T. Duyckaerts *et al.* (2017) "Universality of Blow up Profile for Small Blow up Solutions to the Energy Critical Wave Map Equation,"
International Mathematics Research Notices, Vol. 2018, No. 22, pp. 6961–7025
Advance Access Publication May 16, 2017
doi: 10.1093/imrn/rnx073

Universality of Blow up Profile for Small Blow up Solutions to the Energy Critical Wave Map Equation

Thomas Duyckaerts[1]*, Hao Jia[2], Carlos Kenig[3], and Frank Merle[4]

[1]LAGA, Institut Galilée, Université Paris 13, Université Sorbonne Paris Cité, 99, avenue Jean-Baptiste Clément, Villetaneuse, 93430, FR and Institut Universitaire de France, [2]School of Mathematics, Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA, [3]Department of Mathematics, University of Chicago, 5734 University Avenue, Chicago, IL 60637-1514, USA, and [4]Laboratoire AGM, UFR sciences et techniques, 2 av. Adolphe Chauvin, 95302 Cergy-Pontoise Cedex, France

*Correspondence to be sent to: e-mail: *duyckaer@math.univ-paris13.fr*

We show that an energy-critical wave map into the sphere that has energy just above the degree one harmonic maps and that blows up in finite time asymptotically decouples into a regular part plus a traveling wave with small momentum, in the energy space. In particular, the only possible form of energy concentration is through the concentration of traveling waves. This is often called "quantization of energy" at blow up. The main new tool is a channel of energy type inequality for outgoing small energy wave maps similar to the one proved in our previous work [14] on outgoing solutions of the energy-critical wave equation. We also give a brief review of important background results in the subcritical and critical regularity theory for the two dimensional wave maps from [35–37, 47, 48, 51, 56, 57].

Received March 8, 2017; Revised March 8, 2017; Accepted March 9, 2017

1 Introduction

We consider the Cauchy problem for wave map u from R^{2+1} with Minkowski metric to the standard 2-sphere $S^2 \subset R^3$:

$$\partial_{tt}u - \Delta u = (|\nabla u|^2 - |\partial_t u|^2)u, \text{ in } R^2 \times R \tag{1.1}$$

with initial data $\overrightarrow{u}(0) := (u, \partial_t u)(0) = (u_0, u_1)$. We shall only consider initial data (u_0, u_1) that satisfies the "compatibility condition" that $|u_0| \equiv 1$ and $u_0 \cdot u_1 \equiv 0$. For simplicity, we also assume that the initial data $\overrightarrow{u}(0)$ is smooth, u_1 is compactly supported, and that u equals a fixed constant u_∞ for large x. We call such wave maps classical, following the usual convention. Wave maps from the Minkowski space to a general Riemannian manifold $\mathcal{M}$ arise naturally as the hyperbolic counterpart of harmonic maps, and are given as critical points of the Lagrangian

$$\mathcal{L}(u) := \int_{R^3} |\nabla u|^2 - |\partial_t u|^2 \mathrm{d}x\mathrm{d}t$$

for $u : R^3 \to \mathcal{M}$. It is sometimes more convenient to adopt the more geometric notation: set for $\alpha = 0, 1, 2$ that $\partial_\alpha = \partial_t$ if $\alpha = 0$, $\partial_\alpha = \partial_{x_\alpha}$ if $\alpha = 1, 2$, and that $\partial^\alpha = -\partial_\alpha$ if $\alpha = 0$, $\partial^\alpha = \partial_\alpha$ if $\alpha = 1, 2$. This is of course just using the Minkowski metric to lower or upper the index. We adopt the Einstein summation convention with repeated indices and view u as a column vector. We also use the standard notation that $x^0 = x_0 = t$, $x^j = x_j$ for $j = 1, 2$. Then equation (1.1) can be written as

$$-\partial_\alpha \partial^\alpha u = u\, \partial^\alpha u^\dagger \partial_\alpha u, \tag{1.2}$$

where $u^\dagger$ is the transpose of u.

The wave map equation has been intensively studied, as a natural geometric wave equation and as models from physics—including general relativity and gauge theories. The study of the Cauchy problem and the dynamics of solutions was initiated in the works of Shatah and Tahvildar-Zadeh [45, 46], Christodoulou and Tahvildar-Zadeh [4, 5], and Struwe [49], in the equivariant setting. In these works, many deep and interesting regularity and dynamical properties of equivariant wave maps were revealed. In general, the wave map can develop a singularity in finite time by concentrating energy in a small region. Indeed, singular solutions in the form of a shrinking soliton plus a residue term have been constructed for the 2 + 1 dimensional equivariant wave map equation by Krieger *et al.* [**?**] with prescribed rate, by Rodnianski and Sterbenz [41] in a stable regime for high equivariance wave maps, and by Raphaël and Rodnianski [40] for co-rotational wave maps. We also refer to the recent survey [43] for further discussion. The Cauchy problem for the wave maps without equivariant symmetry is more complicated. Recall that equation (1.1) is invariant under the natural scaling

$$u \to u_\lambda(x,t) = u(\lambda x, \lambda t), \quad (u_0, u_1) \to (u_{0\lambda}(x), u_{1\lambda}(x)) = (u_0(\lambda x), \lambda u_1(\lambda x)) \tag{1.3}$$

and the conserved energy

$$\mathcal{E}(\overrightarrow{u}) := \int_{R^2} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} \right) (x,t)\,\mathrm{d}x \tag{1.4}$$

is invariant under the scaling (1.3). Scale invariance of the equation plays an essential role in both the Cauchy problem and the dynamics of solutions. We note that for equation (1.1), the natural initial data space invariant under the scaling (1.3) is the energy space $\dot{H}^1 \times L^2$, and hence the equation is called *energy critical*. The works of Klainerman and Machedon [32–34], and subsequently Klainerman and Selberg [35, 36], and Selberg [44] established wellposedness in the subcritical space $\dot{H}^{s-1} \times H^{s-1}$ with $s > 1$, and introduced important ideas on the bilinear and null form estimates that also played an important role in the critical theory. The Cauchy problem for the wave map equation in the critical space $\dot{H}^1 \times L^2$ is more difficult, and was addressed in the breakthrough work of Tao [51], using the important null frame spaces introduced by Tataru [57] (In this work, Tataru established small data global wellposedness in a critical space that is slightly stronger than the energy space.) and Tao's idea of gauge transform [50]. The global wellposedness for the energy critical wave maps has been solved, independently in the works of Krieger and Schlag [37], Sterbenz and Tataru [47, 48], and Tao [52–55]. We will mainly rely on the Sterbenz–Tataru approach to the large data theory. [47, 48] proves that if a wave map blows up in finite time or is global and does not scatter, then after suitable transformation using symmetry, it must converge locally, along a sequence of times, to a harmonic map. This is a first step in the proof of the so-called "soliton resolution conjecture", predicting that the solution asymptotically decouples into a finite sum of harmonic maps (modulated by the transformations of the equation) plus a regular part in the finite time blow up case or a linear wave in the global existence case.

Many recent progresses were made on this conjecture in the equivariant setting. It was proved by Côte [7] for co-rotational wave maps and by the second and third authors [26] for all equivariant wave maps that the decomposition holds along a sequence of times. If one imposes certain energy constraint that effectively rules out multi-soliton configuration, then the restriction to a sequence of times can be removed, and the full soliton resolution holds, as proved in [9, 10] (see also [29, 31] for the wave maps equation outside a ball). Note that even in this equivariant setting, the question of proving the full resolution remains open for general solutions. To answer this question, it seems that one needs to understand the interaction of solitons that are separated by scales, which appears to be a challenging task. We refer to [24] for the existence of solutions decoupling into more than one bubble.

The new ingredient in [9, 10, 29, 31] is the *channel of energy method* first introduced in [15, 16] in the context of the energy-critical semilinear wave equation, which provides strong decoupling mechanism between the dispersion and solitary waves. This method consists in proving that the energy of any solution that is dispersive (in a weak sense that has to be made precise) can be bounded from below outside the wave cone in at least one time direction. For the energy critical wave equation, this *channel of energy inequality* was proved by the first, third and fourth authors for small solutions, in odd space dimensions [16] and for any radial, nonstationary solution in three space dimensions [17]. See [8, 18, 27, 28, 42] for other applications to semilinear energy critical wave equations. Note that this channel of energy inequality is very sensitive to the dimensions (see [12, 30]) and (at least for large solutions) depend crucially on the radial assumption.

Going back to equation (1.1), one can ask what happens if we remove the equivariance assumption. Recently, Grinis [21] proved that along a well-chosen sequence of times, all the energy concentration strictly inside the lightcone (That is, in the region $\{|x| < a|T_+ - t|\}$ for any $a < 1$, where T_+ is the blow up time, for a solution blowing-up at $x = 0$.) must be in the form of traveling waves, by showing that there is no energy in the so called "neck region". (The region strictly inside the lightcone and away from the solitons.) It is natural to ask if one can prove the soliton resolution conjecture along a sequence of times, as in the equivariant case. However, a new difficulty appears that is not present in the equivariant setting, where a classical result of Christodoulou and Tahvildar-Zadeh (see [4, 5]) asserts that there is asymptotically no energy accumulation in the so called "self similar" region. In particular, in the equivariant case there cannot be any energy concentration near the boundary of the singularity lightcone $|x| < T_+ - t$ as $t \to T_+$, assuming that the solution blows up at time T_+. As far as the authors know, it is an open question how to rule out energy concentration in this region in the general case. Another question that is left open in the work of Grinis is to obtain a complete characterization for general times $t \to T_+$, and not only for a sequence of times.

Our work addresses both questions, in the restricted case where the energy is only slightly higher than the energy of the degree one co-rotational harmonic maps. We believe that the methods used here apply to wave maps into more general targets without any size restrictions. However it appears that one has to overcome some nontrivial obstacles in the current perturbative setup to achieve this goal. We also believe that the case $T_+ = \infty$, when the solution does not scatter, can also be addressed by the methods developed here. We plan to address these questions in future work.

Let us briefly summarize our main results.

Our first main goal is to introduce the channel of energy argument to the study of wave map equations. We use a new point of view, developed in our recent work [14] on the energy critical wave equation, namely that the channel of energy inequality holds for well-prepared initial data that satisfy an outgoing condition. Unlike previously known channel of energy inequalities, the version for outgoing waves turns out to be rather robust in that it works for nonradial solutions in all dimensions, and we think it is applicable to a wide range of problems. The outgoing condition is natural. For instance, any linear wave at large time will satisfy such outgoing conditions. More interestingly, the dispersed energy that might concentrate near the boundary of the singularity lightcone of a blow up solution also satisfies the outgoing condition (for both energy-critical wave maps and semilinear wave equation).

In [14], the proof of the channel of energy inequality for the energy-focusing semilinear wave equation relies on the corresponding inequality for the linear equation and a straightforward perturbation argument. However, the Cauchy problem for the wave map equations is much more complicated and the current perturbation results are not as precise as in the case of the energy critical wave equation. At this time we can only extend the results from [14] partially, and prove the channel of energy inequality for small data.

Theorem 1.1. Fix $\beta \in (0,1)$. There exists a small $\delta = \delta(\beta) > 0$ and sufficiently small $\epsilon_0 = \epsilon_0(\beta) > 0$, such that if u is a classical wave map with energy $\mathcal{E}(\overrightarrow{u}) < \epsilon_0$ satisfying

$$\|(u_0,\, u_1)\|_{\dot{H}^1\times L^2\left(B^c_{1+\delta}\cup B_{1-\delta}\right)} + \|\not\partial u_0\|_{L^2} + \|\partial_r u_0 + u_1\|_{L^2} \le \delta \|(u_0,\, u_1)\|_{\dot{H}^1\times L^2}, \tag{1.5}$$

then for all $t \ge 0$, we have

$$\int_{|x|>\beta+t} |\nabla_{x,t} u|^2(x,t)\,\mathrm{d}x \ge \beta\, \|(u_0,\, u_1)\|^2_{\dot{H}^1\times L^2}. \tag{1.6}$$

□

As an application for the channel of energy inequality (1.6), we obtain the following classification of finite time blow up wave maps u with energy

$$\mathcal{E}(\overrightarrow{u}) < \mathcal{E}(Q,0) + \epsilon_0^2, \tag{1.7}$$

where Q is the harmonic map with the least energy (which is equal to 4π), and $\epsilon_0 > 0$ is small. Denote $\mathcal{M}_1$ as the space of degree one harmonic maps, (These harmonic maps are all co-rotational with respect to certain axis of symmetry.) and let

$$\mathcal{M}_{\ell,1} := \{Q_\ell : Q \in \mathcal{M}\},$$

where

$$Q_\ell(x,t) = Q\left(x - \frac{\ell \cdot x}{|\ell|^2}\ell + \frac{\frac{\ell\cdot x}{|\ell|^2}\ell - \ell t}{\sqrt{1-|\ell|^2}}\right) \tag{1.8}$$

is the Lorentz transformation of the harmonic map Q. Then we have

Theorem 1.2. Let u be a classical wave map with energy $\mathcal{E}(\overrightarrow{u}) < \mathcal{E}(Q,0)+\epsilon_0^2$, that blows up at a finite time T_+ and at the origin. Assume that ϵ_0 is sufficiently small. Then there exists $\ell \in R^2$ with $|\ell| \ll 1$, $x(t) \in R^2$, $\lambda(t) > 0$ with

$$\lim_{t\to T_+} \frac{x(t)}{T_+ - t} = \ell, \;\; \lambda(t) = o\left(T_+ - t\right)$$

and $(v_0, v_1) \in \dot{H}^1 \times L^2 \cap C^\infty(R^2\backslash\{0\})$ with $(v_0 - u_\infty, v_1)$ being compactly supported, such that

$$\text{(i)}\inf\left\{\left\|\overrightarrow{u}(t) - (v_0, v_1) - (Q_\ell, \partial_t Q_\ell)\right\|_{\dot{H}^1\times L^2} : Q_\ell \in \mathcal{M}_{\ell,1}\right\} \to 0, \text{ as } t \to T_+;$$

$$\text{(ii)}\left\|\overrightarrow{u}(t) - (v_0, v_1)\right\|_{\dot{H}^1\times L^2\left(R^2\backslash B_{\lambda(t)}(x(t))\right)} \to 0 \text{ as } t \to T_+,$$

where $B_{\lambda(t)}(x(t)) = \left\{x \in R^2 \,:\, |x - x(t)| < \lambda(t)\right\}$. □

Heuristically speaking, the above theorem says that at blow up time, the wave map essentially consists of two parts, one regular part outside the lightcone $|x| > T_+ - t$, and a traveling wave with small velocity ℓ that concentrates in a small region (in comparison with the size of the cone) near the point $\ell(T_+ - t)$. In addition, there are no other types of energy concentration. It is an interesting question to ask about the *finer* dynamics of the traveling wave in the region $|x - x(t)| < \lambda(t)$, such as how the axis of rotation of the wave map evolves. It is not clear to us at this moment if the axis of rotation could fail to stabilize. We believe though that it is impossible, at the level of energy regularity, to say more about the finer dynamics of the scale $\lambda(t)$ and the center $x(t)$, due to the symmetries of the equation.

Let us very briefly explain the strategy of the proof. The proof of the channel of energy inequality in Theorem 1.1 uses the extension of the linear channel of energy inequality for outgoing waves from [14] to two dimensions. For the wave maps however, we need to also show that these outgoing conditions are in some sense stable (for *most frequencies*) with respect to frequency projections, in order to use the perturbative results for wave maps which deal with each frequency piece of the map separately. In addition, it is well known from Tao's work [51] that the nonlinearity cannot be treated as small perturbations directly even in the small energy case, and one has to perform a gauge transform to treat the nonlinearity. The key point here is that the gauge transform, which although changes the wave map quite significantly, does not significantly change the energy distribution.

To prove Theorem 1.2, let us take a wave map as in Theorem 1.2. Then by the result of Tataru and Sterbenz [48], along a sequence of times, we can extract a traveling wave from the wave map. A little more effort also shows that there is no other possible energy concentration strictly inside the lightcone except in the neck region, thanks to the energy constraint (1.7). By Grinis' result [21] there is no energy in the neck region either. Thus all residue energy has to concentrate near the boundary of the singularity lightcone $|x| < T_+ - t$. In addition, such residue energy has to be small, again thanks to the energy constraint. We apply the channel of energy inequality to rule out this residue energy. This is a crucial step and the main new point of our article. Hence inside the lightcone (not only *strictly inside* the lightcone) the amount of energy is asymptotically just the energy of the traveling wave. Then by the coercivity of energy near the traveling wave, we conclude that in fact the wave map is trapped in smaller and smaller neighborhoods of the traveling wave, and thus has to stay close to the traveling wave for all times $t < T_+$, not just along a sequence of times. This completes the proof of the main Theorem 1.2.

Our article is organized as follows:

- In Section 2, we recall the necessary subcritical and critical regularity results for the wave equation;
- In Section 3, we prove the channel of energy inequality for small wave maps;
- In Section 4, we recall the Morawetz estimates;

- In Section 5, we prove the decomposition into regular part and traveling wave along a sequence of times; and
- In Section 6, we prove certain coercive property of energy in the neighborhood of the traveling wave and establish the decomposition for all times.

Throughout the article, we shall use the notation

$$\|f\|_{\dot{H}^1(E)} := \|\nabla f\|_{L^2(E)}$$

for any measurable set E. If $s > 1$, we will write

$$\|f\|_{\dot{H}^s} := \||D|^s f\|_{L^2(\mathbb{R}^2)},$$

where $|D|^s$ is the Fourier multiplier with symbol $|\xi|^s$, and say that a distribution f is in $\dot{H}^s$ when $f \in H^s_{\mathrm{loc}}(\mathbb{R}^2)$ and the above seminorm is finite.

2 Preliminaries

In this section, we briefly review the subcritical and critical regularity results for the two-dimensional wave maps into the sphere, that will be needed below.

2.1 Local wellposedness in H^s for $s > 1$

It is well known from the works of Klainerman and Machedon [32–34], Klainerman and Selberg [35, 36], and Selberg [44], that the wave map equation (1.2) is locally wellposed in $H^s \times H^{s-1}$ for $s > 1$. In this subsection we recall the necessary regularity results from these works without giving proofs and refer the reader to the above cited works, and especially the survey [36] for details.

Since the spaces in which one can prove existence and uniqueness involves spacetime Fourier transforms even when one only considers local in time solutions, we have to be more precise on the Banach spaces which are used to hold the solutions and the nonlinearities.

We shall denote $\mathcal{F}(u)$ as the spacetime Fourier transform of u. For s, $b \in R$, and tempered distribution $u \in \mathcal{S}'(R^3)$, define

$$\|u\|_{X^{s,b}(R^3)} := \left(\int_{R^3} (1+|\xi|^2)^s \,(1+||\xi|-|\tau||)^{2b}\, |\mathcal{F}(u)(\xi,\tau)|^2 \mathrm{d}\xi\, \mathrm{d}\tau \right)^{\frac{1}{2}} \tag{2.1}$$

and set

$$X^{s,b}(R^3) := \{u \in \mathcal{S}'(R^3) : \|u\|_{X^{s,b}(R^3)} < \infty\}. \tag{2.2}$$

We record the following wellposedness result for equation (1.2) in the subcritical space $\dot{H}^s \times H^{s-1}$. We shall always assume that the initial data (u_0, u_1) for (1.2) satisfies the "admissibility condition" that $|u_0| \equiv 1$ and $u_0^\dagger u_1 \equiv 0$.

Theorem 2.1. For $s > 1$ and $\frac{1}{2} < b < \min\{s - \frac{1}{2}, 1\}$. Suppose that $(u_0, u_1) \in \dot{H}^s \times H^{s-1}$ and that u_0 equals a constant $u_\infty \in S^2$ for large x. Then for $T = T(\|(u_0 - u_\infty, u_1)\|_{H^s \times H^{s-1}}) > 0$ sufficiently small, there exists a unique solution u to equation (1.2) with initial data (u_0, u_1) on $R^2 \times (-T, T)$ in the sense of distributions, which satisfies the following properties

(1) $u - u_\infty \in C(I, H^s \times H^{s-1})$;

(2) there exists $\overline{u} \in L^2(R^3)$ with $\overline{u}|_{R^2 \times I} \equiv u - u_\infty$ and $\nabla_{x,t}\overline{u} \in X^{s-1,b}$,

where $I = (-T, T)$. □

Remark. The above theorem provides a rigorous definition of solutions to equation (1.2). (2) is important, as (1) by itself is not sufficient to guarantee uniqueness when s is close to 1. One could of course choose to work directly with smooth wave maps, instead of these low-regularity wave maps. However, below we shall need to extend a locally (in space) defined map to a global one, and it is much more convenient to have such extensions in the framework of H^s solutions, rather than smooth solutions. □

Solutions from Theorem 2.1 can be extended to a maximal interval of existence, more precisely, we have

Corollary 2.1. For $s > 1$. Suppose that $(u_0, u_1) \in \dot{H}^s \times H^{s-1}$ and that u_0 equals a constant $u_\infty \in S^2$ for large x. Then there exists $T_+ \in (0, \infty]$, $T_- \in [-\infty, 0)$, such that for any $T_- < T_1 < T_2 < T_+$, u is a distributional solution to equation (1.2), satisfying (1) and (2) on $I = (T_1, T_2)$, and that if $T_+ < \infty$, then

$$\lim_{t \to T_+} \|\overrightarrow{u}(t)\|_{\dot{H}^s \times H^{s-1}} = \infty. \tag{2.3}$$

Similar conclusion holds for T_-. Such u is unique. In addition, if $(u_0, u_1) \in \dot{H}^{s_1} \times H^{s_1-1}$ for some $s_1 > s$, then u satisfies (2.3) and (2.3) with s being replaced by s_1 on any

$I = (T_1, T_2) \Subset (T_-, T_+)$. T_+ and T_- are called the maximal time of existence for the solution u. □

2.2 Critical wellposedness results

Perhaps not surprisingly, our work depends crucially on the regularity results of Tao [50, 51], Tataru [57], and Sterbenz-Tataru [47, 48]. See also the work of Krieger and Schlag [37]. In this section, we recall some important results for wave maps in the energy space from [47, 51, 57], that will be needed below.

In order to control the solution at the $\dot{H}^1 \times L^2$ level of regularity, we need to use more sophisticated spaces. The precise definitions of these spaces are not very important for us, but we shall need the following properties that we briefly review below.

Fix a radial function $\Phi \in C_c^\infty(R^2)$ with $\Phi|_{B_1} \equiv 1$ and supp $\Phi \Subset B_2$. Let $\Psi(x) := \Phi(x) - \Phi(2x)$, and $\Psi_k(x) = \Psi(x/2^k)$ for each $k \in \mathbb{Z}$. Then supp $\Psi \Subset B_2 \backslash B_{1/2}$, and

$$\sum_{k\in\mathbb{Z}} \Psi_k \equiv 1, \quad \text{for } |\xi| \neq 0.$$

Recall that the Littlewood–Paley projection P_k and $P_{<k}$ are defined as

$$\widehat{P_k f}(\xi) = \Psi_k(\xi)\widehat{f}(\xi)$$

and

$$P_{<k} f = \sum_{k'<k} P_{k'} f.$$

We will also use the notations $u_k := P_k u$ and $u_{<k} = P_{<k} u$. Then

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

for all $f \in L^2(R^2)$. We use the same definitions as in [51] for the spaces $S[k]$, $N[k]$, which are translation invariant Banach spaces of distributions on $\mathbb{R}^2_x \times \mathbb{R}_t$ containing Schwartz functions whose partial Fourier transform in the x variable is supported in $\{2^{k-3} \leq |\xi| \leq 2^{k+3}\}$, $\{2^{k-4} \leq |\xi| \leq 2^{k+4}\}$, respectively. For each k, we shall use the space $S[k]$ to hold the frequency localized piece $P_k u$ of the solution u, and use the space $N[k]$ to hold the frequency localized piece $P_k f$ of the nonlinearity $f := u\,\partial_\alpha u^\dagger \partial^\alpha u$. Define the $S(1)$ norm as

$$\|f\|_{S(1)} := \|f\|_{L^\infty} + \sup_k \|P_k f\|_{S[k]}. \tag{2.4}$$

The spaces $S[k]$ and $N[k]$ satisfy the following properties.

Theorem 2.2. There exists a small universal constant $\kappa > 0$, such that

(1) **(Algebra property)** For Schwartz functions ϕ, ψ with $\psi \in S[k_2]$, we have

$$\|P_k(\phi\psi)\|_{S[k]} \lesssim 2^{-\kappa(k_2-k)_+}\|\phi\|_{S(1)}\|\psi\|_{S[k_2]}; \tag{2.5}$$

(2) **(Product property)** For Schwartz functions f, ψ with $f \in N[k_2]$, we have

$$\|P_k(f\psi)\|_{N[k]} \lesssim 2^{-\kappa(k_2-k)_+}\|\psi\|_{S(1)}\|f\|_{N[k_2]}; \tag{2.6}$$

(3) **(Null form estimate)** For Schwartz functions ϕ, ψ with $\phi \in S[k_1]$, $\psi \in S[k_2]$, we have

$$\|P_k(\partial^\alpha\phi\,\partial_\alpha\psi)\|_{N[k]} \lesssim 2^{-\kappa(\max\{k_1,k_2\}-k)_+}\|\phi\|_{S[k_1]}\|\psi\|_{S[k_2]}; \tag{2.7}$$

(4) **(Trilinear estimate)** For Schwartz functions ϕ, φ, ψ with $\phi \in S[k_1]$, $\varphi \in S[k_2]$ and $\psi \in S[k_3]$, we have

$$\begin{aligned}&\|P_k(\phi\,\partial^\alpha\varphi\partial_\alpha\psi)\|_{N[k]}\\ &\quad\lesssim 2^{-\kappa(k_1-\min\{k_2,k_3\})_+}2^{-\kappa(\max\{k_1,k_2,k_3\}-k)_+}\|\phi\|_{S[k_1]}\|\varphi\|_{S[k_2]}\|\psi\|_{S[k_3]}.\end{aligned} \tag{2.8}$$

(5) **(Linear wave estimate)** For solution u^L to the linear wave equation

$$\partial_{tt}u^L - \Delta u^L = f$$

with initial data (u_0, u_1), we have

$$\|P_k u^L\|_{S[k]} \lesssim \|P_k(u_0, u_1)\|_{\dot{H}^1\times L^2} + \|P_k f\|_{N[k]}. \tag{2.9}$$

(6) **($S[k]$ controls energy)** For $u \in S[k]$, we have

$$\|\nabla_{x,t}u\|_{L^\infty_t L^2_x} \lesssim \|u\|_{S[k]}. \tag{2.10}$$

□

Remark. These estimates were proved in [51], and some of them are slightly more general than those stated in the main summary of the properties of $S[k]$, $N[k]$ in Theorem 3 from [51]. However, they can be found elsewhere in that paper. More precisely, the algebra estimate (2.5) is a consequence of equation (125) and (126) at page 516; the product estimate (2.6) is a consequence of (119) at page 510; the null form estimate (2.7) is (134) at page 523; the trilinear estimate (2.8) is taken from the first formula at page 529. We also note that $P_{k'}$ is bounded from $S[k]$ to $S[k]$, by the translation invariance of the Banach space $S[k]$. (6) implies that $\|u\|_{L^\infty} \lesssim \|u\|_{S[k]}$. Another useful property of $S[k]$ is the weak stability of $S[k]$: if $u_i \to u$ in the sense of distributions and $u_i \in S[k]$ with $\|u_i\|_{S[k]} \leq 1$, then $\|u\|_{S[k]} \leq 1$. See a similar statement in (vii) of page 323 in [58]. We will use these estimates extensively below. □

Tao [51] introduced a very useful notation to keep track of multilinear expressions. More precisely, for scalar functions $\phi_1, \ldots, \phi_l$, we use $L(\phi_1, \ldots, \phi_l)$ to denote multilinear expression of the form

$$L(\phi_1, \ldots, \phi_l) := \int K(y_1, \ldots, y_l)\, \phi_1(x - y_1) \cdots \phi_l(x - y_l)\, \mathrm{d}y_1 \cdots \mathrm{d}y_l$$

with a measure K of bounded mass. In many cases, $\phi_1, \ldots, \phi_l$ could also be expressions involving components $\phi_1^{j_1}, \ldots, \phi_l^{j_l}$ and in such cases, we also assume that K depends on $j_1, \ldots, j_l$, but for the ease of notations, we shall suppress this dependence. By the translation invariance of the spaces $S[k]$, $N[k]$, the estimates in Theorem 2.2 extend to expressions of the form $L(\phi, \psi)$, $L(\partial^\alpha \phi, \partial_\alpha \psi)$ instead of just $\phi\,\psi$ and $\partial^\alpha \phi\, \partial_\alpha \psi$.

Let us record here the following useful Lemma from [51].

Lemma 2.1. For Schwartz functions f, g, we have

$$P_k(f\,g) - P_k f \cdot g = 2^{-k} L(f, \nabla g). \tag{2.11}$$

□

Proof. This Lemma is taken from [51], we include the short proof for the convenience of readers. We have

$$\begin{aligned}
&P_k(f\,g)(x) - P_k f(x)\, g(x) \\
&\quad = \int 4^k \check{\Psi}(2^k y) f(x-y) g(x-y) \mathrm{d}y - \int 4^k \check{\Psi}(2^k y) f(x-y) g(x) \mathrm{d}y \\
&\quad = \int_0^1 \int_{R^2} -4^k \check{\Psi}(2^k y) f(x-y)\, y_j \partial^j g(x - ty) \mathrm{d}t \mathrm{d}y \\
&\quad = -2^{-k} \int_0^1 \int_{R^2} 4^k (2^k y_j)\, \check{\Psi}(2^k y) f(x-y)\, \partial^j g(x - ty) \mathrm{d}t \mathrm{d}y \\
&\quad = 2^{-k} L(f, \nabla g).
\end{aligned}$$

The proof is complete. ■

Let us recall the definition of *frequency envelop* introduced in [51]. Fix positive ϑ such that $\vartheta \le \frac{\kappa}{100}$, where κ is as in Theorem 2.2.

Definition 2.1. $(c_k) \in \ell^2$ is called a frequency envelop if $c_k > 0$ and $c_{k_1} \le 2^{\vartheta|k_1-k_2|} c_{k_2}$. □

For any frequency envelop $c = (c_k)$, define the norm $S(c)$ as

$$\|\phi\|_{S(c)} := \|\phi\|_{L^\infty} + \sup_k c_k^{-1} \|P_k \phi\|_{S[k]} \tag{2.12}$$

and the space $S(c)$ as

$$S(c) := \{f \in L^\infty : \|f\|_{S(c)} < \infty\}. \tag{2.13}$$

Note that $1 \in S(c)$. The main property of the space $S(c)$ that we shall use below is that $S(c)$ a Banach algebra.

Lemma 2.2. $S(c)$ is a Banach algebra. □

Proof. This was proved in [51]. We include the short proof for the convenience of readers. We need to prove

$$\|\phi\, \psi\|_{S(c)} \lesssim \|\phi\|_{S(c)}\, \|\psi\|_{S(c)}. \tag{2.14}$$

We note that $\|\phi\|_{S(1)} \lesssim_c \|\phi\|_{S(c)}$ and $\|\phi_{<k}\|_{S(c)} \lesssim \|\phi\|_{S(c)}$. We can normalize $\|\phi\|_{S(c)} = \|\psi\|_{S(c)} = 1$. For each $k \in \mathbb{Z}$, we have

$$\begin{aligned} &\|P_k(\phi\, \psi)\|_{S[k]} \\ &\quad = \left\| P_k\left(\phi_{>k-10}\psi\right) + P_k\left(\phi_{\le k-10}\, \psi_{>k-10}\right) + P_k\left(\phi_{\le k-10}\psi_{\le k-10}\right)\right\|_{S[k]}. \end{aligned}$$

Note that

$$P_k\left(\phi_{\le k-10}\psi_{\le k-10}\right) \equiv 0.$$

We get that

$$\begin{aligned} \|P_k(\phi\, \psi)\|_{S[k]} &\lesssim \sum_{k_1 > k-10} \left\|P_k\left(P_{k_1}\phi\, \psi\right)\right\|_{S[k]} + \sum_{k_2 > k-10} \left\|P_k\left(\phi_{\le k-10}\, P_{k_2}\psi\right)\right\|_{S[k]} \\ &\lesssim \sum_{k_1 > k-10} 2^{-\kappa(k_1-k)_+} \left\|P_{k_1}\phi\right\|_{S[k_1]} \|\psi\|_{S(1)} \end{aligned}$$

$$
\begin{aligned}
&+ \sum_{k_2 > k-10} 2^{-\kappa(k_2-k)_+} \left\| P_{k_2} \psi \right\|_{S[k_2]} \|\phi_{\leq k-10}\|_{S(1)} \\
\lesssim & \sum_{k_1 > k-10} 2^{-\kappa(k_1-k)} c_{k_1} + \sum_{k_2 > k-10} 2^{-\kappa(k_2-k)} c_{k_2} \\
\lesssim & \sum_{k' > k-10} 2^{-(\kappa-\vartheta)(k'-k)} c_k \lesssim c_k
\end{aligned}
$$

and this finishes the proof. ■

Let us recall the following global wellposedness theorem for wave maps from Tao [51].

Theorem 2.3. There exists an $\varepsilon > 0$ sufficiently small such that the following is true. Suppose that (u_0, u_1) is smooth, $u_0 - u_\infty$, u_1 are compactly supported, and that $u_1^\dagger \cdot u_0 \equiv 0$. Assume that $\left(\|P_k(u_0, u_1)\|_{\dot{H}^1 \times L^2}\right)$ lies under a frequency envelop $c = (c_k)$ (For non-negative sequences (a_k) and (b_k), we say that (a_k) lies below (b_k) if $a_k \leq b_k$ for each k.) with

$$\|c_k\|_{\ell^2} \leq \varepsilon.$$

Then the wave map u with initial data (u_0, u_1) is global, and moreover

$$\|P_k u\|_{S[k]} + \sup_{t \in R} \|P_k \overrightarrow{u}(t)\|_{\dot{H}^1 \times L^2} \leq C c_k \tag{2.15}$$

for some universal C. □

Remark. By approximations by smooth maps, and the wellposedness for equation (1.2) in $\dot{H}^s \times H^{s-1}$ for $s > 1$, we can relax the smoothness requirement for the initial data in the above theorem to $(u_0, u_1) \in \dot{H}^s \times H^{s-1}$.

Fix $\epsilon_* > 0$ be sufficiently small, so that classical wave maps with energy smaller than $C\epsilon_*$ exists globally for a sufficiently large universal $C > 1$. In later sections, we shall need the following local-in-space smoothness result, when the initial data are locally but not globally smooth. □

Lemma 2.3. Let $(u_0, u_1) \in \dot{H}^s \times H^{s-1}$ for some $s > 1$ and that $(u_0 - u_\infty, u_1)$ is compactly supported, with $|u_0| \equiv 1$ and $u_0 \cdot u_1 \equiv 0$. Assume that (u_0, u_1) is smooth in B_1, and that $\|(u_0, u_1)\|_{\dot{H}^1 \times L^2} \leq \epsilon_*$. Then the global solution u is smooth in $\{(x, t) : |x| < 1 - |t|\}$. □

By the small energy global existence result and the subcritical Cauchy theory, we get that

$$\sup_{0\leq t<1}\|u(t)\|_{\dot{H}^s} \leq C\left(\|(u_0-u_\infty,\, u_1)\|_{H^s\times H^{s-1}}\right). \tag{2.16}$$

Denote $\lambda := C\left(\|(u_0-u_\infty,\, u_1)\|_{H^s\times H^{s-1}}\right)$. Fix $\overline{r}>0$ small, and take $B_{\overline{r}}(\overline{x})\subset B_1$. Set

$$\overline{u}_0 = \frac{1}{\pi\overline{r}^2}\int_{B_{\overline{r}}(\overline{x})} u.$$

Then by Sobolev inequality, we obtain that

$$\|u_0-\overline{u}_0\|_{L^\infty(B_{\overline{r}}(\overline{x}))} \lesssim_\lambda \overline{r}^{s-1}. \tag{2.17}$$

Take a smooth cutoff function η such that $\eta\equiv 1$ in $B_{\overline{r}-\overline{r}^{1+\delta}}(\overline{x})$ with some $\delta\in(0,\,2(s-1))$, and $\eta\equiv 0$ outside $B_{\overline{r}}(\overline{x})$. In addition, we can require that

$$|\nabla\eta| \lesssim \overline{r}^{-1-\delta}. \tag{2.18}$$

Define

$$(\widetilde{u}_0,\,\widetilde{u}_1) = \left(P\left[\eta\,(u_0-\overline{u}_0)+\overline{u}_0\right],\, \eta\, u_1\right),$$

where for each vector $v\neq 0$

$$Pv = \frac{v}{|v|}.$$

Then $\widetilde{u}_0$, $\widetilde{u}_1$ are smooth, and

$$(\widetilde{u}_0,\,\widetilde{u}_1)\equiv(u_0,\,u_1), \quad \text{in } B_{\overline{r}-\overline{r}^{1+\delta}}(\overline{x}). \tag{2.19}$$

Moreover, we can verify by direct computation thanks to (2.17) and (2.18) that

$$\|(\widetilde{u}_0,\,\widetilde{u}_1)\|_{\dot{H}^1\times L^2} \lesssim \epsilon_*,$$

if $\overline{r}$ is chosen sufficiently small. Hence the solution $\widetilde{u}$ to the wave map equation with the initial data $(\widetilde{u}_0,\,\widetilde{u}_1)$ is smooth and global. By (2.19), $u\equiv\widetilde{u}$ for $|x-\overline{x}|<\overline{r}-\overline{r}^{1+\delta}-|t|$, and is thus smooth for $|x-\overline{x}|<\overline{r}-\overline{r}^{1+\delta}-|t|$. By moving around $\overline{x}$ and finite speed of propagation, we conclude that u is smooth in $\{(x,t) : |x|<1-2\overline{r}^{1+\delta}-|t|,\, |t|<\overline{r}\}$. We can apply the same technique at $|t|=\overline{r}$, $2\overline{r}$ and so on, and conclude recursively that u

is smooth in $\{(x,t): |x| - 2k\bar{r}^{1+\delta} - |t|, |t| < k\bar{r}\}$ for $k = 1, 2, \dots$ with $(k+2)\bar{r} < 1$. Hence, u is smooth in $\{(x,t): |x| < 1 - C\bar{r}^{\delta}, |t| < 1 - 3\bar{r}\}$. Since $\bar{r}$ can be taken arbitrarily small, the lemma follows.

Since the global regularity result for small energy requires that the initial data belongs to a subcritical space $H^s \times H^{s-1}$ for some $s > 1$, (See however Tataru [58] where a notion of finite energy solution was introduced.) we shall need the following lemma when we deal with some initial data which is $C^\infty(R^2\backslash\{0\})$ but may fail to be in $H^s \times H^{s-1}$ globally for any $s > 1$.

Lemma 2.4. Suppose that $(u_0, u_1) \in C^\infty(R^2\backslash\{0\})$, and that $(u_0 - u_\infty, u_1)$ is compactly supported. Assume that

$$\|(u_0, u_1)\|_{\dot{H}^1 \times L^2} \leq \epsilon_*. \tag{2.20}$$

Then there exists a unique smooth $u \in C^\infty((x,t): |x| > |t|\})$ such that u solves the wave map equation in $\{(x,t): |x| > |t|\}$. Moreover

$$\lim_{t\to 0} \|\vec{u}(\cdot, t) - (u_0, u_1)\|_{\dot{H}^1 \times L^2(|x|>|t|)} = 0. \tag{2.21}$$

Similar results hold if we assume instead that $(u_0, u_1) \in H^s_{\text{loc}} \times H^{s-1}_{\text{loc}}(R^2\backslash\{0\})$, and in this case, $\vec{u} \in H^s_{\text{loc}} \times H^{s-1}_{\text{loc}}(|x| > |t|)$. □

Proof. We shall prove only the first part of the lemma. The proof of the second part is clear from the same argument. Let us firstly prove the existence of u claimed in the lemma. For any $r > 0$, since

$$\int_{B_r\backslash B_{\frac{r}{2}}} |\nabla u_0|^2 + |u_1|^2 \, dx \leq \epsilon_*^2,$$

we can find $\bar{r} \in \left(\frac{r}{2}, r\right)$ with

$$\int_{|x|=\bar{r}} |\not\partial u_0|^2 d\sigma \lesssim \frac{\epsilon_*^2}{\bar{r}}.$$

Denote

$$\bar{u}_0 = \frac{1}{2\pi\bar{r}} \int_{\partial B_{\bar{r}}} u_0.$$

Then by Sobolev inequality, we get that

$$\|u_0 - \bar{u}_0\|_{L^\infty(\partial B_{\bar{r}})} \lesssim \epsilon_*.$$

Thus from the fact that $|u_0| \equiv 1$, we see that $|\overline{u}_0| \gtrsim 1$. Take smooth cutoff function η such that $\eta \equiv 1$ for $|x| \geq \overline{r}$ and $\eta \equiv 0$ for $|x| < \frac{\overline{r}}{2}$ with $|\nabla \eta| \lesssim (\overline{r})^{-1}$. Define

$$(\widetilde{u}_0, \widetilde{u}_1) = \begin{cases} (u_0, u_1) & \text{in } B^c_{\overline{r}}; \\ (P\left[\eta(r)(u_0(\overline{r}\theta) - \overline{u}_0) + \overline{u}_0\right], 0) & \text{in } B_{\overline{r}}. \end{cases}$$

Then

$$(\widetilde{u}_0 - u_\infty, \widetilde{u}_1) \in H^s \times H^{s-1}$$

for $s < \frac{3}{2}$, and direct computation shows that

$$\|(\widetilde{u}_0, \widetilde{u}_1)\|_{\dot{H}^1 \times L^2} \lesssim \epsilon_*.$$

Note also that $(\widetilde{u}_0, \widetilde{u}_1)$ is smooth for $|x| > \overline{r}$. Hence by small data theory and Lemma 2.3 the solution $\widetilde{u}$ to the wave map equation with initial data $(\widetilde{u}_0, \widetilde{u}_1)$ is global, and is smooth in $|x| > \overline{r} + |t|$. By taking $r \to 0+$ and the finite speed of propagation, we see that

$$u = \lim_{r \to 0+} \widetilde{u}$$

exists in $|x| > |t|$ and is smooth. We now turn to the proof of (2.21). Let $\widetilde{u}$ be the solution as before, corresponding to $\overline{r}$, then

$$\widetilde{u} \equiv u, \ \text{ for } \ |x| > \overline{r} + |t|, \tag{2.22}$$

and $\widetilde{u}$ is continuous in $\dot{H}^1 \times L^2$ for $t \in (0, 1]$. For any $\epsilon > 0$, we can choose $\overline{r}$ sufficiently small, such that

$$\|(u_0, u_1)\|_{\dot{H}^1 \times L^2(B_{4\overline{r}})} < \epsilon.$$

Then by energy flux identity (say for $t > 0$ and any $\epsilon > 0$),

$$\begin{aligned} &\int_{t+\epsilon<|x|<4\overline{r}-t} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2}\right)(x,t)\mathrm{d}x \\ &\quad + \frac{1}{\sqrt{2}} \int_0^t \int_{|x|=4\overline{r}-t} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} - \frac{x}{|x|} \cdot \nabla u \, \partial_t u\right) \mathrm{d}\sigma \mathrm{d}s \\ &\quad + \frac{1}{\sqrt{2}} \int_0^t \int_{|x|=t+\epsilon} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} + \frac{x}{|x|} \cdot \nabla u \, \partial_t u\right) \mathrm{d}\sigma \mathrm{d}s \\ &= \int_{B_{4\overline{r}} \setminus B_\epsilon} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2}\right)(x,0)\mathrm{d}x, \end{aligned}$$

we see that

$$\|\vec{u}(t)\|_{\dot{H}^1\times L^2(B_{2\bar{r}}\setminus B_{|t|})} \leq \|(u_0, u_1)\|_{\dot{H}^1\times L^2(B_{4\bar{r}})} < \epsilon, \text{ for } |t| < \bar{r}. \tag{2.23}$$

Since $\widetilde{u}$ is continuous in the energy space and $\vec{\widetilde{u}}(x,0) = (u_0, u_1)(x)$ for $|x| > \bar{r}$, we see that for sufficiently small $t_1 \in (0, \bar{r})$ and $|t| < t_1$,

$$\|\vec{\widetilde{u}}(t) - (u_0, u_1)\|_{\dot{H}^1\times L^2(|x|>\bar{r})} < \epsilon. \tag{2.24}$$

Combining (2.23), (2.24), and (2.22), we conclude that for $|t| \leq t_1$

$$\begin{aligned}
&\|\vec{u}(t) - (u_0, u_1)\|_{\dot{H}^1\times L^2(|x|>|t|)} \\
&\quad\leq \|\vec{u}(t) - (u_0, u_1)\|_{\dot{H}^1\times L^2(|x|>2\bar{r})} + \|\vec{u}(t) - (u_0, u_1)\|_{\dot{H}^1\times L^2(|t|<|x|<2\bar{r})} \\
&\quad\leq \|\vec{\widetilde{u}}(t) - (u_0, u_1)\|_{\dot{H}^1\times L^2(|x|>2\bar{r})} + 2\epsilon \\
&\quad\leq 3\epsilon.
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, the lemma is proved. ■

By finite speed of propagation and small data global existence, understanding the energy concentration is important for studying the dynamics of the wave maps. To measure the energy concentration, let us define for a wave map u the "energy concentration radius"

$$\begin{aligned}
&r(\epsilon_*, t) := \\
&\inf\left\{r > 0 : \text{ there exists } \bar{x} \text{ such that } \int_{B_r(\bar{x})} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2}\right)(x,t)\mathrm{d}x > \epsilon_*\right\}.
\end{aligned} \tag{2.25}$$

We adopt the convention that if the set is empty, then the infimum is infinity. The small energy global existence result, Theorem 2.3, and the finite speed of propagation imply that if wave map u blows up at a finite time T_+, then $r(\epsilon_*, t) \to 0+$ as $t \to T_+$. This is a very important piece of information that allows us to zoom in a small region near the blow up point and study the details of the blow up there. Unfortunately, knowing only that the energy concentrates in the small scales does not in itself allow one to "extract" a nontrivial blow up profile in the limit, as we zooms in more and more. This is because a priori the energy can be concentrated in quite an arbitrary way, given that we do not (and it is probably not possible) to obtain control any higher order regularity beyond the energy when the time is close to the blow up time. To obtain a nontrivial blow up profile, the following result due to Sterbenz-Tataru [47] plays an essential role. (More precisely, this result is used to rule out the situation that all energy near the blow up point concentrates near the boundary of lightcone. The control inside the lightcone turns out to be quite favorable.)

Theorem 2.4. There exists a function $\epsilon(E)$ with $0 < \epsilon(E) \ll 1$ of the energy E such that if u is a classical solution to (1.2) in $I \times R^2 = [a, b] \times R^2$, with energy E and

$$\sup_{t\in I} \sup_k \left\|(P_k u, 2^{-k} P_k \partial_t u)(t)\right\|_{L^\infty \times L^\infty} < \epsilon(E), \tag{2.26}$$

then the energy concentration radius $r(\epsilon_*, t)$ has a uniform lower bound on I:

$$\inf_{t\in I} r(\epsilon_*, t) \geq r_0 > 0. \tag{2.27}$$

□

3 Channel of Energy Inequality for Wave Maps with Small Energy

In this section, we prove the channel of energy inequality for small wave maps. Let us begin with the following linear channel of energy inequality for outgoing waves, which is a slightly more quantitative two-dimensional version of the channel of energy inequality that played a decisive role in [14].

Lemma 3.1. Fix $\gamma \in (0, 1)$. There exists $\mu = \mu(\gamma) > 0$ sufficiently small such that the following statement is true. Let v be a finite energy solution to the linear wave equation

$$\partial_{tt} v - \Delta v = 0, \text{ in } R^2 \times [0, \infty)$$

with initial data $(v_0, v_1) \in \dot{H}^1 \times L^2$ satisfying

$$\|(v_0, v_1)\|_{\dot{H}^1 \times L^2(B^c_{1+\mu} \cup B_{1-\mu})} + \|\partial\!\!\!/ v_0\|_{L^2} + \|\partial_r v_0 + v_1\|_{L^2} \leq \mu \|(v_0, v_1)\|_{\dot{H}^1 \times L^2}. \tag{3.1}$$

We also assume that $v_0 \equiv v_\infty$ for some constant v_∞ for large x. Then for all $t \geq 0$, we have

$$\int_{|x|\geq \gamma+t} |\nabla_{x,t} v|^2(x, t)\mathrm{d}x \geq \gamma \|(v_0, v_1)\|^2_{\dot{H}^1 \times L^2}. \tag{3.2}$$

□

We can normalize the initial data so that $\|(v_0, v_1)\|_{\dot{H}^1\times L^2} = 1$. Let $\alpha = \int_{\{1+\mu\le|x|\le 2(1+\mu)\}} v_0(x)\mathrm{d}x$. By Poincaré inequality

$$\int_{1+\mu\le|x|\le 2(1+\mu)} |v_0(x)-\alpha|^2\mathrm{d}x \lesssim \int_{1+\mu\le|x|\le 2(1+\mu)} |\nabla v_0(x)|^2\mathrm{d}x, \tag{3.3}$$

where the implicit constant is independent of $\mu\le 1$.

Take a non-negative radial $\eta\in C_c^\infty(R^2)$ with $\eta\equiv 1$ on $\overline{B_{1+\mu}}$ and $\operatorname{supp}\eta\Subset B_{1+\mu^{1/2}}$ satisfying $|\nabla\eta|\lesssim\mu^{-1/2}$. Define

$$(\widetilde{v}_0,\widetilde{v}_1)=\eta(x)\,(v_0(x)-\alpha, v_1(x)).$$

Using (3.3), the bound $|\nabla\eta|\lesssim\mu^{-1/2}$ and (3.1), we obtain:

$$\|(\nabla(\widetilde{v}-v_0),\widetilde{v}_1-v_1))\|^2_{L^2\times L^2}\lesssim\frac{1}{\mu}\int_{|x|\ge 1+\mu}|\nabla v_0|^2+\int_{|x|\ge 1+\mu}|u_1|^2\lesssim\mu. \tag{3.4}$$

By Sobolev and Hölder inequalities

$$\||D|^{\frac12}\widetilde{v}_0\|_{L^2}\lesssim\|\nabla\widetilde{v}_0\|_{L^{\frac43}}\lesssim\|\nabla\widetilde{v}_0\|_{L^2(\{|x|\le 1-\mu\})}+\mu^{\frac18}\|\nabla\widetilde{v}_0\|_{L^2(\{1-\mu\le|x|\le 1+\mu^{1/2}\})}\lesssim\mu^{\frac18}$$
$$\||D|^{-\frac12}\widetilde{v}_1\|_{L^2}\lesssim\|v_1\|_{L^{\frac43}}\lesssim\|v_1\|_{L^2(\{|x|\le 1-\mu\})}+\mu^{\frac18}\|v_1\|_{L^2(\{1-\mu\le|x|\le 1+\mu^{1/2}\})}\lesssim\mu^{\frac18}.$$

By conservation of the $\dot{H}^{1/2}\times\dot{H}^{-1/2}$ norm for the linear wave equation, we obtain the for all $t\in\mathbb{R}$,

$$\left|\int\widetilde{v}_t(x,t)\widetilde{v}(x,t)\mathrm{d}x\right|\lesssim\left\||D|^{1/2}\widetilde{v}\right\|_{L^2}\left\||D|^{-1/2}\widetilde{v}_t\right\|_{L^2}\lesssim\mu^{1/4}. \tag{3.5}$$

Let $\widetilde{v}$ be the solution to the linear wave equation with initial data $(\widetilde{v}_0,\widetilde{v}_1)$. By direct computation, we see that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{R^2}-\widetilde{v}_t\left(x\cdot\nabla\widetilde{v}+\frac12\widetilde{v}\right)(x,t)\mathrm{d}x=\mathcal{E}(\widetilde{v}):=\mathcal{E}_0. \tag{3.6}$$

Hence, by (3.5) and the outgoing condition (3.1), we get that

$$\begin{aligned}\int_{R^2}-\widetilde{v}_t\,x\cdot\nabla\widetilde{v}(x,t)\mathrm{d}x&=\mathcal{E}_0\,t+\int_{R^2}-\widetilde{v}_1\left(x\cdot\nabla\widetilde{v}_0+\frac12\widetilde{v}_0\right)(x)\mathrm{d}x\\&\quad+O(\mu^{1/4})\\&=\mathcal{E}_0\,(t+1)+O(\mu^{1/4}).\end{aligned}$$

On the other hand, by the finite speed of propagation, supp $\widetilde{v}(\cdot, t) \Subset B_{1+\mu^{1/2}+t}$ for all $t \geq 0$, and thus

$$\int_{R^2} -\widetilde{v}_t\, x\cdot\nabla\widetilde{v}(x,t)\mathrm{d}x \leq \int_{|x|>\gamma+t} (1+\mu^{1/2}+t)\left(\frac{|\widetilde{v}_t|^2}{2}+\frac{|\nabla\widetilde{v}|^2}{2}\right)(x,t)\mathrm{d}x$$
$$+\ (\gamma+t)\int_{|x|<\gamma+t}\left(\frac{|\widetilde{v}_t|^2}{2}+\frac{|\nabla\widetilde{v}|^2}{2}\right)(x,t)\mathrm{d}x$$
$$=(\gamma+t)\mathcal{E}_0-(\gamma+t)\int_{|x|>\gamma+t}\left(\frac{|\widetilde{v}_t|^2}{2}+\frac{|\nabla\widetilde{v}|^2}{2}\right)(x,t)\mathrm{d}x+$$
$$+\ (1+\mu^{1/2}+t)\int_{|x|>\gamma+t}\left(\frac{|\widetilde{v}_t|^2}{2}+\frac{|\nabla\widetilde{v}|^2}{2}\right)(x,t)\mathrm{d}x.$$

Combining this and the above, we see that

$$(1+\mu^{1/2}-\gamma)\int_{|x|>\gamma+t}\left(\frac{|\widetilde{v}_t|^2}{2}+\frac{|\nabla\widetilde{v}|^2}{2}\right)(x,t)\mathrm{d}x$$
$$\geq (1-\gamma)\mathcal{E}_0+O(\mu^{1/4}).$$

By choosing μ sufficiently small, we obtain the channel of energy inequality for $\widetilde{v}$, and consequently also for v, by (3.4).

As mentioned in the introduction, one of the main goals of this article is to extend the channel of energy arguments to the wave map setting. As a first step towards understanding the implications of the channel of energy property of linear wave equations on the wave maps, we prove the following result for small energy wave maps. The extension to large energy case seems to require nontrivial improvement in the perturbative techniques for the wave maps.

Theorem 3.1. Fix $\beta \in (0,1)$. There exist a small $\delta = \delta(\beta) > 0$ and sufficiently small $\epsilon_0 = \epsilon_0(\beta) > 0$, such that if u is a classical wave map with energy $\mathcal{E}(\overrightarrow{u}) < \epsilon_0^2$ satisfying

$$\|(u_0,\, u_1)\|_{\dot{H}^1\times L^2\left(B^c_{1+\delta}\cup B_{1-\delta}\right)}+\|\partial\!\!\!/ u_0\|_{L^2}+\|\partial_r u_0+u_1\|_{L^2}\leq \delta\|(u_0,\, u_1)\|_{\dot{H}^1\times L^2}, \tag{3.7}$$

then for all $t \geq 0$, we have

$$\int_{|x|>\beta+t}|\nabla_{x,t}u|^2(x,t)\mathrm{d}x \geq \beta\,\|(u_0,\, u_1)\|^2_{\dot{H}^1\times L^2}. \tag{3.8}$$

□

Proof. Denote $\epsilon := \|(u_0, u_1)\|_{\dot{H}^1\times L^2} \lesssim \epsilon_0$. To apply Theorem 2.3, let us define the following frequency envelop

$$c_k := \sup_{j\in\mathbb{Z}} 2^{-\vartheta|k-j|}\|(P_j u_0, P_j u_1)\|_{\dot{H}^1\times L^2}. \tag{3.9}$$

Then one can verify that $c = (c_k)$ is a frequency envelop and that $\left(\|P_k(u_0, u_1)\|_{\dot{H}^1\times L^2}\right)$ lies below it. In addition,

$$\|(c_k)\|_{\ell^2} \lesssim \epsilon.$$

By Theorem 2.3, if ϵ_0 is chosen sufficiently small, then the wave map u is globally defined, and satisfies (2.15).

Since the proof is a bit lengthy, we divide the arguments in several steps. ■

Step 1 : *Reduction to proving channel of energy inequality for frequency pieces.* In this step, our main goal is to show that there exists a set $\mathcal{K}$ of *good frequencies*, such that

$$\sum_{m\in\mathcal{K}} \|P_m(u_0, u_1)\|^2_{\dot{H}^1\times L^2} \geq (1 - C\delta^{\frac{1}{12}})\|(u_0, u_1)\|^2_{\dot{H}^1\times L^2} \tag{3.10}$$

and that for any $m \in \mathcal{K}$, 2^m is "high frequency", and that it suffices to prove the channel of energy inequality for each $m \in \mathcal{K}$.

Substep (1): Control of the low frequency component.

Fix k_0 large, whose precise value is to be determined below. We shall show that the total energy with frequency $\leq 2^{k_0}$ is small in a suitable sense. Assume firstly that $2^{-k_0} > C\delta$. Let us bound the low frequency energy of (u_0, u_1), that is

$$\left\|P_{\leq k_0}(u_0, u_1)\right\|_{\dot{H}^1\times L^2}.$$

We can write

$$\begin{aligned}\nabla u_0 &= (\nabla u_0)\chi_{B^c_{1+\delta}\cup B_{1-\delta}} + (\nabla u_0)\chi_{B_{1+\delta}\setminus B_{1-\delta}},\\ u_1 &= u_1\chi_{B^c_{1+\delta}\cup B_{1-\delta}} + u_1\chi_{B_{1+\delta}\setminus B_{1-\delta}}.\end{aligned}$$

By the assumption on (u_0, u_1),

$$\begin{aligned}\|(\nabla u_0)\chi_{B^c_{1+\delta}\cup B_{1-\delta}}\|_{L^2} + \|u_1\chi_{B^c_{1+\delta}\cup B_{1-\delta}}\|_{L^2}&\\ \lesssim \delta\|(u_0, u_1)\|_{\dot{H}^1\times L^2}.&\end{aligned}$$

Thus,

$$\left\|P_{<k_0}\left(\nabla u_0\chi_{B^c_{1+\delta}\cup B_{1-\delta}}\right)\right\|_{L^2} + \left\|P_{<k_0}\left(u_1\chi_{B^c_{1+\delta}\cup B_{1-\delta}}\right)\right\|_{L^2} \lesssim \delta\|(u_0, u_1)\|_{\dot{H}^1\times L^2}. \tag{3.11}$$

Denote $f = (\nabla u_0)\chi_{B_{1+\delta}\setminus B_{1-\delta}}$. Then f is compactly supported in $\overline{B_{1+\delta}}\setminus B_{1-\delta}$, and

$$\|f\|_{L^2} \leq \|(u_0, u_1)\|_{\dot{H}^1\times L^2}.$$

By Bernstein's inequality, then Cauchy–Schwarz

$$\|P_{\leq k_0}f\|_{L^2} \lesssim 2^{k_0}\|f\|_{L^1} \lesssim 2^{k_0}\delta^{\frac{1}{2}}\|f\|_{L^2} \lesssim 2^{k_0}\delta^{\frac{1}{2}}\|(u_0, u_1)\|_{\dot{H}^1\times L^2}.$$

Choosing $2^{-k_0} \sim \delta^{\frac{1}{6}}$, then $\|P_{\leq k_0}f\|_{L^2} \lesssim \delta^{\frac{1}{3}}\|(u_0, u_1)\|_{\dot{H}^1\times L^2}$, that is,

$$\left\|P_{\leq k_0}\left[(\nabla u_0)\chi_{B_{1+\delta}\setminus B_{1-\delta}}\right]\right\|_{L^2} \lesssim \delta^{\frac{1}{3}}\|(u_0, u_1)\|_{\dot{H}^1\times L^2}. \tag{3.12}$$

We can prove similarly that

$$\left\|P_{\leq k_0}\left[u_1\chi_{B_{1+\delta}\setminus B_{1-\delta}}\right]\right\|_{L^2} \lesssim \delta^{\frac{1}{3}}\|(u_0, u_1)\|_{\dot{H}^1\times L^2}. \tag{3.13}$$

Combining (3.11)–(3.13), we conclude that

$$\left\|P_{\leq k_0}(u_0, u_1)\right\|_{\dot{H}^1\times L^2} \lesssim \delta^{\frac{1}{3}}\|(u_0, u_1)\|_{\dot{H}^1\times L^2}. \tag{3.14}$$

Thus the low frequency energy is small.

Substep (2): persistence of condition (3.7) for most high frequencies.

Let us now consider $P_k(\nabla u_0, u_1)$ for high frequency $2^k \geq 2^{k_0}$. Fix small $\lambda > 10\delta^{\frac{1}{6}} \sim 2^{-k_0}$, whose value is to be determined below. Let us firstly bound

$$\|P_k(\nabla u_0, u_1)\|_{L^2\left(B^c_{1+\lambda}\cup B_{1-\lambda}\right)}.$$

We can decompose as before

$$\nabla u_0 = (\nabla u_0)\chi_{B^c_{1+\delta}\cup B_{1-\delta}} + (\nabla u_0)\chi_{B_{1+\delta}\setminus B_{1-\delta}},$$

$$u_1 = u_1\chi_{B^c_{1+\delta}\cup B_{1-\delta}} + u_1\chi_{B_{1+\delta}\setminus B_{1-\delta}}.$$

Denote

$$\sigma_k := \left\| P_k \left[(\nabla u_0,\, u_1) \chi_{B^c_{1+\delta} \cup B_{1-\delta}} \right] \right\|_{L^2 \times L^2},$$

then it follows from (3.7) that

$$\sum_k \sigma_k^2 \lesssim \delta^2 \|(\nabla u_0,\, u_1)\|^2_{L^2 \times L^2}. \tag{3.15}$$

Now let us consider $P_k \left[(\nabla u_0,\, u_1) \chi_{B_{1+\delta} \setminus B_{1-\delta}} \right]$ for x with $||x| - 1| > \lambda$. Denote

$$f := (\nabla u_0) \chi_{B_{1+\delta} \setminus B_{1-\delta}},$$

then

$$P_k f(x) = 4^k \int_{R^2} \check{\Psi}(2^k(x - y)) f(y) \mathrm{d}y.$$

Since f is supported in $1 - \delta \le |y| \le 1 + \delta$, and $||x| - 1| > \lambda \gg \delta$, we get that

$$|P_k f(x)| \lesssim \frac{4^k}{(2^k||x| - 1|)^M} \delta^{\frac{1}{2}} \|(u_0,\, u_1)\|_{\dot{H}^1 \times L^2}.$$

Hence

$$\begin{aligned} \|P_k f\|^2_{L^2\left(B^c_{1+\lambda} \cup B_{1-\lambda}\right)} &\lesssim 4^{(2-M)k} \delta \|(u_0,\, u_1)\|^2_{\dot{H}^1 \times L^2} \int_{\left||x|-1\right| \ge \lambda} \frac{1}{(|x| - 1)^{2M}} \mathrm{d}x \\ &\lesssim 4^{(2-M)k} \delta \lambda^{-2M} \|(u_0,\, u_1)\|^2_{\dot{H}^1 \times L^2}. \end{aligned}$$

Fix $M = 3$. Then we conclude

$$\|P_k f\|_{L^2\left(B^c_{1+\lambda} \cup B_{1-\lambda}\right)} \le 2^{-k} \delta^{\frac{1}{2}} \lambda^{-3} \|(u_0,\, u_1)\|_{\dot{H}^1 \times L^2}. \tag{3.16}$$

Take $\lambda = \delta^{\frac{1}{12}}$, then

$$\|P_k f\|_{L^2\left(B^c_{1+\lambda} \cup B_{1-\lambda}\right)} \lesssim 2^{-k} \delta^{\frac{1}{4}} \|(u_0,\, u_1)\|_{\dot{H}^1 \times L^2},$$

that is,

$$\left\| P_k \left[(\nabla u_0) \chi_{B_{1+\delta} \setminus B_{1-\delta}} \right] \right\|_{L^2\left(B^c_{1+\lambda} \cup B_{1-\lambda}\right)} \lesssim 2^{-k} \delta^{\frac{1}{4}} \|(u_0,\, u_1)\|_{\dot{H}^1 \times L^2}. \tag{3.17}$$

Similarly, we can prove that

$$\left\|P_k\left[u_1\chi_{B_{1+\delta}\backslash B_{1-\delta}}\right]\right\|_{L^2\left(B^c_{1+\lambda}\cup B_{1-\lambda}\right)}\lesssim 2^{-k}\delta^{\frac{1}{4}}\|(u_0,\,u_1)\|_{\dot{H}^1\times L^2}. \tag{3.18}$$

Now let us control

$$\|\partial_r P_k u_0+P_k u_1\|_{L^2(B_{1+\lambda}\backslash B_{1-\lambda})}+\|\not\partial P_k u_0\|_{L^2(B_{1+\lambda}\backslash B_{1-\lambda})}$$

for $k\geq k_0$ with $2^{-k_0}\sim\delta^{\frac{1}{6}}$. We have

$$\begin{aligned}
\partial_r\int_{R^2}4^k\check{\Psi}(2^k y)u_0(x-y)\,dy&=\int_{R^2}4^k\check{\Psi}(2^k y)\frac{x}{|x|}\cdot\nabla u_0(x-y)\mathrm{d}y\\
&=\int_{R^2}4^k\check{\Psi}(2^k y)\frac{x-y}{|x-y|}\cdot\nabla u_0(x-y)\mathrm{d}y\\
&\quad+\int_{R^2}4^k\check{\Psi}(2^k y)\left[\frac{x}{|x|}-\frac{x-y}{|x-y|}\right]\cdot\nabla u_0(x-y)\mathrm{d}y\\
&=I_k+II_k.
\end{aligned}$$

Note that

$$I_k+P_k u_1=\int_{R^2}4^k\check{\Psi}(2^k y)(\partial_r u_0+u_1)(x-y)\mathrm{d}y.$$

Thus

$$\|I_k+P_k u_1\|_{L^2}\lesssim\|P_k(\partial_r u_0+u_1)\|_{L^2}. \tag{3.19}$$

Note also that, for $x\in B_{1+\lambda}\backslash B_{1-\lambda}$,

$$\left|\nabla\frac{x}{|x|}\right|\lesssim 1,$$

thus,

$$\begin{aligned}
|II_k|&\leq\int_{R^2}4^k|\check{\Psi}|(2^k y)\left|\frac{x}{|x|}-\frac{x-y}{|x-y|}\right|\cdot|\nabla u_0(x-y)|\mathrm{d}y\\
&\leq\int_{|y|<2^{-\frac{k}{2}}}+\int_{|y|>2^{-\frac{k}{2}}}\\
&\lesssim\int_{|y|<2^{-\frac{k}{2}}}2^{-\frac{k}{2}}4^k|\check{\Psi}|(2^k y)\cdot|\nabla u_0(x-y)|\mathrm{d}y\\
&\qquad+\int_{|y|>2^{-\frac{k}{2}}}4^k\left|2^k y\right|^{-M}|\nabla u_0(x-y)|\mathrm{d}y.
\end{aligned}$$

Then simple computation shows that

$$\begin{aligned}
&\|\Pi_k\|_{L^2(B_{1+\lambda}\setminus B_{1-\lambda})}\\
&\lesssim \int_{|y|<2^{-\frac{k}{2}}} 2^{-\frac{k}{2}}\, 4^k|\check{\Psi}|(2^k y)\cdot \|\nabla u_0\|_{L^2}\,dy + \int_{|y|>2^{-\frac{k}{2}}} 4^k \left|2^k y\right|^{-M} \|\nabla u_0\|_{L^2}\,dy\\
&\lesssim 2^{-\frac{k}{2}} \|(u_0, u_1)\|_{\dot H^1\times L^2}.
\end{aligned} \tag{3.20}$$

Thus combining (3.17)–(3.20), we get that

$$\|\partial_r P_k u_0 + P_k u_1\|_{L^2(B_{1+\lambda}\setminus B_{1-\lambda})} \lesssim 2^{-\frac{k}{2}} \|(u_0, u_1)\|_{\dot H^1\times L^2} + \|P_k(\partial_r u_0 + u_1)\|_{L^2}. \tag{3.21}$$

The bound

$$\|\not\partial P_k u_0\|_{L^2(B_{1+\lambda}\setminus B_{1-\lambda})} \lesssim 2^{-\frac{k}{2}} \|(u_0, u_1)\|_{\dot H^1\times L^2} + \|P_k \not\partial u_0\|_{L^2} \tag{3.22}$$

follows similarly from the previous arguments.

Substep (3): Summary of estimates from substep (1) and substep (2) and the definition of good frequencies.

From (3.14),(3.17),(3.18),(3.21), and (3.22), we have, for $\delta^{\frac{1}{6}} \sim 2^{-k_0}$

$$(1)\ \|P_{<k_0}(\nabla u_0, u_1)\|_{L^2\times L^2} \lesssim \delta^{\frac{1}{3}} \|(u_0, u_1)\|_{\dot H^1\times L^2}; \tag{3.23}$$

$$(2)\ \sum_{k\geq k_0} \left\| P_k \left[(\nabla u_0, u_1)\chi_{B^c_{1+\delta}\cup B_{1-\delta}}\right]\right\|^2_{L^2\times L^2} \lesssim \delta^2 \|(u_0, u_1)\|^2_{\dot H^1\times L^2}; \tag{3.24}$$

$$\begin{aligned}
&(3)\ \sum_{k\geq k_0} \left\| P_k \left[(\nabla u_0, u_1)\chi_{B_{1+\delta}\setminus B_{1-\delta}}\right]\right\|_{L^2\times L^2(B^c_{1+\lambda}\cup B_{1-\lambda})}\\
&\quad\lesssim \delta^{\frac{1}{4}} \|(u_0, u_1)\|_{\dot H^1\times L^2};
\end{aligned} \tag{3.25}$$

$$\begin{aligned}
&(4) \|\partial_r P_k u_0 + P_k u_1\|_{L^2(B_{1+\lambda}\setminus B_{1-\lambda})} + \|\not\partial P_k u_0\|_{L^2(B_{1+\lambda}\setminus B_{1-\lambda})}\\
&\quad\lesssim 2^{-\frac{k}{2}} \|(u_0, u_1)\|_{\dot H^1\times L^2} + \|P_k(\partial_r u_0 + u_1)\|_{L^2} + \|P_k \not\partial u_0\|_{L^2}.
\end{aligned} \tag{3.26}$$

By (3.23), we can focus on the high frequencies $2^k \geq 2^{k_0}$. Indeed, we have

$$\left\|P_{k\geq k_0}(\nabla u_0, u_1)\right\|_{L^2\times L^2} \geq \left(1 - C\delta^{\frac{1}{4}}\right) \|(u_0, u_1)\|_{\dot H^1\times L^2}. \tag{3.27}$$

By (3.24) and (3.25), we see that

$$\sum_{k\geq k_0} \|P_k(\nabla u_0, u_1)\|^2_{L^2\times L^2\left(B^c_{1+\lambda}\cup B_{1-\lambda}\right)}$$

$$+ P_k\left[(\nabla u_0, u_1)\chi_{B_{1+\delta}\setminus B_{1-\delta}}\right]\Big\|^2_{L^2\times L^2(B^c_{1+\lambda}\cup B_{1-\lambda})}$$

$$\lesssim \sum_{k\geq k_0}\left\|P_k\left[(\nabla u_0, u_1)\chi_{B^c_{1+\delta}\cup B_{1-\delta}}\right]\right\|^2_{L^2\times L^2}$$

$$+ \sum_{k\geq k_0}\left\|P_k\left[(\nabla u_0, u_1)\chi_{B_{1+\delta}\setminus B_{1-\delta}}\right]\right\|^2_{L^2\times L^2\left(B^c_{1+\lambda}\cup B_{1-\lambda}\right)}$$

$$\lesssim \delta^{\frac{1}{2}}\|(u_0, u_1)\|^2_{\dot{H}^1\times L^2}.$$

Then by the above calculation and (3.26), we get that

$$\sum_{k\geq k_0}\left[\|\partial_r P_k u_0 + P_k u_1\|^2_{L^2} + \|\not\partial P_k u_0\|^2_{L^2}\right]$$

$$\lesssim \sum_{k\geq k_0}\|P_k(\nabla u_0, u_1)\|^2_{L^2\times L^2\left(B^c_{1+\lambda}\cup B_{1-\lambda}\right)} +$$

$$+ \sum_{k\geq k_0}\left[\|\partial_r P_k u_0 + P_k u_1\|^2_{L^2(B_{1+\lambda}\setminus B_{1-\lambda})} + \|\not\partial P_k u_0\|^2_{L^2(B_{1+\lambda}\setminus B_{1-\lambda})}\right]$$

$$\lesssim \delta^{\frac{1}{2}}\|(u_0, u_1)\|^2_{\dot{H}^1\times L^2} + \sum_{k\geq k_0}2^{-k}\|(u_0, u_1)\|^2_{\dot{H}^1\times L^2} +$$

$$+ \sum_{k\geq k_0}\left(\|P_k(\partial_r u_0 + u_1)\|^2_{L^2} + \|P_k\not\partial u_0\|^2_{L^2}\right)$$

$$\lesssim \delta^{\frac{1}{6}}\|(u_0, u_1)\|^2_{\dot{H}^1\times L^2}.$$

Hence, if we define the set

$$\mathcal{K} := \Big\{k\geq k_0 :\ \|(P_k u_0, P_k u_1)\|_{\dot{H}^1\times L^2\left(B^c_{1+\lambda}\cup B_{1-\lambda}\right)} + \|\partial_r P_k u_0 + P_k u_1\|_{L^2}$$

$$+ \|\not\partial P_k u_0\|_{L^2} \leq \delta^{\frac{1}{100}}\|P_k(u_0, u_1)\|_{\dot{H}^1\times L^2}\Big\},$$

we can estimate that

$$\sum_{k\geq k_0,\, k\notin\mathcal{K}}\|(P_k u_0, P_k u_1)\|^2_{\dot{H}^1\times L^2}$$

$$\lesssim \delta^{-\frac{1}{50}}\sum_{k\geq k_0,\, k\notin\mathcal{K}}\Big[\|(P_k u_0, P_k u_1)\|^2_{\dot{H}^1\times L^2\left(B^c_{1+\lambda}\cup B_{1-\lambda}\right)}$$

$$+ \|\partial_r P_k u_0 + P_k u_1\|^2_{L^2} + \|\not\partial P_k u_0\|^2_{L^2}\Big]$$

$$\lesssim \delta^{-\frac{1}{50}}\delta^{\frac{1}{6}}\|(u_0, u_1)\|^2_{\dot{H}^1\times L^2} \lesssim \delta^{\frac{1}{12}}\|(u_0, u_1)\|^2_{\dot{H}^1\times L^2}.$$

Hence the total energy at frequencies $\sim 2^k$ with $k \geq k_0$, $k \notin \mathcal{K}$ is negligible, and we will focus on the high frequency pieces $P_k(u_0, u_1)$ with $2^k \geq 2^{k_0}$ and $k \in \mathcal{K}$ below.

Substep (4): Reduction to channel of energy inequality for frequencies in $\mathcal{K}$.

Fix $m \in \mathcal{K}$, then

$$\|(P_m u_0, P_m u_1)\|_{\dot{H}^1 \times L^2\left(B^c_{1+\lambda} \cup B_{1-\lambda}\right)} + \|\partial_r P_m u_0 + P_m u_1\|_{L^2} + \|\not{\partial} P_m u_0\|_{L^2}$$
$$\leq \delta^{\frac{1}{100}} \|P_m(u_0, u_1)\|_{\dot{H}^1 \times L^2}. \tag{3.28}$$

We claim that if we can show for each $m \in \mathcal{K}$ that

$$\int_{|x| \geq \frac{1+\beta}{2}+t} |\nabla_{x,t} P_m u|^2(x,t)\, dx \geq \frac{1+\beta}{2} \|P_m(u_0, u_1)\|^2_{\dot{H}^1 \times L^2} - C\epsilon^2 c_m^2 \tag{3.29}$$

for all $t \geq 0$, then we will be done. Indeed, write for each $t \geq 0$,

$$P_m \nabla_{x,t} u = P_m \left[(\nabla_{x,t} u)\chi_{|x| > \beta+t}\right] + P_m \left[(\nabla_{x,t} u)\chi_{|x| \leq \beta+t}\right].$$

We can estimate, for $|x| > \frac{\beta+1}{2} + t$, that

$$\begin{aligned}
&\left|P_m \left[(\nabla_{x,t} u)\chi_{|y| \leq \beta+t}\right](x)\right| \\
&\quad \leq 4^m \int_{|x-y| \leq \beta+t} \left|\check{\Psi}(2^m y)\right|\left|\nabla_{x,t} u(x-y,t)\right| \mathrm{d}y \\
&\quad \leq 4^m \int_{|y| > \frac{1-\beta}{2}} \left|\check{\Psi}(2^m y)\right||\nabla_{x,t} u(x-y,t)| \mathrm{d}y \\
&\quad \leq 4^m \int_{|y| > \frac{1-\beta}{2}} \left|2^m y\right|^{-M} |\nabla_{x,t} u(x-y,t)| \mathrm{d}y.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\left\|P_m \left[|\nabla_{x,t} u|\chi_{|x| \leq \beta+t}\right]\right\|_{L^2\left(|x| > \frac{\beta+1}{2}+t\right)} \\
&\quad \lesssim 4^m \int_{|y| > \frac{1-\beta}{2}} |2^m y|^{-M}\, \mathrm{d}y\, \|(u_0, u_1)\|_{\dot{H}^1 \times L^2} \\
&\quad \lesssim C(\beta)\, 2^{-(M-2)m} \|(u_0, u_1)\|_{\dot{H}^1 \times L^2}.
\end{aligned}$$

Consequently, we get that

$$
\begin{aligned}
&\sum_{m\geq k_0}\|P_m\nabla_{x,t}u\|^2_{L^2\left(|x|>\frac{1+\beta}{2}+t\right)}\\
&\quad\leq \left(1+\delta^{\frac{1}{50}}\right)\sum_{m\geq k_0}\left\|P_m\left[(\nabla_{x,t}u)\chi_{|y|\geq\beta+t}\right]\right\|^2_{L^2}+\\
&\qquad+\ 2\delta^{-\frac{1}{50}}\sum_{m\geq k_0}\left\|P_m\left[(\nabla_{x,t}u)\chi_{|y|\leq\beta+t}\right]\right\|^2_{L^2\left(|x|>\frac{\beta+1}{2}+t\right)}\\
&\quad\leq \left(1+\delta^{\frac{1}{50}}\right)\int_{|x|>\beta+t}|\nabla_{x,t}u|^2(x,t)\mathrm{d}x+\\
&\qquad+\ C(\beta)\delta^{-\frac{1}{50}}\sum_{m\geq k_0}2^{-2(M-2)m}\|(u_0,\,u_1)\|^2_{\dot H^1\times L^2}\\
&\quad\leq \left(1+\delta^{\frac{1}{50}}\right)\int_{|x|>\beta+t}|\nabla_{x,t}u|^2(x,t)\mathrm{d}x+C(\beta)\delta^{-\frac{1}{50}}4^{-(M-2)k_0}\|(u_0,\,u_1)\|^2_{\dot H^1\times L^2}\\
&\quad\leq \left(1+\delta^{\frac{1}{50}}\right)\int_{|x|>\beta+t}|\nabla_{x,t}u|^2(x,t)\mathrm{d}x+C(\beta)\delta^{\frac{1}{3}}\|(u_0,\,u_1)\|^2_{\dot H^1\times L^2},
\end{aligned}
$$

if we choose $M=4$. Therefore if (3.29) holds, then by the choice of $\mathcal{K}$, (3.27), $\|(c_k)\|_{l^2}\lesssim\epsilon$, and the above calculation, we see that

$$
\begin{aligned}
&(1-C\delta^{\frac{1}{12}}-C\epsilon^2)\frac{1+\beta}{2}\|(u_0,\,u_1)\|^2_{\dot H^1\times L^2}\\
&\quad\leq\sum_{m\in\mathcal{K}}\left(\frac{1+\beta}{2}\|P_m(u_0,\,u_1)\|^2_{\dot H^1\times L^2}-C^2\epsilon^2c_m^2\right)\\
&\quad\leq\sum_{m\in\mathcal{K}}\|P_m\nabla_{x,t}u\|^2_{L^2\left(|x|>\frac{1+\beta}{2}+t\right)}\\
&\quad\leq\left(1+\delta^{\frac{1}{50}}\right)\int_{|x|>\beta+t}|\nabla_{x,t}u|^2(x,t)\mathrm{d}x+C(\beta)\delta^{\frac{1}{3}}\|(u_0,\,u_1)\|^2_{\dot H^1\times L^2}.
\end{aligned}
$$

The channel of energy inequality (3.8) follows if $\delta=\delta(\beta)$ and $\epsilon=\epsilon_0(\beta)$ are taken sufficiently small. Our goal is thus reduced to proving (3.29).

Step 2: *Control of the perturbative part of the nonlinearity.* It is proved in [14] that (3.29) holds for solution to the linear wave equation with this type of outgoing initial data for dimension ≥ 3, although the results we need here are more quantitative, see Lemma 3.1 above. Ideally one would like to say that the nonlinearity is negligible as we have small solutions. However, as is now well known, even in small energy case, the nonlinearity for the wave map equation cannot be treated entirely perturbatively. Rather, we need to perform a gauge transform to modify the nonlinearity so that it becomes perturbative.

Thus it is important to understand how the Gauge transform affects the channel of energy inequality. The arguments we use here are mostly from Tao [51] and Tataru [58]. We shall present the details of the proof below, partly for the convenience of the reader, and partly as those works did not explicitly quantify the nonlinear effects (which are implicit in the proofs). In this step however, we shall firstly control the part of the nonlinearity that is perturbative.

Let

$$\psi := P_m u.$$

Then ψ verifies

$$\begin{cases} \partial_{tt}\psi - \Delta\psi & = & P_m\left(u\,\partial^\alpha u^\dagger \partial_\alpha u\right) \\ \overrightarrow{\psi}(0) & = & (P_m u_0,\, P_m u_1). \end{cases} \tag{3.30}$$

Let us rewrite the nonlinearity $P_m\left(u\,\partial^\alpha u^\dagger \partial_\alpha u\right)$ as

$$\begin{aligned}
&P_m\left(u\,\partial^\alpha u^\dagger \partial_\alpha u\right) \\
&\quad = P_m\left(u_{\geq m-10}\,\partial^\alpha u^\dagger \partial_\alpha u\right) \\
&\qquad + P_m\left(u_{<m-10}\,\partial^\alpha u^\dagger_{>m+10} \partial_\alpha u\right) \\
&\qquad + P_m\left(u_{<m-10}\,\partial^\alpha u^\dagger_{m-10\leq\cdot\leq m+10} \partial_\alpha u_{\geq m-10}\right) \\
&\qquad + P_m\left(u_{<m-10}\,\partial^\alpha u^\dagger_{m-10\leq\cdot\leq m+10} \partial_\alpha u_{<m-10}\right) \\
&\qquad + P_m\left(u_{<m-10}\,\partial^\alpha u^\dagger_{<m-10} \partial_\alpha u_{>m+10}\right) \\
&\qquad + P_m\left(u_{<m-10}\,\partial^\alpha u^\dagger_{<m-10} \partial_\alpha u_{m-10\leq\cdot\leq m+10}\right) \\
&\qquad + P_m\left(u_{<m-10}\,\partial^\alpha u^\dagger_{<m-10} \partial_\alpha u_{<m-10}\right) \\
&\quad = I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7.
\end{aligned}$$

Denote

$$\epsilon := \|(u_0,\, u_1)\|_{\dot H^1\times L^2} \leq \epsilon_0.$$

We firstly peel off the perturbative part of the nonlinearity. We shall call h *disposable* if

$$\sup_{m'=m+O(1)} \|P_{m'} h\|_{N[m']} \lesssim \epsilon\, c_m.$$

Here the $O(1)$ term is a number of size ~ 10. The main use of this term is to deal with some technical "frequency leakage" issues. (On a technical level, to apply the estimates from Theorem 2.2, we need the right-hand sides to carry the frequency localization operator P_k.) We shall call h *disposable in the generalized sense* if there exists a sequence of disposable h_k with $h_k \to h$ in the sense of distributions. Note that the notion of being disposable and that of being disposable in the generalized sense are not the same, due to the technical issue with the space $N[m]$, see for example, page 324 of [58] for more discussions.

Note that $I_4 = I_6$. Furthermore, analysing the support of the trilinear expressions in frequencies, we obtain that $I_5 = I_7 = 0$. We claim that I_1, I_2, I_3 are disposable, that is,

Claim 3.1. For $j = 1, 2, 3$ we have

$$\sup_{m'=m+O(1)} \left\|P_{m'} I_j\right\|_{N[m']} \lesssim \epsilon c_m. \tag{3.31}$$

□

We also claim that

Claim 3.2. For $m' = m + O(1)$,

$$\left\|P_{m'}\left[P_m\left(u_{<m-10}\, \partial^\alpha u^\dagger_{m-10<\cdot<m+10} \partial_\alpha u_{<m-10}\right) - u_{<m-10}\, \partial^\alpha \psi^\dagger \partial_\alpha u_{<m-10}\right]\right\|_{N[m']} \lesssim \epsilon c_m, \tag{3.32}$$

where ψ is defined in (3.30); similarly,

$$\left\|P_{m'}\left[P_m\left(\partial_\alpha u_{<m-10}\, u^\dagger_{<m-10}\, \partial^\alpha u_{m-10<\cdot<m+10}\right) - \partial_\alpha u_{<m-10}\, u^\dagger_{<m-10}\, \partial^\alpha \psi\right]\right\|_{N[m']} \lesssim \epsilon c_m. \tag{3.33}$$

□

We postpone the proof of Claims 3.1 and 3.2 to the end of this section.

Hence, by (3.32) we can rewrite the equation for ψ as

$$\partial_{tt}\psi - \Delta\psi = \widetilde{f} + 2u_{<m-10}\, \partial_\alpha u^\dagger_{<m-10} \partial^\alpha \psi, \tag{3.34}$$

where

$$\sup_{m'=m+O(1)} \left\|P_{m'}\widetilde{f}\right\|_{N[m']} \lesssim \epsilon c_m.$$

Let us note the relation

$$u^\dagger\, \partial_\alpha u = 0.$$

It follows that

$$\begin{aligned}0 &= P_m\left(\partial^\alpha u_{<m-10} u^\dagger\, \partial_\alpha u\right)\\ &= P_m\left(\partial^\alpha u_{<m-10} u^\dagger_{\geq m-10} \partial_\alpha u\right) + P_m\left(\partial^\alpha u_{<m-10} u^\dagger_{<m-10} \partial_\alpha u_{m-10\leq\cdot\leq m+10}\right)\\ &= I + II.\end{aligned}$$

We can estimate the I term, by using the trilinear estimate, as for $m' = m + O(1)$

$$\begin{aligned}&\left\|P_{m'}\left(\partial^\alpha u_{<m-10} u^\dagger_{\geq m-10} \partial_\alpha u\right)\right\|_{N[m']}\\ &\quad\lesssim \sum_{k_1<m-10,\, m-10\leq k_2\leq m+10,\, k_3\leq m+O(1)} \left\|P_{m'}\left(\partial^\alpha u_{k_1} u^\dagger_{k_2}\, \partial_\alpha u_{k_3}\right)\right\|_{N[m']} +\\ &\qquad + \sum_{k_1<m-10,\, k_2>m+10,\, k_3=k_2+O(1)} \left\|P_{m'}\left(\partial^\alpha u_{k_1} u^\dagger_{k_2}\, \partial_\alpha u_{k_3}\right)\right\|_{N[m']}\\ &\quad\lesssim \sum_{k_1<m-10,\, k_3\leq m+O(1)} 2^{-\kappa(m-\min\{k_1,k_3\})} c_{k_1} c_m c_{k_3} +\\ &\qquad + \sum_{k_1<m-10,\, k_2>m+10} 2^{-\kappa(k_2-m)} 2^{-\kappa(k_2-k_1)} c_{k_1} c_{k_2}^2\\ &\quad\lesssim \epsilon^2 c_m.\end{aligned}$$

Consequently, by the boundedness of P_m in $S[m']$, we see that

$$\sup_{m'=m+O(1)} \|P_{m'} II\|_{N[m']} \lesssim \epsilon^2 c_m$$

and thus II is disposable. Thus by (3.33) we can rewrite the equation for ψ as

$$\partial_{tt}\psi - \Delta\psi = f + 2\left(u_{<m-10}\, \partial^\alpha u^\dagger_{<m-10} - \partial^\alpha u_{<m-10}\, u^\dagger_{<m-10}\right) \partial_\alpha \psi, \tag{3.35}$$

where f satisfies

$$\sup_{m'=m+O(1)} \|f\|_{N[m']} \lesssim \epsilon c_m. \tag{3.36}$$

Step 3 : *Construction of the micro-local gauge.* To deal with the non-perturbative part of the nonlinearity, we need to use the idea of Tao [51].

We have

$$\partial_{tt}\psi - \Delta\psi = f + 2\left(u_{<m-10}\, \partial^\alpha u^\dagger_{<m-10} - \partial^\alpha u_{<m-10}\, u^\dagger_{<m-10}\right) \partial_\alpha \psi, \tag{3.37}$$

where f satisfies

$$\sup_{m'=m+O(1)} \|f\|_{N[m']} \lesssim \epsilon c_m. \tag{3.38}$$

Let $w = U_{<m-10}\psi$ for some matrix $U_{<m-10}$ to be determined below, then (3.37) implies that

$$\begin{aligned}
-\partial^\alpha \partial_\alpha w &= -\partial^\alpha \partial_\alpha U_{<m-10}\, \psi - 2\partial^\alpha U_{<m-10} \partial_\alpha \psi \\
&\quad + U_{<m-10} \left[f + 2\left(u_{<m-10}\, \partial^\alpha u^\dagger_{<m-10} - \partial^\alpha u_{<m-10}\, u^\dagger_{<m-10} \right) \partial_\alpha \psi \right] \\
&= (\Box U_{<m-10})\ \psi + U_{<m-10} f + \\
&\quad + 2\left[U_{<m-10} \left(u_{<m-10}\, \partial^\alpha u^\dagger_{<m-10} - \partial^\alpha u_{<m-10}\, u^\dagger_{<m-10} \right) - \partial^\alpha U_{<m-10} \right] \partial_\alpha \psi.
\end{aligned}$$

Then

$$\begin{aligned}
\partial_{tt} w - \Delta w &= (\Box U_{<m-10})\, \psi + U_{<m-10} f \\
&\quad + 2\left[U_{<m-10} \left(u_{<m-10}\, \partial^\alpha u^\dagger_{<m-10} - \partial^\alpha u_{<m-10}\, u^\dagger_{<m-10} \right) - \partial^\alpha U_{<m-10} \right] \partial_\alpha \psi.
\end{aligned} \tag{3.39}$$

Ideally we would like to choose $U_{<m-10}$ so that

$$\partial^\alpha U_{<m-10} = U_{<m-10} \left(u_{<m-10} \partial^\alpha u^\dagger_{<m-10} - \partial^\alpha u_{<m-10}\, u^\dagger_{<m-10} \right)$$

for all α, then the term on the right-hand side of (3.39) containing $\partial_\alpha \psi$ would be eliminated, and we would be in a truly semilinear case. However, this is impossible due to compatibility issues, see the discussions in [51]. Instead we will follow Tataru's modification of Tao's idea in [58] to construct a micro-local approximate solution.

Fix large $N > 1$. Define inductively

$$\begin{aligned}
U^N_{-N} &= I; \\
U^N_k &= U^N_{<k-10} \left(u_{<k-10} u^\dagger_k - u_k u^\dagger_{<k-10} \right),
\end{aligned}$$

where $U^N_{<k-10} = \sum_{-N<j<k-10} U^N_j + I$ if $k > -N + 11$ and $U^N_{<k-10} = I$ otherwise. In the end we will pass $N \to \infty$, but we need to obtain uniform in N estimates for U^N_k in order to do that. We claim the following properties for U^N_k and $U^N_{<k}$ with $-N < k \le N$:

Claim 3.3. For ϵ sufficiently small,

$$U^N_k \text{ has frequency support } 2^{k-2} \le |\xi| \le 2^{k+2}; \tag{3.40}$$

$$\sup_{k'=k+O(1)} \left\| P_{k'} U_k^N \right\|_{S[k']} \lesssim c_k; \tag{3.41}$$

$$\left\| U_{<k}^N \left(U_{<k}^N\right)^\dagger - I \right\|_{S(c)} \lesssim \sqrt{\epsilon}. \tag{3.42}$$

□

We shall prove the claim inductively. For $k = -N+1$, the claim follows from the property that by Theorem 2.3

$$\|u\|_{S(c)} \lesssim 1.$$

Suppose the claim is true up to $k-1$, let us prove it holds also for k. A crucial point is the following important algebraic identity:

$$U_k^N \left(U_{<k-10}^N\right)^\dagger + U_{<k-10}^N (U_k^N)^\dagger \tag{3.43}$$

$$= U_{<k-10}^N \left(u_{<k-10} u_k^\dagger - u_k u_{<k-10}^\dagger\right) \left(U_{<k-10}^N\right)^\dagger + \tag{3.44}$$

$$+ \; U_{<k-10}^N \left(u_k u_{<k-10}^\dagger - u_{<k-10} u_k^\dagger\right) \left(U_{<k-10}^N\right)^\dagger \tag{3.45}$$

$$= 0. \tag{3.46}$$

We also note that U_j^N is anti-symmetric if $-N < j \le -N+11$, which is an easy consequence of the definition of U_j^N.

Thus by the anti-symmetry of U_j^N for $-N < j \le -N+11$, we get that

$$\begin{aligned}
U_{<k}^N \left(U_{<k}^N\right)^\dagger &= \left(\sum_{-N\le j<k} U_j^N\right)\left(\sum_{-N\le j<k} \left(U_j^N\right)^\dagger\right) \\
&= \sum_{-N\le j<j'-10<j'<k} U_j^N \left(U_{j'}^N\right)^\dagger + \sum_{-N\le j'<j-10<j<k} U_j^N \left(U_{j'}^N\right)^\dagger + \\
&\quad + \sum_{|j-j'|\le 10,\, -N<j,j'<k} U_j^N \left(U_{j'}^N\right)^\dagger + \sum_{-N<j\le -N+10} \left[\left(U_j^N\right)^\dagger + U_j^N\right] + I \\
&= \sum_{-N+10<j'<k} U_{<j'-10}^N \left(U_{j'}^N\right)^\dagger + \sum_{-N+10<j<k} U_j^N \left(U_{<j-10}^N\right)^\dagger + I +
\end{aligned}$$

$$+ \sum_{|j-j'|\le 10,\, -N<j,j'<k} U_j^N \left(U_{j'}^N\right)^\dagger$$

$$= I + \sum_{-N<j,j'<k,\, |j-j'|\le 10} U_j^N \left(U_{j'}^N\right)^\dagger .$$

Simplifying the above, we get that

$$U_{<k}^N \left(U_{<k}^N\right)^\dagger - I = \sum_{-N<j,j'<k,\, |j-j'|\le 10} U_j^N \left(U_{j'}^N\right)^\dagger .$$

Hence by (3.41) from induction,

$$\begin{aligned}
&\left\| U_{<k}^N \left(U_{<k}^N\right)^\dagger - I \right\|_{L^\infty} \\
&\lesssim \sum_{-N<j,j'<k,\, |j-j'|\le 10} \left\| U_j^N \right\|_{L^\infty} \left\| U_{j'}^N \right\|_{L^\infty} \\
&\lesssim \sum_{-N<j,j'<k,\, |j-j'|\le 10} \ \sum_{j_1=j+O(1),\, j_2=j'+O(1)} \left\| P_{j_1} U_j^N \right\|_{S[j_1]} \left\| P_{j_2} U_{j'}^N \right\|_{S[j_2]} \\
&\lesssim \sum_{-N<j<k} c_j^2 \lesssim \epsilon^2 .
\end{aligned}$$

In the second inequality above, we used the fact that $U_j^N = \sum_{j_1=j+O(1)} P_{j_1} U_j^N$ which follows from the frequency support property of U_j^N. We shall use this trick often, as a replacement of bound on $\|U_j^N\|_{S[j]}$ which we do not have. Below we will omit the routine details when we use the same trick. In particular, combining the above with the induction bound (3.41), we see that $\|U_{<k}^N\|_{S(1)} \le C$ for some universal constant (by choosing ϵ small).

Similarly, for each $k' < k + O(1)$, by the property of $S[k]$ spaces and induction,

$$\begin{aligned}
&\left\| P_{k'} \left[U_{<k}^N \left(U_{<k}^N\right)^\dagger \right] \right\|_{S[k']} \\
&\lesssim \sum_{-N<j,j'<k,\, |j-j'|\le 10} \left\| P_{k'} \left[U_j^N \left(U_{j'}^N\right)^\dagger \right] \right\|_{S[k']} \\
&\lesssim \sum_{O(1)+k'<j<k} 2^{-\kappa(j-k')_+} c_j^2 \\
&\lesssim \sum_{O(1)+k'<j<k} 2^{-(\kappa-\vartheta)(j-k')_+} c_j c_{k'} \lesssim \epsilon\, c_{k'} .
\end{aligned}$$

Combining the above two estimates, (3.42) follows.

The estimate for $\sup_{k'=k+O(1)} \left\|P_{k'} U_k^N\right\|_{S[k']}$ then follows from the definition and the fact that $\|u_{<k-10}\|_{S(c)}$, $\|U_{<k-10}^N\|_{S(1)}$ are universally bounded. The support property is obvious.

Using these uniform estimates, we can pass $N \to \infty$, and obtain a limit along a subsequence of N, so that

$$U_k := \lim_{N_i \to \infty} U_k^{N_i}, \quad U_{<k} := \lim_{N_i \to \infty} U_{<k}^{N_i}$$

exist in the sense of distributions, for each k. Since U_k^N are frequency localized and have bounded overlap in frequency support, we can conclude that

$$U_{<k} = \sum_{k'<k} U_{k'} + I, \text{ and } U_k = U_{<k-10}(u_{<k-10} u_k^{\dagger} - u_k u_{<k-10}^{\dagger}). \tag{3.47}$$

In addition, U_k, $U_{<k}$ satisfies the same estimates claimed for U_k^N, $U_{<k}^N$ above. As a consequence, we have

$$\sup_{k'=k+O(1)} \|P_{k'} U_k\|_{S[k']} \lesssim c_k, \quad \text{and} \quad \|U_{<k}\|_{S(c)} \lesssim 1. \tag{3.48}$$

This is a direct consequence of the property of $S[k]$ under weak convergence, see the remark below Theorem 2.2.

Step 4 : *Control of the nonlinearity after applying the gauge transform.* We shall show that the terms on the right-hand size of (3.39) are all disposable.

Substep (1): the terms involving $\Box U_k$.

To control the terms $(\Box U_{<m-10})\,\psi$, we need to control $\Box U_{<m-10}^N$ uniformly for all large N. By definition,

$$\begin{aligned}
\Box U_k^N &= \left(\Box U_{<k-10}^N\right)\left(u_{<k-10} u_k^{\dagger} - u_k u_{<k-10}^{\dagger}\right) \\
&\quad - 2\partial^{\alpha} U_{<k-10}^N \partial_{\alpha}\left(u_{<k-10} u_k^{\dagger} - u_k u_{<k-10}^{\dagger}\right) \\
&\quad + U_{<k-10}^N\left(\Box u_{<k-10} u_k^{\dagger} + u_{<k-10} \Box u_k^{\dagger} - \Box u_k\, u_{<k-10}^{\dagger} - u_k \Box u_{<k-10}^{\dagger}\right) \\
&\quad + 2U_{<k-10}^N\left(\partial^{\alpha} u_k \partial_{\alpha} u_{<k-10}^{\dagger} - \partial^{\alpha} u_{<k-10} \partial_{\alpha} u_k^{\dagger}\right) = I + II + III + IV.
\end{aligned}$$

We claim that for $\nu = \frac{\kappa}{32}$, and uniformly for all large N.

Claim 3.4.

$$\sup_{j'=j+O(1)} \left\| P_{j'}\left(\Box U^N_k \phi\right)\right\|_{N[j']} \lesssim 2^{-\nu(j-k)} c_k \|\phi\|_{S[j]} \tag{3.49}$$

for all ϕ with frequency support in $2^{j-5/2} \le \cdot \le 2^{j+5/2}$ and $k < j-7$. □

Assuming this claim for a moment, then we can estimate for $m' = m + O(1)$

$$\begin{aligned}
&\left\| P_{m'}\left[\Box U^N_{<m-10}\,\psi\right]\right\|_{N[m']} \\
&\quad\lesssim \sum_{k<m-10} \left\| P_{m'}\left[\Box U^N_k\,\psi\right]\right\|_{N[m']} \\
&\quad\lesssim \sum_{k<m-10} 2^{-\nu(m-k)} c_k c_m \lesssim \epsilon\, c_m,
\end{aligned}$$

and thus the first term on the right-hand side of (3.39) is disposable in the generalized sense.

We shall prove (3.49) inductively. It is clear that (3.49) holds for $k = -N$. Suppose (3.49) holds for $k' < k$, let us prove that it holds for k. The bound for I term:

$$\begin{aligned}
&\left\| P_{j'}\left[\Box U^N_{<k-10}\left(u_{<k-10}u_k^{\dagger} - u_k\, u^{\dagger}_{<k-10}\right)\phi\right]\right\|_{N[j']} \\
&\quad\lesssim \sum_{k'<k-10,\, |j-j''|\le 3} \left\| P_{j'}\left[\Box U^N_{k'} P_{j''}\left\{\left(u_{<k-10}u_k^{\dagger} - u_k\, u^{\dagger}_{<k-10}\right)\phi\right\}\right]\right\|_{N[j']} \\
&\quad\lesssim \sum_{k'<k-10} 2^{-\nu(j-k')} c_{k'} c_k \|\phi\|_{S[j]} \\
&\quad\lesssim 2^{-\nu(j-k)}\epsilon\, c_k \|\phi\|_{S[j]}
\end{aligned}$$

follows from the inductive hypothesis and the property of $S[k]$ spaces. The projection $P_{j''}$ was used to deal with the frequency leakage, which is a minor technical issue.

Let us consider the II term $\partial^{\alpha} U^N_{<k-10}\partial_{\alpha}\left(u_{<k-10}u_k^{\dagger} - u_k u^{\dagger}_{<k-10}\right)\phi$. By (3.48) and the trilinear estimate, we have

$$\begin{aligned}
&\left\| P_{j'}\left[\partial^{\alpha} U^N_{<k-10}\partial_{\alpha}\left(u_{<k-10}u_k^{\dagger} - u_k u^{\dagger}_{<k-10}\right)\phi\right]\right\|_{N[j']} \\
&\quad\lesssim \sum_{k'<k-10} \left\| P_{j'}\left[\partial^{\alpha} U^N_{k'}\partial_{\alpha}\left(u_{<k-10}u_k^{\dagger} - u_k u^{\dagger}_{<k-10}\right)\phi\right]\right\|_{N[j']} \\
&\quad\lesssim \sum_{k'<k-10} 2^{-\kappa(j-k')} c_{k'} c_k \|\phi\|_{S[j]} \\
&\quad\lesssim 2^{-\kappa(j-k)}\epsilon\, c_k \|\phi\|_{S[j]}.
\end{aligned}$$

Let us now consider the term

$$P_{j'}\left[\left(U^N_{<k-10}\partial^\alpha u_{<k-10}\partial_\alpha u_k^\dagger\right)\phi\right]$$

from term IV. We have, by (3.48) and the trilinear estimate,

$$\begin{aligned}&\left\|P_{j'}\left[\left(U^N_{<k-10}\partial^\alpha u_{<k-10}\partial_\alpha u_k^\dagger\right)\phi\right]\right\|_{N[j']}\\ &\quad\lesssim \sum_{k'<k-10} 2^{-\kappa(j-k')}c_k c_{k'}\|\phi\|_{S[j]}\lesssim 2^{-\kappa(j-k)}\epsilon\, c_k\|\phi\|_{S[j]}\end{aligned}$$

for $j'=j+O(1)$.

The term

$$P_{j'}\left[\left(U^N_{<k-10}\partial^\alpha u_k\partial_\alpha u^\dagger_{<k-10}\right)\phi\right]$$

can be controlled similarly.

It remains to control term III. For this, we need to use the equation for u. Since u satisfies the wave map equation, we see that

$$\Box u_{k'}=P_{k'}\left(u\,\partial^\alpha u^\dagger\partial_\alpha u\right),\ \text{for each } k'\le k. \tag{3.50}$$

It suffices to show that, for any φ with Fourier support $2^{j-3}\le|\xi|\le 2^{j+3}$ and $k'<j-6$,

$$\left\|P_{j'}\left[P_{k'}\left(u\,\partial^\alpha u^\dagger\partial_\alpha u\right)\varphi\right]\right\|_{N[j']}\lesssim 2^{-\nu(j-k')}\epsilon^{\frac12}c_{k'}\|\varphi\|_{S[j]} \tag{3.51}$$

for $j'=j+O(1)$. Indeed, from (3.51), it follows that

$$\begin{aligned}&\left\|P_{j'}\left[U^N_{<k-10}\Box u_{<k-10}u_k^\dagger\phi\right]\right\|_{N[j']}\\ &\quad\lesssim\sum_{k'<k-10}\left\|P_{j'}\left[U^N_{<k-10}\Box u_{k'}u_k^\dagger\phi\right]\right\|_{N[j']}\\ &\quad\lesssim\sum_{k'<k-10}2^{-\nu(j-k')}c_{k'}c_k\|\phi\|_{S[j]}\\ &\quad\lesssim\epsilon\, c_k 2^{-\nu(j-k)}\|\phi\|_{S[j]}\end{aligned}$$

and that

$$\begin{aligned}&\left\|P_{j'}\left[U^N_{<k-10}\Box u_k\, u^\dagger_{<k-10}\phi\right]\right\|_{N[j']}\\ &\quad\lesssim 2^{-\nu(j-k)}c_k\,\epsilon^{\frac12}\|\phi\|_{S[j]}.\end{aligned}$$

These estimates are sufficient for the completion of the induction, due to the presence of the extra $\epsilon^{\frac{1}{2}}$ factor, which can be used to absorb various constants in the inequalities.

To prove (3.51), let us decompose $P_{k'}\left(u\,\partial^\alpha u^\dagger \partial_\alpha u\right)\phi$ as

$$
\begin{aligned}
P_{k'}\left(u\,\partial^\alpha u^\dagger \partial_\alpha u\right)\varphi &= P_{k'}\left(u_{>\frac{j+k'}{2}}\,\partial^\alpha u^\dagger \partial_\alpha u\right)\varphi \\
&+ P_{k'}\left(u_{\leq\frac{j+k'}{2}}\,\partial^\alpha u^\dagger \partial_\alpha u\right)\varphi = I_1 + I_2.
\end{aligned}
$$

For I_1, by the trilinear estimates and symmetry, we can estimate as follows

$$
\begin{aligned}
&\left\|P_{j'}\left[P_{k'}\left(u_{>\frac{j+k'}{2}}\,\partial^\alpha u^\dagger \partial_\alpha u\right)\varphi\right]\right\|_{N[j']} = \left\|\sum_{k_2,k_3,k_1>\frac{j+k'}{2}} P_{j'}\left[P_{k'}\left(u_{k_1}\,\partial^\alpha u^\dagger_{k_2}\partial_\alpha u_{k_3}\right)\varphi\right]\right\|_{N[j']} \\
&\lesssim \|\varphi\|_{S[j]}\left(\sum_{k_1>\frac{j+k'}{2},\,k_3\geq k_1+O(1),\,k_3=k_2+O(1)} 2^{-\kappa(\max_{1\leq i\leq 3}k_i-k')}2^{-\kappa(k_1-\min\{k_2,k_3\})_+}c_{k_1}c_{k_2}c_{k_3}\right. \\
&\quad\left.+\sum_{k_1>\frac{j+k'}{2},\,k_2<k_3-C,\,k_3=k_1+O(1)} 2^{-\kappa(\max_{1\leq i\leq 3}k_i-k')}2^{-\kappa(k_1-\min\{k_2,k_3\})_+}c_{k_1}c_{k_2}c_{k_3}\right) \\
&\lesssim \epsilon^2\,\|\varphi\|_{S[j]}\left(\sum_{k_1>\frac{j+k'}{2}} c_{k_1}2^{-\kappa(k_1-k')} + \sum_{k_1>\frac{j+k'}{2}} c_{k_1}2^{-\kappa(k_1-k')}\right) \\
&\lesssim \epsilon^2\,2^{-\frac{\kappa}{2}(j-k')}\|\varphi\|_{S[j]}c_{k'}.
\end{aligned}
$$

Now let us deal with the term $I_2 = P_{k'}\left(u_{\leq\frac{j+k'}{2}}\,\partial^\alpha u^\dagger \partial_\alpha u\right)\varphi$. In this case, we can insert $P_{<\frac{j+k'}{2}+C}$ in front of $\partial^\alpha u^\dagger\,\partial_\alpha u$, use symmetry, and obtain that

$$
\begin{aligned}
&\left\|P_{j'}\left[P_{k'}\left(u_{\leq\frac{j+k'}{2}}\,\partial^\alpha u^\dagger \partial_\alpha u\right)\varphi\right]\right\|_{N[j']} \\
&= \left\|P_{j'}\left[P_{k'}\left(u_{\leq\frac{j+k'}{2}}\,P_{<\frac{j+k'}{2}+C}\left(\partial^\alpha u^\dagger \partial_\alpha u\right)\right)\varphi\right]\right\|_{N[j']} \\
&\lesssim \left\|\sum_{k_1\leq k_2,\,k_2=k_1+O(1)} P_{j'}\left[P_{k'}\left(u_{\leq\frac{j+k'}{2}}\,P_{<\frac{j+k'}{2}+C}\left(\partial^\alpha u^\dagger_{k_1}\,\partial_\alpha u_{k_2}\right)\right)\varphi\right]\right\|_{N[j']} + \\
&\quad+\left\|\sum_{k_1\leq k_2-C} P_{j'}\left[P_{k'}\left(u_{\leq\frac{j+k'}{2}}\,P_{<\frac{j+k'}{2}+C}\left(\partial^\alpha u^\dagger_{k_1}\,\partial_\alpha u_{k_2}\right)\right)\varphi\right]\right\|_{N[j']},
\end{aligned}
$$

which can be estimated as

$$
\begin{aligned}
&\lesssim \left\| \sum_{k_1\le k_2,\, k_2=k_1+O(1),\, k_1>\frac{3j+k'}{4}} P_{j'}\left[P_{k'}\left(u_{\le\frac{j+k'}{2}}\, P_{<\frac{j+k'}{2}+C}\left(\partial^\alpha u^\dagger_{k_1}\,\partial_\alpha u_{k_2}\right)\right)\varphi\right]\right\|_{N[j']}\\
&\quad+\left\| \sum_{k_1\le k_2,\, k_2=k_1+O(1),\, k_1\le\frac{3j+k'}{4}} P_{j'}\left[P_{k'}\left(u_{\le\frac{j+k'}{2}}\, P_{<\frac{j+k'}{2}+C}\left(\partial^\alpha u^\dagger_{k_1}\,\partial_\alpha u_{k_2}\right)\right)\varphi\right]\right\|_{N[j']}\\
&\quad+\sum_{k_1\le k_2-C,\, k_2\le\frac{j+k'}{2}+C} 2^{-\kappa(j-k_1)}\|\varphi\|_{S[j]}c_{k_1}c_{k_2}
\end{aligned}
$$

which is

$$
\begin{aligned}
&\lesssim \sum_{k_1>\frac{3j+k'}{4}} 2^{-\kappa\left(k_1-\frac{j+k'}{2}\right)}\|\varphi\|_{S[j]}\cdot c^2_{k_1}+\sum_{k_1\le\frac{3j+k'}{4}} c^2_{k_1}\cdot 2^{-\kappa(j-k_1)}\|\varphi\|_{S[j]}\\
&\qquad+\sum_{k_1\le k_2-C,\, k_2\le\frac{j+k'}{2}+C} 2^{-\kappa(j-k_1)}\|\varphi\|_{S[j]}c_{k_1}c_{k_2}\\
&\lesssim 2^{-\frac{\kappa}{8}(j-k')}\epsilon\, c_{k'}\|\varphi\|_{S[j]}.
\end{aligned}
$$

Combining the above estimates for I, II, III, IV terms, the claim follows.

Substep (2): Control of the term containing $\partial\psi$.

Now we address the main term in the nonlinearity that forced us to use the gauge transform

$$
\widetilde{h}=\left[U_{<m-10}\left(u_{<m-10}\partial^\alpha u^\dagger_{<m-10}-\partial^\alpha u_{<m-10}u^\dagger_{<m-10}\right)-\partial^\alpha U_{<m-10}\right]\partial_\alpha\psi.
$$

Note that by (3.47), we have

$$
\begin{aligned}
-\widetilde{h}&=\left[\partial^\alpha U_{<m-10}-U_{<m-10}\left(u_{<m-10}\partial^\alpha u^\dagger_{<m-10}-\partial^\alpha u_{<m-10}u^\dagger_{<m-10}\right)\right]\partial_\alpha\psi\\
&=\sum_{k<m-10}\left[\partial^\alpha U_k-U_{<m-10}\left(u_{<m-10}\partial^\alpha u^\dagger_k-\partial^\alpha u_k u^\dagger_{<m-10}\right)\right]\partial_\alpha\psi\\
&=\sum_{k<m-10}\left[\partial^\alpha U_k-U_{<k-10}\left(u_{<k-10}\partial^\alpha u^\dagger_k-\partial^\alpha u_k u^\dagger_{<k-10}\right)\right]\partial_\alpha\psi\\
&\quad-\sum_{k<m-10}U_{k-10\le\cdot<m-10}\left(u_{<m-10}\partial^\alpha u^\dagger_k-\partial^\alpha u_k\, u^\dagger_{<m-10}\right)\partial_\alpha\psi\\
&\quad-\sum_{k<m-10}U_{<k-10}\left(u_{k-10\le\cdot<m-10}\partial^\alpha u^\dagger_k-\partial^\alpha u_k\, u^\dagger_{k-10\le\cdot<m-10}\right)\partial_\alpha\psi
\end{aligned}
$$

$$= \sum_{k<m-10} \left[\partial^\alpha U_{<k-10}\left(u_{<k-10}u_k^\dagger - u_k u_{<k-10}^\dagger\right) \right.$$
$$\left. - U_{<k-10}\left(\partial^\alpha u_{<k-10}u_k^\dagger - u_k \partial^\alpha u_{<k-10}^\dagger\right)\right]\partial_\alpha\psi + \mathcal{R}.$$

To estimate the $\mathcal{R}$ term, let us firstly bound for $m' = m + O(1)$,

$$\begin{aligned}
&\left\|P_{m'}\left[U_{k-10\le\cdot<m-10}u_{<m-10}\partial^\alpha u_k^\dagger\partial_\alpha\psi\right]\right\|_{N[m']} \\
&\lesssim \sum_{k-10\le k'<m-10} \left\|P_{m'}\left[U_{k'}u_{<m-10}\partial^\alpha u_k^\dagger\partial_\alpha\psi\right]\right\|_{N[m']} \\
&\lesssim \sum_{k-10\le k'<m-10} \sup_{m''=m+O(1)} \left\|P_{m''}\left[U_{k'}\partial^\alpha u_k^\dagger\partial_\alpha\psi\right]\right\|_{N[m'']} \\
&\lesssim \sum_{k-10\le k'<m-10} 2^{-\kappa(k'-k)}c_{k'}c_k c_m \lesssim \sum_{k-10\le k'<m-10} 2^{-(\kappa-\vartheta)(k'-k)}c_k^2 c_m \\
&\lesssim c_k^2 c_m.
\end{aligned}$$

Other terms in $\mathcal{R}$ can be treated similarly. Thus

$$\sup_{m'=m+O(1)} \|P_{m'}[\mathcal{R}]\|_{N[m']} \lesssim \sum_k c_k^2 c_m \lesssim \epsilon^2 c_m, \tag{3.52}$$

and consequently $\mathcal{R}$ is disposable.

We can estimate for $m' = m + O(1)$

$$\begin{aligned}
&\left\|P_{m'}\left[\partial^\alpha U_{<k-10}\left(u_{<k-10}u_k^\dagger - u_k u_{<k-10}^\dagger\right)\partial_\alpha\psi\right]\right\|_{N[m']} \\
&\lesssim \sum_{k'<k-10} \left\|P_{m'}\left[\partial^\alpha U_{k'}\left(u_{<k-10}u_k^\dagger - u_k u_{<k-10}^\dagger\right)\partial_\alpha\psi\right]\right\|_{N[m']} \\
&\lesssim \sum_{k'<k-10} 2^{-\kappa(k-k')}c_k c_{k'}\|\psi\|_{S[m]} \\
&\lesssim \sum_{k'<k-10} 2^{-(\kappa-\vartheta)(k-k')}c_k^2\|\psi\|_{S[m]} \lesssim c_k^2\|\psi\|_{S[m]}
\end{aligned}$$

and

$$\begin{aligned}
&\left\|P_{m'}\left[U_{<k-10}\partial^\alpha u_{<k-10}u_k^\dagger\partial_\alpha\psi\right]\right\|_{N[m']} \\
&\lesssim \sum_{k'<k-10} \left\|P_{m'}\left[U_{<k-10}\partial^\alpha u_{k'}u_k^\dagger\partial_\alpha\psi\right]\right\|_{N[m']} \\
&\lesssim \sum_{k'<k-10} 2^{-\kappa(k-k')}c_{k'}c_k\|\psi\|_{S[m]} \\
&\lesssim c_k^2\|\psi\|_{S[m]}.
\end{aligned}$$

Thus in summary, we can estimate

$$\sup_{m'=m+O(1)} \|\widetilde{h}\|_{N[m']} \lesssim \sum_{k<m-10} c_k^2 \|\psi\|_{S[m]} \lesssim \epsilon^2 c_m$$

and consequently $\widetilde{h}$ is disposable.

Substep (3): $U_{<m-10}f$ term is disposable.

This follows directly as f is disposable.

Step 5 : *Proof of the channel of energy inequality for the good frequency piece.* Take $m \in \mathcal{K}$. By the estimates from *Step 4*, we can write the equation for w in Step 3 as

$$\partial_{tt} w - \Delta w = h \tag{3.53}$$

with h being disposable in the generalized sense, that is, $h = \lim_{k\to\infty} h_k$ in the sense of distributions and $\sup_{m'=m+O(1)} \|P_{m'} h_k\|_{N[m']} \lesssim \epsilon c_m$ uniformly in k. Let us now study how the outgoing condition (3.28) on the initial data of ψ has been transformed. Recall that

$$w = U_{<m-10}\psi.$$

Hence

$$\nabla_{x,t} w = \nabla_{x,t} U_{<m-10}\psi + U_{<m-10}\nabla_{x,t}\psi.$$

Thus at time $t = 0$, by (3.48) and the outgoing condition for ψ,

$$\begin{aligned}
&\|\nabla_{x,t} w(0)\|_{L^2_x(B^c_{1+\lambda}\cup B_{1-\lambda})} \\
&\quad\lesssim \|\nabla_{x,t} U_{<m-10}(0)\|_{L^2} \|\psi(0)\|_{L^\infty_x} + \|U_{<m-10}(0)\|_{L^\infty_x} \|\nabla_{x,t}\psi(0)\|_{L^2(B^c_{1+\lambda}\cup B_{1-\lambda})} \\
&\quad\lesssim \left(\sum_{k<m-10} \|\nabla_{x,t} U_k(0)\|^2_{L^2_x}\right)^{\frac{1}{2}} \|P_m(u_0, u_1)\|_{\dot H^1\times L^2} + \delta^{\frac{1}{100}} \|P_m(u_0, u_1)\|_{\dot H^1\times L^2} \\
&\quad\lesssim \left(\sum_{k<m-10} c_k^2\right)^{\frac{1}{2}} \|P_m(u_0, u_1)\|_{\dot H^1\times L^2} + \delta^{\frac{1}{100}} \|P_m(u_0, u_1)\|_{\dot H^1\times L^2} \\
&\quad\lesssim \left(\epsilon + \delta^{\frac{1}{100}}\right) \|P_m(u_0, u_1)\|_{\dot H^1\times L^2}.
\end{aligned}$$

Similar calculations show that

$$\|\not\partial w_0\|_{L^2} + \|\partial_r w_0 + w_1\|_{L^2} \lesssim \left(\epsilon + \delta^{\frac{1}{100}}\right) \|P_m(u_0, u_1)\|_{\dot H^1\times L^2}$$

and

$$\|(w_0, w_1)\|_{\dot H^1\times L^2} \geq (1-\gamma(\epsilon))\|P_m(u_0, u_1)\|_{\dot H^1\times L^2}$$

with a suitable $\gamma \to 0$ as $\epsilon \to 0$. If δ and ϵ are chosen sufficiently small, then by the channel of energy inequality for the linear wave equation and the bound on h, we conclude using (2.9) and (2.10) that for all $t \geq 0$,

$$\int_{|x|\geq \frac{\beta+1}{2}+t} |\nabla_{x,t} w|^2(x,t)\mathrm{d}x \geq \left|\frac{3+\beta}{4}\right| \|P_m(u_0, u_1)\|^2_{\dot H^1\times L^2} - C\epsilon^2 c_m^2. \tag{3.54}$$

Since $\nabla_{x,t} U_{<m-10}\,\psi$ is small in L^2 (smaller than $C\epsilon c_m$) and $U_{<m-10}$ is almost orthogonal by (3.42), the channel of energy inequality (3.29) for ψ follows (again by choosing ϵ, δ sufficiently small depending on β).

This finishes the proof of the theorem.

It remains to prove Claims 3.1 and 3.2.

Proof of Claim 3.1. We need to control I_1, I_2, I_3.

For I_1, by the trilinear estimate and symmetry, we get that for $m' = m + O(1)$

$$\begin{aligned}
&\left\|P_{m'}\left(u_{\geq m-10}\partial^\alpha u^\dagger \partial_\alpha u\right)\right\|_{N[m']} \\
&\quad = \left\|\sum_{k_1\geq m-10,\, k_2,\, k_3} P_{m'}\left(u_{k_1}\partial^\alpha u^\dagger_{k_2}\partial_\alpha u_{k_3}\right)\right\|_{N[m']} \\
&\quad \lesssim \sum_{k_1\geq m-10,\, k_2\geq k_3} 2^{-\kappa\left(\max\{k_1,k_2,k_3\}-m\right)_+} 2^{-\kappa(k_1-\min\{k_2,k_3\})_+} \times \\
&\qquad \times\ \|u_{k_1}\|_{S[k_1]}\|u_{k_2}\|_{S[k_2]}\|u_{k_3}\|_{S[k_3]} \\
&\quad \lesssim \sum_{k_1\geq m-10,\, k_2\geq k_3} 2^{-\kappa\left(\max\{k_1,k_2\}-m\right)_+} 2^{-\kappa(k_1-k_3)_+} c_{k_1}\epsilon^2 \\
&\quad \lesssim \epsilon^2 c_m \sum_{k_1\geq m-10,\, k_2\geq k_3} 2^{-(\kappa-\vartheta)\left(\max\{k_1,k_2\}-m\right)_+} 2^{-\kappa(k_1-k_3)_+} \lesssim \epsilon^2 c_m.
\end{aligned}$$

For I_2, by the product property and null form estimate, we get that for $m' = m + O(1)$

$$\begin{aligned}
&\left\|P_{m'}\left(u_{<m-10}\partial^\alpha u^\dagger_{>m+10}\partial_\alpha u\right)\right\|_{N[m']} \\
&\quad \lesssim \sum_{k_1>m+10,\, k_2=k_1+O(1)} \left\|P_{m'}\left(u_{<m-10}\partial^\alpha u^\dagger_{k_1}\partial_\alpha u_{k_2}\right)\right\|_{N[m']} \\
&\quad \lesssim \sum_{k_1>m+10,\, k_2=k_1+O(1)} 2^{-\kappa(k_1-m)}\|u_{k_1}\|_{S[k_1]}\|u_{k_2}\|_{S[k_2]}
\end{aligned}$$

$$\lesssim \sum_{k_1>m+10,\, k_2=k_1+O(1)} 2^{-\kappa(k_1-m)} c_{k_1} c_{k_2}$$
$$\lesssim \epsilon\, c_m \sum_{k_1>m+10,\, k_2=k_1+O(1)} 2^{-(\kappa-\vartheta)(k_1-m)} \lesssim \epsilon\, c_m.$$

For I_3, by the product property and null form estimate, we get that for $m' = m + O(1)$

$$\begin{aligned}
&\left\| P_{m'} \left(u_{<m-10} \partial^\alpha u^\dagger_{m-10\le\cdot\le m+10} \partial_\alpha u_{\ge m-10} \right) \right\|_{N[m']} \\
&\quad \lesssim \sum_{k\ge m-10} \left\| P_{m'} \left(u_{<m-10} \partial^\alpha u^\dagger_{m-10\le\cdot\le m+10} \partial_\alpha u_k \right) \right\|_{N[m']} \\
&\quad \lesssim \sum_{k\ge m-10} 2^{-\kappa(k-m)} \epsilon\, c_m \lesssim \epsilon\, c_m.
\end{aligned}$$

Thus the terms I_1, I_2, I_3 are all disposable. The claim is proved. ∎

Proof of Claim 3.2. Noting that

$$P_m \left(u_{m-10\le\cdot\le m+10} \right) = P_m u = \psi$$

by Lemma 2.1, we get that

$$\begin{aligned}
&P_m \left(u_{<m-10} \partial_\alpha u^\dagger_{<m-10} \partial^\alpha u_{m-10\le\cdot\le m+10} \right) - u_{<m-10} \partial_\alpha u^\dagger_{<m-10} \partial^\alpha \psi \\
&\quad = 2^{-m} L \left(\nabla \left(u_{<m-10} \partial_\alpha u^\dagger_{<m-10} \right), \partial^\alpha u_{m-10\le\cdot\le m+10} \right) \\
&\quad = 2^{-m} L \left(\nabla u_{<m-10} \partial_\alpha u^\dagger_{<m-10}, \partial^\alpha u_{m-10<\cdot<m+10} \right) + \\
&\qquad + 2^{-m} L \left(u_{<m-10} \partial_\alpha \nabla u^\dagger_{<m-10}, \partial^\alpha u_{m-10<\cdot<m+10} \right).
\end{aligned}$$

Thus, noting that

$$\|\nabla u_k\|_{S[k]} \lesssim 2^k \|u_k\|_{S[k]}$$

by the trilinear estimate for the first term in the above and the product estimate and null form estimate for the second, we get that for $m' = m + O(1)$

$$\begin{aligned}
&\left\| P_{m'} \left[P_m \left(u_{<m-10} \partial_\alpha u^\dagger_{<m-10} \partial^\alpha u_{m-10<\cdot<m+10} \right) - u_{<m-10} \partial_\alpha u^\dagger_{<m-10} \partial^\alpha \psi \right] \right\|_{N[m']} \\
&\quad \lesssim \sum_{k_1<m-10,\, k_2<m-10} 2^{-m} \left\| P_{m'} \left[L \left(\nabla u_{k_1} \partial_\alpha u^\dagger_{k_2}, \partial^\alpha u_{m-10<\cdot<m+10} \right) \right] \right\|_{N[m']}
\end{aligned}$$

$$+ \sum_{k<m-10} 2^{-m} \left\| P_{m'} \left[L \left(u_{<m-10} \partial_\alpha \nabla u_k^\dagger,\, \partial^\alpha u_{m-10<\cdot<m+10} \right) \right] \right\|_{N[m']}$$
$$\lesssim 2^{-m} \sum_{k_1<m-10,\, k_2<m-10} 2^{k_1} 2^{-\kappa(k_1-k_2)_+} c_m \epsilon^2 + \sum_{k<m-10} 2^{-m} 2^k \epsilon\, c_m \lesssim \epsilon c_m.$$

The first part of the claim is proved. The proof of the second part is similar. ■

4 Morawetz Estimates and Applications

In the previous sections, the main tools we use are all perturbative in nature. In order to understand the dynamics of large wave maps, we need some global control on the solution. Such global control is often achieved with help of suitable monotonicity formulae. The most important monotonicity formula here are the energy flux identity and the Morawetz estimate (see for example [20]). This section follows similar arguments in Sterbenz–Tataru [48].

For notational convenience, we shall work with a classical wave map u defined on $R^2 \times (0,1]$, that equals $u_\infty \in S^2$ for large x. Let us firstly look at the energy flux identity for u. We thus have

$$\partial_{tt} u - \Delta u = (|\nabla u|^2 - |u_t|^2)u, \text{ in } R^2 \times (0,1].$$

Noting that $u^\dagger \cdot u_t \equiv 0$, we have the identity

$$\left(\partial_{tt} u^\dagger - \Delta u^\dagger \right) \cdot u_t = 0, \text{ in } R^2 \times (0,1]. \tag{4.1}$$

Take $0 < t_1 < t_2 < 1$, and integrate the identity (4.1) in the truncated lightcone $\{(x,t) : |x| < t,\ t_1 < t < t_2\}$, we obtain that

$$\int_{|x|<t_2} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} \right)(x,t_2)\mathrm{d}x - \int_{|x|<t_1} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} \right)(x,t_1)\mathrm{d}x$$
$$- \frac{1}{\sqrt{2}} \int_{t_1}^{t_2} \int_{|x|=t} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} + \frac{x}{t} \cdot \nabla u^\dagger\, \partial_t u \right) \mathrm{d}\sigma\, \mathrm{d}t = 0.$$

Denote

$$\mathrm{Flux}(t_1,t_2) := \frac{1}{\sqrt{2}} \int_{t_1}^{t_2} \int_{|x|=t} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} + \frac{x}{t} \cdot \nabla u^\dagger\, \partial_t u \right) \mathrm{d}\sigma\, \mathrm{d}t$$

as the "energy flux" through the lateral boundary of the lightcone. We see that

$$\begin{aligned}\mathrm{Flux}(t_1,t_2) = &\int_{|x|<t_2} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2}\right)(x,t_2)\mathrm{d}x \\ &- \int_{|x|<t_1} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2}\right)(x,t_1)\mathrm{d}x.\end{aligned}$$

Since $\mathrm{Flux}(t_1,t_2) \geq 0$, it follows that

$$\int_{|x|<t} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2}\right)(x,t)\mathrm{d}x$$

is nondecreasing, and has a limit as $t \to 0+$. Thus $\mathrm{Flux}(t_1,t_2) \to 0+$ as t_1, $t_2 \to 0+$.

The control of energy flux plays an essential role in the following Morawetz estimate.

Theorem 4.1. Let u be a classical wave map with energy E on $R^2 \times (0,1]$ and $\epsilon \in (0,1)$. For each $0 < \bar{t} < 1$, if $\mathrm{Flux}(0,\bar{t}) < \epsilon E$, then

$$\int_{\epsilon\bar{t}}^{\bar{t}} \int_{|x|<t} \rho_{\epsilon\bar{t}}^3 \left(X^\alpha \partial_\alpha u\right)^2 \mathrm{d}x\mathrm{d}t + \int_{|x|<\bar{t}} \bar{t}\rho_{\epsilon\bar{t}} \left(\frac{|\nabla u|^2}{2} + \frac{|u_t|^2}{2} + \frac{x}{\bar{t}}\cdot \nabla u^\dagger\, u_t\right)(x,\bar{t})\mathrm{d}x \lesssim E, \tag{4.2}$$

where we set $\rho_{\epsilon\bar{t}} := \left((t+\epsilon\bar{t})^2 - |x|^2\right)^{-\frac{1}{2}}$ and $X^\alpha = x^\alpha$ if $\alpha = 1,2$, $X^0 = t+\epsilon\bar{t}$. □

Proof. By rescaling, we can assume without loss of generality that $\bar{t} = 1$. (Then u is rescaled to $R^2 \times \left(0, \frac{1}{\bar{t}}\right]$) Let us integrate the identity

$$\partial^\alpha \partial_\alpha u^\dagger\, \rho_\epsilon\, X^\beta \partial_\beta u = 0$$

on $\{(x,t) : |x| < t,\ \epsilon < t < 1\}$. We have

$$\begin{aligned}0 &= \int_\epsilon^1 \int_{|x|<t} \partial^\alpha \partial_\alpha u^\dagger\, \rho_\epsilon X^\beta \partial_\beta u \mathrm{d}x\mathrm{d}t \\ &= \int_\epsilon^1 \int_{|x|<t} \rho_\epsilon\, X^\beta \partial^\alpha (\partial_\alpha u^\dagger\, \partial_\beta u) - \rho_\epsilon\, X^\beta \partial_\beta \frac{\partial^\alpha u^\dagger \partial_\alpha u}{2} \mathrm{d}x\mathrm{d}t, \\ &= B + I,\end{aligned}$$

where the boundary term B and the interior term I are

$$B = \int_\epsilon^1 \int_{|x|=t} \rho_\epsilon X^\beta n^\alpha \partial_\alpha u^\dagger \partial_\beta u - \rho_\epsilon X^\beta n_\beta \frac{\partial^\alpha u^\dagger \partial_\alpha u}{2} \mathrm{d}\sigma \mathrm{d}t$$

$$- \int_{|x|<1} \rho_\epsilon X^\beta \partial_t u^\dagger \partial_\beta u(x,1)\mathrm{d}x - \int_{|x|<1} (1+\epsilon)\rho_\epsilon \frac{\partial^\alpha u^\dagger \partial_\alpha u}{2}(x,1)\mathrm{d}x$$

$$+ \int_{|x|<\epsilon} \rho_\epsilon X^\beta \partial_t u^\dagger \partial_\beta u(x,\epsilon)\mathrm{d}x + \int_{|x|<\epsilon} 2\epsilon\, \rho_\epsilon \frac{\partial^\alpha u^\dagger \partial_\alpha u}{2}(x,\epsilon)\mathrm{d}x;$$

and

$$I = -\int_\epsilon^1 \int_{|x|<t} \partial^\alpha \left(\rho_\epsilon X^\beta\right) \partial_\alpha u^\dagger \partial_\beta u - \partial_\beta \left(\rho_\epsilon X^\beta\right) \frac{\partial^\alpha u^\dagger \partial_\alpha u}{2} \mathrm{d}x\, \mathrm{d}t.$$

In the above we use the notation $n = \frac{1}{\sqrt{2}}\left(\frac{x}{|x|}, -1\right)$, $n^j = n_j = \frac{1}{\sqrt{2}}\frac{x_j}{|x|}$ for $j = 1,2$ and $n^0 = -n_0 = \frac{1}{\sqrt{2}}$. Hence $X^\beta n_\beta = -\frac{\epsilon}{\sqrt{2}}$ on $|x| = t$. We can compute

$$\partial_j \rho_\epsilon = x_j \rho_\epsilon^3, \;\; \partial_t \rho_\epsilon = -t\rho_\epsilon^3 - \epsilon\, \rho_\epsilon^3.$$

Hence

$$X^\beta \partial_\beta \rho_\epsilon = -\rho_\epsilon.$$

We also note that $\epsilon\, \rho \leq 1$ when $|x| < t$, and record the following simple bound when $|x| = t$

$$|\rho_\epsilon| = \left(2\epsilon\, t + \epsilon^2\right)^{-\frac{1}{2}} \leq \epsilon^{-\frac{1}{2}} t^{-\frac{1}{2}}.$$

We can simplify the B, I terms as

$$B = \frac{1}{\sqrt{2}} \int_\epsilon^1 \int_{|x|=t} \rho_\epsilon \left(X^\beta \partial_\beta u\right) \cdot \frac{(x^\alpha \partial_\alpha u)}{t} + \epsilon\, \rho_\epsilon \frac{\partial^\alpha u^\dagger \partial_\alpha u}{2} \mathrm{d}\sigma \mathrm{d}t$$

$$- \int_{|x|<1} \rho_\epsilon \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} + x \cdot \nabla u^\dagger \partial_t u\right)(x,1)\mathrm{d}x + O(E)$$

and

$$-I = \int_\epsilon^1 \int_{|x|<t} \Big[X^\beta \partial_\beta u^\dagger \, \partial^\alpha \rho_\epsilon \, \partial_\alpha u + \rho_\epsilon \, \left(|\nabla u|^2 - |\partial_t u|^2 \right) \\ - \frac{3}{2} \rho_\epsilon \, \left(\partial^\alpha u^\dagger \, \partial_\alpha u \right) - X^\beta \partial_\beta \rho_\epsilon \, \frac{\partial^\alpha u^\dagger \, \partial_\alpha u}{2} \Big] \mathrm{d}x \mathrm{d}t \\ = \int_\epsilon^1 \int_{|x|<t} \rho_\epsilon^3 \left| X^\beta \partial_\beta u \right|^2 \mathrm{d}x \mathrm{d}t.$$

We can estimate

$$\left| \frac{1}{\sqrt{2}} \int_\epsilon^1 \int_{|x|=t} \rho_\epsilon \, \left(X^\beta \partial_\beta u \right) \cdot \frac{(x^\alpha \partial_\alpha u)}{t} + \epsilon \, \rho_\epsilon \frac{\partial^\alpha u^\dagger \, \partial_\alpha u}{2} \mathrm{d}\sigma \, \mathrm{d}t \right| \\ \lesssim \int_\epsilon^1 \int_{|x|=t} \rho_\epsilon \, t \left| \partial_t u + \frac{x}{t} \cdot \nabla u \right|^2 \mathrm{d}\sigma \, \mathrm{d}t + \\ + \int_\epsilon^1 \int_{|x|=t} \epsilon \, \rho_\epsilon \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} + \frac{x}{t} \cdot \nabla u^\dagger \, \partial_t u \right) \mathrm{d}\sigma \, \mathrm{d}t \\ \lesssim \epsilon^{-\frac{1}{2}} \mathrm{Flux}(0,1) \leq \epsilon^{\frac{1}{2}} E.$$

Hence, combining the B and I terms, we conclude that

$$\int_{|x|<1} \rho_\epsilon \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} + x \cdot \nabla u^\dagger \, \partial_t u \right) (x,1) \mathrm{d}x + \int_\epsilon^1 \int_{|x|<1} \rho_\epsilon^3 \left| X^\beta \partial_\beta u \right|^2 \mathrm{d}x \mathrm{d}t \\ \lesssim E.$$

The theorem is proved. ■

Theorem 4.1 has the following corollary.

Corollary 4.1. Let u be as above. For any $\tau_n \to 0+$, $\gamma_n \to 1-$ as $n \to \infty$, we have that

$$\int_{B_{\tau_n} \setminus B_{\gamma_n \tau_n}} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} + \frac{x}{t} \cdot \nabla u^\dagger \, \partial_t u \right) (x, \tau_n) \mathrm{d}x = o_n(1). \tag{4.3}$$

□

Proof. Let $\epsilon_n := 2\,\mathrm{Flux}(0, \tau_n)/E$, then $\epsilon_n \to 0$ as $n \to \infty$. Theorem 4.1 implies that

$$\int_{B_{\tau_n}} \tau_n \, \rho_{\epsilon_n \tau_n} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} + \frac{x}{t} \cdot \nabla u^\dagger \, \partial_t u \right) (x, \tau_n) \mathrm{d}x \lesssim E.$$

Note that

$$\tau_n\, \rho_{\epsilon_n \tau_n} \gtrsim \left((1+\epsilon_n)^2 - \gamma_n^2\right)^{-\frac{1}{2}} \to \infty$$

for $t = \tau_n$, $|x| \in (\gamma_n \tau_n, \tau_n)$, we conclude that (4.3) holds. ■

5 Universal Blow up Profile Along a Sequence of Times

Our goal in this section is to prove Theorem 1.2 along a sequence of times. Again for the ease of notations, we shall consider classical wave map u defined on $R^2 \times (0,1]$ that blows up at time $t = 0$. Recall from (2.25), the definition of $r(\epsilon_*, t)$. By the small data theory and finite speed of propagation, we have $\lim_{t\to 0+} r(\epsilon_*, t) = 0$. That is, energy concentrates in smaller and smaller regions as $t \to 0+$. By the definition of $r(\epsilon_*, t)$, we can find $x_*(t)$ such that

$$\|\vec{u}\,(t)\|_{\dot{H}^1 \times L^2 (B_{2r(\epsilon_*,t)}(x_*(t)))} > \frac{\epsilon_*}{2} \tag{5.1}$$

for t close to 0. Again by small data global existence and finite speed of propagation, $x_*(t)$ remains in a bounded region for $t \in (0,1]$. Assume without loss of generality that $x_*(t_n) \to 0$ as $t_n \to 0$ along a sequence of times t_n. Since $r(\epsilon_*, t) \to 0$, we see that for any $r > 0$,

$$\liminf_{t\to 0} \left\|\vec{u}\,(t)\right\|_{\dot{H}^1 \times L^2(B_r)} > \frac{\epsilon_*}{2}. \tag{5.2}$$

In general, we call a point $\bar{x}$ *singular* if for any $r > 0$

$$\limsup_{t\to 0} \|\vec{u}\,(t)\|_{\dot{H}^1 \times L^2(B_r(\bar{x}))} > \frac{\epsilon_*}{2}.$$

By finite speed of propagation and energy flux identity, this is equivalent to requiring that for any $r > 0$,

$$\liminf_{t\to 0} \|\vec{u}\,(t)\|_{\dot{H}^1 \times L^2(B_r(\bar{x}))} > \frac{\epsilon_*}{2}.$$

As the energy is conserved and finite, there can only be finitely many singular points, and in particular the singular points are isolated.

For the singular point $x_* = 0$, since singular points are isolated, there exists $r_1 > 0$ such that for any $\bar{x} \in B_{r_1} \backslash \{0\}$, $\bar{x}$ is not a singular point. Hence we can find $\tilde{r} > 0$ with

$$\|\vec{u}\,(\tau_n)\|_{\dot{H}^1 \times L^2(B_{\tilde{r}}(\bar{x}))} < \epsilon_* \tag{5.3}$$

along a sequence of times $\tau_n \to 0$. In particular, we have

$$\|\partial\!\!\!/ u(\cdot,\, \tau_n)\|_{L^2(B_{\tilde{r}}(\overline{x}))} < \epsilon_*.$$

Hence there exists $\overline{r}_n \in \left(\frac{\tilde{r}}{2},\, \tilde{r}\right)$ with

$$\int_{\partial B_{\overline{r}_n}(\overline{x})} |\partial\!\!\!/ u|^2(x,\tau_n)\mathrm{d}\sigma \lesssim \frac{\epsilon_*^2}{\overline{r}_n}.$$

Denoting $\overline{u}_n$ as the average of $u(\tau_n)$ over $\partial B_{\overline{r}_n}(\overline{x})$, that is

$$\overline{u}_n = \frac{1}{2\pi\overline{r}_n}\int_{\partial B_{\overline{r}_n}(\overline{x})} u(\tau_n)\mathrm{d}\sigma.$$

By Sobolev inequality, we get that

$$\|u(\tau_n) - \overline{u}_n\|_{L^\infty(|x-\overline{x}|=\overline{r}_n)} \lesssim \epsilon_*. \tag{5.4}$$

Take smooth cutoff function $\eta_n \in C_c^\infty\left(B_{2\overline{r}_n}(\overline{x})\right)$ with $\eta_n|_{B_{\overline{r}_n}(\overline{x})} \equiv 1$ and $|\nabla\eta_n| \lesssim (\overline{r}_n)^{-1}$. Recall that for any $v \in R^2$ with $v \neq 0$,

$$Pv = \frac{v}{|v|}.$$

Define

$$(u_{0n},\, u_{1n}) = \begin{cases} (u,\, \partial_t u)(\tau_n) & \text{in } B_{\overline{r}_n}(\overline{x}); \\ (P[\eta_n(u(\overline{r}_n\theta,\tau_n) - \overline{u}_n) + \overline{u}_n],\, 0) & \text{in } \left(B_{\overline{r}_n}(\overline{x})\right)^c. \end{cases}$$

By (5.3) and (5.4), direct computation shows that

$$\|(u_{0n},\, u_{1n})\|_{\dot{H}^1\times L^2} \lesssim \epsilon_*. \tag{5.5}$$

Hence by small energy global existence theory and finite speed of propagation, we see that the solution u_n to the wave map equation with $\vec{u}_n(\tau_n) = (u_{0n},\, u_{1n})$ is global and that

$$u_n \equiv u \;\text{ for }\; |x-\overline{x}| < \frac{\overline{r}_n}{4} \;\text{ and }\; t \in (0,\, \tau_n] \tag{5.6}$$

for sufficiently large n. Since $u_n \in C\left([0,\, \tau_n],\, \dot{H}^1 \times L^2\right)$ and (5.6) holds, we conclude that u can be extended to $t = 0$ so that $u \in C\left([0,\, \tau_n],\, \dot{H}^1 \times L^2(B_{\tilde{r}/8}(\overline{x}))\right)$. Since $\overline{x} \in$

$B_{r_1}\backslash\{0\}$ is arbitrary, we conclude that u can be extended to $t = 0$ in B_{r_1} with $u \in C\left([0, 1], \dot{H}^1 \times L^2(B_{r_1}\backslash B_r)\right)$ for each $0 < r < r_1$.

In addition, by the regularity of u_n, we also have the additional (qualitative) regularity condition that $u \in C^\infty\left(B_{r_1} \times [0, 1]\backslash\{(0,0)\}\right)$. One can of course apply the same argument to other singular points. As a result, we see that $u \in C^\infty(R^2 \times [0,1]\backslash\{(x_j, 0)\})$ where x_j are the singular points.

On the other hand, since $\vec{u}(t)$ is bounded in $\dot{H}^1 \times L^2$ and $|u| \equiv 1$, we can extract a weak limit $(v_0, v_1) \in \dot{H}^1 \times L^2$ along a sequence of times $t_n \to 0+$. This limit is in fact a strong limit outside an arbitrarily small neighborhood of the finitely many singular points. From the above analysis, $(v_0, v_1) \in C^\infty(R^2\backslash\{x_j\})$. Let

$$v = u$$

for $\inf_j |x - x_j| > t$. Then $v \in C^\infty\left(R^2 \times [0,1]\backslash \bigcup_j\{|x - x_j| \le t,\ t \in [0, 1]\}\right)$, and by the same arguments as in the proof of Lemma 2.4,

$$\lim_{t\to 0+} \left\| \vec{v}(\cdot, t) - (v_0, v_1)\right\|_{\dot{H}^1 \times L^2\left(\bigcap_j\{|x-x_j|>t\}\right)} = 0. \tag{5.7}$$

We shall call v the *regular part* of the wave map u. The main issue is to understand the behavior of the wave map u inside singularity lightcones $\bigcup_j\{|x - x_j| \le t,\ t \in [0, 1]\}$.

We shall prove

Theorem 5.1. Let u be a classical wave map with energy

$$\mathcal{E}(\vec{u}) < \mathcal{E}(Q,0) + \epsilon_0^2, \tag{5.8}$$

where Q is a harmonic map with degree 1, defined on $R^2 \times (0, 1]$ that blows up at time $t = 0$ with the origin being a singular point. Assume that ϵ_0 is sufficiently small. Then there exists a sequence of times $t_n \to 0+$, $\ell \in R^2$ with $|\ell| \ll 1$, $x_n \in R^2$, $\lambda_n > 0$ with

$$\lim_{n\to\infty} \frac{x_n}{t_n} = \ell, \quad \lambda_n = o(t_n)$$

and $(v_0, v_1) \in \dot{H}^1 \times L^2 \cap C^\infty(R^2\backslash\{0\})$, such that

$$\vec{u}(t_n) = (v_0, v_1) + \left(Q_\ell,\ \lambda_n^{-1}\partial_t Q_\ell\right)\left(\frac{x - x_n}{\lambda_n},\ \frac{t - t_n}{\lambda_n}\right) + o_{\dot{H}^1\times L^2}(1), \tag{5.9}$$

as $n \to \infty$. □

Remark. As we discussed in the introduction, the main new point in Theorem 5.1 is that we eliminate any possible energy concentration near the boundary of singularity lightcone $|x| < t$. By the energy constraint (5.8) and the proof below, there is only one singularity point. Hence u is regular outside $\{(x,t) : |x| \le t\}$. The main task is to understand the behavior of u inside $\{(x,t) : |x| \le t\}$. □

Proof. Our starting point is the work of Grinis [21], which completely characterized the concentration of energy in $\{(x, \tau_n) : |x| < a\tau_n\}$ for any $a \in (0,1)$ as traveling waves, for a suitable time sequence $\tau_n \to 0+$. See Theorem 1.1 and Theorem 1.2 in [21]. In our case, due to the energy constraint (5.8), there can only be one traveling wave. Hence, as a particular consequence of a rescaled version of the *asymptotic decomposition* in Theorem 1.2 of [21], we have for $|x| < \tau_n$,

$$\vec{u}(\tau_n) := \left(Q_\ell\left(\frac{x-x_n}{r_n}, 0\right), r_n^{-1}\nabla\partial_t Q_\ell\left(\frac{x-x_n}{r_n}, 0\right)\right) + (w_{0n}, w_{1n}) + o_{\dot H^1\times L^2}(1), \tag{5.10}$$

as $n \to \infty$, where $|\ell| \ll 1$, $r_n = o(\tau_n)$, $\ell = \lim_{n\to\infty} \frac{x_n}{\tau_n}$ and

$$\int_{|x|<a\tau_n} |\nabla w_{0n}|^2 + |w_{1n}|^2 dx \to 0, \tag{5.11}$$

as $n \to \infty$ for any $a \in (0,1)$. Our main task is to show that

$$\int_{|x|<\tau_n} |\nabla w_{0n}|^2 + |w_{1n}|^2 dx \to 0$$

as $n \to \infty$. By (5.11), we have to prove that that for any $\gamma_n \to 1-$,

$$\limsup_{n\to\infty} \int_{B_{\tau_n}\setminus B_{\gamma_n\tau_n}} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2}\right)(x,\tau_n)dx = 0 \tag{5.12}$$

assuming that

$$\limsup_{n\to\infty} \int_{B_{\gamma_n\tau_n}} |\nabla w_{0n}|^2 + |w_{1n}|^2 dx = 0. \tag{5.13}$$

We now apply the channel of energy inequality and prove (5.12). ■

Suppose that (5.12) is not true. Then there exists $\epsilon_2 > 0$, such that, by passing to a subsequence if necessary, we have for all sufficiently large n,

$$\mathcal{E}_n^2 := \int_{B_{\tau_n}\setminus B_{\gamma_n\tau_n}} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2}\right)(x,\tau_n)dx \ge \epsilon_2^2. \tag{5.14}$$

By the energy constraint, we must also have

$$\int_{B^c_{\frac{\tau_n}{2}} \cap B_{\tau_n}} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} \right)(x, \tau_n) \mathrm{d}x \lesssim \epsilon_0^2. \tag{5.15}$$

Corollary 4.1 implies that

$$\int_{B_{\tau_n} \backslash B_{\gamma_n \tau_n}} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} + \partial_t u^\dagger \, \partial_r u \right)(x, \tau_n) \mathrm{d}x = o_n(1). \tag{5.16}$$

Since u is regular for $|x| > t$, we have for any $r > 0$,

$$\limsup_{t \to 0+} \int_{B_{2r} \backslash B_t} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} \right)(x, t) \mathrm{d}x \leq \delta(r) \to 0, \text{ as } r \to 0+. \tag{5.17}$$

Fix a small $r > 0$ whose value is to be determined below. We can find $r_{1n} \in \left(\frac{r}{2}, r\right)$, $r_{2n} \in \left(\frac{\tau_n}{2}, \frac{3}{4}\tau_n\right)$, such that

$$\int_{\partial B_{r_{1n}}} |\not{\partial} u|^2(\tau_n) \mathrm{d}\sigma \lesssim \frac{\delta(r)}{r_{1n}}, \text{ and } \int_{\partial B_{r_{2n}}} |\not{\partial} u|^2(\tau_n) \mathrm{d}\sigma = \frac{o_n(1)}{r_{2n}}.$$

Let

$$\overline{u}_n^1 = \frac{1}{2\pi r_{1n}} \int_{\partial B_{r_{1n}}} u(\tau_n) \mathrm{d}\sigma, \text{ and } \overline{u}_n^2 = \frac{1}{2\pi r_{2n}} \int_{\partial B_{r_{2n}}} u(\tau_n) \mathrm{d}\sigma.$$

Fix radial $\eta_{1n} \in C_c^\infty(B_{2r_{1n}})$ with $\eta_{1n}|_{B_{r_{1n}}} \equiv 1$, and radial $1 - \eta_{2n} \in C_c^\infty(B_{r_{2n}})$ with $1 - \eta_{2n}|_{B_{\frac{r_{2n}}{2}}} \equiv 1$. Define

$$(u_{0n}, u_{1n}) = \begin{cases} \left(P\left[\eta_{1n} \left(u(r_{1n}\theta, \tau_n) - \overline{u}_n^1\right) + \overline{u}_n^1\right], 0\right) & \text{in } B^c_{r_{1n}}; \\ \vec{u}\,(\tau_n) & \text{in } B_{r_{1n}} \backslash B_{r_{2n}}; \\ \left(P\left[\eta_{2n} \left(u(r_{2n}\theta, \tau_n) - \overline{u}_n^2\right) + \overline{u}_n^2\right], 0\right) & \text{in } B_{r_{2n}}. \end{cases} \tag{5.18}$$

Then for sufficiently large n, in view of (5.13) and (5.17),

$$\|(u_{0n}, u_{1n})\|_{\dot{H}^1 \times L^2(B^c_{\tau_n} \cup B_{\tau_n \gamma_n})} \lesssim \delta(r)$$

and

$$\epsilon_0 \gtrsim \|(u_{0n}, u_{1n})\|_{\dot{H}^1 \times L^2} > \mathcal{E}_n + O(\delta(r)) \geq \epsilon_2 + O(\delta(r)).$$

In addition, by (5.16), for sufficiently large n

$$\|u_{1n} + \partial_r u_{0n}\|_{L^2} + \|\not{\partial} u_{0n}\|_{L^2} \lesssim \delta(r).$$

Let u_n be the solution to the wave map equation with $\vec{u}(\tau_n) = (u_{0n}, u_{1n})$. Then if r is taken sufficiently small so that $\delta(r)$ is much smaller than ϵ_2 by (a rescaled and time translated version of) Theorem 3.1 we conclude that for $t \geq \tau_n$

$$\int_{|x|>t-\frac{\tau_n}{8}} \left|\nabla_{x,t} u_n\right|^2 (x,t)\mathrm{d}x \gtrsim \mathcal{E}_n^2. \tag{5.19}$$

Take $t = \frac{r}{8}$ in (5.19), we get that for all sufficiently large n,

$$\int_{|x|>\frac{r}{8}-\frac{\tau_n}{8}} \left|\nabla_{x,t} u_n\right|^2 (x,\frac{r}{8})\mathrm{d}x \gtrsim \mathcal{E}_n^2. \tag{5.20}$$

By the energy inequality, (5.17) and the definition of u_n, we see that for $t \leq \frac{r}{8}$,

$$\int_{|x|>t} \left|\nabla_{x,t} u_n\right|^2 (x,t)\mathrm{d}x \lesssim \delta(r). \tag{5.21}$$

By finite speed of propagation, we also have $u \equiv u_n$ for $t - \frac{\tau_n}{4} < |x| < \frac{r}{4}$ and $t \leq \frac{r}{8}$. Combining with (5.20), we conclude that

$$\int_{\frac{r}{8}>|x|>\frac{r}{8}-\frac{\tau_n}{4}} \left|\nabla_{x,t} u\right|^2 (x,\frac{r}{8})\mathrm{d}x \gtrsim \mathcal{E}_n^2 \geq \epsilon_2^2 > 0, \tag{5.22}$$

if we choose r sufficiently small, so that $\delta(r)$ is much smaller than ϵ_2^2. However, (5.22) contradicts with the fact that $\vec{u}\left(\frac{r}{8}\right) \in \dot{H}^1 \times L^2$ for sufficiently large n.

Therefore, combining the above with the regular part outside the singularity lightcone, we get that along the sequence τ_n,

$$\vec{u}(\tau_n) = (v_0, v_1) + (Q_\ell, r_n^{-1}\partial_t Q_\ell)\left(\frac{x - x_n}{r_n}, 0\right) + o_{\dot{H}^1\times L^2}(1), \text{ as } n \to \infty. \tag{5.23}$$

The theorem is proved.

6 Coercivity and Universal Profile for All Times

Our next task is to use a rigidity property of the energy to extend the decomposition we obtained from the last section to all times. One important tool is the following coercivity property of the energy functional near traveling waves.

Theorem 6.1. Let $\mathcal{M}_1$ be the space of harmonic maps from R^2 to S^2 with topological degree 1. Fix $\ell \in R^2$ with $|\ell| < 1$ and for any $Q \in \mathcal{M}_1$, let Q_ℓ be the Lorentz transform of

Q with velocity ℓ, that is,

$$Q_\ell(x,t) = Q\left(x - \frac{\ell\cdot x}{|\ell|^2}\ell + \frac{\frac{\ell\cdot x}{|\ell|^2}\ell - \ell t}{\sqrt{1-|\ell|^2}}\right). \tag{6.1}$$

Denote $\mathcal{M}_{1,\ell}$ as the space of Q_ℓ with $Q \in \mathcal{M}_1$. For $0 < \epsilon < \epsilon_0$ and ϵ_0 sufficiently small, suppose that $(v_0, v_1) \in \dot{H}^1 \times L^2$, with $|v_0(x)| \equiv 1$ and $v_0^\dagger \cdot v_1 \equiv 0$, satisfies

$$\deg(v_0) = 1; \tag{6.2}$$

$$\left|\int_{R^2} \partial_{x_j} v_0^\dagger\, v_1 \mathrm{d}x - \int_{R^2} \partial_{x_j} Q_\ell^\dagger\, \partial_t Q_\ell \mathrm{d}x\right| < \epsilon; \tag{6.3}$$

$$\int_{R^2}\left(\frac{|\nabla v_0|^2}{2} + \frac{|v_1|^2}{2}\right)\mathrm{d}x \le \int_{R^2}\left(\frac{|\partial_t Q_\ell|^2}{2} + \frac{|\nabla Q_\ell|^2}{2}\right)\mathrm{d}x + \epsilon; \tag{6.4}$$

$$\inf_{Q\in\mathcal{M}_1} \|(v_0,\, v_1) - (Q_\ell,\, \partial_t Q_\ell)\|_{\dot{H}^1\times L^2} < \epsilon_0. \tag{6.5}$$

Then there exists $\delta(\epsilon) > 0$ with $\delta(\epsilon) \to 0$ as $\epsilon \to 0$, such that

$$\inf_{Q\in\mathcal{M}_1} \|(v_0,\, v_1) - (Q_\ell,\, \partial_t Q_\ell)\|_{\dot{H}^1\times L^2} < \delta(\epsilon). \tag{6.6}$$

□

Remark. As we will see in the proof

$$\int_{R^2} \frac{|\nabla Q_\ell|^2}{2} + \frac{|\partial_t Q_\ell|^2}{2}\mathrm{d}x = \frac{4\pi}{\sqrt{1-|\ell|^2}},$$
$$-\int_{R^2} \partial_t Q_\ell^\dagger\, \partial_{x_j} Q_\ell \mathrm{d}x = \frac{4\pi\,\ell_j}{\sqrt{1-|\ell|^2}}$$

for any $Q \in \mathcal{M}_1$. The conditions (6.3) and (6.4) are thus independent of the choice of $Q \in \mathcal{M}_1$. □

The definition of degree $\deg(f)$ for mappings between manifolds is classical. For the definition with $f : R^2 \to S^2 \subset R^3$ and $f \in \dot{H}^1$ used here, we refer to [3], see in particular (1) in page 205 of [3]. We also remark that the harmonic maps in $\mathcal{M}_1$ have been completely characterized as degree 1 rational functions (Möbius transforms), see [19] and a more recent discussion in [39]. By elementary geometric properties of Möbius transforms, it is easy to see that degree one harmonic maps from R^2 to $S^2 \subset R^3$ are unique up to the symmetries of R^2 and S^2. More precisely in an appropriate coordinate system, the harmonic maps in $\mathcal{M}_1$ are co-rotational.

Proof. Without loss of generality, let us assume that $\ell = le_1 = (l, 0)$. Suppose that (6.6) is false, then for each $n = 1, 2, \ldots$, by symmetry, we can assume that there exist $(v_{0n}, v_{1n}) \in \dot{H}^1 \times L^2$ with $|v_{0n}| \equiv 1$ and $v_{0n}^\dagger v_{1n} \equiv 0$, such that

$$\deg(v_{0n}) = 1; \tag{6.7}$$

$$\left| \int_{R^2} \partial_{x_j} v_{0n}^\dagger v_{1n} \mathrm{d}x - \int_{R^2} \partial_{x_j} Q_\ell^\dagger \, \partial_t Q_\ell \mathrm{d}x \right| < \frac{1}{n}; \tag{6.8}$$

$$\int_{R^2} \left(\frac{|\nabla v_{0n}|^2}{2} + \frac{|v_{1n}|^2}{2} \right) \mathrm{d}x \le \int_{R^2} \left(\frac{|\partial_t Q_\ell|^2}{2} + \frac{|\nabla Q_\ell|^2}{2} \right) \mathrm{d}x + \frac{1}{n}; \tag{6.9}$$

$$\inf_{Q \in \mathcal{M}_1} \|(v_{0n}, v_{1n}) - (Q_\ell, \partial_t Q_\ell)\|_{\dot{H}^1 \times L^2} < \epsilon_0. \tag{6.10}$$

In addition,

$$\inf_{Q \in \mathcal{M}_1} \|(v_{0n}, v_{1n}) - (Q_\ell, \partial_t Q_\ell)\|_{\dot{H}^1 \times L^2} > \delta_0 > 0. \tag{6.11}$$

For fixed $(v_0, v_1) \in \dot{H}^1 \times L^2$, with $|v_0(x)| \equiv 1$ and $v_0^\dagger \cdot v_1 \equiv 0$, assume without loss of generality that v_0 is positively oriented, that is,

$$\deg(v_0) = -\frac{1}{4\pi} \int_{R^2} v_0^\dagger \cdot (\partial_1 v_0 \times \partial_2 v_0). \tag{6.12}$$

Consider the following algebraic identity

$$\begin{aligned} &\int_{R^2} \left(\frac{|\nabla v_0|^2}{2} + \frac{|v_1|^2}{2} \right) \mathrm{d}x \\ &= \frac{1}{2} \int_{R^2} |v_1 + l\partial_1 v_0|^2 \,\mathrm{d}x + \frac{1}{4} \int_{R^2} \left| \sqrt{1-l^2}\partial_1 v_0 - v_0 \times \partial_2 v_0 \right|^2 \mathrm{d}x \\ &\quad + \frac{1}{4} \int_{R^2} \left| \partial_2 v_0 + \sqrt{1-l^2} v_0 \times \partial_1 v_0 \right|^2 dx - \sqrt{1-l^2} \int_{R^2} v_0^\dagger \cdot (\partial_1 v_0 \times \partial_2 v_0) \,\mathrm{d}x \\ &- l \int_{R^2} \partial_1 v_0^\dagger \, v_1 \mathrm{d}x. \end{aligned} \tag{6.13}$$

(6.13) is a modified form of the remarkable decomposition of energy in [1], see also the illuminating discussion in page 3 of [41]. The modification here is necessary in order to take into account the momentum part.

Direct calculations show that

$$\int_{R^2} \left(\frac{|\partial_t Q_\ell|^2}{2} + \frac{|\nabla Q_\ell|^2}{2} \right) \mathrm{d}x = \frac{4\pi}{\sqrt{1-l^2}} \quad \text{and} \quad -\int_{R^2} \partial_t Q_\ell^\dagger \, \partial_1 Q_\ell \mathrm{d}x = \frac{4\pi l}{\sqrt{1-l^2}}. \tag{6.14}$$

We can assume, after rotation, that (v_{0n}, v_{1n}) has the same momentum as $(Q_{\ell_n}, \partial_t Q_{\ell_n})$ with $\ell_n = l_n e_1$. Then $|l_n - l| \lesssim \frac{1}{n}$. Applying (6.13) to (v_{0n}, v_{1n}) and using the assumptions on (v_{0n}, v_{1n}), we get that

$$\begin{aligned}
&\int_{R^2}\left(\frac{|\nabla v_{0n}|^2}{2}+\frac{|v_{1n}|^2}{2}\right)\mathrm{d}x\\
&\quad=\frac{1}{2}\int_{R^2}|v_{1n}+l_n\partial_1 v_{0n}|^2\,\mathrm{d}x+\frac{1}{4}\int_{R^2}\left|\sqrt{1-l_n^2}\partial_1 v_{0n}-v_{0n}\times\partial_2 v_{0n}\right|^2\mathrm{d}x\\
&\qquad+\frac{1}{4}\int_{R^2}\left|\partial_2 v_{0n}+\sqrt{1-l_n^2}v_{0n}\times\partial_1 v_{0n}\right|^2\mathrm{d}x-\sqrt{1-l_n^2}\int_{R^2}v_{0n}^{\dagger}\cdot(\partial_1 v_{0n}\times\partial_2 v_{0n})\,\mathrm{d}x\\
&\qquad-l_n\int_{R^2}\partial_1 v_{0n}^{\dagger}\,v_{1n}\mathrm{d}x\\
&\quad=\frac{1}{2}\int_{R^2}|v_{1n}+l\partial_1 v_{0n}|^2\,\mathrm{d}x+\frac{1}{4}\int_{R^2}\left|\sqrt{1-l^2}\partial_1 v_{0n}-v_{0n}\times\partial_2 v_{0n}\right|^2\mathrm{d}x\\
&\qquad+\frac{1}{4}\int_{R^2}\left|\partial_2 v_{0n}+\sqrt{1-l^2}v_{0n}\times\partial_1 v_{0n}\right|^2\mathrm{d}x+\frac{4\pi}{\sqrt{1-l^2}}+O\left(\frac{1}{n}\right).
\end{aligned}$$

In the above we used the expression for degree and momentum. From (6.9) and (6.14), we conclude that

$$\begin{aligned}
&\frac{1}{2}\int_{R^2}|v_{1n}+l\partial_1 v_{0n}|^2\,\mathrm{d}x+\frac{1}{4}\int_{R^2}\left|\sqrt{1-l^2}\partial_1 v_{0n}-v_{0n}\times\partial_2 v_{0n}\right|^2\mathrm{d}x\\
&\qquad+\frac{1}{4}\int_{R^2}\left|\partial_2 v_{0n}+\sqrt{1-l^2}v_{0n}\times\partial_1 v_{0n}\right|^2\mathrm{d}x\\
&\quad=O\left(\frac{1}{n}\right).
\end{aligned}\tag{6.15}$$

By (6.10), applying suitable symmetry transformation to (v_{0n}, v_{1n}) if necessary, we can assume that for suitable $\widetilde{Q}_\ell\in\mathcal{M}_{\ell,1}$,

$$(v_{0n}, v_{1n})=(\widetilde{Q}_\ell,\,\partial_t\widetilde{Q}_\ell)(x,0)+(r_{0n}, r_{1n})\tag{6.16}$$

with

$$\|(r_{0n}, r_{1n})\|_{\dot{H}^1\times L^2}\leq 2\epsilon_0.$$

Passing to a subsequence, we can assume that $(v_{0n}, v_{1n})\rightharpoonup(v_0, v_1)$ as $n\to\infty$ with

$$\left\|(v_0, v_1)-(\widetilde{Q}_\ell,\,\partial_t\widetilde{Q}_\ell)(x,0)\right\|_{\dot{H}^1\times L^2}\leq 2\epsilon_0.$$

Hence by the continuity of topological degree (This is a direct consequence of the definition (6.12) of degree, and can be proved by noting that $\int_{R^2}\partial_x u\times\partial_y u\mathrm{d}x\mathrm{d}y=0$ for any $\dot{H}^1$ mapping from $R^2\to S^2$, and the dominated convergence theorem.) in $\dot{H}^1$ and the fact that degree only takes value in integers (see [3]), we see that if ϵ_0 is taken small enough, then

$$\deg(v_0) = 1.$$

(6.15) implies that (v_0, v_1) satisfies the *first order* "Bogomol'nyi equations" (see [2]):

$$\begin{aligned} &v_1 + l\partial_1 v_0 = 0; \\ &\sqrt{1-l^2}\partial_1 v_0 - v_0 \times \partial_2 v_0 = 0; \\ &\partial_2 v_0 + \sqrt{1-l^2} v_0 \times \partial_1 v_0 = 0. \end{aligned} \tag{6.17}$$

Equations (6.17) can be reduced by an obvious change of variable to the case $l = 0$, in which case they can be explicitly solved as harmonic maps. Hence we see that there exists $\tilde{\tilde{Q}} \in \mathcal{M}_1$ such that

$$(v_0, v_1) = \left(\tilde{\tilde{Q}}_\ell, \partial_t \tilde{\tilde{Q}}_\ell\right)(x, 0).$$

Thus we can write

$$(v_{0n}, v_{1n}) = \left(\tilde{\tilde{Q}}_\ell, \partial_t \tilde{\tilde{Q}}_\ell\right)(x, 0) + (\tilde{r}_{0n}, \tilde{r}_{1n})$$

with $(\tilde{r}_{0n}, \tilde{r}_{1n}) \rightharpoonup 0$ as $n \to \infty$. Then the energy expansion for (v_{0n}, v_{1n}) around $\left(\tilde{\tilde{Q}}_\ell, \partial_t \tilde{\tilde{Q}}_\ell\right)(x, 0)$ gives

$$\begin{aligned} \mathcal{E}(Q_\ell, \partial_t Q_\ell) + \frac{1}{n} &\geq \int_{R^2} \left(\frac{|\nabla v_{0n}|^2}{2} + \frac{|v_{1n}|^2}{2}\right) \mathrm{d}x \\ &= \frac{1}{2}\int_{R^2} \left|\nabla \tilde{\tilde{Q}}_\ell\right|^2 + \left|\partial_t \tilde{\tilde{Q}}_\ell\right|^2 \mathrm{d}x \\ &\quad + \int_{R^2} \nabla \tilde{\tilde{Q}}_\ell^{\dagger} \nabla \tilde{r}_{0n} + \partial_t \tilde{\tilde{Q}}_\ell^{\dagger} \tilde{r}_{1n} \mathrm{d}x \\ &\quad + \int_{R^2} \frac{|\nabla \tilde{r}_{1n}|^2}{2} + \frac{|\tilde{r}_{1n}|^2}{2} \mathrm{d}x \\ &= \mathcal{E}\left(\vec{\tilde{\tilde{Q}}}_\ell\right) + \int_{R^2} \frac{|\nabla \tilde{r}_{1n}|^2}{2} + \frac{|\tilde{r}_{1n}|^2}{2} \mathrm{d}x + o_n(1). \end{aligned}$$

By (6.9) and the fact that $\mathcal{E}\left(\vec{\tilde{\tilde{Q}}}_\ell\right) = \mathcal{E}(\vec{Q}_\ell)$, we see that

$$(\tilde{r}_{0n}, \tilde{r}_{1n}) \to 0, \text{ in } \dot{H}^1 \times L^2.$$

This is a contradiction to (6.11). The theorem is proved. ■

Now we turn to the proof of the second main theorem in the article.

Theorem 6.2. Let u be a classical wave map defined on $R^2 \times (0, 1]$ with energy $\mathcal{E}(\vec{u}) < \mathcal{E}(Q, 0) + \epsilon_0^2$, where Q is a harmonic map of degree 1, that blows up at time 0 and at the origin. Assume that ϵ_0 is sufficiently small. Then there exist $\ell \in R^2$ with $|\ell| \ll 1$, $x(t) \in R^2$, $\lambda(t) > 0$ with

$$\lim_{t\to 0} \frac{x(t)}{t} = \ell, \ \ \lambda(t) = o(t)$$

and $(v_0, v_1) \in \dot{H}^1 \times L^2 \cap C^\infty(R^2\backslash\{0\})$ with $(v_0 - u_\infty, v_1)$ being compactly supported, such that

$$\text{(i)} \inf\left\{ \left\|\vec{u}(t) - (v_0, v_1) - (Q_\ell, \partial_t Q_\ell)\right\|_{\dot{H}^1\times L^2} : Q_\ell \in \mathcal{M}_{\ell,1}\right\} \to 0, \text{ as } t \to 0;$$

$$\text{(ii)} \left\|\vec{u}(t) - (v_0, v_1)\right\|_{\dot{H}^1\times L^2\left(R^2\backslash B_{\lambda(t)}(x(t))\right)} \to 0 \text{ as } t \to 0.$$ □

Proof. We have already proved that along a sequence of times $t_n \to 0+$,

$$\vec{u}(t_n) = \left(Q_\ell\left(\frac{x - x_n}{\lambda_n}, 0\right), \frac{1}{\lambda_n}\partial_t Q_\ell\left(\frac{x - x_n}{\lambda_n}, 0\right)\right) + (v_0, v_1) + o_{\dot{H}^1\times L^2}(1), \tag{6.18}$$

where $(v_0, v_1) \in \dot{H}^1 \times L^2 \cap C^\infty(R^2\backslash\{0\})$ and

$$\lim_{n\to\infty} \frac{x_n}{t_n} = \ell, \text{ and } \lambda_n = o(t_n), \text{ as } n \to \infty.$$

Since

$$\epsilon_n := \int_{B_{2t_n}\backslash B_{t_n}} \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2}\right)(x, t_n)dx \to 0, \text{ as } n \to \infty, \tag{6.19}$$

we can find $r_n \in (t_n, 2t_n)$ such that

$$\int_{\partial B_{r_n}} \frac{|\not\partial u|^2}{2}(x, t_n)d\sigma \lesssim \frac{\epsilon_n}{r_n}.$$

Let

$$\overline{u}_n = \frac{1}{2\pi r_n}\int_{\partial B_{r_n}} u(t_n)\mathrm{d}\sigma. \qquad \blacksquare$$

Take a smooth cutoff function η_n with $\eta_n \equiv 1$ on B_{r_n}, $\operatorname{supp}\eta_n \Subset B_{2r_n}$ and $|\nabla\eta_n| \lesssim r_n^{-1}$. Define

$$(u_{0n},\, u_{1n}) = \begin{cases} \vec{u}\,(x, t_n) & \text{for } |x| < r_n; \\ (P\left[\eta_n(r)(u(r_n\theta, t_n) - \overline{u}_n) + \overline{u}_n\right],\, 0) & \text{for } |x| > r_n. \end{cases}$$

One can check that $(u_{0n},\, u_{1n}) \in \dot{H}^s \times H^{s-1}$ for $s < \frac{3}{2}$, and $u_{0n} \equiv P(\overline{u}_n)$ for large x. Moreover,

$$\|(u_{0n},\, u_{1n})\|^2_{\dot{H}^1\times L^2(B^c_{t_n})} \lesssim \epsilon_n. \tag{6.20}$$

Let u_n be the solution to the wave map equation with $\vec{u}_n(t_n) = (u_{0n},\, u_{1n})$. (The local existence of u_n follows from subcritical wellposedness theory.) Then by finite speed of propagation, $u_n \equiv u$ for $|x| < t$ and $t \in (0,\, t_n]$, assuming that u_n is defined in $[t,\, t_n]$. In addition, by (6.20) and energy flux identity, since the energy flux of u_n is equal to that of u on $|x| = t$, $t \in (0,\, t_n]$ which decays to zero as $n \to \infty$, we get that

$$\int_{|x|>t} |\nabla_{x,t}u_n|^2(x,t)\mathrm{d}x \lesssim \epsilon_n + o_n(1) \tag{6.21}$$

for $t \le t_n$, again assuming that u_n is defined in $[t,\, t_n]$. As u_n is identical to u in the singularity light cone $|x| < t$, $0 < t \le t_n$ and u_n has small energy for $|x| \ge t$, $0 < t \le t_n$, we conclude that u_n is defined for $t \in (0,\, t_n]$. From (6.18), it is easy to verify that

$$\begin{aligned} &\deg(u_n(t_n)) = 1; \\ &\mathcal{E}(\vec{u}_n) \le \mathcal{E}(\vec{Q}_\ell) + o_n(1); \\ &\left|\mathcal{M}(\vec{u}_n) - \mathcal{M}(\vec{Q}_\ell)\right| = o_n(1), \end{aligned}$$

where $\mathcal{M}(\vec{u})$ denotes the momentum of u. Hence by Theorem 6.1, $\vec{u}_n(t)$ stays in a $\delta(\epsilon_n)$ neighborhood of $\mathcal{M}_{\ell,1}$ for $t \le t_n$ with $\delta(\epsilon_n) \to 0$ as $n \to \infty$. It follows that

$$\lim_{t\to 0}\, \inf\left\{ \left\|\vec{u}\,(t) - (v_0,\, v_1) - (Q_\ell,\, \partial_t Q_\ell)\right\|_{\dot{H}^1\times L^2} :\, Q_\ell \in \mathcal{M}_{\ell,1}\right\} = 0. \tag{6.22}$$

Part (i) of the theorem is proved. The fact that all degree 1 harmonic maps are co-rotational implies that $\mathcal{M}_{\ell,1}$ is a compact set in the energy space, modulo translations and dilations. Hence, by the regularity of u outside the singularity lightcone, we can find $x(t)$ and $\lambda(t)$ with $\lambda(t) = o(t)$ and

$$\limsup_{t \to 0+} \frac{|x(t)|}{t} \leq 1.$$

The main remaining task is to show that

$$\lim_{t \to 0+} \frac{x(t)}{t} = \ell. \tag{6.23}$$

Without loss of generality, let us assume that $\ell = l\, e_1$ By (6.22), it follows that

$$\lim_{t \to 0} \int_{|x|<t} -\partial_t u\, \partial_{x_1} u(x,t) \mathrm{d}x = \frac{4l\pi}{\sqrt{1-l^2}}. \tag{6.24}$$

Direct computation shows

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} \int_{|x|<t} x_1 \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} \right)(x,t) \mathrm{d}x \\
&= \int_{|x|=t} x_1 \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} \right)(x,t) \mathrm{d}\sigma + \int_{|x|=t} x_1 \frac{x}{|x|} \cdot \nabla u^\dagger\, \partial_t u \mathrm{d}\sigma \\
&\quad - \int_{|x|<t} \partial_{x_1} u\, \partial_t u(x,t) \mathrm{d}x.
\end{aligned}$$

Integrating the above identity from $t = 0$ to t, we get that

$$\int_{|x|<t} x_1 \left(\frac{|\nabla u|^2}{2} + \frac{|\partial_t u|^2}{2} \right)(x,t) \mathrm{d}x = O\left(\mathrm{Flux}(0,t)\right) t + \frac{4l\pi\, t}{\sqrt{1-l^2}} + o(t). \tag{6.25}$$

As $\mathrm{Flux}(0,t) \to 0$ as $t \to 0$, by (6.25) and (6.22), (6.23) follows straightforwardly. The theorem is proved.

Funding

T.D. was partially supported by ERC Grant Blowdisol 291214. H.J. was partially supported by NSF grant DMS-1600779, and grant DMS-1128155 through IAS. C.K. was partially supported by NSF Grants DMS-1265429 and DMS-1463746. F.M. was supported by ERC Grant Blowdisol 291214.

References

[1] Belavin, A. and A. Polyakov. "Metastable states of two-dimensional isotropic ferromagnets." *JETP Letters* 22 (1975): 245–7 (Russian).

[2] Bogomolnyi, E. "The stability of classical solutions." *Soviet Journal of Nuclear Physics* 24, no. 4 (1976): 449–54 (Russian).

[3] Brezis, H. and J-M. Coron. "Large Solutions for Harmonic Maps in Two Dimensions." *Communications in Mathematical Physics* 92 (1983): 203–15.

[4] Christodoulou, D. and A. Tahvildar-Zadeh. "On the asymptotic behavior of spherically symmetric wave maps." *Duke Mathematical Journal* 71, no. 1 (1993): 31–69.

[5] Christodoulou, D. and A. Tahvildar-Zadeh. "On the regularity of spherically symmetric wave maps." *Communications on Pure and Applied Mathematics* 46, no. 7 (1993): 1041–91.

[6] Côte, R. "Instability of nonconstant harmonic maps for the (1+2)-dimensional equivariant wave map system." *International Mathematics Research Notices* 2005, no. 57 (2005): 3525–49.

[7] Côte, R. "On the soliton resolution for equivariant wave m aps to the sphere." *Communications on Pure and Applied Mathematics* 68, no. 11 (2015): 1946–2004.

[8] Côte, R., C. Kenig, A. Lawrie, and W. Schlag. "Profiles for the radial focusing 4d energy-critical wave equation." preprint 2014, arXiv:1402.2307.

[9] Côte, R., C. Kenig, A. Lawrie, and W. Schlag. "Characterization of large energy solutions of the equivariant wave map problem: I." *American Journal of Mathematics* 137, no. 1 (2015): 139–207.

[10] Côte, R., C. Kenig, A. Lawrie, and W. Schlag. "Characterization of large energy solutions of the equivariant wave map problem: II." *American Journal of Mathematics* 137, no. 1 (2015): 209–50.

[11] Côte, R., C. Kenig, and F. Merle. "Scattering below critical energy for the radial 4D Yang-Mills equation and for the 2D corotational wave map system. *Communications in Mathematical Physics* 284, no. 1 (2008): 203–25.

[12] Côte, R., C. Kenig, and W. Schlag. "Energy partition for the linear wave equation." *Mathematische Annalen* 358 (2014): 573–607.

[13] Donninger, R. and J. Krieger. "Nonscattering solutions and blowup at infinity for the critical wave equation." *Mathematische Annalen* 357, no. 1 (2013): 89–163.

[14] Duyckaerts, T., H. Jia, C. Kenig, and F. Merle. "Soliton resolution along a sequence of times for the focusing energy critical wave equation." preprint 2016, arXiv:1601.01871.

[15] Duyckaerts, T., C. Kenig, and F. Merle. "Universality of blow-up profile for small radial type II blow-up solutions of the energy-critical wave equation." *Journal of the European Mathematical Society (JEMS)* 13, no. 3 (2011): 533–99.

[16] Duyckaerts, T., C. Kenig, and F. Merle. "Universality of the blow-up profile for small type II blow-up solutions of the energy-critical wave equation: the nonradial case." *Journal of the European Mathematical Society (JEMS)* 14, no. 5 (2012): 1389–454.

[17] Duyckaerts, T., C. Kenig, and F. Merle. "Classification of radial solutions of the focusing, energy-critical wave equation." *Cambridge Journal of Mathematics* 1, no. 1 (2013): 75–144.

[18] Duyckaerts, T., C. Kenig, and F. Merle. "Scattering for radial, bounded solutions of focusing supercritical wave equations." *International Mathematics Research Notices* no. 1 (2014) 224–58.

[19] Eells, J. and J. C. Wood. "Restrictions on harmonic maps of surfaces." *Topology* 15, no. 3 (1976): 263–6.

[20] Grillakis, M. "Energy estimates and the wave map problem." *Communications in Partial Differential Equations* 23, no. 5–6 (1998): 887–911.

[21] Grinis, R. "Quantization of Time-Like Energy for Wave Maps into Spheres." *Communications in Mathematical Physics* (2016). doi:10.1007/s00220-016-2766-9.

[22] Hillairet, M. and P. Raphaël. "Smooth type II blow-up solutions to the four-dimensional energy-critical wave equation." *Analysis & PDE* 5, no. 4 (2012): 777–829.

[23] Jendrej, J. "Construction of type II blow-up solutions for the energy-critical wave equation in dimension 5." *Journal of Functional Analysis* 272, no. 3 (2017): 866–917.

[24] Jendrej, J. "Construction of two-bubble solutions for energy-critical wave equations." preprint 2016, arXiv:1602.06524.

[25] Jia, H. "Soliton resolution along a sequence of times with dispersive error for type II singular solutions to focusing energy critical wave equation." preprint 2015, 42 pages, arXiv:1510.00075.

[26] Jia, H. and C. Kenig. "Asymptotic decomposition for semilinear wave and equivariant wave map equations." To appear in *American Journal of Mathematics*, see also arXiv 1503.06715.

[27] Jia, H., B. P. Liu, W. Schlag, and G.X. Xu. "Generic and non-generic behavior of solutions to the defocusing energy critical wave equation with potential in the radial case." *International Mathematics Research Notices*, doi: 10.1093/imrn/rnw181, see also arXiv:1506.04763.

[28] Jia, H., B. P. Liu, and G. X. Xu. "Long time dynamics of defocusing energy critical 3 + 1 dimensional wave equation with potential in the radial case." *Communications in Mathematical Physics* 339, no. 2 (2015): 353–84.

[29] Kenig, C., A. Lawrie, B. P. Liu, and W. Schlag. "Stable soliton resolution for exterior wave maps in all equivariance classes." *Advances in Mathematics* 285 (2015): 235–300.

[30] Kenig, C., A. Lawrie, B. P. Liu, and W. Schlag. "Channels of energy for the linear radial wave equation." *Advances in Mathematics* 285 (2015): 877–936.

[31] Kenig, C., A. Lawrie, and W. Schlag, "Relaxation of wave maps exterior to a ball to harmonic maps for all data." *Geometric and Functional Analysis* 24, no. 2 (2014): 610–47.

[32] Klainerman, S. and M. Machedon. "Smoothing estimates for null forms and applications." *Duke Mathematical Journal* 81 (1995): 99–133.

[33] Klainerman, S. and M. Machedon. "On the optimal local regularity for gauge field theories." *Differential and Integral Equations* 10 (1997): 1019–30.

[34] Klainerman, S. and M. Machedon. "On the algebraic properties of the $H^{n/2,1/2}$ spaces." *International Mathematics Research Notices* 15 (1998): 765–74.

[35] Klainerman, S. and S. Selberg. "Remark on the optimal regularity for equations of wave maps type." *Communications in Partial Differential Equations* 22 (1997): 901–18.

[36] Klainerman, S. and S. Selberg. "Bilinear estimates and applications to nonlinear wave equations." *Communications in Contemporary Mathematics* 4, no. 2 (2002): 223–95.

[37] Krieger, J. and W. Schlag. "Concentration Compactness for Critical Wave Maps." EMS Monographs in Mathematics. Zürich: European Mathematical Society (EMS), 2012. vi+484 pp. ISBN: 978-3-03719-106-4.

[38] Krieger, J., W. Schlag, and W. Tataru. "Renormalization and blow up for the critical Yang-Mills problem." *Advances in Mathematics* 221, no. 5 (2009): 1445–1521.

[39] Lawrie, A. and S. J. Oh. "A refined threshold theorem for (1+2)-dimensional wave maps into surfaces." *Communications in Mathematical Physics* 342, no. 3 (2016): 989–99.

[40] Raphaël, P. and I. Rodnianski. "Stable blow up dynamics for the critical co-rotational wave maps and equivariant Yang-Mills problems." *Publications Mathématiques. Institut des Hautes Études Scientifiques* 115 (2012): 1–122.

[41] Rodnianski, I. and J. Sterbenz. "On the formation of singularities in the critical O(3) -model." *Annals of Mathematics. Second Series* 172, no. 1 (2010): 187–242.

[42] Rodriguez, C. "Profiles for the radial focusing energy-critical wave equation in odd dimensions." *Differential and Integral Equations* 21, no. 5–6 (2016): 505–70.

[43] Schlag, W. "Semilinear wave equations." *Proceedings of the ICM*, Seoul, South Korea, 2014. Volume III, p. 425.

[44] Selberg, S. "Multilinear space-time estimates and applications to local existence theory for non-linear wave equations." Thesis, Princeton University, 1999.

[45] Shatah, J. and A. Tahvildar-Zadeh. "Regularity of harmonic maps from the Minkowski space into rotationally symmetric manifolds." *Communications on Pure and Applied Mathematics* 45, no. 8 (1992): 947–71.

[46] Shatah, J. and A. Tahvildar-Zadeh. "On the Cauchy problem for equivariant wave maps." *Communications on Pure and Applied Mathematics* 47, no. 5 (1994): 719–54.

[47] Sterbenz, J. and D. Tataru. "Energy dispersed large data wave maps in 2+1 dimensions." *Communications in Mathematical Physics* 298, no. 1 (2010): 139–230.

[48] Sterbenz, J. and D. Tataru. "Regularity of wave-maps in dimension 2+1." *Communications in Mathematical Physics* 298, no. 1 (2010): 231–64.

[49] Struwe, M. "Equivariant wave maps in two space dimensions." *Communications on Pure and Applied Mathematics* 56, no. 7 (2003): 815–23.

[50] Tao, T. Global regularity of wave maps I. Small critical Sobolev norm in high dimension. *International Mathematics Research Notices* 7 (2001): 299–328.

[51] Tao, T. "Global regularity of wave maps II. Small energy in two dimensions." *Communications in Mathematical Physics* 224 (2001): 443–544.

[52] Tao, T. "Global regularity of wave maps III. Large energy from R^{1+2} to hyperbolic spaces." arXiv:0805.4666, 2008.

[53] Tao, T. "Global regularity of wave maps IV. Absence of stationary or self-similar solutions in the energy class." arXiv:0806.3592, 2008.

[54] Tao, T. "Global regularity of wave maps V. Large data local well-posedness in the energy class." arXiv:0808.0368, 2008.

[55] Tao, T. "Global regularity of wave maps VI. Minimal energy blowup solutions." arXiv:0906.2833, 2009.

[56] Tataru, D. "Local and global results for wave maps I." *Communications in Partial Differential Equations* 23, (1998): 1781–93.

[57] Tataru, D. "On global existence and scattering for the wave maps equation." *American Journal of Mathematics* 123, no. 1 (2001): 37–77.

[58] Tataru, D. "Rough solutions for the wave maps equation." *American Journal of Mathematics* 127, no. 2 (2005): 293–377.