ROUGH SOLUTIONS FOR THE WAVE MAPS EQUATION

DANIEL TATARU

ABSTRACT. We consider the wave maps equation with values into a Riemannian manifold which is isometrically embedded in $\mathbb{R}^m$. Our main result asserts that the Cauchy problem is globally well-posed for initial data which is small in the critical Sobolev spaces. This extends and completes recent work of Tao and other authors.

1. INTRODUCTION

Given a Riemannian manifold (M, g) we consider M valued functions in the Minkowski space $\mathbb{R} \times \mathbb{R}^n$,

$$\phi : \mathbb{R} \times \mathbb{R}^n \to M, \qquad (t, x) \to \phi(t, x)$$

Its first order derivatives take values in the tangent space of M,

$$\partial_\alpha \phi : \mathbb{R}^n \to T_\phi M$$

Thus they are sections of the pull-back bundle

$$\phi^*(TM) = \cup_{x \in \mathbb{R}^n} \{x\} \times T_{\phi(x)} M$$

Wave maps are critical points of the Lagrangian

$$L(\phi) = \int_{\mathbb{R}^n} -|\partial_t \phi|_g^2 + |\partial_x \phi(x)|_g^2 dx$$

In order to describe the corresponding Euler-Lagrange equations it is convenient to use covariant differentiation. To the metric g we associate the natural covariant differentiation on TM defined by its Levi-Civita connection. This induces a connection $\mathbf{D}$ on the pull-back bundle $\phi^*(TM)$. If v is a section of $\phi^*(TM)$ then we set

$$\mathbf{D}_X V = (\nabla_{\phi_* X} V)$$

With this notation, the wave maps equations have the form

$$(1.1) \qquad\qquad -\mathbf{D}^\alpha \partial_\alpha \phi = 0$$

Here we lift the indices with respect to the Minkowski metric,

$$(ds)^2 = -(dt)^2 + (dx)^2$$

1

If the image of ϕ is supported in the domain of a local coordinate patch on M then we can express this in local coordinates,

$$\Box\psi^i = -\Gamma^i_{jk}(u)\partial^\alpha u^j \partial_\alpha u^k$$

This is of little use in this paper, as we consider discontinuous, unbounded solutions for the wave maps equation.

Instead, here we work with manifolds M which are isometrically embedded into $\mathbb{R}^d$ with the euclidean metric. More precisely,

Definition 1.1. We say that a Riemannian manifold M is uniformly isometrically embedded into $\mathbb{R}^d$ if there is $r > 0$ so that

a) For each $y_0 \in M$ the intersection $M \cap B(y,r)$ is the graph of a smooth function $y^1 = \phi_{y_0}(y')$ in a suitable orthonormal frame (y_0, y').

b) The derivatives of the functions ϕ_{y_0} are bounded uniformly in $y_0 \in M$.

We assume that this holds throughout the paper. If M is compact then such an embedding always exists by Nash's theorem [14] (see also Gromov [5] and Günther [6]). More generally, a uniform isometric embedding exists provided that the curvature tensor and its covariant derivatives are uniformly bounded on M.

In this case one can derive another form of the equations which involves the second fundamental form S of M. Here we view S as a symmetric bilinear form

$$S : TM \times TM \to NM, \qquad \langle S(X,Y), N \rangle = \langle \partial_X N, Y \rangle$$

Then a simple computation shows that the wave maps equation has the form

$$(1.2) \qquad\qquad \Box\phi^i = -S^i_{jk}(\phi)\partial^\alpha \phi^j \partial_\alpha \phi^k$$

Thus the Cauchy problem for the wave maps equation has the form

$$(1.3) \quad \begin{cases} \Box\phi^i = -S^i_{jk}(\phi)\partial^\alpha \phi^j \partial_\alpha \phi^k & \text{in } \mathbb{R} \times \mathbb{R}^n \\[2mm] \phi(0,x) = \phi_0(x), \ \partial_t\phi(0,x) = \phi_1(x) & \text{in } \mathbb{R}^n \end{cases}$$

The initial data (ϕ_0, ϕ_1) must be chosen so that

$$\phi_0(x) \in M, \quad \phi_1(x) \in T_{\phi_0(x)}M, \quad x \in \mathbb{R}^n$$

We shall use the notation

$$\phi[t] = (\phi(t), \partial_t\phi(t))$$

A natural question is whether the Cauchy problem is locally well-posed for initial data in Sobolev spaces,

$$\phi[0] \in H^s \times H^{s-1}$$

One has to be careful here with the definition of the Sobolev spaces because we are dealing with functions which take values in a Riemannian manifold (see for instance Brezis-Nirenberg [2], and also the discussion in Hélein [7]).

Some intuition about which are the interesting values for s can be gained from scaling. The wave maps equation is invariant with respect to the dimensionless scaling

$$\phi(t, x) \rightarrow \phi(\lambda t, \lambda x) \qquad \lambda \in \mathbb{R}$$

The scale invariant initial data space corresponds to $s = \frac{n}{2}$. Of course in order to have a scale invariant problem one has to use the homogeneous Sobolev spaces,

$$\phi[0] \in \dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$$

This is precisely the problem we consider in this article. Since it is left unchanged by scaling, the size of the initial data becomes important. We prove here that for initial data which is small in the above space the wave maps equation is globally well-posed. The similar problem for large data remains largely open.

Before we state our main result, we need to clarify the definition of the above Sobolev spaces. Since we work with $M \subset \mathbb{R}^d$, we think of the M-valued functions as $\mathbb{R}^d$ valued functions which take values in M a.e. For test functions $u \in \mathcal{D}(\mathbb{R}^n, \mathbb{R}^d)$ we use the Fourier transform to define

$$\|u\|_{\dot{H}^s} = \||\xi|^s \hat{\phi}_1\|_{L^2}$$

Then we define $\dot{H}^s$ as the completion of $\mathcal{D}$ with respect to this norm.

If $s < \frac{n}{2}$ this is a space of distributions which have locally integrable Fourier transform. However, for $s = \frac{n}{2}$ the space $\dot{H}^{\frac{n}{2}}$ is not a space of distributions. Instead it can be identified with a subspace of a quotient space,

$$\dot{H}^{\frac{n}{2}} \subset BMO \subset \mathcal{D}'(\mathbb{R}^n)/\text{Const}$$

where by Const we have denoted the space of constant functions. Since the wave maps equation is nonlinear, the constants are important. Hence we add them back in, with the disadvantage that instead of a norm we now have a seminorm which vanishes precisely on a one dimensional subspace. A side effect of this is that while the $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$ norm gives a good measurement for the size of a solution, it cannot be used as well in order to compare different solutions.

Theorem 1.2. *Assume that the manifold M is uniformly isometrically embedded into $\mathbb{R}^d$. Then the wave maps equation (1.3) is well-posed for small initial data in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$ in the following sense:*

(i) (smooth solutions) If the initial data is smooth, then there is a global smooth solution ϕ which for $s \geq \frac{n}{2}$ satisfies the global bounds:

$$(1.4) \qquad \|\phi[t]\|_{L^\infty(\dot{H}^s \times \dot{H}^{s-1})} \lesssim \|\phi[0]\|_{\dot{H}^s \times \dot{H}^{s-1}}$$

in the sense that the left hand side is finite and equality holds whenever the right hand side is finite.

(ii) (rough solutions as limits of smooth solutions) For each small initial data set in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$ there is a solution ϕ satisfying

$$(1.5) \qquad \|\phi[t]\|_{L^\infty(H^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1})} \lesssim \|\phi[0]\|_{H^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}}$$

which is the unique limit of smooth solutions in the $[L^\infty(H^{\frac{n}{2}} \times H^{\frac{n}{2}-1})]_{loc}$ topology.

(iii) (weak stability) Let $\phi^{(1)}, \phi^{(2)}$ be as in (i). Then for $s < \frac{n}{2}$ and close to it we have

$$(1.6) \qquad \|\phi^{(1)}[t] - \phi^{(2)}[t]\|_{L^\infty(\dot{H}^s \times \dot{H}^{s-1})} \lesssim \|\phi[0]\|_{\dot{H}^s \times \dot{H}^{s-1}}$$

(iv) (continuous dependence) The solution ϕ depends continuously on the initial data in the sense that for $s < \frac{n}{2}$ and close to it

$$\phi_n[0] - \phi[0] \to 0 \quad in \ \dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1} \cap \dot{H}^s \times \dot{H}^{s-1}$$

implies

$$\phi_n - \phi_0 \to 0 \ in \ L^\infty(\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}) \cap L^\infty(\dot{H}^s \times \dot{H}^{s-1})$$

It is also useful to state a local version of the result. We denote

$$B_R = \{|x| \leq R\}, \qquad Q_R = \{|t| + |x| \leq R\}$$

For functions in B_R, respectively Q_R we define their Sobolev norms as the best norms of their extensions to $\mathbb{R}^n$, respectively $\mathbb{R} \times \mathbb{R}^n$.

Theorem 1.3. *Consider the wave maps equation (1.3) with initial data $\phi[0]$ in B_R which is small in $[\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}](B_R)$. Then*

(i) (smooth solutions) If the initial data is smooth, then there is a smooth solution ϕ in Q_R which for $s \geq \frac{n}{2}$ satisfies the bounds:

$$(1.7) \qquad \|\phi[t]\|_{[L^\infty(\dot{H}^s \times \dot{H}^{s-1})](Q_R)} \lesssim \|\phi[0]\|_{[\dot{H}^s \times \dot{H}^{s-1}](B_R)}$$

(ii) (rough solutions as limits of smooth solutions) For each initial data there is a solution ϕ satisfying

$$(1.8) \qquad \|\phi[t]\|_{[L^\infty(\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1})](Q_R)} \lesssim \|\phi[0]\|_{[\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}](B_R)}$$

which is the unique limit of smooth solutions in the $L^\infty(H^{\frac{n}{2}} \times H^{\frac{n}{2}-1})$ topology.

(iii) (continuous dependence) The solution ϕ depends continuously on the initial data in the sense that $\phi_n[0] \to \phi[0]$ in $[H^{\frac{n}{2}} \times H^{\frac{n}{2}-1}](B_R)$ implies $\phi_n \to \phi$ in $[L^\infty(H^{\frac{n}{2}} \times H^{\frac{n}{2}-1})](Q_R)$.

The local well-posedness problem for the wave maps equation has received considerable attention in recent years. A natural approach is to use the method of energy estimates, but this only gives results when s is at least one unit above scaling. Some additional improvement can be gained by using the Strichartz estimates for the wave equation. This applies equally to generic equations of the form

$$\Box\phi = \mathbf{S}(\phi)(\nabla\phi^2)$$

However, the wave maps equation does not belong to the generic class. Instead its quadratic nonlinearity

$$Q_0(\phi,\phi) = \partial^\alpha\phi\partial_\alpha\phi$$

exhibits a cancellation property called the null condition, see Klainerman [8]. This was first taken advantage of in work of Klainerman-Machedon [9] ($n \geq 3$) and Klainerman-Selberg [11] ($n = 2$) who proved local well-posedness for all s above scaling, $s > \frac{n}{2}$. A key element of their approach is to use the $X^{s,b}$ spaces for bilinear estimates. These are Sobolev type spaces which are adapted to the wave operator.

At scaling the problem becomes considerably more delicate. A naive attempt to use homogeneous versions of the $X^{s,b}$ spaces quickly runs into trouble because of numerous logarithmic divergences. The first scale invariant result was obtained in work of the author [19] ($n \geq 4$) followed by [20] for the much more difficult low dimensional case $n = 2, 3$. These results are for initial data which is not in the scale invariant Sobolev space but instead are in the slightly smaller Besov space

$$\dot{B}_{2,1}^{\frac{n}{2}} \times \dot{B}_{2,1}^{\frac{n}{2}-1}$$

In the high dimensional case this is achieved by using a mix of homogeneous $X^{s,b}$ spaces and Strichartz type norms for the bilinear estimates. In the low dimensional case the construction of the function spaces for the solution becomes much more intricate and involves the use of energy spaces with respect to rotating null frames. Just as in earlier works [9], [11], the well-posedness in [19], [20] is established using the contraction principle. This gives not only existence and uniqueness but also smooth dependence of the solution on the initial data. Because of the embedding $\dot{B}_{2,1}^{\frac{n}{2}} \subset L^\infty$ the solutions are small in L^∞ therefore one can work in local coordinates in the target space M. The null condition is used, but not the form of the nonlinear coefficients $\Gamma_{jk}^i(u)$.

5

At the time it was not clear whether one can also obtain the result in the scale invariant Sobolev spaces by improving the choice of the function spaces in [19], [20]. However, we now know that this is not the case. Indeed, in recent work D'Ancona-Georgiev [3] not only show that the wave maps equation is ill-posed below scaling, but they also prove that there is no uniformly continuous dependence on the initial data in the scale invariant Sobolev space.

The next main step was achieved by Tao in [17] ($n \geq 5$) and [18] ($n \geq 2$). He considers the case when the target manifold is a sphere and proves that if the initial data is smooth and small in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$ then there is a global smooth solution. This corresponds to part (i) of Theorem 1.2. Tao's function spaces are based on the ones in [19], [20]; however, some modifications are necessary in order to to gain a key algebra property. To still be able to use linear estimates for the wave equation, he introduces a renormalization argument. This essentially eliminates the genuinely nonlinear part of the equation, leaving only a semilinear part which can be properly estimated in a semilinear fashion. This idea comes from previous work on harmonic maps (see Hélein's book [7] and references therein) but its implementation for the wave maps equation is much more delicate.

Following Tao's work, a number of authors thought to extend the result to a larger class of target manifolds. This includes work of Klainerman-Rodnianskii [10] ($n \geq 5$), Nahmod-Stefanov-Uhlenbeck [13] and Shatah-Struve [16] ($n \geq 4$) and Krieger ($n = 3$, hyperbolic space).

An alternate venue pursued by some authors was to obtain weak global energy solutions using various penalization techniques, see for instance [15], [12] and [4]. It would be interesting to understand whether in two dimensions these must coincide with the unique solutions obtained here.

The present article is focused on the two dimensional case. The results are also true in higher dimensions but in those cases many simplifications are possible. In addition, we not only consider the behavior of smooth solutions, but we also obtain rough solutions depending continuously on the initial data.

In what follows we give a brief overview of the paper. Our above definition of the initial data spaces is provided in terms of the isometric embedding of M into $\mathbb{R}^d$. But such an embedding is not unique, therefore one would like to know that our spaces do not depend on the embedding. In Section 3 we provide an invariant definition of the initial data spaces, and we show that it is equivalent to the embedding related one.

In Section 4 we summarize the properties of the function spaces S, N which we use for the solutions and for the right hand side of the wave equation. In this we largely follow Tao's setup. Theorem 4.1, which summarizes linear, bilinear and trilinear estimates for the S and N spaces, is essentially (a part of) Tao's [18], Theorem 3. We supplement this with the new nonlinear Moser type estimates in Theorem 4.3, which are essential if one wants to move from the a spherical target manifold to a general one.

The main bootstrap estimate on the regularity of wave maps is proved in Section 6. This includes the renormalization (gauge change) argument, which extends and simplifies the similar argument in Tao's paper. A related bound for the linearized equations is obtained in Section 7. Estimates for the linearized equations are essential in the study of the dependence of the solutions on the initial data.

Following some preliminary work in Section 5, our main result in Theorem 1.2 is proved in Section 8. The smooth solutions are obtained using a continuity argument in which the bootstrap estimate plays the main role. For the rough solutions and the continuous dependence on the initial data we combine this with the bounds for the linearized equations.

Finally the last section contains the proof of the Moser estimates in Theorem 4.3. This argument is quite involved and is based on certain multilinear paradifferential type expansions.

2. NOTATIONS

To the extent possible we follow Tao's [18] notations in order to make the subject easier to follow for the reader.

2.1. **Inequalities.** The expression $A \lesssim B$ means $A \leq cB$ for a universal constant c. $A \gtrsim B$ has a similar meaning. The relation $A \approx B$ is equivalent to $A \lesssim B$ and $A \gtrsim B$.

2.2. **Fourier variables.** We denote the spatial Fourier variable by ξ and the time Fourier variable by τ. The symbol of the wave equation is

$$\tau^2 - \xi^2$$

Solutions to the wave equation are concentrated near the characteristic cone

$$K = \{(\tau, \xi) \in \mathbb{R} \times \mathbb{R}^n;\ \tau^2 - \xi^2 = 0\}$$

We call ξ the frequency and the distance $||\tau| - |\xi||$ to the cone, the modulation.

2.3. **Multipliers.** First of all we use a dyadic Littlewood-Paley decomposition in frequency

$$1 = \sum_{k \in \mathbb{Z}} P_k$$

where the symbols $p_k(\xi)$ of P_k are localized in the region $\{2^{k-1} < |\xi| < 2^{k+1}\}$ and coincide up to scaling,

$$p_k(\xi) = p_0(2^{-k}\xi)$$

We also set

$$P_{<k} = \sum_{j<k} P_j, \qquad P_{>k} = \sum_{j>k} P_j$$

By $\tilde{P}_k$ we denote a multiplier with slightly larger support than P_k whose symbol equals 1 near the support of P_k. A similar convention is used for all "tilde" operators such as $\tilde{P}_{<k}$, $\tilde{P}_{>k}$, etc.

Alternatively we need a continuous Littlewood-Paley decomposition

$$1 = \int_{\mathbb{R}} P_h dh$$

where the symbols $p_h(\xi)$ of P_h are localized in the region $\{2^{h-1} < |\xi| < 2^{h+1}\}$ and coincide up to scaling,

$$p_h(\xi) = p_0(2^{-h}\xi)$$

In this case we define

$$P_{<h} = \int_{-\infty}^{h} P_l dl, \qquad P_{>h} = \int_{h}^{\infty} P_l dl$$

There is no danger of confusion as we do not use both notations in the same section.

We also use a similar decomposition with respect to the modulation,

$$1 = \sum_{k \in \mathbb{Z}} Q_k$$

where the symbols $q_k(|\xi|, |\tau|)$ of Q_k are localized in the region $\{2^{k-1} < ||\xi| - |\tau|| < 2^{k+1}\}$. Some care is required in order for these symbols to be smooth at $\tau = 0$. We similarly set

$$Q_{<k} = \sum_{j<k} Q_j, \qquad Q_{>k} = \sum_{j>k} Q_j$$

On occasion we replace this with a continuous Littlewood-Paley decomposition

$$1 = \int_{\mathbb{R}} Q_h dh$$

as above.

8

2.4. Frequency localized functions.
Given a function ϕ we use the notation ϕ_h for $P_h\phi$. Similarly we use the notations $\phi_{<h}$, $\phi_{>h}$. For the gradient potentials A and for the gauge transformations U we use the same notation with a slightly different meaning; precisely, A_k and U_k are simply functions which are localized at frequency $|\xi| \approx 2^k$.

2.5. The $\dot{X}^{s,b}$ spaces.
are Sobolev type spaces adapted to the wave operator. For a test function ϕ we define

$$\|\phi\|_{\dot{X}^{s,b}} = \|\hat{\phi}(|\tau| + |\xi|)^s||\tau| - |\xi||^b\|_{L^2}$$

To define the $\dot{X}^{s,b}$ space we take the completion with respect to this norm. This is well defined as a space of distributions if

$$b < \frac{1}{2}, \qquad s + b < \frac{n+1}{2}$$

and contains all the test functions provided that

$$b > -\frac{1}{2}, \qquad s + b > -\frac{n+1}{2}$$

Unfortunately we would like to work with the space $\dot{X}^{\frac{n}{2},\frac{1}{2}}$ for solutions to the wave equation, respectively $\dot{X}^{\frac{n}{2}-1,-\frac{1}{2}}$ for the inhomogeneous term, neither of which are in the above range. This shows that we cannot fully rely on these spaces. However, we can use some modifications of them as a rough framework on which we superimpose a finer structure.

For $1 \le q \le \infty$ we add a Besov type structure with respect to the modulation and set

$$\|\phi\|_{\dot{X}^{s,b,q}}^2 = \sum_{k\in\mathbb{Z}} \left(\sum_{j\in\mathbb{Z}} \|P_kQ_j\phi\|_{\dot{X}^{s,b}}^q \right)^{\frac{2}{q}}$$

Then we want to use a function space S for solutions to the wave equation which satisfies

$$\dot{X}^{\frac{n}{2}-1,\frac{1}{2},1} \subset \nabla S \subset \dot{X}^{\frac{n}{2}-1,\frac{1}{2},\infty}$$

while for the inhomogeneous term we would like to have a space N so that

$$\dot{X}^{\frac{n}{2}-1,-\frac{1}{2},1} \subset N \subset \dot{X}^{\frac{n}{2}-1,-\frac{1}{2},\infty}$$

2.6. **Translation invariant operators.** Following Tao [18] we denote by L any linear or multilinear translation invariant operator with bounded mass,

$$L(\phi_1, \cdots, \phi_k)(x) = \int \phi_1(x + y_1) \cdots \phi_k(x + y_k) d\mu(y), \qquad \int |d\mu| \lesssim 1$$

Since all function spaces used in this paper are translation invariant, operators of this type do not affect multilinear multiplicative estimates. Precisely, if for some translation invariant function spaces $X, X_1, \cdots, X_k$ one has a multiplicative estimate

$$\|\phi_1 \cdots \phi_k\|_X \lesssim \|\phi_1\|_{X_1} \cdots \|\phi_k\|_{X_k}$$

then for any such operator L we also have

$$\|L(\phi_1, \cdots, \phi_k)\|_X \lesssim \|\phi_1\|_{X_1} \cdots \|\phi_k\|_{X_k}$$

Operators of this type are called disposable, and can be neglected in estimates.

2.7. **Constants.** We reserve the letter ϵ for the size of the wave maps in the main space S introduced later on. The initial data will be chosen so that its $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$ size is less than ϵ^2.

We denote by $\delta_0 > 0$ the gain in several of the bilinear estimates. Ideally one would hope to have δ_0 close to $\frac{1}{4}$, but it is not clear that this can be achieved. The constants δ_1, δ_2 satisfy $0 < \delta_2 < \delta_1 \ll \delta$ and are used to define the class of admissible weights for Sobolev spaces of initial data.

By N we denote a sufficiently large integer. Its size depends on δ_0 and on the dimension n.

Also C is a sufficiently large integer. For most of the paper, $C = 4$ is sufficiently large. However, in the proof of the Moser estimates in the last section the choice of C depends on N and (therefore on δ_0 and n). There we roughly need $C \gg \log N$.

3. Initial data sets

The definition of the critical Sobolev spaces for the initial data sets in the previous section is very convenient for our analysis, but it has the disadvantage that it may depend on the choice of the isometric embedding. In this section we remedy this by providing an invariant definition of the Sobolev spaces. We consider as well the related questions of approximating rough initial data sets with smooth initial data sets and of extending local data to global data.

3.1. **Invariantly defined initial data sets.** We interpret the rough initial data sets as trace spaces for families of smooth functions. We consider smooth functions

$$\tilde{\phi} : \mathbb{R} \times \mathbb{R}^n \to TM$$

which we think of as one parameter families of initial data sets,

$$h \to \tilde{\phi}(h) = (\tilde{\phi}_0(h), \tilde{\phi}_1(h)), \qquad \tilde{\phi}_1(h, x) \in T_{\tilde{\phi}_0(h,x)}M$$

By $\mathcal{A}$ we denote the set of such functions which satisfy bounds of the form

$$|\mathbf{D}_x^\beta \mathbf{D}_h^j \tilde{\phi}_0(h, x)| \leq 2^{|\beta|h} c_{j,\beta}, \qquad j = 0, 1, \quad |\beta| + j \geq 1$$

$$|\mathbf{D}_x^\beta \mathbf{D}_h^j \tilde{\phi}_1(h, x)| \leq 2^{(|\beta|-1)h} c_{j,\beta}, \qquad j = 0, 1$$

Let N be a large positive integer. We consider the subset $\mathcal{H}$ of the functions in $\mathcal{A}$ for which the following "norm" is finite,

$$\|\tilde{\phi}\|_{\mathcal{H}}^2 = \int_{\mathbb{R}} \|\mathbf{D}_h \tilde{\phi}(h)\|_{H,h}^2 dh$$

where

$$\|\mathbf{D}_h \tilde{\phi}(h)\|_{\mathcal{H},h}^2 = \int_{\mathbb{R}^n} \sum_{0 \leq |\beta| \leq N} 2^{2(\frac{n}{2}-|\beta|-1)h}(|\mathbf{D}_x^\beta \mathbf{D}_x \mathbf{D}_h \tilde{\phi}_0|^2 + |\mathbf{D}_x^\beta \mathbf{D}_h \tilde{\phi}_1|^2)$$
$$+ \quad 2^n |\mathbf{D}_h \tilde{\phi}_0|^2 dx, \qquad\qquad n \geq 3$$

respectively

$$\|\mathbf{D}_h \tilde{\phi}(h)\|_{\mathcal{H},h}^2 = \int_{\mathbb{R}^n} \sum_{0 \leq |\beta| \leq N} 2^{2(\frac{n}{2}-|\beta|-1)h}(|\mathbf{D}_x^\beta \mathbf{D}_x \mathbf{D}_h \tilde{\phi}_0|^2 + |\mathbf{D}_x^\beta \mathbf{D}_h \tilde{\phi}_1|^2)$$
$$+ \quad 2^{nh}|\mathbf{D}_h \tilde{\phi}_0|^2 dx + 2^{nh}\|\mathbf{D}_h \tilde{\phi}_1(h)\|_{\mathbf{D}L^2+L^1}^2 \qquad n = 2$$

In the last formula we set

$$\|\mathbf{D}_h \tilde{\phi}_1(h)\|_{\mathbf{D}L^2+L^1} = \inf_{\mathbf{D}_h \tilde{\phi}_1(h)=\mathbf{D}_x V + W} \|V\|_{L^2} + \|W\|_{L^1}$$

where $V_1, \cdots, V_n, W$ are assumed to be smooth sections of the pull-back bundle $\phi_0(h)^*(TM)$.

If $\tilde{\phi} \in \mathcal{H}$ and $s > \frac{n}{2}$ we also denote

$$\|\tilde{\phi}\|_{\mathcal{H}^s}^2 = \int_{\mathbb{R}} 2^{(2s-n)h} \|D_h \tilde{\phi}(h)\|_{\mathcal{H},h}^2 dh$$

We want to define the initial data sets as the limits as $h \to \infty$ of functions in $\mathcal{H}$. However, before we do that we need to show that these limits exist.

11

Proposition 3.1. *Let $\tilde{\phi} \in \mathcal{H}$. Then there is an unique measurable function*

$$\phi : \mathbb{R}^n \to TM$$

so that

(3.1)
$$\lim_{h \to \infty} \int_{\mathbb{R}^n} d(\phi(x), \tilde{\phi}(h, x))^2 dx = 0$$

where the distance is measured with respect to the induced Riemmanian metric in TM.

This shows that the following definition is meaningful.

Definition 3.2. a) An initial data set $\phi = (\phi_0, \phi_1)$ belongs to the space $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)$ if there is some $\tilde{\phi} \in \mathcal{H}$ so that

$$\phi(x) = \lim_{h \to \infty} \tilde{\phi}(h, x)$$

in the sense of (3.1).
  b) The $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)$ size of initial data set is

$$\|\phi\|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)} = \inf\{\|\tilde{\phi}\|_{\mathcal{H}};\ \phi = \lim_{h \to \infty} \tilde{\phi}(h)\ \}$$

  c) For $s > \frac{n}{2}$ the $\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, M)$ size of the initial data set is

$$\|\phi\|_{\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, M)} = \inf\{\|\tilde{\phi}\|_{\mathcal{H}^s};\ \phi = \lim_{h \to \infty} \tilde{\phi}(h)\ \}$$

The proof of Proposition 3.1 is quite simple for $n \geq 3$, but it becomes more difficult for $n = 2$. As in the latter case we need some preliminaries, we begin with the high dimensional case.

Proof of Proposition 3.1, $n \geq 3$. For this we only use the bound

$$\int_{\mathbb{R}} \int_{\mathbb{R}^n} 2^{\frac{n}{2}h} |\mathbf{D}_h \tilde{\phi}_0|^2 + 2^{(\frac{n}{2}-1)h} |\mathbf{D}_h \tilde{\phi}_1|^2 dx dh \lesssim \|\tilde{\phi}\|_{\mathcal{H}}^2$$

This shows that the integral

$$\int_0^\infty |\mathbf{D}_h \tilde{\phi}_0(h, x)| + |\mathbf{D}_h \tilde{\phi}_1(h, x)| dh$$

converges exponentially in L^2, therefore converges pointwise a.e. Let x be a point of convergence. Then the curve

$$\tilde{\phi}(h, x) \qquad h \in [0, \infty)$$

has finite length therefore the limit

$$\phi(x) = \lim_{h \to \infty} \tilde{\phi}(h, x)$$

exists and we can compute

$$d(\phi(x), \tilde{\phi}(h_0, x)) \lesssim \int_{h_0}^{\infty} |\mathbf{D}_h \tilde{\phi}_0(h, x)| + |\mathbf{D}_h \tilde{\phi}_1(h, x)| dh$$

$$\lesssim 2^{(1-\frac{n}{2})h} \left(\int_{h_0}^{\infty} 2^{\frac{n}{2}h} |\mathbf{D}_h \tilde{\phi}_0|^2 + 2^{(\frac{n}{2}-1)h} |\mathbf{D}_h \tilde{\phi}_1|^2 dh \right)^{\frac{1}{2}}$$

We square and integrate to obtain

$$\int_{\mathbb{R}^n} d(\phi(x), \tilde{\phi}(h_0, x))^2 dx \lesssim 2^{(2-n)h} \|\tilde{\phi}\|_{\mathcal{H}}^2$$

which concludes the proof. □

Now we take $\tilde{\phi} \in \mathcal{H}$ and establish L^p bounds for covariant derivatives of $\mathbf{D}_h \tilde{\phi}$. This is nothing but a modification of the usual Sobolev embeddings.

Lemma 3.3. *Let $\tilde{\phi} \in \mathcal{H}$ and $2 \le p \le \infty$. Then for $|\beta| < N - O(n)$ we have*

(3.2) $$\|\mathbf{D}_x^{\beta} \mathbf{D}_h \tilde{\phi}_0(h)\|_{L^p} \lesssim 2^{-(\frac{n}{p}-|\beta|)h} \|\mathbf{D}_h \tilde{\phi}\|_{\mathcal{H}, h}$$

(3.3) $$\|\mathbf{D}_x^{\beta} \mathbf{D}_h \tilde{\phi}_1(h)\|_{L^p}^{\lesssim} 2^{-(\frac{n}{p}-1-|\beta|)h} \|\mathbf{D}_h \tilde{\phi}\|_{\mathcal{H}, h}$$

Proof. Both inequalities follow from a bound for fixed h of the type

(3.4) $$\|\mathbf{D}_x^{\beta} V\|_{L^p} \lesssim 2^{(-\frac{n}{p}+|\beta|)h} \sum_{|\gamma|=0}^{N} 2^{(|\gamma|-\frac{n}{2})} \|\mathbf{D}_x^{\gamma} V\|_{L^2}$$

For this we use the Sobolev inequalities

$$\|v\|_{L^p} \lesssim \|\nabla v\|_{L^q}, \quad \frac{n}{p} = \frac{n}{q} - 1, \quad 1 \le q < n, \quad v \in L^q$$

respectively

$$\|v\|_{BMO} \lesssim \|\nabla v\|_{L^n}$$

We apply this to the function

$$v = |V|, \qquad \nabla v = |V|^{-1} \langle \mathbf{D}_x V, V \rangle$$

to get

$$\|V\|_{L^p} \lesssim \|\mathbf{D}_x V\|_{L^q}, \quad \frac{n}{p} = \frac{n}{q} - 1, \quad 1 \le q < n, \quad |V| \in L^q$$

respectively

$$\||V|\|_{BMO} \lesssim \|\nabla V\|_{L^n}$$

If $n \ge 3$ then we apply this with V replaced by $\mathbf{D}_x^{\beta} V$ for $q = 2$ yields the correct bound for $p = \frac{2n}{n-2}$. Interpolation with L^2 gives the range $2 \le p \le p = \frac{2n}{n-2}$. Iterating this argument we get all the values of p.

13

If $n = 2$ then we use the second inequality to obtain a BMO bound instead. This we interpolate with L^2 to get all indices $2 \le p < \infty$.

Finally, the L^∞ bound follows from an interpolation inequality of the form

$$\|v\|_{L^\infty} \lesssim \|v\|_{L^p}^\theta \|\nabla v\|_{L^q}^{1-\theta}, \qquad \theta \frac{n}{p} + (1-\theta)(\frac{n}{q} - 1) = 0$$

$\square$

The second step is to obtain bounds for the functions $\tilde{\phi}(h, \cdot)$. To bound the covariant derivatives of $\tilde{\phi}(h, \cdot)$ we use L^p spaces, as before. However, to measure $\tilde{\phi}_0(h, \cdot)$ itself we introduce the set $BMO(\mathbb{R}^n, M)$ of measurable functions

$$\phi : \mathbb{R}^n \to M$$

for which the following quantity is finite:

$$\|\phi\|_{BMO(\mathbb{R}^n, M)} = \sup_{R>0} \sup_{x_0 \in \mathbb{R}^n} \inf_{y \in M} \left(\frac{1}{|B(x_0, R)|} \int_{B(x_0, R)} d(\phi(x), y)^2 dx \right)^{\frac{1}{2}}$$

Here $d(.,.)$ is the distance function on M. In what follows we denote by $A(\tilde{\phi})$ a constant which may depend at most polynomially on $\|\tilde{\phi}\|_{\mathcal{H}}$.

Proposition 3.4. *Let $\tilde{\phi} \in \mathcal{H}$. Then for $|\beta| < N - O(n)$, $p \ge 2$ and $|\beta| \ge \frac{n}{p}$ we have*

(3.5) $$\|\mathbf{D}_x^\beta \tilde{\phi}_0(h)\|_{L^p} \lesssim 2^{(|\beta| - \frac{n}{p})h} A(\tilde{\phi}) \|\tilde{\phi}\|_{\mathcal{H}}, \qquad |\beta| \ge 1$$

(3.6) $$\|\mathbf{D}_x^\beta \tilde{\phi}_1(h)\|_{L^p} \lesssim 2^{(|\beta| + 1 - \frac{n}{p})h} A(\tilde{\phi}) \|\tilde{\phi}\|_{\mathcal{H}},$$

In addition,

(3.7) $$\|\tilde{\phi}_0(h)\|_{BMO} \lesssim A(\tilde{\phi}) \|\tilde{\phi}\|_{\mathcal{H}}$$

Proof. It suffices to prove the bounds for $\tilde{\phi}_0$, those for $\tilde{\phi}_1$ are obtained in a similar fashion. We consider three increasingly difficult cases.

Case 1: $|\beta| > n/p$. We use induction with respect to $|\beta|$ and begin with the case $|\beta| = 1$. Then we have

$$\mathbf{D}_h \mathbf{D}_x \tilde{\phi}_0 = \mathbf{D}_x \mathbf{D}_h \tilde{\phi}_0$$

Hence

(3.8) $$|\mathbf{D}_x \tilde{\phi}_0(h)| \le \int_{-\infty}^h |\mathbf{D}_x \mathbf{D}_h \tilde{\phi}_0(l)| dl$$

Then (3.5) follows directly from (3.2) and the Cauchy-Schwartz inequality.

14

It is also useful to consider the case $|\beta| = 2$. Then

$$\mathbf{D}_h \mathbf{D}_x^2 \tilde{\phi}_0 = \mathbf{D}_x^2 \mathbf{D}_h \tilde{\phi}_0 + R(\mathbf{D}_h \tilde{\phi}_0, \mathbf{D}_x \tilde{\phi}_0) \mathbf{D}_x \tilde{\phi}_0$$

We integrate this and use the L^∞ bound for $\mathbf{D}_x \tilde{\phi}_0$ to obtain

$$\|\mathbf{D}_h \mathbf{D}_x^2 \tilde{\phi}_0(h)\|_{L^p} \lesssim \int_{-\infty}^{h} \|\mathbf{D}_x^2 \mathbf{D}_h \tilde{\phi}_0(l)\|_{L^p} + 2^{2l} A(\tilde{\phi}) \|\mathbf{D}_h \tilde{\phi}_0(l)\|_{L^p} dl$$

Now we can use (3.2) again.

Finally, we do the induction step. Repeated commutations lead to

$$|\mathbf{D}_h \mathbf{D}_x^\beta \tilde{\phi}_0| \lesssim |\mathbf{D}_x^\beta \mathbf{D}_h \tilde{\phi}_1| + \sum_{\substack{\beta_0 + \cdots + \beta_k = \beta}}^{|\beta_i| < |\beta|} |\mathbf{D}_x^{\beta_0} \mathbf{D}_h \tilde{\phi}_0| |\mathbf{D}_x^{\beta_1} \tilde{\phi}_0| \cdots |\mathbf{D}_x^{\beta_k} \tilde{\phi}_0|$$

We bound all factors but the one containing $\mathbf{D}_h$ in L^∞ and use again (3.2) for $\mathbf{D}_h \tilde{\phi}_0$.

Case 2: $|\beta| = n/p$, $p > 2$. Here we can no longer use the Cauchy-Schwartz inequality to take advantage of (3.2) because we loose the exponential decay. Instead we use the Littlewood Paley theory. The main bound we need is summarized in the following

Lemma 3.5. *Let $2 \leq q < p < \infty$ and $\theta \in (0,1)$ so that*

$$\frac{n}{p} = \frac{n}{q} - \theta$$

Then

$$\left\| \int_{-\infty}^{\infty} f(h) dh \right\|_{L^p}^2 \lesssim \int_{-\infty}^{\infty} 2^{2\theta h} \|f(h)\|_{L^q}^2 + 2^{2(\theta-1)h} \|\nabla_x f(h)\|_{L^q}^2$$

Consider the case when $|\beta| = 1$ which corresponds to $p = n \geq 3$. Then we apply the above lemma for the integral in (3.8), with

$$f(h) = |\mathbf{D}_x \mathbf{D}_h \tilde{\phi}_0(l)|, \qquad |\nabla f(h)| = |\mathbf{D}_x^2 \mathbf{D}_h \tilde{\phi}_0(l)|$$

The choice of q is not important, any q satisfying the relations in the Lemma is appropriate. A similar argument applies for other values of $|\beta|$.

Case 3: $|\beta| = n/2$, $p = 2$. This requires n to be even. In order to keep the argument simple we focus on the most interesting case $n = 2$ and leave the rest for the reader. In this case we prove the L^2 bound for $\tilde{\phi}_1$; the L^2 bound for $\mathbf{D} \tilde{\phi}_0$ is obtained in an identical fashion.

We denote by $S(h, l)$ the parallel transport from l to h along the curves $h \to \tilde{\phi}_0(h, x)$. Then we write

$$
\begin{aligned}
\frac{1}{2}\frac{d}{dh}|\tilde{\phi}_1(h)|^2 &= \langle \mathbf{D}_h \tilde{\phi}_1(h), \tilde{\phi}_1(h) \rangle \\
&= \langle \mathbf{D}_h \tilde{\phi}_1(h), \int_{-\infty}^{h} \frac{d}{dl}[S(h, l)\tilde{\phi}_1(l)]dl \rangle \\
&= \int_{-\infty}^{h} \langle \mathbf{D}_h \tilde{\phi}_1(h), S(h, l)\mathbf{D}_l\tilde{\phi}_1(l) \rangle dl \\
&= \int_{-\infty}^{h} \langle S(l, h)\mathbf{D}_h \tilde{\phi}_1(h), \mathbf{D}_l\tilde{\phi}_1(l) \rangle dl
\end{aligned}
$$

This implies that

$$
\|\tilde{\phi}_1(h)\|_{L^2}^2 = 2 \int_{\mathbb{R}^n} \int_{l<m<h} \langle S(l, m)\mathbf{D}_m\tilde{\phi}_1(m), \mathbf{D}_l\tilde{\phi}_1(l) \rangle dl\, dm\, dx
$$

This is easily estimated using the Cauchy-Schwartz inequality provided we can obtain some off-diagonal decay,

$$
| \int_{\mathbb{R}^n} \langle S(l, m)\mathbf{D}_m\tilde{\phi}_1(m), \mathbf{D}_l\tilde{\phi}_1(l) \rangle dx | \lesssim
$$

(3.9)
$$
A(\tilde{\phi})2^{-\frac{1}{2}|m-l|}\|\mathbf{D}_m\tilde{\phi}_1(m)\|_{\mathcal{H},m}\|\mathbf{D}_l\tilde{\phi}_1(l)\|_{\mathcal{H},l}
$$

It remains to prove (3.9). By symmetry we assume $l < m$. According to the two dimensional definition of $\|\cdot\|_{\mathcal{H},m}$ we have the representation

$$
\mathbf{D}_m\tilde{\phi}_1(m) = \mathbf{D}_x V + W, \qquad \|V\|_{L^2} + \|W\|_{L^1} \lesssim 2^{-m}\|\mathbf{D}_m\tilde{\phi}_1(m)\|_{\mathcal{H},m}
$$

The W component of (3.9) is easy to estimate,

$$
\begin{aligned}
| \int_{\mathbb{R}^n} \langle S(l, m)W, \mathbf{D}_l\tilde{\phi}_1(l) \rangle dx | &\lesssim \|W\|_{L^1}\|\mathbf{D}_l\tilde{\phi}_1(l)\|_{L^\infty} \\
&\lesssim 2^{-|m-l|}\|\mathbf{D}_m\tilde{\phi}_1(m)\|_{\mathcal{H},m}\|\mathbf{D}_l\tilde{\phi}_1(l)\|_{\mathcal{H},l}
\end{aligned}
$$

It remains to consider the V component of (3.9). We compute

$$
\begin{aligned}
\partial_k S(l, k)\mathbf{D}_x S(k, m)V &= S(l, k)\mathbf{D}_k\mathbf{D}_x S(k, m)V \\
&= S(l, k)[\mathbf{D}_k, \mathbf{D}_x]S(k, m)V \\
&= S(l, k)R(\mathbf{D}_k\tilde{\phi}_0(k), \mathbf{D}_x\tilde{\phi}_0(k))S(k, m)V
\end{aligned}
$$

Integrating this we obtain

$$
\begin{aligned}
S(l, m)\mathbf{D}_x V = &-\int_{l}^{m} S(l, k)R(\mathbf{D}_k\tilde{\phi}_0(k), \mathbf{D}_x\tilde{\phi}_0(k))S(k, m)V\, dk \\
&+ \mathbf{D}_x S(l, m)V
\end{aligned}
$$

16

The contribution of the second term to (3.9) has the form

$$\left| \int_{\mathbb{R}^n} \langle \mathbf{D}_x S(l,m)V, \mathbf{D}_l \tilde{\phi}_1(l) \rangle dx \right| = \left| \int_{\mathbb{R}^n} \langle S(l,m)V, \mathbf{D}_x \mathbf{D}_l \tilde{\phi}_1(l) \rangle dx \right|$$

and we can bound both V and $\mathbf{D}_x \mathbf{D}_l \tilde{\phi}_1(l)$ in L^2. The contribution of the first term is

$$\left| \int_l^m \int_{\mathbb{R}^n} \langle S(l,k)R(\mathbf{D}_k\tilde{\phi}_0(k), \mathbf{D}_x\tilde{\phi}_0(k))S(k,m)V, \mathbf{D}_l\tilde{\phi}_1(l) \rangle dx dk \right|$$

This is estimated by

$$\int_l^m \|\mathbf{D}_k\tilde{\phi}_0(k)\|_{L^\infty} \|\mathbf{D}_x\tilde{\phi}_0(k)\|_{L^4} \|V\|_{L^2} \|\mathbf{D}_l\tilde{\phi}_1(l)\|_{L^4} dk$$

For the first and the last factor we use (3.3) while for $\mathbf{D}_x\tilde{\phi}_0(k)$ we use the nonsharp case of (3.5). Then the above quantity is estimated by

$$2^{-m}A(\tilde{\phi})\|\mathbf{D}_m\tilde{\phi}_1\|_{\mathcal{H},m} 2^{\frac{l}{2}} \|\mathbf{D}_l\tilde{\phi}_1(l)\|_{\mathcal{H},l} \int_l^m 2^{\frac{k}{2}} \|\mathbf{D}_k\tilde{\phi}_1(k)\|_{\mathcal{H},k} dk$$

After using Cauchy Schwartz with respect to k this is bounded by the right hand side of (3.9).

Case 4: The BMO bound. By rescaling it suffices to consider a cube Q of size 1. Integrating the L^2 bound for $\mathbf{D}_h\tilde{\phi}_1$ yields the bound

$$\|d(\tilde{\phi}_0(h), \tilde{\phi}_0(0))\|_{L^2} \lesssim \|\tilde{\phi}\|_{\mathcal{H}}, \qquad h \geq 0$$

Hence without any restriction in generality we can take $h \leq 0$. But then we control the Lipschitz constant of $\tilde{\phi}(h)$, and this concludes the proof. $\qquad\square$

We are now able to conclude the proof of Proposition 3.1.

Proof of Proposition 3.1, $n = 2$. To show that the limit

$$\lim_{h \to \infty} \tilde{\phi}(h, x)$$

exists in the L^2 sense with respect to the distance on TM we verify the Cauchy property in the metric space $L^2(TM)$. We have

$$d(\tilde{\phi}(h, x), \tilde{\phi}(l, x)) \lesssim \int_l^h |\mathbf{D}_h\tilde{\phi}_0(l)| dl + |\tilde{\phi}_1(h) - S(h,l)\tilde{\phi}_1(l)|$$

We square, integrate in x and use Cauchy-Schwartz for the first term to obtain

$$\int_{\mathbb{R}^n} d(\tilde{\phi}(h, x), \tilde{\phi}(l, x))^2 dx \lesssim 2^{-2l} \|\tilde{\phi}\|_{\mathcal{H}} + \int_{\mathbb{R}^n} |\tilde{\phi}_1(h) - S(h,l)\tilde{\phi}_1(l)|^2 dx$$

17

The first term decays as $h, l \to \infty$, it remains to consider the second. Arguing as in Case 3 of the previous Proposition, we represent it as

$$2 \int_{\mathbb{R}^n} \int_{l<k<m<h} \langle S(k,m) \mathbf{D}_m \tilde{\phi}_1(m), \mathbf{D}_k \tilde{\phi}_1(k) \rangle \, dk \, dm \, dx$$

By (3.9) combined with the Cauchy-Schwartz inequality we bound this by

$$A(\tilde{\phi}) \int_l^h \| \mathbf{D}_k \tilde{\phi}_1(k) \|_{\mathcal{H},h}^2 \, dk$$

For $\tilde{\phi} \in \mathcal{H}$ this converges to 0 as $h, l \to \infty$. $\qquad \square$

3.2. **Regularizations and the equivalence of definitions.** Our task here is to relate the above definition of $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$ with the earlier definition in the first section. As it turns out, as long as the initial data is small in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$ the two definitions remain equivalent.

Proposition 3.6. *Suppose that M is uniformly isometrically embedded in $\mathbb{R}^d$. Then*

a) Any initial data set (ϕ_0, ϕ_1) which is in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)$ is also in the space $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, \mathbb{R}^d)$ and

$$\| (\phi_0, \phi_1) \|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, \mathbb{R}^d)} \lesssim A(\phi) \| (\phi_0, \phi_1) \|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)}$$

If in addition $(\phi_0, \phi_1) \in \dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, M)$ then it is also in the space $\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, \mathbb{R}^d)$ and

$$\| (\phi_0, \phi_1) \|_{\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, \mathbb{R}^d)} \lesssim A(\phi) \| (\phi_0, \phi_1) \|_{\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, M)}$$

In both cases $A(\phi)$ depends polynomially on the $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)$ size of ϕ.

b) Any initial data set (ϕ_0, ϕ_1) which is small in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, \mathbb{R}^d)$ is also in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)$ and

$$\| (\phi_0, \phi_1) \|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)} \lesssim \| (\phi_0, \phi_1) \|_{H^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, \mathbb{R}^d)}$$

If in addition (ϕ_0, ϕ_1) is in $\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, \mathbb{R}^d)$ then it is also in the space $\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, M)$ and

$$\| (\phi_0, \phi_1) \|_{\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, M)} \lesssim \| (\phi_0, \phi_1) \|_{\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, \mathbb{R}^d)}$$

Proof. a) Let $\tilde{\phi}$ be as in Definition 3.2, with almost optimal $\mathcal{H}$ size. Our first task is to express regular derivatives in terms of covariant derivatives.

18

Lemma 3.7. *Let V be a smooth section of $\tilde{\phi}_0^*(TM)$. Then there is a representation*

$$\mathbf{D}^\beta V(h) = \partial^\beta V(h) + \sum_{k \geq 1} \sum_{\beta_0 + \cdots + \beta_k = \beta}^{\beta_1, \cdots \beta_k \neq 0} g(\tilde{\phi}(h)) \partial^{\beta_0} V(h) \partial^{\beta_1} \tilde{\phi}(h) \cdots \partial^{\beta_k} \tilde{\phi}(h)$$

and

$$\partial^\beta V(h) = \mathbf{D}^\beta V(h) + \sum_{k \geq 1} \sum_{\beta_0 + \cdots + \beta_k = \beta}^{\beta_1, \cdots \beta_k \neq 0} g(\tilde{\phi}(h)) \mathbf{D}^{\beta_0} V(h) \mathbf{D}^{\beta_1} \tilde{\phi}(h) \cdots \mathbf{D}^{\beta_k} \tilde{\phi}(h)$$

where g stands for smooth bounded functions on M with uniformly bounded derivatives.

Proof. This follows from the representation of the covariant derivatives in local coordinates,

$$\mathbf{D}_i V = \partial_i V - S_{\tilde{\phi}_0(h)}(\nabla \tilde{\phi}_0(h), V)$$

where S is the second fundamental form. In order to be able to iterate this, we extend $S_{\tilde{\phi}_0(h)}$ to a smooth bilinear form on all of $\mathbb{R}^d$. This yields the first representation. For the second we invert the first. $\quad\square$

Lemma 3.8. *Let $\tilde{\phi} \in \mathcal{H}$. Then for small β we have*

$$\|\partial_x^\beta \partial_h \tilde{\phi}_0(h)\|_{L^2} \lesssim 2^{(|\beta| - \frac{n}{2})h} A(\tilde{\phi}) \|\mathbf{D}_h \tilde{\phi}(h)\|_{\mathcal{H},h}$$

$$\|\partial_x^\beta \partial_h \tilde{\phi}_1(h)\|_{L^2} \lesssim 2^{(|\beta| + 1 - \frac{n}{2})h} A(\tilde{\phi}) \|\mathbf{D}_h \tilde{\phi}(h)\|_{\mathcal{H},h}$$

In addition, if $n = 2$ then

$$\|\tilde{\phi}_1(h)\|_{\partial L^2 + L^1} \lesssim 2^{-h} A(\tilde{\phi}) \|\mathbf{D}_h \tilde{\phi}(h)\|_{\mathcal{H},h}$$

Proof. This follows from the representation in Lemma 3.7 combined with L^2 bounds for $\mathbf{D}_h \tilde{\phi}$ and its covariant derivatives, and the L^∞ bounds on the covariant derivatives $\tilde{\phi}$ in Proposition 3.4. $\quad\square$

Now we can conclude the proof of part (a) of Proposition 3.6. We write

$$\partial_x \tilde{\phi}_0(h_0) = \int_{-\infty}^h \partial_x \partial_h \tilde{\phi}_0(h) dh$$

$$\tilde{\phi}_1(h_0) = \int_{-\infty}^h \partial_h \tilde{\phi}_1(h) dh$$

For $h \in [j, j+1]$ the functions $\partial_x \partial_h \phi_0(h), \partial_h \phi_1(h)$ are essentially concentrated at frequency 2^j, therefore using the previous Lemma we obtain the bound

$$\|\partial_x \tilde{\phi}_0(h_0)\|_{\dot{H}^{\frac{n}{2}-1}} + \|\tilde{\phi}_0(h_0)\|_{\dot{H}^{\frac{n}{2}-1}} \lesssim \|\tilde{\phi}\|_{\mathcal{H}}$$

19

Letting $h_0 \to \infty$ we obtain the desired bound. A similar argument yields the $\dot{H}^s \times \dot{H}^{s-1}$ estimate.

Proof of Proposition 3.6,(b). We note first that if $M = \mathbb{R}^d$ then each initial data set $(\phi_0, \phi_1) \in \dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$ is the trace as $h \to \infty$ of the $\mathcal{H}$ function

$$\tilde{\phi}(h) = P_{<h}\phi$$

If $M \subset \mathbb{R}^d$ then this construction fails because there is no guarantee that the regularized functions ϕ_h take values in M. To remedy this we need to project these regularized functions into M. This is the aim of the next lemma.

Lemma 3.9. *Let $(\phi_0, \phi_1) \in \dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, \mathbb{R}^d)$ be a small initial data set for the wave maps equation. Then there exists a smooth function*

$$\tilde{\phi} : \mathbb{R} \times \mathbb{R}^n \to M \subset \mathbb{R}^d$$

satisfying

(3.10)
$$|\partial_h^j \partial_x^\beta \tilde{\phi}_0(h, x)| \leq 2^{|\beta|h} c_{j,\beta}, \quad |\beta| + j \geq 1$$

(3.11)
$$|\partial_h^j \partial_x^\beta \tilde{\phi}_1(h, x)| \leq 2^{(|\beta|-1)h} c_{j,\beta}, \quad |\beta| + j \geq 0$$

so that

(i) For any nonnegative multiindex β we have

$$\int_{\mathbb{R}^n} 2^{(\frac{n}{2}-|\beta|)h} |\partial^\beta \partial_h \tilde{\phi}_0|^2 + 2^{(\frac{n}{2}-|\beta|-1)h} |\partial^\beta \partial_h \tilde{\phi}_1|^2 dx \leq$$

$$c_\beta \|P_h(\phi_0, \phi_1)\|^2_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, \mathbb{R}^d)}$$

In addition, if $n = 2$ then

$$2^h \|\partial_h \tilde{\phi}_1\|^2_{\partial L^2 + L^1} \lesssim \|P_h(\phi_0, \phi_1)\|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, \mathbb{R}^d)}$$

(ii) $(\tilde{\phi}_0(h), \tilde{\phi}_1(h))$ are uniformly small in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, \mathbb{R}^d)$ and

$$\lim_{h\to\infty} (\tilde{\phi}_0(h), \tilde{\phi}_1(h)) = (\phi_0, \phi_1) \quad in \quad \dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$$

Proof. We begin with the regularized functions

$$(P_{<h}\phi_0, P_{<h}\phi_1)$$

which in general do not take values into the the manifold M. However, we claim that

(3.12)
$$d(P_{<h}\phi_0, M) \lesssim \|\phi_0\|_{\dot{H}^{\frac{n}{2}}}$$

Indeed, it is easy to see that

$$\|P_{<h}\phi_0 - \phi_0\|_{L^2} \lesssim 2^{-\frac{nh}{2}} \|\phi_0\|_{\dot{H}^{\frac{n}{2}}}$$

Combining this with

$$\|\partial P_{<h}\phi_0\|_{L^\infty} \lesssim 2^{-h}\|\phi_0\|_{\dot{H}^{\frac{n}{2}}}$$

we obtain the BMO type bound

$$\fint_{B(x,2^{-h})} |P_{<h}\phi_0(x) - \phi_0(y)|^2 \lesssim 2^{-nh}\|\phi_0\|_{\dot{H}^{\frac{n}{2}}}$$

which implies (3.12).

In a small tubular neighborhood $\tilde{M}$ of M we define the projection operator

$$\Pi : T\tilde{M} \to TM$$

which is the normal projection onto M in the physical space, respectively the euclidean orthogonal projection in the fiber. Then Π is uniformly smooth with respect to both variables, and linear in the second variable. Hence we set

$$(\tilde{\phi}_0(h), \tilde{\phi}_1(h)) = \Pi(P_{<h}\phi_0, P_{<h}\phi_1)$$

Linearizing Π we get

$$\partial_h\tilde{\phi}_0(h) = g(P_{<h}\phi_0)P_h\phi_0,$$

(3.13) $$\partial_h\tilde{\phi}_1(h) = g(P_{<h}\phi_0)P_h\phi_1 + g(P_{<h}\phi_0)P_{<h}\phi_1 P_h\phi_0$$

where g stands for smooth bounded functions with bounded derivatives. The derivatives of $\partial_h((\tilde{\phi}_0(h), \tilde{\phi}_1(h))$ are computed using the chain rule,

$$\partial^\beta \partial_h\tilde{\phi}_0(h) = \sum_{k\geq 0} \sum_{\beta_0+\cdots+\beta_k=\beta}^{\beta_1,\cdots,\beta_k\neq 0} g(P_{<h}\phi_0)\partial^{\beta_0}P_h\phi_0\,\partial^{\beta_1}P_{<h}\phi_0\cdots\partial^{\beta_k}P_{<h}\phi_0$$

$$\partial^\beta \partial_h\tilde{\phi}_1(h) = \sum_{k\geq 0} \sum_{\beta_0+\cdots+\beta_k=\beta}^{\beta_1,\cdots,\beta_k\neq 0} g(P_{<h}\phi_0)\partial^{\beta_0}P_h\phi_1\,\partial^{\beta_1}P_{<h}\phi_0\cdots\partial^{\beta_k}P_{<h}\phi_0$$

$$+ \sum_{k\geq 1} \sum_{\beta_0+\cdots+\beta_k=\beta}^{\beta_2,\cdots,\beta_k\neq 0} g(P_{<h}\phi_0)\partial^{\beta_0}P_h\phi_0\,\partial^{\beta_1}P_{<h}\phi_1\cdots\partial^{\beta_k}P_{<h}\phi_0$$

The L^∞ bounds in (3.10) follow easily by estimating each factor in L^∞. For the L^2 bounds in (i) we use the L^2 bounds for the frequency localized functions $P_h(\phi_0, \phi_1)$ and L^∞ bounds for everything else.

Consider also the two dimensional estimate in (ii). The second term in (3.13) estimated in L^1,

$$\|g(P_{<h}\phi_0)P_{<h}\phi_1 P_h\phi_0\|_{L^1} \lesssim \|P_{<h}\phi_1\|_{L^2}\|P_h\phi_0\|_{L^2} \lesssim \|P_h\phi_0\|_{L^2}$$

For the first term in (3.13) we write

$$g(P_{<h}\phi_0)P_h\phi_1 = \partial_i(g(P_{<h}\phi_0)\Delta^{-1}\partial_i P_h\phi_1) - g'(P_{<h}\phi_0)\partial_i P_{<h}\phi_0\Delta^{-1}\partial_i P_h\phi_1$$

The first part is the divergence of an L^2 function,

$$\|g(P_{<h}\phi_0)\Delta^{-1}\partial_i P_h\phi_1\|_{L^2} \lesssim 2^{-h}\|P_h\phi_1\|_{L^2}$$

The second is in L^1,

$$\begin{aligned}
\|g'(P_{<h}\phi_0)\partial_i P_{<h}\phi_0\Delta^{-1}\partial_i P_h\phi_1\|_{L^1} &\lesssim 2^{-h}\|\partial P_{<h}\phi_0\|_{L^2}\|\|P_h\phi_1\|_{L^2} \\
&\lesssim 2^{-h}\|P_h\phi_1\|_{L^2}
\end{aligned}$$

(ii) From (i) we see that the limit of $(\tilde\phi_0(h), \tilde\phi_1(h))$ as h tends to infinity exists in $\dot H^{\frac{n}{2}} \times \dot H^{\frac{n}{2}-1}$. Thus in order to see that the limit is (ϕ_0, ϕ_1) it suffices to show that this is the limit in the sense of distributions. For ϕ_0 we note that

$$\|\phi_0 - P_{<h}\phi_0\|_{L^2} \lesssim \lambda^{-\frac{n}{2}}\|\phi_0\|_{\dot H^{\frac{n}{2}}}$$

which by the definition on Π leads to

$$\|\phi_0 - \tilde\phi_0(h)\|_{L^2} \lesssim \lambda^{-\frac{n}{2}}\|\phi_0\|_{\dot H^{\frac{n}{2}}}$$

On the other hand, if k_h is the kernel of P_h then

$$P_{<h}\phi_1(x) = \int_{\mathbb{R}^n} \phi_1(y)k_h(x-y)dy.$$

The angle between $T_{\phi_0(y)}M$ and $T_{\phi(h,x)}M$ is at most $|\phi_0(y) - \tilde\phi_0(h,x)|$, therefore

$$\begin{aligned}
|\tilde\phi_1(h,x) - P_{<h}\phi_1(x)| &\lesssim \int_{\mathbb{R}^n} |\phi_1(y)||\phi_0(y) - \tilde\phi_0(h,x)||k_h(x-y)|dy \\
&\lesssim \int_{\mathbb{R}^n} |\phi_1(y)||\phi_0(y) - \tilde\phi_0(h,y)||k_h(x-y)|dy \\
&\quad + \int_{\mathbb{R}^n} |\phi_1(y)||\tilde\phi_0(h,x) - \tilde\phi_0(h,y)||k_h(x-y)|dy \\
&\lesssim \int_{\mathbb{R}^n} |\phi_1(y)||\phi_0(y) - P_{<h}\phi_0(y)||k_h(x-y)|dy \\
&\quad + \int_{\mathbb{R}^n} |\phi_1(y-z)||P_{<h}\phi_0(x) - P_{<h}\phi_0(x-z)||k_h(z)|dz
\end{aligned}$$

Then

$$\begin{aligned}
&\|\tilde\phi_1(h) - P_{<h}\phi_1\|_{L^{\frac{n}{2}}} \\
&\lesssim \|\phi_1\|_{L^n}(\|P_{>h}\phi_0\|_{L^n} + \int \|\phi_0(x) - \phi_0(x+z)\|_{L^n}|k_h(z)|dz) \\
&\lesssim 2^{-h} + \int |z||k_h(z)|dz \lesssim 2^{-h}
\end{aligned}$$

which proves the desired convergence. $\qquad \square$

To conclude the proof of Proposition 3.6,(b) we simply use Lemma 3.7 to switch from differentiation in $\mathbb{R}^d$ to covariant differentiation in M. The $\dot{H}^s \times \dot{H}^{s-1}$ follows in a similar fashion. $\qquad \square$

3.3. **Initial data sets for the linearized equation.** Consider a smooth initial data set (ϕ_0, ϕ_1) for the wave maps equation which is small in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)$. A smooth initial data set for the linearized equation is a a pair of smooth functions

$$(\psi_0, \psi_1) : \mathbb{R}^n \to T_{\phi_0} M \times T_{\phi_0} M$$

For ϕ as above we denote by $\mathcal{A}(\phi)$ the set of its extensions $\tilde{\phi}$ which are small in $\mathcal{H}$. For $s < \frac{n}{2}$ and $\tilde{\phi} \in \mathcal{A}(\phi)$ we denote by $\mathcal{A}_{\tilde{\phi}}^s(\psi)$ the set of smooth extensions $\tilde{\psi}$ of ψ which satisfy

$$\tilde{\psi}_0(h, x), \tilde{\psi}_1(h, x) \in T_{\tilde{\phi}_0(h,x)} M$$

and, for $j = 0, 1$,

$$|\mathbf{D}_h^j \mathbf{D}_x^\beta \tilde{\psi}_0(h, x)| \leq 2^{(|\beta| + \frac{n}{2} - s)h} c_{j,\beta}, \quad |\mathbf{D}_h^j \mathbf{D}_x^\beta \psi_1(h, x)| \leq 2^{((|\beta| + 1 + \frac{n}{2} - s)h} c_{j,\beta},$$

For $\tilde{\psi} \in \mathcal{A}_{\tilde{\phi}}^s(\psi)$ we denote

$$\|\tilde{\psi}(h)\|_{\mathcal{H}_{lin}^s}^2 = \int_{-\infty}^\infty 2^{(n-2s)h} \|\mathbf{D}_h \psi(h)\|_{\mathcal{H}, h}^2 dh$$

The $\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, TM)$ size of ψ is defined as

$$\|\psi\|_{\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, TM)} = \inf_{\tilde{\phi} \in \mathcal{A}(\phi)} \inf_{\tilde{\psi} \in \mathcal{A}_{\tilde{\phi}}^s(\psi)} \|\tilde{\psi}\|_{\mathcal{H}_{lin}^s}$$

The next result relates this to its euclidean counterpart in the case when M is isometrically embedded in $\mathbb{R}^d$.

Proposition 3.10. *Suppose that M is uniformly isometrically embedded in $\mathbb{R}^d$. Let ϕ be a smooth initial data set which is small in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$. Then for any smooth initial data set ψ for the linearized equation we have*

$$(3.14) \qquad \|\psi\|_{\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, TM)} \approx \|\psi\|_{\dot{H}^s \times \dot{H}^{s-1}(\mathbb{R}^n, \mathbb{R}^d)}$$

Proof. (i) The "$\gtrsim$" part. Let $\tilde{\psi} \in \mathcal{H}_{lin}^s$ be an almost optimal extension of ψ. Arguing as in the proof of Proposition 3.6,(a) it is easy to establish the uniform bounds

$$(3.15) \qquad \|\mathbf{D}_x^\beta \tilde{\psi}(h)\|_{L^\infty} \lesssim 2^{(|\beta| + \frac{n}{2} - s)h} \|\tilde{\psi}\|_{\mathcal{H}_{lin}^s}$$

Using this and the representation formula in Lemma 3.7 we obtain the analogue of Lemma 3.8, namely

Lemma 3.11. *Let $\tilde\psi \in \mathcal{H}^s$. Then for small β we have*

$$
\begin{aligned}
\|\partial^\beta \partial_h \tilde\psi_0(h)\|_{L^2} &\lesssim\ 2^{(|\beta|-\frac{n}{2})h}\|\mathbf{D}_h\tilde\psi_0(h)\|_{\mathcal{H},h} \\
&+\ 2^{(|\beta|-s)h}\|\tilde\psi\|_{\mathcal{H}^s_{lin}}\|\mathbf{D}_h\tilde\phi(h)\|_{\mathcal{H},h}
\end{aligned}
$$

$$
\begin{aligned}
\|\partial^\beta \partial_h \tilde\psi_1(h)\|_{L^2} &\lesssim\ 2^{(|\beta|+1-\frac{n}{2})h}\|\mathbf{D}_h\tilde\psi_1(h)\|_{\mathcal{H},h} \\
&+\ 2^{(|\beta|+1-s)h}\|\tilde\psi\|_{\mathcal{H}^s_{lin}}\|\mathbf{D}_h\tilde\phi(h)\|_{\mathcal{H},h}
\end{aligned}
$$

In addition, if $n = 2$ then

$$
\|\tilde\phi_1(h)\|_{\mathbf{D}_x L^2 + L^1} \lesssim 2^{-h}\|\mathbf{D}_h\tilde\psi_1(h)\|_{\mathcal{H},h} + 2^{-sh}\|\tilde\psi\|_{\mathcal{H}^s_{lin}}\|\mathbf{D}_h\tilde\phi(h)\|_{\mathcal{H},h}
$$

This immediately gives the "$\gtrsim$" part of (3.14).

(ii) The "$\lesssim$" part. This is done exactly as in part (b) of Proposition 3.6. We use the same function $\tilde\phi$ as there and the function $\tilde\psi$ is constructed as follows. We begin with the mollified function

$$
(P_{<h}\psi_0, P_{<h}\psi_1)
$$

Then we orthogonally project both components on $T_{\tilde\phi(h)}M$. $\qquad\square$

3.4. **Comparing different initial data sets.** Now we consider two initial data sets and seek to connect them with a one parameter family of initial data sets.

Definition 3.12. Consider two initial data sets $\phi^{(0)}$ and $\phi^{(1)}$ which are small in $\dot H^{\frac{n}{2}} \times \dot H^{\frac{n}{2}-1}(\mathbb{R}^n, M)$. The $\dot H^s \times \dot H^{s-1}$ distance between them is

$$
d_M(\phi^{(1)}, \phi^{(2)})_{\dot H^s \times \dot H^{s-1}} = \inf \int_0^1 \|\mathbf{D}_\theta\phi\|_{\dot H^s \times \dot H^{s-1}} d\theta
$$

where the infimum is taken over all initial data paths $\phi(\theta)$, $\theta \in [0,1]$ with the property that
  (i) $\phi(\theta)$ is smooth in x, θ for $\theta \in (0,1)$.
  (ii) $\phi(\theta)$ is small in $\dot H^{\frac{n}{2}} \times \dot H^{\frac{n}{2}-1}(\mathbb{R}^n, M)$ uniformly in θ.
  (iii) $\lim_{\theta\to 0}\phi(\theta) = \phi^{(0)}$ and $\lim_{\theta\to 1}\phi(\theta) = \phi^{(1)}$ in the L^2 sense.

It is not difficult to see that this is a distance. Note however, that it might be infinite. The next result shows that in the context of the isometric embedding $M \subset \mathbb{R}^d$ this distance is equivalent to the corresponding euclidean norm.

24

Proposition 3.13. *Suppose that M is uniformly isometrically embedded in $\mathbb{R}^d$. Consider two initial data sets $\phi^{(1)}$ and $\phi^{(2)}$ which are small in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)$. Then*

$$d_M(\phi^{(0)}, \phi^{(1)})_{\dot{H}^s \times \dot{H}^{s-1}} \approx \|\phi^{(0)} - \phi^{(1)}\|_{\dot{H}^s \times \dot{H}^{s-1}}$$

Proof. By Proposition 3.10 the "$\gtrsim$" inequality is straightforward. For the converse one is tempted to begin with the straight line between $\phi^{(1)}$ and $\phi^{(2)}$ and then project it onto M. Such an argument cannot work because there is no guarantee that the intermediate points are close to M.

Instead, we first regularize $\phi^{(1)}$ and $\phi^{(2)}$ up to a point, and only then we take a straight line which we project onto M. Precisely, let $\epsilon > 0$ be a small with the property that

$$\|\phi^{(1)}\|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)}, \|\phi^{(2)}\|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)} \leq \epsilon$$

We choose h_0 so that

$$\|\phi^{(1)} - \phi^{(2)}\|_{\dot{H}^s \times \dot{H}^{s-1}} = \epsilon 2^{-h_0(\frac{n}{2}-s)}$$

Estimating an L^∞ norm by the $\dot{H}^s$ norm we obtain

(3.16)
$$\|P_{<h_0}\phi_0^{(1)} - P_{<h_0}\phi_0^{(2)}\|_{L^\infty} \lesssim \epsilon$$

We now construct a path between $\phi^{(1)}$ and $\phi^{(2)}$ as follows:

(i) Beginning from $\phi^{(1)}$ we move along $\Pi P_{<h}\phi^{(1)}$ from $h = \infty$ to $h = h_0$.

(ii) We continue on the projection of the straight line

$$\Pi((1 - \theta)P_{<h_0}\phi_0^{(1)} + \theta P_{<h_0}\phi_0^{(2)}), \qquad \theta \in [0, 1]$$

(iii) We move along $\Pi P_{<h}\phi^{(1)}$ from $h = h_0$ to $h = \infty$.

From the proof of Proposition 3.6(b) we know that steps (i) and (iii) are well defined. On the other hand, (3.16) also shows that the straight line in (ii) stays in an ϵ neighbourhood of M, so that its projection is well defined.

If we disregard the projection Π then the $\dot{H}^s \times \dot{H}^{s-1}$ length of this path is easy to determine. Indeed, for the first part (i) we have

$$
\begin{aligned}
\int_{h_0}^{\infty} \|\partial_h P_{<h}\phi^{(1)}\|_{\dot{H}^s \times \dot{H}^{s-1}} dh \quad &\lesssim \quad \int_{h_0}^{\infty} 2^{-h(\frac{n}{2}-s)} \|\phi^{(1)}\|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}} dh \\
&\lesssim \quad \epsilon 2^{-h(\frac{n}{2}-s)} \\
&\lesssim \quad \|\phi^{(1)} - \phi^{(2)}\|_{\dot{H}^s \times \dot{H}^{s-1}}
\end{aligned}
$$

and similarly for the third part (iii). The second part (ii) is straight, therefore the similar bound follows.

Finally, one can argue as in the proof of Proposition 3.6 to see that these bounds survive after applying the projection Π.

$\square$

3.5. **Extension results.** Consider a ball B in $\mathbb{R}^n$. Then we can define

$$\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(B, M)$$

as the space of those initial data sets which have extensions in

$$\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)$$

with the induced notion of "size". Similarly, we define

$$\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(B, \mathbb{R}^d)$$

as the space of those functions which have extensions in

$$\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, \mathbb{R}^d)$$

with the induced norm.

Suppose that M embeds isometrically into $\mathbb{R}^d$. Then Proposition 3.6 shows that

$$\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(B, M) \subset \dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(B, \mathbb{R}^d)$$

Is the converse true ? This is investigated in the next result.

Proposition 3.14. *Suppose that M is uniformly isometrically embedded into $\mathbb{R}^n$. Let ϕ be an initial data set which is small in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(B, \mathbb{R}^d)$. Then ϕ is in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(B, M)$ and*

$$\|\phi\|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(B,M)} \lesssim \|\phi\|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(B,\mathbb{R}^d)}$$

The extension of ϕ in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, M)$ can be taken to be constant outside $2B$. Furthermore, this extension is continuous in all better topologies $\dot{H}^s \times \dot{H}^{s-1}$, $s > n/2$.

Proof. After rescaling we can assume that B has size 1. Let $\phi \in \dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(B, \mathbb{R}^d)$. Then it has an extension to $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}(\mathbb{R}^n, \mathbb{R}^d)$, which we still denote by ϕ. The problem is, this extension needs not take values into TM outside B.

The strategy of the proof is to first construct an extension which stays close to M, and then project it on TM. We cannot do this with an arbitrary extension, as there is no guarantee it stays close to M.

We consider a discrete Littlewood-Paley decomposition for ϕ,

$$\phi = \sum_{k \in \mathbb{Z}} P_k \phi$$

We want to truncate $P_k\phi$ at distance 2^{-k} from B. However, because of the uncertainty principle this produces tails at all other frequencies. This is not a problem with the higher frequencies, but some care is required for the lower frequencies. We choose a large integer k and set

$$\phi^c = \sum_{k \in \mathbb{N}} \Delta^N(\chi_k \Delta^{-N} P_k \phi) + P_{<0}\phi(0) + \chi_0(P_{<0}\phi - P_{<0}\phi(0))$$

Here χ_k is a cutoff function supported in $B_{1+2^{-k+1}}$ and which equals 1 in $B_{1+2^{-k}}$ if $k \geq 0$. For $k < 0$ we simply set $\chi_k = 1$. Then it is easy to verify that ϕ^c satisfies

$$\|\phi^c\|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}} \lesssim \|\phi\|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}}$$

We claim that we also have

(3.17) $$d(\phi_0^c, M) \lesssim \epsilon \qquad \text{in } 2B$$

Take $k > 0$ and $x \in 2B$ so that $d(x, B) \in [2^{-k-1}, 2^{-k}]$. Then

$$\phi_0^c(x) = \Delta^N(\chi_k \Delta^{-N} P_k \phi) + P_{<k}\phi_0(x)$$

The first term is small in L^∞ and can be neglected. The second one on the other hand is smooth on the 2^{-k} scale,

$$\|\partial P_{<k}\phi_0(x)\|_{L^\infty} \ll 2^k$$

We can compare it with ϕ_0,

$$\|\phi_0 - P_{<k}\phi_0\|_{L^2} \ll 2^{-\frac{nk}{2}}$$

This implies we can find $y \in B \cap B(x, 2^{-k+1})$ so that

$$|\phi_0(y) - P_{<k}\phi_0(y)\| \ll 0$$

which further yields

$$d(P_{<k}\phi_0(y), M) \ll 1$$

But $P_{<k}\phi_0(y)$ is also smooth on the 2^{-k} scale,

$$\|\partial P_{<k}\phi_0(x)\|_{L^\infty} \ll 2^k$$

This implies that

$$d(P_{<k}\phi_0(x), M) \ll 1$$

Take now $x \in 2B$ so that $d(x, B) \in [\frac{1}{2}, 1]$. Then

$$\phi_0^c(x) = (P_{<0}\phi_0)(x) + (1 - \chi_0(x))(P_{<0}\phi_0)(0) - P_{<0}\phi_0)(x))$$

The first term is treated as in the case $k > 0$ and the second is small.

This proves (3.17). To conclude the proof of the Lemma it remains to replace ϕ^c with $\Pi\phi^c$, which can be done as in the proof of Proposition 3.6.

<div align="right">□</div>

4. Function spaces and Tao's estimates

Here we describe the properties of the function spaces which we use to estimate the solution ϕ and (good parts of) the right hand side of the equation.

The dyadic components of our function spaces are identical to those used in Tao's paper [18]. However, we assemble these dyadic pieces differently, and also we rely less on L^∞ bounds, which become somewhat of a luxury for unbounded target manifolds. In the next theorem δ_0 is a sufficiently small positive constant.

Theorem 4.1. *There exist translation invariant and scale invariant Banach spaces*

$$S \subset \mathcal{S}'(\mathbb{R}^{n+1})/\{const\}, \qquad N \subset \mathcal{S}'(\mathbb{R}^{n+1})$$

which satisfy the following properties:

(i) (Energy estimates)

$$\tag{4.1} \|\nabla_{x,t}\phi\|_{C(\dot{H}^{\frac{n}{2}-1})} \lesssim \|\phi\|_S \lesssim \|\Box\phi\|_N + \|\phi[0]\|_{\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}}$$

(ii) (dyadic structure)

$$\tag{4.2} \|\phi\|_S^2 \approx \sum \|\phi_k\|_S^2, \qquad \|f\|_N^2 \approx \sum \|f_k\|_N^2$$

(iii) (algebra property) $S \cap L^\infty$ is an algebra and

$$\tag{4.3} \|\phi_k \psi_{<k}\|_S \lesssim \|\phi_k\|_S \|\psi_{<k}\|_{S \cap L^\infty}$$

$$\tag{4.4} \|P_i(\phi_k \psi_j)\|_S \lesssim 2^{-|k-i|}\|\phi_k\|_S \|\psi_j\|_S \qquad k = j + O(1)$$

(iv) (product estimates)

$$\tag{4.5} \|\phi_{<k} f_k\|_N \lesssim \|\phi_{<k}\|_{S \cap L^\infty} \|f_k\|_N$$

$$\tag{4.6} \|P_i(\phi_k f_j)\|_N \lesssim 2^{-\delta_0(k-i)_+}\|\phi_k\|_S \|f_j\|_N, \qquad j \leq k$$

(v) (null form estimates)

$$\tag{4.7} \|P_i(\partial^\alpha \phi_k \partial_\alpha \psi_j)\|_N \lesssim 2^{-\delta(k-i)_+}\|\phi_k\|_S \|\psi_j\|_S$$

(vi) (trilinear estimate)

$$\left\|P_k\left(\phi_{k_1}^{(1)} \partial^\alpha \phi_{k_2}^{(2)} \partial_\alpha \psi_{k_3}^{(3)}\right)\right\|_N \lesssim 2^{-\delta_0[(k_1 - \min\{k_2,k_3\})_+ + (\max\{k_1,k_2,k_3\}-k)_+]}$$

$$\tag{4.8} \|\phi_{k_1}^{(1)}\|_S \|\phi_{k_2}^{(2)}\|_S \|\psi_{k_3}^{(3)}\|_S$$

(vii) (weakly closed spaces) If $\phi_i \to \phi$ in the sense of distributions and $\|\phi_i\|_S \leq 1$ then $\phi \in S$ and $\|\phi\|_S \leq 1$. The same property holds for N.

This result is essentially[1] contained in Theorem 3 in [18]. The construction of the spaces S and N is quite intricate. It combines an $X^{s,b}$ type structure with a more sophisticated combination of frequency localized energy type spaces with respect to varying characteristic frames. This type of spaces have originally been introduced in [20]. In [18] the space N is essentially the same as in [20]; however, the space S has been enlarged in order to gain the crucial algebra property, previously available only at the dyadic level.

There are many possible variations on this structure. Indeed, one is tempted to modify it in order to gain various technical simplifications and improvements in the estimates. However, the disadvantage would be that one would have to repeat many of the arguments in [20] and [18], thus lengthening the paper considerably without a major conceptual breakthrough. Fortunately the proof of the main theorem does not use directly the structure of the S and N spaces; instead it only relies on their properties in Theorems 4.1,4.3. This will make it easier to simplify the current setup in the future. In what follows we comment on each of the properties above.

The space S is used to bound solutions with initial data in $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$. The space N is where the right hand side of the equation is estimated. Part (i) shows that indeed one can solve the linear wave equation in these spaces, and that they are compatible with the initial data spaces.

An immediate consequence of (i) is the uniform bound

$$(4.9) \qquad \|\nabla \phi_k\|_{L^\infty} \lesssim 2^k \|\phi\|_S$$

A corollary of it says that dyadic pieces of functions in S are bounded.

$$(4.10) \qquad \|\phi_k\|_{L^\infty} \lesssim \|\phi_k\|_S$$

On the contrary, the embedding $S \subset L^\infty$ is false.

Part (ii) simply says that the S spaces inherit the natural dyadic structure from the initial data. One can take this as the definition of the spaces S and N in this paper, where the dyadic pieces are the same ones introduced in [18].

The dyadic version of the algebra property (iii) was already proved in [20]. Gaining the full algebra property was a crucial step of the arguments in [18]. While not explicitly stated in [18], the gain in the high $\times$ high $\rightarrow$ low frequency interactions in (4.4) is in effect obtained in the proof. The one unit gain is not essential here, but we do need more that a half unit gain. The algebra property is no longer sufficient

---

[1]see the comments below

in this article. Instead it only serves as a stepping stone toward the stronger Moser estimates in Theorem 4.12 below.

The product estimates in (iv) are from [18]. They are not very far from the algebra property in (iii), as the spaces S and N have to a large extent dual structures. Perhaps the most significant disadvantage of the spaces S and N in [18], which we also use here, is the lack of a gain in (4.6) in the case $j < k$. This is essentially what makes the trilinear estimate in (vi) necessary. Such a gain existed in [20] and can be regained here by modifying the definition of the spaces S and N.

The dyadic null form estimates in (v) also include a gain in the high $\times$ high $\to$ low frequency interactions. There is however no gain in the high $\times$ low $\to$ high frequency interaction. This is independent of the choice of the spaces, and reflects the fact that the wave maps equation in the critical Sobolev spaces is genuinely nonlinear, as opposed to semilinear.

The trilinear estimates (vi) show that the only genuinely nonlinear interactions occur when the undifferentiated factor has the lowest frequency. This is one of the basic tenets of paradifferential calculus, and is what makes a paradifferential type approach possible for the wave maps equation. Ideally one would like to think of it as a consequence of bilinear estimates, but this is not possible without changing the spaces S and N.

The weak closedness property (vii) for S reduces to the similar property for the dyadic pieces of S. But then this is a consequence of the definition of the dyadic spaces in [18]. The similar property for the dyadic pieces of the space N in [18] is not true. However this can be remedied by replacing N with its weak completion, without harming any of the properties (i)-(vi). Precisely, for any distribution f we set

$$\|f\|_N = \inf\{\liminf \|f_n\|_N;\ f_n \in N,\ f_n \to f \text{ in } \mathcal{D}'\}$$

and the completion consists of those distributions f for which the above norm is finite.

In order to measure more (or slightly less) regular solutions to the wave maps equation we adopt Tao's notion of frequency envelope. We fix a sufficiently large constant M and choose small positive constants $0 < \delta_2 < \delta_1 \ll \delta_0$.

Definition 4.2. a) A frequency envelope $c = \{c_k\}_{k\in\mathbb{Z}}$ is admissible if it satisfies

$$2^{-\delta_1(k-j)}c_k \le c_j \le 2^{M(k-j)}c_k \qquad j < k$$

b) A frequency envelope $c = \{c_k\}_{k\in\mathbb{Z}}$ is admissible for the linearized equation if it satisfies

$$2^{-\delta_1(k-j)}c_k \leq c_j \leq 2^{-\delta_2(k-j)}c_k \qquad j < k$$

If c is an admissible frequency envelope then we denote by S_c the Banach space which is the closure of $\mathcal{D}(\mathbb{R}^{n+1})$ with respect to the norm

$$\|\phi\|_{S_c} = \sup_{k\in\mathbb{Z}} c_k^{-1}\|\phi_k\|_S$$

As in the case of S, depending on the choice of c this may be a quotient space of $\mathcal{D}'$ by the space of polynomials up to a certain order.

The reader is cautioned that while our definition of the S_c spaces coincides with Tao's, the class of admissible frequency envelopes does not.

For an admissible frequency envelope c we introduce the spaces H_c and H_c^{-1} with norm

$$\|u\|_{H_c} = \sup_{j\in\mathbb{Z}} c_j^{-1}2^{\frac{nj}{2}}\|u_j\|_{L^2}$$

$$\|u\|_{H_c^{-1}}^2 = \sup_{j\in\mathbb{Z}} c_j^{-1}2^{(\frac{n}{2}-1)j}\|u_j\|_{L^2}^2$$

The initial data space corresponding to S_c is denoted by $S_c[0]$ and has norm

$$\|\phi[0]\|_{S_c[0]} = \|\phi_0\|_{H_c} + \|\phi_1\|_{H_c^{-1}}$$

This may also be a quotient space.

Given $\delta > -\delta_1$ we also introduce the space S_δ with norm

$$\|\phi\|_{S_\delta}^2 = \sum_{k\in\mathbb{Z}} 2^{\delta k}\|\phi_k\|_S^2$$

Then $S_\delta[0]$ is the corresponding initial data space.

In the sequel we need to use the spaces S, S_c etc. in a bounded time interval I. Then the corresponding norms are defined using the smallest extensions to $\mathbb{R}^{n+1}$. For instance

$$\|\phi\|_{S(I\times\mathbb{R}^n)} = \inf\{\|\tilde{\phi}\|_S;\ \tilde{\phi} \text{ is an extension of } \phi\}$$

For convenience we drop the domain $I \times \mathbb{R}^n$ from the notation and use simply S. Its meaning will be clear from the context.

The linear and multilinear estimates for the S and S_c spaces are generally handled using Theorem 4.1. However, given the form of the wave maps equation, here we also need to consider fully nonlinear expressions. This is achieved in the following supplement to Theorem 4.1.

Theorem 4.3. *There is a positive integer N so that F is a smooth function with uniformly bounded derivatives then $\phi \in S$ implies $F(\phi) \in S$ and*

$$(4.11) \qquad \|F(\phi)\|_S \lesssim \|\phi\|_S (1 + \|\phi\|_S^N)$$

If in addition $\phi \in S_c$ for some admissible frequency envelope c then $F(\phi) \in S_c$ and

$$(4.12) \qquad \|F(\phi)\|_{S_c} \lesssim \|\phi\|_{S_c} (1 + \|\phi\|_S^N)$$

5. Preliminary results

In the next sections we prove various estimates in S and S_c spaces for smooth solutions for the wave maps equation. Before we do that, however, we need to show that smooth solutions are indeed in these spaces. This is the aim of this section. The arguments here are rather standard and are presented only for the sake of completeness. They are independent of the rest of the paper.

For this type of analysis we do not not take advantage of the precise form of the wave maps equation. Consequently, we consider a generic semilinear wave equation of the form

$$(5.1) \qquad \Box\phi = \mathbf{S}(\phi)(\nabla\phi)^2$$

In what follows we fix $T > 0$, and consider the equation (5.1) in the time interval $[-T, T]$ with initial data $\phi[0]$.

We would like to add extra regularity to the initial data but without altering its global behavior. Consequently we choose some large integer k and define the spaces $H_{(1,\infty)}^k$ respectively $H_{0,\infty}^k$ of functions in $\mathbb{R}^n$ with norms

$$\|u\|_{H_{(0,\infty)}^k} = \sup_{y \in \mathbb{R}^n} \sum_{|\beta|=0}^{k} \int_{B(y,1)} |\partial^\beta u|^2 dx$$

$$\|u\|_{H_{(1,\infty)}^k} = \sup_{y \in \mathbb{R}^n} \sum_{|\beta|=1}^{k} \int_{B(y,1)} |\partial^\beta u|^2 dx$$

The space $H_{(0,\infty)}^k$ is a subspace of L^∞. The $H_{(1,\infty)}^k$ can be viewed as a subspace of the quotient space $L^\infty/\{const\}$. We begin with a result which has a local nature:

Theorem 5.1. *a) For each initial data $\phi[0] \in H_{(1,\infty)}^k \times H_{(0,\infty)}^{k-1}$ there is some $T_0 > 0$, depending only on the size of the initial data, and a solution $\phi \in C(-T_0, T_0; H_{(1,\infty)}^k \times H_{(0,\infty)}^{k-1})$. Furthermore, the solution depends smoothly on the initial data in the above topologies.*

b) If the initial data $\phi[0]$ is sufficiently small in $H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)}$ then we can take $T_0 \geq T$.

c) If in addition the initial data is smooth then the solution is smooth up to time T_0.

Because of the finite speed of propagation this result reduces to a corresponding result for compactly supported initial data. But in this case the proof is straightforward, using a fixed point argument in the space $C(H^k \times H^{k-1})$. The details are left to the reader. Note that the fact that the space $H^k_{(1,\infty)}$ does not see the constants is compensated for by the fact that all constant functions solve the equation.

We continue with a result concerning the global Sobolev regularity of smooth solutions:

Theorem 5.2. *Assume that for some initial data $\phi[0] \in H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)}$ there is a solution $\phi \in C(-T,T; H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$.*

a) If $\phi[0] \in S[0]$ then $\phi \in S$ and

$$\|\phi\|_S \lesssim A \|\phi[0]\|_{S[0]}$$

where A depends only on the $C(-T,T; H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$ norm of ϕ and on the $S[0]$ norm of $\phi[0]$.

b) Let c be an admissible frequency envelope. If in addition the initial data $\phi[0]$ is in $S_c[0]$ then $\phi \in S_c$ and

$$\|\phi\|_{S_c} \lesssim B \|\phi[0]\|_{S_c[0]}$$

where B depends only on the $C(-T,T; S[0] \cap H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$ norm of ϕ.

Above we considered solutions for the wave maps equation with initial data in $S[0] \cap H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)}$. However, when we seek to prove continuous dependence on the initial data we need a stronger topology. Precisely, we shall measure the difference between two initial data in the smaller space $S[0] \cap H^k_{(0,\infty)} \times H^{k-1}_{(0,\infty)}$.

Theorem 5.3. *Let $\phi \in S \cap C(-T,T; H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$ be a solution to (5.1). Then:*

a) For any initial data $\tilde{\phi}[0]$ which is close to ϕ in $S[0] \cap H^k_{(0,\infty)} \times H^{k-1}_{(0,\infty)}$ there is an unique solution $\tilde{\phi}$ which satisfies

$$\tilde{\phi} - \phi \in S \cap C(-T,T; H^k_{(0,\infty)} \times H^{k-1}_{(0,\infty)})$$

depending smoothly on the initial data in the above topology.

b) Let c be an admissible frequency envelope for the linearized equation. Then for each initial data $\psi[0] \in S_c[0] \cap H^k_{(0,\infty)} \times H^{k-1}_{(0,\infty)}$ there

is a solution $\psi \in C(-T, T; H_{(0,\infty)}^{k} \times H_{(0,\infty)}^{k-1})$ for the linearized equation which satisfies

$$\|\psi\|_{S_c} \leq B\|\psi[0]\|_{S_c[0]}$$

where B depends only on the $C(-T, T; S[0] \cap H_{(1,\infty)}^{k} \times H_{(0,\infty)}^{k-1})$ norm of ϕ.

Before we prove Theorems 5.2, 5.3 we remark that they are trivial if the initial data $\phi[0]$ is supported in a fixed compact set. The interest comes from the fact that we make no decay assumption at infinity. However, this is not so essential because we work on a bounded time interval and the wave equation has a finite speed of propagation.

We begin with some preparatory steps. First we define the Besov space $\dot{B}_{\infty,\infty}^{0}$ as the space of temperate distributions in $\mathbb{R}^n$ for which the following norm is finite,

$$\|\phi\|_{\dot{B}_{\infty,\infty}^{0}} = \sup_{h \in \mathbb{R}} \|\phi_h\|_{L^\infty}$$

and which satisfy

(5.2) $$\lim_{h \to -\infty} \|\partial_x \phi_{<h}\|_{L^\infty} = 0$$

This is a space of distributions modulo constant functions. The role of the latter condition is to exclude higher order polynomials. It implies that

(5.3) $$\partial_x \phi_{<h} = \int_{-\infty}^{h} \partial_x \phi_l dl$$

where the integral converges exponentially in L^∞.

Lemma 5.4. *Let F be a smooth bounded function with uniformly bounded derivatives. Then for $\phi \in \dot{B}_{\infty,\infty}^{0}$ we have*

$$\|\partial^\beta F(\phi_{<h})\|_{L^\infty} \lesssim A_\beta 2^{\beta h}$$

where A_β depends on β and on the $\dot{B}_{\infty,\infty}^{0}$ norm of ϕ.

The proof is a simple application of the chain rule. We continue with a Moser type estimate,

Lemma 5.5. *Let F be a smooth bounded function with uniformly bounded derivatives. Let $s > 0$. Then for all $\phi \in \dot{H}^s \cap \dot{B}_{\infty,\infty}^{0}$ we have*

$$\|F(\phi)\|_{\dot{H}^s} \lesssim A\|\phi\|_{\dot{H}^s}$$

where A depends only on the $\dot{B}_{\infty,\infty}^{0}$ norm of ϕ.

Proof. We begin with the relation

$$\partial_x F(\phi_{<h_1}) - \partial_x F(\phi_{<h_0}) = \int_{h_0}^{h_1} \partial_x(\phi_h F'(\phi_{<h}))dh$$

We claim that we can let $h_0 \to -\infty$ and $h_1 \to \infty$ to obtain

(5.4) $$\partial_x F(\phi) = \int_{-\infty}^{\infty} \partial_x(\phi_h F'(\phi_{<h}))dh$$

in the sense of distributions. For the limit as $h_0 \to -\infty$ it suffices to use (5.2). For the limit as $h_1 \to \infty$ we observe that

$$\|\phi_h\|_{L^2} \lesssim 2^{-sh}\|\phi\|_{\dot{H}^s}$$

while

$$\|F(\phi_{<h}) - F(\phi)\|_{L^2} \lesssim \|\phi_{<h} - \phi\|_{L^2} \lesssim 2^{-sh}$$

From (5.4) we obtain

$$P_k F(\phi) = \int_{-\infty}^{\infty} P_k(\phi_h F'(\phi_{<h}))dh$$

For a fixed large C we split the integral in two,

$$P_k F(\phi) = \int_{-\infty}^{h-C} P_k(\phi_h P_{>h-C} F'(\phi_{<h}))dh + \int_{h-C}^{\infty} P_k(\phi_h F'(\phi_{<h}))dh$$

We bound this in L^2, using Lemma 5.4. This yields

$$\|P_k F(\phi)\|_{L^2} \lesssim A \int_{-\infty}^{k-C} 2^{-N|k-h|}\|\phi_h\|_{L^2}dh + \int_{h-C}^{\infty} \|\phi_h\|_{L^2}dh$$

and further

$$2^{sk}\|P_k F(\phi)\|_{L^2} \lesssim \int 2^{(k-h)s} L(k,h)2^{sh}\|\phi_h\|_{L^2}dh$$

where

$$L(k,h) = \begin{cases} 2^{-N|k-h|} & h < k \\ 1 & h > k \end{cases}$$

But the kernel $2^{(k-h)s}L(k,h)$ is translation invariant and integrable, therefore it maps L^2 into L^2. The conclusion follows. $\qquad\square$

Some immediate applications to our problem are collected in the next Lemma:

Lemma 5.6. *Let F be a smooth bounded function with uniformly bounded derivatives.*
  a) Let $\phi \in H^k_{(1,\infty)}$. Then $F(\phi) \in H^k_{(1,\infty)}$ and

$$\|F(\phi)\|_{H^k_{(1,\infty)}} \le A\|\phi\|_{H^k_{(1,\infty)}}$$

35

where A depends only on the quantity

$$\max\{\|\nabla\phi_{<0}\|_{L^\infty}, \sup_{h\geq 0}\|\phi_h\|_{L^\infty}\}$$

b) Let $\phi \in \dot{H}^{\frac{n}{2}}$. Then $F(\phi) \in \dot{H}^{\frac{n}{2}}$ and

$$\|F(\phi)\|_{\dot{H}^{\frac{n}{2}}} \lesssim \|\phi\|_{\dot{H}^{\frac{n}{2}}}$$

where A depends only on the $\dot{B}^0_{\infty,\infty}$ size of ϕ.

c) Suppose $\phi \in H_c \cap \dot{B}^0_{\infty,\infty}$ for some admissible frequency envelope c. Then $F(\phi) \in H_c$ and

$$\|F(\phi)\|_{H_c} \lesssim A\|\phi\|_{H_c}$$

where A depends only on the $\dot{B}^0_{\infty,\infty}$ size of ϕ.

Proof. a) It suffices to estimate $F(\phi)$ inside the unit ball B. Let χ be a cutoff function supported in $2B$ which equals 1 in B. We substitute

$$v = \chi(\phi - \phi_{<0}(0)), \qquad G = F(\cdot + \phi_{<0}(0)) - F(\phi_{<0}(0))$$

and note that

$$\|v\|_{H^k} \lesssim \|\phi\|_{H^k_{(1,\infty)}}$$

$$\|v\|_{\dot{B}^0_{\infty,\infty}} \lesssim \max\{\|\nabla\phi_{<0}\|_{L^\infty}, \sup_{j\geq 0}\|\phi_j\|_{L^\infty}\}$$

Then it suffices to show that

$$\|G(v)\|_{H^k} \lesssim A\|v\|_{H^k}$$

for F satisfying $F(0) = 0$ and A depending only on the $\dot{B}^0_{\infty,\infty}$ size of v. The low frequencies are easy to estimate,

$$\|F(\phi)\|_{L^2} \lesssim \|\phi\|_{L^2}$$

and the $\dot{H}^k$ norms are given by Lemma 5.5.

b) This follows directly from Lemma 5.5.

c) Arguing as in Lemma 5.5, we need to show that

$$\sum_{j\in\mathbb{Z}} 2^{\frac{n}{2}(k-j)} L(k,j)c_j \lesssim c_k$$

But this is a consequence of the admissibility condition for c. $\qquad\square$

Next we consider estimates for the nonlinear term in the equation,

$$N(\phi) = \mathbf{S}(\phi)(\nabla\phi)^2$$

To describe some of them we introduce a slight modification of the space $S[0]$, namely

$$\tilde{S}[0] = \dot{H}^{\frac{n}{2}} \cap \dot{H}^{\frac{n}{2}-1} \times \dot{H}^{\frac{n}{2}-1}$$

Lemma 5.7. *Let $\phi[t] \in S[0] \cap H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)}$. Then*

a) $N(\phi)(t) \in \dot{H}^{\frac{n}{2}-1}$ and

(5.5)
$$\|N(\phi)(t)\|_{\dot{H}^{\frac{n}{2}-1}} \lesssim A \|\phi[t]\|_{S[0]}$$

where A depends only the $S[0] \cap H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)}$ norm of $\phi[0]$.

b) If in addition $\phi[0] \in S_c[0]$ then

(5.6)
$$\|N(\phi)\|_{H_c^{-1}} \lesssim A \|\phi[t]\|_{S_c[0]}$$

c) Consider a second initial data set $\psi[t] \in \tilde{S}[0] \cap H^k_{(0,\infty)} \times H^{k-1}_{(0,\infty)}$. Then

(5.7)
$$\|N(\phi+\psi) - N(\phi)\|_{\dot{H}^{\frac{n}{2}-1}} \lesssim B \|\psi[t]\|_{\dot{H}^{\frac{n}{2}} \cap \dot{H}^{\frac{n}{2}-1} \times \dot{H}^{\frac{n}{2}-1}}$$

where B depends on the $S[0] \cap H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)}$ norm of $\phi[t]$ and the $H^k_{(0,\infty)} \times H^{k-1}_{(0,\infty)}$ norm of $\psi[t]$.

Proof. a) We use Lemma 5.6(b) for the nonlinear factor $\mathbf{S}(\phi)$. Then it suffices to prove two scale invariant bilinear estimates, namely

(5.8)
$$\|\phi^{(1)}\phi^{(2)}\|_{\dot{H}^{\frac{n}{2}-1}} \lesssim \|\phi^{(1)}\|_{L^\infty}\|\phi^{(2)}\|_{\dot{H}^{\frac{n}{2}-1}} + \|\phi^{(2)}\|_{L^\infty}\|\phi^{(1)}\|_{\dot{H}^{\frac{n}{2}-1}}$$

(5.9)
$$\|\phi^{(1)}\phi^{(2)}\|_{\dot{H}^{\frac{n}{2}-1}} \lesssim \|\phi^{(1)}\|_{L^\infty \cap \dot{H}^{\frac{n}{2}}}\|\phi^{(2)}\|_{\dot{H}^{\frac{n}{2}-1}}$$

These are routine and are left for the reader.

b) We use Lemma 5.6(b),(c) for the nonlinear term. The expression $\nabla\phi$ is in a bounded set in the space

$$\dot{H}^{\frac{n}{2}-1} \cap W^{1,\infty} \subset W^{\frac{n}{4},4} \cap L^\infty$$

Then it suffices to prove two other bilinear estimates,

$$\|\phi^{(1)}\phi^{(2)}\|_{H_c^{-1}} \lesssim \|\phi^{(1)}\|_{W^{\frac{n}{4},4} \cap L^\infty}\|\phi^{(2)}\|_{H_c^{-1}} + \|\phi^{(2)}\|_{W^{\frac{n}{4},4} \cap L^\infty}\|\phi^{(1)}\|_{H_c^{-1}},$$

$$\|\phi^{(1)}\phi^{(2)}\|_{H_c^{-1}} \lesssim \|\phi^{(1)}\|_{\dot{H}^{\frac{n}{2}} \cap L^\infty}\|\phi^{(2)}\|_{H_c^{-1}} + \|\phi^{(2)}\|_{\dot{H}^{\frac{n}{2}-1}}\|\phi^{(1)}\|_{H_c}$$

These are also standard and are left for the reader.

c) For the expression $\mathbf{S}(\psi)[\nabla(\phi+\psi)^2 - (\nabla\phi)^2]$ we argue as in case (a) above. It remains to consider the term

$$(\mathbf{S}(\phi+\psi) - \mathbf{S}(\phi))(\nabla\phi)^2$$

We represent the difference

$$\mathbf{S}(\phi+\psi) - \mathbf{S}(\phi) = \psi\mathbf{S}_1(\phi,\psi)$$

where $\mathbf{S}_1$ is smooth, bounded with uniformly bounded derivatives.

For ψ in a bounded set we can use Fourier series to expand the second factor in a rapidly convergent series of terms of the form

$$\mathbf{S}_2(\psi)\mathbf{S}_3(\phi)$$

where we take $\mathbf{S}_2$ to be compactly supported. Hence we can replace $\mathbf{S}(\phi + \psi) - \mathbf{S}(\phi)$ by
$$\psi \mathbf{S}_2(\psi) \mathbf{S}_3(\phi)$$
The nonlinear term $\psi \mathbf{S}_2(\psi)$ is estimated in $\dot{H}^{\frac{n}{2}-1} \cap L^\infty$ using Lemma 5.5, while the remaining product $\mathbf{S}_3(\phi)(\nabla \phi)^2$ is estimated in $H^{\frac{n}{2}-1} \cap L^\infty$ using part (a) of the lemma. To multiply them we use (5.8).

□

Finally we consider the linearization of the nonlinearity,
$$DN(\phi)\psi = 2\mathbf{S}(\phi)\nabla\phi\nabla\psi + \mathbf{S}'(\phi)\psi(\nabla\phi)^2$$

Lemma 5.8. *Suppose that* $\phi[t] \in S[0] \cap H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)}$.

a) If $\psi[t] \in S[0] \cap H^k_{(0,\infty)} \times H^{k-1}_{(0,\infty)}$ *then* $DN(\phi)\psi \in \dot{H}^{\frac{n}{2}-1}$ *and*

$$\tag{5.10} \|DN(\phi)\psi\|_{\dot{H}^{\frac{n}{2}-1} \cap H^{k-1}_{(0,\infty)}} \lesssim \|\psi[t]\|_{S[0] \cap H^k_{(0,\infty)} \times H^{k-1}_{(0,\infty)}}$$

b) Let c *be an admissible frequency envelope for the linearized equation. If* $\psi[t] \in S_c[0]$ *then* $DN(\phi)\psi \in H^{-1}_c$ *and*

$$\tag{5.11} \|DN(\phi)\psi\|_{H^{-1}_c} \lesssim \|\psi[t]\|_{S_c[0]}$$

Proof. For part (a) we can use the bounds (5.8), (5.9) in the previous Lemma. The same bounds allow us to reduce part (b) to two bilinear estimates, namely
$$\|\phi^{(1)}\phi^{(2)}\|_{H^{-1}_c} \lesssim \|\phi^{(1)}\|_{H_c} \|\phi^{(2)}\|_{\dot{H}^{\frac{n}{2}-1}}$$
$$\|\phi^{(1)}\phi^{(2)}\|_{H^{-1}_c} \lesssim \|\phi^{(1)}\|_{H^{-1}_c} \|\phi^{(2)}\|_{\dot{H}^{\frac{n-1}{2}, \frac{2n}{n-1}} \cap L^\infty}$$
where we have used the fact that for large k we have
$$\dot{H}^{\frac{n}{2}-1} \cap H^{k-1}_{(0,\infty)} \subset \dot{H}^{\frac{n-1}{2}, \frac{2n}{n-1}} \cap L^\infty$$
Again the proof of the two estimates above is routine.

□

Proof of Theorem 5.2. **(i) A short time estimate.** Here we prove that there is some small time T_0, possibly depending on $\|\phi[0]\|_{S[0]}$, so that the solution ϕ is in $C(-T_0, T_0; S[0] \cap H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$.

Denote by $\tilde{\phi} \in C(-T, T; S[0] \cap H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$ the solution to
$$\Box \tilde{\phi} = 0, \qquad \tilde{\phi}[0] = \phi[0]$$
The difference
$$\psi = \phi - \tilde{\phi}$$
should satisfy
$$\Box \psi = \mathbf{S}(\psi + \tilde{\phi})(\nabla\psi + \nabla\tilde{\phi})^2, \qquad \psi[0] = 0$$

To solve this we use a contraction argument in the space

$$C(-T_0, T_0; X), \qquad X = S[0] \cap H_{(0,\infty)}^k \times H_{(0,\infty)}^{k-1}$$

Set

$$Y = \dot{H}^{\frac{n}{2}-1} \cap H_{(0,\infty)}^{k-1}$$

Due to the linear bound

$$\|\psi\|_{C(-T_0, T_0; X)} \lesssim \|\psi[0]\|_X + \|\Box\psi\|_{L^1(-T_0, T_0; Y)}$$

it suffices to prove that the nonlinearity

$$\psi \to N(\psi + \tilde{\phi})$$

is locally Lipschitz from $C(-T_0, T_0; X)$ into $L^1(-T_0, T_0; Y)$. To obtain the smallness in the contraction argument we replace the above L^1 norm by an L^∞ norm and assume that T_0 is small. Then the proof is concluded using (5.10), which yields the fixed time bound

$$\|DN(\psi + \tilde{\phi})\tilde{\psi}(t)\|_Y \leq C\|\tilde{\psi}[t]\|_X$$

with a constant C which may depend on $\|\psi[t]\|_X$.

(ii) A long time estimate. According to the previous step, we know that either the solution remains in $C(S[0])$ up to time T, or there exists some time $T_{max} \leq T$ so that

$$\phi[\cdot] \in C([0, T_{max}); S[0]), \qquad \lim_{t \to T_{max}} \|\phi[t]\|_{S[0]} = \infty$$

Here we prove that the latter alternative cannot occur, and that

$$(5.12) \qquad \|\phi[\cdot]\|_{C(0,T; S[0])} + \|\Box\phi\|_{L^1(0,T; \dot{H}^{\frac{n}{2}-1})} \leq A\|\phi[0]\|_S$$

where A depends only on the $C(-T, T; H_{(1,\infty)}^k \times H_{(0,\infty)}^{k-1})$ norm of ϕ. The same argument applies with T replaced by $-T$. Then the S bound for ϕ follows from (4.1).

For this we prove an a-priori bound. We choose $t < \min\{T_{max}, T_1\}$ where T_1 will be chosen later. We also let λ be a large parameter which will be chosen later. Energy estimates for the function ψ defined in (i) yield

$$\|\psi[\cdot]\|_{C(0,t; S[0])} \lesssim \int_0^t \|N(\psi + \tilde{\phi})(s)\|_{\dot{H}^{\frac{n}{2}-1}} ds$$

One additional integration gives a trivial improvement of it,

$$\|\psi[\cdot]\|_{C(0,t; \tilde{S}[0])} \lesssim \int_0^t \|N(\psi + \tilde{\phi})(s)\|_{\dot{H}^{\frac{n}{2}-1}} ds$$

This in turn implies that

$$\|e^{-\lambda t}\psi\|_{L^\infty(0,t; \tilde{S}[0])} \lesssim \lambda^{-1}\|e^{-\lambda t}N(\psi + \tilde{\phi})\|_{L^\infty(0,t; \dot{H}^{\frac{n}{2}-1})}$$

We use (5.5) for $N(\tilde{\phi})$ and (5.7) for the difference $N(\psi + \tilde{\phi}) - N(\tilde{\phi})$ to obtain

$$\|e^{-\lambda t}\psi\|_{L^\infty(0,t;\tilde{S}[0])} \lesssim A\lambda^{-1}(\|\phi[0]\|_{S[0]} + \|e^{-\lambda t}\psi\|_{L^\infty(0,t;\tilde{S}[0])})$$

where A depends only on the $S[0]$ size of $\phi[0]$ and the $C(-t,t; H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$ norm of ϕ. Hence if we choose $\lambda \gg A^{-1}$ we obtain

$$\|e^{-\lambda t}\psi\|_{C(-t,t;S[0])} \lesssim \|\phi[0]\|_{S[0]}$$

and further

$$\|e^{-\lambda t}\phi\|_{C(-t,t;S[0])} \lesssim \|\phi[0]\|_{S[0]}, \qquad t < \min\{T_{max}, T\}$$

This leads to contradiction if $T_{max} \leq T$. Hence we must have $T_{max} > T$ and (5.12) holds.

(iii) The S_c bound. As above we want to prove an estimate of the form

$$\|e^{-\lambda t}\phi\|_{L^\infty(0,T;S_c[0])} + \|e^{-\lambda t}\Box\phi\|_{L^\infty(0,T;H_c^{-1})} \lesssim B\|\phi[0]\|_{S_c[0]}$$

with B depending only on the $C(S[0] \cap H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$ norm of ϕ. It suffices to prove this in the case when $c \in L^2$; indeed, for any admissible frequency envelope c we can find admissible frequency envelopes $c^{(j)} \in L^2$ so that

$$\lim_{j \to \infty} c_i^{(j)} = c_i \qquad \forall\, i \in \mathbb{Z}$$

Then the estimates for c follow from the estimates for c^j. Energy estimates yield the bound

$$\|e^{-\lambda t}\phi\|_{L^\infty(0,T_1 S_c[0])} \lesssim \|\phi[0]\|_{S_c[0]} + \lambda^{-1}\|e^{-\lambda t}N(\phi)\|_{L^\infty(0,T;H_c^{-1})}$$

Using (5.6) we get

$$\|e^{-\lambda t}\phi\|_{L^\infty(0,T_1 S_c[0])} \lesssim \|\phi[0]\|_{S_c[0]} + B\lambda^{-1}\|e^{-\lambda t}\phi\|_{L^\infty(0,T;S_c[0])}$$

Choosing λ large enough we obtain

$$\|e^{-\lambda t}\phi\|_{L^\infty(0,T_1 S_c[0])} \lesssim \|\phi[0]\|_{S_c[0]}$$

while the bound for $\Box\phi$ comes from (5.6). $\qquad\qquad\square$

Proof of Theorem 5.3. a) Denote by $\psi = \tilde{\phi} - \phi$. Then ψ solves

$$\Box\psi = \mathbf{S}(\psi + \phi)(\nabla\psi + \nabla\phi)^2 - \mathbf{S}(\phi)(\nabla\phi)^2, \qquad \psi[0] = \tilde{\phi}[0] - \phi[0]$$

We solve by the contraction principle for ψ in a sufficiently small ball of the space $C_\lambda(0,T;X)$ with norm

$$\|\psi\|_{C_\lambda(0,T;X)} = \|e^{-\lambda t}\psi\|_{L^\infty(X)}$$

where we use the estimate (5.10) for the linearization of $N(\phi + \psi)$.

b) The local theory shows that a solution ψ exists and is unique in the space $C(-T, T; H_{0,\infty}^k \times H_{0,\infty}^{k-1})$. It remains to show that in effect $\psi \in S_c$. For this we note that ψ can be obtained via a fixed point argument in the space $C_\lambda(0, T; S_c[0])$. This is a consequence of the estimate

$$\|e^{-\lambda t} DN(\phi)\psi\|_{L^\infty(-T,T;H_c^{-1})} \lesssim \|e^{-\lambda t}\psi\|_{C(0,T;S_c[0])}$$

which in turn follows from (5.11).

A byproduct of the argument is that the right hand side $DN(\phi)\psi$ of the linearized equation is in $L^\infty H^{c-1}$. By (4.1) this implies that $\psi \in S_c$. $\qquad\square$

6. Bounds for smooth solutions.

Here we show how to bootstrap the regularity of smooth wave maps in S, respectively S_c. Since we do not a-priori know that the solutions are global, we prove the results on a fixed time interval $[-T, T]$. However by rescaling they hold uniformly with respect to T.

Theorem 6.1. *There is $\epsilon > 0$ so that for any $u \in S$ a smooth solution to the wave maps equation in $[-T, T]$ we have*

(6.1)
$$\left.\begin{array}{c} \|\phi[0]\|_{S[0]} \leq \epsilon^2 \\[2mm] \|\phi\|_S \leq 2\epsilon \end{array}\right\} \implies \|\phi\|_S \leq \epsilon$$

If in addition, $u \in S_c$ for some admissible frequency envelope c then

(6.2)
$$\left.\begin{array}{c} \|\phi\|_S \ll \epsilon \\[2mm] \|\phi\|_{S_c} \leq 2M \\[2mm] \|\phi[0]\|_{S_c[0]} \leq \epsilon M \end{array}\right\} \implies \|\phi\|_{S_c} \leq M$$

Proof of Theorem 6.1. The core estimate is (6.2), so we prove it first. Later we shall see that (6.1) is a straightforward consequence of it.

Let ϕ be a smooth wave map in $[-T, T]$ which satisfies the bounds on the left in (6.2). The first step is to replace the equation for ϕ with appropriate paradifferential like equations for its frequency localized components ϕ_k. The idea, going back to work of Bony [1] (see also Taylor's book [21] and references therein) is that the high-high frequency interactions are weaker and can be neglected in a first approximation. Thus, modulo acceptable errors one can replace a nonlinear equation

$$P(u, \partial u) = 0$$

with a sequence of linear equations for frequency localized components of the solution

$$P_{lin}(u_{<k}, \partial u_{<k})u_k \approx 0$$

in which P_{lin} is the linearized operator near the lower frequency part of the solution. This idea was used by Tao [18] in the case of the sphere, and our implementation of it here is developed along similar lines.

In what follows, we fix k and denote by "error" any function which satisfies

$$\|\text{error}\|_N \lesssim \epsilon M c_k$$

According to the above philosophy, the paradifferential formulation of the wave maps equation should be

$$\Box \phi_k^l = -2\mathbf{S}_{ij}^l(\phi)_{<k-C}\partial^\alpha \phi_{<k-C}^i \partial_\alpha \phi_k^j + \text{error}$$

While this is true, it would not be sufficient for our purposes. Instead we observe that according to the definition of the second fundamental form $\mathbf{S}$, we have $\mathbf{S}_{il}^j(\phi)\partial^\alpha \phi^j = 0$. Passing to a paradifferential formulation, this suggests that the expression

$$\mathbf{S}_{il}^j(\phi))_{<k-C}\partial_\alpha \phi_k^j$$

should also be better behaved. Then we replace the previous formulation of the paradifferential equations with

$$\Box \phi_k^l = 2(\mathbf{S}_{il}^j(\phi)_{<k-C} - \mathbf{S}_{ij}^l(\phi)_{<k-C})\partial^\alpha \phi_{<k-C}^i \partial_\alpha \phi_k^j + \text{error}$$

The gain in doing this is that we obtain an antisymmetric coefficient

$$(A_{lj}^\alpha)_{<k-C} = (\mathbf{S}_{ij}^l(\phi)_{<k-C} - \mathbf{S}_{il}^j(\phi)_{<k-C})\partial^\alpha \phi_{<k-C}^i$$

The next lemma makes this frequency localization rigorous.

Lemma 6.2. *We have*

$$\Box \phi_k = 2A_{<k-C}^\alpha \partial_\alpha \phi_k + error$$

Proof. We need to show that

$$P_k(\mathbf{S}_{ij}^l(\phi)\partial^\alpha \phi^i \partial_\alpha \phi^j) - 2\mathbf{S}_{ij}^l(\phi)_{<k-C}\partial^\alpha \phi_{<k-C}^i \partial_\alpha \phi_k^j = \text{error}$$

$$\partial^\alpha \phi_{<k-C}^i \left(P_k(\mathbf{S}_{il}^j(\phi)\phi^j) - \mathbf{S}_{il}^j(\phi)_{<k-C}\partial_\alpha \phi_k^j \right) = \text{error}$$

We only prove the first of the two, the second is similar but somewhat simpler. Redenote

$$\phi^{(1)} = \mathbf{S}_{ij}^l(\phi), \qquad \phi^{(2)} = \phi^i, \qquad \phi^{(3)} = \phi^j$$

Using the Moser type estimates (4.11), (4.12) for the first term it follows that all three satisfy the bounds

$$\|\phi^{(i)}\|_S \lesssim \epsilon, \quad \|\phi^{(i)}\|_{S_c} \lesssim M, \quad i = 1, 2, 3$$

Also note the trivial but important fact that

$$\|\phi^{(1)}\|_{L^\infty} \lesssim 1$$

We use a dyadic decomposition for the first trilinear expression,

$$P_k(\phi^{(1)}\partial^\alpha\phi^{(2)}\partial_\alpha\phi^{(3)}) = \sum_{k_1,k_2,k_3} P_k(\phi^{(1)}_{k_1}\partial^\alpha\phi^{(2)}_{k_2}\partial_\alpha\phi^{(3)}_{k_3})$$

In order for this expression to be nonzero we need

$$k \leq \max\{k_1,k_2,k_3\} + C$$

where C is a fixed large constant. For the highest frequency factor we use the S_c norm, while for the rest we use the S norm. The trilinear estimate (4.8) gives the desired bound for each triplet (k_1,k_2,k_3), and also insures the summation with respect to k_1,k_2,k_3 in the case when $k_1 \geq k - C$.

Consider now the case when $k_1 < k - C \leq k_2 = k_3 + O(1)$. Then

$$P_k\left(\phi^{(1)}_{<k-C}\partial^\alpha\phi^{(2)}_{k_2}\partial_\alpha\phi^{(3)}_{k_3}\right) = P_k\left(\phi^{(1)}_{<k-C}\tilde{P}_k(\partial^\alpha\phi^{(2)}_{k_2}\partial_\alpha\phi^{(3)}_{k_3})\right)$$

therefore we can use (4.7) for the null form estimate and (4.5) for the remaining product.

The above considerations imply that we obtain

$$P_k\left(\phi^{(1)}\partial^\alpha\phi^{(2)}\partial_\alpha\phi^{(3)} - \phi^{(1)}_{<k-C}\partial^\alpha\phi^{(2)}_{<k-C}\partial_\alpha\phi^{(3)} - \phi^{(1)}_{<k-C}\partial^\alpha\phi^{(2)}\partial_\alpha\phi^{(3)}_{<k-C}\right)$$
$$= \text{error}$$

It remains to consider the commutator estimate,

$$\left[P_k, \phi^{(1)}_{<k-C}\partial^\alpha\phi^{(2)}_{<k-C}\right]\tilde{P}_k\partial_\alpha\phi^{(3)} = \text{error}$$

and the similar relation with the indices $2,3$ interchanged. Here we have added the harmless operator $\tilde{P}_k$. As in Lemma 2 in [18] this expression can be rewritten as

$$2^{-k}L\left(\nabla_x(\phi^{(1)}_{<k-C}\partial^\alpha\phi^{(2)}_{<k-C}), \tilde{P}_k\partial_\alpha\phi^{(3)}\right)$$

and further as

$$2^{k_2-k}\sum_{k_2<k-C} L(\phi^{(1)}_{<k_2}, \partial^\alpha\phi^{(2)}_{k_2}, \tilde{P}_k\partial_\alpha\phi^{(3)}) + 2^{k_1-k}\sum_{k_2\leq k_1<k-C} L(\phi^{(1)}_{k_1}, \partial^\alpha\phi^{(2)}_{k_2}, \tilde{P}_k\partial_\alpha\phi^{(3)})$$

For the first term we use (4.7) for the null form estimate and (4.5) for the remaining product. The second term is estimated using the trilinear estimate (4.8).

□

43

It remains to consider the linear equation for ϕ_k. Just as in [18], the difficulty is that we cannot treat the first order term as a negligible error. However, we can do so after conjugation by a gauge transformation $U_{<k-C}$,

$$\phi_k \to U_{<k-C}\phi_k$$

The next Lemma is the counterpart of the Step 2 of Proposition 1 in [18]. We state it in greater generality, as we will also need to use it later in the analysis of the linearized flow.

Lemma 6.3. *Let ϕ be a solution to the wave maps equation which satisfies*

$$\|\phi\|_S \le \epsilon$$

Then for each k there is a gauge transformation $U_{<k-C} \in S$ satisfying

$$\|U_{<k-C}\|_S \lesssim \epsilon, \qquad |U_{<k-C}U_{<k-C}^t - I|_{L^\infty} \lesssim \epsilon$$

so that for all $k \in \mathbb{Z}$ and all $\psi \in S$ we have

$$\|\Box(U_{<k-C}\psi_k) - U_{<k-C}(\Box\psi_k - 2A^\alpha_{<k-C}\partial_\alpha\psi_k)\|_N \lesssim \epsilon\|\psi\|_S$$

We note that one can go one step further and find a gauge transformation U so that $U_{<k} = P_{<k}U$. We do not pursue this as it is not needed for the arguments in this paper.

Proof. The proof is to a certain extent similar to Step 2 of the proof of Proposition 1 in [18]. However, in the context here there are some additional difficulties. Also we are able to simplify the argument somewhat. Consequently, we present a complete proof.

Ideally we would like to have U solve the equation

$$(6.3) \qquad\qquad\qquad \partial^\alpha U = -UA^\alpha$$

However, this equation is overdetermined and has no solutions unless a compatibility condition is satisfied,

$$\partial^\beta A^\alpha - \partial^\alpha A^\beta = [A^\beta, A^\alpha]$$

This is too much to ask, so instead we shall settle for a weaker formulation. To motivate it, recall the form of the A^α's,

$$A^\alpha_{<k-C} = B(\phi)_{<k-C}\partial^\alpha\phi_{<k-C}, \qquad B^l_{ij}(\phi) = \mathbf{S}^l_{ij}(\phi) - \mathbf{S}^j_{il}(\phi)$$

Intuitively the largest contribution to A^α comes from the case when the differentiated factor has the highest frequency. Hence we are led to neglect the spatial dependence of both the U factor and of the undifferentiated factor of A^α in (6.3). Then we arrive at a paradifferential type inductive construction for U, which is similar to the one introduced by

44

Tao [18] in the case of the sphere. What we would naively like to do is to inductively set

$$U = I + \sum_{j=-\infty}^{\infty} U_k \qquad U_k = -U_{<k-C}B(\phi)_{<k-C}\partial^\alpha \phi_k$$

However, this inductive construction is not well defined as it begins at $k = -\infty$. This is not a trivial issue, for even if we apriori knew that U_k are localized at frequency 2^k and

$$\sum_{k=-\infty}^{\infty} \|U_k\|_S^2 < \infty$$

there is no guarantee that the sum $\sum_k U_k$ converges at $-\infty$ in the sense of distributions.

To remedy this we consider first the case when the construction begins at some finite index i. Hence we set

$$U_{<j}^{(i)} = I + \sum_{k=i}^{j} U_k^{(i)} \qquad U_j^{(i)} = -U_{<j-C}^{(i)}B^\alpha(\phi)_{<j-C}\partial^\alpha \phi_j, \quad j \geq i$$

We use induction with respect to j to prove that $U^{(i)}$ satisfies the bounds

(6.4) $\qquad \|U_j^{(i)}\|_S \leq \epsilon^{-\frac{1}{2}}\|\phi_j\|_S, \qquad \|U_{<j}^{(i)}(U_{<j}^{(i)})^t - I\|_{L^\infty} \leq \dfrac{1}{2}$

Suppose this holds up to j and prove it for $j + 1$. We have

$$U_{j+1}^{(i)} = -U_{<j+1-C}^{(i)}B(\phi)_{<j+1-C}\phi_{j+1}$$

The induction hypothesis for U on one hand and the Moser estimates (4.11) for B show that

$$\|U_{<j+1-C}^{(i)}\|_{S\cap L^\infty} \lesssim 1, \qquad \|B(\phi)_{<j+1-C}\|_{S\cap L^\infty} \lesssim 1,$$

Hence, using the the algebra property for $S \cap L^\infty$ we obtain

$$\|U_{j+1}^{(i)}\|_S \lesssim \|\phi_{j+1}\|_S$$

On the other hand, since B is antisymmetric we have

$$U_k^{(i)}(U_{<k-C}^{(i)})^t = 0$$

therefore

$$U_{<j+1}^{(i)}(U_{<j+1}^{(i)})^t - I = \sum_{\substack{k,l<j+1}}^{|k-l|\leq C} U_k^{(i)}(U_l^{(i)})^t$$

45

From Cauchy-Schwartz we obtain

$$\|U_{<j+1}^{(i)}(U_{<j+1}^{(i)})^t - I\|_{L^\infty} \lesssim \sum_{k=i}^{j} \|U_k^{(i)}\|_{L^\infty}^2 \lesssim \sum_{k=i}^{j} \|U_k^{(i)}\|_S^2 \lesssim \epsilon$$

If ϵ is small enough then this concludes the inductive proof of (6.5), and also shows that in effect we have the better bound

(6.5) $$\|U_j^{(i)}\|_S \lesssim \|\phi_j\|_S$$

We can also use this to reevaluate

(6.6) $$\|U_{<j}^{(i)}(U_{<j}^{(i)})^t - I\|_{L^\infty} \lesssim \|\phi_{<j}\|_S^2 \lesssim \epsilon^2$$

Now we can obtain the functions $U_{<j}$ by letting $i \to -\infty$. For each $j \in \mathbb{Z}$ the functions $U_{<j}^{(i)}$ are equibounded and equicontinuous, therefore on a subsequence, which for convenience we still denote by $U_{<j}^{(i)}$, they will simultaneously converge uniformly on compact sets to a family of functions $U_{<j}$. If we set $U_j = U_{<j+1} - U_{<j}$ then also $U_j^{(i)}$ converge to U_j uniformly on compact sets. Hence U_j must be localized at frequency $\approx 2^j$ and, by Theorem 4.1(vii) and (6.5), $U_j \in S$ and

$$\|U_j\|_S \lesssim \|\phi_j\|_S$$

Also we get the recurrence relation

$$U_k = -U_{<k-C}B(\phi)_{<k-C}\partial^\alpha \phi_k$$

Now we evaluate the difference in the Lemma,

$$
\begin{aligned}
R_k &= \Box(U_{<k-C}\psi_k) - U_{<k-C}(\Box\psi_k - 2A_{<k-C}^\alpha \partial_\alpha \psi_k) \\
&= 2(\partial^\alpha U_{<k-C} + U_{<k-C}A_{<k-C}^\alpha)\partial_\alpha \psi_k + \psi_k \Box U_{<k-C}
\end{aligned}
$$

The first part of R_k is rewritten as

$$\sum_{j<k-C} (\partial^\alpha U_j + U_{<j-C}A_j^\alpha)\partial_\alpha \psi_k + U_{j-C\leq\cdot<k-C}A_j^\alpha \partial_\alpha \psi_k =$$

$$\sum_{j<k-C} \partial^\alpha(U_{<j-C}B(\phi)_{<j-C})\phi_j\partial_\alpha \psi_k + U_{j-C\leq\cdot<k-C}B(\phi)_{<j-C}\partial^\alpha \phi_j \partial_\alpha \psi_k$$

For the first term we bound the product $U_{<j-C}B(\phi)_{<j-C}$ in S and then use the trilinear estimate (4.8). For the second we neglect the harmless low frequency factor $B(\phi)_{<j-C}$ and estimate the remaining trilinear form by (4.8).

It remains to estimate the second part of R_k, namely $\psi_k\Box U_{<k-C}$. For this we would like to inductively prove the bound

(6.7) $\|\psi_k\Box U_j\|_N \lesssim \epsilon 2^{-\delta(k-j)}\|\psi_k\|_S$, $\qquad j < k - C$, $\qquad 0 < \delta < \delta_0$

Unfortunately we do not have a good starting point for the induction. However, this difficulty is avoided by proving uniform bounds for $U_j^{(i)}$ and passing to the limit. In the case of $U_j^{(i)}$ the bound is trivial for $j < i$, so it remains to do the induction step. We assume that the above bound holds up to $j - 1$ and prove it for j.

We have

$$
\begin{aligned}
-\Box U_j^{(i)} &= \Box U_{<j-C}^{(i)} B(\phi)_{<j-C} \phi_j + 2\partial^\alpha U_{<j-C}^{(i)} \partial_\alpha (B(\phi)_{<j-C} \phi_j) \\
&+ 2U_{<j-C}^{(i)} \partial^\alpha B(\phi)_{<j-C} \partial_\alpha \phi_j + U_{<j-C}^{(i)} \Box B(\phi)_{<j-C} \phi_j \\
&+ U_{<j-C}^{(i)} B(\phi)_{<j-C} \Box \phi_j
\end{aligned}
$$

For the first term we need to bound

$$
\|\Box U_{<j-C}^{(i)} B(\phi)_{<j-C} \phi_j \psi_k\|_N \lesssim \epsilon 2^{-\delta(k-j+C)} \|\psi_k\|_S
$$

Due to the algebra property (4.3) we can include $B(\phi)_{<j-C} \phi_j$ into ψ_k and then conclude using the induction hypothesis.

For the second we use (4.3) to include $B(\phi)_{<j-C}$ in ϕ_j and then finish with the trilinear estimate (4.8). For the third we include $U_{<j-C}^{(i)}$ into ψ_k and then use the trilinear estimate. In both cases it is important that the undifferentiated factor has the highest frequency.

For the last two terms we use again (4.3) to include all undifferentiated factors in ψ_k. Writing down the equation for $B(\phi)$, we need the bound

$$
\|\psi_k P_j[S(\phi)\partial^\alpha \phi \partial_\alpha \phi]\|_N \lesssim \epsilon 2^{-\delta(k-j)} \|\psi\|_S, \qquad j < k - C
$$

where S is a smooth bounded function with uniformly bounded derivatives. This is essentially the bound (68) in [18], but for the sake of completeness we also discuss it here. If we do a dyadic decomposition for the last three factors then we need to estimate

$$
\sum_{k_2, k_3 \in \mathbb{Z}} (\|\psi_k P_j[S(\phi)_{<j+C} \partial^\alpha \phi_{k_2} \partial_\alpha \phi_{k_3}]\|_N + \sum_{k_1 \geq j+C} \|\psi_k P_j[S(\phi)_{k_1} \partial^\alpha \phi_{k_2} \partial_\alpha \phi_{k_3}]\|_N)
$$

The first expression is rewritten as

$$
\|\psi_k P_j[S(\phi)_{<j+C} P_{<j+2C}(\partial^\alpha \phi_{k_1} \partial_\alpha \phi_{k_2})]\|_N
$$

We discard P_j and include $S(\phi)_{<j+C}$ in ψ_k. If $k_2, k_3 \leq \frac{j+k}{2}$ then we discard $P_{<j+2C}$ and use the trilinear estimate (4.8). Else we use (4.7) for the null form and finish with (4.5).

Consider now the second expression. If $\max\{k_1, k_2, k_3\} > \frac{k+j}{2}$ then we use the trilinear estimate (4.8), and then multiply with ψ_k by (4.5). Else, we discard P_j, include $S(\phi)_{k_1}$ in ψ_k and use (4.8) for the remaining trilinear form.

47

$\square$

We can now conclude the proof of Theorem 6.1. From Lemma 6.2 we have

$$\|\square\phi_k - 2A^\alpha_{<k-C}\partial_\alpha\phi_k\|_N \lesssim \epsilon M c_k$$

On the other hand, Lemma 6.3 gives a gauge $U_{<k-C}$ so that

$$\|\square(U_{<k-C}\phi_k) - U_{<k-C}(\square\phi_k - 2A^\alpha_{<k-C}\partial_\alpha\phi_k)\|_N \lesssim \epsilon\|\phi_k\|_S \lesssim \epsilon M c_k$$

Since $U_{<k-C} \in S \cap L^\infty$, we can use (4.5) to combine the two relations and obtain

$$\|\square(U_{<k-C}\phi_k)\|_N \lesssim \epsilon M c_k$$

On the other hand, it is easy to see that the initial data for $U_{<k-C}\phi_k$ satisfies

$$\begin{aligned}\|(U_{<k-C}\phi_k)[0]\|_{S[0]} &\lesssim (\|U_{<k-C}\|_{L^\infty} + 2^{-k}\|\partial_t U_{<k-C}\|_{L^\infty})\|\phi_k[0]\|_{S[0]} \\ &\lesssim \epsilon M c_k\end{aligned}$$

Then we can use (4.1) for $U_{<k-C}\phi_k$ to obtain

$$\|U_{<k-C}\phi_k\|_S \lesssim \epsilon M c_k$$

For small enough ϵ this gives (6.2).

To prove (6.1) we define a weight c by

$$c_k = \sum_{j\in\mathbb{Z}} 2^{-\delta_1|j-k|}(\epsilon^{-1}\|\phi_k\|_S + \epsilon^{-2}\|\phi_k[0]\|_{S[0]})$$

Then c is admissible while

$$\|c\|_{l^2} \lesssim \delta^{-1}$$

and

$$\|\phi\|_{S_c} \lesssim \epsilon, \qquad \|\phi[0]\| \lesssim \epsilon^2$$

We apply (6.2) with $M \approx \epsilon\delta^{-1}$ to obtain

$$\|\phi\|_{S_c} \lesssim \epsilon^2$$

which implies

$$\|\phi\|_S \lesssim \epsilon^2\delta^{-1}$$

For ϵ small enough this gives (6.1).

$\square$

7. The linearized equations and stability estimates

If we differentiate (1.3) then we obtain the linearized equations

$$(7.1) \qquad \Box \psi^l = -(\partial_m \mathbf{S}^l_{ij})(\phi))\psi^m \partial^\alpha \phi^i \partial_\alpha \phi^j - 2\mathbf{S}^l_{ij}(\phi)\partial^\alpha \phi^i \partial_\alpha \psi^j$$

The function ψ must satisfy the compatibility condition

$$(7.2) \qquad\qquad\qquad \psi(t,x) \in T_{\phi(t,x)}M$$

Understanding the behavior of these equations is the key to comparing different solutions of the wave maps equation. Our main result asserts that

Theorem 7.1. *There is $\epsilon > 0$ so that for any smooth solution ϕ to the wave maps equation in $[-T, T]$ with uniformly bounded derivatives and which satisfies*

$$\|\phi\|_S \lesssim \epsilon^2$$

the linearized equation (7.1), is well-posed for initial data in $S_c[0]$ which satisfies the above compatibility condition, and

$$\|\psi\|_{S_c} \lesssim \|\psi[0]\|_{S_c[0]}$$

Here c is any admissible frequency envelope for the linearized equation.

Corollary 7.2. *Under the assumptions in the theorem the linearized equation is well-posed for initial data in $S_{-\delta}[0]$ for sufficiently small positive δ and the solution ψ satisfies*

$$\|\psi\|_{S_{-\delta}} \lesssim \|\psi[0]\|_{S_{-\delta}[0]}$$

Proof of Theorem 7.1. Since this time we want to solve a linear equation, one might expect to use a fixed point argument based on the constant coefficient d'Allembertian. The difficulty with this approach is that our estimates use the compatibility condition in an essential way, and it seems difficult to construct approximate solutions which satisfy it. Consequently, we fall back on the type of bootstrap arguments which have served us well in the estimates for the full wave maps equation.

Theorem 5.3 (b) already shows that the linearized equation is well-posed in $S_c[0]$. However, the bounds it provides are not uniform with respect to wave maps ϕ whose initial data is bounded in $S[0]$; instead they depend on uniform bounds for higher derivatives of ϕ. Nevertheless, we can use it as a starting point for our estimates and bootstrap the S_c norm of ψ.

An intermediate step is to obtain paradifferential type approximate equations for the frequency localized components ψ_k of ψ. Since ψ has

lower Sobolev regularity than ϕ, there is an additional gain every time a low frequency in ψ contributes to a higher frequency output. Hence we keep only the dyadic factors where ψ has the largest frequency. On the other hand, the first term on the right has ψ undifferentiated. Hence by the trilinear estimate there is some gain unless the ψ factor has the lowest frequency. Thus we should be able to eliminate the first term altogether. This would suggest the paradifferential formulation

$$\Box \psi_k^l \approx -2\mathbf{S}_{ij}^l(\phi)_{<k-C}\partial^\alpha \phi_{<k-C}^i \partial_\alpha \psi_k^j$$

However, as before, we would like to have an antisymmetric gradient potential. For this we observe that $S_{il}^j(\phi)\psi^j = 0$ which yields

$$(\partial_m \mathbf{S}_{il}^j(\phi))\partial_\alpha \phi^m \psi^j + \mathbf{S}_{il}^j(\phi)\partial_\alpha \psi^j = 0$$

The first term corresponds to the first term in (7.1) and can be neglected because ψ is undifferentiated. Thus we conclude that the expression

$$\mathbf{S}_{il}^j(\phi)_{<k-C}\partial_\alpha \psi_k^j$$

should be better behaved. This leads to the modified paradifferential formulation

$$\Box \psi_k^l \approx 2(\mathbf{S}_{il}^j(\phi)_{<k-C}\partial^\alpha - \mathbf{S}_{ij}^l(\phi)_{<k-C})\partial^\alpha \phi_{<k-C}^i \partial_\alpha \psi_k^j$$

which has an antisymmetric gradient potential (the same in fact as for the wave maps equation). This is made precise in the following Lemma.

Lemma 7.3. *Let c be an admissible weight for the linearized equation. Suppose $\psi \in S_c$ is a smooth solution for the linearized equation (7.1) satisfying the compatibility condition (7.2). Then for small enough ϵ the bounds*

$$\|\phi\|_S \lesssim \epsilon, \qquad \|\psi\|_{S_c} \lesssim M \qquad \|\psi[0]\|_{S_c[0]} \lesssim \epsilon M$$

imply

(7.3) $$\Box \psi_k = 2A_{<k-C}^\alpha \partial_\alpha \psi_k + error, \qquad \|error\|_N \lesssim \epsilon M c_k$$

Proof. Comparing (7.3) with (7.1) we need to show that

(7.4) $$P_k\left((\partial_m \mathbf{S}_{ij}^l)(\phi)\psi^m \partial^\alpha \phi^i \partial_\alpha \phi^j\right) = error$$

(7.5) $$P_k\left(\mathbf{S}_{ij}^l(\phi)\partial^\alpha \phi^i \partial_\alpha \psi^j\right) - 2\mathbf{S}_{ij}^l(\phi)_{<k-C}\partial_{<k-C}^\alpha \phi^i \partial_\alpha \psi_k^j = error$$

(7.6) $$\partial_{<k-C}^\alpha \phi^i P_k\left(\partial_m \mathbf{S}_{il}^j(\phi)\partial_\alpha \phi^m \psi^j\right) = error$$

(7.7) $$\partial_{<k-C}^\alpha \phi^i \left(P_k(\mathbf{S}_{il}^j(\phi)\partial_\alpha \psi^j) - \mathbf{S}_{il}^j(\phi)_{<k-C}\partial_\alpha \psi_k^j\right) = error$$

50

The proof of (7.6), respectively (7.7) is a simplified version of the proof of (7.4), respectively (7.5). Hence we focus our attention on the first two and leave the rest to the reader. For (7.4) we redenote

$$(\partial_m \mathbf{S}^l_{ij})(\phi) = \phi^{(1)}, \quad \phi^i = \phi^{(2)}, \quad \phi^j = \phi^{(3)}, \quad \psi^m = \psi$$

Then we need to estimate the expression

$$P_k(\phi^{(1)}\psi\partial^\alpha\phi^{(2)}\partial_\alpha\phi^{(3)}) = I + II$$

where

$$I = P_k(\phi^{(1)}_{<k-C}\psi\partial^\alpha\phi^{(2)}\partial_\alpha\phi^{(3)}), \quad II = P_k(\phi^{(1)}_{\geq k-C}\psi\partial^\alpha\phi^{(2)}\partial_\alpha\phi^{(3)})$$

We rewrite I as

$$I = P_k\left(\phi^{(1)}_{<k-C}\tilde{P}_k(\psi\partial^\alpha\phi^{(2)}\partial_\alpha\phi^{(3)})\right)$$

We discard P_k and use (4.5) to also discard $\phi^{(1)}_{<k-C}$. For the rest we can use the trilinear estimate (4.8).

For II we do a full dyadic decomposition,

$$II = \sum_{k_0,k_1,k_2,k_3}^{k_1 > k-C} P_k(\phi^{(1)}_{k_1}\psi_{k_0}\partial^\alpha\phi^{(2)}_{k_2}\partial_\alpha\phi^{(3)}_{k_3})$$

If $k_0 < k - C$ then we rewrite the summand as

$$P_k\left(\psi_{k_0}\tilde{P}_k(\phi^{(1)}_{k_1}\partial^\alpha\phi^{(2)}_{k_2}\partial_\alpha\phi^{(3)}_{k_3})\right)$$

We discard P_k, use (4.5) to deal with ψ_{k_0} and the trilinear estimate for the remaining factor.

If $k_0 \geq k - C$ and $k_1 < k + C$ then we write the summand as

$$\sum_{j \leq k+2C} P_k\left(\phi^{(1)}_{k_1}P_j(\psi_{k_0}\partial^\alpha\phi^{(2)}_{k_2}\partial_\alpha\phi^{(3)}_{k_3})\right)$$

Then we use the trilinear estimate for $P_j[\psi_{k_0}\partial^\alpha\phi^{(2)}_{k_2}\partial_\alpha\phi^{(3)}_{k_3}])$ and (4.5) for the last product.

If $k_0 \geq k - C$ and $k_1 \geq k + C$ then we write the summand as

$$P_k\left(\phi^{(1)}_{k_1}\tilde{P}_{k_1}(\psi_{k_0}\partial^\alpha\phi^{(2)}_{k_2}\partial_\alpha\phi^{(3)}_{k_3})\right)$$

Then we use the trilinear estimate for $P_{k_1}[\psi_{k_0}\partial^\alpha\phi^{(2)}_{k_2}\partial_\alpha\phi^{(3)}_{k_3}]$ and (4.5) for the last product.

Next we consider (7.5). This time we redenote

$$(\mathbf{S}^l_{ij})(\phi) = \phi^{(1)}, \quad \phi^i = \phi^{(2)}, \quad \psi^m = \psi$$

Then we need to estimate the expression

$$P_k(\phi^{(1)}\partial^\alpha\phi^{(2)}\partial_\alpha\psi) - 2\phi^{(1)}_{<k-C}\partial^\alpha_{<k-C}\phi^{(2)}\partial_\alpha\psi_k = I + II$$

where
$$I = 2[P_k, \phi^{(1)}_{<k-C} \partial^\alpha \phi^{(2)}_{<k-C}] \tilde{P}_k \psi,$$
$$II = P_k[(\phi^{(1)} \partial^\alpha \phi^{(2)} - \phi^{(1)}_{<k-C} \partial^\alpha \phi^{(2)}_{<k-C}) \partial_\alpha \psi]$$

Observe that the commutator estimate in I was already done in the proof of Lemma 6.2. We split the low and high $\phi^{(1)}$ frequencies in II, $II = II(a) + II(b)$ where

$$II(a) = P_k \left(\phi^{(1)}_{<k-C} \tilde{P}_k \left(\sum_{k_0,k_2}^{k_2 \geq k-C} \partial^\alpha \phi^{(2)}_{k_2} \partial_\alpha \psi_{k_0} \right) \right)$$

$$II(b) = \sum_{k_0,k_1,k_2}^{k_1,k_2 \geq C-k} P_k(\phi^{(1)}_{k_1} \partial^\alpha \phi^{(2)}_{k_2} \partial_\alpha \psi_{k_0})$$

The frequency localizations in $II(a)$ force either $k_0 = k_2 + O(1) \geq k + O(1)$ or $k_0 \leq k_2 + O(1) = k + O(1)$. In both cases we can discard P_k, disregard $\phi^{(1)}_{<k-C}$ (by (4.5)) and use (4.7) for the null form.

Finally, $II(b)$ can be dealt with using the trilinear estimate.

$\square$

We can now conclude the proof of Theorem 7.1. Suppose that
$$\|\phi\|_S \lesssim \epsilon^2, \quad \|\psi[0]\|_{S_c[0]} = \epsilon, \quad \|\psi\|_{S_c} = M_0$$

Then Lemma 7.3 with $M = M_0 + 1$ shows that
$$\|\Box\psi_k - 2A^\alpha_{<k-C} \partial_\alpha \psi_k\|_N \lesssim \epsilon M c_k$$

Using the gauge change U given by Lemma 6.3 this implies that
$$\|\Box(U_{<k-C}\psi_k)\|_N \lesssim \epsilon M c_k$$

On the other hand, for the initial data we have the straightforward bound
$$\|(U_{<k-C}\psi_k)[0]\|_{S[0]} \lesssim \epsilon M c_k$$

Using the linear estimate (4.1) we conclude that
$$\|U_{<k-C}\psi_k\|_S \lesssim \epsilon M c_k$$

Inverting U in $S \cap L^\infty$ yields
$$\|\psi_k\|_S \lesssim \epsilon M c_k$$

Then
$$M_0 = \|\psi\|_{S_c} \lesssim \epsilon M = \epsilon(M_0 + 1)$$

which implies
$$M_0 \lesssim \epsilon$$

$\square$

8. Proof of Theorem 1.2

8.1. A continuity argument.

We consider some arbitrary $T > 0$ and an initial data set $\phi[0]$ with

$$\|\phi[0]\|_{S[0]} \leq \epsilon^2$$

For this initial data we consider the approximations $\phi^h[0]$ constructed in the proof of Theorem 3.9. These are smooth functions, depending smoothly on h. Furthermore, they satisfy the uniform bounds

$$\|\phi^h[0]\|_{S[0]} \lesssim \epsilon^2,$$

and the h dependent bounds

$$\|\phi^h[0]\|_{S_m[0]} \lesssim 2^{hm} \qquad m > 0$$

From Sobolev embeddings we get the uniform bounds

$$\|\partial^\beta \phi^h[0]\|_{\dot{W}^{1,\infty} \times L^\infty} \lesssim c_\beta 2^{h(1+|\beta|)}$$

Theorem 5.1(b) shows that for small enough h the wave maps equation with initial data $\phi^h[0]$ has an unique small smooth solution ϕ^h in $C(-T, T; H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$. By Theorem 5.2(a) this solution satisfies

$$\tag{8.1} \|\phi^h\|_S \leq \epsilon$$

We claim that in effect smooth solutions satisfying (8.1) exist for all $h \in \mathbb{R}$. To prove this we use a continuity argument. We denote by $A \subset (-\infty, \infty)$ the set of those h for which there is a smooth solution ϕ^h in $[-T, T]$ satisfying (8.1). As argued before, A is nonempty. If we prove that it is both open and close then it follows that $A = (-\infty, \infty)$.

A **is open**. Let $h_0 \in A$. Then the wave maps equation with initial data $\phi^{h_0}[0]$ has a smooth solution

$$\phi^{h_0} \in S \cap C(-T, T; H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$$

Since $\phi^h[0]$ depends smoothly on h in $S[0] \cap H^k_{(0,\infty)} \times H^{k-1}_{(0,\infty)}$, by Theorem 5.3(a) it follows that for h close to h_0 there is also a smooth solution

$$\phi^h \in S \cap C(-T, T; H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$$

depending smoothly on h in the $S \cap C(-T, T; H^k_{(0,\infty)} \times H^{k-1}_{(0,\infty)})$ topology. Hence for h close to h_0 we have

$$\|\phi^h\|_S \leq 2\epsilon$$

Then theorem 6.1 shows that in effect (8.1) must hold.

A **is closed**. Let $h_j \in A$, $h_j \to h$. Then the corresponding solutions ϕ^{h_j} are uniformly bounded in S. Furthermore, their initial data are uniformly bounded in $S_N[0]$ for large N. By Theorem 5.2(b)

53

the solutions ϕ^{h_j} are in S_N for all N. Then Theorem 6.1 shows that the solutions u^{h_j} are in effect uniformly bounded in S_N. This implies that their derivatives are uniformly bounded. Since $\phi^{h_j}[0] \to \phi^h[0]$ uniformly, by Arzela-Ascoli we have a subsequence which converges uniformly on compact sets, together with all its derivatives. The limit is in $C(-T, T; H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$ for some large k. It must also solve the wave maps equation in $[-T, T]$ and has initial data $\phi^h[0]$. Then this solution is also in S, and the above argument shows that in effect ϕ^h is the strong limit of ϕ^{h_j} in S.

8.2. **Construction of rough solutions.** Consider an initial data set $\phi[0]$ which is small in $S[0]$. The previous step shows that the approximate solutions ϕ^h with initial data $\phi^h[0]$ are uniformly small in S. Since $\phi[0] \in S[0]$ there must be some admissible frequency envelope $c \in l^2$ so that $\phi[0]$ is small in $S_c[0]$. Then $\phi^h[0]$ are uniformly small in $S_c[0]$ therefore by Theorem 5.2(b) we have $\phi^h \in S_c$. We can bootstrap this with Theorem 6.1 to obtain that ϕ^h are uniformly small in S_c.

On the other hand the approximate solutions ϕ^h depend smoothly on h in the $S \cap C(-T, T; H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$ topology. Then the functions $\frac{d}{dh}\phi^h$ also belong to $S \cap C(-T, T; H^k_{(1,\infty)} \times H^{k-1}_{(0,\infty)})$ and solve the linearized equation. For small positive δ their initial data $\frac{d}{dh}\phi^h[0]$ satisfies

$$\|\frac{d}{dh}\phi^h[0]\|_{S^{-\delta}[0]} \lesssim 2^{-\delta h}$$

By Theorem 5.3(b) the functions $\frac{d}{dh}\phi^h$ are in $S^{-\delta}[0]$. We can use Theorem 7.1 to bootstrap their regularity and obtain

$$\|\frac{d}{dh}\phi^h\|_{S^{-\delta}} \lesssim 2^{-\delta h}$$

Integrating this, it follows that there exists a limit function ϕ so that

$$\|\phi - \phi^h\|_{S_{-\delta}} \lesssim 2^{-\delta h}$$

Since ϕ^h are also uniformly bounded in S_c, it also follows that

$$\phi^h \to \phi \text{ in } S$$

It is now easy to pass to the limit and see that ϕ solves the wave maps equation in the sense of distributions.

8.3. **Sobolev regularity of solutions for the wave maps equation.** Let $\phi[0]$ which is small in S, and $\phi \in S$ the solution constructed

above. Suppose that in addition $\phi[0] \in S_c[0]$ for some admissible frequency envelope c. Then the approximate initial data $\phi^h[0]$ are uniformly bounded in $S_c[0]$,

$$\|\phi^h[0]\|_{S_c[0]} \lesssim \|\phi[0]\|_{S_c[0]}$$

By Theorem 5.2(b) it follows that ϕ^h are in S_c. We can bootstrap this with Theorem 6.1 and show that ϕ^h are uniformly bounded in S_c,

$$\|\phi^h\|_{S_c} \lesssim \|\phi^h[0]\|_{S_c[0]} \lesssim \|\phi[0]\|_{S_c[0]}$$

Passing to the limit we obtain

$$\|\phi\|_{S_c} \lesssim \|\phi[0]\|_{S_c[0]}$$

8.4. **Smooth solutions to the wave maps equations.** Let $\phi[0]$ which is smooth and small in $S[0]$. Given $R > 0$ we use Proposition 3.14 to construct a smooth initial data set $\phi^R[0]$ which is small in $S[0]$, is constant outside B_{2R} and agrees with $\phi[0]$ in $B(0, R)$. Then $\phi^R[0] \in S^c[0]$ for all admissible frequency envelopes c. This implies that the corresponding solution ϕ^R is in S^c. Since its initial data is smooth, the solutions ϕ^R must be smooth. Because of the finite speed of propagation these solutions must coincide on the common cone of influence. Letting $R \to \infty$ we obtain a global smooth solution for the initial data $\phi[0]$.

8.5. **Weak stability estimates.** Consider two smooth initial data sets $\phi^{(1)}[0]$ and $\phi^{(2)}[0]$ which are small in $S[0]$ and so that

$$\phi^{(1)}[0] - \phi^{(2)}[0] \in S_{-\delta}[0]$$

Following Proposition 3.13, we construct a smooth one parameter family $\phi^{(\theta)}[0]$, $\theta \in [0, 1]$ of initial data sets which joins them so that $\phi^{(\theta)}$ are uniformly small in $S[0]$ and

$$\int_0^1 \|\frac{d}{d\theta}\phi^{(\theta)}[0]\|_{S_{-\delta}[0]} \lesssim \|\phi^{(1)}[0] - \phi^{(2)}[0]\|_{S_{-\delta}[0]}$$

The corresponding solutions are smooth and small in S, depending smoothly on $\theta \in (0, 1)$ in the S topology.

The solutions $\frac{d}{d\theta}\phi^{(\theta)}$ to the linearized equation are also smooth. By Theorem 5.3(b) they must be in $S_{-\delta}$, and by Theorem 7.1 we can in effect estimate

$$\|\frac{d}{d\theta}\phi^{(\theta)}\|_{S_{-\delta}} \lesssim \|\frac{d}{d\theta}\phi^{(\theta)}[0]\|_{S_{-\delta}[0]}$$

Integrating this we conclude that

$$\|\phi^{(1)} - \phi^{(2)}\|_{S_{-\delta}} \lesssim \|\phi^{(1)}[0] - \phi^{(2)}[0]\|_{S_{-\delta}[0]}$$

We have proved this for smooth solutions, but rough solutions are constructed as limits of smooth solutions in $S \cap S_{-\delta}$, therefore the above inequality is easily transferred to rough solutions.

8.6. Global continuous dependence.

Consider a sequence $\phi^{(k)}[0]$ of initial data sets which is small in $S[0]$ so that

$$\phi^{(k)}[0] - \phi[0] \to 0 \quad \text{in } S[0] \cap S_{-\delta}[0]$$

Then it must be equibounded, in the sense that

$$\lim_{j \to \infty} \sup_k \|P_{>j}\phi^{(k)}[0]\|_{S[0]} = 0$$

Hence we can choose a slowly decreasing sequence a_j so that

$$(8.2) \qquad \|P_{>j}\phi^{(k)}[0]\|_{S[0]} \le a_j, \qquad \lim_{j \to \infty} a_j = 0$$

Here slowly decreasing means

$$i < j \implies a_i > a_j > 2^{-\delta|i-j|}a_i$$

where δ is a sufficiently small positive constant. If $\phi[0] \in S[0]$ then $\phi[0] \in S_c[0]$ for some admissible frequency envelope $c \in l^2$. If in addition (8.2) holds then it is easy to see that c can also be chosen so that it satisfies

$$\sum_{i>j} c_i^2 \lesssim a_j^2$$

Since

$$\|\phi^{(k)}\|_{S_c} \lesssim \|\phi^{(k)}[0]\|_{S_c}$$

we also get

$$(8.3) \qquad \|P_{>j}\phi^{(k)}\|_{S} \lesssim a_j$$

which applies uniformly to all solutions $\phi^{(k)}$.

On the other hand the weak stability result shows that

$$\phi^{(k)} - \phi \to 0 \text{ in } S_{-\delta}$$

Combining this with the equiboundedness property in (8.3) we conclude that

$$\phi^{(k)} \to \phi \text{ in } S$$

8.7. **Local continuous dependence.** Let $R > 0$. Consider small initial data sets $\phi^{(k)}[0], \phi[0]$ in $\dot{H}^{\frac{n}{2}} \times H^{\frac{n}{2}-1}(B_R)$ so that

$$\lim_{k \to \infty} \phi^{(k)}[0] = \phi[0] \quad \text{in } \dot{H}^{\frac{n}{2}} \times H^{\frac{n}{2}-1}(B_R)$$

We can extend them to functions in $S = \dot{H}^{\frac{n}{2}} \times H^{\frac{n}{2}-1}(\mathbb{R}^n)$ so that the above property still holds. Furthermore, by Proposition 3.14 we can also arrange that the extensions are initial data sets which are equal and constant outside B_{2R}. For the extension we apply the global stability result. Then we use the finite speed of propagation to conclude that the initial solutions satisfy

$$\|\nabla(\phi^{(k)} - \phi)\|_{L^\infty \dot{H}^{\frac{n}{2}-1}(\{|t|+|x| \leq R\})} \to 0$$

9. Moser type estimates

This section is devoted to the proof of the nonlinear estimates in Theorem 4.3. This result is similar to Moser type estimates, which are proved using the paradifferential calculus. In our case simply using the paradifferential calculus is not good enough. Instead we shall use an iterated version of it.

As discussed before, given $\phi \in S$ we can find an admissible frequency envelope c with $\|c\|_{L^2} \lesssim 1$ so that

$$\|\phi\|_{S_c} \lesssim \|\phi\|_S$$

Hence the estimate (4.11) is a consequence of (4.12). For (4.12) we need to show that

$$(9.1) \qquad \|P_k F(\phi)\|_S \lesssim c_k \|\phi\|_{S_c} (1 + \|\phi\|_S)^N$$

This requires more information about the space S, which is described in the next two subsections.

9.1. **The structure of the S spaces.** Here we summarize Tao's definition of the S spaces. Fortunately we do not need to explicitly use their fine structure. Hence we can omit some of the details, for which we refer the reader to [18].

Due to the estimate (4.2), it suffices to define the S norm of dyadic pieces of S. Following [18] we set

$$(9.2) \quad \|\phi_k\|_S = \|\nabla \phi_k\|_{L^\infty \dot{H}^{\frac{n}{2}-1}} + \|\nabla \phi_k\|_{X^{\frac{n}{2}-1,\frac{1}{2},\infty}} + \sup_{j<k-20} \|Q_{<j}\phi_k\|_{S[k,j]}$$

where the $S[k,j]$ norm above is defined by

$$\|Q_{<j}\phi_k\|_{S[k,j]} = \sup_{\pm} \Big(\sum_{\kappa \in K_l} \|P_{k,\pm\kappa} Q^{\pm}_{<k-2l}\phi\|^2_{S[k,\kappa]} \Big)^{\frac{1}{2}}, \qquad k-j = 2l$$

57

The last term is taken from [18],(79), with the notations used there. Here we do not use the precise form of the $S[k,j]$ norms. In effect most of our estimates can be handled using the bound

(9.3) $$\|Q_{<j}\phi_k\|_{S[k,j]} \lesssim \|\nabla Q_{<j}\phi_k\|_{X^{\frac{n}{2}-1,\frac{1}{2},1}}$$

(see Lemma 8 in [18]). If we combine this with the inequality

$$\|\nabla\phi_k\|_{X^{\frac{n}{2}-1,\frac{1}{2},\infty}} \leq \|\phi_k\|_S$$

we see that at a fixed size for frequency and modulation the S spaces coincide with the the above $\dot{X}^{s,b}$ spaces. The $S[k,j]$ norms come into play only when it comes to describing how these dyadic pieces are fit together.

It is only at the very end where we need an extra piece of information, which roughly says that the $S[k,j]$ type norms are essentially stable with respect to multiplication by bounded functions with sufficiently small frequency and modulation. Precisely,

Lemma 9.1. *Let ψ be a bounded function with Fourier transform supported in $\{|\xi| + |\tau| \lesssim 2^{j-2C}\}$. Then*

$$\|\psi Q_{<j-2C}\phi_k\|_{S[k,j]} \lesssim \|Q_{<j-2C}\phi_k\|_{S[k,j-20]}\|\psi\|_{L^\infty}$$

Proof. With trivial modifications this is essentially the content of [Case 2(b).3(b).2(b)] of Section 16, [18]]. However, it is perhaps not made clear there that only the L^∞ norm of the low frequency factor ψ is needed, therefore we sketch the proof. This uses the notations and the set-up in [18].

The norm on the left is defined as

$$\|\psi Q_{<j-2C}\phi_k\|_{S[k,j]} = \left(\sum_{\kappa \in K_l} \|P_{\kappa,k}(\psi Q_{<j-2C}\phi_k)\|_{S[k,\kappa]}^2\right)^{\frac{1}{2}}, \qquad l = \frac{k-j}{2}$$

Given the frequency localizations of ϕ and ψ, for $\kappa \in K_l$ we can write

$$P_{k,\kappa}(\psi Q_{<j-2C}\phi_k) = \sum_{\substack{\kappa' \in K_{l+10} \\ \kappa' \subset \kappa}}^{\kappa' \subset \kappa} P_{k,\kappa}Q_{<j}(\psi Q_{<j-2C}P_{k,\kappa'}\phi)$$

Then

$$\|P_{k,\kappa}(\psi Q_{<j-2C}\phi_k)\|_{S[k,\kappa]} \lesssim \sum_{\substack{\kappa' \in K_{l+10} \\ \kappa' \subset \kappa}}^{\kappa' \subset \kappa} \|P_{k,\kappa}Q_{<j}(\psi Q_{<j-2C}P_{k,\kappa'}\phi)\|_{S[k,\kappa]}$$

The multiplier $P_{k,\kappa}Q_{<j}$ is disposable, therefore using also [18],(67) we obtain

$$\|P_{k,\kappa}\psi\phi\|_{S[k,\kappa]} \lesssim \|\psi\|_{L^\infty} \sum_{\kappa'\in K_{l+10}}^{\kappa'\subset\kappa} \|Q_{<j-2C}P_{k,\kappa'}\phi\|_{S[k,\kappa']}$$

The conclusion of the Lemma follows after square summation with respect to $\kappa \in K_l$. $\qquad\square$

9.2. Bilinear estimates.

Here we summarize the bilinear estimates which are needed in the proof of the nonlinear estimates. We consider separately the two possible types of bilinear interaction.

(i) high × high → low interaction, $k \le h_1 + O(1) = h_2$**.** The main dyadic bilinear estimate has the form

$$(9.4) \qquad \|P_k(\phi_{h_1}\phi_{h_2})\|_{S_k} \lesssim 2^{-|h_1-k|}\|\phi_{h_1}\|_{S_{h_1}}$$

Closer to the cone there is an improvement, namely

$$(9.5) \qquad \|P_kQ_j\phi_{h_1}\phi_{h_2}\|_{X^{\frac{n}{2},\frac{1}{2},1}} \lesssim 2^{-|h_1-k|-\delta_0|j-k|}\|\phi_{h_1}\|_{S_{h_1}}\|\phi_{h_2}\|_{S_{h_2}},$$

Another improvement occurs when considering input with large modulation but output with small modulation.

$$(9.6)$$
$$\|P_kQ_j[(Q_{j_1}\phi_{h_1})\phi_{h_2}]\|_{X^{\frac{n}{2},\frac{1}{2},1}} \lesssim 2^{-|h_1-k|-(j_1-j)-\delta_0|j_1-k|}\|\phi_{h_1}\|_{S_{h_1}}\|\phi_{h_2}\|_{S_{h_2}},$$

(ii) high × low → high interaction, $k = h_1 + O(1) > h_2$**.** The main dyadic bilinear estimate has the form

$$(9.7) \qquad \|P_k(\phi_{h_1}\phi_{h_2})\|_{S_k} \lesssim \|\phi_{h_1}\|_{S_{h_1}}$$

The improvement near the cone is now

$$(9.8) \qquad \|P_kQ_j\phi_{h_1}\phi_{h_2}\|_{X^{\frac{n}{2},\frac{1}{2},1}} \lesssim 2^{-\epsilon(h_2-j)_+}\|\phi_{h_1}\|_{S_{h_1}}\|\phi_{h_2}\|_{S_{h_2}},$$

Finally, large modulation input yields better small modulation output,

$$(9.9) \qquad \|P_kQ_j[(Q_{j_1}\phi_{h_1})\phi_{h_2}]\|_S \lesssim 2^{-(j_1-j)-\delta_0(h_2-j)_+}\|\phi_{h_1}\|_S\|\phi_{h_2}\|_S,$$

$$(9.10)$$
$$\|P_kQ_j[\phi_{h_1}(Q_{j_1}\phi_{h_2})]\|_S \lesssim 2^{-\frac{\rho+1}{2}(j_1-j)-\delta_0(h_2-j)_++\rho|h_1-h_2|}\|\phi_{h_1}\|_S\|\phi_{h_2}\|_S,$$

where $1/4 \le \rho \le 1$.

The estimates (9.4) and (9.7) are taken from Theorem 4.1 (ii).

The bounds (9.5) and (9.8) come from Lemma 13 in Tao [18]. Although the factor $2^{-|h_1-k|}$ in (9.5) is not explicitly stated in there, it is fairly easy to see that it is in effect obtained in the proof.

The estimates (9.6) and (9.9) are direct consequences of Lemma 12 in Tao [18]. The same applies to (9.10) for $\rho = 1$. On the other hand, (9.10) is very easy to prove if $\rho = 0$, $\delta_0 = 0^2$:

$$
\begin{aligned}
\|\phi_{h_1}(Q_{j_1}\phi_{h_2})\|_{L^2} &\lesssim \|\phi_{h_1}\|_{L^\infty L^2}\|Q_{j_1}\phi_{h_2}\|_{L^2 L^\infty} \\
&\lesssim 2^{\frac{nh_2}{2} - \frac{nh_1}{2}}\|\phi_{h_1}\|_S\|Q_{j_1}\phi_{h_2}\|_{L^2} \\
&\lesssim 2^{-\frac{j}{2} - \frac{nh_1}{2}}\|\phi_{h_1}\|_S\|\phi_{h_2}\|_S
\end{aligned}
$$

Hence in order to obtain the intermediate range of ρ we interpolate and readjust δ_0.

Note that ideally one would like to have (9.10) with $\frac{\rho+1}{2}$ replaced by 1 and ρ replaced by 0. Unfortunately, in the current set-up of the S spaces this seems difficult to obtain.

9.3. **Multilinear paradifferential expansions.** Instead of using a discrete Littlewood-Paley decomposition of u it is more convenient here to use the continuous one. For $h \in \mathbb{R}$ we can write

$$
\frac{d}{dh}F(\phi_{<h}) = \phi_h F'(\phi_{<h})
$$

or in integral form

$$
F(\phi) = F(\phi_{<l}) + \int_l^\infty \phi_h F'(\phi_{<h})dh \tag{9.11}
$$

This suffices for energy estimates, but not for estimates in the S type spaces. Hence we iterate this computation to obtain

$$
\begin{aligned}
F(\phi) &= F(\phi_{<l}) + \sum_{j=1}^{K-1} F^{(j)}(\phi_{<l}) \int_{[l,\infty)^j} \chi(h)\phi_{h_0} \cdots \phi_{h_j}dh \\
&\quad + \int_{[l,\infty)^K} \chi(h)\phi_{h_0} \cdots \phi_{h_K} F^{(K)}(\phi_{h_K})dh
\end{aligned} \tag{9.12}
$$

where by $\chi(h)$ we denote the ordering function

$$
\chi(h) = 1_{h_j \leq h_{j-1} \leq \cdots \leq h_0}
$$

where $j = 1, N$ depending on the context. It is easy to see that if $\phi \in S$ then these integrals converge in $L^\infty L^2$.

---

^{2}Even a $\delta_0 > 0$ is not so difficult to gain using the simplest sector decomposition.

9.4. **An L^∞ bound.** Here we establish the correct L^∞ bounds for the nonlinear terms arising in the above multilinear expansions.

Lemma 9.2. *Assume that G is a smooth, bounded function, with uniformly bounded derivatives. Then for each nonzero multiindex β*

(9.13) $$\|\partial_x^\beta G(\phi_{<h})\|_{L^\infty} \leq c_\beta 2^{\beta h}(\|\phi\|_S + \|\phi\|_S^{|\beta|})$$

(9.14) $$\|\partial^\beta G(Q_{<h}\phi_{<h})\|_{L^\infty} \leq c_\beta 2^{\beta h}(\|\phi\|_S + \|\phi\|_S^{|\beta|})$$

Observe that in effect one can replace $\|\phi\|_S$ by $\sup_{h\in\mathbb{R}}\|\phi_h\|_{L^\infty}$ in the above inequality.

Proof. For the nonlinear expression in (9.13) we use the chain rule to get

$$\partial_x^\beta G(\phi_{<h}) = \sum_{j=1}^{|\beta|} \sum_{\substack{\beta_1+\ldots+\beta_j=\beta}}^{\beta_i\neq 0} (\partial_x^{\beta_1}\phi_{<h})\cdots(\partial_x^{\beta_j}\phi_{<h})G^{(j)}(\phi_{<h})$$

Then the conclusion follows from the trivial linear bound

$$\|\partial_x^\beta \phi_{<h}\|_{L^\infty} \lesssim 2^{\beta h}\|\phi\|_S$$

The same argument applies for (9.14), except that now we need the L^∞ bound

$$\|\partial^\beta Q_{<h}\phi_{<h}\|_{L^\infty} \lesssim 2^{\beta h}\|\phi\|_S$$

$\square$

9.5. **The energy estimates.** The energy estimate part of (9.1) represents a minor extension of the classical Moser estimates.

Lemma 9.3. *Assume that F is a smooth, bounded function, with uniformly bounded derivatives. Then for large enough N we have*

(9.15) $$\|\nabla P_k F(\phi)\|_{L^\infty L^2} \lesssim (1+\|\phi\|_S)^N c_k\|\phi\|_S$$

Proof. We use the representation in (9.11) with small l. We have

$$\begin{aligned}
\|P_k F(\phi_{<l})\|_{L^\infty} &\lesssim 2^{-k}\|\nabla F(\phi_{<l})\|_{L^\infty} \lesssim 2^{-k}\|\nabla\phi_{<l}\|_{L^\infty} \\
&\lesssim 2^{l-k}\sup_{j<l}\|\phi_j\|_{L^\infty}
\end{aligned}$$

therefore

$$\lim_{l\to-\infty} P_k F(\phi_{<l}) = 0 \qquad \text{in } L^\infty$$

Hence in the sense of distributions we can write

$$\nabla P_k F(u) = \int_{-\infty}^{\infty} \nabla P_k[\phi_h F'(\phi_{<h})]dh$$

For fixed h we compute

(9.16) $\quad \nabla P_k[\phi_h F'(\phi_{<h})] = P_k[\nabla \phi_h F'(\phi_{<h})] + P_k[\phi_h \nabla \phi_{<h} F''(\phi_{<h})]$

From here on, we only need to prove fixed time estimates; the time variable is therefore neglected. We consider three cases, depending on how h and k compare.

(i) *Low frequency input, $h < k - 2C$*. Then we can write

$$\nabla P_k[\phi_h F'(\phi_{<h})] = P_k[\nabla \phi_h \tilde{P}_k F'(\phi_{<h})] + P_k[\phi_h \nabla \phi_{<h} \tilde{P}_k F''(\phi_{<h})]$$

We discard P_k and use the straightforward bounds

$$\|\phi_h\|_{L^2} \lesssim 2^{-h} \|\nabla \phi_h\|_{L^2}, \qquad \|\nabla \phi_{<h}\|_{L^\infty} \lesssim 2^h \|\phi\|_S$$

together with (9.13) to obtain

$$\|\nabla P_k[\phi_h F'(\phi_{<h})]\|_{L^2} \lesssim c_N 2^{-N(k-h)} \|\nabla \phi_h\|_{L^2}(1 + \|\phi\|_S^{N+1})$$

This suffices if N is large enough.

(ii) *Intermediate frequency input, $k - 2C < h < k + 2C$*. In this case the h integration is trivial. We bound again the ϕ_h terms in L^2 and everything else in L^∞ to obtain

$$\|\nabla P_k[\phi_h F'(\phi_{<h})]\|_{L^2} \lesssim \|\nabla \phi_h\|_{L^2}(1 + \|\phi\|_S)$$

(ii) *High frequency input, $h > k + 2C$*. If we argue as before, estimating one factor in L^2 and the rest in L^∞, then we fail to gain a small power of 2^{k-h}, which is necessary in order to accommodate all admissible frequency envelopes c. To remedy this we need an $L^2 \times L^2 \to L^2$ bound,

(9.17) $\quad \|P_k(\phi_h \psi)\|_{L^2} \lesssim 2^{\frac{nk}{2}} \|\phi_h\|_{L^2} \|\psi\|_{L^2}, \quad h > k + 2C$

whose proof is left for the reader. In order to be able to use it we need to process further the relation (9.16):

$$\nabla P_k[\phi_h F'(\phi_{<h})] = P_k[\nabla \phi_h \tilde{P}_h F'(\phi_{<k})] + P_k[\phi_h \nabla \phi_{<k} \tilde{P}_h F''(\phi_{<k})]$$
$$+ \int_k^{h-C} P_k[(\nabla \phi_h \phi_l + \phi_h \nabla \phi_l) \tilde{P}_h F''(\phi_{<l}) + \phi_h \nabla \phi_{<l} \phi_l \tilde{P}_h F'''(\phi_{<l})]dl$$
$$+ \int_{h-C}^{h} P_k[(\nabla \phi_h \phi_l + \phi_h \nabla \phi_l) F''(\phi_{<l}) + \phi_h \nabla \phi_{<l} \phi_l F'''(\phi_{<l})]dl$$

where the operators $P_{>h-C}$ are freely inserted based on the relative size of the interacting frequencies.

For the first two terms we bound the $\nabla \phi_h$, respectively the ϕ_h factor in L^2 and the rest in L^∞, using (9.13) for the nonlinear factor. The same works for the first integral term.

For the second integral term, on the other hand, we bound $\nabla\phi_h$, respectively ϕ_h in L^2 and the remaining products also in L^2. Then we use (9.17).

Summing up the results in all cases we get

$$\|\nabla P_k[\phi_h F'(\phi_{<h})]\|_{L^2} \lesssim 2^{\frac{n(k-h)}{2}} \|\nabla\phi_h\|_{L^2}(1 + \|\phi\|_S)^N$$

where it is important that the coefficient of $k - h$ is strictly positive.

$\square$

9.6. **Multilinear estimates.** The key element in our nonlinear analysis is the following multilinear estimate:

Lemma 9.4. *a) If $h_K < h_{K-1} < h_1 < h_0$ then*

$$(9.18) \qquad \|P_k[\phi_{h_0}\cdots\phi_{h_K}]\|_S \lesssim 2^{-|h_0-k|} \prod_{i=0}^{K} \|\phi_{h_i}\|_S$$

b) Assume in addition that $\delta_0(K + 1) < \frac{2}{3}$. Then for $j < \min\{h_1, k\}$ we obtain the following improvement:

$$(9.19) \quad \|P_k Q_j[\phi_{h_0}\cdots\phi_{h_K}]\|_S \lesssim 2^{-|h_0-k|-\delta_0 \sum_{i=1}^{K}(\min(h_i,k)-j)_+} \prod_{i=0}^{K} \|\phi_{h_i}\|_S$$

c) For $\frac{1}{2} < \delta_0(K - 1)$ and $j > h_K$ we have

$$(9.20) \qquad \|P_{<h_K}[\phi_{h_0}\cdots\phi_{h_K}]\|_{L^2} \lesssim 2^{-|h_0-k|-\frac{(n+1)h_K}{2}} \prod_{i=0}^{K} \|\phi_{h_i}\|_S$$

If in addition $j > h_K$ then

$$(9.21) \qquad \|P_{<h_K}Q_j[\phi_{h_0}\cdots\phi_{h_K}]\|_{L^2} \lesssim 2^{-|h_0-k|-\frac{(n-2)h_K+3j}{2}} \prod_{i=0}^{K} \|\phi_{h_i}\|_S$$

Proof. a) The bound (9.18) follows easily from (9.4) and (9.7).

b) To prove (9.19) we use induction with respect to K. For $K = 2$ this follows from the bilinear estimates (9.8) and (9.5). Suppose it holds for K and prove it for $K + 1$. We consider three cases:

Low frequency output, $k < h_{K+1} - C$. Then only the frequency h_{K+1} part of the product $\phi_{h_0}\cdots\phi_{h_K}$ contributes,

$$
\begin{aligned}
P_k Q_j[\phi_{h_0}\cdots\phi_{h_K}] &= P_k Q_j[\phi_{h_{K+1}} \tilde{P}_{h_{K+1}}(\phi_{h_0}\cdots\phi_{h_K})] \\
&= P_k Q_j[\phi_{h_{K+1}} \tilde{P}_{h_{K+1}} Q_{>h_{K+1}}(\phi_{h_0}\cdots\phi_{h_K})] \\
&\quad + P_k Q_j[\phi_{h_{K+1}} \sum_{j_1 < h_{K+1}} \tilde{P}_{h_{K+1}} Q_{j_1}(\phi_{h_0}\cdots\phi_{h_K})]
\end{aligned}
$$

63

The first term is easy to estimate, we simply bound $\phi_{h_{K+1}}$ in L^∞ and the remaining product in L^2, using (9.18). It remains to consider the sum, which we split in two.

For $j_1 \leq j$ we use (9.5) and the induction hypothesis,

$$
\begin{aligned}
\|P_k Q_j [\phi_{h_{K+1}} \tilde{P}_{h_{K+1}} Q_{j_1} (\phi_{h_0} \cdots \phi_{h_K})]\|_S & \\
\lesssim\ & 2^{-|h_{K+1}-k|-\delta_0|j-k|} \|\phi_{h_{K+1}}\|_S \|\tilde{P}_{h_{K+1}} Q_{j_1} (\phi_{h_0} \cdots \phi_{h_K})\|_S \\
\lesssim\ & 2^{-|h_{K+1}-k|-\delta_0|j-k|-|h_0-h_{K+1}|-K\delta_0(h_{K+1}-j_1)_+} \prod_{i=0}^{K+1} \|\phi_{h_i}\|_S \\
\lesssim\ & 2^{-|h_0-k|-(K+1)\delta_0|j-k|-K\delta_0|j-j_1|} \prod_{i=0}^{K+1} \|\phi_{h_i}\|_S
\end{aligned}
$$

We note that the summation with respect to j_1 is trivial.

For $j_1 > j$ we combine (9.6) instead with the induction hypothesis to obtain

$$
\begin{aligned}
\|P_k Q_j [\phi_{h_{K+1}} \tilde{P}_{h_{K+1}} Q_{j_1} (\phi_{h_0} \cdots \phi_{h_K})]\|_S & \\
\lesssim\ & 2^{-|h_{K+1}-k|-(j_1-j)-\delta_0|j-k|} \|\phi_{h_{K+1}}\|_S \|\tilde{P}_{h_{K+1}} Q_{j_1} (\phi_{h_0} \cdots \phi_{h_K})\|_S \\
\lesssim\ & 2^{-|h_{K+1}-k|-(j_1-j)-\delta_0|j-k|} 2^{-|h_0-h_{K+1}|} 2^{-K\delta_0(h_{K+1}-j_1)_+} \prod_{i=0}^{K+1} \|\phi_{h_i}\|_S \\
\lesssim\ & 2^{-|h_0-k|-(K+1)\delta_0|j-k|-(1-K\delta_0)|j-j_1|} \prod_{i=0}^{K+1} \|\phi_{h_i}\|_S
\end{aligned}
$$

The summation with respect to j_1 is again trivial.

Intermediate frequency output, $h_{K+1} - C \leq k < h_{K+1} + C$. Then all frequencies up to $k + 2C$ in the product $\phi_{h_0} \cdots \phi_{h_K}$ contribute,

$$
\begin{aligned}
P_k Q_j [\phi_{h_0} \cdots \phi_{h_{K+1}}] &= P_k Q_j [\phi_{h_{K+1}} P_{\leq k+2C} (\phi_{h_0} \cdots \phi_{h_K})] \\
&= P_k Q_j [\phi_{h_{K+1}} \sum_{h \leq k+2C} P_h Q_{>h} (\phi_{h_0} \cdots \phi_{h_K})] \\
&\quad + P_k Q_j [\phi_{h_{K+1}} \sum_{j_1 < h \leq k+2C} P_h Q_{j_1} (\phi_{h_0} \cdots \phi_{h_K})]
\end{aligned}
$$

64

For the first term we use the energy bound for $\phi_{h_{K+1}}$ and the L^2 bound from (9.18) for the remaining product,

$$\|\phi_{h_{K+1}} P_h Q_{>h}(\phi_{h_0} \cdots \phi_{h_K})\|_{L^2}$$
$$\lesssim \|\phi_{h_{K+1}}\|_{L^\infty L^2} \|P_h Q_{>h}(\phi_{h_0} \cdots \phi_{h_K}))\|_{L^2 L^\infty}$$
$$\lesssim 2^{\frac{nh}{2}} \|\phi_{h_{K+1}}\|_{L^\infty L^2} \|P_h Q_{>h}(\phi_{h_0} \cdots \phi_{h_K}))\|_{L^2}$$
$$\lesssim 2^{-\frac{nh_{K+1}}{2} - \frac{h}{2}} \|\phi_{h_{K+1}}\|_S \|P_h Q_{>h}(\phi_{h_0} \cdots \phi_{h_K}))\|_S$$
$$\lesssim 2^{-\frac{nh_{K+1}}{2} - \frac{h}{2} - |h_0 - h|} \prod_{i=0}^{K+1} \|\phi_{h_i}\|_S$$
$$\lesssim 2^{-\frac{(n+1)h_{K+1}}{2} - |h_0 - h_{K+1}| - \frac{1}{2}|h_{K+1} - h|} \prod_{i=0}^{K+1} \|\phi_{h_i}\|_S$$

The second term is split in two as before. For $j_1 \leq j$ we use (9.9) and the induction hypothesis,

$$\|P_k Q_j[\phi_{h_{K+1}} P_h Q_{j_1}(\phi_{h_0} \cdots \phi_{h_K})]\|_S$$
$$\lesssim 2^{-\delta_0 (h-j)_+} \|\phi_{h_{K+1}}\|_S \|P_h Q_{j_1}(\phi_{h_0} \cdots \phi_{h_K})\|_S$$
$$\lesssim 2^{-\delta_0 (h-j)_+ - |h_0 - h| - K\delta_0 (h-j_1)} \prod_{i=0}^{K+1} \|\phi_{h_i}\|_S$$
$$\lesssim 2^{-|h_0 - k| - (K+1)\delta_0 |k-j| - K\delta_0 |j-j_1| - (1-(K+1)\delta_0)|h-k|} \prod_{i=0}^{K+1} \|\phi_{h_i}\|_S$$

which sums up easily with respect to j_1 and h.

For $j_1 > j$ we use (9.10) instead to obtain

$$\|P_k Q_j[\phi_{h_{K+1}} P_h Q_{j_1}(\phi_{h_0} \cdots \phi_{h_K})]\|_S$$
$$\lesssim 2^{-\delta_0 (h-j)_+ - \frac{\rho+1}{2}(j_1-j) + \rho|k-h|} \|\phi_{h_{K+1}}\|_S \|P_h Q_{j_1}(\phi_{h_0} \cdots \phi_{h_K})\|_S$$
$$\lesssim 2^{-\frac{\rho+1}{2}(j_1-j) + \rho|k-h| - \delta_0|j-h| - |h_0 - h| - K\delta_0(h-j_1)} \prod_{i=0}^{K+1} \|\phi_{h_i}\|_S$$
$$\lesssim 2^{-|h_0 - k| - (K+1)\delta_0 |k-j| - (\frac{\rho+1}{2} - K\delta_0)|j-j_1| - (1-\rho-(K+1)\delta_0)|h-k|} \prod_{i=0}^{K+1} \|\phi_{h_i}\|_S$$

We choose $\rho = \frac{1}{3}$ and sum up with respect to h, j_1.

High frequency output, $k > h_{K+1} + C$. Then only frequencies of size $k + O(1)$ in the product $\phi_{h_0} \cdots \phi_{h_K}$ contribute,

$$
\begin{aligned}
P_k Q_j [\phi_{h_0} \cdots \phi_{h_{K+1}}] &= P_k Q_j [\phi_{h_{K+1}} \tilde{P}_k (\phi_{h_0} \cdots \phi_{h_K})] \\
&= P_k Q_j [\phi_{h_{K+1}} \tilde{P}_k Q_{>k} (\phi_{h_0} \cdots \phi_{h_K})] \\
&+ P_k Q_j [\phi_{h_{K+1}} \sum_{j_1 < k+C} \tilde{P}_k Q_{j_1} (\phi_{h_0} \cdots \phi_{h_K})]
\end{aligned}
$$

For the first term we simply bound $\phi_{h_{K+1}}$ in L^∞ and the remaining product in L^2. The second term is again split in two. For $j_1 \leq j$ we use (9.8) and the induction hypothesis to get

$$
\begin{aligned}
\| &P_k Q_j [\phi_{h_{K+1}} \tilde{P}_k Q_{j_1} (\phi_{h_0} \cdots \phi_{h_K})] \|_S \\
&\lesssim 2^{-\delta_0 (h_{K+1}-j)_+} \| \phi_{h_{K+1}} \|_S \| \tilde{P}_k Q_{j_1} (\phi_{h_0} \cdots \phi_{h_K}) \|_S \\
&\lesssim 2^{-\delta_0 (h_{K+1}-j)_+ - \delta_0 \sum_{i=1}^K (\min(h_i,k)-j_1)_+ - |h_0 - k|} \prod_{i=0}^{K+1} \| \phi_{h_i} \|_S \\
&\lesssim 2^{-\delta_0 \sum_{i=1}^{K+1} (\min(h_i,k)-j)_+ - |h_0 - k| - \delta_0 K |j - j_1|} \prod_{i=0}^{K+1} \| \phi_{h_i} \|_S
\end{aligned}
$$

This we can sum with respect to j_1.

For $j_1 > j$ we use (9.9) instead to obtain

$$
\begin{aligned}
\| &P_k Q_j [\phi_{h_{K+1}} \tilde{P}_k Q_{j_1} (\phi_{h_0} \cdots \phi_{h_K})] \|_S \\
&\lesssim 2^{-\delta_0 (h_{K+1}-j)_+ - (j_1-j)} \| \phi_{h_{K+1}} \|_S \| \tilde{P}_k Q_{j_1} (\phi_{h_0} \cdots \phi_{h_K}) \|_S \\
&\lesssim 2^{-\delta_0 (h_{K+1}-j)_+ - (j_1-j) - \delta_0 \sum_{i=1}^K (\min(h_i,k)-j_1)_+ - |h_0 - k|} \prod_{i=0}^{K+1} \| \phi_{h_i} \|_S \\
&\lesssim 2^{-\delta_0 \sum_{i=1}^{K+1} (\min(h_i,k)-j)_+ - |h_0 - k| - (1 - \delta_0 (K+1 |j - j_1|)} \prod_{i=0}^{K+1} \| \phi_{h_i} \|_S
\end{aligned}
$$

This concludes the proof of part (b).

c) Taking into account (9.19), it suffices to estimate only the low frequencies, namely $P_{<h_K - C}[\phi_{h_0} \cdots \phi_{h_K}]$. For this we can write

$$
P_{<h_K - C}[\phi_{h_0} \cdots \phi_{h_K}] = P_{<h_K - C}[\phi_{h_K} \tilde{P}_{h_K} (\phi_{h_0} \cdots \phi_{h_{K-1}})]
$$

Then

$$
\| P_{<h_K - C}[\phi_{h_0} \cdots \phi_{h_K}] \|_{L^2} \lesssim \| \phi_{h_K} \|_{L^\infty} \| \tilde{P}_{h_K} (\phi_{h_0} \cdots \phi_{h_{K-1}}) \|_{L^2}
$$

The L^∞ norm of the first factor is controlled by its S norm, while for the second we use the result in part (b) with K replaced by $K - 1$. This gives (9.20). For (9.21) we note in addition that either ϕ_{h_K} or

66

$P_{h_K}(\phi_{h_0} \cdots \phi_{h_{K-1}})$ must contribute with modulations at least 2^j, and conclude as before. $\qquad\square$

9.7. Nonlinear low $\to$ high frequency bounds.
Such interactions cannot occur in a bilinear or multilinear setting. While they do occur in our context, the next lemma shows that their output is not too large.

Lemma 9.5. *Assume that F is a smooth, bounded function, with uniformly bounded derivatives. Then for large enough N we have*

$$2^{\frac{n+1}{2}k}\|P_k F(\phi_{<j})\|_{L^2} + \|\nabla P_k F(\phi_{<j})\|_{X^{\frac{n}{2}-1,\frac{1}{2},\infty}} \lesssim$$
$$(9.22) \qquad 2^{-N|j-k|}(1 + \|\phi\|_S)^{3N} \sup_{h \leq j} 2^{-N|h-j|}\|\phi_h\|_S$$

for all $\phi \in S$ and $j \leq k - 2C$.

Combining this with Lemma 9.3 immediately yields

Corollary 9.6. *Under the same assumptions as above we have*

$$(9.23) \qquad \|P_k F(\phi_{<j})\|_S \lesssim 2^{-N|j-k|}(1 + \|\phi\|_S)^{3N} \sup_{h \leq j} 2^{-N|h-j|}\|\phi_h\|_S$$

for all $\phi \in S$ and $j \leq k - 2C$.

Proof of Lemma 9.5. We use the expansion (9.12) for the nonlinear term, but this time we replace $+\infty$ with j. This gives

$$P_k F(\phi_{<j}) = P_k F(\phi_{<l}) + P_k \left(\sum_{i=1}^{N-1} F^{(i)}(\phi_{<l}) \int_{[l,j)^{i+1}} \chi(h)\phi_{h_0} \cdots \phi_{h_i} dh \right.$$
$$\left. + \int_{[l,j)^{N+1}} \chi(h)\phi_{h_0} \cdots \phi_{h_N} F^{(N)}(\phi_{h_N}) dh \right)$$

We claim that, as l approaches $-\infty$, all terms but the last one converge to 0 in L^∞. Indeed, since $j \leq k - 2C$ we can write

$$P_k \left(F^{(i)}(\phi_{<l}) \int_{[l,j)^{i+1}} \chi(h)\phi_{h_0} \cdots \phi_{h_i} dh \right) =$$
$$P_k \left(\tilde{P}_k F^{(i)}(\phi_{<l}) \int_{[l,j)^{i+1}} \chi(h)\phi_{h_0} \cdots \phi_{h_i} dh) \right)$$

Using (9.13) we bound this in L^∞,

$$\left\| P_k \left(F^{(i)}(\phi_{<l}) \int_{[l,j)^{i+1}} \chi(h)\phi_{h_0} \cdots \phi_{h_i} dh \right) \right\|_{L^\infty}$$
$$\lesssim |k-l|^{i+1}\|\phi\|_S^{i+1}\|\tilde{P}_k F^{(i)}(\phi_{<l})\|_{L^\infty}$$
$$\lesssim 2^{-N|k-l|}|k-l|^{i+1}\|\phi\|_S^{i+1}(1 + \|\phi\|_S^N)$$

which decays exponentially at $l = -\infty$. Hence we conclude that

$$(9.24) \qquad P_k F(\phi_{<j}) = \int_{[-\infty,j)^{N+1}} \chi(h)\phi_{h_0}\cdots\phi_{h_N}(\tilde{P}_k F^{(N)}(\phi_{<h_N}))dh$$

We can bound this in L^2 using the multilinear estimates (9.19), (9.21) and (9.20),

$$\|P_k F(\phi_{<j})\|_{L^2} \lesssim \int_{(-\infty,j]^{N+1}} \chi(h)2^{-\frac{(n+1)h_N}{2}}\|\phi_{h_0}\|_S\cdots\|\phi_{h_N}\|_S$$
$$\|\tilde{P}_k F^{(N)}(\phi_{h_N})\|_{L^\infty}dh$$

Integrate with respect to all but h_N, redenote h_N by h and use (9.13) to obtain

$$\|P_k F(\phi_{<j})\|_{L^2} \lesssim \int_{-\infty}^{j} |h-j|^N 2^{-\frac{(n+1)h}{2}} 2^{-2N|j-h|}\|\phi_h\|_S dh(1 + \|\phi\|_S^{3N})$$
$$\lesssim 2^{-\frac{(n+1)k}{2}} 2^{-N|k-j|}(1 + \|\phi\|_S^{3N}) \sup_{h<j} 2^{-N|j-h|}\|\phi_h\|_S$$

This gives (9.22) up to modulation 2^k. It remains to obtain an estimate in the region away from the cone. Hence take $j > k$ and seek an L^2 bound for

$$P_k Q_j[\phi_{h_0}\cdots\phi_{h_N}\tilde{P}_k F^{(N)}(\phi_{<h_N})]$$

At least one of the factors in the above product must have modulation at least 2^{j-C}. The easy case is when this is one of the ϕ_{h_i}'s. Then we can bound it in L^2 and everything else in L^∞. The remaining case is when the ϕ_{h_i}'s have low modulation, which implies that the nonlinear factor must have modulation about 2^j. Hence it remains to consider the expression

$$v = P_k Q_j[\phi_{h_0}\cdots\phi_{h_N}\tilde{Q}_j\tilde{P}_k F^{(N)}(\phi_{<h_N})]$$

Now the trivial L^∞ bound for the nonlinear factor is insufficient, therefore we need to process it further. We write

$$F^{(N)}(\phi_{<h_N}) = F^{(N)}(Q_{<h_N}\phi_{<h_N}) + \int_{h_N}^{\infty} Q_h\phi_{<h_N} F^{(N+1)}(Q_{<h}\phi_{<h_N})dh$$

which leads to a corresponding decomposition of v,

$$v = v_0 + \int_{h_N}^{\infty} v(h)dh$$

where

$$v_0 = P_k Q_j \left(\phi_{h_0}\cdots\phi_{h_N}\tilde{Q}_j\tilde{P}_k F^{(N)}(Q_{<h_N}\phi_{<h_N})\right)$$

$$v(h) = P_k Q_j \left(\phi_{h_0}\cdots\phi_{h_N}\tilde{Q}_j\tilde{P}_k(Q_h\phi_{<h_N}P_{>k-C}F^{(N+1)}(Q_{<h}\phi_{<h_N}))\right)$$

68

To estimate v_0 we use the multilinear L^2 bound (9.20) and (9.14) to obtain

$$
\begin{aligned}
\|v_0\|_{L^2} &\lesssim \|\phi_{h_0}\cdots\phi_{h_N}\|_{L^2}\|\tilde{Q}_j\tilde{P}_k F^{(N)}(Q_{<h_N}\phi_{<h_N})\|_{L^\infty} \\
&\lesssim 2^{-\frac{n+1}{2}h_N}2^{-2N(j-h_N)}(1+\|\phi\|_S)^{3N}\|\phi_{h_N}\|_S
\end{aligned}
$$

This can easily be integrated with respect to $h_0, ..., h_N$.

To estimate the integral term $v(h)$ we note the bound

$$
(9.25)\qquad 2^{-h_N}\|Q_h\phi_{<h_N}\|_{L^2L^\infty} \lesssim \|Q_h\phi_{<h_N}\|_{L^2\dot{H}^{\frac{n}{2}-1}} \lesssim 2^{-\frac{3h}{2}}\|\phi\|_S
$$

If $h \geq j - 2C$ then we discard the disposable operators P_kQ_j and $\tilde{Q}_j\tilde{P}_k$ and use (9.25) and (9.13) to estimate

$$
\begin{aligned}
\|v(h)\|_{L^2} &\lesssim \prod_{i=0}^{N-1}\|\phi_{h_i}\|_{L^\infty}\|\phi_{h_N}\|_{L^\infty L^2}\|Q_h\phi_{<h_N}\|_{L^2L^\infty} \\
&\quad \|P_{>k-C}F^{(N+1)}(Q_{<h}\phi_{<h_N})\|_{L^\infty} \\
\\
&\lesssim 2^{-(\frac{n}{2}-1)h_N}2^{-\frac{3h}{2}}2^{-2N(k-h_N)}(1+\|\phi\|_S^{3N})\|\phi_{h_N}\|_S
\end{aligned}
$$

This can be integrated with respect to $h, h_0, ..., h_N$.

If $h < j - 2C$ then we need to consider only modulations of at least 2^j in the nonlinear term,

$$
v(h) = P_kQ_j\Big(\phi_{h_0}\cdots\phi_{h_N}\tilde{Q}_j\tilde{P}_k(Q_h\phi_{<h_N}Q_{>j-C}P_{>k-C}F^{(N+1)}(Q_{<h}\phi_{<h_N}))\Big)
$$

We discard again the disposable operators P_kQ_j and $\tilde{Q}_j\tilde{P}_k$. However, this time we have a better bound for the nonlinear term,

$$
\|Q_{>j-C}P_{>k-C}F^{(N+1)}(Q_{<h}\phi_{<h_N})\|_{L^\infty} \lesssim 2^{-(2N-3)(k-h_N)}2^{-3(j-h)}(1+\|\phi\|_S^{2N})
$$

which is a variation on the theme of (9.13) and (9.14). We use this instead of (9.13) and proceed as above.

$\square$

9.8. **Nonlinear low $\rightarrow$ high modulation bounds.** Since we are working with a semilinear wave equation, all the interesting action should happen near the characteristic cone. However, due to the particular choice of the space S, some attention must be given to what happens away from the cone. The next Lemma examines the nonlinear output away from the cone arising from input near the cone.

Lemma 9.7. *Assume that F is a smooth, bounded function, with uniformly bounded derivatives. Then for large enough N we have*

$$
(9.26)\qquad 2^{\frac{(n+1)k}{2}}\|Q_kP_{<k}F(Q_{<j}\phi_{<j})\|_{L^2} \lesssim 2^{-N|j-k|}(\|\phi\|_S + \|\phi\|_S)^{3N}
$$

for all $\phi \in S$ and $j \le k - 2C$.

Proof. We proceed as in the proof of the previous Lemma, except that now we reduce both the frequency and the modulation of the input at the same time. Observe that

$$\frac{d}{dh}Q_{<h}\phi_{<h} = Q_h\phi_{<h} + Q_{<h}\phi_h$$

For these components we use the L^∞ bounds

(9.27)
$$\|Q_{<h}\phi_h\|_{L^\infty} + \|Q_h\phi_{<h}\|_{L^\infty} \lesssim \|\phi\|_S$$

the energy bound

(9.28)
$$\|Q_{<h}\phi_h\|_{L^\infty L^2} \lesssim 2^{-\frac{nh}{2}}\|\phi\|_S$$

and the $L^2 L^\infty$ bound

(9.29)
$$\|Q_h\phi_{<h}\|_{L^2 L^\infty} \lesssim 2^{\frac{s}{2}}\|Q_h\phi_{<h}\|_{L^2 H^{\frac{n}{2}-1}} \lesssim 2^{-\frac{s}{2}}\|\phi\|_S$$

Another variation on this theme is the bound

(9.30)
$$\|Q_h\phi_{<h}\|_{L^{2n}} \lesssim 2^{-\frac{s}{2}}\|\phi\|_S$$

Arguing as in the previous Lemma but using (9.14) instead of (9.13) we establish the following analogue of (9.24)

$$Q_k P_{<k} F(Q_{<j}\phi_{<j}) = \int_{[-\infty,j)^{N+1}} \chi(h) Q_k P_{<k} \left((Q_{h_0}\phi_{<h_0} + Q_{<h_0}\phi_{h_0}) \cdots\right.$$
$$\left.(Q_{h_N}\phi_{<h_N} + Q_{<h_N}\phi_{h_N})\tilde{Q}_k F^{(N)}(Q_{<h_N}\phi_{<h_N})\right) dh$$

To estimate this we consider the following cases:

(i) All factors have the form $Q_{<h}\phi_h$. Then we are in the situation of the previous Lemma, and we can use the multilinear estimate (9.20).

(ii) There is at least one factor of the form $Q_{<h}\phi_h$ and one of the form $Q_h\phi_{<h}$. Then we bound the former in $L^2 L^\infty$ and the latter in $L^\infty L^2$. For everything else we use L^∞ bounds.

(iii) All factors have the form $Q_h\phi_{<h}$. Then we use (9.30) exactly n times, and L^∞ bounds for everything else.

$\square$

9.9. Nonlinear high modulation bounds.
Here we consider an extension of the previous Lemma, where we also allow high modulation input but we get less in return.

Lemma 9.8. *Let $k > j + 2C$. Then*
(9.31)
$$2^{\frac{k}{2}}\|Q_k P_{<k} F(\phi_{<j})\|_{L^2 L^\infty} \lesssim 2^{\frac{3k}{2}}\|Q_k P_{<k} F(\phi_{<j})\|_{L^2 L^n} \lesssim (\|\phi\|_S + \|\phi\|_S)^{3N}$$

Proof. We write

$$P_{<k}Q_kF(\phi_{<j}) = P_{<k}Q_kF(Q_{<j}\phi_{<j}) + P_{<k}Q_k\int_j^\infty Q_h\phi_{<j}F(Q_{<h}\phi_{<j})dh$$

The bound for the first term comes from Lemma 9.7.

For the second term we note the estimate

$$\|Q_h\phi_{<j}\|_{L^2L^n} \lesssim \|Q_h\phi_{<j}\|_{L^2\dot{H}^{\frac{n}{2}-1}} \lesssim 2^{\frac{3h}{2}}\|\phi_{<j}\|_S$$

This suffices for $h > k - C$. Otherwise, we rewrite the integrand as

$$P_{<k}Q_k[Q_h\phi_{<j}F(Q_{<h}\phi_{<j})] = P_{<k}Q_k[Q_h\phi_{<j}\tilde{Q}_kF(Q_{<h}\phi_{<j})]$$

and also use the decay of the L^∞ norm for the last factor,

$$\|\tilde{Q}_kF(Q_{<h}\phi_{<j})\|_{L^\infty} \lesssim 2^{-N|h-k|}(\|\phi\|_S + \|\phi\|_S^N)$$

$\square$

9.10. **The frequency localized L^2 bounds.** Here we obtain the core estimate in the proof of Theorem 4.3, which shows that for the most part the nonlinear expression $\nabla F(u)$ can be estimated in the better space $\dot{X}^{\frac{n}{2}-1,\frac{1}{2},\infty}$.

Lemma 9.9. *a) For all $\phi \in S \cap S_c$, large enough N and $j \leq k$ we have*

$$\|P_kQ_{<j}F(\phi) - Q_{<j-2C}\phi_kQ_{<j-2C}P_{<j-2C}F'(\phi_{<j-3C})\|_{\dot{X}^{\frac{n}{2},\frac{1}{2},1}}$$

$$(9.32) \qquad \lesssim c_k\|\phi\|_{S_c}(1 + \|\phi\|_S)^N$$

and

$$(9.33) \qquad \|\nabla P_kF(u)\|_{\dot{X}^{\frac{n}{2}-1,\frac{1}{2},\infty}} \lesssim c_k\|\phi\|_{S_c}(1 + \|\phi\|_S)^N$$

Proof. We begin with the expansion in (9.11) for $F(u)$ with $l = k - 2C$:

$$P_kF(u) = P_kF(\phi_{<k-2C}) + P_k\int_{k-2C}^\infty \phi_hF'(\phi_{<h})dh$$

For the first term we use the bound (9.22). It remains to consider the integral term, for which we distinguish two cases, namely $h > k + 2C$ respectively $k - 2C < h < k + 2C$.

(a): High frequency input ($h > k + 2C$). We prove a slightly better bound than we need,
(9.34)
$$\|P_k[\phi_hF'(\phi_{<h})]\|_{\dot{X}^{\frac{n}{2}+\delta_0,\frac{1}{2}-\delta_0,\infty}\cap\dot{X}^{\frac{n}{2},\frac{1}{2},\infty}} \lesssim 2^{-|k-h|}\|\phi_h\|_S(1 + \|\phi\|_S)^N$$

(a1) High modulation output, $j > h + 2C$. Here we estimate

$$Q_jP_k[\phi_hF'(\phi_{<h})], \qquad j > h + 2C$$

71

At least one of the factors must have a large modulation, therefore we write it as

$$P_k Q_j[Q_{>j-C}\phi_h \tilde{P}_h F'(\phi_{<h})] + P_k Q_j[Q_{<j-C}\phi_h \tilde{Q}_j \tilde{P}_h F'(\phi_{<h})]$$

For the first term we use (9.15) for the nonlinear factor,

$$
\begin{aligned}
\|P_k Q_j[Q_{>j-C}\phi_h \tilde{P}_h F'(\phi_{<h})]\|_{L^2} &\lesssim 2^{\frac{nk}{2}} \|[Q_{>j-C}\phi_h \tilde{P}_h F'(\phi_{<h})]\|_{L^2 L^1} \\
&\lesssim 2^{\frac{nk}{2}} \|Q_{>j-C}\phi_h\|_{L^2} \|\tilde{P}_h F'(\phi_{<h})]\|_{L^\infty L^2} \\
&\lesssim 2^{-\frac{3j}{2}-(\frac{n}{2}-1)k-(n-1)(h-k)} \|\phi_h\|_S (1+\|\phi\|_S^N)
\end{aligned}
$$

For the second term we use (9.31) for the nonlinear factor,

$$
\begin{aligned}
\|P_k Q_j[Q_{<j-C}\phi_h \tilde{Q}_j \tilde{P}_h F'(\phi_{<h})]\|_{L^2} &\lesssim 2^k \|Q_{<j-C}\phi_h \tilde{Q}_j \tilde{P}_h F'(\phi_{<h})\|_{L^2 L^{\frac{2n}{n+2}}} \\
&\lesssim 2^k \|\phi_h\|_{L^\infty L^2} \|\tilde{Q}_j \tilde{P}_h F'(\phi_{<h})]\|_{L^2 L^n} \\
&\lesssim 2^{-\frac{3j}{2}-(\frac{n}{2}-1)k-\frac{n}{2}(h-k)} \|\phi_h\|_S (1+\|\phi\|_S^N)
\end{aligned}
$$

(a2) Low modulation output, $j \leq h + 2C$.

Using again (9.11) we rewrite the left hand side function in (9.34) as

$$(9.35) \qquad P_k[\phi_h \tilde{P}_h F'(\phi_{<h-2C})] + P_k \int_{h-2C}^h \phi_h \phi_{h_1} F''(\phi_{<h_1})dh_1$$

Here we have added the harmless multiplier $\tilde{P}_h$. This is allowed because of the balance of frequencies in the formula.

(a2)(i) The first term in (9.35) can be estimated in L^2 using (9.22):

$$
\begin{aligned}
\|P_k[\phi_h \tilde{P}_h F'(\phi_{<h-2C})]\|_{L^2} &\lesssim \|\phi_h\|_{L^\infty} \|\tilde{P}_h F'(\phi_{<h-2C})\|_{L^2} \\
&\lesssim 2^{\frac{(n+1)h}{2}} \|\phi_h\|_S (1+\|\phi\|_S)^N
\end{aligned}
$$

(a2)(ii) The second term in (9.35) localized at modulation 2^j has the form

$$v = P_k Q_j \int_{h-2C}^h \phi_h \phi_{h_1} F''(\phi_{<h_1})dh_1$$

For this we use the expansion (9.12) down to the frequency $l = j - 3C$,

$$v = \sum_{i=2}^{N-1} v_i + z$$

where

$$v_i = P_k Q_j \int_{D_i} \phi_h \phi_{h_1} \cdots \phi_{h_i} F^{(i)}(\phi_{<j-3C})dh_1 \cdots dh_i, \quad i = 1, N-1$$

$$z = P_k Q_j \int_{D_N} \phi_h \phi_{h_1} \cdots \phi_{h_N} F^{(N)}(\phi_{<h_N})dh_1 \cdots dh_N$$

$$D_i = \{j - 3C < h_i < \cdots < h_1, \ h - 2C < h_1 < h\}$$

We split v_i into

$$v_i = v_i^{(1)} + v_i^{(2)} + v_i^{(3)}$$

with

$$v_i^{(1)} = P_k Q_j \int_{D_i} \phi_h \phi_{h_1} \cdots \phi_{h_i} P_{>j-C} F^{(i)}(\phi_{<j-3C}) dh_1 \cdots dh_i$$

$$v_i^{(2)} = P_k Q_j \int_{D_i} \tilde{P}_k Q_{>j-C}(\phi_h \phi_{h_1} \cdots \phi_{h_i}) P_{<j-C} F^{(i)}(\phi_{<j-3C}) dh_1 \cdots dh_i$$

$$v_i^{(3)} = P_k Q_j \int_{D_i} \tilde{P}_k Q_{<j-C}(\phi_h \phi_{h_1} \cdots \phi_{h_i}) P_{<j-C} \tilde{Q}_j F^{(i)}(\phi_{<j-3C}) dh_1 \cdots dh_i$$

For $v_i^{(1)}$ we bound the nonlinear term in S using (9.23) and then we apply the multilinear estimate (9.19).

To bound $v_i^{(2)}$ we use the multilinear estimate (9.19) for the product $\phi_{h_1} \cdots \phi_{h_i}$ and the trivial L^∞ bound for the nonlinear term.

For $v_i^{(3)}$ we get an improved energy bound for $P_k Q_{<j-C}(\phi_h \phi_{h_1} \cdots \phi_{h_i})$ using the multilinear estimate (9.19). If $j \leq k$ then this is combined with the $L^2 L^\infty$ bound for $P_{>j-C} F^{(i)}(\phi_{<j-3C})$ given by (9.31). If $k \leq j$ then it is more efficient to apply the $L^2 L^n$ bound in (9.31) for the nonlinear term and then to use a Sobolev inequality at frequency 2^k.

Next we decompose z,

$$z = z^{(1)} + z^{(2)} = z^{(1)} + z^{(3)} + z^{(4)}$$

where

$$z^{(1)} = P_k Q_j \int_{D_N} \phi_h \phi_{h_1} \cdots \phi_{h_N} P_{>h_N+C} F^{(N)}(\phi_{<h_N}) dh_1 \cdots dh_N$$

$$z^{(2)} = P_k Q_j \int_{D_N} \phi_h \phi_{h_1} \cdots \phi_{h_N} P_{<h_N+C} F^{(N)}(\phi_{<h_N}) dh_1 \cdots dh_N$$

$$z^{(3)} = P_k Q_j \int_{D_N} Q_{>j-C}(\phi_h \phi_{h_1} \cdots \phi_{h_N}) P_{<h_N+C} F^{(N)}(\phi_{<h_N}) dh_1 \cdots dh_N$$

$$z^{(4)} = P_k Q_j \int_{D_N} Q_{<j-C}(\phi_h \phi_{h_1} \cdots \phi_{h_N}) \tilde{Q}_j P_{<h_N+C} F^{(N)}(\phi_{<h_N}) dh_1 \cdots dh_N$$

For $z^{(1)}$ we bound the nonlinear term in S using (9.23) and then we apply the multilinear estimate (9.19).

For $z^{(2)}$ we use the multilinear estimate (9.19) to estimate the product in L^2 and trivially bound the nonlinear term in L^∞. This gives the correct bound only if $j \leq h_N + C$.

For $j > h_N + C$ we use the second decomposition. For $z^{(3)}$ we combine the L^2 multilinear bound (9.19) for the product with the trivial L^∞ bound for the nonlinear term.

Finally, for $z^{(4)}$ we use (9.19) to derive an improved $L^\infty L^2$ bound for the multilinear expression. This we combine with either the $L^2 L^\infty$ bound in (9.31) for the nonlinear term (if $j < k$) or the $L^2 L^n$ bound in (9.31) for the nonlinear term (if $j \geq k$).

(b) The intermediate frequency bounds $(k-2C < h < k+2C)$. Since the range of h is bounded here, we can disregard the h integration.

(b1) High modulation output, $j > h + 3C$. Then we write

$$\begin{aligned}
P_k Q_j(\phi_h(F'(\phi_{<h})) &= P_k Q_j(\phi_h P_{>h+C} F'(\phi_{<h})) \\
&+ P_k Q_j(Q_{>j-C}\phi_h P_{<h+C} F'(\phi_{<h})) \\
&+ P_k Q_j(Q_{<j-C}\phi_h \tilde{Q}_j P_{<h+C} F'(\phi_{<h}))
\end{aligned}$$

The bound for the first term follows from (9.22) and the bilinear estimate (9.5). For the second term we use the L^2 bound for $Q_{>j-C}\phi_h$ and the trivial L^∞ bound for the nonlinearity. Finally, for the last term we use the $L^\infty L^2$ bound for the linear factor and the $L^2 L^\infty$ estimate (9.31) for the nonlinear one.

(b2) Low modulation output, $j < h + 3C$. In this case we need to use the full multilinear expansion in (9.12). Our contention is that the genuinely multilinear terms in the expansion are better behaved, therefore we can peel them off. For $l < h + 3C$ we write

$$\begin{aligned}
P_k Q_{<l}[\phi_h F'(\phi_{<h})] &= P_k Q_{<l}(\phi_h(F'(\phi_{<h}) - F'(\phi_{<l-3C}))) \\
&+ P_k Q_{<l}(\phi_h P_{>l-2C} F'(\phi_{<l-3C})) \\
&+ Q_{<l}[P_k, P_{<l-2C} F'(\phi_{<l-3C})]\phi_h \\
\text{(9.36)} \quad &+ Q_{<l}(P_k \phi_h Q_{>l-2C} P_{<l-2C} F'(\phi_{<l-3C})) \\
&+ Q_{<l}(Q_{>l-2C} P_k \phi_h Q_{<l-2C} P_{<l-2C} F'(\phi_{<l-3C})) \\
&+ Q_{<l-2C} P_k \phi_h Q_{<l-2C} P_{<l-2C} F'(\phi_{<l-3C})
\end{aligned}$$

For the first term we need to use the full multilinear expansion in (9.12). The idea is is that it is genuinely multilinear (i.e. contains at least two factors besides the nonlinearity) therefore it is better behaved.

The second, fourth and fifth term contain only interactions in which at least one term has high modulation, which are milder in our set-up. Finally, the third term would be as bad as the main one (i.e. the last one) if it were not for the extra gain coming from the commutator. Next we discuss each of these terms separately.

(b2)(i) The genuinely multilinear part. For $l < h + 3C$ we estimate the first term in (9.36). Precisely, for $j \leq l$ we prove that
(9.37)
$$\|P_k Q_j(\phi_h(F'(\phi_{<h}) - F'(\phi_{<l-3C})))\|_{\dot{X}^{\frac{n}{2},\frac{1}{2},1}} \lesssim 2^{-\delta_0(l-j)}\|\phi_h\|_S(1 + \|\phi\|_S)^N$$

We use the expansion (9.12) to write the left hand side function in the form

$$P_k Q_j(\phi_h(F'(\phi_{<h}) - F'(\phi_{<j-3C}))) = \sum_{i=1}^{N-1} v_i + z$$

where

$$v_i = P_k Q_j \phi_h \int_{D_i} \phi_{h_1} \cdots \phi_{h_i} F^{(i)}(\phi_{<j-3C}) dh_1 \cdots dh_i$$

$$z = P_k Q_j \phi_h \int_{D_N} \phi_{h_1} \cdots \phi_{h_N} F^{(N)}(\phi_{<h_N}) dh_1 \cdots dh_N$$

and

$$D_i = \{j - 3C \leq h_N < \cdots < h_1 < h\}$$

Then we proceed as in case (a2)(ii).

(b2)(ii) High modulation input. Here we consider the second, fourth and fifth terms in (9.36). We begin with

$$P_k Q_{<l}(\phi_h P_{>l-2C} F'(\phi_{<l-3C})) = \int_{l-2C}^{\infty} P_k Q_{<l}(\phi_h P_{h_1} F'(\phi_{<l-3C})) dh_1$$

By (9.23) we bound the nonlinearity in S and then we conclude using the bilinear estimate (9.8).

Secondly we estimate

$$Q_{<l}(P_k \phi_h Q_{>l-2C} P_{<l-2C} F'(\phi_{<l-3C}))$$

in L^2 using the energy bound for ϕ_h and the $L^2 L^\infty$ bound in (9.31) for the nonlinear factor.

Finally, the fifth term in (9.36),

$$Q_{<l}(Q_{>l-2C} P_k \phi_h Q_{<l-2C} P_{<l-2C} F'(\phi_{<l-3C}))$$

is also estimated in L^2 using the L^2 bound for the first factor and the trivial L^∞ bound for the nonlinearity.

(b2)(iii) The commutator term. For $j \leq l$ we estimate the modulation 2^j part of the commutator term. For the commutator we use the representation in [18], Lemma 2:

$$Q_j[P_k, P_{<l-2C} F'(\phi_{<l-3C})]\phi_h = 2^{-h} Q_j L(\partial_x P_{<l-2C} F'(\phi_{<l-3C}), \phi_h)$$
$$= \int_{j-2C}^{l-2C} 2^{h_1-h} Q_j L(P_{h_1} F'(\phi_{<l-3C}), \phi_h) dh_1$$
$$+ 2^{j-h} Q_j L(P_{<j-2C} F'(\phi_{<l-3C}), \phi_h)$$

We first estimate the integrand. We write

$$\begin{aligned}
Q_j L(P_{h_1} F'(\phi_{<l-3C}), \phi_h) &= Q_j L(P_{h_1}(F'(\phi_{<l-3C}) - F'(\phi_{<h_1-3C})), \phi_h) \\
&+ Q_j L(P_{h_1} F'(\phi_{<h_1-3C}), \phi_h)
\end{aligned}$$

For the first term we discard L and P_{h_1} and use (9.37). This gives a $2^{-\delta_0(h_1-j)}$ gain, which combines with the 2^{h_1-h} gain from the commutator. For the second term we use (9.23) to bound the nonlinearity in S and then the bilinear estimate (9.8).

For the last part of the commutator we have

$$\begin{aligned}
2^{j-h} Q_j L(P_{<j-2C} F'(\phi_{<l-3C}), \phi_h) &= 2^{j-h} Q_j L(P_{<j-2C}(F'(\phi_{<l-3C}) \\
&- F'(\phi_{<j-3C})), \phi_h) + 2^{j-h} Q_j L(P_{<j-2C} F'(\phi_{<j-3C})), \phi_h)
\end{aligned}$$

In the first term we discard L and $P_{<j-2C}$ and use (9.37). We decompose the second part further as

$$\begin{aligned}
&2^{j-h} Q_j L(P_{<j-2C} F'(\phi_{<j-3C})), Q_{>j-2C}\phi_h) \\
+ \; &2^{j-h} Q_j L(Q_{>j-2C} P_{<j-2C} F'(\phi_{<j-3C})), Q_{<j-2C}\phi_h)
\end{aligned}$$

For the first term we use the L^2 bound for $Q_{>j-2C}\phi_h$ and the L^∞ bound for the nonlinearity. For the second we use the energy bound for $Q_{<j-2C}\phi_h$, respectively the $L^2 L^\infty$ bound in (9.31) for the nonlinear factor. $\square$

9.11. Conclusion. We consider now all the components of

$$\|P_k F(\phi)\|_{S_c}$$

(i) The energy norm

$$2^{(\frac{n}{2}-1)k} \|\nabla P_k F(\phi)\|_{L^\infty L^2}$$

was estimated in Lemma 9.3.

(ii) The norm

$$\|P_k \nabla F(\phi)\|_{X^{\frac{n}{2}-1, \frac{1}{2}, \infty}}$$

is controlled due to Lemma 9.9.

(iii) It remains to bound the quantities

$$\|P_k Q_{<j} F(\phi)\|_{S[k,j]}, \qquad j < k$$

Since

$$\|\|P_k Q_{<j}\psi\|_{S[k,j]} \lesssim \|\psi\|_{X^{\frac{n}{2}, \frac{1}{2}, 1}}$$

Lemma 9.9 allows us to replace the above expression by

$$\|Q_{<j-2C}\phi_k P_{<j-2C} Q_{<j-2C} F'(\phi_{<j-3C})\|_{S[k,j]}$$

But this is estimated using Lemma 9.1.

References

[1] 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.

[2] H. Brezis and L. Nirenberg. Degree theory and BMO. I. Compact manifolds without boundaries. *Selecta Math. (N.S.)*, 1(2):197–263, 1995.

[3] Piero D'Ancona and Georgiev. Ill posedness results for the two dimensional wave maps equation. preprint.

[4] Alexandre Freire, Stefan Müller, and Michael Struwe. Weak compactness of wave maps and harmonic maps. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 15(6):725–754, 1998.

[5] M. L. Gromov. Isometric imbeddings and immersions. *Dokl. Akad. Nauk SSSR*, 192:1206–1209, 1970.

[6] Matthias Günther. Isometric embeddings of Riemannian manifolds. In *Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990)*, pages 1137–1143, Tokyo, 1991. Math. Soc. Japan.

[7] Frédéric Hélein. *Harmonic maps, conservation laws and moving frames*, volume 150 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, second edition, 2002.

[8] 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.

[9] S. Klainerman and M. Machedon. Space-time estimates for null forms and the local existence theorem. *Comm. Pure Appl. Math.*, 46(9):1221–1268, 1993.

[10] Sergiu Klainerman and Igor Rodnianski. On the global regularity of wave maps in the critical Sobolev norm. *Internat. Math. Res. Notices*, (13):655–677, 2001.

[11] Sergiu Klainerman and Sigmund Selberg. Remark on the optimal regularity for equations of wave maps type. *Comm. Partial Differential Equations*, 22(5-6):901–918, 1997.

[12] Stefan Müller and Michael Struwe. Global existence of wave maps in $1 + 2$ dimensions with finite energy data. *Topol. Methods Nonlinear Anal.*, 7(2):245–259, 1996.

[13] Andrea Nahmod, Atanas Stefanov, and Karen Uhlenbeck. On the wellposedness of the wave map problem in high dimensions. preprint.

[14] John Nash. The imbedding problem for Riemannian manifolds. *Ann. of Math. (2)*, 63:20–63, 1956.

[15] Jalal Shatah. Weak solutions and development of singularities of the SU(2) σ-model. *Comm. Pure Appl. Math.*, 41(4):459–469, 1988.

[16] Jalal Shatah and Michael Struwe. The Cauchy problem for wave maps. *Int. Math. Res. Not.*, (11):555–571, 2002.

[17] Terence Tao. Global regularity of wave maps. I. Small critical Sobolev norm in high dimension. *Internat. Math. Res. Notices*, (6):299–328, 2001.

[18] Terence Tao. Global regularity of wave maps. II. Small energy in two dimensions. *Comm. Math. Phys.*, 224(2):443–544, 2001.

[19] Daniel Tataru. Local and global results for wave maps. I. *Comm. Partial Differential Equations*, 23(9-10):1781–1793, 1998.

[20] Daniel Tataru. On global existence and scattering for the wave maps equation. *Amer. J. Math.*, 123(1):37–77, 2001.

[21] Michael E. Taylor. *Pseudodifferential operators and nonlinear PDE*, volume 100 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1991.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY

E-mail address: `tataru@math.berkeley.edu`