posedness result for the linear paradifferential equation in order to obtain a good bound for $\tilde{v}$. The source terms are perturbative by (9.19) and Theorem 4.12, so the solution $\tilde{v}$ must satisfy the bound

$$\boxed{\texttt{tva}} \qquad (9.28) \qquad \|\tilde{v}\|_{S_{AIT}^{\frac{1}{2}}} + \|\partial_t \tilde{v}\|_{S_{AIT}^{-\frac{1}{2}}} \lesssim \|\tilde{v}[0]\|_{\mathcal{H}^{\frac{1}{2}}} + c\|v\|_{S_{AIT}^{\frac{1}{2}}} + \|\partial_t v\|_{S_{AIT}^{-\frac{1}{2}}}, \qquad c \ll 1.$$

It remains to show that the Strichartz estimates carry over to v. For this, it suffices to show that

$$\boxed{\texttt{tvb}} \qquad (9.29) \qquad \|v_2\|_{S_{AIT}^{\frac{1}{2}}} + \|\partial_t v_2\|_{S_{AIT}^{-\frac{1}{2}}} \ll \|v\|_{L^\infty \mathcal{H}^{\frac{1}{2}}}.$$

If this is true, then combining the last two bounds with the norm equivalence (9.17) we obtain the desired bound for the linearized evolution (9.3), namely

$$\boxed{\texttt{v-Str}} \qquad (9.30) \qquad \|v\|_{S_{AIT}^{\frac{1}{2}}} + \|\partial_t v\|_{S_{AIT}^{-\frac{1}{2}}} \ll \|v[0]\|_{\mathcal{H}^{\frac{1}{2}}}$$

with a universal implicit constant. This concludes the proof of Theorem 9.1 in dimension $n \geq 3$. The case $n = 2$ is virtually identical.

It remains to prove (9.29). The energy norm for v_2 has already been estimated in part (i) of Proposition 9.5, so it remains to consider the $L^2 L^\infty$ norm in three and higher dimensions. This is a soft bound, where we only need to use the energy bound for v on the right, and not the full Strichartz norm, as we would also have been allowed. There are eight norms to estimate; most of them are similar, so we consider a representative sample, leaving the rest for the reader.

For a streamlined unbalanced bound we consider the term

$$\|\langle D_x \rangle^{-\frac{n-2}{2} - \frac{1}{4} - \delta} T_{T_{\partial^\gamma u} v} \partial_\gamma u\|_{L^2 L^\infty} \lesssim \|T_{\partial^\gamma u} v\|_{L^\infty L^{\frac{2n}{n-1}}} \|\langle D_x \rangle^{\frac{1}{4} - \delta} \partial u\|_{L^\infty} \lesssim \mathcal{A} \|v\|_{L^\infty \mathcal{H}^{\frac{1}{2}}},$$

where we have used Bernstein's inequality twice and the Strichartz bound for u. This pattern is followed for all unbalanced terms.

For the worst balanced case, we apply the time derivative to v in the next to last term in v_2. Then we have to estimate

$$\|\langle D_x \rangle^{-\frac{n}{2}-\frac{1}{4}-\delta} \Pi(T_{\partial^\gamma u} \partial^2 v, u)\|_{L^2 L^\infty} \lesssim \|\Pi(T_{\partial^\gamma u} \partial^2 v, u)\|_{L^2 L^{\frac{4n}{2n-1}}}$$

$$\lesssim \mathcal{A}\|\partial^2 v\|_{L^\infty H^{-\frac{3}{2}}} \|\langle D_x \rangle^{\frac{3}{2}+\frac{1}{4}} u\|_{L^2 L^\infty} \lesssim \mathcal{A}\|v\|_{L^\infty \mathcal{H}^{\frac{1}{2}}},$$

where we have used Bernstein's inequality twice, Lemma 9.3 and the Strichartz bound for u.

10. Short time Strichartz estimates

s:ST

The aim of this section is to provide a more detailed overview of the local well-posedness result in [38], and at the same time to provide a formulation of this result which applies in a large data setting, but for a short time. Instead of working with the equation (1.13), here it is easier to work with the problem

ST-eqn (10.1)
$$g^{\alpha\beta}(\mathbf{u})\partial_\alpha \partial_\beta \mathbf{u} = h^{\alpha\beta}(\mathbf{u})\partial_\alpha \mathbf{u}\, \partial_\beta \mathbf{u}$$

for a possibly vector-valued function $\mathbf{u}$. This is exactly the set-up of [38], and has the advantage that it is scale invariant. We recall that the scaling exponent for this problem is $s_c = \frac{n}{2}$. In our problem, we will apply the results in this section to the function $\mathbf{u} = \partial u$.

We begin with a review of the local well-posedness result in [38], but where we describe also the structure of the Strichartz estimates:

t:ST **Theorem 10.1** (Smith-Tataru [38]). *Consider the problem* (10.1) *with initial data satisfying*

reg-data (10.2)
$$\|\mathbf{u}[0]\|_{\mathcal{H}^{s_1}} \ll 1.$$

where

reg-s2 (10.3)
$$s_1 > s_c + \frac{3}{4}, \qquad n = 2,$$

respectively

reg-s3 (10.4)
$$s_1 > s_c + \frac{1}{2}, \qquad n = 3, 4, 5.$$

Then the solution exists on the time interval $[0, 1]$, *and satisfies the following Strichartz estimates*

se2 (10.5)
$$\|\langle D_x \rangle^{\delta_0} \partial \mathbf{u}\|_{L^4 L^\infty} \lesssim 1, \qquad n = 2,$$

respectively

se3 (10.6)
$$\|\langle D_x \rangle^{\delta_0} \partial \mathbf{u}\|_{L^2 L^\infty} \lesssim 1, \qquad n = 3, 4, 5,$$

with a small $\delta_0 > 0$.

In addition, another conclusion of the work in [38], which is used as an intermediate step in the proof of the theorem above, is that the linearized problem around the solutions in Theorem 10.1 is well-posed in a range of Sobolev spaces, and almost lossless Strichartz estimates hold for them. Precisely, we have the following:

t:ST-lin | **Theorem 10.2** ([38]). *Let $\mathbf{u}$ be a solution for (10.1) in the time interval $[0,1]$ as in Theorem 10.1. Then the linear equation*

box-g-gen | (10.7)
$$\begin{cases} g^{\alpha\beta}(\mathbf{u})\partial_\alpha\partial_\beta v = 0 \\ v[0] = (v_0, v_1) \end{cases}$$

is well-posed in $\mathcal{H}^r$ in the same time interval for $1 \leq r \leq s_1 + 1$, and the solutions satisfy the uniform and Strichartz estimates (4.33) for the same range of r.

We note that in [38] it is also assumed that $g^{00} = -1$, akin to our metric $\tilde{g}$; but it is clear that such an assumption is not needed in the above theorems, as one can simply divide the equation by g^{00}.

We also remark that the equation (10.7) is not the same as the linearized equation. The reason (10.7) is preferred in [38] is the extended upper bound for r. It is also noted in [38] that for a range of r with a lower upper bound, the conclusion of the last theorem is also valid for the full linearized equation; this is a straightforward perturbative argument. From below, the Sobolev exponent $r = 1$ suffices in dimension $n \geq 3$ in [38], though it is also clear that this is not optimal. Indeed, for dimension $n = 2$ the above result is extended in [38] to the range $\frac{3}{4} \leq r \leq s_1 + 1$, and the linearized equation is shown to be well-posed in $\mathcal{H}^{\frac{3}{4}}$; see [38, Lemma A4]; the same method also works in higher dimension.

We also remark that if the linearized equation is in divergence form, (which can be arranged in the present paper, see (3.24)), then, by duality, (forward/backward) well-posedness in $\mathcal{H}^r$ implies (backward/forward) well-posedness in $\mathcal{H}^{1-r}$, with the center point at $r = \frac{1}{2}$. This motivates why, in the context of the present paper, it is easiest to study the linearized equation exactly in $\mathcal{H}^{\frac{1}{2}}$. Unfortunately our argument runs into a technical obstruction in dimension $n = 2$, which is why we make a slight adjustment there and work instead in $\mathcal{H}^{\frac{5}{8}}$.

To summarize, in the present paper we will not need directly the conclusion of Theorem 10.2, but rather a minor variation of it where we also consider the divergence form equation and its associated paradifferential flow, and we lower the range for r in order to include the space $\mathcal{H}^{\frac{1}{2}}$ ($\mathcal{H}^{\frac{5}{8}}$ in dimension two).

In the proof of the main result of this paper, we will need to use this result for solutions that are not small in $\mathcal{H}^s$, so we cannot apply it directly. Instead, we will seek to rephrase it and use it in a large data setting via a scaling argument.

The difficulty we face is that rescaling keeps homogeneous Sobolev norms unchanged, rather than the inhomogeneous ones. A first step in this direction is to consider smooth solutions, but which may be large at low frequency:

t:smooth | **Theorem 10.3.** *Consider the problem (10.1) with initial data satisfying*

g-data-hom | (10.8)
$$\|\mathbf{u}[0]\|_{\dot{\mathcal{H}}^N \cap \dot{\mathcal{H}}^{s_c}} + \|\mathbf{u}(0)\|_{L^\infty} \ll 1.$$

Then the solution exists up to time 1, and satisfies the uniform bound

reg-point | (10.9)
$$\|\mathbf{u}\|_{L^\infty} \ll 1.$$

and the Sobolev bound

eg-sobolev | (10.10)
$$\|\mathbf{u}\|_{L^\infty([0,1];\dot{\mathcal{H}}^N \cap \dot{\mathcal{H}}^{s_c})} \lesssim \|\mathbf{u}[0]\|_{\dot{\mathcal{H}}^N + \dot{\mathcal{H}}^{s_c}}.$$

In addition,

(10.11)
$$\|\mathbf{u}[\cdot]\|_{L^\infty([0,1];\mathcal{H}^1)} \lesssim \|\mathbf{u}[0]\|_{\mathcal{H}^1}$$

whenever the right hand side is finite.

Proof. Locally, after subtracting a constant, the data is small in $\mathcal{H}^N$ so the existence of regular solutions is classical. It remains to establish energy estimates in homogeneous Sobolev norms. The problem reduces to the case of the paradifferential flow, and, by conjugation with a power of $\langle D_x \rangle$, to bounds in $\mathcal{H}^1$ that are straightforward. $\qquad \square$

A second step is the following variation of Theorem 10.1, where we consider a small $\mathcal{H}^{s_1}$ perturbation of a small and smooth data:

Theorem 10.4. *Consider the problem (10.1) with initial data $\mathbf{u}[0]$ of the form*

$$\mathbf{u}[0] = \mathbf{u}^{lo}[0] + \mathbf{u}^{hi}[0], \tag{10.12}$$

where the two components satisfy

$$\|\mathbf{u}^{lo}[0]\|_{\mathcal{H}^N} \ll 1, \qquad \|\mathbf{u}^{hi}[0]\|_{\mathcal{H}^{s_1}} \leq \epsilon \ll 1. \tag{10.13}$$

Then the solution u exists on the time interval $[0,1]$, and satisfies the following Strichartz estimates

$$\|\langle D_x \rangle^{\delta_0} \partial(\mathbf{u} - \mathbf{u}^{lo})\|_{L^4 L^\infty} \lesssim \epsilon, \qquad n = 2 \tag{10.14}$$

respectively

$$\|\langle D_x \rangle^{\delta_0} \partial(\mathbf{u} - \mathbf{u}^{lo})\|_{L^2 L^\infty} \lesssim \epsilon, \qquad n \geq 3. \tag{10.15}$$

with a small $\delta_0 > 0$.

We remark that the solutions in this second theorem are still covered by Theorem 10.1. The only difference is that the constant in the Strichartz bound depends only on the $\mathbf{u}^{hi}[0]$ bound.

Proof. This follows by a direct application of the results in Theorem 10.1 and Theorem 10.2. We write an equation for $\mathbf{u}^{hi} = \mathbf{u} - \mathbf{u}^{lo}$,

$$g^{\alpha\beta}(\mathbf{u})\partial_\alpha \partial_\beta \mathbf{u}^{hi} = -(g^{\alpha\beta}(\mathbf{u}) - g^{\alpha\beta}(\mathbf{u}^{lo}))\partial_\alpha \partial_\beta \mathbf{u}_{lo} := \mathbf{f}^{hi},$$

where the source term $\mathbf{f}^{hi}$ can be estimated at fixed time by

$$\|\mathbf{f}^{hi}\|_{H^{s_1-1}} \lesssim \|\mathbf{u}^{hi}\|_{H^{s_1}},$$

and thus it is perturbative. Then we apply the Strichartz estimates in Theorem 10.2 to $\mathbf{u}^{hi}$, and the desired conclusion follows.

$\qquad \square$

Now we consider the large data problem, where we show local well-posedness by a scaling argument. The price to pay will be that the time interval for which we have the solutions will be shorter. Precisely, we will show that

Theorem 10.5. *For any s_1 as in (10.3), (10.4) there exists $\delta_0 > 0$ so that the following holds: For any $M > 0$ and any solution $\mathbf{u}$ to the problem (1.13) with initial data satisfying*

$$\|\mathbf{u}[0]\|_{\dot{\mathcal{H}}^{s_1}} \ll M, \qquad \|\mathbf{u}[0]\|_{\dot{\mathcal{H}}^{s_c}} \ll 1. \tag{10.16}$$

we have:

a) The solution exists up to time T_M given by

$$T_M^\sigma = M^{-1}, \qquad \sigma = s_1 - s_c, \tag{10.17}$$

with uniform bounds

(10.18) $$\|\mathbf{u}[\cdot]\|_{C([0,T_M];\mathcal{H}^{s_1})} \lesssim M, \qquad \|\mathbf{u}[\cdot]\|_{C([0,T_M];\dot{\mathcal{H}}^{sc})} \lesssim 1,$$

as well as

(10.19) $$\|\mathbf{u}[\cdot]\|_{C([0,T_M];\mathcal{H}^1)} \lesssim \|\mathbf{u}[0]\|_{\mathcal{H}^1},$$

whenever the right hand side is finite.

    b) The solution $\mathbf{u}$ *satisfies the following Strichartz estimates in* $[0, T_M]$:

(10.20) $$\|\langle T_M D'\rangle^{\delta_0} \partial \mathbf{u}\|_{L^4 L^\infty} \lesssim T_M^{-\frac{3}{4}}, \qquad n = 2,$$

respectively

(10.21) $$\|\langle T_M D'\rangle^{\delta_0} \partial \mathbf{u}\|_{L^2 L^\infty} \lesssim T_M^{-\frac{1}{2}}, \qquad n \geq 3.$$

    c) Furthermore, the homogeneous Strichartz estimates (4.33) *also hold in* $\mathcal{H}^r$ *for the associated linear equations* (10.7), *on the same time intervals for* $r \in [1, s_1]$.

Proof. As stated, the result is invariant with respect to scaling. Precisely M plays the role of a scaling parameter, and by scaling we can set it to 1.

    It remains to prove the result for $M = 1$ in which case $T_M = 1$. In a nutshell, the idea of the proof is to use the finite proof of propagation to localize the problem and, by scaling, to reduce it to the case when Theorems 10.3, 10.4 can be applied. To fix the notations, we will consider the case $n \geq 3$ in what follows; the two dimensional case is identical after obvious changes in notations.

    On the Fourier side we split the initial data into two components,

$$\mathbf{u}[0] = \mathbf{u}^{lo}[0] + \mathbf{u}^{hi}[0], \qquad \mathbf{u}^{lo}[0] = P_{<0}\mathbf{u}[0],$$

and we denote by $\mathbf{u}$ and $\mathbf{u}^{lo}$ the corresponding solutions.

    On the other hand on the physical side we partition the initial time slice $t = 0$ into cubes Q of size 1, and consider a partition of unity associated to the covering by $8Q$,

$$1 = \sum \chi_Q,$$

and define the localized initial data

$$\mathbf{u}_Q[0] = (\chi_Q(\mathbf{u}_0 - \bar{\mathbf{u}}_{0,Q}^{lo}), \chi_Q u_1), \qquad \bar{\mathbf{u}}_{0,Q}^{lo} = \fint_Q \mathbf{u}_0^{lo}\, dx,$$

which agrees with $\mathbf{u}[0]$ in $6Q$ up to a constant. The speed of propagation for solutions $\mathbf{u}$ with $|\mathbf{u}| \ll 1$ is close to 1, therefore the corresponding solutions w_Q agree with $\mathbf{u}$ in $4Q$ (again, up to a constant) in $[0, 1]$, assuming both exist up to this time.

    Next we consider the existence and properties of the solutions $\mathbf{u}_Q$ in the time interval $[0, T_M]$. For $\mathbf{u}_Q[0]$ we have a low-high decomposition,

$$\mathbf{u}_Q[0] = (\chi_{T_M^{-1}Q}(\mathbf{u}_0^{lo} - \bar{\mathbf{u}}_{0,Q}^{lo}), \chi_{T_M^{-1}Q} u_1^{lo}) + \chi_{T_M^{-1}Q}(\mathbf{u}_0^{hi}, Q u_1^{hi}) := \mathbf{u}_Q^{lo}[0] + \mathbf{u}_Q^{hi}[0]$$

    Now we consider energy bounds for the initial data. For $\mathbf{u}^{lo}$ we have

(10.22) $$\|\mathbf{u}^{lo}[0]\|_{\dot{H}^{sc} \cap \dot{\mathcal{H}}^N} \ll 1.$$

Since $s_1 - 1 < n/2$, after localization this also implies that the low frequency components satisfy

$$\|\mathbf{u}_Q^{lo}[0]\|_{\mathcal{H}^N} \ll 1, \tag{10.23}$$

which is exactly as in Theorem 10.3, respectively Theorem 10.4.

On the other hand, for the high frequency bounds we have the almost orthogonality relation

$$\sum_Q \|\mathbf{u}_Q[0]\|_{\mathcal{H}^{s_1}}^2 \ll 1. \tag{10.24}$$

By Theorem 10.1, it follows that the solutions $\mathbf{u}_Q$ exist up to time 1, and satisfy the Strichartz bounds

$$\|\langle D_x\rangle^{\delta_0}\partial\mathbf{u}_Q\|_{L^2 L^\infty} \lesssim 1. \tag{10.25}$$

Theorem 10.4 allows us to improve this to

$$\|\mathbf{u}_Q - \mathbf{u}_Q^{lo}\|_{L^\infty \mathcal{H}^{s_1}} + \|\langle D_x\rangle^{\delta_0}\partial(\mathbf{u}_Q - \mathbf{u}_Q^{lo})\|_{L^2 L^\infty} \lesssim \|\mathbf{u}_Q^{hi}[0]\|_{\mathcal{H}^{s_1}}. \tag{10.26}$$

The solutions $\mathbf{u}_Q$, respectively $\mathbf{u}_Q^{lo}$ agree with $\mathbf{u}$, $\mathbf{u}^{lo}$ in $[0,1] \times 4\tilde{Q}$. Then we can recombine the $\mathbf{u}_Q$ bounds using a partition of unity on the unit spatial scale. We obtain a $\mathbf{u}$ bound, namely

$$\|\mathbf{u} - \mathbf{u}^{lo}\|_{L^\infty \mathcal{H}^{s_1}} + \|\langle D_x\rangle^{\delta_0}\partial(\mathbf{u} - \mathbf{u}^{lo})\|_{L^2 L^\infty} \lesssim \|\mathbf{u}_Q^{hi}[0]\|_{\mathcal{H}^{s_1}} \ll 1. \tag{10.27}$$

On the other hand for $\mathbf{u}_{lo}$ we have the bounds given by Theorem 10.3.

The energy bounds for $\mathbf{u} - \mathbf{u}^{lo}$ and $\mathbf{u}^{lo}$ combined yield the desired energy bound (10.18) in the theorem. In terms of the Strichartz bounds (10.20), we already have them for $\mathbf{u} - \mathbf{u}^{lo}$ so it remains to prove them for $\mathbf{u}^{lo}$. But there we trivially use Sobolev embeddings and Holder's inequality in time.

It remains to consider the Strichartz estimates for $\mathcal{H}^1$ solutions to the linearized equation. By the same finite speed of propagation argument as above, it suffices to prove them for the linearization around the localized solutions $\mathbf{u}_Q$. But this follows by Theorem 10.2. $\square$

To conclude this section we reinterpret the above result in the context of the minimal surface equation, exactly in the for it will be used in the last section. We keep the same notations, with the only change that now $s_c = \frac{n}{2} + 1$:

Theorem 10.6. *For any s_1 as in (10.3), (10.4) there exists $\delta_0 > 0$ so that the following holds: For any $M > 0$ and any solution u to the problem (1.7) with initial data satisfying*

$$\|u[0]\|_{\dot{\mathcal{H}}^{s_1}} \ll M, \qquad \|u[0]\|_{\dot{\mathcal{H}}^{s_c}} \ll 1. \tag{10.28}$$

We have:

a) The solution exists up to time T_M given by

$$T_M^\sigma = M^{-1}, \qquad \sigma = s_1 - s_c, \tag{10.29}$$

with uniform bounds

$$\|u[\cdot]\|_{C([0,T_M];\mathcal{H}^{s_1})} \lesssim M, \qquad \|u[\cdot]\|_{C([0,T_M];\dot{\mathcal{H}}^{s_c})} \lesssim 1. \tag{10.30}$$

as well as

$$\|u[\cdot]\|_{C([0,T_M];\mathcal{H}^1)} \lesssim \|u[0]\|_{\mathcal{H}^1}, \tag{10.31}$$

whenever the right hand side is finite.

b) The solution u satisfies the following Strichartz estimates in $[0, T_M]$:

<div style="text-align: right;">Str-M2-ms</div>

(10.32)
$$\|\langle T_M D'\rangle^{\delta_0} \partial^2 u\|_{L^4 L^\infty} \lesssim T_M^{-\frac{3}{4}}, \qquad n = 2,$$

respectively

<div style="text-align: right;">str-M3-ms</div>

(10.33)
$$\|\langle T_M D'\rangle^{\delta_0} \partial^2 u\|_{L^2 L^\infty} \lesssim T_M^{-\frac{1}{2}}, \qquad n \geq 3.$$

c) Furthermore, the homogeneous Strichartz estimates (4.33) also hold in $\mathcal{H}^r$ for the associated linear equations (10.7), on the same time intervals for $r \in [1, s_1]$. Also the the full Strichartz estimates (4.42) with $S = S_{ST}$ hold for the linear paradifferential equation hold in $\mathcal{H}^r$ on the same time intervals for all real r.

The theorem is obtained by applying the previous theorem to $\mathbf{u} = \partial u$. For the Strichartz estimates for the linear paradifferential equation we observe in addition that we have the bound
$$\|\partial^2 g\|_{L^1(0,T_M;L^\infty)} \lesssim 1.$$
Then the $r = 1$ case of the Strichartz estimates for the linear equations (10.7) together with Proposition 4.8 imply the desired conclusion.

11. Conclusion: proof of the main result

<div style="text-align: left;">s:final</div>

After using the finite speed of propagation to reduce to the small data problem, here we combine our balanced energy estimates with the short time Strichartz bounds in order to complete the proof of our main result in Theorem 1.3. Our rough solutions are constructed as limits of smooth solutions obtained by regularizing the initial data, so the emphasis is on obtaining favourable estimates for these smooth solutions.

11.1. **Reduction to small data.** By Sobolev embeddings, the initial data satisfies
$$\|u_0\|_{C^{1,\sigma}} + \|u_1\|_{C^\sigma} \lesssim 1, \qquad \sigma = s - \frac{n}{2} - 1.$$
Then given $x_0 \in \mathbb{R}^n$, within a small ball $B(x_0, 4r)$ we have
$$|u_0(x) - (u_0(x_0) + (x - x_0)\partial u(x_0))| + |u_1 - u_1(x_0)| \lesssim r^\sigma.$$
This allows us to truncate the above differences near x_0 to obtain the localized data
$$u_0^{r,x_0}(x) = (u_0(x_0) + (x - x_0)\partial u(x_0)) + \chi(r^{-1}(x - x_0))u_0(x) - (u_0(x_0) + (x - x_0)\partial u(x_0)),$$
$$u_1^{r,x_0}(x) = u_1(x_0) + \chi(r^{-1}(x - x_0))(u_1 - u_1(x_0)),$$
where $\chi \in \mathcal{D}(\mathbb{R}^n)$ is equal to 1 in $B(0, 2)$ and 0 outside $B(0, 4)$.

Let $\epsilon > 0$. Then for small enough r, depending on ϵ, these initial data are close to the initial data for the linear solution to the minimal surface equation given by
$$\tilde{u}^{x_0,r}(t, x) = (u_0(x_0) + (x - x_0)\partial u(x_0)) + t u_1(x_0),$$
in the sense that

<div style="text-align: left;">small-data</div>

(11.1)
$$\|u^{x_0,r}[0] - \tilde{u}^{x_0,r}[0]\|_{\mathcal{H}^s} \leq \epsilon \ll 1.$$

This will be our smallness condition for the initial data, with $\partial \mathbf{u}^{x_0,r}$ in a compact subset of the set described in (1.12).

<div style="text-align: center;">114</div>

To reduce the problem to the case when the initial data satisfies instead the simpler smallness condition

<div style="text-align: right;">mall-data+</div>

(11.2)
$$\|u^{x_0,r}[0]\|_{\mathcal{H}^s} \leq \epsilon \ll 1$$

it suffices to apply a linear transformation in the Minkowski space $\mathbb{R}^{n+2}$ that preserves the time slices but maps our linear solution $\tilde{u}^{x_0,r}$ to the zero solution. The price we pay for this is that the background Minkowski metric is then changed to another Lorentzian metric. But the new metric belongs to a compact set in the space of flat Lorenzian metrics for which the time slices are uniformly space-like and the graph of the zero function is uniformly time-like. Hence our small data result applies uniformly to these localized solutions, see Remark $\overset{\texttt{r:all-Lorentz}}{3.1}$. Then, due to the finite speed of propagation, we also obtain solutions up to time $O(r)$ for the original problem.

11.2. **Uniform bounds for regularized solutions.** Let s be as in Theorem $\overset{\texttt{t:main}}{1.3}$. Given an initial data $u[0] \in \mathcal{H}^s$ that is small,

<div style="text-align: right;">-data-last</div>

(11.3)
$$\|u[0]\|_{\mathcal{H}^s} \leq \epsilon \ll 1,$$

we consider a continuous family of frequency localizations

$$u^h[0] = P_{<h}u[0]$$

to frequencies $\leq 2^h$. For fixed h and a short time which may depend on h, these solutions exist by Theorem $\overset{\texttt{t:ST-large}}{10.5}$. Further, they are smooth and also depend smoothly on h. Finally, we consider the functions

$$v^h = \frac{d}{dh}u^h.$$

These functions solve the linearized equation around u^h, with initial data

(11.4)
$$v^h[0] = P_h u[0],$$

which is localized at frequency 2^h. The functions v^h will be measured in $\mathcal{H}^{\frac{1}{2}}$ in dimension $n \geq 3$ and in $\mathcal{H}^{\frac{5}{8}}$ in dimension $n = 2$. Thus the initial data for v^h satisfies the bound

<div style="text-align: right;">vh-data</div>

(11.5)
$$\|v^h[0]\|_{\mathcal{H}^{\frac{1}{2}}} \lesssim 2^{-(s-\frac{1}{2})h}\epsilon \qquad n \geq 3,$$
$$\|v^h[0]\|_{\mathcal{H}^{\frac{5}{8}}} \lesssim 2^{-(s-\frac{5}{8})h}\epsilon \qquad n = 2.$$

Our first objective will be to show that these solutions exist on a time interval that does not depend on h, and satisfy uniform bounds:

<div style="text-align: right;">t:uh</div>

Theorem 11.1. *The above solutions u^h have the following properties:*

a) Uniform lifespan and uniform bounds: *The solutions u^h exist up to time 1, with uniform bounds*

<div style="text-align: right;">uh-s</div>

(11.6)
$$\|u^h[\cdot]\|_{C([0,1];\mathcal{H}^s)} \lesssim \epsilon,$$

and higher regularity bounds

<div style="text-align: right;">uh-s+1</div>

(11.7)
$$\|u^h[\cdot]\|_{C([0,1];\mathcal{H}^{s+1})} \lesssim 2^h\epsilon.$$

<div style="text-align: center;">115</div>

b) Bounds for the linearized flow: *The linearized equation around u^h is well-posed in $\mathcal{H}^{\frac{1}{2}}$, with uniform estimates in $[0,1]$, uniformly in h,*

$$
\begin{aligned}
\|v\|_{L^\infty([0,1];\mathcal{H}^{\frac{1}{2}})} &\lesssim \|v[0]\|_{\mathcal{H}^{\frac{1}{2}}}, && n \geq 3, \\
\|v\|_{L^\infty([0,1];\mathcal{H}^{\frac{5}{8}})} &\lesssim \|v[0]\|_{\mathcal{H}^{\frac{5}{8}}}, && n = 2,
\end{aligned}
$$

(11.8)

and uniform Strichartz estimates with loss of derivatives,

$$
\begin{aligned}
\|\langle D_x\rangle^{-\frac{n}{2}-\frac{1}{4}-\delta}\partial v\|_{L^2 L^\infty} &\lesssim \|v[0]\|_{\mathcal{H}^{\frac{1}{2}}}, && n \geq 3, \\
\|\langle D_x\rangle^{-\frac{n}{2}-\frac{1}{4}-\delta}\partial v\|_{L^4 L^\infty} &\lesssim \|v[0]\|_{\mathcal{H}^{\frac{5}{8}}}, && n = 2,
\end{aligned}
$$

(11.9)

for any $\delta > 0$.

The exponent $s + 1$ in ($\overset{\text{uh-s+1}}{11.7}$) is chosen so that it falls into the range of existing theory, where we already have well-posedness and continuous dependence. We remark that, as a corollary of part (b), we also obtain uniform bounds for the functions v_h, namely

$$
\begin{aligned}
\|v^h\|_{L^\infty([0,1];\mathcal{H}^{\frac{1}{2}})} &\lesssim \epsilon 2^{(s-\frac{1}{2})h}, && n \geq 3, \\
\|v^h\|_{L^\infty([0,1];\mathcal{H}^{\frac{5}{8}})} &\lesssim \epsilon 2^{(s-\frac{5}{8})h}, && n = 2.
\end{aligned}
$$

$\boxed{\texttt{vh-uniform}}$ (11.10)

11.3. The bootstrap assumptions. Our proof of the main result in Theorem $\overset{\text{t:uh}}{11.1}$ will be formulated as a bootstrap argument. Then the question is what is a good bootstrap assumption. Having the bounds for the linearized equation as part of the bootstrap assumption would be technically complicated. On the other hand, not having any assumptions at all related to the linearized equation would introduce too many difficulties in getting the argument started. As it turns out, there is a good middle ground, which is to have the uniform energy bounds on both u^h and v^h as part of the bootstrap assumptions, which are then set as follows:

i) Uniform $\mathcal{H}^s$ bounds:

$\boxed{\texttt{boot-uh}}$ (11.11)
$$\|u^h[\cdot]\|_{C([0,1];\mathcal{H}^s)} \leq 1,$$

ii) Higher regularity bounds:

$\boxed{\texttt{boot-uh-hi}}$ (11.12)
$$\|u^h[\cdot]\|_{C([0,1];\mathcal{H}^{s+1})} \leq 2^h,$$

iii) Difference bounds,

$\boxed{\texttt{boot-vh}}$ (11.13)
$$
\begin{aligned}
\|v_h\|_{L^\infty([0,1];\mathcal{H}^{\frac{1}{2}})} &\leq 2^{-(s-\frac{1}{2})h}, && n \geq 3, \\
\|v_h\|_{L^\infty([0,1];\mathcal{H}^{\frac{5}{8}})} &\leq 2^{-(s-\frac{5}{8})h}, && n = 2.
\end{aligned}
$$

The v_h bootstrap bound will be useful in particular in order to obtain good low frequency bounds for differences of the u^h functions,

(11.14)
$$\|u^h - u^k\|_{L^\infty([0,1];\mathcal{H}^{\frac{1}{2}})} \lesssim 2^{-(s-\frac{1}{2})h}, \qquad h \leq k,$$

with the obvious change in two dimensions.

To avoid having a bootstrap assumption on a noncompact set of functions, we may freely restrict the range of h. Precisely, given an arbitrary threshold h_0, we assume the bootstrap assumption to hold for all $h \leq h_0$ and show that the desired bounds hold in the same range. Since h_0 plays no role in the analysis, we will simply drop it in the proofs.

11.4. **Short time Strichartz estimates for u^h and v^h.** Our goal here is to use the results in Theorem 10.6 together with our bootstrap assumption in order to obtain short time Strichartz estimates for both u^h and v^h.

By the bootstrap assumptions (11.11) and (11.12), we may bound the local well-posedness norm $\mathcal{H}^{s_1}$ of the solution u^h by

(11.15)
$$\|u^h[\cdot]\|_{L^\infty \mathcal{H}^{s_1}} \lesssim M_h := 2^{h(s_1-s)}.$$

Then the result of Theorem 10.5 is valid on time intervals I_h of length

$$|I_h| = T_h := M_h^{-\frac{1}{\sigma}} = 2^{-\frac{s_1-s}{s_1-s_c}h}.$$

In practice, s_1 will be chosen as close as possible to the threshold in (10.3),(10.4). This will insure that in all dimensions we have

$$\frac{s_1 - s}{s_1 - s_c} < \frac{1}{2}.$$

In particular, by Theorem 10.5 it follows that the solution u^h satisfies full Strichartz estimates on such intervals,

(11.16)
$$\|\langle D_x \rangle^{1+\delta_0} \partial u^h\|_{L^2(I_h;L^\infty)} \lesssim T_h^{-\frac{1}{2}}, \qquad n \geq 3,$$

respectively

(11.17)
$$\|\langle D_x \rangle^{1+\delta_0} \partial u^h\|_{L^4(I_h;L^\infty)} \lesssim T_h^{-\frac{3}{4}}, \qquad n = 2.$$

Also the linearized problem and the linear paradifferential flow will be well-posed in $\mathcal{H}^{\frac{1}{2}}$ and will satisfy Strichartz estimates on similar time intervals,

(11.18)
$$\|\langle D_x \rangle^{-\frac{n}{2}-\delta} \partial v\|_{L^2(I_h;L^\infty)} \lesssim \|v\|_{L^\infty(I_h;\mathcal{H}^{\frac{1}{2}})}, \qquad n \geq 3,$$

respectively

(11.19)
$$\|\langle D_x \rangle^{-\frac{n}{2}-\frac{1}{8}-\delta} \partial v\|_{L^4(I_h;L^\infty)} \lesssim \|v\|_{L^\infty(I_h;\mathcal{H}^{\frac{5}{8}})}, \qquad n = 2,$$

where the L^∞ norm on the right may be replaced by the same $\mathcal{H}^{\frac{1}{2}}$ norm evaluated at some fixed time within I_h. The last set of bounds may be in particular applied to v^h, which, in view of our bootstrap assumption, yields

(11.20)
$$\|\langle D_x \rangle^{-\frac{n}{2}-\delta} \partial v^h\|_{L^2(I_h;L^\infty)} \lesssim 2^{-(s-\frac{1}{2})h}, \qquad n \geq 3,$$

respectively

(11.21)
$$\|\langle D_x \rangle^{-\frac{n}{2}-\frac{1}{8}-\delta} \partial v^h\|_{L^4(I_h;L^\infty)} \lesssim 2^{-(s-\frac{5}{8})h}, \qquad n = 2.$$

11.5. **Long time Strichartz estimates for u^h and v^h.** Our objective now is to obtain long time Strichartz bounds by simply adding up the short time bounds. Some care is needed when using (11.21) and (11.20) because, as h increases, we gain on one hand in the bound on the right, but we loose in the size of the interval I_h. However, the gain overrides the loss, so integrating in h we arrive at the difference bound

(11.22)
$$\|\langle D_x \rangle^{-\frac{n}{2}-\delta} \partial(u^h - u^k)\|_{L^2(I_h;L^\infty)} \lesssim 2^{-(s-\frac{1}{2})h}, \qquad n \geq 3, \quad h < k,$$

respectively

(11.23)
$$\|\langle D_x \rangle^{-\frac{n}{2}-\frac{1}{8}-\delta} \partial(u^h - u^k)\|_{L^4(I_h;L^\infty)} \lesssim 2^{-(s-\frac{5}{8})h}, \qquad n = 2, \quad h < k.$$

Now we are able to obtain Strichartz bounds for u^h on the full time interval $[0,1]$, simply by adding the short time bounds. Precisely, we claim that for some small universal $\delta_1 > 0$ we have

(11.24)
$$\|\langle D_x \rangle^{\frac{1}{2}+\delta_1} P_k \partial u^h\|_{L^2([0,1];L^\infty)} \lesssim 1, \qquad n \geq 3,$$

respectively

(11.25)
$$\|\langle D_x \rangle^{\frac{1}{2}+\delta_1} P_k \partial u^h\|_{L^4(0,1;L^\infty)} \lesssim 1 \qquad n = 2.$$

To see this, we differentiate cases depending on how k and h compare. We fix the dimension to $n \geq 3$ for clarity.

a) If $k \geq h$, then we simply apply ($\overset{\text{uh-short3}}{11.16}$) or ($\overset{\text{uh-short2}}{11.17}$), taking the loss from the number of intervals. For instance in three and higher dimensions we get for $\delta_1 \leq \delta_0$

$$\|\langle D_x \rangle^{\frac{1}{2}+\delta_1} P_k \partial u^h\|_{L^2([0,1];L^\infty)} \lesssim T_h^{-\frac{1}{2}} \sup_{I_h} \|\langle D_x \rangle^{\frac{1}{2}+\delta_1} P_k \partial u^h\|_{L^2(I_h;L^\infty)}$$

$$\lesssim T_h^{-\frac{1}{2}} 2^{-\frac{k}{2}} \sup_{I_h} \|\langle D_x \rangle^{1+\delta_0} P_k \partial u^h\|_{L^2(I_h;L^\infty)}$$

$$\lesssim T_h^{-\frac{1}{2}} 2^{-\frac{h}{2}} T_h^{-\frac{1}{2}} = 2^{(\frac{s_1-s}{s_1-s_c}-\frac{1}{2})h} \leq 1,$$

for a favourable choice of s_1; for instance $s_1 = s + \frac{1}{4}$ suffices, as then $\frac{s_1-s}{s_1-s_c} < \frac{1}{2}$. The two dimensional argument is similar.

b) if $k < h$ instead, then we first write

$$P_k u^h = P_k u^k + P_k(u^h - u^k).$$

Here the first term was already estimated before, while for the second we use ($\overset{\text{du-short3}}{11.22}$) or ($\overset{\text{du-short2}}{11.23}$), where the loss from the interval size is only in terms of k and not h. In dimension three and higher this yields

$$\|\langle D_x \rangle^{\frac{1}{2}+\delta_1} P_k \partial(u^h - u^k)\|_{L^2([0,1];L^\infty)} \lesssim T_k^{-\frac{1}{2}} \sup_{I_h} \|P_k \partial(u^h - u^k)\|_{L^2(I_h;L^\infty)}$$

$$\lesssim T_k^{-\frac{1}{2}} 2^{-(s-\frac{1}{2})k} 2^{(\frac{n+1}{2}+\delta)k} = 2^{(\frac{s_1-s_c}{2\sigma}-(s-\frac{1}{2})+\frac{n+1}{2}+\delta)k}$$

$$\lesssim 2^{(\frac{s_1-s_c}{2\sigma}-(s-s_c)+\delta)k} \leq 1,$$

again for a good choice of s_1 (same as above) and a small enough δ.

In particular, the estimates ($\overset{\text{Bdelta3}}{11.24}$), respectively ($\overset{\text{Bdelta2}}{11.25}$) allow us to estimate our control parameter $\mathcal{B}$ as follows:

(11.26)
$$\|\mathcal{B}\|_{L^2[0,1]} \lesssim 1, \qquad n \geq 3,$$

respectively

(11.27)
$$\|\mathcal{B}\|_{L^4[0,1]} \lesssim 1, \qquad n = 2.$$

This in turn allows us to use Theorem $\overset{\text{t:ee}}{8.1}$ to control the energy growth for the full equation, and in particular to prove the bounds ($\overset{\text{uh-s}}{11.6}$) and ($\overset{\text{uh-s+1}}{11.7}$), thus closing part of the bootstrap loop, namely for the bounds ($\overset{\text{boot-uh}}{11.11}$) and for ($\overset{\text{boot-uh-hi}}{11.12}$).

11.6. **Strichartz estimates for the paradifferential flow.** Our objective here is to establish Strichartz estimates with loss of derivatives for the linear paradifferential flow around u^h. Thus, we consider an $\mathcal{H}^{\frac{1}{2}}$ solution v for the paradifferential flow around u^h, and we seek to estimate it dyadic pieces in the Strichartz norm, with frequency losses:

<div style="border:1px solid">p:Str-v</div> **Proposition 11.2.** *Under the bootstrap assumptions* ($\overset{\texttt{boot-uh}}{11.11}$), ($\overset{\texttt{boot-uh-hi}}{11.12}$) *and* ($\overset{\texttt{boot-vh}}{11.13}$), $\mathcal{H}^r$ *solutions v for the linear paradifferential equation*

$$(11.28) \qquad \partial_\alpha T_{g^{\alpha\beta}(\partial u^h)} \partial_\beta v = f$$

satisfy the Strichartz estimates ($\overset{\texttt{Str-full+}}{4.43}$) *with $S = S_{AIT}$ for all $r \in \mathbb{R}$.*

Compared with the full Strichartz bounds, here we have a loss of $1/4$ derivative in dimension 3 and higher, respectively $1/8$ derivative in dimension 2.

Proof. Our starting point is Theorem $\overset{\texttt{t:Str-move-around}}{4.12}$, which allows us to reduce the problem to proving the homogeneous Strichartz estimates ($\overset{\texttt{Str-hom}}{4.33}$) for the corresponding homogeneous equation, again for all real r. To prove the proposition in this case, we have two tools at our disposal:

(i) The energy estimates of Theorem $\overset{\texttt{t:para-wp}}{7.1}$. In view of the bounds ($\overset{\texttt{BB2}}{11.27}$) and ($\overset{\texttt{BB3}}{11.26}$), these give uniform $\mathcal{H}^r$ bounds for v,

$$\|v[\cdot]\|_{L^\infty \mathcal{H}^r} \lesssim \|v[0]\|_{\mathcal{H}^r}.$$

(ii) The short time Strichartz estimates ($\overset{\texttt{Str-hom}}{4.33}$) with $S = S_{ST}$ on the T_h time scale, provided by Theorem $\overset{\texttt{t:ST-large-ms}}{10.6}$. Adding these with respect to the time intervals, we arrive at

<div style="border:1px solid">v-long3-</div> $$(11.29) \qquad \||D|^{-\frac{d}{2}-\delta} \partial v\|_{L^2([0,1];L^\infty)} \lesssim T_h^{-\frac{1}{2}} \|v[0]\|_{\mathcal{H}^{\frac{1}{2}}}, \qquad n \geq 3,$$

respectively

<div style="border:1px solid">v-long2-</div> $$(11.30) \qquad \||D|^{-\frac{9}{8}-\delta} \partial v\|_{L^4(I_h;L^\infty)} \lesssim T_h^{-\frac{1}{4}} \|v[0]\|_{\mathcal{H}^{\frac{5}{8}}}, \qquad n = 2.$$

Now we want to use these tools in order to prove the long term bounds ($\overset{\texttt{Str-hom}}{4.33}$) with $S = S_{AIT}$ on the unit time scale. Given the expression for T_h, our first observation is that the estimates ($\overset{\texttt{v-long3-}}{11.29}$), respectively ($\overset{\texttt{v-long2-}}{11.30}$) suffice for our bounds at frequencies $\geq h$, but not below that.

Thus, consider a lower frequency $k < h$, and seek to estimate $P_k v$. At this frequency, we have the correct estimate for the solution $\tilde{v}$ to the linear paradifferential equation around u_k. It remains to compare v and $\tilde{v}$. For this we use the $T_{\tilde{P}(u^k)}$ flow, and we think of $P_k v$ as an approximate solution for this flow,

$$T_{P(u^k)} P_k v = [T_{P(u^k)}, P_k] v + P_k \partial_\alpha T_{g_h^{\alpha\beta} - g_k^{\alpha\beta}} \partial_\beta v.$$

We can bound the source terms as follows, fixing the dimension to $n \geq 3$:

$$\|T_{P(u^k)} P_k v\|_{L^1(I_k, H^{-\frac{1}{2}})} \lesssim \left(\|\partial^2 u^k\|_{L^1(I_k, L^\infty)} + 2^k \|P_{<k}(g(\partial u^h) - g(\partial u^k))\|_{L^1(I_k, L^\infty)} \right) \|v\|_{L^\infty \mathcal{H}^{\frac{1}{2}}}.$$

To conclude it suffices to estimate

<div style="border:1px solid">str-uk</div> $$(11.31) \qquad \|\partial^2 u^k\|_{L^1(I_k, L^\infty)} \lesssim 1,$$

<div style="border:1px solid">str-duk</div> $$(11.32) \qquad \|P_{<k}(g(\partial u^h) - g(\partial u^k))\|_{L^1(I_k, L^\infty)} \lesssim 2^{-k}.$$

The first bound follows from our earlier Strichartz estimates for u^k, see (11.17), (11.16). For the second bound, we expand and then it suffices to have

want-vj   (11.33) $$\|P_{<k}(g'(\partial u^j)\partial v^j)\|_{L^1(I_k,L^\infty)} \lesssim 2^{-k}2^{-c(j-k)}, \qquad j > k,$$

with a positive constant c in order to allow for integration in j. We expand paradifferentially, depending on the frequencies of the two factors above. It suffices to consider the following two cases:

a) v^j has the frequency below 2^k. Then we use the Strichartz bounds for v_j over intervals I_j, and then sum over such intervals. For instance in dimension $n \geq 3$ we get

$$\|P_{<k}\partial v_j\|_{L^1(I_k;L^\infty)} \lesssim |I_k||I_j|^{-\frac{1}{2}} \sup_{I_j} \|P_{<k}\partial v_j\|_{L^2(I_j;L^\infty)}$$

$$\lesssim 2^{(\frac{n}{2}+\delta)k}2^{-(s-\frac{1}{2})j}\frac{|I_k|}{|I_j|}|I_j|^{\frac{1}{2}}$$

$$= 2^{[\frac{n}{2}+\delta-\frac{s_1-s}{s_1-s_c}]k}2^{-(s-\frac{1}{2}-\frac{s_1-s}{2(s_1-s_c)})j}$$

$$= 2^{-k}2^{[(s_c-s)(1-\frac{1}{2(s_1-s_c)})+\delta]k}2^{-(s-\frac{1}{2}-\frac{s_1-s}{2(s_1-s_c)})(j-k)}.$$

Here the coefficient of $j - k$ is negative by a large margin, while the coefficient of k in the middle factor is also negative since $s_1 - s_c > \frac{1}{2}$ and δ is arbitrarily small. Hence we obtain a bound as desired in (11.33).

b) The balanced case, where both frequencies have size 2^l with $l \geq k$. This is easier, as we have a better energy bound for the first factor. Hence in this case it is more efficient to estimate the output by applying Bernstein's inequality first,

$$\|P_{<k}[P_l g'(\partial u^j)P_l\partial v^j]\|_{L^1(I_k,L^\infty)} \lesssim |I_k|^{\frac{1}{2}}2^{\frac{nk}{2}}\|P_l g'(\partial u^j)\|_{L^\infty(I_k,L^2)}\|P_l\partial v^j\|_{L^2(I_k,L^\infty)}$$

$$\lesssim |I_k||I_j|^{-\frac{1}{2}}2^{\frac{nk}{2}}\|P_l g'(\partial u^j)\|_{L^\infty L^2}\sup_{I_j}\|P_l\partial v^j\|_{L^2(I_j,L^\infty)}$$

$$\lesssim |I_k||I_j|^{-\frac{1}{2}}2^{\frac{nk}{2}}2^{-(s-1)l}2^{(\frac{n}{2}+\delta)l}2^{-(s-\frac{1}{2})j}$$

$$= 2^{[\frac{n}{2}-\frac{s_1-s}{s_1-s_c}]k}2^{(\frac{n}{2}-s+1+\delta)l}2^{-(s-\frac{1}{2}-\frac{s_1-s}{2(s_1-s_c)})j}$$

$$= 2^{-k}2^{[(s_c-s)(2-\frac{1}{2(s_1-s_c)})+\delta]k}2^{(\frac{n}{2}-s+1+\delta)(l-k)}2^{-(s-\frac{1}{2}-\frac{s_1-s}{2(s_1-s_c)})(j-k)},$$

which is better than in case (a). $\qquad\square$

11.7. Strichartz estimates for the linearized flow. Our aim here is to show that that we have $\mathcal{H}^{\frac{1}{2}}$ well-posedness and Strichartz estimates with loss of derivatives for the linearized flow around u^h in $\mathcal{H}^{\frac{1}{2}}$ (respectively $\mathcal{H}^{\frac{5}{8}}$ if $n = 2$):

:Str-v-lin   **Proposition 11.3.** *Under the bootstrap assumptions* (11.11), (11.12) *and* (11.13), *the linearized equation around u^h is well-posed in $\mathcal{H}^{\frac{1}{2}}$ (respectively $\mathcal{H}^{\frac{5}{8}}$ if $n = 2$), and its solutions satisfy the full Strichartz estimates* (4.43) *with $S = S_{AIT}$.*

Here we use the analysis in Section 9. Precisely Theorem 9.1 there shows that the above proposition follows directly from the similar result in Proposition 11.2 for the linear paradifferential equation.

120

11.8. **Closing the bootstrap argument.** Combining the Strichartz estimates for the linear paradifferential equation in Proposition 11.2 with the result of Theorem 9.1, it follows that the linearized flow around u^h is wellposed in $\mathcal{H}^{\frac{1}{2}}$ (respectively $\mathcal{H}^{\frac{5}{8}}$ if $n = 2$), with the same Strichartz estimates as in Proposition 11.2, which is exactly part (b) of Theorem 11.1. As a consequence, the initial data bound (11.5) for v implies the uniform bound (11.10), which in turn closes the bootstrap assumption (11.13).

11.9. **The well-posedness result.** In order to be able to obtain a complete well-posedness argument, we follow the outline in [21], and measure the size of the functions u^h and v^h in terms of frequency envelopes. Precisely, we consider a normalized frequency envelope ϵc_h for $u[0]$ in $\mathcal{H}^s$. Then for the localized initial data we have the bounds

$$(11.34) \qquad\qquad \|u^h[0]\|_{\mathcal{H}^s} \lesssim \epsilon,$$

$$(11.35) \qquad\qquad \|u^h[0]\|_{\mathcal{H}^{s+1}} \lesssim 2^h \epsilon c_h.$$

On the other hand, fixing the dimension to $n \geq 3$, we will measure v^h in $\mathcal{H}^{\frac{1}{2}}$, where for the initial data we have

$$(11.36) \qquad\qquad \|v^h[0]\|_{\mathcal{H}^{\frac{1}{2}}} \lesssim 2^{-(s-\frac{1}{2})h} c_h, \qquad n \geq 3.$$

Then by Theorem 11.1, we obtain corresponding uniform bounds for the solutions on the time interval $[0,1]$,

$$(11.37) \qquad\qquad \|u^h[\cdot]\|_{L^\infty([0,1];\mathcal{H}^s)} \lesssim \epsilon,$$

$$(11.38) \qquad\qquad \|u^h[\cdot]\|_{L^\infty([0,1];\mathcal{H}^{s+1})} \lesssim 2^h \epsilon c_h.$$

Similarly, the linearized increments v_h satisfy the uniform bounds

$$(11.39) \qquad\qquad \|v_h[\cdot]\|_{L^\infty([0,1];\mathcal{H}^{\frac{1}{2}})} \lesssim \epsilon 2^{(s-\frac{1}{2})s} c_h.$$

Integrating the last bound with respect to h, we obtain the difference bounds

$$(11.40) \qquad\qquad \|(u^h - u^k)[\cdot]\|_{L^\infty([0,1];\mathcal{H}^{\frac{1}{2}})} \lesssim \epsilon 2^{(s-\frac{1}{2})s} c_h, \qquad h < k.$$

This implies that the limit

$$u = \lim_{h\to\infty} u^h$$

exists in $C([0,1];\mathcal{H}^{\frac{1}{2}})$. In view of (11.37), the limit u will also satisfy

$$\|u[\cdot]\|_{L^\infty([0,1];\mathcal{H}^s)} \lesssim \epsilon.$$

We can also prove that we have the previous convergence in this stronger topology. To see this, we consider unit increments in h, and compare u_h with u_{h+1}, using (11.38) on one hand, and (11.40) on the other hand. This yields

$$(11.41) \qquad\qquad \|u^h - u^{h+1}\|_{C([0,1];\mathcal{H}^{s+1})} \lesssim 2^h \epsilon c_h,$$

respectively

$$(11.42) \qquad\qquad \|u^h - u^{h+1}\|_{C([0,1];\mathcal{H}^{\frac{1}{2}})} \lesssim \epsilon 2^{(s-\frac{1}{2})s} c_h, \qquad h < k.$$

121

These two bounds balance exactly at frequency 2^h, and measure the $\mathcal{H}^s$ norm but with decay away from frequency 2^k. Hence the differences are almost orthogonal in $\mathcal{H}^s$, and, summing them up, we obtain

$$\boxed{\texttt{uh-diff=}} \qquad (11.43) \qquad \|u^h - u^k\|_{C([0,1];\mathcal{H}^s)} \lesssim \epsilon c_{[h,k]}.$$

This implies uniform convergence in $\mathcal{H}^s$. Thus our solution u is uniquely identified as the strong $\mathcal{H}^s$ uniform limit of u^h.

The continuous dependence and the weak Lipschitz dependence follow exactly as in [21].

References

[1] Albert Ai, Mihaela Ifrim, and Daniel Tataru. Two dimensional gravity waves at low regularity I: Energy estimates, 2019.

[2] Thomas Alazard and Jean-Marc Delort. Global solutions and asymptotic behavior for two dimensional gravity water waves. *Ann. Sci. Éc. Norm. Supér. (4)*, 48(5):1149–1238, 2015.

[3] Thomas Alazard and Jean-Marc Delort. Sobolev estimates for two dimensional gravity water waves. *Astérisque*, (374):viii+241, 2015.

[4] Hajer Bahouri and Jean-Yves Chemin. Équations d'ondes quasilinéaires et effet dispersif. *Internat. Math. Res. Notices*, (21):1141–1178, 1999.

[5] Hajer Bahouri and Jean-Yves Chemin. Équations d'ondes quasilinéaires et estimations de Strichartz. *Amer. J. Math.*, 121(6):1337–1377, 1999.

[6] Jean-Michel Bony. Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires. *Ann. Sci. École Norm. Sup. (4)*, 14(2):209–246, 1981.

[7] Jean-Michel Bony. Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires. *Ann. Sci. École Norm. Sup. (4)*, 14(2):209–246, 1981.

[8] Simon Brendle. Hypersurfaces in Minkowski space with vanishing mean curvature. *Comm. Pure Appl. Math.*, 55(10):1249–1279, 2002.

[9] Michael Christ and Alexander Kiselev. Maximal functions associated to filtrations. *J. Funct. Anal.*, 179(2):409–425, 2001.

[10] Roland Donninger, Joachim Krieger, Jérémie Szeftel, and Willie Wong. Codimension one stability of the catenoid under the vanishing mean curvature flow in Minkowski space. *Duke Math. J.*, 165(4):723–791, 2016.

[11] Boris Ettinger. *Well-posedness of the three-form field equation and the minimal surface equation in Minkowski space.* ProQuest LLC, Ann Arbor, MI, 2013. Thesis (Ph.D.)–University of California, Berkeley.

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

[13] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi*, 3:e6, 75, 2015.

[14] Massimiliano Gubinelli, Herbert Koch, and Tadahiro Oh. Paracontrolled approach to the three-dimensional stochastic nonlinear wave equation with quadratic nonlinearity. *arXiv e-prints*, page arXiv:1811.07808, November 2018.

[15] J. Hoppe. Some classical solutions of relativistic membrane equations in 4-space-time dimensions. *Phys. Lett. B*, 329(1):10–14, 1994.

[16] Lars Hörmander. Pseudo-differential operators of type 1, 1. *Comm. Partial Differential Equations*, 13(9):1085–1111, 1988.

[17] Lars Hörmander. *Lectures on nonlinear hyperbolic differential equations*, volume 26 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 1997.

[18] 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.

HIT	[19]	John K. Hunter, Mihaela Ifrim, and Daniel Tataru. Two dimensional water waves in holomorphic coordinates. *Comm. Math. Phys.*, 346(2):483–552, 2016.
BH	[20]	John K. Hunter, Mihaela Ifrim, Daniel Tataru, and Tak Kwong Wong. Long time solutions for a Burgers-Hilbert equation via a modified energy method. *Proc. Amer. Math. Soc.*, 143(8):3407–3412, 2015.
IT-primer	[21]	Mihaela Ifrim and Daniel Tataru. Local well-posedness for quasilinear problems: a primer. *arXiv e-prints*, page arXiv:2008.05684, August 2020.
Jerrard	[22]	Robert Jerrard. Defects in semilinear wave equations and timelike minimal surfaces in Minkowski space. *Anal. PDE*, 4(2):285–340, 2011.
Kap	[23]	L. V. Kapitanskiĭ. Estimates for norms in Besov and Lizorkin-Triebel spaces for solutions of second-order linear hyperbolic equations. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 171(Kraev. Zadachi Mat. Fiz. i Smezh. Voprosy Teor. Funktsiĭ. 20):106–162, 185–186, 1989.
Keel-Tao	[24]	Markus Keel and Terence Tao. Endpoint Strichartz estimates. *Amer. J. Math.*, 120(5):955–980, 1998.
null	[25]	S. Klainerman. The null condition and global existence to nonlinear wave equations. In *Nonlinear systems of partial differential equations in applied mathematics, Part 1 (Santa Fe, N.M., 1984)*, volume 23 of *Lectures in Appl. Math.*, pages 293–326. Amer. Math. Soc., Providence, RI, 1986.
KR	[26]	S. Klainerman and I. Rodnianski. Improved local well-posedness for quasilinear wave equations in dimension three. *Duke Math. J.*, 117(1):1–124, 2003.
KRS	[27]	Sergiu Klainerman, Igor Rodnianski, and Jeremie Szeftel. The bounded L^2 curvature conjecture. *Invent. Math.*, 202(1):91–216, 2015.
UV	[28]	Herbert Koch and Daniel Tataru. Dispersive estimates for principally normal pseudodifferential operators. *Comm. Pure Appl. Math.*, 58(2):217–284, 2005.
KL	[29]	Joachim Krieger and Hans Lindblad. On stability of the catenoid under vanishing mean curvature flow on Minkowski space. *Dyn. Partial Differ. Equ.*, 9(2):89–119, 2012.
L-blow	[30]	Hans Lindblad. Counterexamples to local existence for quasilinear wave equations. *Math. Res. Lett.*, 5(5):605–622, 1998.
L-small	[31]	Hans Lindblad. A remark on global existence for small initial data of the minimal surface equation in Minkowskian space time. *Proc. Amer. Math. Soc.*, 132(4):1095–1102, 2004.
Metivier	[32]	Guy Métivier. *Para-differential calculus and applications to the Cauchy problem for nonlinear systems*, volume 5 of *Centro di Ricerca Matematica Ennio De Giorgi (CRM) Series*. Edizioni della Normale, Pisa, 2008.
MSS	[33]	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.
2d-counter	[34]	Gaspard Ohlmann. Ill-posedness of a quasilinear wave equation in two dimensions for data in $H^{7/4}$. *arXiv e-prints*, page arXiv:2107.03732, July 2021.
Sm	[35]	Hart F. Smith. A parametrix construction for wave equations with $C^{1,1}$ coefficients. *Ann. Inst. Fourier (Grenoble)*, 48(3):797–835, 1998.
SS-counter	[36]	Hart F. Smith and Christopher D. Sogge. On Strichartz and eigenfunction estimates for low regularity metrics. *Math. Res. Lett.*, 1(6):729–737, 1994.
ST-counter	[37]	Hart F. Smith and Daniel Tataru. Sharp counterexamples for Strichartz estimates for low regularity metrics. *Math. Res. Lett.*, 9(2-3):199–204, 2002.
ST	[38]	Hart F. Smith and Daniel Tataru. Sharp local well-posedness results for the nonlinear wave equation. *Ann. of Math. (2)*, 162(1):291–366, 2005.
Sogge	[39]	Christopher D. Sogge. *Lectures on non-linear wave equations*. International Press, Boston, MA, second edition, 2008.
S-small	[40]	Atanas Stefanov. Global regularity for the minimal surface equation in Minkowskian geometry. *Forum Math.*, 23(4):757–789, 2011.
ST-wm1	[41]	Jacob Sterbenz and Daniel Tataru. Energy dispersed large data wave maps in $2 + 1$ dimensions. *Comm. Math. Phys.*, 298(1):139–230, 2010.
Tao-wm2d	[42]	Terence Tao. Global regularity of wave maps. II. Small energy in two dimensions. *Comm. Math. Phys.*, 224(2):443–544, 2001.
T:nlw1	[43]	Daniel Tataru. Strichartz estimates for operators with nonsmooth coefficients and the nonlinear wave equation. *Amer. J. Math.*, 122(2):349–376, 2000.

`T:nlw2`	[44] Daniel Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. II. *Amer. J. Math.*, 123(3):385–423, 2001.
`T:ICM`	[45] Daniel Tataru. Nonlinear wave equations. In *Proceedings of the International Congress of Mathematicians, Vol. III (Beijing, 2002)*, pages 209–220. Higher Ed. Press, Beijing, 2002.
`T:nlw3`	[46] Daniel Tataru. Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients. III. *J. Amer. Math. Soc.*, 15(2):419–442, 2002.
`T:wm2`	[47] Daniel Tataru. Rough solutions for the wave maps equation. *Amer. J. Math.*, 127(2):293–377, 2005.
`MR3730736`	[48] Willie Wai Yeung Wong. Global existence for the minimal surface equation on $\mathbb{R}^{1,1}$. *Proc. Amer. Math. Soc. Ser. B*, 4:47–52, 2017.
`MR3786768`	[49] Willie Wai Yeung Wong. Singularities of axially symmetric time-like minimal submanifolds in Minkowski space. *J. Hyperbolic Differ. Equ.*, 15(1):1–13, 2018.
`Zwiebach`	[50] Barton Zwiebach. *A first course in string theory.* Cambridge University Press, Cambridge, second edition, 2009. With a foreword by David Gross.

Department of Mathematics, University of Wisconsin, Madison
Email address: `aai@math.wisc.edu`

Department of Mathematics, University of Wisconsin, Madison
Email address: `ifrim@wisconsin.edu`

Department of Mathematics, University of California at Berkeley
Email address: `tataru@math.berkeley.edu`